syi.hu/cse

Cselekvéselmélet 1.0

Syi

i@syi.hu

November 20, 2013

Contents

Előszó 1

1 Logikai, módszertani alapok **7**

1.1 Metafogalmak 10
1.2 Halmazelmélet 14
1.3 Kijelentéslogika 16
1.3.1 Kijelentéslogikai műveletek 17
1.3.2 Kijelentéskalkulus 23
1.4 Predikátumlogika 27
1.4.1 Predikátumlogikai alapfogalmak 27
1.4.2 Predikátumkalkulus 29
1.5 Relációelmélet 30
1.5.1 Relációelméleti fogalmak 30
1.5.2 Relációalgebra 46
1.6 Kategoriális logika 47
1.7 Ontológia 58
1.7.1 JEPD-elv 58
1.7.2 Reifikáció 59
1.8 Másodrendű logika 60
1.9 Dinamikus logika 62
1.10 Modális logika 63
1.11 A tudás és a vélekedés logikája 74
1.11.1 Episztemikus logika 75
1.11.2 Doxasztikus logika 79
1.11.3 Kollektív tudás 81
1.11.4 Közös tudás 84
1.11.5 Dinamikus episztemikus logika 85
1.11.6 Tudástipológiák 86
1.12 Preferencialogika 86
1.12.1 Intrinzikus preferencia 89
1.12.2 Az intrinzikus preferencia tulajdonságai 91
1.12.3 Monadikus preferenciarelációk 119
1.12.4 Preferencia és modális logika 119
1.12.5 További elméletek 128
1.13 Deontikus logika 128
1.13.1 Dinamikus logika és deontika 143

2 A társadalom modelljei **145**
2.1 Ego . . . 145
2.1.1 Szelf . . . 145
2.2 A világ . . . 150

3 Egyéni döntések **153**
3.1 Intencionalitás . . . 153
3.1.1 Brentano . . . 153
3.1.2 Husserl . . . 154
3.1.3 Anscombe, Dennet, Fodor, Searle . . . 155
3.1.4 egyéb . . . 157
3.1.5 Neurológiai alapok . . . 158
3.1.6 Értelem, tudat . . . 158
3.2 Reprezentáció . . . 158
3.3 Teleológia vs. teleonómia . . . 158
3.4 Érvényesség . . . 159
3.5 Intenció . . . 159
3.5.1 Bratman előtti korszak: vágy-vélekedés modellek . . . 162
3.5.2 Bratman: BDI-elmélet . . . 164
3.5.3 más elméletek . . . 166
3.5.4 Az intencionalitás értékelése . . . 167
3.6 Vágy . . . 167
3.6.1 Castelfranchi: vélekedés és vágy . . . 167
3.6.2 Cross: diszkrepancia alapú vágyelmélet . . . 169
3.6.3 magasabbrendű vágyak . . . 169
3.7 Másodrendű vágy: érték . . . 170
3.7.1 Osgood . . . 171
3.7.2 Axiológiai hatszög . . . 171
3.8 Akarat . . . 171
3.9 Egyéni döntések . . . 173
3.9.1 Racionalitásértelmezések . . . 173
3.9.2 döntési szabályok . . . 188
3.9.3 Viselkedésgazdaságtan: “döntési irracionalitások” . . . 189

4 Közösségi döntések **191**
4.1 Aggregálás mint közösségi döntési technika . . . 191
4.1.1 Preferenciaprofilok tipológiája . . . 197
4.2 Preferenciaaggregálás . . . 202
4.2.1 Manipuláció . . . 203
4.2.2 Algoritmusmanipuláció . . . 204
4.2.3 Condorcet-paradoxon . . . 207
4.2.4 Arrow lehetetlenségi tétele . . . 208
4.2.5 Lehetségességi tételek . . . 215
4.2.6 No-show paradox . . . 222
4.2.7 Reprezentációs paradoxon . . . 224
4.2.8 Szavazatmanipuláció . . . 224
4.2.9 Szavazatcsere . . . 226

4.2.10 Csalásbiztos eljárás . . . 229
4.2.11 Gibbard-Satterthwaite tétel . . . 231
4.2.12 napirend . . . 236
4.2.13 AI-válasz . . . 236
4.2.14 Párosításelmélet . . . 239
4.2.15 Körzethatár-manipuláció . . . 243
4.2.16 Többkörzetes paradoxon . . . 244
4.2.17 Többesetes paradoxon . . . 245
4.2.18 Anscombe-paradoxon . . . 248
4.2.19 Osztrogorszkij paradoxon . . . 250
4.3 Véleményaggregálás . . . 252
4.3.1 Doktrinális paradoxon . . . 252
4.3.2 Diszkurzív paradoxon, List-Pettit lehetetlenségi tétele . . . 254
4.3.3 Aggregálás általában . . . 257
4.3.4 Liberális paradoxon . . . 257
4.3.5 Multiple elections . . . 257
4.4 Másodrendű preferenciák . . . 258
4.4.1 Archimédeszi preferencia . . . 259
4.4.2 Lexikografikus preferencia . . . 260
4.4.3 Formális modellek . . . 260
4.4.4 egyéb . . . 265

5 Társadalmi cselekvés 267
5.1 Cselekvés . . . 268
5.1.1 Szent Anzelm cselekvéstipológiája . . . 268
5.1.2 Tevés és tartózkodás . . . 270
5.1.3 Magasabbrendű cselekvés . . . 273
5.2 Cselekvő . . . 273
5.2.1 Ego és Alter . . . 274
5.3 Formális modellek . . . 274
5.3.1 stit-elmélet . . . 274
5.3.2 program-elmélet . . . 275
5.3.3 Davidson . . . 275
5.4 Interdependens cselekvések . . . 275

6 A társadalmi cselekvés típusai 279
6.1 Fiske relációs elmélete . . . 279
6.2 Szabályszerűségek . . . 279
6.3 Ego és a világ . . . 280
6.3.1 Célracionalitás . . . 280
6.3.2 Értékracionalitás . . . 280
6.3.3 Érzelem . . . 281
6.3.4 Habitus . . . 284
6.3.5 Miért kell a racionalitás fogalma és miért hamis a mítosza? . . . 288

7 Társadalmi kapcsolat **291**
7.1 Ágensre irányultság: társas preferencia . . . 292
7.2 Magasabbrendű beállítódás: büntetni akarás . . . 294
7.3 Társas értékorientáció . . . 295
7.3.1 Egoizmus . . . 302
7.3.2 Altruizmus . . . 304
7.3.3 Kooperáció . . . 306
7.3.4 Versengés . . . 307
7.3.5 Agresszió . . . 307
7.3.6 Autoagresszió . . . 309
7.3.7 Önfeláldozás . . . 310
7.3.8 Destrukció . . . 310
7.4 Dinamikus kapcsolat: reciprocitás . . . 310
7.5 Méltányosságelméletek . . . 310
7.6 Szolidaritás . . . 311
7.7 Dominancia, státus . . . 311
7.8 A társadalmi kapcsolat típusai . . . 311
7.9 Neurológiai folyamatok . . . 314

8 Kényszeren alapuló társadalmi kapcsolat **315**
8.1 Érdek . . . 315
8.2 Hatalom . . . 315
8.2.1 Hatalom az állatok közt . . . 316
8.2.2 Hatalom emberek közt . . . 316
8.3 A kooperáció biológiai mechanizmusai . . . 318
8.3.1 Rokoni altruizmus . . . 318
8.3.2 Zöld szakáll hatás . . . 319
8.3.3 Spite . . . 319
8.3.4 Csoportszelekció . . . 320
8.3.5 A costly signaling és a handicap principle . . . 320

9 Kényszermentes, interdependens társadalmi kapcsolat **321**
9.1 Strukturális kényszerek: játékelmélet . . . 321
9.2 Magjátékok . . . 323
9.3 A játékok statikus kiterjesztése . . . 337
9.3.1 Aszimmetrikus játékok . . . 337
9.3.2 Többalternatívás játékok . . . 339
9.3.3 Többszereplős játékok . . . 340
9.3.4 Nem szigorúan rendezett játékok . . . 341
9.4 A játékok magasabbrendű kiterjesztése: büntetés . . . 341
9.4.1 Másodrendű játékok – költséges büntetés . . . 342
9.4.2 Szankcióparadoxon . . . 343
9.4.3 Aszinkron játékok . . . 345
9.4.4 Alkujátékok . . . 346
9.5 A játékok dinamikus kiterjesztése: társadalmi emlékezet . . . 350
9.5.1 Reciprocitás . . . 351
9.5.2 Direkt reciprocitás játékok . . . 352

9.6 A játékok közösségi kiterjesztése: reputáció 364
9.6.1 Indirekt reciprocitás játékok 364
9.6.2 Közös minták a direkt és indirekt reciprocitás sikeres stratégiáiban 371
9.6.3 Értékelési rendszerek 373
9.7 Koordinációs problémák, közös tudás, aszimmetrikus információ 373
9.7.1 Szaliencia, fókuszpont, precedens 373
9.7.2 Közös tudás 374
9.8 minimal social situation 374
9.9 Kollektív cselekvés, kollektív döntés 374
9.10 Kooperatív játékok: koalícióelmélet 375
9.11 Társadalmi dilemmák feloldási lehetőségei 376

10 Társadalmi pozíció **377**
10.1 Perszonális deontika 377
10.2 Normatív pozíciók, jogosultságok és kötelezettségek 380
10.2.1 Hohfeld elsőrendű jogviszonyfogalmai 380
10.2.2 Az elsőrendű jogviszonyfogalmak formalizálása 395
10.2.3 Hohfeld másodrendű jogviszonyfogalmai 404
10.2.4 A másodrendű jogviszonyfogalmak formalizálása 410
10.2.5 A jogviszonyfogalmak és szabályfogalmak viszonya 417
10.2.6 A jogviszonyfogalmak teljes logikai rendszere 429
10.2.7 A jogviszonyfogalmak kiterjesztése 430
10.2.8 pozíciók 434
10.2.9 Jogi szemantika 434
10.2.10 Deontikus logika és az informatikusok 434
10.3 Társadalmi szerep 435
10.4 Privacy 435

11 Uralom **437**
11.1 Érvényesség, legitimitás 438
11.1.1 Legitimitástipológia 439
11.2 Uralomtípusok 440
11.2.1 Tradicionális uralom 440
11.2.2 Racionális uralom 440
11.2.3 Karizmatikus uralom 440
11.2.4 Átértelmezett karizmatikus uralom 441
11.2.5 Áttekintés 441

12 Norma **443**
12.1 A norma leírása 444
12.1.1 A norma jellemzői 444
12.1.2 A norma típusai 453
12.2 Játékelméleti magyarázat a normák keletkezésére 461
12.3 Rítus, narratíva 461

13 Normavezérelt társadalmi kapcsolat **463**
13.1 Hatalom és uralom viszonya 463
13.2 Szervezet 463
13.3 norma és érdek 464

14 Egyéb **465**
14.1 erkölcsi elvek 465
14.2 Antinómiák, önreferencia 465
14.3 Hibák, javítandók 465

Előszó

Mario Bunge,

Harrison White: Identity and control 2. kiadás, 2008 getting action vs. blocking action

replika White-különszám, 2011 varga attila szerk. berger viktor

olasz sociologica online különszám Harrison White catnet fogalmának geneziséről

Stephan Fuchs, Aganist essentialism, Lunhmant értelmezi újra kapcsolatházati szempontokból

Steve Martin, Social structures

Ronald Breiger, blokkmodell elemzés alapjai Duality of persons and groups

Ronald Burt, Structural holes

John Scott, Handbook of Social Networks

Linton Freemann, History of Social Network, 2004

Charles Kadushin, Social Network http://www.charleskadushin.com/

Handbook of Social Networks, 3 kötet

Angelusz, Tardos Társadalmak rejtett hálózata

> „There is only one social science." Jack Hirshleifer [358]

> „Dr. Felix Smith, a Stanford Kutatóintézet Molekuláris Fizika Tanszékének vezetője mesélt egy fizikus barátjáról, aki Los Alamos ban dolgozott a II. világháború után. Egy bonyolult probléma megoldása végett elment a híres magyar matematikushoz, Neumann Jánoshoz, aki Los Alamosban volt konzultáns.
> 'A karakterisztikák módszerével egyszerűen meg lehet oldani.' – mondta Neumann.
> 'Félek, nem értem a karakterisztikák módszerét.' – felelte a fizikus.
> 'Fiatalember,' – válaszolta Neumann – 'a matematikában az ember nem megérti a dolgokat, hanem megszokja.'." [843, 208])

> „Mélységesen téveteg közhely, amelyet minden jegyzetfüzet és igen kitűnő emberek is szakadatlanul ismételgetnek, arról prédikálva, fejlesszék ki azt a szokást, hogy tudatosan gondolják el azt, amit tesznek. Ennek pontosan az ellenkezője az igaz. A civilizáció azáltal halad előre, hogy gyarapítja ama fontos műveletek számát, amelyek végrehajthatók a róluk való gondolkodás nélkül. A gondolkodási műveletek a csaták lovasrohamaihoz hasonlítanak – számuk szigorúan korlátozott, megkívánják a friss lovakat, és csak a döntő pillanatokban vethetők be." A.N.Whitehead [815, 61]

> „A szabályok létét az indokolja, hogy az emberek nem képesek erőfeszítések nélkül és hibátlanul alkalmazni az utilitariánus elvet; szükség van rá, hogy időt takarítsanak meg, és útjelzőket tűzzenek ki." J. Rawls [595, 387]

A cselekvéselmélet a társadalmi cselekvések leírásával, megértésével, magyarázatával

foglalkozik. Ebben a minőségében – definíció szerint – a módszertani individualizmus talaján áll. A cselekvéselmélet fogalmát azonban lehet szűkebb és tágabb értelemben használni. Szűkebb értelemben véve beszélhetünk azokról a modellekről, amelyek azt a kérdést próbálják megválaszolni, hogy mikor tekinthető valami cselekvésnek. A fogalom tágabb értelemben minden olyan szempontot és problémát ide sorolhatunk, amely a társadalmi cselekvések szabályszerűségeinek jobb megértéséhez segíthet minket. Ebben a könyvben a cselekvéselmélet fogalmának kiterjesztett jelentését fogadom el, de szentelek egy külön fejezetet a szűkebb értelmezésnek is.

Jack Hirshleifer egy 1985-ös cikkéből származik az oldal tetején olvasható idézet [358], aminek tartalmával egyetértek. A cikk a közgazdaságtan imperializmusával és mindenhatóságával szemben megfogalmazott programként is értelmezhető. A közgazdász szerző egyrészt rámutat arra, miért tudott minden társadalomtudományi területre behatolni a közgazdaságtani megközelítés, másrészt elemzi azt is, hogy milyen hiányosságai vannak a közgazdaságtan világképének, hol és miért kell kibővíteni az eszközkészletét, fogalmi modelljeit, harmadrészt felvázolja azt az egységes társadalomtudományi keretrendszert, amelyben a "kigészített" közgazdaságtan és a biológia által teremtett alapokra lehet a további szaktudásterületeket felépíteni.

Egy ilyen programmal óvatosan egyetértve egyetlen ponton egészíteném csak ki azt: a társadalmiságunk döntő elemének tartom a társadalmi valóság konstruált jellegét, amit sem a biológia, sem a közgazdaságtan nem képes megragadni, ezért ezt önálló, programalkotó építőelemnek tartom.

Közel két éven át írtam ezt a könyvet. Talán a legfontosabb célom az volt, hogy az érték, a norma, az érdek, a racionalitás, a racionális cselekvés fogalmait minél pontosabban definiálni, értelmezni tudjam. Azt gondoltam, hogy ezek azok az alapvető fogalmak, amelyekre támaszkodva a leginkább képesek lehetünk magyarázni a társadalmi cselekvések szabályszerűségeit. Mire a könyvem végére értem sok minden átértékelődött bennem e céllal kapcsolatban. Ma már úgy látom, hogy az embert leginkább az érzelmein és a szokásain keresztül érthetjük meg.

Nem könnyű belátni, még nehezebb elfogadni, de én – ma már – úgy gondolom, hogy választásainkat az érzelmeink határozzák meg. Ha meg akarjuk érteni a döntéseinket, cselekvéseinket, akkor bizony az érzelmek kihagyhatatlanok, tehát módszeresen fel kell tárnunk az érzelmeink működését, hatásmechanizmusait. De talán még ennél is fontosabb ténye az életünknek, hogy a cselekvéseink igen nagy részét (amit csak lehet) szokásokká, reflexekké alakítjuk (habituáljuk), és ezek után, illetve ezek alapján sok esetben "egyszerű" cselekvésautomataként működünk. Ingereket, inputokat kapunk, és ezekre reflexszerűen válaszolunk. Sokkoló élmény volt ráébredni arra, amire két híres ember, Neumann János és Alfred North Whitehead mondatai utalnak: még a racionális működésmódunk is – jelentős részben – megszokáson alapul. Az agyunk számára így gazdaságos, tehát meta szinten, evolúciós értelemben (?) racionálisnak mondható, de akkor is nehezen emészthető.

Ez az elmozdulás még nem igazán érezhető a könyvemben.[1] Amíg írtam, mindvégig az eredeti célok – ha tetszik, a racionalitás paradigma – mentén haladtam. Mindazt, amit

[1] Egy időközben megírt cikkemben azért már tetten érhető ez a felismerés, lásd [726].

itt megírtam, szükségesnek gondolom a társadalmi cselekvések szabályszerűségeinek magyarázatához. Szükségesnek, de messze nem elégségesnek. Ma már úgy látom, hogy sokkal kisebb – vagy legalábbis más – szerepe van az életünkben a racionalitásnak, a mérlegelésnek, a kalkulációnak, mint ahogy korábban hittük, hittem. De legalább még van mit feltárni.

♩♪♩♪♫ ♫

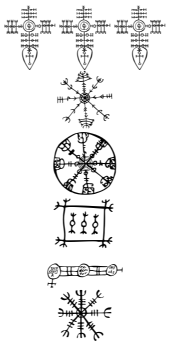
Tisztelgés Max Weber előtt. [724]

♫ ♫♩♪♩♪

a társadalomtudományi gondolkodást meghatározó lényeges változás a kognitív tudományi alapok kiépülése Sok mindent más megvilágításba tesz, sok mindent megerősít és megalapoz.

♩♩♪♫ ♪♫

Végezetül meg kell jegyezzem, hogy a könyv (pontosabban a könyv egy részének) megírásakor egy nagyon is gyakorlati cél lebegett a szemem előtt. Az imperszonális deontika kidolgozásától magamban kezdettől fogva azt reméltem, hogy valamiféle kiindulási pontot, logikai lapot jelenthet egy olyan jogi szemantika (jogi ontológia) kidolgozásához, amely normatív, jogi szövegek gépi feldolgozását lehetővé teszi.



Chapter 1

Logikai, módszertani alapok

„ . . . »a matematika egyik fegfontosabb hozzájárulása a tudományos problémák megoldásához az lehet, hogy az adott problémával kapcsolatban a logikailag létező lehetőségek kimerítő elemzését adja«. Egy ilyen elemzés eredménye a logikailag létező lehetőségek halmazának felosztása egymást kölcsönösen kizáró, de együttesen kimerítő alternatívákra." (Kemeny, Snell és Thompson könyvéből idéz [449, 63])

Ebben a fejezetben bemutatom, értelmezem azokat a módszertani, logikai fogalmakat és elméleteket, amelyeket használni fogok a cselekvéselmélet felépítése során. Nem az a célom ezzel a fejezettel, hogy kimerítő és minden szempontból megalapozott logikai ismereteket adjak át az olvasónak. Csak azt szeretném, hogy egyfelől a könyvem gondolatmenete kellő mértékben megalapozott legyen, másfelől akik birtokában vannak az itt használt logikai apparátusoknak, azok tudják kötni ezeket a fogalmakat az általuk már ismert szerzőkhöz, diszciplinákhoz, akik számára pedig mindezek újdonságot jelentenek, azok tudják, hová kell fordulniuk, ha mélyebb tudásra vágynak.[1]

A logika eszköz ahhoz, hogy a gondolatainkat következetes és egyértelmű módon fejezhessük ki. Az egyértelműség és ezzel a – relatív – konzisztencia lehetősége abból fakad, hogy az elméleteink fogalmait valamilyen formális nyelven, formulák segítségével is reprezentálni tudjuk. Ez azonban nem jelenti azt, hogy csak formulákon keresztül fejtenénk ki mondanivalónkat. A tudományos kommunikációban lényegszerűen két nyelvet használunk. Az emberhez természetes nyelven szólunk, az elméleteinkhez használt formális nyelv(ek) mindig csak kiegészítik a természetes nyelvű kommunikációt. De ha szükséges, az egyértelműsítés kedvéért a formális nyelvekhez fordulhatunk.

További előnye lehet még a logika (matematika) alkalmazásának az a "megnyugtató" érzés, amely akkor keletkezhet bennünk, amikor választ várunk arra a kérdésre, vajon az éppen adott elmélet építése közben lehetne-e más irányban továbbmenni, más fogalmakat használni, más téziseket keresni. A tudományos gondolkodók versenyében mondanak ezt, mondanak azt, és mindig ott lapul a kérdés, nem lehet-e még további fogalmakat, állításokat, elméleteket az addigiak mellé tenni. A fejezet mottájában idézett gondolat arról szól, hogy miután már elköteleztük magunkat bizonyos fogalmak, axiómák,

[1] A logika iránt érdeklődő olvasóknak magyar nyelven Ruzsa Imre könyveit ajánlom (például ezt: [627]), a matematikai logikai kérdésekben Ferenczi Miklós könyvére támaszkodtam [250].

tézisek, elméletek mellett, utána már gyakran kerülhetünk olyan helyzetbe, amikor a logika, matematika összefüggéseire támaszkodva mondhatunk olyat, hogy ennyi és ennyi, ez és ez a lehetőség van elméletileg, megvizsgálhatjuk az összeset, és abban az értelemben megnyugodhatunk, hogy azon a ponton további alternatívákat már nem kell, mert nem is lehet keresnünk.

Az egyértelműség és a konzisztencia kialakításának és fenntartásának lehetőségén túl van még egy érv a logikai apparátus alkalmazása mellett. Egyfajta takarékossági, gazdaságossági szempontból is gondolhatunk a logikára, amikoris az elméletépítéseink során újra és újra alkalmazzuk a logika különböző következtetési szabályait, kalkulusait, szillogizmusait. A következtetési szabályok folyamatos alkalmazása (újrahasznosítása) nagyon fontos hozadéka a logikának.[2]

A gazdaságossági szempontokat azonban nem minden áron érdemes/lehet figyelembe venni. A rokonsági fogalmak (apja, anyja, fia, lánya, testvére, húga, nagyapja, ükapja, sógora, ángya stb.) többszáz tételből álló rendszerét kifejezhetnénk 5-6 alapfogalom segítségével, de nem tesszük ezt[3]. Terminológiai szempontból gazdaságos lenne, hiszen a többszáz helyett csak néhány terminust kellene felvenni a modellünkbe, viszont ennek az lenne az ára, hogy hosszú kifejezéseket kellene képeznünk az alapterminusainkból. 'Ükapád' helyett az 'apád apja apjának apja' alakot (formulát) kellene mondanunk. Ez se kimondani, se befogadni nem lenne jó. Állandóan dolgozni kellene azon, hogy az adott kifejezés mit is jelent. Könnyebb nekünk sok terminust és kevesebb feldolgozási műveletet kezelnünk, mint fordítva. Érdemes hallgatnunk Whitehead fenti – az előszó mottójába foglalt – figyelmeztetésére.

A logika egyfelől egy nagy hagyományokkal bíró, önálló tudományág. Ebben a minőségében behatárolható, közös vizsgálati tárggyal, sajátos elemzési szempontokkal, módszerekkel rendelkezik. Másfelől logikának mondhatunk minden formális nyelven kifejezett elméletet is. A kommunikálni kívánt fogalmakat az adott nyelvnek megfelelő terminusok (nevek és predikátumok) segítségével fejezzük ki, ezek révén kijelentéseket (mondatokat, formulákat) képzünk, melyekhez aztán igazságértékeket rendelhetünk. Mindeközben figyelembe vesszük a nyelvhez kapcsolható szintaktikai szabályok összességét, és mindezt pedig azért csináljuk, hogy a terminusainkkal, mondatainkkal szemantikalag (tehát jelentésüket tekintve) is megfelelő üzeneteket hozhassunk létre. Ebben az értelemben sokféle logikát hozhatunk létre, melyeket aztán különféle szempontok alapján természetesen tipizálhatunk, ezáltal "korlátozhatjuk" a használatba vehető logikák számosságát.

Minden logika számára fontos a **logikai kalkulus** (röviden: kalkulus) fogalma. A kalkulus azon univerzális szabályok (kijelentések, formulák, tételek) összefüggő rendszere, amelyek mindig igaznak bizonyulnak, bármilyen, az adott logikai nyelvhez tartozó kijelentést (formulát), predikátumot vagy nevet helyettesítünk beléjük. A kalkulus tehát egy axiómarendszer, amely az adott logikai nyelv lehetséges formulái közül részben azokat a logikai igazságokat írja le, amelyek tautológiák (mindig igazak), részben azokat a szabályokat adja meg, amelyek segítségével további formulákat lehet levezetni

[2] Meg kell itt jegyezzem, hogy a következtetési szabályok folyamatos alkalmazása másrészről az elméleti konzisztencia fenntartásához is hozzájárul.

[3] A rokonsági fogalmakról szóló tanulmányomat meghivatkozhatóvá kell tenni.

az adott nyelven. E kettősséggel voltaképpen azt ígéri minden logika, hogy a kalukulusokra támaszkodva minden további állítást, formulát ki lehet fejezni (le lehet vezetni), ami másképp szólva azt jelenti, hogy ami az az adott nyelven "elmondható", kifejezhető, az elmondható, kifejezhető csak a kalkulusba tartozó axiómák használatával – legfeljebb nagyon bonyolult, nagyon hosszú kijelentések, formulák megfogalmazására lehet szükség. Mivel azonban ez a bonyolultság megnehezíti a nyelv – emberek általi – használatát, ezért az adott nyelven nemcsak lehet, de kívánatos is új és új formulák, terminusok bevezetése, ami logikai értelemben redundáns, praktikus szempontból viszont gazdaságos. Említettem már azt a köznapi példát, miszerint hivakozhatnánk valakire úgy is, mint az 'apám apjának az apjának az apja', illetve mint az 'ükapám', mindkétszer ugyanarra a személyre mutathatnánk rá (tekintsünk most el az apai és anyai vonal különbségétől). Ha az első megoldást választom, akkor hosszabban kell beszélnem és nehezebben lehet befogadni az üzenetet, viszont ezen a nyelven nem kell új terminust bevezetnem, megtanulnom és megtanítanom. A második megoldás gyorsabb és könnyebben folytatható kommunikációt kínál, aminek az az ára, hogy új terminust kell a nyelvbe felvenni. Ennek az a hátránya, hogy az ember számára nyilván korlátos a befogadható terminusok száma.[4] Változik a helyzet akkor, amikor számítógépeket is bevonunk a kommunikációba, mert azok számára nem jelentkeznek ilyen feldolgozási problémák. A gépek számára sokszor elhagyható ez a terminológiai redundancia, bár azt mindig mérlegelni kell, hogy a gépek valamilyen ponton már emberekkel is kapcsolatba kerülnek is ilyenkor váltani kell az ember számára kényelmesebb redundáns "beszédmódra".

Minden logika fontos tulajdonsága a logika kifejező- vagy leíróereje, amivel azt fejezhetjük ki, hogy milyen típusú állításokat, fogalmakat tudunk megragadni az adott formális nyelv segítségével. Ez más szempontból azt jelenti, hogy az adott nyelven kifejezett elméletbe milyen ontológiai elköteleződések mentén veszünk fel új és új terminusokat (fogalmakat). Ha olyan logikánk van, amely csak mondatokat (kijelentéseket) tud kezelni, akkor tartalmilag nagyon keveset állíthatunk arról, mi van a világban, és természetesen nagyon keveset (gyakorlatilag semmit sem) tudunk elmondani arról, mit gondolunk a világról. Ha a rokonsági rendszerek világát akarjuk egy elméletben leírni, akkor fel kell vennünk pár alapfogalmat (GYEREKE, FIATALABB, FÉRFI, HÁZASTÁRSA), és ezek segítségével már minden ismert rokonsági fogalmat kifejezhetünk. Ám amíg nem vesszük be a modellünkbe a férfi kategóriáját (vagyis nem köteleződünk el az ember kettős szexuális nembelisége mellett), addig ugyan definiálni tudjuk a szülő, a testvér, a nagyszülő stb. fogalmait, de nem tudunk különbséget tenni az apa és anya, a fia és lánya, a húga és bátyja fogalmak között. Ez a problémakör kapcsolódik a logikai nyelv pontossága, finomsága, granularitása kérdéséhez, ami kijelöli a nyelv helyét az általánosság vs. specifikusság tengelyen.

A következtető rendszer használhatóságát felülről korlátozza a használt ontológia nyelvének kifejezőereje, az a tény, hogy au adott nyelv segítségével mennyire komplex összefüggések írhatóak le. Azonban minél nagyobb egy nyelv kifejezőereje, annál szűkebb az olyan kérdések köre, melyre (matematikailag) egyáltalán lehetséges a válaszadás. Így találni kell egy egészséges egyensúlyt, melynél már kellően bonyolult összefüggéseket le tudunk írni, ugyanakkor a kérdések megválaszolhatósága még nem reménytelen.

A formális nyelvek bonyolultsága, számíthatósága, eldönthetősége nagyon fontos prob-

[4] Ez a probléma nemcsak a formális, de a természetes nyelvek esetében is ugyanúgy (és ugyanazért) létezik.

lémakört alkot, a nyelvek használhatósága szempontjából is kiemelt kérdésnek számít, ám e könyv mondanivalója szempontjából elhagyhatónak tűnik, ezért nem tárgyalom.

Az általam használni kívánt logikák egymásra épülését két dimenzióhoz igazodva lehet kifejezni. Figyelni lehet arra, hogy *(i)* hanyadrendű logikáról van szó, illetve *(ii)* extenzionális vagy intenzionális logikát vizsgálunk-e [627]. Az extenzionális logikákat úgy veszem sorba, hogy kezdem a nulladrendű propozicionális logikával, majd áttérek az elsőrendű predikátumlogikára, valamint az utóbbi alá sorolható kategoriális logika néhány kérdésére, végül röviden kifejtem, hogy hogyan lehet definálni a másodrendű (magasabbrendű) logikákat, és miért fontos figyelembe venni azt, hogy magasabbrendű logikákról (és fogalmakról) van-e szó. Az intenzionális logikák közül csak a modális logikát veszem elő, azt is csak azért és olyan mélységben tárgyalom, hogy érthetővé tudjam tenni azt, miért van szükségünk a deontikus logikára a társadalmi cselekvések vizsgálata során.

A cselekvéselmélet tárgyalásához szükséges fogalmakat – ahol csak tudom – megpróbálom úgy bemutatni, hogy formális nyelven reprezentált terminusokkal fejezem ki a fogalom jelentését. Ezt úgy teszem, hogy megmutatom az új, éppen definiálandó fogalom helyét a már létező, addig már bemutatott fogalmi modell struktúrájában. Ezáltal – remélhetőleg – átláthatóbbá válik, hogy az éppen felvett új fogalomhoz milyen ontológiai elköteleződések szükségesek. Az egyes – jól elkülöníthető – elméleteket önálló logikákkal írom le. A cselekvéselmélet egész belül az alábbi logikákat (és a hozzájuk tartozó jelöléseket) fogom használni.

(AL.xx)	aggregációlogika (HUN)	aggregation logic (ENG)
(BA.xx)	preferenciaalgebra (HUN)	betterness algebra (preferential algebra) (ENG)
(BL.xx)	preferencialogika (HUN)	betterness logic (preference logic) (ENG)
(CL.xx)	választáslogika (HUN)	choice logic (ENG)
(DL.xx)	deontikus logika (HUN)	deontic logic (ENG)
(EL.xx)	episztemikus logika (HUN)	epistemic logic (ENG)
(FL.xx)	formális logika (HUN)	formal logic (ENG)
(GL.xx)	kategórialogika (HUN)	category logic (ENG)
(HL.xx)	halmazlogika (HUN)	set logic (ENG)
(IL.xx)	interakció logika (HUN)	interaction logic (ENG)
(ML.xx)	modális logika (HUN)	modal logic (ENG)
(PC.xx)	kijelentéskalkulus (HUN)	propositional calculus (ENG)
(PL.xx)	kijelentéslogika (HUN)	propositional logic (ENG)
(QL.xx)	predikátumlogika (HUN)	predicate logic (ENG)
(QC.xx)	predikátumkalkulus (HUN)	predicate calculus (ENG)
(RA.xx)	relációalgebra (HUN)	relation algebra (ENG)
(RL.xx)	relációlogika (HUN)	relation logic (ENG)
(SL.xx)	érzelemlogika (HUN)	sentiment logic (ENG)
(ZL.xx)	cselekvéslogika (HUN)	action, power-authority logic (ENG)

1.1 Metafogalmak

A formális nyelveket mindig a természetes nyelvek analógiájára építjük fel abban az értelemben, hogy szavakat (kifejezéseket, terminusokat, fogalmakat) definálunk, amelyekkel a világban létező vagy létezőnek hitt dolgainak valamely halmazára mutatunk rá,

illetve ezen szavak egymás mellé illesztésével (relációba állításával) mondatokat (kijelentéseket) állítunk elő azzal az igénnyel, hogy azokkal érvényes állításokat tegyünk a világról. Eközben észrevétlenül használunk olyan fogalmakat, amelyeket magától értetődőnek tekintünk, de nem – vagy csak ritkán – reflektálunk rá. A legfontosabbakat ezek közül bemutatom itt.

A nyelvhasználat során – a dolgokat kissé elegyszerűsítve – kétfajta dolgot teszünk. Egyfelől szószerű nyelvi jeleket képzünk, hogy azokkal létező vagy tételezett dolgok halmazára rámutathassunk, referálhassunk rájuk, a dolgokat reprezentálhassuk a "szavakkal". Ezekkel a szószerű nyelvi jelekkel kevés dolgot fejezhetünk ki, amikor használjuk őket. Talán csak annyit, hogy az adott nyelvi jelet kibocsátva (leírva, kimondva) tételezzük a világ bizonyos dolgainak létezését – legalább annyira, hogy aztán kijelentéseket tehessünk róluk. Az ilyen megnyilatkozások érvényességét csak annyiban lehet vitatni, hogy magát a létezési feltételt támadhatjuk, semmi mást. Akkor tudunk tartalmasabb, mélyebb és szerteágazóbb érvényességvizsgálatba keveredni, ha mondatszerű nyelvi jeleket képzünk. Ezt úgy tehetjük meg, ha a szószerű nyelvi jeleket kapcsolatba állítjuk egymással. Ez a kapcsolatba állítás (valamilyen reláció képzése) az a művelet, amelynek révén szószerű jelekből állításokat keletkeztethetünk, ezért a reláció a mondatszerű megnyilatkozások kulcsa. Mivel a szószerű jelek egymáshoz illesztését rengetegféleképpen tehetjük meg, a mondatszerű megnyilatkozások szintjén megnő a kifejezés gazdagsága, mélysége, szélessége és vitathatósága egyaránt. Amikor egy elméletet építünk fel, akkor ezt a kettősséget kell mindvégig szem előtt tartani.

A logika alapfogalma az állítás. Azt természetesen nem állítom, hogy ez a fogalom ne lenne reflektált a logika világában, de azt fontosnak tartom hangsúlyozni, hogy sokszor zavart okoz az a tény, hogy a logika nyelvek túlnyomó részében nincs önálló jele, reprezentációja ennek a fogalomnak.[5]

(FL.0) PROPOSITION $\equiv A, B, C$

Az **állítás** igazságra vonatkozó érvényességigénnyel fellépő, adott nyelv szabályai szerint érvényes nyelvi megnyilatkozást kibocsátó aktus eredménye. Nagyon sok szinonimája van (jelen írás keretein belül): **kijelentés**, **mondat**, **asszerció**, **propozíció**, **ítélet**, **affirmáció**, **formula**. A tényleges gyakorlatban az állítás fogalmának nincs külön jele. Az állítások logikai értéke igaz vagy hamis lehet, amit szöveges megjelenítésben 'igaz' vagy 'hamis' karakteres formák, táblázatokban a '0' és az '1' értékek feltüntetésével fejezhetünk ki.

Minden nyelvben fontos a tagadás fogalma, amely az állítás "párja". Leggyakrabban olyan műveletként definálják, amely valamely állításhoz képest előállít egy újabb állítást úgy, hogy az eredeti és az új állítás logikai értéke ellentétes egymással.[6]

(FL.1) NEGATION $\equiv \neg \equiv \sim$

A **tagadás** (negáció) az a nyelvi művelet, amely során valamely állítás logikai értékét ellentétére változtatjuk. A tagadás állításra irányul függetlenül attól, hogy egyszerű vagy összetett állításokról van-e szó.

[5] Frege, a fogalomírás és a modern logika egyik megteremtője még használt külön jelet (ítéletvonalat) az állításra, de a követői mind elhagyták ezt a szokást.

[6] Itt hivatkozni kell a metarelációs fejezet művelet fogalmára.

Az állítással, illetve annak tagadásával a mondatszerű nyelvi megnyilatkozásokat reprezentálhatjuk. Ahhoz, hogy a kijelentéseink belső tartalmát, tehát a szószerű jeleket is meg tudjuk ragadni, szükség van a terminus fogalmára. Minden szószerű nyelvi jelre, amelyekkel még nem tudunk állításokat kifejezni, ráhúzhatjuk a terminus kategóriáját. A terminussal a világ dolgaira mutatunk rá, és ezt a kapcsolatot (a terminusok és a világ között) egyfelől a terminus terjedelmébe sorolt dolgokkal, másfelől a terminus jelentésével jellemezhetjük.

(FL.2) TERM $\equiv a, b, c, x, y, z,$ **P,R,Q**

A **terminus** a világ adott dolgaiból álló halmaz létezését reprezentáló nyelvi megnyilatkozás.

A tagadás mellett használjuk gyakran a komplementálás fogalmát is, aminek szoros kapcsolata van ugyan a tagadás fogalmával, sokban hasonlít hozzá, de mégsem azonos vele.

(FL.3) COMPLEMENT $\equiv \neg \equiv \sim$

komplementálás az a logikai művelet, amely során a terminus terjedelme alapján vesszük az univerzális halmazhoz képest komplementer terjedelemmel rendelkező terminust.

A komplementaritás a halmazelmélet gyakran használt fogalma, és mivel valamely halmaznak az univerzumhoz képest vett kiegészítését jelenti, a komplementer halmaz képzése a tagadás műveletével mutat szoros rokonságot. Egy halmaz és annak komplementere úgy viszonyul egymáshoz, mint egy állítás a tagadásához. Előfordul, hogy a tagadás, negálás terminusait is a komplementálás értelmében alkalmazzák, és a **komplementerképzés** terminusa is szinoním a **komplementálás** terminusával, míg a komplementaritás fogalma az egész jelenségegyüttest írja le. A halmazelméletben a komplementer halmaznak (és így áttételesen a komplementálás műveletének) van saját jelölése, ez olykor felbukkan a formális nyelvekben is, de gyakraban nem alkalmazzák a tagadás műveletét jelző szimbólumo(ka) erre a műveletre.

A halmazelméleti eredet és analógia rámutat arra, hogy ott lehet alkalmaznunk a komplementálás fogalmát, ahol elemekkel, előfordulásokkal, terjedelemmel van dolgunk, ez pedig a terminusok világa. A komplementálás művelete a terminusok terjedelmébe sorolt előfordulások univerzumhoz képest vett kiegészítése. A tagadás művelet viszont – elsősorban – kijelentésekre vonatkozik, ennyiben szemben áll a komplementálás fogalmával.

A formális nyelv elemei

a nyelv szimbólumai
a nyelv **jól-formált formulái** (well-formed formulas)
a nyelv **interpretáció** (interpretation)
a nyelv **szemantikája** (semantics)
faktuális értékek, kiértékelés, igazság,
modell stuktúra struktúraosztály formula formulaosztály

Ha a logikai kalkulusok és a logikai rendszerek közti viszonyt vizsgáljuk, akkor a köztük levő megfelelésre vonatkozóan tehetünk állításokat. Ezt minősítést a **konzisztencia**, illetve az **inkonzisztencia** fogalmaival tehetjük meg (ami a kielégíthetőség, illetve a kielégíthetetlenség szemantikai fogalmával állítható párhuzamba).

(FL.4) $\text{CONSISTENT} \stackrel{\text{def}}{=}$

konzisztens (HUN) – consistent (ENG)

(FL.5) $\text{INCONSISTENT} \stackrel{\text{def}}{=}$

inkonzisztens (HUN) – inconsistent (ENG)

Ha van két következményfogalmunk, a szintaktikai és a szemantikai következmény reláció, akkor fel lehet tenni a kérdést, hogy milyen a két fogalom közti viszony. E viszony jellemzésére többféle fogalmat is használnak a logikában, de nekünk nem lesz mindegyikre szükségünk, csak a három legfontosabbat (a helyesség, a teljesség és az adekvátság kategóriáit) említem meg közülük [627, 128].

(FL.6) $\text{SOUND} \stackrel{\text{def}}{\Leftrightarrow} \Gamma \vdash A \Rightarrow \Gamma \models A$

helyes (HUN) – sound (ENG)

(FL.7) $\text{COMPLETE} \stackrel{\text{def}}{\Leftrightarrow} \Gamma \models A \Rightarrow \Gamma \vdash A$

teljes (HUN) – complete (ENG)

(FL.8) $\text{ADEQUATE} \equiv \text{SOUND} \wedge \text{COMPLETE} \stackrel{\text{def}}{\Leftrightarrow} \Gamma \vdash A \Leftrightarrow \Gamma \models A$

adekvát (HUN) – adequate (ENG)

Következtetések

Amikor egy nyelvi megnyilatkozásnak ismerjük valamilyen belső szerkezetét, akkor mindig felmerülhet az a kérdés, vajon értelmes-e és ha igen, milyen eredményeket produkál az, ha a belső szerkezetét (vagyis az összetevők egymáshoz való viszonyát) megváltoztatjuk. A későbbi tárgyalandó elméletekben is elő-előkerülnek ilyen-olyan változtatási, transzformációs műveletek, úgyhogy röviden érdemes kitérni erre a kérdéskörre. Már csak azért is érdekesek ezek a szabályok, mert közülük több is a *közvetlen következtetések* vagyis az egyetlen állításból levonható következtetések körébe tartozik, és mint ilyenek a következtetési szabályok egyik fontos csoportját jelentik.

Egy nyelvi megnyilatkozásnak kétféle értelemben lehet belső szerkezete. Egyfelől ha mondatszerű megnyilatkozásokat vizsgálunk, akkor a "mondatok" belső szerkezetét nem ismerhetjük, viszont elemezhetjük a mondatok összekapcsolását, amikoris az egyes mondatok mint összetevők már a megnyilatkozás mint egész részeiként, belső szerkezeti egységeiként értelmezhetők. Másfelől viszont tekinthetjük a mondatszerű megnyilatkozást (vagy csak annak egy részét) is olyan egységnek, amelynek a belső szerkezete érdekel minket, amikor a szószerű összetevőit és azok egymáshoz, illetve a mondategészhez (vagy csak a mondat egy részéhez) való viszonyát akarjuk megérteni.

Jelölési konvenciók

A tudományos diskurzusban fogalmakat, tételeket, elméleteket gyártunk, és igyekszünk minél pontosabban kifejezni ezeket. Ebebn nagy segítséget jelent minél több formalizmus alkalmazása. De még ha így teszünk, akkor is szavakkal kommunikálunk egymás felé, és ilyenkor mindig könnyen pongyolává válhatunk. Egyszerűen azért, mert amikor az esetleg egyértelműen formalizált fogalmainkat természetes nyelvi szinten fejezzük ki

(terminusok, kifejezések, nyelvi konstrukciók alkalmazásával), folyamatosan belebonyolódunk a szinonimitás és homonimitás hálójába. Még a tudományos kommunikációban is tele vagyunk hononím és szinoním szavakkal, kifejezésekkel. Ezen a nehézségen azzal próbálok felülkerekedni, hogy amikor egy-egy új fogalmat vezetek be, akkor mindig megadom azt, hogy hogy hivatkozom rá fogalomként, illetve hogy hivatkoznak rá a különböző tudományterületek képviselői, a különböző szerzők természetes nyelvi konstrukciókkal.

A fogalmakat KISEBB MÉRETŰ, VERZÁL BETŰTÍPUSSAL jelzem és angol nyelvű terminus reprezentálom. A fogalmi síkon megpróbálom teljesen kizárni a homonimitás lehetőségét, és bár ezen a szinten lehetségesnek vélem a szinonimitás jelenségét, de igyekszem ezeket is elkerülni. A természetes nyelvi konstrukciók, terminusok jelölésére – első és/vagy lényeges felbukkanásuk során – **vastagított, talpatlan betűt** használok. A konstrukciók szintjén mind a homonimitás, mind a szinonimitás jelensége gyakran előfordul. A logikai formalizmusokkal támogatott modellépítés egyik fontos célja pont az, hogy ahol csak lehetséges meg lehessen találni és rá lehessen mutatni olyan helyzetekre, amelyekben a tudományos kommunikációban szinoním vagy homoním terminusok használatát tapasztalhatjuk. Amikor a terminusok, konstrukciók szintjén az egyértelműsítés miatt ez szükségesnek látszik, akkor zárójelben megadom, hogy az adott konstrukció melyik fogalomhoz kapcsolódik. Ezt természetesen akkor teszem csak meg, amikor egy-egy új terminust (fogalmat) bevezetek a modell tárgyalásakor, illetve amikor a szövegkörnyezetből nem derül ki egyértelműen, hogy az éppen használt kifejezés melyik jelentéséről van szó. s Minden nyelvhez tartoznak központozást segítő jelek, amelyeket nem tüntetünk fel a nyelv definiálásakor, de mindvégig használjuk őket. Ilyenek a nyitó és csukó zárójelek (‘(’ és ‘)’), valamint az elemek felsorolását segítő vessző ‘,’ jele. Ugyancsak “univerzálisan”, azaz minden nyelvben használok két ekvivalenciajelet: egyfelől egy új fogalom bevezetésekor a definiálás jelzésére alkalmazom a ‘$\stackrel{\text{def}}{=}$’, illetve a $\stackrel{\text{def}}{\Leftrightarrow}$ **definíciós ekvivalencia** műveletjeleit (attól függően, hogy tárgynyelven vagy a trágynyelv és metanyelv közötti kapcsolatról van-e szó), másfelől a terminusok szintjén létező szinonimitás kezelésére használom a ‘$\equiv$’ **terminológiai ekvivalencia** műveletjelét.

1.2 Halmazelmélet

A halmazoknak, a halmazelméletnek kardinális szerepe van minden formális elméletben. Nincs szükség itt a halmazelmélet tárgyalására, de arra igen, hogy felsoroljam azokat a legfontosabb halmazelméleti fogalmakat (illetve ezek jeleit), amelyeket a könyvben mindvégig használni fogok.

(HL.0) $\mathscr{L}_{SL}(\mathbb{X}, \wp(\mathbb{X}), \varnothing, \mathbb{U}, \mathbb{N}, \mathscr{R}, \{a,b,c\}, \in, =, |\mathbb{X}|, \cup, \cap, \overline{\mathbb{X}}, \setminus, \subset, \subseteq)$, ahol

$\mathbb{X}$	tetszőleges halmaz
$\wp(\mathbb{X})$	$\mathbb{X}$ hatványhalmaz ($\mathbb{X}$ részhalmazainak halmaza)
$\varnothing$	üres halmaz
$\mathbb{U}$	univerzális (teljes) halmaz
$\mathbb{N}(\mathbb{N}^+)$	természetes számok halmaza (nullával együtt vagy nulla nélkül)
$\mathscr{R}$	valós számok halmaza
$\{a,b,c\}$	halmaz elemei
$\in$	eleme reláció
$=$	egyenlőség reláció
$\lvert\mathbb{X}\rvert$	halmaz számossága (kardinalitása) függvény
$\cup$	unióképzés halmazművelet
$\cap$	metszetképzés halmazművelet
$\overline{\mathbb{X}}$	komplementerképzés halmazművelet
$\setminus$	különbségképzés halmazművelet
$\subset$	valódi részhalmaza reláció
$\subseteq$	részhalmaza reláció

A halmazelméleti műveletek tulajdonságait, a köztük levő azonosságokat, a halmazelmélet paradoxonait nem mutatom be, de igazából nem is fogok ezekre támaszkodni a későbbiekben. Egy kivételt teszek csak itt azzal, hogy bevezetem a **partíció** (particionálás) fogalmát, mert erre mindvégig szükségem lesz. A halmazelméletben ezt a fogalmat akkor használjuk, amikor egy halmazt részhalmazokra akarunk bontani, miközben az az elvárásunk, hogy az részek ne legyenek átfedésben, és összeségükben fedjék le teljesen az eredeti teljes halmazt. Ha így osztunk fel egy halmazt, akkor beszélünk partícióról. Ezt a műveletet így lehet definiálni.

(HL.1) PARTITION$(\mathbb{X})$

1) $\forall i \in \mathbb{N}(\mathbb{S}_i \subset \mathbb{X}) \wedge \mathbb{S}_i \in \wp(\mathbb{X}))$
2) $\varnothing \notin \wp(\mathbb{X})$
3) $\bigcup_{\mathbb{S}_i \in \wp} \mathbb{S}_i = \mathbb{X}$
4) $\forall i \forall j((\mathbb{S}_i, \mathbb{S}_j \in \wp(\mathbb{X}) \wedge \mathbb{S}_i \neq \mathbb{S}_j) \rightarrow \mathbb{S}_i \cap \mathbb{S}_j = \varnothing)$

A fenti formulák segítségével biztosíthatjuk, hogy az ($\mathbb{X}$) halmaz olyan részhalmazokra legyen felosztva, amelyek egymástól teljesen elkülönülnek, viszont együttesen kiadják az egész halmazt. Lehetnek, akiknek furcsának tűnhet, mégis megkockáztatom, hogy a megfelelő particionálás a megalapozott fogalmi gondolkodás egyik legfontosabb biztosítéka (↬58.o.). A későbbiekben erre a kijelentésemre még visszatérek.

——————— hatványhalmaz powerset
mégiscsak kéne a halmazkalkulus is
a műveleteket nagyon be kéne mutatni

(HL.2) UNION $\equiv \cup \equiv \bigcup \stackrel{\text{def}}{=}$

(halmaz)unió (HUN) – (set) union (ENG)

(HL.3) INTERSECTION $\equiv \cap \equiv \bigcap \stackrel{\text{def}}{=}$

(halmat)metszet (HUN) – (set) intersection (ENG)

(HL.4) COMPLEMENT $\equiv\equiv \stackrel{\text{def}}{=}$

komplementum (HUN) – complement(ENG)

(HL.5) $? \equiv \setminus \stackrel{\text{def}}{=}$

(halmaz)kivonás (HUN) – (set) (ENG)

(HL.6) PROPER_SUBSET $\equiv \subset \stackrel{\text{def}}{=}$

(valódi) részhalmaza (HUN) – (proper) subset (ENG)

(HL.7) SUBSET $\equiv \subseteq \stackrel{\text{def}}{=}$

részhalmaza (HUN) – subset (ENG)

1.3 Kijelentéslogika

A logikai elméletek közül elsőként a nulladrendű extenzionális logikát veszem elő [627]. A későbbiekben inkább **kijelentéslogikaként** fogok hivatkozni rá, de ebben a könyvben szinonímaként kezelem még az 'állításlogika', illetve a 'propozicionális logika' terminusokat is. A kijelentéslogika a mondatszerű nyelvi megnyilatkozásokra fókuszál azáltal, hogy a mondatokba foglalt elemi állítások egymáshoz fűzésének lehetőségeit, szabályait vizsgálja. Ebből következően a kijelentések belsejével, összetevőivel, vagyis a szószerű nyelvi megnyilatkozásokkal nem foglalkozik. A kijelentéslogika lényegét a kijelentéseket összekapcsoló logikai műveletek adják, ezért először ezeket kell felvenni a fogalmi modellünkbe. A következő fejezetben ezeket veszem sorba. Előtte azonban rögzítenem kell azt, milyen nyelvre van mindehhez szükség.

(PL.0) $\mathscr{L}_{PL}(A, B, C, \neg, \rightarrow)$, ahol

A, B, C	propozíció formula
$\neg$	tagadás logikai művelet
$\rightarrow$	kondicionális logikai művelet

A propozicionális logika kevés ontológiai elkötelezettséget kíván meg, hiszen, láthatjuk, csak a kijelentések és a pár logikai művelet fogalmára van szükség hozzá. A kijelentéslogikát fel lehet építeni a kijelentéskalkulusra támaszkodva, de mielőtt ez utóbbit bemutatnám, egy hosszabb kitérőt teszek a logikai műveletek világába.

1.3.1 Kijelentéslogikai műveletek

Amennyiben veszünk állításokat és ezeket valamilyen művelet segítségével összekapcsol-juk, akkor egy újabb kijelentéshez jutunk, és a kijelentéslogika azt vizsgálja, hogy ezeknek az új állításoknak milyen lesz az igazságértéke. A modellünkbe fel kell vennünk a logikai művelet fogalmát,[7] ami kijelentések, állítások összekapcsolására vonatkozik. Mivel a fogalom nagyon széles körben használt, rengeteg szinonimája termelődött ki.

(PL.1) LOGICAL OPERATION

logikai művelet (HUN) – funktor (HUN) – junktor (HUN) – logikai kapcsoló (HUN) – logikai konnektívum (HUN) – konnektívum (HUN) – mondatoperátor (HUN) – propozicionális kapcsoló (HUN) – mondatkapcsoló (HUN) – propozicionális operátor (HUN) – logical operation (ENG) – logical connective (ENG)

A kijelentéslogikában a műveletek változóit, összetevőit, operandusait az ítéletek alkotják. Ha nem okoz zavart, akkor az ítéletekre mint a műveletek változóira operandusként fogok hivatkozni. Ha a kijelentéslogikában igazából csak azt vizsgálhatjuk, hogy milyen módon lehet összekapcsolni a kijelentéseket, és ezekben az összekapcsolásokban milyen szabályszerűségeket találhatunk, akkor kiemelt kérdés, hogy hányféleképpen és hogyan kapcsolhatjuk össze a kijelentéseket egymással. A válasz rászben függ attól, hogy milyen műveletről is beszélünk. Nulláris, azaz nullaváltozós műveletből kettő, unáris, azaz egyváltozós műveletből négy, míg bináris, azaz kétváltozós műveletből tizenhat van.

A két nulláris művelet, a tautológia és az ellentmondás jellemzője az, hogy igazságértékük független az általuk összekapcsolt állítások igazságértékeitől. A tautológia mindig igaz, az ellentmondás mindig hamis. Ezért is nevezzük nullaváltozós műveletnek őket, mivel "érzéketlenek" a bennük levő kijelentések mint változók értékeire.[8]

Az unáris műveletek értelmezésekor gyakran mondják azt, hogy a két nulláris művelet, a tautológia és az ellentmondás mellett ide tartozik még az állítás, a tagadás. A tagadásról természetesen lehet azt mondani, hogy az egy unáris művelet, aminek a bemenete egy állítás, és a kimenete egy – az állítással ellentétes igazságértékű – másik állítás. Mindig szem előtt kell azonban tartanunk azt a tényt, hogy a tagadás műveletét tetszőleges – tehát bármilyen összetettségű – állításra lehet vonatkoztatni. Beszélhetünk műveletek által előállított, összetett kijelentések tagadásáról is. A tagadás mindig egyváltozós művelet, de ettől még lehet többargumentumú. Ugyanaz igaz az állításra is természetesen, csak sokkal kevésbé tűnik intuitívnek azt mondani, hogy az állítás maga unáris művelet. Emiatt a fogalmi "kavarodás" miatt én ezen a ponton a 'projekció', illetve ennek tagadásaként a kontraprojekció fogalmát alkalmazom. A projekció olyan kétargumentumú, de egyváltozós művelet, amely a két argumentum közül csak az egyik értékeit veszi figyelembe, tehát csak attól az egy – projektált – változótól függ. A tizenhat bináris logikai művelet közé sorolhatjuk természetesen a négy unáris műveletet is, és ezeken túl van még 12 "szabályos", vagyis két-argumentumú és kétváltozós művelet.

A műveletek jellemzésekor bemutatom azokat a logikai jeleket, amelyeket valamilyen formális logikai nyelvben elterjedt módon használnak, azokat a – magyar – ter-

[7] A művelet fogalmát itt nem definiáljuk, metanyelvi fogalomként használjuk. A tárgynyelvben használt művelet fogalmáról lásd a relációról szóló fejezetet (↬42.o.).

[8] Mivel az alacsonyabb változószámú műveletek mindig besorolhatóak a magasabb változószámú műveletek közé, ezért a nullaváltozós műveletek egyben egy- és kétváltozós műveletnek is tekinthetők.

mészetes nyelvű szavakat, kifejezéseket, amelyeket a fontosabb elméletekben gyakrabban felbukkantak, megmutatom az adott művelet ellentétét, valamint megadok néhányat a műveletet leíró formulák közül. Végül minden leírás mellett feltüntetek egy ábrát, ami a műveleti igazságtábla geometriai reprezentációjának tekinthető

Minden művelet esetében igaz, hogy a különböző logikai nyelveken olykor más műveleti jeleket alkalmaztak, ezért még itt is, a formális nyelvek világában is létezik a szinonimitás jelensége (sőt, a homonimitás is megfigyelhető, mert elő-előfordul, hogy ugyanazokat a logikai jelekkel más műveleteket reprezentálnak). Még gyakoribb azonban a szinonimitás (és homonimitás) előfordulása a formális elméleteket körbevevő, azokat interpretáló természetes nyelvek területén (főként, ha ide soroljuk a különböző természetes nyelvek által használt terminusrétegeket is). A különböző kontextusokban kitermelődő alternativitás, illetve stilisztikai kényszerek miatt igencsak kiterjedt szinonimitással (homonimitással) kell együtt élnünk.

A műveletekre vonatkozóan sokféle megnevezés verseng egymással, különböző kontextusokban más és más terminusokat használnak ugyanarra a műveletre. A műveletek leírásakor felsorolom majd a leggyakrabban használt neveiket. Amikor döntenem kellett, ezek közül melyiket válasszam magam számára, akkor azt a szempontot vettem figyelembe, hogy a megnevezések önmagukban is minél jobban utaljanak a köztük levő kapcsolatokra. Ezért – ahol csak lehetett – az egymásból származtatható műveletek esetében azokat a neveket választottam, amelyek utaltak két művelet közti konverziós kapcsolatra vagy a köztük levő tagadási viszonyra. Előbbit a ‘retro’, utóbbit a ‘kontra’ előtag használata jelzi. Nem lehettem azonban teljesen következetes ezen a téren, mert bizonyos relációk esetében a széles körben használt terminusokat nem lehetett nem figyelembe venni. Ilyen esetekben “önkényesen” választottam a versengő nevek közül.

Logikai műveletek

A tizenhat bináris logikai műveletet úgy mutatom be egymás után, hogy megadom a cselekvéselméleti modellben használt nevüket, a gyakran használt megnevezéseiket, valamit adok egy rövid szöveges jellemzést is rájuk. Megmutatom azt is, hogy miként lehet őket definálni más műveletek segítségével, de itt nem törekszem arra, hogy ugyanazt a logikai definíciót adjam meg, amivel a kijelentéskalkulusban meghatározhatjuk őket. A cél itt az, hogy értelmezni lehessen az egyes műveleteket. Akkor hát, vegyük sorba őket!

(PL.2) $\text{CONTRADICTION} \equiv \bot \stackrel{\text{def}}{=} A \wedge \neg A$

ellentmondás (HUN) – kontradikció (HUN) – sosem igaz (HUN) – falsum (LAT) – contradiction (ENG)

Az ellentmondás művelete egyfajta konzisztenciát biztosít a kijelentéseinknek abban az értelemben, hogy létével jelzi és kizárja az egymást tagadó kijelentések együttes fennállását. Mondhatjuk azt is, hogy ‘Esik az eső.’ és azt is, hogy ‘Nem esik az eső.’, de ha a kettőt egyszerre állítjuk, akkor ellentmondáshoz jutunk. Az ellentmondás művelete mint két – egymást tagadó – állítás összekapcsolása sosem lehet igaz. Ez a művelet nulláris, konstans.

(PL.3) $\text{CONEGATION} \equiv \not\vee \equiv \downarrow \equiv \ \nabla \ \equiv || \stackrel{\text{def}}{=} \neg A \wedge \neg B$

konnegáció (HUN) – sem-sem (HUN) – NOR (LOG) – not OR (LOG) – Webb-művelet (HUN) – egyesített tagadás (HUN) – conegation (ENG) – Webb-operation (ENG)

A konnegáció a diszjunkció tagadása. Ha a hétköznapi nyelvben is használatos ‘sem-sem’ formulát nézzük (‘Sem utódja, sem boldog őse.’), ezen az alakon látszik csak igazán, hogy ez a művelet "közel áll" az ellentmondáshoz, hiszen igazságtáblázataik csak egyetlen esetben térnek el egymástól: amikor mindkét operandus hamis, akkor a konnegáció igaz, míg a kontradikció hamis. Ez egyfajta szimmetricitást kölcsönöz a műveletnek. Az ellentmondáshoz való “közelség” azonban kicsit látszólagos – abban az értelemben legalábbis, hogy a művelet két összetevőjének együttes hamissága a logikai univerzum jóval nagyobb részét fedi le, mint az összes többi opció.

(PL.4) $\text{RETROSUBTRACTION} \equiv \nleftarrow \stackrel{\text{def}}{=} \neg A \wedge B$

retroszubtrakció (HUN) – konverz szubtrakció (HUN) – konverz nem-implikáció (HUN) – retrosubtraction (ENG) – converse subtraction (ENG)

Ez a művelet a szubtrakció megfordítása, konverze. A művelet jellemzésére ugyanazok mondhatók el, mint a szubtrakció esetében, a különbség csak a két operandus közti viszony irányultságában van.

(PL.5) $\text{CONTRAPROJECTION} \equiv \not\lhd \stackrel{\text{def}}{=} \neg A$

kontraprojekció (HUN) – contraprojection (ENG)

Ez a művelet a projekció ellentéte, tartamilag megegyezik az A operandusba foglalt kijelentés tagadásával. Azért nevezem kontraprojekciónak, mert a kontra előtaggal valaminek az ellentétét szokás kifejezni, és ebben az esetben pont erről, a projekció tagadásáról van szó.

(PL.6) $\text{SUBTRACTION} \equiv \not\rightarrow \stackrel{\text{def}}{=} A \wedge \neg B$

szubtrakció (HUN) – kivonás (HUN) – nem-kondicionális (HUN) – abjunkció (HUN) – subtraction (ENG) – abjunction (ENG) – non-conditional (ENG)

A szubtrakció a kondicionális művelet tagadása, amely csak akkor igaz, amikor A “önmagában” igaz, vagyis amikor B nem igaz, de A igaz. Ez utóbbi összefüggés magyarázza a “kivonás” megnevezést, hiszen a művelet igazságtartományát úgy kapjuk meg, ha A igazságtartományából kivonjuk B igazságtartományát.

(PL.7) $\text{CONTRA-RETROPROJECTION} \equiv \not\rhd \stackrel{\text{def}}{=} \neg B$

kontra-retroprojekció (HUN) – contra-retroprojection (ENG)

Ez a művelet a retroprojekció ellentéte, tartamilag megegyezik a B operandus tagadásával.

(PL.8) BISUBTRACTION $\equiv \nleftrightarrow \equiv \oplus \equiv \veebar \stackrel{\text{def}}{=} \neg(A \leftrightarrow B) \equiv (A \wedge \neg B) \vee (\neg A \wedge B)$

biszubtrakció (HUN) – exkluzív diszjunkció (HUN) – kizáró diszjunkció (HUN) – erős diszjunkció (HUN) – exkluzív OR (LOG) – XOR (LOG) – vagy-vagy (HUN) – kizáró vagy (HUN) – kontravalencia (HUN) – antivalencia (HUN) – kontradiktórius (HUN) – bisubtraction (ENG) – strong disjunction (ENG) – contravalence (ENG) – antivalence (ENG)

A biszubtrakció művelete a bikondicionális tagadása, de tartalmilag kapcsolatba hozható a diszjunkció műveletével is, amit félreérthetetlenül jelez több névváltozata is. Tételezhető még a szubtrakció és annak konverze, a retroszubtrakció diszjunkciójaként is. A művelet legfontosabb tulajdonsága az, hogy csak akkor igaz, ha a két operandus közül valamelyik igaz, de úgy, hogy egyszerre sosem lehetnek igazak. Ezt fejezi ki a kizáró vagy exkluzív diszjukció megnevezés. A mindennapi beszédben is gyakran használjuk ezt a műveletet. Gondoljunk csak a "csak egy maradhat", a "pénzt vagy életet" vagy a "mindent vagy semmit" szófordulatok felbukkanására. A művelet költői leírása pedig:

> „Vagy bolondok vagyunk s elveszünk egy szálig,
> Vagy ez a mi hitünk valóságra válik."

(PL.9) EXCLUSION $\equiv \not\wedge \equiv \uparrow \equiv | \equiv || \stackrel{\text{def}}{=} \neg A \vee \neg B$

exklúzió (HUN) – kizárás (HUN) – NAND (LOG) – not AND (LOG) – nem-és (HUN) – Sheffer-művelet (HUN) – vagylagos tagadás (HUN) – alternatív tagadás (HUN) – kontrárius (HUN) – exclusion (ENG) – Sheffer-operation (ENG)

Az exklúzió a konjunkció tagadása. Bizonyos értelemben ez a művelet van legközelebb a tautológiához, hiszen csak a két operandus egyidejű igazsága esetén hamis, minden más esetben igaz. A két összetevő, A és B együttes fennállása erős feltételt jelent, és amikor ez nem teljesül, tehát az esetek nagyobb részében, a művelet igaz. Lehet valami barna, lehet zöld, lehet egyik sem, de egyszerre nem lehet barna is, zöld is.

(PL.10) CONJUNCTION $\equiv \wedge \equiv \& \stackrel{\text{def}}{=} A \wedge B$

konjunkció (HUN) – AND (LOG) – és (HUN) – conjunction (ENG)

A konjunkció abban az értelemben erős állítás (legalábbis a diszjunkcióhoz képest), hogy a művelet igazsága egyszerre függ a két operandus igazságától, vagyis ha fennáll a konjunkció igazsága, akkor tudhatjuk azt is, hogy ez annak "köszönhető", hogy mindkét operandus egyszerre igaz. A művelet által lefedett esetek tartománya kisebb, de az így kifejezett tudásunk "tartalmasabb", hiszen kér dolgot is tudunk (A és B igazságát). Ez az operáció is a széleskörben használt műveletek közé tartozik.

(PL.11) BICONDITIONAL $\equiv \leftrightarrow \equiv \Leftrightarrow \equiv \text{`} \equiv \prime \stackrel{\text{def}}{=} A \leftrightarrow B \equiv (A \wedge \neg B) \vee (\neg A \wedge B)$

bikondicionális (HUN) – ekvivalencia (HUN) – materiális ekvivalencia (HUN) – biimplikáció (HUN) – XNOR (LOG) – akkor és csak akkor (HUN) – csakkor (HUN) – a.cs.a (HUN) – iff (LOG) – biconditional (ENG) – equivalence (ENG) – material equivalence (ENG) – biimplication (ENG)

A bikondicionális művelete a kondicionális és annak konverze, a retrokondicionális egyidejű fennállását jelenti, azaz az implikatív jelleg mindkét irányban érvényes. Fontos művelet a tudományos gondolkodás számára, de a hétköznapi kommunikációban is gyakran előfordul. Példa lehet rá a következő idézet: „A garancia akkor és csak akkor érvényes, amennyiben a terméket az adott termékre vonatkozóan jóváhagyott és a használati útmutatóban feltüntetett környezetben használják."

(PL.12) $\text{RETROPROJECTION} \equiv \rhd \stackrel{\text{def}}{=} B$

retroprojekció (HUN) – második projekció (HUN) – jobb projekció (HUN) – right projection (ENG)

A retroprojekció csak annyiban tér el a projekciótól, hogy a két operandus közül nem a bal-, hanem a "jobboldalit" emeli ki, vagyis A és B közül B-t tételezi, míg A "állításától" eltekint. Ebben az értelemben a két műveletet egymás konverzeiként értelmezhetjük. A 'retroprojekció' megnevezés analóg a retrokondicionális név képzésével: a 'retro' és a 'kondicionális' szavak összetételében az előtaggal a megfordítást, a konverzitást fejezzük ki, amit – értelemszerűen – az utótagra vonatkoztatunk. A retroprojekció tartalmilag megegyezik a jobboldali operandus (vagyis B) "egyedüli" állításával.

(PL.13) $\text{CONDITIONAL} \equiv \rightarrow \equiv \supset \equiv \Rightarrow \stackrel{\text{def}}{=} \neg A \vee B$

kondicionális (HUN) – materiális kondicionális (HUN) – implikáció (HUN) – materiális implikáció (HUN) – ha, akkor (HUN) – szubaltern (HUN) – conditional (ENG) – implication (ENG) – material implication (ENG) – if, then (ENG)

A kondicionális a hétköznapi kommunikáció talán legfontosabb művelete. Az indoklásainkban, magyarázatainkban gyakran használjuk ezt a műveletet, amely a hipotetikus állítások mintázatát adja. A hétköznapi praxis nem igazán reflektál arra a filozófiai problémára, amely abból fakad, hogy A hamissága esetén a művelet mindig igaz eredményt ad.

(PL.14) $\text{PROJECTION} \equiv \lhd \stackrel{\text{def}}{=} A$

projekció (HUN) – első projekció (HUN) – balprojekció (HUN) – projection (ENG) – left projection(ENG)

Ritkán szokták a projekció műveletét ebben a kontextusban használni, helyette valamelyik operandus mint állítás tételezését alkalmazzák – miközben ezt a tételezést unáris műveletként értelmezik. Ez természetesen sok szempontból elfogadható, azonban bináris műveletek teljeskörű bemutatásakor jobbnak tartom azt a megoldást, amikor nem emelünk be unáris műveleteket az elemzésünkbe, csak bináris műveletekkel foglalkozunk. A projekció az a művelet, amely két operandus közül kiemeli az egyiket úgy, hogy a művelet igazságértékei minden ponton megegyeznek a kiemelt állítás igazságértékeivel. Ez természetesen azt is jelenti, hogy a másik operandust ilyenkor nem vesszük figyelembe, úgy tételezzük, mintha nem is létezne, nem is állítanánk.[9] Amennyiben a bal oldali állítást

[9] A projekció hasonlít a reifikáció műveletére, hiszen mindkét esetben kiválasztunk egyet két összekapcsolt elemből: logikai műveletek esetén az operandusok, reifikáció esetén a relátumok közül (↬60.o.).

(tehát A-t) projektáljuk, akkor "egyszerű" projekcióról beszélünk. Ez tartalmilag megegyezik a bal oldali propozíció tételezésével.

(PL.15) retroconditionalRETROCONDITIONAL $\equiv \leftarrow \equiv \subset \equiv \Leftarrow \stackrel{\text{def}}{=} A \vee \neg B$

retrokondicionális (HUN) – konverz kondicionális (HUN) – konverz implikáció (HUN) – retroconditional (ENG) –converse conditional (ENG) – converse implication (ENG)

Ez a kondicionális művelet konverze. Tartamilag, formailag mindenben hasonlatos a kondicionálisra, csak éppen minden fordítva, "jobbról-balra van". A megszokás (vélhetőleg a balról-jobbra olvasás dominanciája) miatt a kondicionálist érezzük természetesnek. Enek ellenére a hétköznapi beszédben is előfordul. Példa rá a következő két kifejezés: „Kétszer ad, ki gyorsan ad.", illetve „Akkor szép az erdő, mikor zöld."

(PL.16) DISJUNCTION $\equiv \vee \stackrel{\text{def}}{=} A \vee B$

diszjunkció (HUN) – OR (LOG) – vagy (HUN) – alternáció (HUN) – megengedő vagy (HUN) – gyenge diszjunkció (HUN) – inkluzív diszjunkció (HUN) – disjunction (ENG) – alternation (ENG) – inclusive disjunction (ENG) –

A diszjunkció gyenge művelet, amely két állítást úgy kapcsol össze, hogy a művelet igazsága fennállásához "elegendőnek tartja" az egyik operandus igazságát. Abban az értelemben kell ezt gyengének tartanunk, hogy a művelet eredményeként nem tudjuk, hogy a világról szóló igaz tudásunk mennyiben köszönhető a két operandus által kifejezhető igaz tudásnak. Kicsit másként: ha csak a diszjunkció igazságát ismerjük, nem tudhatjuk, hogy ez A vagy B, esetleg mindkettő egyifejű igazságának köszönhető-e. A diszjunkció a hétköznapi kommunikáció során leggyakrabban használt műveletek közé tartozik.

(PL.17) TAUTOLOGY $\equiv \top \stackrel{\text{def}}{=} A \vee \neg A$

tautológia (HUN) – mindig igaz (HUN) – tautology (ENG)

A tautológia az a művelet, amely úgy kapcsol össze két állítást, hogy azok együttese mindig igaz legyen. Ebben a minőségében a kontradikció ellentéte. Ha azt mondjuk, 'Esik az eső, vagy nem esik az eső.', akkor ez mindképpen igaz. Persze, kérdés, hogy mi az értelme vagy inkább mi a haszna az ilyen műveletnek. A példamondat alapján nem igazán érezni, pedig igenis nagy jelentősége van a tautológiáknak: a logikai következtetési szabályaink nagy részét ezek adják. A művelet nulláris, konstans.

A tizenhat bináris művelet áttekintése után érdemes őket egy táblázatban egyszerre felsorolni (az ellentéteikkel és igazságtáblázataikkal együtt).
A műveletek igazságtáblázatai alapján észrevehetünk pár érdekességet. Nyolc olyan művelet van, amely a két összetevő igazsága esetén igaz, és nyolc olyan, amely ilyenkor hamis. Nyolc művelet másik nyolcnak a tagadása. Az egymás tagadásaként tekinthető

művelet	jele	formula	11	10	01	00	ellentéte
ellentmondás	$\bot$	$A \wedge \neg A$	0	0	0	0	tautológia
konnegáció	$\not\vee, \downarrow, \nabla, \|\|$	$\neg A \wedge \neg B$	0	0	0	1	diszjunkció
retroszubtrakció	$\nleftarrow$	$\neg A \wedge B$	0	0	1	0	retrokondicionális
kontraprojekció	$\not\lhd$	$\neg A$	0	0	1	1	projekció
szubtrakció	$\nrightarrow$	$A \wedge \neg B$	0	1	0	0	kondicionális
kontra-retroprojekció	$\not\rhd$	$\neg B$	0	1	0	1	retroprojekció
biszubtrakció	$\nleftrightarrow, \oplus, \veebar$	$(A \wedge \neg B) \vee (\neg A \wedge B)$	0	1	1	0	bikondicionális
exklúzió	$\not\wedge, \uparrow, \mid, \|\|$	$\neg A \vee \neg B$	0	1	1	1	konjunkció
konjunkció	$\wedge$, &	$A \wedge B$	1	0	0	0	exklúzió
bikondicionális	$\leftrightarrow, \Leftrightarrow, \equiv$	$(\neg A \vee B) \wedge (A \vee \neg B)$	1	0	0	1	biszubtrakció
retroprojekció	$\rhd$	B	1	0	1	0	kontra-retroprojekció
kondicionális	$\rightarrow, \supset, \Rightarrow$	$\neg A \vee B$	1	0	1	1	szubtrakció
projekció	$\lhd$	A	1	1	0	0	kontraprojekció
retrokondicionális	$\leftarrow, \subset, \Leftarrow$	$A \vee \neg B$	1	1	0	1	retroszubtrakció
diszjunkció	$\vee$, v	$A \vee B$	1	1	1	0	konnegáció
tautológia	$\top, \vDash$	$A \vee \neg A$	1	1	1	1	ellentmondás

műveletek igazságértékei olyanok, hogy minden ponton egymás ellentétei. Ez szemléletesen megnyilvánul abban a mintázatban, hogy ahol az egyik művelet igaz, ott a másik hamis. Nyolc szimmetrikus és nyolc aszimmetrikus művelet van. A nyolc aszimmetrikus műveletből négy a maradék négy művelet konverzének tekinthető. Nyolc olyan művelet van, amelynél mindkét operandus hamissága esetén a művelet egésze is hamis, míg a másik nyolcművelet esetében a művelet egésze ilyenkor igaz. A logikai műveletek struktúráját, az egyes műveleteken belül az összekapcsolt kijelentések egymáshoz, illetve a tárgyalási univerzumhoz való viszonyát, valamint az egyes logikai műveletek egymáshoz való viszonyát jobban megérthetjük, ha a műveleteket ábrázoló geometriai ábrákat egyszerre, egymás mellé felrajzoljuk (lásd 1.1. ábra).

1.3.2 Kijelentéskalkulus

A kijelentéslogika a propozíciók és a köztük levő logikai műveletek, logikai kapcsolók struktúrájával foglalkozik. E struktúra alapját a **kijelentéskalkulus** (Propositional Calculus, PC) adja. A propozícionális kalkulus axiómakészletének definiálásához nyilván kevesebb elemre van szükség, ezért a PC nyelve szűkebb a kijelentéslogika nyelvénél. A PC logikája a következő nyelvi alapelemekre építhető fel.

(PC.0) $\mathscr{L}_{PC}(A, B, C, \neg, \rightarrow)$, ahol

A, B, C	formula (propozíció)
$\neg$	tagadás logikai művelet
$\rightarrow$	kondicionális logikai művelet

A fenti nyelv segítségével már definiálhatunk egy axiómkészletet, amiből a kijelentéslogika minden további fogalma, tétele levezethető. A PC a következő tételekből áll:

(PC.1) $A \rightarrow (B \rightarrow A)$

(PC.2) $(A \rightarrow B) \rightarrow ((A \rightarrow (B \rightarrow C)) \rightarrow (A \rightarrow C))$

(PC.3) $(\neg A \rightarrow \neg B) \rightarrow (B \rightarrow A)$

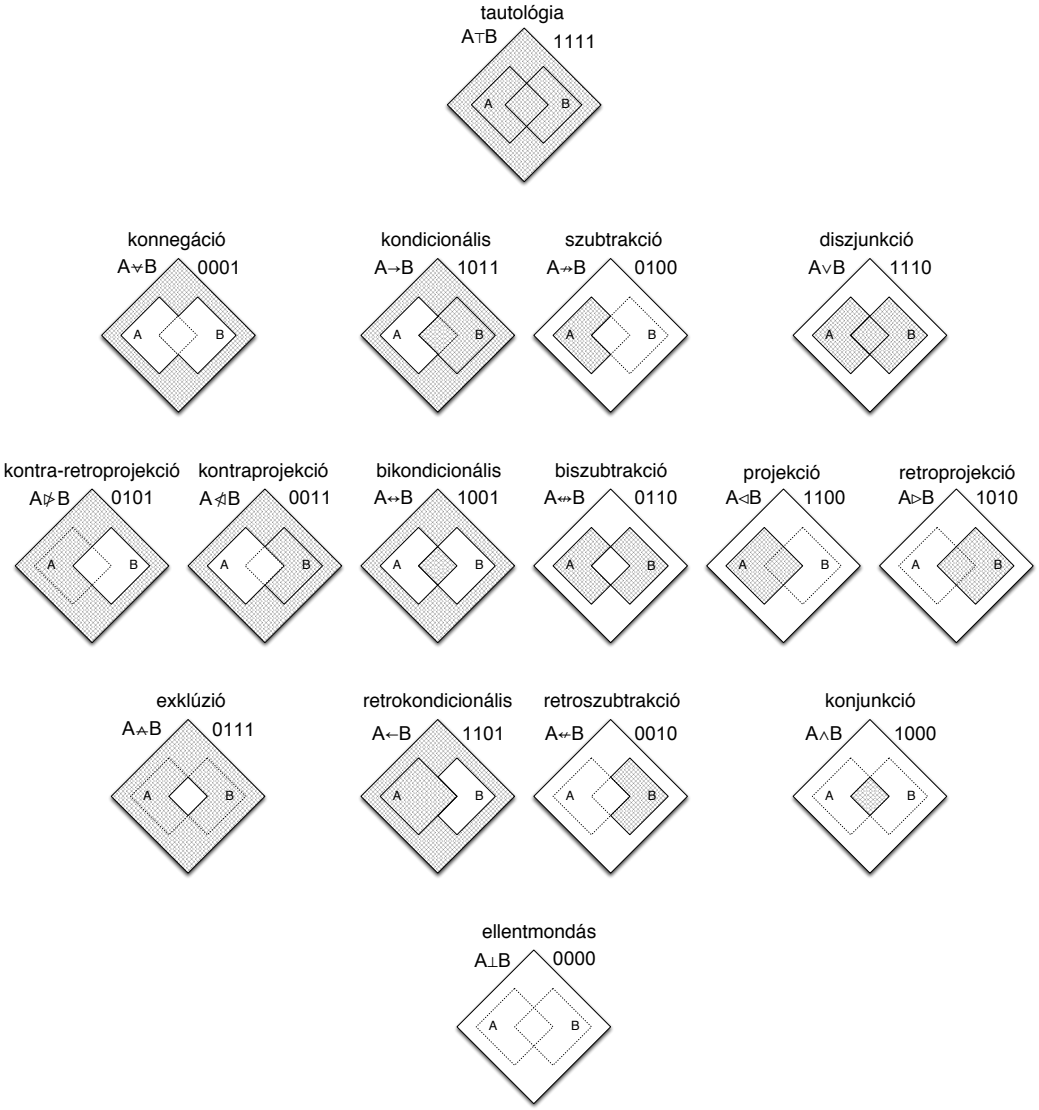


Figure 1.1: logikai műveletek

Ezek az axiómák a logikai műveletek (kapcsolók) bevezetésére és eliminálására szolgálnak. A logika teljessé tételéhez szükség van következtetési szabályokra is. Ezek olyan formulák, melyek esetében igaz formulákat a változókba írva igaz kijelentésekhez jutunk (ez kevesebb, mint az axiómasémák igazsága, mert utóbbiak esetében bármit írhatunk a változókba, mindig igaz marad a séma formulája. Ezeket az összefüggéseket mindig a logikai következmény relációk segítségével tudjuk kifejezni, amelyeket az alábbi szimbólumokkal jelölünk.

(PC.4) LOGICAL_CONSEQUENCE_RELATION $\equiv \Rightarrow \underline{\vee} \Leftarrow \underline{\vee} \Leftrightarrow$

logikai következmény relációk (HUN) – logical consequence relations (ENG)

A három szimbólumot értelemszerűen az egyik, a másik vagy a mindkét irányból vett következmény fennállásának jelzésére használjuk. A logikai következmény reláció nem az adott nyelvhez tartozik, a nyelvhez képest metanyelvi relációnak kell tartanunk. A (PC) következtetési szabálya a modus ponens (MP), ami ami szűkebb értelemben véve nem tartozik a kalkulushoz, de ettől még nélkülözhetetlen.

(PC.5) $(\mathrm{MP}) \stackrel{\mathrm{def}}{=} \{A, (A \to B)\} \Rightarrow B$

modus ponens következtetési szabály (HUN) – modus ponens inference rule (ENG)

A modus ponens úgy értelmezhetjük, hogy ha egy A kijelentés és a belőle vont bármely B utótagú $(A \to B)$ implikáció része igaz az adott nyelven, akkor a kondicionális B utótagja is a igaz. Vannak más következtetési szabályok is, de ezek mind levezethetők a sémaaxiómák és a modus ponens szabály formulái alapján. Ezekre támaszkodva az összes ismert összefüggést, következtetési sémát le lehet vezetni.[10] Csak a szemléltetés kedvéért álljon itt pár további következtetési szabály.

(PC.6) $A \vee B \stackrel{\mathrm{def}}{=} \neg(\neg A \wedge \neg B)$

diszjunkció definiálása (HUN) – definition of conjuction (ENG)

(PC.7) $A \wedge B \Rightarrow A$
$A \wedge B \Rightarrow B$

konjunkció eliminálása (HUN) – elimination of conjuction (ENG)

(PC.8) $A \Rightarrow A \vee B$
$B \Rightarrow A \vee B$

diszjunkció bevezetése (HUN) – introduction of disjunction (ENG)

(PC.9) $\{\neg B, (A \to b)\} \Rightarrow \neg A$

modus tollens következtetési szabály (HUN) – modus tollens inference rule (ENG)

(PC.10) $\neg\neg A \Rightarrow A$

kettős tagadás törvénye (HUN) – double negation law (ENG)

(PC.11) $(A \to B) \Rightarrow (\neg B \to \neg A)$

kondicionális kontrapozíció törvénye (HUN) – law of conditional contraposition (ENG)

[10] A propozícionális kalkulus kapcsán S.C. Kleene vagy negyven formulát bemutat [401], más felfogás mentén Ruzsa Imre 27 szabályt sorol fel [627, 29-45].

(PC.12) $(A \leftrightarrow B) \Rightarrow (A \rightarrow B)$
$(A \leftrightarrow B) \Rightarrow (B \rightarrow A)$

bikondicionális eliminálása (HUN) – elimination of biconditional (ENG)

(PC.13) $A \leftarrow B \stackrel{\text{def}}{=} B \rightarrow A$

retrokondicionális definiálása (HUN) – kondicionális konverzió (HUN) – implikációs konverzió (HUN) – definition of retroconditional (HUN) – conditinal conversion (ENG)

Megállapíthatunk még egy további fontos szabályszerűséget: a nulladrendű kijelentéslogika zárt a logikai operátorokra, ami azt jelenti, hogy minden művelet eredménye propozíció, vagyis a műveletekkel képzett új állítások a logika érvényességi körén belül maradnak. Logikai szempontból még tovább lehetne egyszerűsíteni a kijelentéslogika nyelvét, ha ki akarnánk használni azt a tényt, hogy mind az exklúzió ($|$), mind a konnegáció ($||$) azzal az egyedi tulajdonsággal rendelkezik, hogy segítségével az összes többi logikai művelet kifejezhető. A tagadás – például – az exklúzió segítségével a következőképpen írható fel:

(PC.14) $\neg A \Leftrightarrow A \mid A$

Ez azt jelenti, hogy a propozicionális logika minden művelete átalakítható lenne, és minden formulája felírható lenne e két logikai operátor valamelyikével. A 1.3.2 táblában bemutatom, hogy az exklúzióra támaszkodva hogyan lehet reprezentálni a tizenhat logikai kapcsoló műveletét. Ez a megoldás egyszerűsítené az adott logikai nyelvet. Ha nem lenne

Table 1.1: a logikai műveletek visszavezetése az exklúzió műveletére

művelet	formula ($\mid$)
ellentmondás	$(A \mid (A \mid A)) \mid (A \mid (A \mid A))$
konnegáció	$((A \mid A) \mid (B \mid B)) \mid ((A \mid A) \mid (B \mid B))$
retroszubtrakció	$((A \mid B) \mid B) \mid ((A \mid B) \mid B)$
kontraprojekció	$A \mid A$
szubtrakció	$(A \mid (A \mid B)) \mid (A \mid (A \mid B))$
kontra-retroprojekció	$B \mid B$
biszubtrakció	$((A \mid (A \mid B)) \mid ((B \mid (A \mid B)))$
exklúzió	$A \mid B$
konjunkció	$(A \mid B) \mid (A \mid B)$
bikondicionális	$((A \mid (A \mid B)) \mid ((A \mid B) \mid B)) \mid ((A \mid (A \mid B)) \mid ((A \mid B) \mid B))$
retroprojekció	B
kondicionális	$A \mid (A \mid B)$
projekció	A
retrokondicionális	$(A \mid B) \mid B$
diszjunkció	$(A \mid A) \mid (B \mid B)$
tautológia	$(A \mid A) \mid ((A \mid A) \mid (A \mid A))$

szükség arra, hogy emberek is értelmezzék a nyelv formuláit, akkor gazdaságos lenne ez az átírást megtenni. Amikor számítógépeket lehet használni kijelentéslogikai feladatok megoldására, akkor ezt meg is teszik, a logikai áramköröket gyakran NAND- vagy NOR-kapuk segítségével építik fel, ami azt jelenti, hogy minden logikai műveletet e két operátort reprezentáló fizikai eszközzel oldanak meg. Az egy-műveletes nyelv gazdaságossága azonban hátránnyá válik akkor, amikor embereknek is értelmezniük kell a nyelv

kifejezéseit. Az ember inkább használ több nyelvi terminust, inkább elviseli a redundáns kifejezéseket, mintsem hosszú logikai formulákból bogozza ki állandóan a formula(rész) tényleges jelentését. A fenti táblázatban a leghosszabb formulája a bikondicionálisnak van, ami helyett nekünk könnyebb egy új logikai jelet ($\leftrightarrow$) megtanulni és használni erre a műveletre.

1.4 Predikátumlogika

A nulladrendű logika a kijelentések, a mondatok szintjén keresi a szabályszerűségeket. A mondat, az állítás ilyenkor bonthatatlan egészet alkot, nem tudjuk, nem tudhatjuk, mi van a propozíciók belsejében. Csak az számít, hogy igaz vagy hamis-e az állítás. Ezen a szinten a logika csak 'tudomásul veszi', de érdemben nem tud foglalkozni a logikai igazság/hamisság kérdésével. Az állításlogikában ezért a logikai elemzés fókusza a mondatok közti műveletek vizsgálatára irányul.

1.4.1 Predikátumlogikai alapfogalmak

A logikai apparátus leíróerejét úgy lehet növelni, ha a logikai elemzés hatókörét kiterjesztjük úgy, hogy a kijelentések "belsejét" is vizsgálni tudjuk. Ehhez az kell, hogy "felbontsuk" a mondatot, és megengedjük a kijelentések összetevőkre bontását és azok vizsgálatát. Amikor ezt tesszük, akkor lépünk át az elsőrendű vagy predikátumlogika világába. Ennek leírásához használhatjuk a következő predikátumlogikai nyelvet.

(QL.0) $\mathscr{L}_{QL}(\mathscr{L}_{PL}, m, n, o, t, s, r, \forall, \exists, =)$, ahol

m, n, o	változó
t, s, r	konstans
$\forall$	univerzális kvantor
$\exists$	egzisztenciális kvantor
$=$	egyenlőség

A predikátumlogikában egy mondat egy névre vagy több névre és egy predikátumra bontható fel. A név individuumkonstans: olyan logikai konstans, amely valamely individuumot jelöl. A nevek a predikátumok bemeneteinek, paramétereinek számítanak. A predikátum az a mondatkomponens, amely egy vagy több névre mint bemeneti paraméterekre vonatkozóan fejezi ki az állítás tartalmát (valamilyen cselekvést, tulajdonságot, állapotot, történést, kapcsolatot stb.). Amikor a predikátumlogikát formális nyelvként írjuk le, akkor részben pontosítjuk, részben igazodunk a formális nyelvelmélet terminológiájához, ezért a neveket logikai konstansokkal (individuumok neveivel) fejezzük ki, a predikátumok helyett pedig vagy a relációk vagy a függvények fogalmait használjuk.

A predikátumlogika nyelvébe átvesszük az állításlogika nyelvét (kijelentéseket, formulákat, logikai műveleteket), és ennek kiegészítéseként be kell még vezetnünk további kategóriákat [627]. Az első kettő új komponens az értékét nem változtató név, amely individuumokat azonosító nyelvi megnyilatkozás, valamint a predikátum, ami egy vagy több név mint bemeneti paraméter alapján állítást képező logikai függvény.

(QL.1) $\text{CONSTANT} \stackrel{\text{def}}{=} t, s, r$

konstans (HUN) – név (HUN) – constant (ENG) – name (ENG)

(QL.2) PREDICATE $\stackrel{\text{def}}{=} A, B, C, P, R$

predikátum (HUN) – logikai függvény (HUN) – funktor (HUN) – reláció (HUN) – függvény (HUN) – predicate (ENG) – logical function (ENG) – relation (ENG) – function (ENG) – functor (ENG)

Az elsőrendű logika segítségével mondhatunk olyan mondatokat, amelyekben valamely reláció konkrét előfordulására vonatkozóan teszünk egy állítást úgy, mint a következő példában (tekintsünk most el attól, hogy a 'szürke' minősítést pontosan hogyan is kell érteni, kezelni logikailag):

'Ráró szürke.' = 'A Ráró nevű ló szürke.'

Az ilyen mondatok azonban kevéssé érdekesek a logika számára. Ha csak konstansok és relációk lennének ebben a formális nyelvben, akkor ezt a logikát kevéssé lehetne használni, mert így gyenge általánosítóerővel rendelkezne. Azzal léphetünk nagyot az általánosíthatóság, az absztrakció irányában, ha az individuumokat jelző konstansok (konkrét nevek) helyett **változókat** alkalmazunk. Általánosíthatjuk a mondatainkat, ha a relációkba változókat írunk:

'Az x ló szürke.' = 'LÓ(x) szürke'

Ekkor viszont nem tudjuk egyértelműen, mire (itt: melyik lóra) vonatkozik az állítás. Az ilyen kijelentés nem tartalmas, nem egyértelmű. Azt mondjuk rá, hogy nyitott mondat. Az az izgalmas a logika számára, ha olyan általános kijelentéseket fogalmazhat meg, amelyek tartalmasak, zártak. Ilyen mondatok képzésére valók a **kvantorok**, amelyekkel "le lehet zárni" a nyitott logikai mondatokat.

Két kvantort használunk: az univerzális kvantort (jelentése: 'minden x-re igaz, hogy', jele: $\forall$) és az egzisztenciális kvantort (jelentése: 'van olyan x, amire igaz, hogy', jele: $\exists$). A kvantorokkal "leköthetjük" a szabad változókat, ami azt is jelenti egyben, hogy általános (és tartalmas, zárt) állításokat nyerhetünk az alkalmazásukkal. Ha azt mondjuk, hogy:

'minden ló szürke' – $\forall m$ LÓ(x) szürke

'létezik ló, ami szürke' – $\exists x$ LÓ(x) szürke

akkor olyan állításokhoz jutunk, amelyek már logikailag értékelhetőek (és az első kijelentést hamisnak minősíthetjük, hiszen vannak nem szürke lovak is, a másodikat állítást viszont igaznak tarthatjuk, hiszen láttunk szürke lovat). Az elsőrendű logika nyelvét ki kell egészítenünk a fent bemutatott új kategóriákkal. Szükség van *(i)* a változó fogalmára, ami a logikai kifejezésekben valamely halmaz elemeit mint lehetséges értékhordozókat helyettesítő nyelvi jel, illetve szükség van két – a kategóriák terjedelmét jelentő előfordulásokra vonatkozó – új operátorrra is. Az egyik *(ii)* az egzisztenciális kvantor, amely legalább egy előfordulás létezését tételezi a logikai kifejezésben, a másik *(iii)* az univerzális kvantor, amely az összes előfordulás létezését fejezi ki a logikai kifejezésben.

(QL.3) VARIABLE $\stackrel{\text{def}}{=} m, n, o, p$

változó (HUN) – variable (ENG)

(QL.4) EXISTENTIAL_QUANTIFIER $\stackrel{\text{def}}{=} \exists$

egzisztenciális kvantor (HUN) – existential quantifier (ENG)

(QL.5) UNIVERSAL_QUANTIFIER $\stackrel{\text{def}}{=} \forall$

univerzális kvantor (HUN) – universal quantifier (ENG)

A két kvantor közül elég csak az egyiket felvenni a nyelvbe alapkategóriaként, a másik már definálható a segítségével. A predikátumkalkulusban így is teszünk majd. A következő fejezetben mutatom meg, hogy hogyan.

1.4.2 Predikátumkalkulus

Természetesen a predikátumlogikának is van kalkulusa, a **predikátumkalkulus** (Predicate Calculus – QC). A QC-kalkulust sem elemzem itt alaposabban, elégnek tartom bemutatni az axiómsémáit. A predikátumkalkulus a predikátumlogika nyelvének azon legszűkebb része, amelyre támaszkodva minden predikátumlogikai formulát ki lehet fejezni valahogyan. A predikátumkalkulusnak egyszerűbb nyelv is elégséges, amit a következőféleképpen lehet definiálni.

(QC.0) $\mathscr{L}_{QC}(A, B, C, m, n, o, t, \forall, \neg, \rightarrow, =)$, ahol

A, B, C	formula
m, n, o	változó
t, s, r	konstans
$\forall$	univerzális kvantor
$\neg$	tagadás logikai művelet
$\rightarrow$	kondicionális logikai művelet
$=$	egyenlőség

Ezen a nyelven már definiálhatjuk a predikátumkalkulus axiómaséma-készletét, amely az alábbi nyolc tételből áll [627, 130]. A negyedik és a nyolcadik sémában a kitevőben szereplő '/' törtjeles kifejezést úgy kell értelmezni, hogy bárhogy is szerepel az A formulában a törtjel utáni változó (o), a séma érvényes marad minden olyan esetben, ha az o helyébe a t terminust (nevet), illetve az m vagy n változót helyettesítjük [627, 81].

(QC.1) $A \rightarrow (B \rightarrow A)$

(QC.2) $(A \rightarrow (B \rightarrow C)) \rightarrow ((A \rightarrow B) \rightarrow (A \rightarrow C))$

(QC.3) $(\neg A \rightarrow \neg B) \rightarrow (B \rightarrow A)$

(QC.4) $\forall o(A \rightarrow A^{t/o})$

(QC.5) $\forall o(A \rightarrow B) \rightarrow (\forall o(A) \rightarrow \forall o(B)))$

(QC.6) $A \rightarrow \forall o(A)$

(QC.7) $o = o$

(QC.8) $(m = n) \rightarrow (A^{m/o} \rightarrow A^{n/o})$

A (QC) első három tétele megegyezik a (PC) első három axiómasémájával. Ebből következik, hogy az összes logikai műveletet be lehet vezetni a predikátumkalkulusban. A második három séma a kvantifikációra, az utolsó kettő pedig az azonosságra vonatkozik. A (QC) alapján természetesen további logikai szimbólumokat is be lehet vezetni a propozícionális logika nyelvébe, akár az egzisztenciális kvantort, akár az implikáción túli logikai műveleteket. Az egzisztenciális kvantorra az alábbi definíció adható:

(QC.9) $\text{EXISTENTIAL_QUANTIFIER}(m) \equiv \exists m \stackrel{\text{def}}{=} \neg\forall\neg m$

egzisztenciális kvantor(HUN) – existential quantifier (ENG)

A (QC) levezetési szabálya a (MP) modus ponens, de természetesen minden más állításlogikai következtetési szabályt alkalmazhatunk. A későbbiekben szükségünk lehet arra, hogy amikor az egzisztenciális kvantort alkalmazzuk, akkor mondhassuk azt, hogy abból az individuumból, aminek a létezését állítjuk az egzisztenciális kvantor segítségével, abból csak egyetlen egy van. Ezt a következő formulával fejezhetjük ki.

(QC.10) $\text{UNIQUENESS_QUANTIFIER}(m) \equiv \exists! m \equiv \exists! m(\mathbf{P}(m)) \stackrel{\text{def}}{=} \exists m\forall n(\mathbf{P}(n) \leftrightarrow m = n)$

unicitás kvantor (HUN) – uniqueness quantifier (ENG) – unique existential quantifier (ENG)

A módosított kvantort (operátort) úgy kell értelmezni, hogy ‘létezik egy és csak egy’, és mindig mellé kell képzelni egy predikátumot, amire vonatkoztatjuk a kvantort (aminek az egyeleműségét állítjuk). Mivel az operátor segítségével az egyelemű predikátumokat tudjuk kifejezni, és a fenti formulát szokás unicitássémának is nevezni[627, 98], ezért használhatjuk az unicitás kvantor (operátor) megnevezést.

1.5 Relációelmélet

A formális nyelvek által definiált fogalmak valamilyen struktúrát alkotnak, a struktúrák vázát pedig mindig az adott nyelvbe felvett relációk adják. Bármilyen relációs struktúrát építünk is fel, minden esetben szükség van arra, hogy a relációkat magukat is jellemezni tudjuk. A relációelmélet ezt a feladatot végzi el. Három fogalom van az elmélet központjában: a **reláció**, a **függvény** és a **művelet**. Ezek mindegyikét különféle tulajdonságokkal jellemezhetjük, műveletek értelmezhetünk közöttük, és mindegyik levezethető a másikból, tehát bármelyiket megtehetnénk egy formális elmélet alapfogalmául. Én azt a megközelítést fogadom el, amely szerint a reláció fogalmából kell kiindulni, és ebből lehet levezetni előbb a függvény, majd a művelet fogalmát.

1.5.1 Relációelméleti fogalmak

A relációelméleti fogalmak rövid bemutatását a relációhoz kapcsolódó fogalomcsokor tárgyalásával kezdem. Első lépésként sorba veszem a relációt leíró tulajdonság- és műveletfogalmakat. Ezután módszertanilag annak kéne következnie, hogy bemutassam a reláció fogalmára építhető kalkulust, de ezt addig nem lehet megtenni, amíg be nem vezettük a művelet fogalmát. Ezért a relációs tulajdonságfogalmak bemutatása után rátérek a függvény és a művelet fogalmainak tárgyalására, és csak miután ezekkel végeztem, azután tudom röviden felvázolni a relációelmélet kalkulusát, a relációalgebrát.

Reláció

A **reláció** fogalmával kettő vagy több dolog közti kapcsolatot ragadhatjuk meg. Ez azt jelent, hogy mindig meg kell tudnunk adni, hogy milyen dolgok közötti viszonyt írunk le az adott reláció segítségével. Ehhez meg kell adnunk egy (vagy több) halmazt, amely(ek)nek elemeit összekapcsolhatjuk egymással a reláció keretében. Az összekapcsolt elemeket

relátumoknak nevezzük. Az összekapcsolt elemeket adjuk meg a reláció argumentumában, és attól függően, hogy hány elemet kapcsolunk össze beszélhetünk a reláció argumentumszámáról. A relációk jellemzéséhez kiindulhatunk az alábbi nyelvből.

(RL.0) $\mathscr{L}_{RL}(\mathscr{L}_{QL}, \mathbf{R}, \mathbf{F}, \mathbf{O}, \langle m,n,o\rangle, \mathbb{X}\times\mathbb{X}, \mathbb{Dom}, \mathbb{Im})$, ahol

$\langle m,n,o\rangle$	rendezett n-s
R,P,Q	reláció
F	függvény
O	művelet
$\mathbb{X}\times\mathbb{X}$	Descartes szorzat
$\mathbb{Dom}$	értelmezési tartomány
$\mathbb{Im}$	képtartomány

A relációk minél jobb megragadhatósága érdekében rengeteg tulajdonságot definiáltak, és ezek alapján nevezetes relációkat is elkülönítettek egymástól. Első körben a legfontosabb relációtulajdonságokat, majd a leggyakoribb nevezetes relációtípusokat fogom bemutatni. A relációk egyik első tipizálási lehetőségét az jelent, ha figyelembe vesszük azt a tényt, hogy hány argumentuma van a relációnak. Beszélhetünk unáris, bináris vagy n-áris relációról, ha egy, kettő vagy n elemet kapcsolunk össze a reláció segítségével. A reláció fogalmát halmazelméleti alapon a következőféleképpen határozhatjuk meg.

(RL.1) $\text{RELATION}(m,n) \equiv \langle m,n\rangle \equiv (m,n) \in \mathbf{R} \subset \mathbb{X}\times\mathbb{X} \equiv \mathbf{R}(m,n) \stackrel{\text{def}}{=} \{\{m\},\{m,n\}\}$

reláció

A relációkhoz (függvényekhez) kapcsolódik két fontos osztályfogalom, az értelmezési tartomány és értékkészlet. Ezeket a következő módon emelhetjük be a modellünkbe.

(RL.2) $\text{DOMAIN} \equiv \mathbb{Dom}(\mathbf{R}(m,n)) \equiv \mathbb{Dom}(\mathbf{R}) \equiv \mathbb{Dom} \equiv \mathbb{D} \stackrel{\text{def}}{=} \{m : \exists n(\mathbf{R}(m,n))\}$

értelmezési tartomány (HUN) – domain (ENG)

(RL.3) $\text{IMAGE} \equiv \mathbb{Im}(\mathbf{R}(m,n)) \equiv \mathbb{Im}(\mathbf{R}) \equiv \mathbb{Im} \equiv \mathbb{I} \stackrel{\text{def}}{=} \{n : \exists m(\mathbf{R}(m,n))\}$

képtartomány (HUN) – kép (HUN) – értékkészlet (HUN) – image (ENG) – (ENG) range (ENG) – co-domain (ENG)

A relációk egyik legáltalánosabb tipizálási lehetőségét az adja, ha arra figyeljük, hogy a reláció által összekapcsolt dolgok homogének-e vagy sem, azaz ugyanabból a halmazból vesszük-e őket vagy sem. Ha igen, akkor **homogén** (homogeneous), ha nem, akkor **heterogén** (heterogeneous) relációról beszélünk. Még mindig a legáltalánosabb szinten maradva definiálhatunk négy nevezetes relációt, amelyekre olykor szükségünk lesz majd a későbbiekben.

(RL.4) $\text{UNIVERSAL_RELATION}(m,n) \equiv \underset{\perp}{\mathbf{U}}(m,n) \equiv \underset{\perp}{\mathbf{U}} \equiv \underset{\perp}{\mathbf{F}}(m,n) \equiv \underset{\perp}{\mathbf{F}} \stackrel{\text{def}}{=} \forall m\forall n(R(m,n))$

$\underset{\perp}{\mathbf{F}}$ reláció **teljes** (HUN) – univerzális reláció (HUN) – universal relation (ENG) – full relation (ENG)

(RL.5) $\text{EMPTY_RELATION}(m,n) \equiv \underset{\perp}{\mathbf{E}}(m,n) \equiv \underset{\perp}{\mathbf{E}} \stackrel{\text{def}}{=} \forall m\forall n(\neg R(m,n))$

$\underset{\perp}{\mathbf{E}}$ reláció **üres** reláció (HUN) – null reláció (HUN) – empty relation (ENG) – null relation (ENG)

(RL.6) $\text{IDENTITY_RELATION}(m,n) \equiv \underset{\scriptscriptstyle\perp}{\mathbf{I}}(m,n) \equiv \underset{\scriptscriptstyle\perp}{\mathbf{I}} \stackrel{\text{def}}{=} \forall m \forall n (R(m,n) \leftrightarrow m = n)$

$\underset{\scriptscriptstyle\perp}{\mathbf{I}}$ reláció **azonosság** reláció(HUN) – azonosság (HUN) – egység reláció (HUN) – diagonál(is) reláció (HUN) – egyenlőségi reláció (HUN) – egyenlőség (HUN) – identitás reláció (HUN) – identity (ENG) – diagonal (ENG)

(RL.7) $\text{DIVERSITY_RELATION}(m,n) \equiv \underset{\scriptscriptstyle\perp}{\mathbf{D}}(m,n) \equiv \underset{\scriptscriptstyle\perp}{\mathbf{D}} \stackrel{\text{def}}{=} \forall m \forall n (R(m,n) \leftrightarrow m \neq n)$

$\underset{\scriptscriptstyle\perp}{\mathbf{D}}$ reláció **különbözőség** reláció (HUN) – különbözőség (HUN) – egyenlőtlenségi reláció (HUN) – egyenlőtlenség (HUN) – antidiagonális reláció (HUN) – differencia reláció (HUN) – inkompatibilitási reláció (HUN) – diverzitási reláció (HUN) – difference (ENG) – incompatibility relation (ENG) – antidiagonal relation (ENG) – diversity relation (ENG)

Mivel a fenti négy relációtípus fontos szerepet tölt be a relációelméletben, ezért saját nevet kapott, nem pedig a rá jellemző tulajdonság alapján lehet hivatkozni rájuk. A következő blokkban azokat a relációtulajdonságokat mutatom be, amelyek önmagukban "csak" tulajdonságok, de a relációfogalomhoz kapcsolva őket egy sor nevezetes, gyakran használt relációt azonosíthatunk velük. A különböző korok, diszciplinák, megközelítések, emberek ugyanarra a problémára sokszor nagyon hasonló, de bizonyos pontokon így vagy úgy eltérő tulajdonságokat tartottak fontosnak bevezetni, ezért az egymáshoz közeli fogalmakat csoportokba rendezve mutatom be.

A reflexivitással kapcsolatos tulajdonságok azt a minőséget próbálják megragadni, hogy az elemek között tételezett reláció vonatkoztatható-e az elemekre önmagukra is, vagyis összeköthetők-e az elemek önmagukkal vagy sem. Ezt fejezi ki a reflexivitás tulajdonsága. Hétköznapi példa lehetne a reflexivitásra a SZERETI reláció, hiszen valaki szerethet másokat, de szeretheti magát is, ám ez példa kevés lehet a reflexivitás fennállásához, hiszen nem minden esetben érvényesül. Vannak ugyanis olyan emberek, akik nem szeretik önmagukat. Márpedig a reflexivitás minőségét csak akkor érdemes egy relációhoz kötni, ha az minden elemre vonatkoztatható. Ebben az értelemben elfogadható példa lehet az EGYIDŐS reláció, hiszen egy ember egyidős lehet valaki mással, de mindig egyidős önmagával. Olyan eset is előfordulhat, amikor a reflexivitás sosem érvényesül. Például a SZÜLŐJE vagy a TESTVÉRE relációra hivatkozhatunk, amikoris mondhatjuk, hogy senki sem lehet önmaga szülője vagy testvére.

(RL.8) $\text{REFLEXIVE}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m (\mathbf{R}(m,m))$

reflexív (HUN) – reflexive (ENG) – erősen reflexív (HUN) – strongly reflexive (ENG)

(RL.9) $\text{IRREFLEXIVE}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m (\neg\mathbf{R}(m,m))$

irreflexív (HUN) – antireflexív (HUN) – irreflexive (ENG) – antireflexive (ENG)

(RL.10) $\text{NON-REFLEXIVE}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \exists m (\neg\mathbf{R}(m,m))$

nem-reflexív (HUN) – non-reflexive (ENG)

(RL.11) $\text{SECONDARILY_REFLEXIVE}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n (\mathbf{R}(m,n) \rightarrow \mathbf{R}(n,n))$

másodlagosan reflexív (HUN) – szekunder reflexív (HUN) – secondarily reflexive (ENG) – shift reflexive (ENG)

(RL.12) $\text{INVERSE_SECONDARILY_REFLEXIVE}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n (\mathbf{R}(m,n) \rightarrow \mathbf{R}(m,m))$

fordítottan másodlagosan reflexív (HUN) – feltételesen reflexív (HUN) – conditional reflexive (ENG) – inverse secondarily reflexive (ENG)

(RL.13) $\text{COREFLEXIVE}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n(\mathbf{R}(m,n) \rightarrow m = n)$

koreflexív (HUN) – üres (HUN) – vacuous (ENG) – coreflexive (ENG)

A relációtulajdonságok másik csoportjába tartoznak a szimmetricitással kapcsolatos tulajdonságok. Ekkor arra kérdezünk rá, hogy ha a reláció két elem között fennáll, akkor vajon fennáll-e a fordított irányban is, vagy másként: ha a reláció igaz, akkor igaz-e a konverze. Szimmetrikus a TESTVÉRE reláció, hiszen ha A testvére B-nek, akkor ez fordítva is igaz, vagyis B testvére A-nak. Léteznek azonban olyan relációk is, amelyekre minden esetben kizárhatjuk a szimmetricitást. Ilyen például az IDŐSEBB reláció: ha A idősebb B-nél, akkor B nem lehet sosem idősebb A-nál.

(RL.14) $\text{SYMMETRIC}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n(\mathbf{R}(m,n) \rightarrow \mathbf{R}(n,m))$

szimmetrikus (HUN) – symmetric (ENG)

(RL.15) $\text{ASYMMETRIC}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n(\mathbf{R}(m,n) \rightarrow \neg\mathbf{R}(n,m))$

aszimmetrikus (HUN) – antiszimmetrikus (HUN) – szigorúan antiszimmetrikus (HUN) – asymmetric (ENG) – antisymmetric (ENG)

(RL.16) $\text{ANTISYMMETRIC}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n((\mathbf{R}(m,n) \wedge \mathbf{R}(n,m)) \rightarrow m = n)$

antiszimmetrikus (HUN) – identitív (HUN) – antisymmetric (ENG) – identitive (ENG)

(RL.17) $\text{NON-SYMMETRIC}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \mathbf{R}(m,n) \wedge \neg\mathbf{R}(n,m))$

nem-szimmetrikus

Fontos csoportot képeznek azok a tulajdonságok, amelyek a teljességgel, összehasonlíthatósággal, összekapcsoltsággal, egyértelműséggel kapcsolatosak. Mindig jó tudnunk egy relációval kapcsolatban, hogy a reláció tartóhalmazának minden elemére kiterjed-e a reláció érvénye vagy sem. Különösen fontos kérdés ez az összehasonlíthatóság miatt. A ISMERI relációt például nem tarthatjuk teljesnek, mert egy csoporton belül nem mindenki ismerheti egymást, ezért a csoport tagjai között nem állíthatunk fel egy teljes ismertségi hálót. Az egyértelműség kérdése pedig azért lehet fontos, mert sokszor akarhatjuk azt, hogy egy reláció fennállásához egyértelmű következményeket rendelhessünk. Ennek "kiderítésére" is érdemes tulajdonságokat definiálni.

(RL.18) $\text{WEAKLY_CONNECTED}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n \forall o(\mathbf{R}(m,n) \wedge \mathbf{R}(m,o) \rightarrow (n = o \vee \mathbf{R}(n,o) \vee \mathbf{R}(o,n)))$

gyengén összefüggő (HUN) – weakly connected (ENG) – weakly connex (ENG) – összehasonlítható (HUN) – comparable (ENG)

(RL.19) $\text{CONNECTED}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n(m \neq n \rightarrow (\mathbf{R}(m,n) \vee \mathbf{R}(n,m)))$

összefüggő (HUN) – konnex (HUN) – connex (ENG) – connected (ENG) – lineáris (HUN) – teljes (HUN) – complete (ENG) – totális (HUN) – total (ENG) – összehasonlítható (HUN) – comparable (ENG)

(RL.20) $\text{STRONGLY_CONNECTED}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n(\mathbf{R}(m,n) \vee \mathbf{R}(n,m))$

erősen összefüggő (HUN) – strongly connex (ENG) – strongly connected (ENG) – totally connected (ENG) – totálisan összefüggő (HUN) – teljes (HUN) – complete (ENG) – totális (HUN) – total (ENG) – összehasonlítható (HUN) – comparable (ENG)

(RL.21) $\text{DICHOTOME}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n (m \neq n \rightarrow (\mathbf{R}(m,n) \wedge \neg \mathbf{R}(n,m)) \vee (\neg \mathbf{R}(m,n) \wedge \mathbf{R}(n,m)))$

dichotóm (HUN) – összehasonlítható (HUN) – comparable (ENG) – totális (HUN) – teljes (HUN) – total (ENG) – dichotomous (ENG) – complete (ENG)

(RL.22) $\text{TRICHOTOME}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n ((\mathbf{R}(m,n) \wedge \neg \mathbf{R}(n,m) \wedge m \neq n) \vee (\neg \mathbf{R}(m,n) \wedge \mathbf{R}(n,m) \wedge m \neq n) \vee m = n)$

trichotóm (HUN) – lineáris (HUN) – trichotomous (ENG)

(RL.23) $\text{COMPARABLE}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n \forall o (\mathbf{R}(m,n) \rightarrow (\mathbf{R}(m,o) \vee \mathbf{R}(o,n)))$

összehasonlítható (HUN) – comparable (ENG)

(BL.24) $\text{INDIFFERENT}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n (\mathbf{R}(m,n) \wedge \mathbf{R}(n,m))$

indifferens (HUN) – erősen indifferens (HUN) – erősen közömbös (HUN) – strongly indifferent (ENG) – indifferent (ENG) – egyenlő értékű (HUN) – equal in value (ENG) – ugyanolyan, mint (HUN) – same as (ENG)

(BL.25) $\text{INCOMPARABLE}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n (\neg \mathbf{R}(m,n) \wedge \neg \mathbf{R}(n,m))$

összehasonlíthatatlan (HUN) – gyengén közömbös (HUN) – gyengén indifferens (HUN) – incomparable (ENG) – weakly indifferent (ENG) – ugyanolyan értékes, mint (HUN) – same value as (ENG)

(RL.26) $\text{LEFT-TOTAL}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \exists n (\mathbf{R}(m,n))$

balról totális (HUN) – totális (HUN) – sorozatos (HUN) – szeriális (HUN) – serial (ENG) – sorozat (HUN) – left-total (ENG) – extendable (ENG)

(RL.27) $\text{RIGHT-TOTAL}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall n \exists m (\mathbf{R}(m\mathbf{R}(m,n))$

jobbról totális (HUN) – right-total (ENG)

(RL.28) $\text{LEFT-UNIQUE}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n \forall o ((\mathbf{R}(m,o) \wedge \mathbf{R}(n,o)) \rightarrow m = n)$

balról egyértelmű (HUN) – inverz-funkcionális (HUN) – inverse functional (ENG)

(RL.29) $\text{RIGHT-UNIQUE}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n \forall o ((\mathbf{R}(m,n) \wedge \mathbf{R}(m,o)) \rightarrow n = o)$

jobbról egyértelmű (HUN) – funkcionális (HUN) – parciális függvény (HUN) – partial function (ENG) – functional relation (ENG) – right-unique (ENG) – univalent (ENG)

(RL.30) $\text{DENSE}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \forall o (\mathbf{R}(m,o) \rightarrow \exists n (\mathbf{R}(m,n) \wedge \mathbf{R}(n,o)))$

sűrű (HUN) – dense (ENG) – gyengén sűrű (HUN) – weakly dense (ENG)

(RL.31) $\text{CONVERGENT}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n \forall o \exists p ((\mathbf{R}(o,m) \wedge \mathbf{R}(o,n)) \rightarrow (\mathbf{R}(m,p) \wedge \mathbf{R}(n,p)))$

konvergens (HUN) – convergent (ENG)

A későbbiekben talán a legtöbbet azzal a tulajdonságcsoporttal foglalkozunk majd, amely a tranzitivitással, ciklikussággal, aciklikussággal kapcsolatos. A relációk segítségével való következtetéseinket támogathatja az, ha tudjuk, hogy ha az adott reláció fennáll bizonyos elemek között, akkor ebből következtethetünk más minden elem közti fennállás tényére. Példaként hivatkozhatunk a IDŐSEBB relációra: ha *A* idősebb *B*-nél, és *B* idősebb *C*-nél, akkor már – mivel ez a reláció tranzitív – következtethetünk arra, hogy *A* idősebb *C*-nél. A későbbiekben a tranzitivitás tulajdonsága a preferencia fogalmával kapcsolatban fog gyakran előkerülni. Nem könnyíti meg a fogalom értelmezését az a tény, hogy

fontossága miatt sokféle módon próbálták meg kifejezni a tranzitivitás minőségét, így sok versengő, egymásra hasonlító definíció született e téren.

(RL.32) $\text{TRANSITIVE}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n \forall o((\mathbf{R}(m,n) \wedge \mathbf{R}(n,o)) \rightarrow \mathbf{R}(m,o))$

tranzitív (HUN) – transitive relation (ENG)

(RL.33) $\text{INTRANSITIVE}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n \forall o((\mathbf{R}(m,n) \wedge \mathbf{R}(n,o)) \rightarrow \neg\mathbf{R}(m,o))$

intranzitív (HUN) – atranzitív (HUN) – antitranzitív (HUN) – intransitive (ENG) – atransitive (ENG) – antitransitive (ENG)

(RL.34) $\text{QUASI-TRANSITIVE}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n \forall o((\mathbf{R}(m,n) \wedge \mathbf{R}(n,o) \wedge m \neq o) \rightarrow \mathbf{R}(m,o))$

kvázitranzitív (HUN) – quasi-transitive (ENG)

(RL.35) $\text{ALIO-TRANSITIVE}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n \forall o((\mathbf{R}(m,n) \wedge \mathbf{R}(n,o)) \rightarrow (\mathbf{R}(m,o) \vee m = o))$

aliotranzitív (HUN) – alio-transitive (ENG)

(RL.36) $\text{SEMI-TRANSITIVE}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n \forall o \forall p((\mathbf{R}(m,n) \wedge \mathbf{R}(n,o)) \rightarrow (\mathbf{R}(m,p) \vee \mathbf{R}(p,o)))$

szemitranzitív (HUN) – semi-transitive (ENG)

(RL.37) $\text{FERRERS_PROPERTY}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n \forall o \forall p((\mathbf{R}(m,n) \wedge \mathbf{R}(o,p)) \rightarrow (\mathbf{R}(m,p) \vee \mathbf{R}(o,n)))$

Ferrers tulajdonság (HUN) – Ferrers property (ENG)

(RL.38) $\text{NEGATIVELY_TRANSITIVE}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n \forall o((\mathbf{R}(m,n) \rightarrow (\mathbf{R}(m,o)) \vee \mathbf{R}(o,n)))$

negatívan tranzitív (HUN) – negatively transitive relation (ENG)

(RL.39) $\text{NON-TRANSITIVE}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \exists x \exists n \exists z((\mathbf{R}(m,n) \wedge \mathbf{R}(n,o)) \wedge \neg\mathbf{R}(m,o))$

nem-tranzitív (HUN) – atranzitív (HUN) – intranzitív (HUN) – non-transitive (ENG) – intransitive (ENG) – atransitive (ENG)

(RL.40) $\text{CIRCULAR}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n \forall o((\mathbf{R}(m,n) \wedge \mathbf{R}(n,o)) \rightarrow \mathbf{R}(o,m))$

cirkuláris (HUN) – circular (ENG)

(RL.41) $\text{ACIRCULAR}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n \forall o((\mathbf{R}(m,n) \wedge \mathbf{R}(n,o)) \rightarrow \neg\mathbf{R}(o,m))$

acirkuláris (HUN) – acircular (ENG)

(RL.42) $\text{EUCLIDEAN}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n \forall o((\mathbf{R}(m,n) \wedge \mathbf{R}(m,o)) \rightarrow \mathbf{R}(n,o))$

euklideszi (HUN) – euclidean (ENG)

(RL.43) $\text{ANEUCLIDEAN}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n \forall o((\mathbf{R}(m,n) \wedge \mathbf{R}(m,o)) \rightarrow (\neg\mathbf{R}(n,o) \wedge \neg\mathbf{R}(o,n)))$

aneuklideszi (HUN) – antieuclidean (ENG) – aneuclidean (ENG) – ineuclidean (ENG) – antitransitive (ENG)

(RL.44) $\text{CATENARY}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n \exists o(\mathbf{R}(m,n) \rightarrow ((\mathbf{R}(n,o) \vee \mathbf{R}(o,m))))$

katenáris (HUN) – láncolt reláció (HUN) – nem bipoláris (HUN) – catenary (ENG) – non-bipolar (ENG)

(RL.45) $\text{ACATENARY}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n \forall o(\mathbf{R}(m,n) \rightarrow (\neg\mathbf{R}(n,o) \wedge \neg\mathbf{R}(o,m)))$

akatenáris (HUN) – nem-láncolt (HUN) – bipoláris (HUN) – non-catenary (ENG) – acatenary (ENG) – bipolar (ENG)

(RL.46) $\text{CONFLUENT}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n \forall o \exists p((\mathbf{R}(m,n) \wedge \mathbf{R}(m,o)) \rightarrow (\mathbf{R}(n,p) \wedge \mathbf{R}(o,p)))$

konfluens (HUN) – confluent (ENG) – directed (ENG) – incesztuális reláció (HUN) – incestual relation (ENG)

(RL.47) $\text{ACONFLUENT}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n \forall o \exists p(\mathbf{R}(m,n) \wedge \mathbf{R}(m,o)) \rightarrow \neg(\mathbf{R}(n,p) \wedge \mathbf{R}(o,p)))$

akonfluens (HUN) – aconfluent (ENG)

Ahhoz, hogy további fontos relációtulajdonságokat tudjunk definiálni, szükség van arra, hogy bevezessünk olyan új fogalmakat, amelyek elkülönítik a reláció valamely részét adott tulajdonság mentén, és csak az így lehatárolt részre vonatkoznak.

(RL.48) $\text{SYMMETRIC_PART_OF_}\mathbf{R}(m,n) \equiv \mathbf{I}(\mathbf{R}) \stackrel{\text{def}}{=} \{(m,n) \in \mathbb{X} \times \mathbb{X} | (m,n) \in \mathbf{R} \wedge (n,m) \in \mathbf{R}\}$

R szimmetrikus része (HUN) – symmetric part of **R** (ENG) – **R** szimmetrikus faktora (HUN) – symmetric factor of **R** (ENG)

(RL.49) $\text{ASYMMETRIC_PART_OF_}\mathbf{R}(m,n) \equiv \mathbf{P}(\mathbf{R}) \stackrel{\text{def}}{=} \{(m,n) \in \mathbb{X} \times \mathbb{X} | (m,n) \in \mathbf{R} \wedge (n,m) \notin \mathbf{R}\}$

R aszimmetrikus része (HUN) – asymmetric part of **R** (ENG) – **R** aszimmetrikus faktora (HUN) – asymmetric factor of **R** (ENG)

A későbbiekben még definiálni fogunk további származtatott relációkat (↬45.o.). A fenti relációrész-fogalmak segítségével bevezethetünk olyan további tulajdonságokat, amelyekre komoly mértékben támaszkodnunk majd a későbbiekben, a racionalitás problémájának tárgyalásakor. Adottnak veszünk egy relációt, és annak egy részére írunk elő követelményeket.

(RL.50) $\text{TRIPLE_ACYCLIC}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n \forall o((m,n) \in \mathbf{P}(\mathbf{R}) \wedge (n,o) \in \mathbf{P}(\mathbf{R}) \rightarrow (o,m) \notin \mathbf{P}(\mathbf{R}))$

hármas aciklusos (HUN) – triple acyclic relation

(RL.51) $\text{ACYCLIC}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall t \in \mathbb{N}^+, \forall m^1, m^2, \ldots, m^t \in \mathbb{X}(\forall \tau \in \{1,2,\ldots,t-1\}(m^t, m^{t+1}) \in \mathbf{P}(\mathbf{R})) \rightarrow (m^t, m^1) \notin \mathbf{P}(\mathbf{R})$

aciklusos (HUN) – acyclic relation (ENG)

(RL.52) $\text{CONSISTENT}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall t \in \mathbb{N}^+, \forall m^1, m^2, \ldots, m^t \in \mathbb{X}((m^1, m^2) \in \mathbf{P}(\mathbf{R}) \wedge \forall \tau \in \{2,3,\ldots,t-1\}((m^\tau, m^{\tau+1}) \in \mathbf{R})) \rightarrow (m^t, m^1) \in \mathbf{R}$

(Suzumura) konzisztens (HUN) – (Suzumura) consistent relation (ENG)

A konzisztencia fogalmára később bemutatok másfajta definíciókat is (↬45.o.).

---

felvezetni a jól-fundáltságot is valahogy, vagy eltenni későbbre, amikor már a minimális elem (és társai) definiálva vannak vagy eltenni a halmazelméletbe, Ruzsa halmazaxiómaként definiálja....

(RL.53) $\forall \mathbb{S} \subseteq \mathbb{X}(\mathbb{S} \neq \varnothing \rightarrow \exists m \in \mathbb{S}(s,m) \notin \mathbf{R})$

jólfundált az $\mathbb{X}$ halmazra nézve akkor és csak akkor, ha X minden nem-üres S részhalmazának van minimális eleme R-re nézve.

Mivel a tranzitivitás tulajdonsága sok cselekvéselméleti fogalomhoz kapcsolódik, a gazdasági cselekvések területén pedig a legfontosabb fogalmak közé tartozik, a tranzitivitás

tulajdonságára támaszkodva további fogalmakat határozhatunk meg. Ha van egy **R**(m,n) tranzitív relációnk, akkor definiálhatjuk a következő tranzitivitás-vonatkozású tulajdonságokat [661].

(BL.54) PP-TRANSITIVE$(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n \forall o(((\mathbf{R}(m,n) \wedge \neg\mathbf{R}(n,m)) \wedge (\mathbf{R}(n,o) \wedge \neg\mathbf{R}(o,n))) \rightarrow (\mathbf{R}(m,o) \wedge \neg\mathbf{R}(o,m)))$

PP-tranzitív (HUN) – PP-transitive (ENG) – kvázitranzitív (HUN) – quasi-transitive (ENG)

(BL.55) IP-TRANSITIVE$(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n \forall o(((\mathbf{R}(m,n) \wedge \mathbf{R}(n,m)) \wedge (\mathbf{R}(n,o) \wedge \neg\mathbf{R}(o,n))) \rightarrow (\mathbf{R}(m,o) \wedge \neg\mathbf{R}(m,o)))$

IP-tranzitív (HUN) – IP-transitive (ENG)

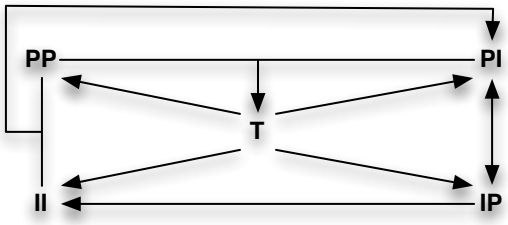
(BL.56) PI-TRANSITIVE$(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n \forall o(((\mathbf{R}(m,n) \wedge \neg\mathbf{R}(n,m)) \wedge (\mathbf{R}(n,o) \wedge \mathbf{R}(o,n))) \rightarrow \mathbf{R}(m,o) \wedge \neg\mathbf{R}(o,m)))$

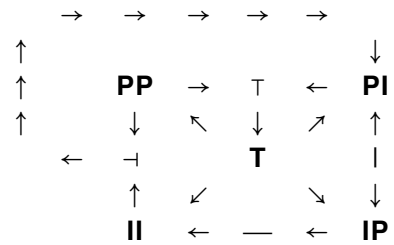
PI-tranzitív (HUN) – PI-transitive (ENG)

(BL.57) II-TRANSITIVE$(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n \forall o(((\mathbf{R}(m,n) \wedge \mathbf{R}(n,m)) \wedge (\mathbf{R}(n,o) \wedge \mathbf{R}(o,n))) \rightarrow (\mathbf{R}(m,o) \wedge \mathbf{R}(o,m)))$

II-tranzitív (HUN) – II-transitive (ENG)

Az így definiált tulajdonságok között fontos összefüggések állapíthatók meg, de ezeket nem mutatom be formális módon, csak megemlítek párat közülük [661], [336, 324]. Igaz például az, hogy ha egy reláció tranzitív (jelöljük ezt itt T-vel), akkor fennáll mind a négy belőle levezethető PP-, II-, IP- és PI-tranzititás tulajdonság is, de fontos tézis az is, miszerint a PP- és PI-, illetve a PP- és II-tulajdonságok együttes fennállásából következik a T tranzitivitás létezése. Ezeket a kapcsolatokat szemléltethetjük egy ábra segítségével.





A preferencia fogalmának tárgyalásakor még visszatérünk ezekre az összefüggésekre, és ott további relációk megnevezésével egyszerűsítünk majd az itt bevezetett tranzitivitási

tulajdonságok formuláin (↬114.o.). Akkor lehet majd rámutatni arra, hogy honnan származnak a most definiált tulajdonságok megnevezései. Az eddig definiált relációtulajdonságok között természetesen nem teljesen függetlenek egymástól, rengeteg összefüggés van közöttük. Itt csak azokat mutatom be, amelyekre a későbbiekben valamiért szükség lesz [718].[11]

(RL.58) ASYMMETRIC → IRREFLEXIVE

(RL.59) STRONGLY_CONNECTED → IRREFLEXIVE

(RL.60) TRANSITIVE → CONSISTENT

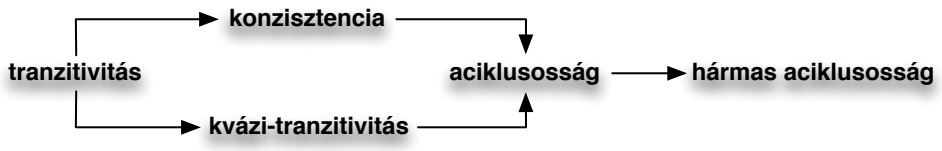
(RL.61) TRANSITIVE → QUASY_TRANSITIVE

(RL.62) CONSISTENT → ACYCLIC

(RL.63) QUASY_TRANSITIVE → ACYCLIC

(RL.64) ACYCLIC → TRIPLE_ACYCLIC

A fenti tételek közül a tranzitivitással kapcsolatos összefüggéseket grafikusan is ábrázolva szemléletesebbé tehetők a fogalmak közti összefüggések. Az ábráról könnyen leolvasható, hogy a legerősebb tulajdonság a tranzitivitás, míg a "leggyengébb" hármas aciklusosság, és köztük vannak a többiek.



Hansson definiál anticiklikusságot is
ciklusmentesség?
Houthakker szemitranzitív relációja [366]

Sen, Suzumura: tranzitív - kvázi-tranzitív - aciklusos relációk fontosak lesznek a társadalmi választási, döntési szabályok megkülönböztetésében Sen egyszerűbben definiálja az aciklusosságot

---

Nevezetes relációk

A legelemibb rendezésfogalmakat a különféle relációtulajdonságok segítségével lehet elkülöníteni egymástól. Sajnos, a rendezési fogalmak világában sincs terminológiai tisztaság, gyakran találkozhatunk mind a homonimitás, mind a szinonimitás jelenségével. Ebben a könyven leginkább a rendezési relációkat fogom használni, de fontosnak tartom megemlíteni itt a többi nevezetes relációt is.

[11] A kvázitranzitivitás terminusa korábban kétféle fogalomhoz kapcsolódva is felbukkant. Az itt bemutatott összefüggésekben a PP-tranzitivitás értelemben szerepel a terminus. Ugyanezt a tartalmat fejezhetjük ki egyébként azzal, ha azt mondjuk, hogy a kvázitranzitív reláció ekvivalens a reláció tanzitív aszimmetrikus részével [105].

(RL.65) $\text{TOLERANCE} \stackrel{\text{def}}{=} \text{REFLEXIVE} \wedge \text{SYMMETRIC}$

tolerancia reláció (HUN) – kompatibilitási reláció (HUN) – hasonlósági reláció (HUN) – compatibility relation (ENG) – similarity relation (ENG) – tolerance relation (ENG)

(RL.66) $\text{PARTIAL_EQUIVALENCE} \stackrel{\text{def}}{=} \text{TRANSITIVE} \wedge \text{SYMMETRIC}$

részleges ekvivalencia reláció (HUN) – partial equivalence relation (ENG) – per (ENG)

(RL.67) $\text{EQUIVALENCE} \stackrel{\text{def}}{=} \text{REFLEXIVE} \wedge \text{SYMMETRIC} \wedge \text{TRANSITIVE}$

ekvivalenciareláció (HUN) – ekvivalencia (HUN) – equivalence (ENG)

(RL.68) $\text{PARTIAL_ORDER} \stackrel{\text{def}}{=} \text{REFLEXIVE} \wedge \text{ANTISYMMETRIC} \wedge \text{TRANSITIVE}$

részben rendezés (HUN) – parciális rendezés (HUN) – részleges rendezés (HUN) – gyenge rendezés (HUN) – gyenge részben rendezés (HUN) – nem szigorú rendezés (HUN) – weak order (ENG) – weak partial order (ENG) – partial order (ENG)

(RL.69) $\text{STRICT_ORDER} \stackrel{\text{def}}{=} \text{IRREFLEXIVE} \wedge \text{TRANSITIVE}$

szigorú rendezés (HUN) – szigorú rendezési reláció (HUN) – erős részben rendezés (HUN) – irreflexív rendezés (HUN) – strict order (ENG) – strict weak order (ENG) – strong order (ENG) – strong weak order (ENG)

(RL.70) $\text{LINEAR_PARTIAL_ORDER} \stackrel{\text{def}}{=} \text{REFLEXIVE} \wedge \text{ANTISYMMETRIC} \wedge \text{TRANSITIVE} \wedge \text{DICHOTOMOUS}$

nem szigorú elrendezés reláció (HUN) – teljes parciális rendezés (HUN) – lineáris parciális rendezés (HUN) – totális parciális rendezés (HUN) – total partial order (ENG) – linear partial order (ENG)

(RL.71) $\text{LINEAR_STRICT_ORDER} \stackrel{\text{def}}{=} \text{IRREFLEXIVE} \wedge \text{TRANSITIVE} \wedge \text{DICHOTOMOUS}$

szigorú elrendezés (HUN) – teljes szigorú rendezés (HUN) – lineáris szigorú rendezés (HUN) – totális szigorú rendezés (HUN) – total strict order (ENG) – linear strict order (ENG) – total strong order (ENG) – linear strong order (ENG)

(RL.72) $\text{PREORDER} \stackrel{\text{def}}{=} \text{REFLEXIVE} \wedge \text{TRANSITIVE}$

előrendezés (HUN) – kvázirendezés (HUN) – preorder (ENG) – quasi order (ENG)

(RL.73) $\text{WEAK_ORDER} \stackrel{\text{def}}{=} \text{REFLEXIVE} \wedge \text{TRANSITIVE} \wedge \text{DICHOTOMOUS}$

gyenge rendezés (HUN) – lineáris előrendezés (HUN) – weak order (ENG) – linear preorder (ENG) – total preorder (ENG)

(RL.74) $\text{WEAK_PARTIAL_ORDER} \stackrel{\text{def}}{=} \text{ANTISYMMETRIC} \wedge \text{TRANSITIVE}$

gyenge részben rendezés (HUN) – partial order (ENG) – weak partial order (ENG)

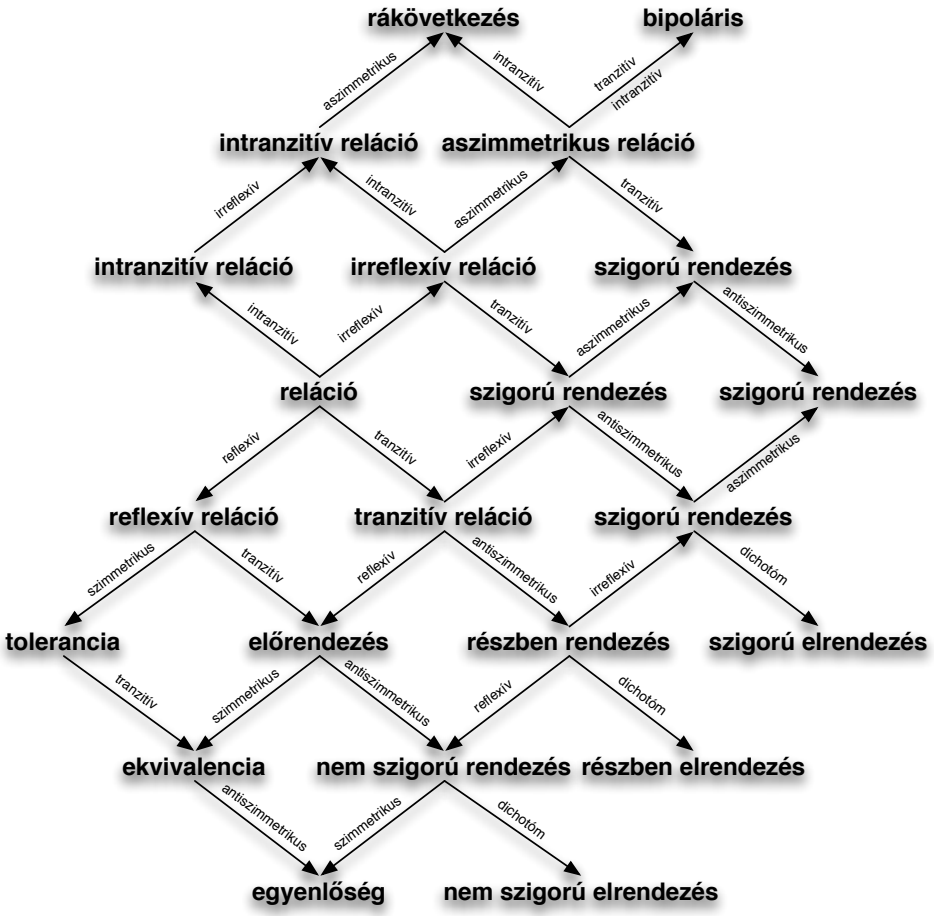
(RL.75) $\text{LINEAR_WEAK_PARTIAL_ORDER} \stackrel{\text{def}}{=} \text{ANTISYMMETRIC} \wedge \text{TRANSITIVE} \wedge \text{DICHOTOMOUS}$

részben elrendezés (HUN) – teljes gyenge részben rendezés (HUN) – total weak partial order (ENG) – linear weak partial order (ENG)

(RL.76) $\text{TOURNAMENT} \stackrel{\text{def}}{=} \text{TRANSITIVE} \wedge \text{DICHOTOMOUS}$

turnament (HUN) – tournament (ENG)

A relációról szóló könyvében a rendezési fogalmakról Jurij A. Srejder készített egy érdekes áttekintő ábrát (jelölve a relációk egymáshoz való viszonyát is) [704], amit korábban további relációkkal egészítettem ki [729], és most még teljesebbé téve mutatok be.[12]



A relációkat mindig a hozzájuk tartozó (tartó)halmazon definiáljuk. Ha veszünk egy halmazt és definiálunk rajta egy vagy több relációt, akkor már reláció **struktúráról** beszélhetünk. A struktúra fogalmának tehát nagyon egyszerű (mert nagyon általános) meghatározása van: egy halmazból és egy (vagy több) reláció együttese.

(RL.77) $\text{STRUCTURE} \equiv \mathfrak{S} \equiv \mathfrak{F} \equiv \mathfrak{T} \stackrel{\text{def}}{=} \left(\mathbb{T}, \{\mathbf{R}_i\}_{i\in\mathbb{N}^+}\right)$

struktúra (ENG) – structure (ENG) – relációs struktúra (ENG) – relation structure (ENG) – modell (HUN) – model (ENG) – algebra (HUN) – algebra (ENG)

Az eddig tárgyalt rendezési fogalmak is felfoghatók valamilyen struktúrának, de persze nemcsak rendezési fogalmak léteznek. Definiálhatunk egy nagyon fontos relációs struktúrát (ami még mindig rendezési fogalom). A neve **fa** vagy fastruktúra [92, 6], de nevezhetjük hierarchiának is [728]. Ezt a struktúrát mind a tudományban, mind a hétköznapi életben rengeteget használjuk, a szervezési, szerveződési technikák egyik legfontosabb típusa.

[12] Összhangba hozni az ábrát és a definíciókat!!!!!!!!!

(RL.78) $\text{TREE}(m,n) \stackrel{\text{def}}{=} \forall n \exists! r \exists! m (\mathbf{R}^*(r,n) \wedge (m \neq r \rightarrow \mathbf{R}(m,n)) \wedge \neg \mathbf{R}^+(n,n))$

fa (HUN) – tree (ENG) – monohierarchia – (HUN) – mono-hierarchy (ENG)

Az első részformula azt határozza meg, hogy a fának egy és csak egy **csúcspontja** (root) van, a formula második összetevője pedig azt definálja, hogy a csúcselemen kívül minden más elemhez egy és csak egy **közvetlenül megelőző** elem (predecessor) tartozik.[13] A harmadik formula az aciklikusságot mondja ki a reláció $\mathbf{R}^+$ tranzitív lezártja segítségével – más formában, mint az aciklikusságra adott korábbi definíció (↬36.o.).

Függvény

A reláció egyik típusának fogható fel a függvény fogalma, de – ahogy azt már említettem – a függvényt is lehetne alapkategóriának tekinteni és le lehetne vezetni belőle a reláció fogalmát. A reláció kategóriatanát azzal folytatom, hogy előbb veszem a függvény mint speciális reláció fogalmát, majd bemutatom, hogy milyen fontosabb függvénytulajdonságok vannak.

(RL.79) $\text{FUNCTION}(m,n) \equiv \mathbf{F}(m,n) \equiv \mathbf{F}(m) = n \equiv \mathbf{F} \equiv \mathbf{F} : \mathbb{Dom} \mapsto \mathbb{Cod} \stackrel{\text{def}}{=} \text{LEFT-TOTAL}(\mathbf{F}) \wedge \text{RIGHT-UNIQUE}(\mathbf{F})$

függvény (HUN) – leképezés (HUN) – function (ENG)

(RL.80) $\text{INVERSE_FUNCTION}(m,n) \equiv \mathbf{F}^{-1}(m,n) \stackrel{\text{def}}{=} \text{RIGHT-TOTAL}(\mathbf{F}) \wedge \text{LEFT-UNIQUE}(\mathbf{F})$

inverz függvény (HUN) – inverz leképezés (HUN) – inverse function (ENG)

(RL.81) $\text{BIUNIQUE}(\mathbf{F}(m,n)) \stackrel{\text{def}}{=} \text{LEFT-UNIQUE}(\mathbf{F}) \wedge \text{RIGHT-UNIQUE}(\mathbf{F})$

egy-egyértelmű (HUN) – egy-egyértelmű (HUN) – biunique (ENG) – unique (ENG)

(RL.82) $\text{BITOTAL}(\mathbf{F}(m,n)) \stackrel{\text{def}}{=} \text{LEFT-TOTAL}(\mathbf{F}) \wedge \text{RIGHT-TOTAL}(\mathbf{F})$

bitotális (HUN) – 1:1-es reláció (HUN) – bitotal (ENG) – bitotal relation (ENG) – 1-1 relation (ENG)

(RL.83) $\text{INJECTIVE}(\mathbf{F}(m,n)) \stackrel{\text{def}}{=} \text{FUNCTION}(\mathbf{F}) \wedge \text{LEFT-UNIQUE}(\mathbf{F})$

injektív (HUN) – egyrétű (HUN) – injective (ENG)

(RL.84) $\text{SURJECTIVE}(\mathbf{F}(m,n)) \stackrel{\text{def}}{=} \text{FUNCTION}(\mathbf{F}) \wedge \text{RIGHT-TOTAL}(\mathbf{F})$

szürjektív (HUN) – ráképzés (HUN) – szuperjektív (HUN) – lefedő (HUN) – surjective (ENG)

(RL.85) $\text{BIJECTIVE}(\mathbf{F}(m,n)) \stackrel{\text{def}}{=} \text{INJECTIVE}(\mathbf{F}) \wedge \text{SURJECTIVE}(\mathbf{F})$

bijektív (HUN) – egyrétűen lefedő (HUN) – bijective (ENG)

(RL.86) $\text{SERIES}(\mathbf{F}(m,n)) \stackrel{\text{def}}{=} \text{FUNCTION}(\mathbf{F}) \wedge \mathbb{D} = \mathbb{N}$

sorozat (HUN) – series (ENG)

Ha speciális relációstruktúraként értelmezünk olyan struktúrát, amelyben ugyanazon a halmazon értelmezünk egy rendezési relációt és egy függvényt, akkor további fontos tulajdonságokat definiálhatunk. Többször lehet fontos az a kérdés, vajon egy rendezés által

[13] Az első részformulában a hierarchia csúcspontjára, illetve a második összetevőben a közvetlenül megelőző elemre megkövetelt egyediséget az unicitás kvantorral biztosítottam (↬30.o.).

kialakított sorrend megőrződik-e akkor, ha definiálunk egy függvényt ugyanazon a halmazon. Ezt írhatjuk le a monotonicitás körébe tartozó tulajdonságokkal.

(RL.87) $\text{MONOTONE}(\mathbf{F}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n (m \leq n \rightarrow \mathbf{F}(m) \preceq \mathbf{F}(n))$

monoton (HUN) – monoton növekvő (HUN) – monotone (ENG) – monotonic (ENG) – monotonically increasing (ENG) – isotone (ENG) – order-preserving (ENG)

(RL.88) $\text{ANTITONE}(\mathbf{F}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n (m \leq n \rightarrow \mathbf{F}(n) \preceq \mathbf{F}(m))$

antiton (HUN) – komonoton (HUN) – monoton csökkenő (HUN) – co-monotone (ENG) – monotonically decreasing (ENG) – antitone (ENG) – anti-monotone (ENG) – order-reserving (ENG)

Művelet

A **műveletek** (operációk) speciális függvényként értelmezhető, de természetesene fordítva is igaz, és definiálható lenne a függvény a művelet fogalmára támaszkodva. A műveletre gyakran gasználják még az operáció kifejezést is, ez magyarázza azt a tényt, hogy a műveletek argumentumaira, változóira operandusként hivatkozunk. A műveletek esetében szinte kizárólagos az ún. infix reprezentációs mód, amikor a művelettel összekapcsolt változókat (operandusokat) a műveleti jel két oldalára írjuk. A következő blokkban bemutatom a művelet és a legfontosabb művelettulajdonságok definícióit.

(RL.89) $\text{OPERATION}(m,n) \equiv \mathbf{O}(m,n) \equiv m\mathbf{O}n \stackrel{\text{def}}{=} \text{FUNCTION}(m,n) \wedge \mathbb{Dom} = \mathbb{Cod}$

művelet (HUN) – operáció (HUN) – belső összekapcsolás (HUN) – operation (ENG)

(RL.90) $\text{ASSOCIATIVE}(\mathbf{O}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n \forall o (m\mathbf{O}(n\mathbf{O}o) = (m\mathbf{O}n)\mathbf{O}o)$

asszociatív (HUN) – társítható (HUN) – associative (ENG)

(RL.91) $\text{COMMUTATIVE}(\mathbf{O}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n (m\mathbf{O}n = n\mathbf{O}m)$

kommutatív (HUN) – felcserélhető (HUN) – commutative (ENG)

(RL.92) $\text{DISTRIBUTIVE}(\mathbf{O}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n \forall o ((m\mathbf{O}(n\mathbf{Q}o) = (m\mathbf{O}n)\mathbf{Q}(m\mathbf{O}o))$

O disztributív művelet Q-ra HUN) – szétoszthető (HUN) – széttagolható (HUN) – distributive (ENG)
$\text{CONCEPT:RELATION:DISTRIBUTIVE}(\mathbf{Q}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n \forall o ((n\mathbf{Q}o)\mathbf{O}m = (n\mathbf{O}m)\mathbf{Q}(o\mathbf{O}m))$

Q disztributív O-ra (HUN) – szétoszthető (HUN) – széttagolható (HUN) – distributive (ENG)

(RL.93) $\text{IDEMPOTENT}(\mathbf{O}(m,n)) \stackrel{\text{def}}{=} \forall m (m\mathbf{O}m = m)$

idempotens (HUN) – önhatványozó (HUN) – azonos hatványú (HUN) – idempotent (ENG)

(RL.94) $\text{ABSORPTIVE}(\mathbf{O}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n (m\mathbf{Q}(m\mathbf{O}n) = m)$

abszorbtív Q-ra (HUN) – elnyelő (HUN) – adjunktív (HUN) – absorptive (ENG)
$\text{ABSORPTIVE}(\mathbf{Q}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n (n\mathbf{O}(m\mathbf{Q}n) = m)$

abszorbtív O-ra (HUN) – elnyelő (HUN) – adjunktív (HUN) – absorptive (ENG)

Relációműveletek

Miután már ismerjük a műveletek tulajdonságait, típusait, visszatérhetünk annak a kérdésnek a tisztázásához, hogy milyen műveletek definiálhatók a relációk között. A relációműveletek azért fontosak, mert ha relációkat kapcsolunk össze relációkkal (műveletek mint speciális relációk segítségével), akkor valamilyen halmazon (a relációk halmazán) értelmezett reláció(k)ról, művelet(ek)ről van szó, vagyis ilyen módon is valamilyen struktúrát állítunk elő. A műveletek között unáris és bináris műveletet definiálunk.

(RL.95) $\text{R-CONVERSION} \equiv \mathbf{R}^{-1}(m,n) \equiv \mathbf{R}^{-c}(m,n) \equiv \mathbf{R}^{-1} \equiv \mathbf{R}^{-c} \stackrel{\text{def}}{=} \forall n \forall n(\mathbf{R}^{-1}(m,n) \leftrightarrow \mathbf{R}(n,m))$

relációkonvertálás (HUN) – konverzképzés (HUN) – inverzképzés (HUN) – megfordítás (HUN) – relation conversion operation (ENG) – relation converse operation (ENG) – relation inverse operation (ENG)

Ha van egy $\mathbf{R}(m,n)$ relációnk, akkor annak a konverze (vagyis a konverzreláció) úgy állítható elő, hogy a relációba kapcsolt elemek, a relátumok sorrendjét felcseréljük egymással. A relációkonvertálást a köznyelvben sokszor új terminus bevezetésével oldjuk meg. Például a GYEREKE(*Béla,Pál*) reláció konverze a GYEREKE^{-1}(*Béla,Pál*), de nem ezt alkalmazzuk, hanem inkább egy új kifejezést alkotunk és a SZÜLŐJE(*Pál,Béla*) relációt használjuk. A definícióból adódik, hogy ha a konverziót kétszer egymás után alkalmazzuk, akkor az eredeti relációhoz jutunk vissza. A konverzreláció a megfordított reláció, a konverzió maga a megfordítási művelet. A konverzió műveletét mind a mondatszerű, mind a szószerű megnyilatkozások szintjén értelmezhetjük. Amikor az összekapcsolt mondatok belső struktúráját vizsgáljuk, akkor **kondicionális konverzióról** beszélünk (↬26.o.), amikor a mondatokon belüli összetevők szerkezetét, akkor **kategoriális konverzióról** (↬56.o.). Az első esetben a művelet értelmezési tartománya a propozíciókból, a második esetben a kategoriális propozíciók alkotórészeiből, predikátumokból, kategóriákból áll.[14]

Timothy Williamson egy tanulmányában kifejtette, hogy a konverzreláció nem tekinthető önálló relációnak, sőt, voltaképpen teljesen megegyezik az alaprelációval [819], vagyis nem tekinthető másnak, mint a szokásostól eltérő, alternatív jelölési konvenciónak. Ahogy a formális nyelvekben ugyanazt a relációt is többféleképpen jelölhetjük (például $\mathbf{R}mn$, $m\mathbf{R}n$ vagy $\mathbf{R}(m,n)$), úgy mutatunk ugyanarra a tartalomra az alapreláció $\mathbf{R}(m,n)$, illetve a konverzreláció $\mathbf{R}(n,m)$ vagy $\mathbf{R}^{-1}(m,n)$ jeleivel, csak éppen más sorrendben rögzítjük a relátumokat – ami csak megállapodás kérdése, semmi más. Ebben az értelemben a konverzrelációt jelző terminusok mind szinonímái az alapreláció terminusainak, tehát a GYEREKE relációnak szinonímája a SZÜLŐJE reláció, mivel tartalmilag teljesen ugyanazt jelentik. Ez természetesen csak logikai szempontból igaz. A hétköznapi nyelvhasználatban ezt az összefüggést nem is érzékeljük, nem is érvényesítjük.

Sok helyen, sokan szinonimaként használják az **inverz** és **konverz** terminusokat, és vannak, akik megkülönböztetik őket egymástól. Én is így teszem, de kiderül majd az is, hogy vannak helyzetek, amikor valóban megegyezik a két művelet eredménye (tehát bizonyos feltételek esetén valóban kezelhetjük őket szinonimaként).

(RL.96) $\text{R-INVERSION} \equiv \mathbf{R}^{-i}(m,n) \equiv \mathbf{R}^{-i} \stackrel{\text{def}}{=} \forall m \forall n(\mathbf{R}^{-i}(m,n) \leftrightarrow \mathbf{R}(\neg m, \neg n))$

relációinvertálás (HUN) – inverzképzés (HUN) – relation inverse operation (ENG)

[14] Lásd a 'Cambridge Dictionary of Philosophy' szótár "converse" lexikai egységét [35].

Ha van egy $\mathbf{R}(m,n)$ relációnk, akkor annak az inverze (vagyis a inverzreláció) úgy állítható elő, hogy a relációba kapcsolt elemek, a relátumok komlementereit (tagadásait) kapcsoljuk össze. A definíció alapján könnyű belátni, hogy két egymás utáni inverzió visszaadja magát az eredeti relációt.
Transzformációs művelet még az **obverzió** (átalakítás) is, de ezt csak később tudom bemutatni és definiálni (↬56.o.), mivel tágyalásához szükség van még további fogalmak ismeretére. Itt elemzem viszont az általános kontrapozíció műveletét, hogy aztán a későbbi fejezetekben kitérjek majd a további típusaira is.

(RL.97) $\text{R-CONTRAPOSITION} \equiv \mathbf{R}^{-p}(m,n) \equiv \mathbf{R}^{-p} \stackrel{\text{def}}{=} \forall m \forall n (\mathbf{R}^{-p}(m,n) \leftrightarrow \mathbf{R}(\neg n, \neg m))$

kontrapozíció (HUN) – transzpozíció (HUN) – contraposition (ENG) – transposition (ENG)

Ha van egy $\mathbf{R}(m,n)$ relációnk, akkor annak a kontrapozíciója (vagyis a reláció kontraponáltja) úgy állítható elő, hogy a relációba kapcsolt elemek, a relátumok komplementereit fordított sorrendben kapcsoljuk össze. A kontrapozíciót is értelmezhetjük mind a mondatszerű, mind a szószerű megnyilatkozások szintjén. Összekapcsolt mondatok szerkezetének átalakításakor kondicionális kontrapozícióról (↬25.o.), mondatokon belüli összetevők átstrukturálásakor kategoriális kontrapozícióról (↬57.o.) beszélünk. A kontrapozíció definíciójából az is látszik, hogy a kontrapozíció nem más, mint a konverzió inverze, illetve az inverzió konverze.[15]

(RL.98) $\text{R-CONTRAPOSITION} \leftrightarrow \text{R-CONVERSION}(\text{R-INVERSION}) \leftrightarrow \text{R-INVERSION}(\text{R-CONVERSION})$

Két további fontos unáris relációműveletet a reláció halmazelméleti kötődése alapján határozhatunk meg. Ha a relációt speciális halmazként fogjuk fel, akkor alkalmazható rá a halmazokon értelmezett komplementálás művelete (↬15.o.), ami az univerzális halmazra vonatkoztatva a reláció tagadásaként is felfogható.

(RL.99) $\text{R-COMPLEMENTATION} \equiv \mathbf{R}^{-t}(m,n) \equiv \mathbf{R}^{-t} \equiv \overline{\mathbf{R}}(m,n)) \equiv \overline{\mathbf{R}} \stackrel{\text{def}}{=} \forall m \forall n (\overline{\mathbf{R}}(m,n) \leftrightarrow \neg \mathbf{R}(m,n)$

relációkomplementálás (HUN) – relation complementation (ENG) – relation transzpozíció (HUN) – relation transposition (ENG)

Ha pedig egymás után alkalmazzuk a reláció konvertálásának és komplementálásának (tagadásának) műveletét, akkor a reláció duális párjához jutunk.[16]

(RL.100) $\text{R-DUAL} \equiv \mathbf{R}^{-d}(m,n)) \equiv \mathbf{R}^{-d} \stackrel{\text{def}}{=} \forall m \forall n (\mathbf{R}^{-d}(m,n) \leftrightarrow \neg \mathbf{R}(n,m) \equiv \overline{\mathbf{R}^{-c}}(m,n)$

duálreláció (HUN) – duális (reláció) (HUN) – dual relation (ENG)

A relációk között bináris műveleteket is értelmezhetünk. A két halmazelméleti művelet (az únió- és metszetképzés), valamint a két aritmetikai művelet (a szorzás és összeadás) mintájára bevezethetjük a következőket.

(RL.101) $\text{R-INTERSECTION} \equiv \times(\mathbf{R}(m,n), \mathbf{Q}(m,n)) \equiv \times(\mathbf{R}, \mathbf{Q}) \equiv \mathbf{R} \times \mathbf{Q} \stackrel{\text{def}}{=} \forall m \forall n (\mathbf{R} \times \mathbf{Q}(m,n) \leftrightarrow (\mathbf{R}(m,n) \wedge \mathbf{Q}(m,n))$

relációmetszés (HUN) – relation intersection (ENG)

[15] Ezen a ponton elméletileg hiányzik még a relációk egymás utáni alkalmazásának vagyis a relációkompozíció műveletének fogalma, de ezt a hiányt még ebben a fejezetben pótlom.

[16] Itt is szükség van a relációkompozíció műveletére.

(RL.102) $\text{R-UNION} \equiv +(\mathbf{R}(m,n),\mathbf{Q}(m,n)) \equiv +(\mathbf{R},\mathbf{Q}) \equiv \mathbf{R}+\mathbf{Q} \stackrel{\text{def}}{=} \forall m \forall n(\mathbf{R}+\mathbf{Q}(m,n) \leftrightarrow (\mathbf{R}(m,n) \vee \mathbf{Q}(m,n))$

relációegyesítés (HUN) – relációúnió (HUN) – relációegyesítés (HUN) – relation union (ENG)

(RL.103) $\text{R-COMPOSITION} \equiv *(\mathbf{R}(m,n),\mathbf{Q}(m,n)) \equiv *(\mathbf{R},\mathbf{Q}) \equiv \mathbf{R}*\mathbf{Q} \stackrel{\text{def}}{=} \forall m \forall n(\mathbf{R}*\mathbf{Q}(m,n)) \leftrightarrow (\mathbf{R}(m,o) \wedge \mathbf{Q}(o,n))$

relációszorzás (HUN) – relatív szorzat (HUN) – kompozíció (HUN) – multiplikáció (HUN) – szuperpozíció (HUN) – relációkompozíció (HUN) – relation product (ENG) – relative multiplication (ENG) – relative product (ENG) – superposition (ENG) – composition (ENG)

(RL.104) $\text{R-ADDITION} \equiv \dagger(\mathbf{R}(m,n),\mathbf{Q}(m,n)) \equiv \dagger(\mathbf{R},\mathbf{Q}) \equiv \mathbf{R}\dagger\mathbf{Q} \stackrel{\text{def}}{=} \forall m \forall n(\mathbf{R}\dagger\mathbf{Q}(m,n)) \leftrightarrow (\mathbf{R}(m,o) \vee \mathbf{Q}(o,n))$

relációösszeadás (HUN) – összeadás (HUN) – relatív összeadás (HUN) – relative addition (ENG) – relative sum (ENG)

Bár a fenti bináris relációműveletek mindegyike szerepet kap a relációalgebra struktúrájában, közülük is "kiemelkedik" a relációszorzat vagy másik gyakori nevén a relációkompozíció művelete. A mindennapi fogalmi készletünkben is előfordulnak olyan fogalmak, amelyek ezt a műveletet használják. Egyszerű példát lehet hozni erre a rokonsági fogalmak köréből. Vegyük a GYEREKE relációt, amivel összeköthetjük Pétert és Katalint (GYEREKE(*Péter,Katalin*)). Ezután vegyük a TESTVÉRE relációt, amivel összekapcsolhatjuk Máriát és Pétert (TESTVÉRE(*Mária,Péter*)). A két relációt összekomponálva a következő eredményt kapjuk: TESTVÉRE(*Mária,Péter*)$*$GYEREKE(*Péter,Katalin*), és hogy ne kelljen ilyen hosszan kifejeznünk ezt a kapcsolatot Mária és Katalin között, használjuk inkább a NAGYNÉNJE relációt.

A relációműveletek segítségével további származtatott relációkat definiálhatunk, amelyek fontosak lesznek a későbbiekben. Az itt definiált tranzitív lezártja reláció voltaképp ugyanúgy relációművelet, mint a fentiek, csak épp itt végtelen sorozat segítségével tudjuk a pontos meghatározást megadni.

(RL.105) $\text{TRANSITIVE_CLOSURE}(\mathbf{R}) \equiv \mathbf{R}^+ \stackrel{\text{def}}{=} (\mathbf{R}^1 = \mathbf{R} \wedge (i > 1 \rightarrow \mathbf{R}^{i+1} = \mathbf{R} * \mathbf{R}^i)) \rightarrow \bigcup_{i \in \mathbb{N}^+} \mathbf{R}^i$

tranzitív lezártja (HUN) – transitive closure (ENG)

(RL.106) $\text{REFLEXIVE_TRANSITIVE_CLOSURE}(\mathbf{R}) \equiv \mathbf{R}^* \stackrel{\text{def}}{=} a$

reflexív tranzitív lezártja (HUN) – reflexive transitive closure (ENG)

Ha egy reláció megegyezik a tranzitív lezártjával, akkor a reláció egyben tranzitív, de a legtöbbször nem ez a helyzet, és a reláció, illetve a tranzitív lezártja különbözik egymástól. A reláció tranzitív lezártja fogalom segítségével a korábbinál jóval egyszerűbben is definiálhatjuk a – racionalitás leírásában oly fontos – konzisztencia fogalmát (↬36.o.), ehhez csak a reláció tranzitívját, illetve aszimmetrikus részét kell összekapcsolni egymással [104].

(RL.107) $\text{CONCISTENT_RELATION} \stackrel{\text{def}}{=} \forall m \forall n(\mathbf{R}^+(m,n) \rightarrow \neg \mathbf{P}(\mathbf{R})(n,m))$

konzisztens reláció (HUN) – consistent relation (ENG)

A formula egyszerűen értelmezhető: egy reláció konzisztens, ha tetszőleges két elemet, m-t és n-t a reláció tranzitív lezártjával összekötünk, akkor a két elem inverz viszonyban (tehát n és m) nem tartozik bele a reláció aszimmetrikus részébe. Ennek alapján Suzumura társaival javasolta bevezetni a reláció konzisztens lezártja reláció fogalmát [104].

(RL.108) $\text{CONCISTENT_CLOSURE} \equiv \mathbf{R}^{\times}(m,n) \stackrel{\text{def}}{=} \forall m \forall n(\mathbf{R}(m,n) \vee (\mathbf{R}^{+}(m,n) \wedge \mathbf{R}(n,m)))$

konzisztens reláció (HUN) – consistent relation (ENG)

Az $\mathbf{R}^{\times}(m,n)$ relációra támaszkodva már kimondható és bizonyítható a tétel, miszerint egy reláció akkor konzisztens, ha megegyezik a konzisztens lezártjával.

Bossert, Sprumont and Suzumura (2005)

1.5.2 Relációalgebra

Minden formális elmélet számára kiemelten fontosak a relációs fogalmak, hiszen a relációk mint predikátumok adják a kijelentéseink általánosítható ismétlődéseit, ezek biztosítják kijelentéseink univerzális jellegét, ezek az állításaink "igéi". A struktúra, a strukturáltság fogalma és a relációra épül rá. Egy ilyen központi fogalom esetében is keresni lehet további absztrakciós lehetőségeket. Felmerül a kérdés: lehet-e a relációk és műveleteik között valamilyen ismert matematikai struktúrát találni? A válasz pedig igenlő, van egy ilyen absztrakt stuktúra, amit **relációalgebrának** neveznek. A reláció és a köztük végezhető műveletek fogalmaira támaszkodva definiálták ezt a struktúrát. Ez azért fontos, mert az **algebra** egy alaposan leírt és nagyon erős tulajdonságokkal rendelkező matematikai struktúra. A matematika nagyon sok mindent tud egy struktúrára nézve mondani, ha kiderül, hogy az algebra. Vigyázni kell azonban, hogy ne tévesszük össze a relációs algebrával, ami a relációs adatbázis-kezelés matematikai alapjait biztosító elmélet. A relációalgebra felépítéséhez a következő relációs fogalmakból álló nyelvre van szükségünk.

(RA.0) $\mathscr{L}_{RA}(\mathbf{R,P,Q}, \mathbf{U}, \mathbf{E}, \mathbf{I}, \mathbf{D}, \neg, \wedge, \rightarrow, =, \mathbf{R}^{-1}, \overline{\mathbf{R}}, \times, +, *, \dagger)$, ahol

$\mathbf{R,P,Q}$	relációk
$\underset{\perp}{\mathbf{F}}$	teljes reláció
$\underset{\perp}{\mathbf{E}}$	üres reláció
$\underset{\perp}{\mathbf{I}}$	azonosság reláció
$\underset{\perp}{\mathbf{D}}$	különbözőség reláció
$\neg$	tagadás logikai művelet
$\wedge$	konjunkció logikai művelet
$\rightarrow$	kondicionális logikai művelet
$=$	egyenlőség reláció
$\mathbf{R}^{-1}$	relációkonvertálás művelet
$\overline{\mathbf{R}}$	relációkomplementálás művelet
$\times$	relációmetszés művelet
$+$	relációegyesítés művelet
$*$	relációszorzás művelet
$\dagger$	relációösszeadás művelet

A fenti nyelvre támaszkodva már definiálhatjuk a relációalgebrát. Ehhez a következő axiómákat kell felvennünk.

(RA.1) $(\mathbf{R} = \mathbf{P} \wedge \mathbf{R} = \mathbf{Q}) \rightarrow \mathbf{P} = \mathbf{Q}$

(RA.2) $\mathbf{R} = \mathbf{P} \rightarrow (\mathbf{R}+\mathbf{Q} = \mathbf{P}+\mathbf{Q} \wedge \mathbf{R}\times\mathbf{Q} = \mathbf{P}\times\mathbf{Q})$

(RA.3) $\mathbf{R}+\mathbf{Q} = \mathbf{Q}+\mathbf{R} \wedge \mathbf{R}\times\mathbf{Q} = \mathbf{Q}\times\mathbf{R}$

(RA.4) $(\mathbf{R}+\mathbf{P})\times\mathbf{Q} = \mathbf{R}\times\mathbf{Q}+\mathbf{P}\times\mathbf{Q} \wedge (\mathbf{R}\times\mathbf{P})+\mathbf{Q} = (\mathbf{R}+\mathbf{Q})\times(\mathbf{P}+\mathbf{Q})$

(RA.5) $\mathbf{R}+\underset{\perp}{\mathbf{E}} = \mathbf{R} \wedge \mathbf{R}\times\underset{\perp}{\mathbf{F}} = \mathbf{R}$

(RA.6) $(\mathbf{R}+\overline{\mathbf{R}} = \underset{\perp}{\mathbf{F}}) \wedge (\mathbf{R}\times\overline{\mathbf{R}} = \underset{\perp}{\mathbf{E}})$

(RA.7) $\neg\underset{\perp}{\mathbf{F}} = \underset{\perp}{\mathbf{E}}$
$\underset{\perp}{\mathbf{F}} = \neg\underset{\perp}{\mathbf{E}}$

(RA.8) $(\mathbf{R}^{-1})^{-1} = \mathbf{R}$

(RA.9) $(\mathbf{R}*\mathbf{Q})^{-1} = \mathbf{R}^{-1}*\mathbf{Q}^{-1}$

(RA.10) $\mathbf{R}*(\mathbf{P}*\mathbf{Q}) = (\mathbf{R}*\mathbf{P})*\mathbf{Q}$

(RA.11) $\mathbf{R}*\underset{\perp}{\mathbf{I}} = \mathbf{R}$

(RA.12) $\mathbf{R}*\underset{\perp}{\mathbf{F}} = \underset{\perp}{\mathbf{F}}$
$\underset{\perp}{\mathbf{F}}*\overline{\mathbf{R}} = \underset{\perp}{\mathbf{F}}$

(RA.13) $((\mathbf{R}*\mathbf{P})\times\mathbf{Q}^{-1} = \underset{\perp}{\mathbf{E}}) \rightarrow ((\mathbf{P}*\mathbf{Q})\times\mathbf{R}^{-1} = \underset{\perp}{\mathbf{E}})$

(RA.14) $\underset{\perp}{\mathbf{D}} = \overline{\underset{\perp}{\mathbf{I}}}$

(RA.15) $\mathbf{R}\dagger\mathbf{Q} = \overline{\overline{\mathbf{R}}*\overline{\mathbf{Q}}}$

hivatkozni, ellenőrizni, magyarázni, példázni, mire jó, miért kell itt

1.6 Kategoriális logika

A kategoriális logika részben a predikátumlogika egyfajta előzményének minősíthetjük, részben a predikátumlogikával párhuzamosan fejlődő, azzal részleges átfedésben levő logikai diszciplinának tekinthetjük. A kategoriális logika a kategorikus ítéletek logikai összefüggéseivel foglalkozik. Arisztotelész dolgozta ki az alapjait. A logika két fő kategóriája a szubjektum és a predikátum. Mivel ezek a terminusok más kontextusokban – legalább részben – más jelentéssel bírnak, ezért itt minősítő jelzőket kapcsolok hozzájuk, és a kategoriális szubjektum és a kategoriális predikátum fogalmát fogom definiálni. A meghatározások előtt azonban tisztázni kell azt a kérdést, mi a kategoriális logika viszonya a predikátumlogikához. A predikátumlogika az általánosabb, mivel bármilyen predikátum használatát megengedi, ezért a kijelentések belső szerkezete is tetszőleges lehet. Vehetünk egy három-argumentumú predikátumot, amit három individuumnév behelyettesítésével lehet megítélhető állítássá formálni.

elad(x,y,v,w) + ‘Laci’, ‘Zoli’, ‘BMW’ – ‘Laci eladta Zolinak a BMW-jét.’

A kategoriális logika speciálisabb, mert nem enged meg bármilyen mondatstruktúrát. Ahogy már említettem, kétféle kategóriát enged meg, a szubjektumot és a predikátumot, és ezeket csak speciális állításszerkezetekbe engedi ‘beletenni”. Például:

Minden veréb [van] madár.

Ez egy lehetséges kategoriális állítás, és az egyik legfontosabb minősége az, ami a magyar nyelvben "láthatatlan": két kategória közti kopulával (a magyar nyelvben elhagyott 'van'-nal) azt fejezzük ki, hogy a két kategória (illetve azok terjedelmébe eső előfordulások) között valamilyen ontológiai kapcsolat van. A kategoriális szubjektumról állítjuk azt, hogy annak valamilyen köze van a kategoriális predikátumhoz. A kategoriális logikában két kategóriát kapcsolatára mondunk valamit úgy, hogy valamilyen módon (mennyiségileg vagy minőségileg) módosítani lehet a kategóriákat, de a két kategóriát mindenképpen a kopulával kapcsoljuk össze.

(GL.0) CATEGORICAL_SUBJECT ≡ LOGICAL_CONSTANT ≡ INDIVIDUAL_NAME ≡ **S**

kategoriális szubjektum: amire a kategoriális kijelentés szerkezetén belüli másik kategória, a kategoriális predikátum vonatkozik.

(GL.1) CATEGORICAL_PREDICATE ≡ RELATION ≡ FUNCTION ≡ **P**

kategoriális predikátum: amivel a kategoriális kijelentés szerkezetén belüli másik kategóriára, a kategoriális szubjektumra vonatkozóan állítunk valamit.

Az nyilvánvaló, hogy a kategoriális predikátum a predikátumlogikai értelemben felfogott predikátum fogalma alá sorolható, a kérdés az, hogy mit mondhatunk a kategoriális szubjektumra vonatkozóan. A válaszhoz tudnunk kell, hogy ilyen szerkezeti feltételek tehetünk szinguláris állításokat, amikor a szubjektum egy individuum neve, illetve általános állításokat, amikor a szubjektum valamilyen predikátumlogikai értelemben vett predikátum. Rögzítenünk kell tehát, hogy mind a kategoriális szubjektum, mind a kategoriális predikátum lehet "hagyományos" predikátum.

A kategoriális logika kidolgozásakor és továbbépítésekor azonban nem ez a kérdés foglalkoztatta a logikusokat. Kezdetekben két szempont szerint osztották fel a kategoriális kijelentéseket, és ez alapján tettek meg fontos – évezredeken át fenntartott – állításokat. Egyfelől figyeltek arra, hogy partikuláris vagy univerzális állításról van-e szó, a másik szempont szerint pedig azt nézték, hogy az állításba az adott kategóriákat állító vagy tagadó módban vették-e föl. Az univerzális kijelentések a kategória terjedelmébe sorolt minden előfordulásra vonatkoztak, míg a partikuláris kijelentések hatóköre csak néhány ilyen elemre irányult. Az állító, pozitív vagy megerősítő kijelentésekben a kopula szerepelt a kategóriák között, míg a tagadó, negatív kijelentésekben a negált, tagadott kopula. A négyféle kijelentést az alábbi táblázatba rendezhetjük:

kijelentéstípus	jele	latin mondat	magyar mondat
egyetemes állító	**(A)**	Omne S est P.	Minden S [van] P.
egyetemes tagadó	**(E)**	Nullum S est P.	Egyetlen S sem P.
partikuláris állító	**(I)**	Quoddam S est P.	Néhány S [van] P.
partikuláris tagadó	**(O)**	Quoddam S non est P.	Néhány S nem P.

Az évezredeken át formálódó közös tudás formális szempontból inkonzisztenciát tartott fent, mert az egyetemes tagadás nem a konzisztens alakot tartotta számon (ami a 'Minden S [van] nem P' lett volna), de tartalmilag mindvégig helyesen kezelték ezeket az állításokat.[17] Arisztotelésznél még nem így volt, de a követői a négyféle állítást már

[17] kell-e ez, ide kell-e ez?
Azt is meg kell itt jegyezzem, hogy akkor is előfordult inkonzisztencia, amikor a kopula tagadása helyett a

egyetlen rendszerbe szervezték szemléletes ábrával reprezentálva az állítások közti összefüggéseket, és innentől kezdve logikai négyszögként hivatkoztak rájuk. Ezt a következő fejezetben bontom ki alaposabban. Előtte azonban felveszem az új fogalmakat az ontológiánkba.

(GL.2) UNIVERSAL_AFFIRMATIVE_PROPOSITION ≡ $\mathbf{S(A)P} \equiv \mathbf{(A)} \stackrel{\text{def}}{=} \forall m(\mathbf{S}(m) \rightarrow \mathbf{P}(m))$

egyetemes állító kijelentés (HUN) – universal affirmative proposition (ENG)

(GL.3) UNIVERSAL_NEGATIVE_PROPOSITION ≡ $\mathbf{S(E)P} \equiv \mathbf{(E)} \stackrel{\text{def}}{=} \forall m(\mathbf{S}(m) \rightarrow \neg\mathbf{P}(m))$

egyetemes tagadó kijelentés (HUN) – universal negative proposition (ENG)

(GL.4) PARTICULAR_AFFIRMATIVE_PROPOSITION ≡ $\mathbf{S(I)P} \equiv \mathbf{(I)} \stackrel{\text{def}}{=} \exists m(\mathbf{S}(m) \rightarrow \mathbf{P}(m))$

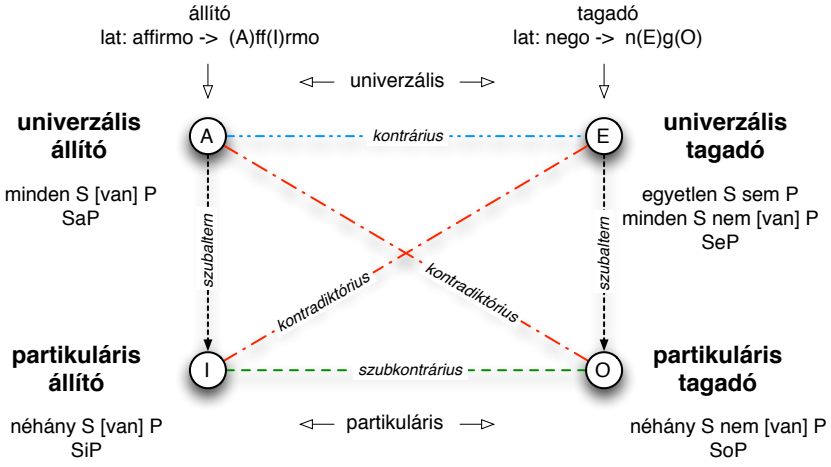
partikuláris állító kijelentés (HUN) – particular affirmative proposition (ENG)

(GL.5) PARTICULAR_NEGATIVE_PROPOSITION ≡ $\mathbf{S(O)P} \equiv \mathbf{(O)} \stackrel{\text{def}}{=} \exists m(\mathbf{S}(m) \rightarrow \neg\mathbf{P}(m))$

partikuláris tagadó kijelentés (HUN) – particular negative proposition (ENG)

Logikai négyszög

A kategoriális logika alapjait Arisztotelész fektette le, de ő még nem rajzolt ábrákat. Pár évszázad múlva Apuleius, majd Boethius volt az, aki a négyféle kijelentés közti kapcsolatokat ábrában kezdte mutatni. Az ilyen az ábrákat szokták Apuleius négyszögének vagy Boetius-féle négyzetnek is nevezni, de ezeken túl gyakran használják még a **logikai négyszög**, a logikai négyzet, valamint az ellentétek négyszöge vagy az oppozíciós négyszög kifejezéseket is [81][520].[18] Az ábra így néz ki:



predikátum (tehát egy kategória) tagadásáról beszéltek (ez nyilvánul meg a modernkori logikai formulákban is egyébként). Abban az értelemben tartom ezt inkonzisztenciának, hogy elméletileg nem szabad összemosni a tagadás műveletét a komplementálással. Előbbi ugyanis a kijelentésekre vonatkozik, utóbbi viszont terminusokra (predikátumokra, kategóriákra). A negált terminus, negált kategória értelmes fogalom, de szigorú értelemben komplementált vagy komplementer terminusnak kellene neveznünk. Amennyiben a kontextusok alapján ez egyértelműen kiderül, akkor természetesen megengedhető ez a szinonimitás.

[18] Mivel a négyszög továbbfejlesztésekor tartalmilag sokféle irányban elmozdultak, az eredeti négyszöget ma már **kvantifikációs négyszögnek** is nevezik.

A négyszög négy sarokpontját elnevezték ((**A**)-, (**E**)-, (**I**)-, (**O**)-pontoknak), és így hivatkoznak rájuk a mai napig. Az ábrán már feltüntettem, de az eddigiekből még nem derült ki, miért is nevezték ellentétek négyszögének ezt az alakzatot. Ugyan Arisztotelész nem használta még ezt a fenti vizuális megjelenítést, azonban ő volt az, aki alaposan leírta a kijelentések közti viszonyrendszert, és kiderült, a kijelentéstípusok közti kapcsolatok majd mindegyike valamifajta ellentétet fejez ki. A logikai négyzet elemei közt évszázadokon keresztül három ellentétes viszonyt tételeztek: *(i)* a **kontradiktórius** (ellentmondó), *(ii)* a **kontrárius** (ellentétes), valamint *(iii)* a **szubkontrárius** (alárendelt ellentétes) kapcsolatot. Ezeken felül rögzítettek még egy alárendeltségi viszonyt kifejező **szubaltern** (alárendelt), **szuperaltern** (fölérendelt) kapcsolatpárt is. Ezen kapcsolatokat, illetve modernkori megfelelőiket az alábbi táblázatban mutatom meg:

ellentét	formulája	logikai művelet neve	jele
szubkontrárius	$\neg(\neg A \wedge \neg B)$	diszjunkció	$A \vee B$
kontradiktórius	$\neg(A \wedge B) \wedge \neg(\neg A \wedge \neg B)$	biszubtrakció	$A \not\leftrightarrow B$
kontrárius	$\neg(A \wedge B)$	exklúzió	$A \not| B$
szubaltern	$A \rightarrow B$	implikáció	$A \rightarrow B$

A különféle oppozíciók meghatározását természetesen nem a táblázatban látható modern képletekkel fejezték ki évszázadokon keresztül, hanem szöveges értelmezést adtak. Kontradiktórikus viszony esetén a két elem nem lehet egyszerre sem hamis, sem igaz, vagyis csak az lehetséges, hogy vagy az egyik igaz és a másik hamis vagy fordítva. A kontrárius ellentét azt jelenti, hogy a két összetevő nem lehet egyszerre igaz, de lehet egyszerre hamis (és az is lehetséges, hogy az egyik igaz és a másik hamis vagy fordítva). A szubkontrárius ellentét azt az esetet fejezi ki, amikor a két elem nem lehet egyszerre hamis, de lehet egyszerre is és külön-külön is igaz. Végül a szubaltern kapcsolat azt fejezi ki, hogy ha valami igaz, akkor annak a szubalternje is igaz (a szuperaltern pedig ennek a konverze).

A logikai négyzethez hozzátartozott még pár átalakítási szabály is (pl. az (**E**)- és (**I**)-sarkokban a konverzió művelete is igaz állítást eredményez), de ezeket itt nem részletezem, inkább külön fejezetben tárgyalom majd.

A logikai négyzet, a maga ellentmondás-fogalmaival nagyon népszerű volt évszázadokon keresztül. Idővel kiterjesztették használatát más – a kvantifikációs kapcsolatok kezelésén túlmutató – célokra is. Erről még bővebben is írni fogok a későbbiekben. Előtte azonban meg kell említenem, hogy a logikai négyszög elemei közti kapcsolatokat a modern logika revízió alá vette, és – legalább részben – felülbírálták az ókor óta hangoztatott téziseket. Kiderült ugyanis, hogy az üres, illetve az univerzális terminusokkal (vagy osztályokkal), valamint az egzisztenciális import jelenségével problémák vannak – olyannyira, hogy e felfedezések óta sokan már csak a kontradiktórius kapcsolatot tartják érvényesnek [563].

A kritikák egyik vonulata úgy szólt [635], hogy ha az (**A**)-pontba egy univerzális osztályt, tehát egy mindig igaz logikai igazságot, azaz tautológiát teszünk (‘Minden négyzet derékszögű.’), akkor arról tudjuk, hogy sosem lehet hamis. Egy tautológiával szemben azonban semmi sem állhat kontrárius viszonyban, hiszen a tautológia sosem hamis, viszont a kontrárius kapcsolat létezéséhez ezt kell feltételeznünk, hogy a két összetevő lehet egyszerre hamis. Ugyancsak ellentmondáshoz jutunk, ha az (**I**)-sarokba egy mindig hamis állítást, tehát egy kontradikciót teszünk (‘Néhány négyzet köralakú’). Egy ellentmondás

sosem lehet igaz, így nem kerülhet szubkontrárius kapcsolatba egy másik kijelentéssel, hiszen utóbbi tételezése azt kívánná meg, hogy a két kijelentés egyszerre legyen igaz.

Gondok adódnak abban az esetben is, ha egy üres osztályt, üres terminust teszünk be az **(A)**-kijelentés szubjektumába [402]. Mivel egy üres terminusról (üres osztályról) minden állítható[19], így az **(A)**-mondat igaz. Az üres terminus esetén viszont nem lehet igaz az **(I)**-sarok, hiszen ha nincs az S terjedelme alatt semmi, akkor nem lehet igaz a 'Néhány S [van] P' kijelentés. Ekkor viszont az **(I)**-pont kontradiktórius párja az **(E)**-mondatnak igaznak kell lenni, amiből viszont az következne, hogy a vele szubaltern viszonyban álló **(O)**-pontnak is igazságot kellene tartalmaznia. Az **(O)**-sarokban viszont **(A)** igazságból következően – és a köztük levő kontradiktórius viszony miatt – hamis állításnak kellene szerepelni, így ellentmondáshoz jutunk. Az üres osztályok megjelenése esetén tehát felborul a logikai négyzet rendje.

———————-

???????????

Egy kijelentéseknek akkor van egzisztenciális importja, ha a kijelentés igazsága "megköveteli", hogy a szubjektum terjedelmébe tartozzon legalább egy indivuduum, előfordulás. Az univerzális kategorikus kijelentések esetében ezt nem tudjuk mindig garanztálni, ezért kimondhatjuk a tézist, hogy ezeknek nincs egzisztenciális importja.

???????????

———————-

a Hansson féle ellentmondásbiztos megoldást alkalmazni itt is

Russel a denotálásról szóló tanulmányában megoldást javasol a denotálás, referálás értelmezésére üres termek, nem létező objektumok esetében [626] Russel megoldását lehet-e itt alkalmazni?

másról szól Roderick Chisholm cikke, de van benne egy használható megoldás az üres termek kiküszöbölésére. Ahelyett, hogy felvegyünk a predikátum mellé egy ürességet kizáró állítást, a predikátumot would operátorral kell értelmezni: vagyis ha van előfordulása a predikátumnak, akkor érvényesíteni lehet magát a predikátumot. [150, 147]

a szubaltern reláció sem mindig érvényes, lásd Ruzsa könyvének 1.2.3 szakasza

———————-

* * *
* * *

A logikai négyzet eredeti formájában kvantifikációs állításokat tartalmazott, ezért is nevezték el idővel kvantifikációs négyzetnek. Azért kellett új nevet kitalálni, mert az évszázadok során másféle kijelentésekkel, kategóriákkal töltötték fel a logikai négyzet sarokpontjait, amelyek esetében ugyanúgy működni látszott a séma, az új négyszögek elemei és a köztük levő kapcsolatok ugyanazt a mintázatot mutatták, mint az eredeti négyzet.

????????????

A modális logika kifejlődésével párhuzamosan az egyes modális logikák kulcskategóriáit is négyszögekbe rendezték.

	(A)	**(E)**	**(I)**	**(O)**
alethikus logika	szükségszerű	lehetetlen	lehetséges	nem-szükségszerű
deontikus logika	kötelező	tilos	megengedett	nem-kötelező

[19] Az egyetemes állítás mindig átalakítható a két kategóriát kondicionális művelettel összakapcsoló kijelentéssé (a 'Minden S [van] P' propozíció ekvivalens azzal, hogy 'Ha S, akkor P'), és ha egy kondicionális előtagja hamis (üres), akkor abból következik az utótag igazsága.

Sok mindenre használták még a logikai négyzetet, például alkalmazták a cselekvéselmélethez egészen közel áll fogalmakra is. Canterburyi Szent Anzelm olyan cselekvéstipológiát dolgozott ki, ami a logikai négyzetre épült. A cselekvéselmélet szempontjából vett fontossága miatt az anzelmiánus elméletet részletesebben bemutatom majd, de nem itt, hanem a cselekvsé fogalmáról szóló fejezetben (↬268.o.)

Logikai hatszög

A logikai négyzettel kapcsolatban felmerültek másfajta problémák is, ezek azonban sokkal inkább a továbbfejlesztési lehetőségek keresésére ösztönözték a kutatókat, mint a négyszög-gondolat feladására. A négyszög hatszögesítését Augustin Sesmat és Robert Blanché egymástól függetlenül, közel egyidőben végezte el az 1950-es években, de közel ötven éven keresztül eredményeiknek nem volt jelentős hatása a tudományos gondolkozásra. Csak a 2000-es években változott ez a helyzet [81].

Kiderült, hogy a logikai négyzet négy sarokpontja egyszerre volt kevés is és sok is a tényleges összefüggések pontos megragadására. Zavaró volt, hogy a hétköznapi és a tudományos-logikai nyelvhasználat több ponton eltért egymástól. Az **(I)**-sarokba tartozó kategóriákat sokszor az **(A)**-sarokba tartozó kategóriákkal ellentétesnek érezzük, szemben a logikai négyzet "állításával", mely szerint a két pont állításai között szubaltern viszony létezik, ami egyáltalán nem számít ellentétes kapcsolatnak (a megengedett vagy a lehetséges fogalmát sokszor sokkal inkább érezzük ellentétesnek, mint alárendeltnek a kötelező vagy a szükségszerű fogalmához képest). Magyarázatra szorult az a furcsa tény is, hogy az **(O)**-pontba került kategóriáknak szinte sosem volt megfelelő hétköznapi terminusa. A modális logikában például a nem-szükségszerű, nem-kötelező "mesterséges" terminusokat kellett használni a hiányzó természetes szavak hiányában.

Az ilyen és ehhez hasonló problémák feloldhatónak látszottak azáltal, hogy két új pólust veszünk fel a négyzethez, és ezáltal olyan logikai hatszöghöz jutunk, amely két háromszögből áll. Első lépésként az "eredeti" logikai négyszög helyett érdemesebb egy olyan logikai háromszöget képezni, amelyben megmarad az **(A)**- és az **(E)**-sarok, és ezek mellé kell egy új pólust felvenni. Ezt az új pólust – Blanché után – **'Y'** karakterrel kezdték el jelölni. Az az új logikai háromszög nagyon hasznosnak ígérkezett. Egyre több olyan fogalomhármast kezdtek felfedezni a gondolkodás, a kategóriaképzés legkülönfélébb területein, amelyek mind egyidejű igazságukat kölcsönösen kizáró, tehát kontrárius háromszöget alkottak egymással. Íme néhány példa rá:

	(A)	**(E)**	**(Y)**
kvantifikáció	minden	semmi	valami
tér	mindenhol	sehol	valahol
idő	mindig	soha	valamikor (néha)
alethikus modaitás	szükségszerű	lehetetlen	kontingens
deontikus modalitás	kötelező	tiltott	megengedett

A kontrárius háromszögek mellett természetesen képezhetők voltak a szubkontrárius háromszögek is. Ehhez új elemként az **(U)**-sarokpontot vették fel, ami az "eredeti négyzet" **(I)** és **(O)** pólusait kötötte össze. A két háromszög egymáshoz illesztésével pedig már felrajzolhatóvá vált az új minőség, a logikai hatszög, amely egyben tartalmazta az eredeti logikai négyszöget is, de benne volt még a két "új" logikai háromszög is. Kiderült az is,

hogy az új pólusok felvételével a hatszög végpontjai mind kifejezhetőek a szomszédos sarokpontok segítségével.

(GL.6) UNIVERSAL_PROPOSITION ≡ **S(U)P** ≡ **(U)** $\stackrel{\text{def}}{=}$ **(A)** ∨ **(E)**

egyetemes kijelentés (HUN) – universal proposition (ENG)

(GL.7) PARTICULAR_PROPOSITION ≡ **S(Y)P** ≡ **(Y)** $\stackrel{\text{def}}{=}$ **(I)** ∧ **(O)**

partikuláris kijelentés (HUN) – particular proposition (ENG)

Az új elemek felvétele után felvetődik a kérdés, vajon milyen összefüggések léteznek az új (és régi) kategóriák között. Kiderült, hogy a már jelzett kapcsolatokon túl a szubaltern relációk még több sarokpont között behúzhatóak. A teljes kapcsolatrendszert mutatja az 1.2. ábra.

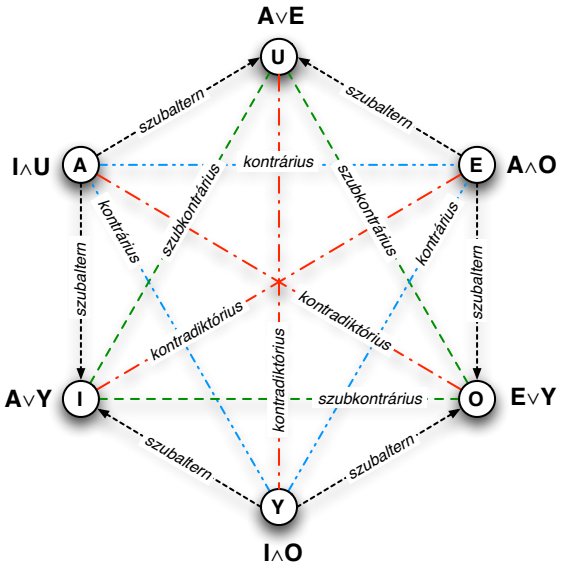


Figure 1.2: logikai hatszög

A hatszög kapcsolatrendszerét felvázolva észrevehetünk még egy további rejtett összefüggést is. A hatszögbe nemcsak egy, hanem összesen három négyszög van "elrejtve", és a két "új" négyszög ugyanolyan belső szerkezettel rendelkezik, mint az "eredeti" logikai négyszög. Az ellentétes pólusaikban kontradiktórius, a szomszédos sarokpontok közt kontrárius vagy szubkontrárius, valamint szubaltern viszonyokat találhatunk. A három négyszöget kiemeltem a hatszögből és eredeti helyzetüket megtartva mutatom be őket az 1.3. ábrán.

Idővel kiderült, hogy Blanché logikai hatszöge nagyon sok területen érvényes. A logikai operátorokat vizsgálva Blanché rájött arra, hogy a tíz legfontosabb operátor ábrázolható két logikai hatszög segítségével. A nullad- és az elsőrendű operátorok (a tautológia és ellentmondás, valamint a négyféle projekciós művelet) nélkül marad tíz műveletünk, ezek a "tisztán" másodrendű logikai operátorok, és ezekkel tudta Blanché egy összefüggő logikai szerkezetbe rendezni (1.4. ábra).

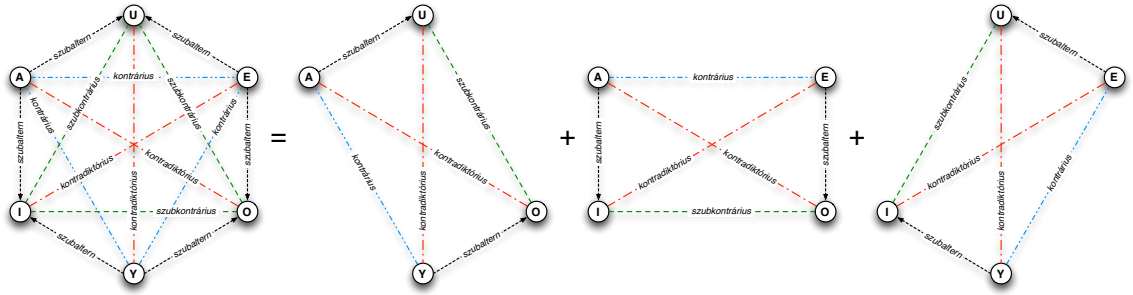


Figure 1.3: logikai hatszög és négyszögek

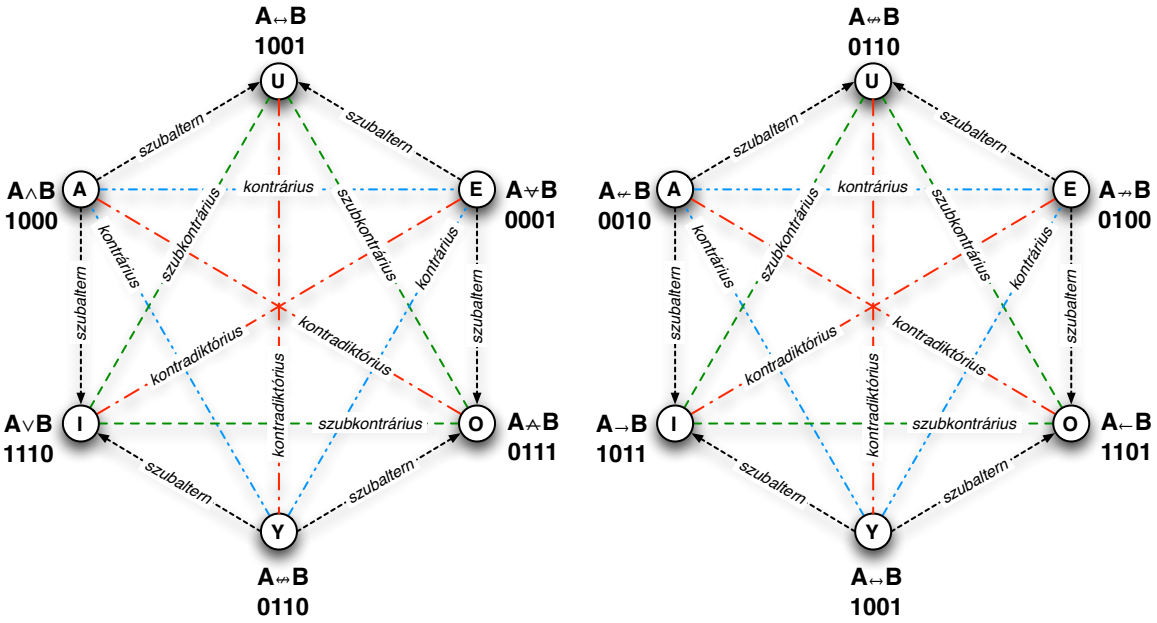


Figure 1.4: Blanché tíz logikai operátora két hatszögbe fogva

A geometriai szemlélet érvényesítését, további hasznosítását azonban itt már nem lehetett leállítani.[20] Amíg a logikai operátorok egymáshoz való viszonyát Blanché még úgy fejezte ki, hogy az operátorok közül tizet két hatszögbe foglalt bele, addig Sauriol már mind a tizenhat operátort egyetlen alakzatban tudta "megmutatni" egy tetrahexahedron segítségével. A tetrahexahedron egy kockából és hat darab piramisból álló geometriai forma, aminek a drótváza hat darab hatszögből is összeállítható. A háromdimenziós térbeli alakzatot az 1.5. ábra mutatja, és hogy könnyebben felismerhetők legyenek a benne "rejtőző" hatszögek, az 1.6. ábrán egyenként kiemelve megmutatom a hatszögek helyzetét a tetrahexahedronon belül.

⁂

A logikai hatszögek, majd a további geometriai formák alkalmazhatósága sikere bizonyos értelemben "kijelölte" a továbblépési lehetőségeket a geometriai általánosítás irányába, és

[20] A logikai relációk háromdimenziós megjelenítésről lásd: [692].

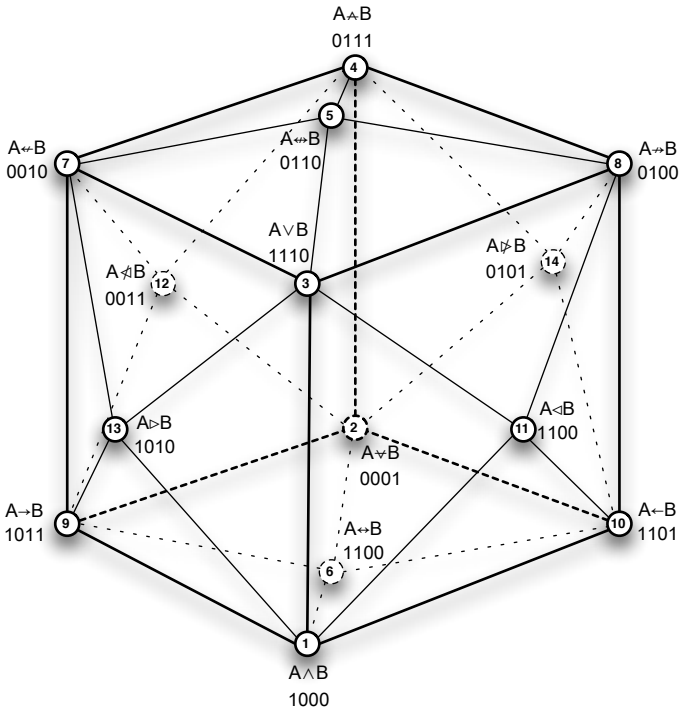


Figure 1.5: logikai operátorok tetrahexahedronja

megjelentek azok a cikkek, amelyek logikai n-szögekről és a problémakör geometrizálásáról szóltak [517], [519]. Ezekkel azonban itt nem foglalkozom.

Közvetlen következtetések, átalakítási műveletek kategoriális kijelentések esetén

A logika négyszög kapcsán is sokat foglalkoztak a kategoriális kijelentések átalakítási lehetőségeivel. Azért volt mindez fontos, mert ezek olykor közvetlen következtetési szabályként is működtek [35, 895]. Előre fontos hangsúlyozni, hogy ezek az átalakítási szabályok nem mindig eredményeznek igaz állításból igaz állítást, tehát nem minden esetben működnek következtetési szabályként. Ahhoz, hogy mind a négy kategorikus állítástípusra egyszerre megfogalmazhassuk majd az átalakítási szabályokat, illetve az érvényes következtetési formulákat, érdemes bevezetni a (**C**) általános kategorikus propozíció fogalmát. Ebből persze egyből kettőt kell definiálnunk, hogy – szükség esetén – jelezni tudjuk, melyik esetről teszünk állítást, négy vagy hat típusról, a logikai négyszögről vagy a logikai hatszög elemeiről beszélünk.

(GL.8) $\text{GENERAL_C-PROPOSITION} \equiv \mathbf{S(C)_4P} \equiv \mathbf{(C)_4} \stackrel{\text{def}}{=} \mathbf{(A)} \veebar \mathbf{(E)} \veebar \mathbf{(O)} \veebar \mathbf{(I)}$

(GL.9) $\text{GENERAL_C-PROPOSITION} \equiv \mathbf{S(C)_6P} \equiv \mathbf{(C)_6} \stackrel{\text{def}}{=} \mathbf{(A)} \veebar \mathbf{(E)} \veebar \mathbf{(O)} \veebar \mathbf{(I)} \veebar \mathbf{(U)} \veebar \mathbf{(Y)}$

Az új fogalom bevezetésének csak annyi értelme van, hogy segítségével egyszerűsíthessünk a későbbi definíciókon, a fogalomnak önálló szemantikája nincs, csak egy belső mintázattal rendelkező változóként érdemes felfogni. A korábban már bevezetett átalakítási műveleteket itt – igaz, új értelmezési tartományon – újraértelmezhetjük. A konverzió meghatározásában a korábbiakhoz képest (↬43.o.) csak annyi változásra van itt

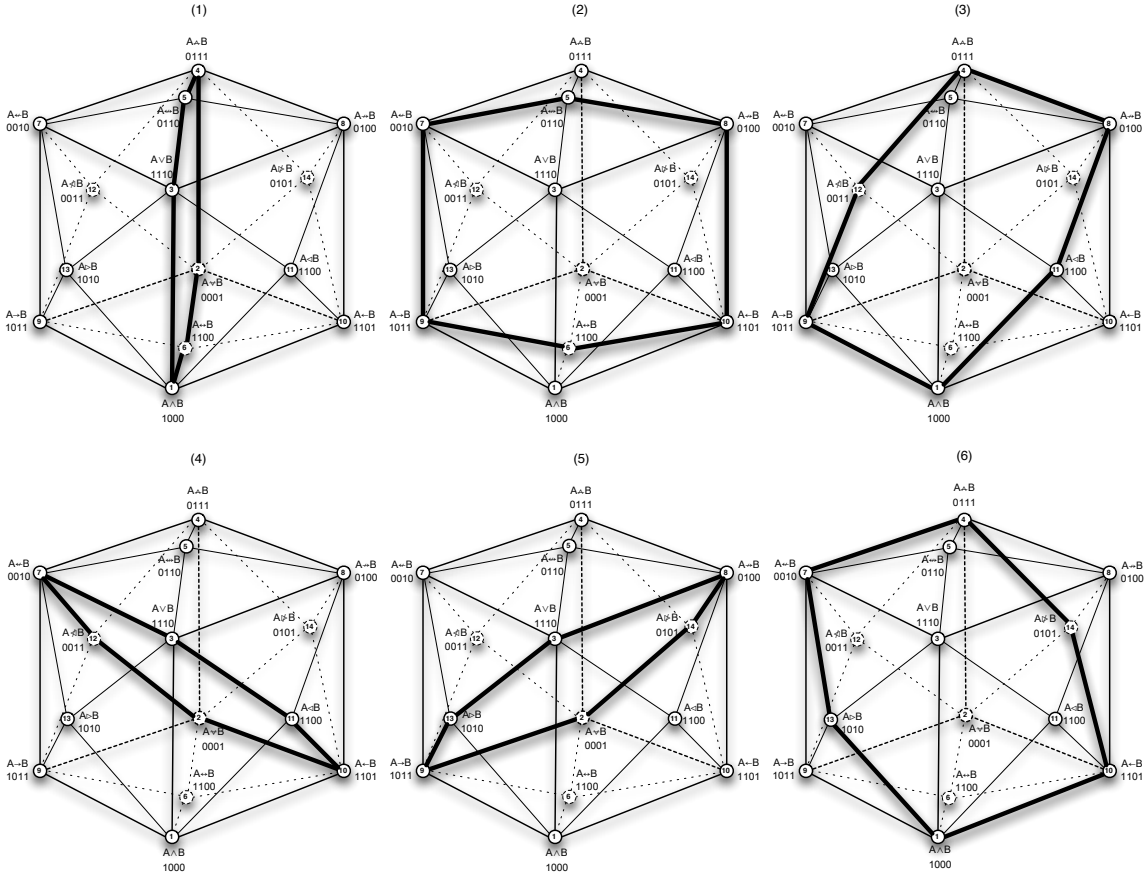


Figure 1.6: hatszögek a tetrahexahedronban

szükség, hogy az értelmezési tartományt le kell szűkíteni a kategoriális kijelentések halmazára. A műveletet az általános kategoriális propozícióváltozó segítségével definiálhatjuk:

(GL.10) $\text{C-CONVERSION} \equiv \mathbf{S(C)}_4^{-1}\mathbf{P} \equiv \mathbf{(C)}_4^{-1} \stackrel{\text{def}}{=} \forall m\big(\mathbf{S(C)}_4^{-1}\mathbf{P} \leftrightarrow \mathbf{P(C)}_4\mathbf{S}\big)$

kategoriális konverzió (HUN) – category conversion (ENG)

A kategoriális konverzió olyan konverzió, amelynek értelmezési tartománya a kategoriális kijelentések halmaza. A kategoriális konverzió praktikusan annyit tesz, hogy a kijelentésben felcseréljük a szubjektumot és a predikátumot. Ha a tradicionális oppozíciós négyzet (**E**)-propozíciója (‘Egyetlen S sem P.’) igaz, akkor annak megfordítása (‘Egyetlen P sem S.’) is igaz, tehát ez a konverzió közvetlen következtetési szabályként is alkalmazható. Ugyanúgy elmondható ez az (**I**)-propozícióra is (‘Néhány S [van] P.’), mert annak a megfordítása (‘Néhány P [van] S.’) megintcsak igaz. A négyzet másik két állítása is megfordítható, de ezekben az esetekben nem áll fenn a konverz kijelentések igazsága.

Az átalakítási műveletek korábbi tárgyalásakor nem mutattam be az **obverzió** műveletét, mert ez igazából csak a kategoriális kijelentések esetében értelmezhető. Az új művelet bevezetéséhez szükség van a kategóriák (relációk) komplementumaira is, és a kategoriális predikátum tagadására.

(GL.11) $\text{OBVERSION} \equiv \mathbf{S(C)}_4^{-o}\mathbf{P} \equiv \mathbf{(C)}_4^{-o} \stackrel{\text{def}}{=} \forall m\big(\mathbf{S(C)}_4^{-o}\mathbf{P} \leftrightarrow \overline{\mathbf{S}}\mathbf{(C)}_4\neg\overline{\mathbf{P}}\big)$

obverzió (HUN) – átalakítás[21] (HUN) – category obversion (ENG)

Az obverzió olyan átalakítási művelet, ami megtartja a szubjektum és predikátum sorrendjét, de kicseréli őket a komplementumaira, valamint a predikátum komplementumát tagadva állítja. A propozicionális logikában az obverzió műveletét nem használják. Vélhetőleg azért nem, mert az állításlogikában hiányzik a pozitív és negatív (állító és tagadó) kijelentések megkülönböztetése, ami viszont a kategoriális logikában adott.

A kategoriális kijelentések színterén is értelemezhető a kontrapozíció művelete, ami itt kifejezhető a konverzió és obverzió segítségével is. Ezt azonban akkor lehet jól szemléltetni, ha használjuk a részleges kontrapozíció műveletét.

(GL.12) $\text{PARTIAL_C-CONTRAPOSITION} \equiv \mathbf{S(C)}_4^{-r}\mathbf{P} \equiv \mathbf{(C)}_4^{-r} \stackrel{\text{def}}{=} \forall m\big(\mathbf{S(C)}_4^{-r}\mathbf{P} \leftrightarrow \neg\overline{\mathbf{P}}\mathbf{(C)}_4\overline{\mathbf{S}}\big)$

részleges kategoriális kontrapozíció (HUN) – partial category contraposition (ENG)

A részleges kategoriális kontrapozíció az a művelet, amely egyszer-egyszer alkalmazza az obverzió és a konverzió műveletét. A művelet "képlete" tehát egyszerű: részleges kontrapozíció = átalakítás + megfordítás.

Ide valami értékelés kell még. 'Minden bogár rovar.'

(GL.13) $\text{C-CONTRAPOSITION} \equiv \mathbf{S(C)}_4^{-p}\mathbf{P} \equiv \mathbf{(C)}_4^{-p} \stackrel{\text{def}}{=} \forall m\big(\mathbf{S(C)}_4^{-p}\mathbf{P} \leftrightarrow \neg\mathbf{P}\mathbf{(C)}_4\neg\mathbf{S}\big)$

teljes kategoriális kontrapozíció (HUN) – kategoriális kontrapozíció (HUN) – categoric contraposition (ENG)

A teljes kategoriális kontrapozíció olyan kontrapozíció, amelynek az értelmezési tartománya a kategoriális kijelentések halmaza. A kategoriális kontrapozíció praktikusan annyit tesz, hogy megcseréli a kijelentés szubjektumának és predikátumának komplementumait. A művelet felírható a következő alakban is: kategoriális kontrapozíció = átalakítás + megfordítás + átalakítás.

(GL.14) $\big((\mathbf{C})_4 = (\mathbf{A}) \veebar (\mathbf{C})_4 = (\mathbf{O})\big) \rightarrow \big(\text{C-CONTRAPOSITION} \leftrightarrow \text{OBVERSION}\big(\text{C-CONVERSION}\big(\text{OBVERSION}\big)\big)\big)$

A kondicionális alak azt jelzi, hogy a fenti formula nem minden esetben igaz: a négyből csak két esetben. Ha az **(A)**-propozíció igaz, akkor annak kategoriális kontraponáltja is igaz, tehát ez az átalakítás használható közvetlen következtetési szabályként. Ugyanez áll az **(O)**-propozícióra is, hiszen a kontraponáltja igaz.

A relációműveletek között szerepelt az inverzió is. Ezt is definiálhatjuk itt, bár a jelentősége, hasznossága kisebb a többi átalakítási szabályhoz képest. A definíciót így adhatjuk meg:

(GL.15) $\text{C-INVERSION} \equiv \mathbf{S(C)}_4^{-i}\mathbf{P} \equiv \mathbf{(C)}_4^{-i} \stackrel{\text{def}}{=} \forall m\big(\mathbf{S(C)}_4^{-i}\mathbf{P} \leftrightarrow \overline{\mathbf{S}}\mathbf{(C)}_4\overline{\mathbf{P}}\big)$

kategoriális inverzió (HUN) – category inversion (ENG)

[21] Az átalakítás terminus használata zavaró lehet, hiszen az átalakítást mint obverzió az átalakítási műveletek egyik típusaként értelmezve részleges, de mindenképpen inkonzisztens homonímához jutunk. Ha a kontextusból egyértelműsíthető, hogy az átalakítás terminusát az obverzió szinonimájaként kell érteni, akkor persze használni lehet ezt a terminust is.

A kategoriális inverzió az a művelet, amelynek az értelmezési tartománya a kategoriális kijelentések halmaza. A kategoriális inverzió kicseréli a kijelentés két kategóriáját azok komplementjeire, de megtartja a sorrendjüket. Az inverziónak nincs direkt következtetési potenciálja.

1.7 Ontológia

MEO-hivatkozások ontológiai elköteleződés kérdése

OntoClean-módszertan, fogalomtulajdonságok, fogalomtípusok

Guarino-hivatkozás entitás-ábrák a rigid és nem-rigid fogalmak különbségére

a szerep nem rigid fogalom a konstituált entitások, terminusok egy része szerepet definiál

1.7.1 JEPD-elv

Amikor elméleti modellt építünk, akkor a generikus alárendeltje reláció segítségével képzünk új fogalmakat úgy, hogy egy már definiált fogalom alá egy tulajdonságnyaláb segítségével szűkebb terjedelmű, viszont bővebb jelentésű alosztály-fogalmakat rendelünk. Az alosztály (altípus, faj) fogalmak képzésekor sokdszor érdemes törekednünk arra, hogy igazodjunk a JEPD-elv elvárásaihoz [90]. Ez az elv azt rögzíti, hogy egy fogalom alá rendelt fogalmak **együttesen kimerítőek** legyenek és **kölcsönösen kizárják egymást** (JEPD – Jointly Exhaustive and Pairwise Disjoint).[22] A generikus struktúra konzisztenciájához nélkülözhetetlen vagy legalábbis fontos, hogy az azonos fogalmi szinten egymás mellett levő, egy közös fölérendelt kategóriához tartozó fogalmak tartalma kölcsönösen kizáró és együttesen kimerítő legyen. Ezáltal elégíthetjük ki azt az elvárást, hogy egy adott fogalom alá kidolgozott fogalmak együttesen a fogalom terjedelmébe tartozó minden előfordulást egyértelműen lefedjenek.

A matematika világából a JEPD-elvnek megfeleltethető egy osztály alatt létrehozott alosztályokból történő **partíció** képzése. Ebben az esetben is az az elvárás, hogy az alosztályok egyetlen példányt tekintve se fedjék át egymást (vagyis az alosztályok partíciók legyenek), illetve az alosztályok terjedelmükkel teljes mértékben lefedjék a fölöttes osztályfogalom (szülőosztály) terjedelmét (vagyis a partíció teljes legyen). Ez ugyanazt jelenti, mint amit a halmazelméletben a **partíció** (partition) fogalma a halmaz és részhalmazai közti viszonyban (↬15.o.). Ha egy fogalom alá olyan generikusan alárendelt fogalmakat hozunk létre, amelyek kielégítik a JEPD-elv elvárásait, akkor azt mondhatjuk, hogy a fogalmat particináltuk, **partíciókra** bontottuk fel. A JPED-elv érvényesül például a logikai hatszög azon kategóriái között, amelyek kontrárius viszonyban állnak egymással, vagyis az **(A)**, az **(E)** és az **(Y)** kategóriák között. Nem teljesül viszont a logikai négyszög elemeire.

A logikai hatszög bevezetése ezért nagyon fontos. Ezáltal juthatunk ugyanis olyan kategóriafelosztáshoz, amely partíciónak minősül, vagyis érvényes rá a JEPD-elv. A későbbiek során, amikor csak lehet, alkalmazni fogom a JEPD-elvet. Az elméleti építkezés során mindig fontos annak tisztázása, vajon a fogalmi modell bővítésekor mindent figyelembe

[22] Angolul szokták még a ‘mutually incompatible and jointly exhaustive’ kifejezést is használni.

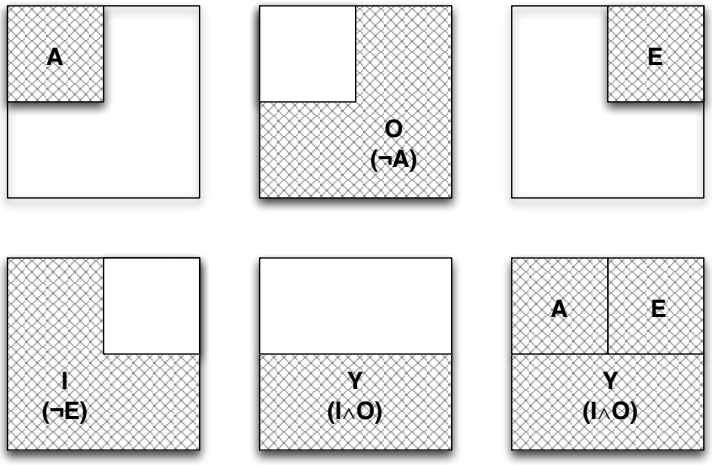


Figure 1.7: kategóriák a JEPD elv szerint

vettünk-e. Ha ilyenkor az új fogalmakra vonatkozóan állítani lehet a JEPD-elv teljesülését, akkor az éppen adott szempont szerint teljesnek mondható az elméleti keret bővítése.

1.7.2 Reifikáció

Az egyértelműséggel és konzisztenciával kapcsolatos elvárások miatt említést kell tennem még egy nagyon fontos jelenségről. A relációk a világ dolgai közti kapcsolatokat fejezik ki. Ez az jelenti, hogy – bináris relációk esetén – egy relációt jelző terminussal mindig két dolog, két relátum közti viszonyt ragadunk meg. A relációfogalmak bevezetése után azonban a fogalom (terminus) használata során előfordulhat, hogy ezt a viszonyfogalmat "tárgyiasítjuk" és valamelyik relátummal megragadott "dologra" vonatkoztatjuk a fogalmat (terminust). Egy hétköznapi példával szemléltetem ezt a jelenséget. Vegyük az 'apja' relációt (terminust), ami egy közvetlen biológiai leszármazási viszonyt ír le két ember között, amikor az "örökítő" pozícióban levő egyed hímnenű. Az 'apja' tehát viszonyt fejez ki. A kommunikáció során azonban nagyon gyakran úgy használjuk ezt a fogalmat, hogy sokkal inkább arra a dologra, arra az emberre gondolunk, aki a kapcsolat egyik felét testesíti meg. Magyarul ekkor az 'apa' terminust használjuk. Amikor így teszünk, akkor reifikáljuk a relációfogalmat, a relációból kiemeljük az egyik relátumot és csak erre hivatkozunk.[23] Természetesen nincs semmi probléma ezzel a jelenséggel, csak arra kell figyelni, hogy mindig tudjuk, éppen a relációfogalmat magát vagy épp annak származtatott változatát, a reifikátumát használjuk. Azt következményt viszont fontos figyelembe venni, hogy amikor reifikálunk egy relációt, akkor információt vesztünk, hiszen a reifikátum már nem mond semmit a reláció másik relátumáról. A reláció másik relátumát általában a reláció inverze segítségével tudjuk könnyen reifikálni.
A bemutatott példák alapján már meghatározhatjuk a reifikáció fogalmát. Mivel egy – bináris – relációt két irányból is tárgyiasíthatunk, ezért a reifikáció műveletét kétféleképpen is definiálhatjuk, így beszélhetünk bal- és jobboldali reifikációról.

(GL.16) $\text{LEFT-REIFICATION} \equiv \text{REIFICATION} \equiv \text{`}\mathbf{R}(m,n) \equiv \text{`}\mathbf{R} \equiv \mathbf{R}' \stackrel{\text{def}}{=} \forall m \forall n (\text{`}\mathbf{R}(m,n) \leftrightarrow \mathbf{R}(m))$

[23] Ez a művelet olyan, mint a logikai műveletek között a projekció (↬21.o.).

állítás	reláció	reifikátum	inverz reláció	inverz reifikátum
Laci apja Katinak.	$apja(Laci, Kati)$	$apa(Laci)$	$gyereke(Kati, Laci)$	$gyerek(Kati)$
Laci apja Lilinek.	$apja(Laci, Lili)$	$apa(Laci)$	$gyereke(Lili, Laci)$	$gyerek(Lili)$
Zoli apja Petinek.	$apja(Zoli, Peti)$	$apa(Zoli)$	$gyereke(Peti, Zoli)$	$gyerek(Zoli)$

bal-reifikáció (HUN) – reifikáció (HUN) – reification (ENG)

(GL.17) RIGHT-REIFICATION ≡ REIFICATION ≡ $\mathbf{R}'(m,n) \equiv \mathbf{R}' \stackrel{\text{def}}{=} \forall m \forall n(\mathbf{R}'(m,n) \leftrightarrow \mathbf{R}(n))$

jobb-reifikáció (HUN) – reifikáció (HUN) – reification (ENG)

A reifikáció művelete – a dolgokat absztraktan szemlélve – hasonlít a propozicionális logika projekció műveleteihez, hiszen itt is, ott is arról van szó, hogy két valamilyen módon összekapcsolt elem közül a művelet végrehajtása során (eredményeként) csak az egyiket vesszük figyelembe. A reifikáció műveletét még tovább absztrahálhatjuk, és tetszőlegse absztrakt individuumokat létrehozhatunk, ami a tudásreprezentáció számára nagyon praktikus előnyökkel szolgálhat [113, 41].

Donald Davidson az események logikai reprezentálhatóságát keresve tett egy javaslatot arra, hogy miként lehetne az eseményeket megfelelő módon logikai formulákkal leírni [189].

Jerry Hobbs reifikációs leíró nyelve ‘Ontológiai promiszkuitás’ címen írott cikkében továbbvitte Davidson javalatát, amivel nagyobb rugalmasságot, jobb kezelhetőséget biztosított a jelenségek logikai leírása számára [359].

mi az, amikor a logikai négyszög elemei közötti kapcsolatot jellemzi (kontadiktorius, kontrárius stb.)

Ez az értelmezés különbözik a mind Berger-Luckman-féle reifikációfelfogástól, mind a Marx-féle tárgyiasulás/tárgyiasítás fogalomtól.

1.8 Másodrendű logika

Abban a pillanatban, hogy a logikai nyelvünkbe felvesszük a változókat és kvantorokat, lehetővé válik, hogy egy új szempont szerint különíthessünk el különböző logikákat, logikai szinteket. Nem véletlenül nevezik a kijelentéslogikát nulladrendű extenzionális logikának, a predikátumlogikát pedig elsőrendű logikának [627]. Mivel a kvantálás, a kvantorok alkalmazásának lehetősége a predikátumlogikában jelenik meg először, ezért ez a logikát (és ezt a logikai szintet) nevezzük elsőrendűnek. Elsőrendű logikákban a kvantorokat az individuumok valamely csoportját jelölő változókra lehet alkalmazni, semmi másra. Ha egyáltalán nem használhatunk kvantorokat (mert az adott logikában ez nincs is implementálva), akkor nulladrendű nyelvről beszélhetünk. Az ítéletlogikában nincsenek kvantorok, ezért az a logika nulladrendű.

A másodrendű logikák felé úgy lehet elmozdulni, hogy kiterjesztjük a változók, illetve a kvantorok hatókörét. Ezt úgy tehetünk meg, hogy egyfelől megengedjük, hogy változókat ne csak individuumokra, hanem predikátumokra (relációkra, függvényekre) vonatkozóan is lehessen képezni, másfelől megengedjük azt is, hogy a kvantorokat az individuumváltozókon túl a predikátumokra irányuló változókra is lehessen alkalmazni

[250, 238], [627, 197]. A formulák szintjén ez úgy ismerhető fel, hogy az individuumváltozók mellett a predikátumváltozókra is érvényesítjük a formulában használt kvantorokat, valahogy így:

$$\forall m \forall P \forall Q(P(x) \rightarrow Q(x))$$

Ez a formula a **generikus alárendeltje** (GENERIC SUBORDINATE) relációt definiálja, amely a predikátumok, fogalmak közti alá-fölérendeltségi viszonyt (vagy másként a faja-neme, ISA relációt) írja le.

Ha a másodrendű logikában megengedett, hogy a predikátumokra, kategóriákra is alkalmazzuk a változókat és kvantorokat, akkor ez azt is jelenti egyben, hogy a fogalmi szinten is képezhetünk másodrendű fogalmakat. Ezek formálisan úgy jeleníthetőek meg, hogy az adott predikátum (fogalom) változóhelyeire, a predikátum paraméterei közé olyan predikátumot (fogalmat) írhatunk be, amely szerepel az elsőrendű predikátumok (fogalmak) között. Ez a beágyazás egyértelműen jelzi a predikátum (fogalom) másodrendű minőségét. Ha vesszük a 'tudja, hogy' predikátumot, aminek két változóhelyet tartunk fent, amikor jelölni akarjuk, hogy 'ki az, aki tud valamit', illetve 'mit tud az a valaki'. Ezt a predikátumot így definiálhatjuk:

$\mathsf{TUD}(x, p)$ – *x tudja, hogy p*

Előfordulhat azonban az is, hogy a p paraméterre újra a 'tudja hogy' predikátumot alkalmazzuk, amikor például azt mondjuk, hogy:

$p = \mathsf{TUD}(y, q)$ – *y tudja, hogy q*

Ha a p tudásra vonatkozó második mondatunkat behelyettesítjük az első mondatba, akkor ilyen formulát kapunk:

$\mathsf{TUD}(x, p) = \mathsf{TUD}(x, \mathsf{TUD}(y, q))$ – *x tudja, hogy y tudja, hogy q*

Ezen a ponton szépen látszik a predikátum önmagába való beágyazottsága. Amikor ilyet tapasztalunk, akkor mindig másodrendű fogalomról van szó, aminek leírásához már másodrendű logikára van szükség. Ilyen esetekben ugyanis ha általános kijelentéseket akarunk tenni (az univerzális kvantor használatával), akkor meg kell engednünk, hogy a változókat és kvantorokat predikátumra is lehessen alkalmazni. A hétköznapi kommunikációban és a társadalomtudományi diskurzusban egyaránt rengeteg másodrendű fogalmat használunk, igaz, reflektálatlanul. A későbbiekben mindig jelezni fogom, amikor másodrendű fogalmakhoz jutunk valamely téma tárgyalása során. Teszem ezt egyfelől a fogalmi egzaktság érdekében, de másfelől teszem ezt azért is, mert a másodrendű fogalmak használata könnyen paradoxonhoz vezet.[24]

Ha már beszélünk másodrendű logikáról, akkor természetesen beszélhetünk harmad-, negyed-, elvileg tetszőleges szintű, magasabbrendű logikáról is. Ezeket a másodrendű logikához hasonló módon lehet képezni, aminek eredményeként egyre mélyebben egymásba ágyazott fogalmakhoz jutunk. A cselekvéselméleten belül leggyakrabban a másodrendű fogalmak bukkanak fel, ritkán kell használni harmadrendű fogalmakat, és ennél magasabb szintre "nem kell felmennünk."

Stanford Philosophy [233]

Foundation without Foundationalism. A Case for Second-order Logic [677]

[24] Ezen a ponton meg kell említenem a predikativitás fogalmát, illetve az ehhez kapcsolódó elméletet, ami az önmagába ágyazott fogalmak viselkedésével, paradoxonaival foglalkozik.

1.9 Dinamikus logika

A **dinamikus logikát** (dynamic logic) eredetileg a számítógépes programok leírására hozták létre, hogy legyen olyan eszköz, amelynek segítségével meg lehet vitatni a programok helyességét, de a nyelv idővel egészen másfajta alkalmazási területeken is hasznosnak bizonyult. A dinamikus logika a modális logika egyik típusa. Központi kategóriája a **program(darab)** fogalma, ami értelmezhető úgy is, hogy valamely cselekvést (action) kell végrehajtania. A program(darab)ot a-val jelöljük. Két deontikus operátort használunk a **szükségszerű** és **lehetséges** operátorok helyett:

$$\begin{array}{lll} \Box p & \text{helyett:} & [a]p \\ \Diamond p & \text{helyett:} & \langle a\rangle p \end{array}$$

ahol az a program(darab) valamely p állapotra (azt leíró propozícióra) vonatkozik. Az első operátor azt fejezi ki, hogy ha végrehajtjuk az a programot, akkor p szükségszerűen igaz kell legyen, vagyis a szükségszerűen létrehozza p-t, míg a második operátor ennek csak a lehetőségét teremti meg. A programokkal (cselekvésekkel) különböző műveleteket lehet végezni (ezeket értelmezhetjük úgy, hogy 'futtasd le' vagy másként 'tedd meg'). A legfontosabb műveleteket a következőféleképpen jelölhetjük:

(ML.0) $\mathscr{L}_{DL}([\,],\langle\rangle,1,0,\cup,\cap,;,*,?)$, ahol

$[a]p$		ha megteszed a-t, kötelező, hogy p
$\langle a\rangle p$		ha megteszed a t, lehetséges, hogy p
1	SKIP/NOP	ne csinálj semmit, állj le a programmal (terminálj)
0	BLOCK, konstans	ne csinálj semmit
$a\cup b$	egyesítés	tedd meg a-t vagy b-t
$a;b$	komponálás	tedd meg a-t, utána b-t
$a*$	iterálás	ismételten tedd meg (véges sokszor) a-t
$a\cap b$	egyszerre futtatás	tedd meg a-t és b-t
$a?$	tesztelés	ha a igaz, akkor tedd meg, ha hamis, akkor állj le

A műveletek logikai definíciójához szükség van néhány modális logikai axiómára (egyre a modális operátorok közti transzformáció, kettőre a következtetési szabályok miatt):

$[a]p \equiv \neg\langle a\rangle\neg p$	operátorok közti transzformáció
$(\vdash p \wedge \vdash p\to q)\to \vdash q$	modus ponens
$\vdash p \to \vdash [a]p$	szükségességi szabály (necessitation rule)

A fent bemutatott dinamikus logikai műveleteket a következő axiómákkal határozhatjuk meg (ahol a és b programot, p tetszőleges kifejezést, x - változót, e konstansokból és változókból álló kifejezést, $\Phi(x)$?????-t jelent):

(DL.1) $[0]p$

(DL.2) $[1]p \equiv p$

(DL.3) $[a\cup b]p \equiv [a]p \wedge [b]p$

(DL.4) $[a;b]p \equiv [a][b]p$

(DL.5) $[a*]p \equiv p \wedge [a][a*]p$

(DL.6) $p \wedge [a*](p \rightarrow [a]p) \rightarrow [a*]p$

(DL.7) $[x := e]\Phi(x) = \Phi(e)$

(DL.8) $[p?]q \equiv p \rightarrow q$

Értelmezés .....

Az általános axiómák és a műveletekre vonatkozó axiómák alapján levezethetőek azok az egyenlőségek is, amelyek a másik dinamikus logikai operátorra állapítanak meg összefüggéseket:

(DL.9) $\neg\langle 0\rangle p$

(DL.10) $\langle 1\rangle p \equiv p$

(DL.11) $\langle a \cup b\rangle p \equiv \langle a\rangle p \vee \langle b\rangle p$

(DL.12) $\langle a; b\rangle p \equiv \langle a\rangle\langle b\rangle p$

(DL.13) $\langle a*\rangle p \equiv p \vee \langle a\rangle\langle a*\rangle p$

(DL.14) $\langle a*\rangle p \rightarrow p \vee \langle a*\rangle(\neg p \wedge \langle a\ \rangle p)$

(DL.15) $\langle p?\rangle q \equiv p \wedge q$

Nem tartozik a könyv témájához, de ha már itt vagyunk, érdemes megjegyezni, hogy a programozás két fontos konstrukciója, az ‘if-then elágazás’ és a ‘while-ciklus’ is egyszerűen definiálható a dinamikus logika operátoraival.

(DL.16) IF p THEN a ELSE $b \stackrel{\text{def}}{=} (p?; a) \cup (\neg p?; b)$

(DL.17) WHILE p DO $a \stackrel{\text{def}}{=} (p?; a)*; \neg p?$

Logikából nagyon sokféle van. A mérnökök, informatikusok ezerfélét kidolgoztak a legkülönfélébb célokra, a legkülönbfélébb szintaxissal, a legkülönfélébb képességekkel. A dinamikus logika csak egy a sok közül, ami csak azért kell számunkra, mert a társadalmi cselekvés szempontjából fontos fogalmak közül többet is dinamikus logikai alapon formalizáltak.

1.10 Modális logika

A társadalmi cselekvés modelljének felépítése során több alkalommal is szükségünk lesz olyan elméletekre, amelyek a deontikus logika területéhez sorolhatók (már a cselekvés fogalmának értelmezéséhez is szükségünk van erre a diszciplinára). A deontikus logikát a modális logika egyik ágaként kezeljük. A modális logika nem elsősorban a kijelentések igazságának kérdéseivel, hanem inkább a kijelentések igazságának modalitásával (minőségével, “erősségével”, megalapozottságával) foglalkozik. Szükségszerű-e egy kijelentés igazsága; mindig fent kell-e tartanunk a kijelentés igazságát vagy azt csak bizonyos feltételek, adott világ, világállapot létezéséhez tudjuk kötni; van-e olyan kijelentés, ami sosem lehet igaz, tehát lehetetlen; melyek azok a kijelentések, amelyek lehetségesen igazak, de nem szükségszerűek; ha valami szűkségszerűen igaz, akkor igaz is-e egyben – ilyen és ezekhez hasonló kérdésekre próbál meg válaszokat találni a modális

logika.[25] Mivel sokféleképpen módosíthatjuk az állításainkat, sokféle modalitást kezelhetünk. Egy *A* kijelentés kapcsán mondhatjuk azt, hogy ‘ tudom, hogy *A*’, ‘azt hiszem, hogy *A*’, ‘gyakran *A*’, ‘biztos, hogy *A*’, ‘tilos *A*’, ‘szükségszerű, hogy *A*’ stb. Ilyen esetekben az állításunkat (*A*-t) valamilyen modalitásértékkel “kiegészítjük”, módosítjuk, amit úgy is kifejezhetünk, hogy az állításhoz egy modális operátort kapcsolunk hozzá. Amikor az állításaink szükségszerűségét vizsgáljuk, akkor alethikus modalitásról van szó. A normativitás, a konstruált kötelezőség világát a deontikus, a kijelentések időbeliségét a temporális, a hitek logikáját a doxasztikus, a tudásokét az episztemikus modalitással próbáljuk megragadni.

A modális logika nyelvébe fel kell vennünk a modális operátorokat és egy olyan relációt, amely a lehetséges világok közti elérhetőségi kapcsolatot képezi le. A különböző modális logikákat azon az alapon különíthetjük meg egymástól, hogy feltételezzük, sokféle világ létezhet, melyek közül némelyikben igazak bizonyos állítások, másokban viszont hamisak, megint más állítások minden lehetséges világban fennállnak, míg vannak olyanok is, amelyek sehol sem igazak. Az elérhetőségi reláció azt a tényt fejezi ki, hogy adott világból szemlélve mely más világok érhetők el (aminek inverze írja le azt, hogy melyek nem). Ennek az elérhetőségi relációnak fontos szerepe van a modális logika tipizálásában.

A lehetséges világok közti elérhetőségi relációra vonatkozóan különböző relációtulajdonságokat lehet megállapítani, és ezek a tulajdonságok egyben jellemzik a különböző modális logikákat is. Egy-egy ilyen relációtulajdonsággal egy-egy kalkulust definiálhatunk, amely az adott modális logika ‘szükségszerű’ (‘erős modálisan igaz’), illetve ‘lehetséges’ (‘gyenge modálisan igaz’) fogalmára vonatkozik. A reflexivitás tulajdonságával például az ún. (T) kalkulust határozhatjuk meg, amelyre a következő formula írható fel:

$$\Box A \rightarrow A \quad \text{(T)}$$

A (T) formulát úgy interpretálhatjuk, hogy ‘ami szükségszerűen igaz, az igaz’. Első “nekifutásra” ez kijelentés magától értetődőnek tűnhet, pedig könnyen lehet olyan értelmezést találni, amikor ez az állítás nem igaz. Jobban meg lehet érteni mindezt, ha az előző mondatot átfogalmazzuk a következő alakra: ‘ez így van, mert így kell lennie’. Ha a ‘kell’ fogalmát egy természettörvényre hivatkozva értelmezzük, akkor a “kellés” (szükségszerűség) valóban maga után vonja az állítás igazságát is: ha feldobunk egy követ, akkor annak – a tömegvonzás törvénye miatt – le kell esnie, és, valóban, a földobott kő le is esik a földre. Ha azonban a ‘kellés’ fogalma alatt a kötelezőséget értjük, akkor közel sem annyira magától értetődő, hogy ami kell (kötelező), annak meg is kell valósulnia, hiszen rengeteg esetet ismerünk, amikor valamilyen kötelező parancsba foglalt állítás a valóságban nem teljesült.

Egy közlekedési példával szemléltetve ezt a különbséget tegyük fel, hogy éppen azt meséljük egy barátunknak, hogy egy közös ismerősünket az úttesten elütötte egy jobbról érkező autó, amire a barátunk ennyit mond:

> Jobbról nem is jöhetett autó azon a helyen.

[25] A modális logika két gyakran hivatkozott forrása Brian Farrell Chellas [145] és G.E. Hughes és M.J. Cresswell műve [367], én Patrick Blackburn, Maarten de Rijke és Yde Venema könyvére [92], valamint Robert A. Bull és Krister Segerberg tanulmányára [131] támaszkodom leginkább.

Ez a mondat kétféleképpen is értelmezhető. Nem jöhetett jobbról autó, mert *(i)* a gázolás egy éles kanyarban történt, és a baleset színhelyétől jobbra házfalak meredeznek, vagy *(ii)* egy egyenes, ám – balról jobbra – egyirányú utcában jobbról nem érkezhetett semmi a kötelező haladási iránnyal szemben. Kicsit másként: nem jöhetett jobbról az autó, mert az fizikailag lehetetlen lett volna, illetve nem jöhetett jobbról az autó, mert a tiltó jogszabály miatt nem lett volna szabad. A fenti megnyilatkozást átfogalmazhatjuk a következő formába:

Nem lehetséges, hogy az adott helyen jobbról jöjjön az autó.

Ebben az esetben ugyanúgy fennáll a kettős jelentés lehetősége, csak ekkor a 'nem lehetséges' kifejezésnek adhatunk kétféle értelmezést: *(i)* 'nem lehetséges, mert fizikailag lehetetlen', illetve *(ii)* 'nem lehetséges, mert nem megengedett'. A 'lehetséges' kétféle interpretációja teljesen eltérő magyarázósémákat és különböző modális logikákat eredményez. Mielőtt bemutatnám, hogy miként is határozhatjuk meg a legfontosabb modális kalkulusokat (modális logikákat), át kell tekintenem a modális logika nyelvnek elemeit.

A modális logikai nyelv elemei

Szükség van tehát egy modális logikai nyelvre, ami a következőből áll [92].

(ML.0) $\mathscr{L}_{ML}(p,q,r,A,B,C,\bot,\neg,\vee,\Diamond)$, ahol

$p,q,r \in \Phi$	propozíció
A,B,C	formula
$\bot$	sosem igaz (ellentmondás)
$\neg$	tagadás logikai művelet
$\vee$	diszjunkció logikai művelet
$\Diamond$	lehetséges modális logikai operátor

A modális nyelv jól-formált formulái a következő induktív felsorolással adhatók meg (ahol az egyes opciókat a 'I' jel választja el).

(ML.1) $A \stackrel{\text{def}}{=} p \mid \bot \mid \neg B \mid B \vee C \mid \Diamond B$

jól-formált formula (HUN) – well-formed formula (ENG)

A felsorolást úgy kell értelmezni, hogy *(i)* jól formált formula bármely propozíció vagy az *(ii)* ellentmondás; *(iii)* ha van egy jól-formált formulánk, akkor jól-formált marad, ha tagadjuk; *(iv)* ha van két jól-formált formulánk, akkor ezek diszjunkciója is jól-formált; *(v)* ha egy jól-formált formulát összekapcsoljuk a lehetséges operátorral, akkor jól-formált formulához jutunk. Ezeket formulákat természetesen tetszőleges formában és mértékben egymásba ágyazhatjuk, így a többi ismert logikai műveletet is bevezethetjük a nyelvbe.

A modális logika nyelvében két modális operátort használunk, hogy kifejezhessük velük a szükségszerűség és lehetségesség fogalmait. A két új operátor között természetesen szigorú összefüggés létezik, egyik a másikból levezethető. A különböző iskolák, szerzők különböznek abban, hogy melyik kategóriát veszik fel definiálatlan alapfogalomként (és ebből következően melyiket vezetik be a másik segítségével). Én a 'lehetséges' kategóriáját vettem fel alapként, a szükségszerű fogalmát a fent ismertetett módszer segítségével már definiálni lehet.

(ML.2) $\textsc{possible}(A) \equiv \Diamond(A) \equiv \Diamond A$

lehetséges (HUN) – gyenge modálisan igaz (HUN) – gyémánt (HUN) – possible (ENG) – diamond (ENG)

(ML.3) $\text{NECESSARY}(A) \equiv \Box(A) \equiv \Box A$

szükségszerű (HUN) – erős modálisan igaz (HUN) – necessary (ENG) – box (ENG)

(ML.4) $(\text{DUAL}) \stackrel{\text{def}}{=} \Box A \leftrightarrow \neg \Diamond \neg A$

(Dual)-séma (HUN) – (Dual)-schema (ENG)

A modális logika következő fontos fogalma a relációs struktúra (↬40.o.), amit azonban itt másként, **keretnek**, neveznek. A keret segítségével léphetünk tovább következő fontos modális kategóriára, a **modell** fogalmára. Ez felfogható egy másodrendű relációstruktúrának, amiben egy keret és egy reláció van definiálva. A modellhez szükséges reláció egy $\mathbf{V}$ értékelési (valuation) függvény, amely a $A \in \Phi$ propozíciókat rendeli hozzá a $\mathbb{W}$ halmaz elemeihez. Ezt a három új fogalmat fel kell vennünk az elméletünkbe.

(ML.5) $\text{FRAME} \equiv \mathfrak{F} \stackrel{\text{def}}{=} (\mathbb{W}, \mathbf{R})$

keret (HUN) – frame (ENG)

(ML.6) $\text{VALUATION} \stackrel{\text{def}}{=} \mathbf{V} : \Phi \mapsto \mathbf{V}(A) \subseteq \mathbb{W}, p \in \Phi$

értékelés(i függvény) (HUN) – valuation (function) ((ENG))

(ML.7) $\text{MODEL} \equiv \mathfrak{M} \stackrel{\text{def}}{=} (\mathfrak{F}, \mathbf{V})$

modell (HUN) – model ((ENG))

A modell kategóriája azért (is) fontos, mert vele lehet az **igazság** vagy kielégíthetőség fogalmait értelmezni, a keret pedig az **érvényesség** meghatározásához szükséges. Nézzük meg előbb, hogy lehet az igazság fogalmát meghatározni. Feltételezzük, hogy ϕ, ψ tetszőleges formulák, $A \in \Phi$ adott propozíció, $w, u \in \mathbb{W}$ adott állapotok ($\mathbb{W}$ elemei). Ekkor az ϕ, ψ formulák igazságát induktív felsorolással lehet meghatározni az alábbi módon.[26]

(ML.8) $\text{SATISFIED} \equiv \vDash \stackrel{\text{def}}{=}$

1) $\mathfrak{M}, w \vDash A \stackrel{\text{def}}{\Leftrightarrow} w \in \mathbf{V}(A), p \in \Phi$
2) $\mathfrak{M}, w \stackrel{\text{def}}{\neg\vDash} \bot$
3) $\mathfrak{M}, w \vDash \neg\phi \stackrel{\text{def}}{\Leftrightarrow} \mathfrak{M}, w \stackrel{\text{def}}{\neg\vDash} \phi$
4) $\mathfrak{M}, w \vDash \phi \vee \psi \stackrel{\text{def}}{\Leftrightarrow} \mathfrak{M}, w \vDash \phi \stackrel{\text{def}}{\vee} \mathfrak{M}, w \vDash \psi$
5) $\mathfrak{M}, w \vDash \Diamond\phi \stackrel{\text{def}}{\Leftrightarrow} \stackrel{\text{def}}{\exists} v \in \mathbb{W}(\mathbf{R}(w, v)) \stackrel{\text{def}}{\wedge} \mathfrak{M}, v \vDash \phi$

kielégítő (HUN) – satisfied (ENG) – igaz (HUN) – true (ENG) – szemantikai következmény (HUN) – semantic consequence (ENG)

A feltételeket meglepően egyszerű értelmezni. Mind az öt feltételt az adott $\mathfrak{M}$ modellre és az adott w állapotra kell értelmezni. Az első feltétel azt mondja ki, hogy az A propozícióra vonatkozóan a modell akkor (és csak akkor) elégíti ki az igazságfeltételt az adott w állapotban, ha az állapot benne van a $\mathbf{V}$ értékelési függvény értékkészletében. A második feltétel szerint az ellentmondás nem igaz az adott modell szerint az adott w állapotban. A harmadik formula azt mondja ki, hogy ha egy ϕ formula igaz (kielégíti az adott modellt, az adott w állapotban), akkor (és csak akkor) a tagadása nem igaz. A negyedik feltétel két

[26] A definiáláshoz használni kell olyan logikai műveleteket, amelyeket nem lehet a tárgynyelvben alkalmazott szimbólumokkl jelölni, még akkor sem, ha az értelmük ugyanaz. Az ilyen metanyelvi műveleteket a műveletek fölé írt 'def' jelzéssel látom el.

formula diszjunkciójára mondja ki, hogy annak igazsága a komponensformulák vagylagos igazságától függ az adott w állapotban. Végül az ötödik formula rögzíti azt, hogy az adott modell szerint egy ϕ formula akkor LEHETSÉGES a w állapotban, ha van egy másik v állapot, amely $\mathbf{R}$-relációban van a w állapottal, és amelyben a formula igaz. Ezekkel a kielégíthetőségi feltételekkel már tetszőleges formuláról eldönthetjük, hogy igaz-e az adott modell szerint a w állapotban. Azért pedig, hogy kifejezhessük azt is, ha valami nem kielégíthető, vagyis hamis, érdemes bevezetni a szemantikai következmény reláció tagadását.

(ML.9) $\text{FALSIFIED} \equiv \nvDash \stackrel{\text{def}}{=} \neg \vDash$

nem kielégítő (HUN) – hamis (HUN) – false (ENG) – falsified (ENG) – refuted (ENG)

Ha ismerjük már a lehetséges operátor használatát, akkor nézzük meg, hogyan lehet a szükségszerű operátort alkalmazni. Az eddigiekből következik, hogy a definícióban a korábbi egzisztenciális kvantifikációt ki kell cserélni univerzálisra.

(ML.10) $\text{LOCALLY_NECESSARY} \stackrel{\text{def}}{=} \mathfrak{M}, w \vDash \Box A \stackrel{\text{def}}{\Leftrightarrow} \stackrel{\text{def}}{\forall} v \in \mathbb{W}(\mathbf{R}(w,v)) \stackrel{\text{def}}{\wedge} \mathfrak{M}, v \vDash A$

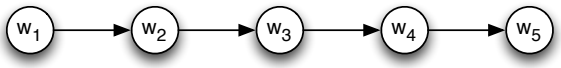
lokálisan szükségszerű (HUN) – locally necessary (ENG)

Ezek lokálisan érvényes kijelentések. Globálisan vagy univerzálisan igaz formula definíciója.

(ML.11) $\text{UNIVERSALLY_TRUE} \stackrel{\text{def}}{=}$

univerzálisan igaz (HUN) – globálisan igaz (HUN) – universally true (ENG) – globally true (ENG)

Az értelmezést megkönnyítendő bemutatok egy példát [92]. Vegyük a következő öt-elemű $\mathbb{W}$-halmazt, amelynek elemei közti relációt a nyilak jelzik. Az elérhetőségi reláció legyen: $\mathbf{R}(w_i, w_j) \leftrightarrow j = i+1$. Ekkor a keretünk így néz ki: $\mathfrak{F} = (\{w_1, w_2, w_3, w_4, w_5\}, \mathbf{R})$. Az állapotok közti kapcsolatot az alábbi grafikonnal lehet szemléltetni.

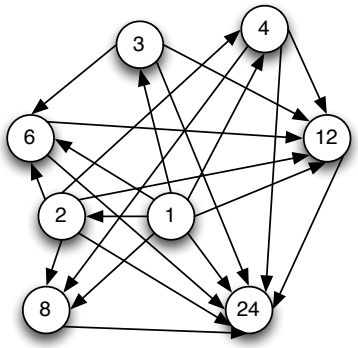


Az elérhetőségi reláció itt azt jelenti, hogy w_1-ből csak w_2, w_2-ből csak w_3, w_3-ból csak w_4 és w_4-ből csak w_5 állapot érhető el, utóbbiból pedig semmi sem. Ezen a kereten vegyük a következő értékelési függvényt: $\mathbf{V}(A) = \{ w_2, w_3\}$, $\mathbf{V}(B) = \{ w_1, w_2, w_3, w_4, w_5\}$, illetve $\mathbf{V}(C) = \varnothing$. Ekkor a $\mathfrak{M} = (\mathfrak{F}, \mathbf{R})$ modellen a következő formulák kielégítőek (igazak).

1) $\mathfrak{M}, w_1 \vDash \Diamond\Box A$
2) $\mathfrak{M}, w_1 \nvDash \Diamond\Box A \rightarrow A$
3) $\mathfrak{M}, w_2 \vDash \Diamond(A \wedge \neg C)$
4) $\mathfrak{M}, w_1 \vDash B \wedge \Diamond(B \wedge \Diamond(B \wedge \Diamond(B \wedge \Diamond B)))$
5) $\mathfrak{M} \vDash \Box B$

Az első formula megértéséhez elsőként meg kell állapítanunk, hogy a $\Box A$ formula igaz a w_1 és w_2 állapotban, mivel mindkettőre igaz, hogy belőlük elérhető a w_2, illetve w_3 állapot, és ezekben az A propozíció igaz. Mivel w_1-ból és w_2-ből más állapotok nem érhető el, ezért igaz az az állítás, hogy a belőlük elérhető összes állapotban fennáll az A propozíció, ez pedig w_1-r és w_2-re is $\Box A$ igazságát jelenti. Ezek után nézzük meg a $\Diamond\Box A$ formula kielégíthetőségét w_1-ben! Jelöljük a formulánkat így: $\Diamond(D)$. Ez akkor igaz, ha van olyan w_1-ből elérhető állapot, ahol a D igaz. Ilyen pedig van: mivel w_1-ből csak a w_2 állapot érhető el, és ott a $D(=\Box A)$ igaz (ahogy az előbb megállapítottuk). Az első formula tehát igaz. A második sor azt fejezi ki, hogy a $\Diamond\Box A \rightarrow A$ formula nem kielégíthető az adott modellen, a w_1 pontban. Ez a formula azért nem igaz, mert a kondicionális előtagja igaz (ezt az előbb néztük át), viszont a kondicionális utótagja (maga az A propozíció) nem igaz, hiszen a **V** értékelési függvény az A propozíciót csak a w_2 és W_3 állapotokban adja meg igaznak. Emiatt a vizsgált formulánk (kondicionálisunk) nem lehet igaz. A harmadik formula azt fejezi ki, hogy a w_2 állapotból elérhető egy olyan állapot (ez a w_3), amelyben igaz az A propozíció és nem igaz a C. Ez így is van, hiszen w_3-ban a **V**(A) igaz, a **V**(C) nem igaz (mivel utóbbi sehol sem igaz). A negyedi formula értelmezéséhez azt kell figyelembe vennünk, hogy az értékelési függvény szerint a B propozíció minden állapotban igaz. A vizsgált – kielégíthető – formula azt fejezi ki, hogy a w_1 állapotban a B propozíció igaz, és van olyan belőle elérhető állapot, amelyben a B propozíció igaz (ez a w_2 állapot), és van olyan állapot, amelyben a w_2-ből elérhető és ott a B igaz és így tovább addig, amíg az utolsó állapotig el nem érünk. Az ötödik formula tartogat egy "meglepetést". A formula ugyanis azt állítja, hogy a B propozíció univerzálisan igaz, vagyis minden állapotban van egy olyan elérhető másik állapot, ahol a propozíció fennáll. Ez triviális az első négy állapotra, hiszen a belőlük elérhető szomszédos állapotokban igaz a B. Az utolsó w_5 állapotból azonban egyetlen más állapot sem érhető el. Erre viszont azt mondhatjuk, hogy a lokális kieléígíthetőséghez csak azt a feltételt kell kielégíteni, hogy az adott állapotból elérhető minden más állapotban igaznak kell lenni a propozíciónak. Mivel egy állapot sem érhető el a w_5-ből, ezért e feltétel üresen (vacuously) teljesül, és emiatt a vizsgált teljes formula igazsága is fennáll.

Érdemes megnéznünk egy másik példát is [92], hogy még közelebb kerülhessünk az eddigi fogalmainhoz. Legyen a halmazunk $\mathbb{W} = \{1,2,3,4,6,8,12,24\}$ és az $\mathbf{R}(m,n)$ legyen az 'n osztható m-mel' reláció ($m \neq n$). Ekkor az 1.-ből elérhető minden más állapot, tehát $\{2,3,4,6,8,12,24\}$, a 2.-ből elérhető a $\{4,6,8,12,24\}$ állapot, a 3.-ból a $\{6,12,24\}$ állapot, a 4.-ből a $\{8,12,24\}$ állapot, a 6.-ból a $\{12,24\}$ állapot, a 8.-ból a $\{24\}$ állapot, a 12.-ből a $\{24\}$ állapot, és a 24.-ből nem érhető el semmi. Ezt a következő ábrával szemléltethetjük.

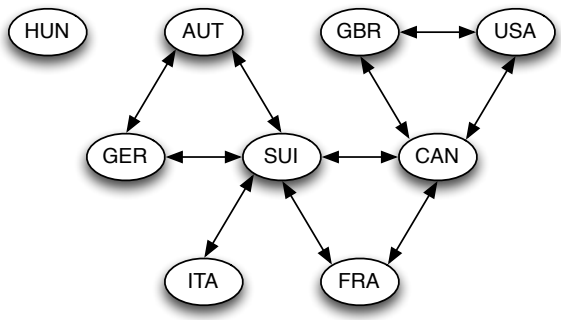


Legyen az értékelési függvényünk a következő: $\mathbf{V}(A) = \{4,8,12,24\}$ és $\mathbf{V}(B) = \{6\}$ (triviális értelmezésként mondhatjuk, hogy A a 4-gyel, B a 6-tal való oszthatóságot jelenti). Ekkor az adott modellünkön az alábbi formulák elégíthetők ki.

1) $\mathfrak{M},4 \vDash \Box A$
2) $\mathfrak{M},6 \vDash \Box A$
3) $\mathfrak{M},2 \nvDash \Box A$
4) $\mathfrak{M},2 \vDash \Diamond(B \wedge \Box A) \wedge \Diamond(\neg B \wedge \Box A)$

Nézzük meg, miért. Az 1) formula azt fejezi ki, hogy a 4. állapotból elérhető a 8., a 12. és a 24. állapot, és ezekben az értékelési függvény szerint az A formula igaz, vagyis minden olyan világállapotban, amely a 4.-ből elérhető, az A formula kielégítő. Ez pedig pont megfelel annak, hogy a $\Box A$ formula kielégíti a modellünket. A 2) sor szerint a 6. állapotból látszik a 12. és 24. állapot, és ezek mindegyikében igaz az A formula, amiből következően a 6. állapotban igaz lesz a $\Box A$ formula is. A 3) formula értelmezése: a 2. állapotból elérhető $\{4,6,8,12,24\}$, amelyek közül a 6. esetében nem teljesül az A propozíció igazsága, ezért a 2. állapotra nézve nem áll fent a $\Box A$ formula igazsága. Az utolsó formula pedig azt fejezi ki, hogy egyszerre igaz két állítás. Egyfelől a 2. állapotból elérhető legalább egy olyan világ, amelyben B is és $\Box A$ is igaz (ilyen a 6. állapot, hiszen ott a B igaz az értékelési függvény alapján, a $\Box A$ formula pedig igaz a fenti 2. formula alapján). Másfelől van legalább egy olyan állapot is, amelyben $\neg B$ és $\Box A$ egyszerre igaz (ilyen a 4. állapot, hiszen ott B nem igaz az értékelési függvényalapjá, viszont $\Box A$ fennáll a fenti 1. formula szerint).

Végül nézzünk át egy olyan példát, amit talán jobban köthetünk a mindennapi életünkhöz, mint az előző kettőt. Vegyünk pár országot: Németország (GER), Ausztria (AUT), Svájc (SUI), Olaszország (ITA), Franciaország (FRA), Magyarország (HUN), Kanada (CAN) és Egyesült Államok (USA). Nézzük meg, hogy ezekben az országokban milyen hivatalos nyelvek vannak: német (N), olasz (O), francia (F), magyar (M), angol (A). Ha az elérhetőségi reláción azt értjük, hogy két ország közt beszélnek-e közös nyelven az emberek, akkor a következő kapcsolatokat kapjuk: $\mathbf{R}(HUN,\varnothing)$, $\mathbf{R}(AUT,\{GER,SUI\})$, $\mathbf{R}(GBR,\{USA,CAN\})$, $\mathbf{R}(USA,\{CAN,GBR\})$, $\mathbf{R}(GER,\{AUT,SUI\})$, $\mathbf{R}(CAN,\{FRA,SUI\})$, $\mathbf{R}(ITA,\{SUI\})$, $\mathbf{R}(SUI,\{GER,AUT,ITA,FRA,CAN\})$, $\mathbf{R}(FRA,\{CAN,SUI\})$. Tételezzük fel azt, hogy az értékelési függvényünkkel azt írjuk le, hogy a különböző országokban (világokban) hivatalos-e az adott nyelv vagy sem. Vagyis: $\mathbf{V}(N) = \{GER,SUI,AUT\}$, $\mathbf{V}(O) = \{SUI,ITA\}$, $\mathbf{V}(F) = \{SUI,FRA,CAN\}$, $\mathbf{V}(M) = \{HUN\}$. Az elérhetőségi reláció az alábbi módon kapcsolja össze a világokat.



Ebben a modellben – sok egyéb mellett – az alábbi formulák kielégíthetőek.

1) $\mathfrak{M}, HUN \vDash \Box M \wedge \Box A \wedge \Box N \wedge \Box O \wedge \Box F$
2) $\mathfrak{M}, SUI \vDash \Box(O \vee F \vee N)$
3) $\mathfrak{M}, USA \vDash \Box A$
4) $\mathfrak{M}, CAN \nvDash \Box A$
5) $\mathfrak{M}, ITA, FRA, GER, GBR, SUI, AUT, CAN, USA \nvDash \Diamond M$

Az 1) formula azt írja le, hogy a minden HUN-világból elérhető más világban szükségszerűen beszélik az összes nyelvet. Ez a logikai igazság azonban üres abban az értelemben. hogy csak azért igaz, mert nincs egyetlen olyan másik világ, amely elérhető lenne ebből az állapotból. Tartalmas jelentése van a 2) formulának, ez ugyanis azt takarja, hogy a Svájból elérhető világokban szükségszerűen beszélnek a német, a francia vagy az olasz nyelv valamelyikén. A 3) sorban az szerepel, hogy az Egyesült Államokból elérhető országokban szükségszerűen (mindenhol) beszélnek angolul. A 4) formula tanulsága szerint azonban ez nem igaz Kanada esetén (mivel van olyan ország, ahol az angol nem hivatalos nyelv, például Svájc). Végül az 5) sor azt a tényt reprezentálja, hogy nincs olyan ország, amelyből elérhető lenne egy olyan világ, ahol a magyart beszélik (azaz a magyar nyelv szigetként működik).

A bemutatott példákból kiderül az is, hogy az **R** elérhetőségi relációnak tetszőleges jelentést tulajdoníthatunk. Fontos kérdés, hogy mit mondhatunk az elérhetőségi relációra vonatkozóan általában, de a válaszhoz még várnunk kell kicsit. Előtte be kell még vezetnünk pár további fogalmat.

———————

állapotok - lehetséges világok
lehetségesen ϕ szükségszerűen ϕ
egységes behelyettesítés
modus ponens
generalizáció törvénye ——————

A keret fogalmának segítségével definiálhatjuk az érvényesség fogalmát, amihez három lépésben juthatunk el. Előbb egy adott keret egy adott állapotában való érvényességet, aztán egy adott keretre vonatkozó érvényességet, végül minden kereten belül való érvényességet mondunk ki.

(ML.12) $\text{VALID_AT_A_STATE}(\phi, w) \equiv \mathfrak{F}, w \vDash \phi \overset{\text{def}}{=} \overset{\text{def}}{\forall} (\mathfrak{F}, \mathbf{V})(\text{TRUE}(\phi))$

érvényes egy keret egy állapotában (HUN) – valid at a state in a frame (ENG)

(ML.13) $\text{VALID_IN_A_FRAME}(\phi) \equiv \mathfrak{F} \vDash \phi \overset{\text{def}}{=} \overset{\text{def}}{\forall} w(\text{VALID_AT_A_STATE}(\phi, w))$

érvényes egy keretben (HUN) – valid in a frame (ENG)

(ML.14) $\text{VALID}(\phi) \equiv \vDash \phi \overset{\text{def}}{=} \overset{\text{def}}{\forall} \mathfrak{F}(\text{VALID_IN_A_FRAME}(\phi))$

érvényes (HUN) – valid (ENG)

ss

kielégíthető kielégíthetetlen
szintaktikai következmény szemantikai következmény
levezethető
modálisan szaturált

biszimuláció
kanonikus normál

(ML.15) $\text{EXISTENTIAL_MODALITY}(\phi) \equiv \overset{\text{m}}{\exists}\phi \overset{\text{def}}{=} \mathfrak{M}, w \vDash \overset{\text{m}}{\exists}\phi \overset{\text{def}}{\Leftrightarrow} \exists w'(\mathfrak{M}, w' \vDash \phi)$

egzisztenciális modalitás (HUN) – existential modality (ENG)

(ML.16) $\text{UNIVERSAL_MODALITY}(\phi) \equiv \overset{\text{m}}{\forall}\phi \overset{\text{def}}{=} \neg \overset{\text{m}}{\exists} \neg\phi$

univerzális modalitás (HUN) – universal modality (ENG)

Modális sémák, kalkulusok

Ahogy korábban már láttuk, a logikai rendszerekhez axiómákat, sémákat definiálhatunk, amelyek alapján a logikai rendszer fogalmai, tételei mind levezethetőek. Ezt megtehetjük a modális logikában is, vagyis modális sémákat definiálhatunk, amelyek alapján modális kalkulusok építhethetünk fel. Egymás után definálhatunk újabb és újabb sémaaxiómákat, amelyek a korábbi axiómákkal együtt egy-egy modális kalkulust alkotnak majd – miközben mind átveszik a kijelentéskalkulus három axiómáját (↬23.o.). Az axiómasémákat pedig jellemezhetjük az elérhetőségi relációra vonatkozóan különböző tulajdonságokkal is. A következőkben sorban bemutatom a (K), a (D), a (T), a (B), az (S4) és az (S5) modális axiómasémák definícióit.

(ML.17) $(\text{K}) \overset{\text{def}}{=} \Box(A \to B) \to (\Box A \to \Box B)$

(K) (LOG) – disztribúciós axióma (HUN) – distribution axiom (ENG)

A (K)-séma jelentése annyi, hogy ha egy $A \to B$ kondicionális szükségszerűen igaz, akkor a kondicionális két tagjának szükségszerűségét is összekapcsolhatjuk a kondicionálissal. Ez tétel a modális operátorra vonatkozó disztributivitás (↬42.o.) tulajdonságát is kimondja egyben (a kondicionális műveletre nézve).

(ML.18) $(\text{D}) \overset{\text{def}}{=} \Box A \to \Diamond A$

(D) (LOG) – szeriális axióma (HUN) – seriality axiom (ENG)

A (D)-séma azt jelenti, hogy ha valami szükségszerű(en igaz), akkor az lehetséges(en igaz). Első pillantásra egy kicsit intuícióellenesnek bizonyulhat sokak számára ez a tétel, hiszen ‘miért ne lenne igaz, ami szüksészerű’. Pedig ez a séma biztosítja a deontikus séma és a deontikus logika speciális minőségét. A (D)-sémában az a lényeg, ami csak implicit módon van benne: azokban a világokban, amelyek erre sémára támaszkodnak, ha valami szükségszerű (kötelező), az nem feltétlenül igaz egyben. Vagyis, ami szükségszerű (kötelező), az nem szükségszeűen teljesül, csak a lehetősége van meg ennek. Ez a séma összakapcsolható az elérhetőségi relációra vonatkozó szerialitás tulajdonságával (↬34.o.).

(ML.19) $(\text{T}) \overset{\text{def}}{=} \Box A \to A$

(T) (LOG) – (M) (LOG) – reflexivitási axióma (HUN) – reflexivity axiom (ENG)

Ami a (D)-sémában nem teljesül, azt mondja ki a (K)-axiómaséma: ha valami szükségszerű(en igaz), akkor az igaz. Azokra a világokra tehát, amelyeket ez a séma jellemez, teljesül az a – mint láthattuk, nem teljesen megalapozott – intuíciónk, hogy valami akkor

lehet szükszerűen igaz, ha igaz is egyben. Ezt a tulajdonságot az elérhetőségi reláció reflexivitásával hozhatjuk párhuzamba (↬32.o.). Ez annyit tesz, hogy a (K)-szükségszerűség fennállása esetén nemcsak az aktuális világból elérhető más világokra vonatkozó elvárásokat az aktuális világra is ki kell terjeszteni.

(ML.20) $(\mathrm{B}) \stackrel{\text{def}}{=} A \to \Box \Diamond A$

(B) (LOG) – szimmetricitási axióma (HUN) – symmetricity axiom (ENG)

A (B)-séma jelentése az, hogy ha valami igaz, akkor szükségszerű(en igaz), hogy lehetséges(en igaz). Ebben az esetben az elérhetőségi reláció szimmetrikus (↬33.o.). Első értelmezésre elfogadhatónak látszik a tétel, hiszen ha egyszer valami már fennáll, akkor ennek a lehetőségét szükségesnek érezzük, de a helyzet mégsem ilyen egyszerű. Tekintsük ugyanis a következő levezetést:

1) $A \to \Box \Diamond A$
2) $\neg \Box \Diamond A \to \neg A$
3) $\neg(\neg \Diamond \neg) \Diamond A \to \neg A$
4) $\Diamond \neg \Diamond A \to \neg A$
5) $\Diamond \neg \Diamond (\neg \neg A) \to (\neg A)$
6) $\Diamond \Box (\neg A) \to (\neg A)$
7) $\Diamond \Box B \to B$

A levezetés tanulsága az, hogy a kondicionális kontrapozíció, a kettős tagadás, a zárójelezési, a behelyttesítési és a (Dual) szabályok segítségével a (B)-sémából levezethető a hetedik sorban látható tétel, amit már közel sem érzünk triviálisnak, sőt inkább intuícióellenesnek tartunk. A tétel ugyanis azt állítja, hogy ha valami lehet, hogy szükségszerű, akkor az igaz. Ennek az igazsága egyáltalán nem nyilvánvaló. A (B) séma elfogadása ezért mégsem annyira magától értetődő, mint az az első látásra feltételezhető lenne. Sokkal inkább elfogadható a következő axiómaséma, bár bizonyos modalitásértelmezések esetén ennek fennállását is meg lehet kérdőjelezni.

(ML.21) $(\mathrm{S4}) \stackrel{\text{def}}{=} \Box A \to \Box \Box A$

(S4) (LOG) – (4) (LOG) – tranzititási axióma (HUN) – transitivity axiom (ENG)

A séma azt rögzíti, hogy ha valami szükségszerű(en igaz), akkor szükségszerű(en igaz), hogy szükségszerű(en igaz). Ez az elérhetőségi reláció tranzitivitását is jelenti egyben (↬35.o.), ami abban nyilvánul meg, hogy ha valami szükségszerű egy aktuális világban, akkor nemcsak belőle elérhető szomszédos világokban, de az azokból elérhető "másodszomszédos" világokban is szükségszerű lesz. Az erre a sémára épülő S4-kalkulusban interpretáljuk úgy – Gödel nyomán – a $\Box A$ formulát, hogy annak az '*A* bizonyított állítás' jelentést tulajdonítjuk. Ezzel az értelmezéssel az S4 kalkulus a bizonyítottság[27] (vagy bizonyíthatóság) logikájának tekinthető [627, 263]. Már csak ezért is, de persze más okok miatt is, az (S4)-sémára szükség lesz a cselekvéselmélet tárgyalása során.

[27] Ruzsa Imre megjegyzi, hogy csak akkor tartható ez az interpretáció, ha a bizonyítottságon nem formalizált elméleten belüli bizonyíthatóságot, levezethetőséget értünk. Ebben az értelemben az S4-es kalkulus inkább a tapasztalati (kísérleti) tudáson alapuló bizonyítottság fogalmához köthető.

(ML.22) $(\mathrm{S5}) \stackrel{\mathrm{def}}{=} \Diamond A \rightarrow \Box \Diamond A$

(S5) (LOG) – (E) (LOG) – euklideszi axióma (HUN) – euclidean axiom (ENG)

A legerősebb S5-kalkulus alapsémája az (S5), amelynek jelentése: ha valami lehetséges(en igaz), akkor szükségszerű(en igaz), hogy lehetséges(en igaz). Ez az értelmezés a sémát az elérhetőségi reláció euklideszi tulajdonságával kapcsolja össze (↬35.o.). Ebben a kalkulusban adhatjuk azt az értelemzést a $\Box A$ formulának, hogy 'analitikusan igaz, hogy A'. Ha így teszünk, akkor a kalkulussal az analicitás, vagyis az analitikus állítások logikáját kapjuk [627, 265]. Vigyázni kell azonban ezzel az értelmezéssel, mert nem jelent minden esetben egyértelmű segítséget akkor, amikor el akarjuk dönteni egy állításról, hogy analitikus-e vagy sem.[28]

Fontosnak tartom újra hangsúlyozni, hogy a fenti definíciók csak egy-egy axiómasémát határoznak meg, és csak úgy jutunk kalkulusokhoz a segítségükkel, ha az "alattuk levő" többi axiómát is beemeljük az adott kalkulusba (vagy ha nem is mindet, de legalábbis a legtöbbet). És az axiómák mellé szükség van még levezetési szabályokra is. Egyfelől minden modális kalkulus átveszi a modus ponens szabályt (↬25.o.), illetve a behelyettesíthetőségre is kimondanak egy szabályt, de ezzel itt nem foglalkozom. A modális kalkulusokhoz azonban kell még egy további levezetési szabály is, amit többféleképpen is neveznek, de persze a definíció minden esetben ugyanaz.

(ML.23) $(\mathrm{NEC}) \stackrel{\mathrm{def}}{=} \vdash A \Rightarrow \vdash \Box A$

szükségességi szabály (HUN) – modális generalizálás szabálya (HUN) – Necessitation Rule (ENG) – Modal Generalization rule (ENG) – (Nec) – (LOG) – (Gen) – (LOG)

Ez a levezetési szabály tehát olyan, mint a modus ponens, ami szükséges az axiómákból és tételekből való következtetések levonásához. Az értelmezése egyszerű: ha egy A formula(osztály) eleme egy logikai nyelvnek, akkor a $\Box A$ formula is eleme lesz.

A fent bemutatott sémát (és kalkulusok) közül a továbbiakban a (D) séma és a hozzá tartozó modális logika a deontikus logikában (↬128.o.) és a doxasztikus logikában (↬79.o.) lesz különösen fontos, a (K), (T), (B), (S4) és (S5) sémákat pedig az episztemikus (↬75.o.) és a preferencialogikában (↬119.o.) fogjuk majd alkalmazni.

A modális logikai fejezet lezárásaként ki kell mégtérnem a modális logika és a logikai négyszög kapcsolatára. Utóbbit ugyanis alkalmazhatjuk a modális logikai operátorokra is. Sőt, a modális operátorokra nemcsak a négyszögbe, de a hatszögbe rendezés is értelmes. Ha a legáltalánosabb modális logikai hatszöget akarjuk felrajzolni, akkor a szükségszerűség és lehetségesség terminusai helyett az erős és gyenge modális operátorokat használva az alábbi ábrát kaphatjuk:
Hasonló ábrát lehetne felrajzolni az alethikus logikára, ami a szükségszerűség és a hozzá kapcsolódó fogalmak összefüggéseit mutatná, és a többi modális logika esetében is ugyanúgy érvényesnek bizonyulna a logikai hatszög az operátorok közti összefüggések szemléltetésére [80][81][520].

[28] „Ennek oka az, hogy az emberiség tapasztalati ismereteinek törzsanyaga fölszívódott a szavak jelentésébe, és ezért nehéz megvonni a határt a tapasztalati ismeretek és a nyelvhasználaton, a szavak jelentésén alapuló ismeretek között." [627, 265].

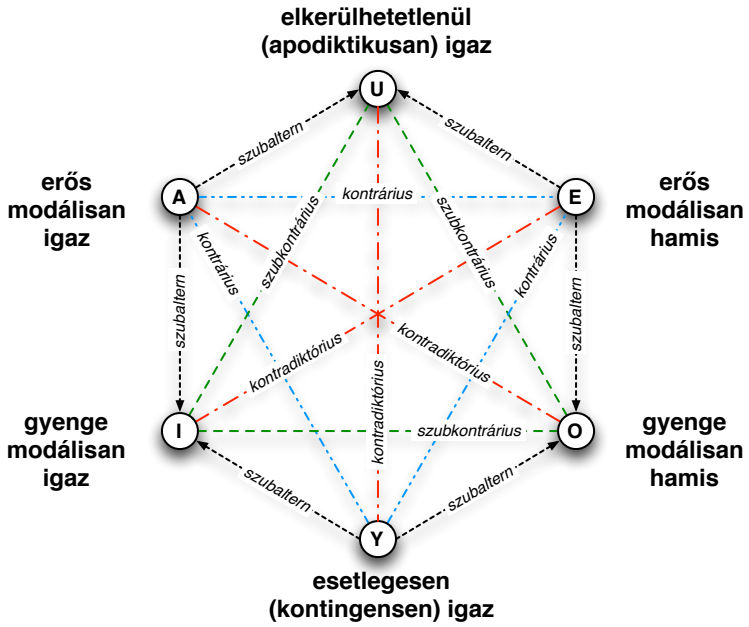


Figure 1.8: modális logikai hatszög

1.11 A tudás és a vélekedés logikája

Négyféle ember van.
Aki nem tudja, és nem tudja, hogy nem tudja: a balga – kerüld őt;
Aki nem tudja, de tudja, hogy nem tudja: az egyszerű – tanítsd őt;
Aki tudja, de nem tudja, hogy tudja: alszik még – ébreszd fel őt;
Aki tudja, és tudja, hogy tudja: ő a bölcs – őt kövesd.
(arab közmondás) [237]

Az idézet egyetlen fogalom, a tudás segítségével különít el négy embertípust egymástól. A balga és bölcs alakja közé kihúzott ívben a balga az, aki “nem tudva nem tud”, és a bölcs az, aki “tudva tud”. Az episztemikus logika, illetve a doxasztikus logika a modális logikának az az ága, amely a vélekedés, tudás, ismeret, hit logikai leírásának lehetőségeivel foglalkozik. A hétköznapi beszédben nem különítjük el élesen ezeket a terminusokat, s bár a **tudás** (knowledge) (ismeret) és **vélekedés** (hit) kettőssége mentén érezni valamilyen elkülönítési szándékot, de a terminológiában inkább zavar van e kérdésben. Az biztos, hogy amiről ennek kapcsán szó van (a tudás és a vélekedés kettőssége), az általában valamilyen mentális állapottal hozható kapcsolatba, olyan mentális állapottal, amelyek tartalma egy propozíció. A tudáshoz képes a vélekedést gyengébbnek, megalapozatlanabbnak érezzük. Egy vélekedés lehet megalapozatlan (hamis) és megalapozott (igaz), a tudásra ez nem mondható. Fontos feladat, hogy megtaláljuk, hogyan tehetünk különbséget e fogalmak között.

Angol nyelvben a ‘knowledge’ és ‘belief’ szavak vannak használatban, ezek közül a ‘belief’ fordítása nem egységes. Az angol ‘belief’ terminust lehetne ‘hit’-ként vagy

'hiedelem'-ként fordítani, de a magyarban mindkét terminusnak olyan konnotációja van, ami itt nem kívánatos értelmezést sugallna, amit szeretnék elkerülni. A magyar nyelvhasználatban ugyanis mind a hit, mind a hiedelem terminusához egyfelől valamiféle megalapozatlanságot, másfelől nem tudományos jellegű világmagyarázó intézményeket (vallást, ezoteriát) társítunk. Angolban a 'belief' terminus ebben az értelemben semleges, amit magyarul jobban kifejezhetünk a 'vélekedés' terminusával. Ennek ellenére használni fogom a 'hit' terminust is, ha az adott nyelvi kontextusban nem sugall transzcendentális kapcsolatot.

„Thoroughly conscious ignorance is the prelude to every real advance in science." (James Clerk Maxwell) Stuart Firestein: The pursuit of ignorance, TED-videó

1.11.1 Episztemikus logika

Az episztemikus logikát úgy építhetjük fel, hog először átértelmezzük a modális logika (K)-kalkulusát, majd azt további axiómákkal bővítve egyre erősebb nyelvhez jutunk [237]. Induljunk ki tehát a (K)-kalkulus nyelvéből.

(EL.0) $\mathscr{L}_{EL}(\mathscr{L}_{(K)C}, \boldsymbol{K})$, ahol

$\mathscr{L}_{(K)C}$	(K)-kalkulus
$\boldsymbol{K}_i$	tudja

Az episztemikus logikában a $\Box$ szükségszerű modális operátort úgy értelmezzük át, hogy annak jelentése az 'i ágens tudja, hogy a ϕ formula igaz', a jelölése pedig a $\boldsymbol{K}_i(\phi)$ szimbólum legyen. A $\boldsymbol{K}$ tudásoperátorhoz kapcsolt alsó index jelzi, hogy ennek a logikának a szemantikai újraértelmezés mellett van még egy másik újdonsága. Ez egy multimodális logika, ami azt jelenti, hogy több modális operátort is fel lehet venni a kalkulusba. Erre azért van szükség, hogy kezelni tudjuk azt a fontos társadalmi tényt, hogy a tudást, a hitet sok esetben nem általában, hanem az egyes ágensekhez rendelve értelmezzük. Természetesen előfordulhat az, hogy úgy akarunk, úgy tudunk mondani valamit a tudásra vonatkozóan, hogy el lehet tekinteni a – tudáshordozó – ágensektől, de nagyon sok más esetben pont arra van szükség, hogy jelezni, kezelni tudjuk azt, hogy melyik ágens tudásáról, hitéről van szó egy társadalmi kapcsolaton belül. Ilyen megfontolások miatt kell ellátni a tudásoperátort – alapértelmezés szerint – az ágensekre utaló indexekkel (az ágensekre itt is a természetes számok segítségével hivatkozunk).

(EL.1) $\text{TUDJA} \equiv \boldsymbol{K}_i(\phi) \equiv \boldsymbol{K}_i\phi \stackrel{\text{def}}{=} \forall i(\Box_i\phi), i \in \mathbb{N}^+$

tudja (HUN) – tudja, hogy (HUN) – knows (ENG) – knows that (ENG)

A modális logika (K)-kalkulusát (↬71.o.) itt úgy vehetjük át, hogy a szükségszerűség operátort a fentiek szerint átértelmezzük. A különböző szerzők többféle alakban is kifejezik az axiómát, amelyek természetesen teljesen ekvivalensek egymással. Kettőt bemutatok közülük: elsőként a modális (K)-kalkulus alakhű átértelmezését, majd utána egy átalakított formulát [237]. Az episztemikus disztribúciós szabályt tehát – kétféle formában – a következőképpen rögzíthetjük.

(EL.2) $(\text{K'}) \stackrel{\text{def}}{=} \boldsymbol{K}_i(\phi \to \psi) \to (\boldsymbol{K}_i\psi \wedge \boldsymbol{K}_i\psi)$

(EL.3) $(\mathrm{K''}) \stackrel{\text{def}}{=} (\boldsymbol{K}_i\phi \wedge \boldsymbol{K}_i(\phi \to \psi)) \to \boldsymbol{K}_i\psi$

episztemikus disztribúciós axióma (HUN) – epistemic distribution axiom (ENG)

A disztibúciós axiómát úgy is interpretálhatjuk, hogy a tudás zárt a logikai következményre (implikációra), azaz ha valaki tuda, hogy ψ állítás a ϕ-nek következménye, és még azt is tudja, hogy ϕ fennáll, akkor tudnia kell azt is, hogy a következmény, vagyis ψ-nak is fenn kell állnia.

Természetesen az episztemikus kalkulushoz is következtetési szabály(oka)t kell rendelnünk. Itt is érvényesnek tarthatjuk a modus ponens szabályt (↬25.o.), a modális generalizálás szabályát (↬76.o.) pedig át kell értelmezzük a $\boldsymbol{K}$ tudásoperátorra vonatkozóan.

(EL.4) $(\mathrm{MP}) \stackrel{\text{def}}{=} \{A, (A \to B)\} \Rightarrow B$

modus ponens következtetési szabály (HUN) – modus ponens inference rule (ENG)

(EL.5) $(\mathrm{NEC}) \stackrel{\text{def}}{=} \vdash \phi \Rightarrow \vdash \boldsymbol{K}_i\phi$

tudásgeneralizáció szabálya (HUN) – (Nec) (LOG) – (Gen) (LOG) – Knowledge Generalization rule (ENG)

Növelni lehet az episztemikus logika erősségét, ha további modális sémákat veszünk fel a kalkulusba. Első lépésként felvehetjük a modális (T)-sémát (↬71.o.) az axiómák közé (természetesen episztemikus módon interpretálva). Ennek az lesz az eredménye, hogy a tudás fogalma összekapcsolódik az igazságéval. További hozadéka lesz még a (T)-séma alkalmazásának az a tény, hogy ennek segítségével meg tudjuk különböztetni a **tudás** és a **vélekedés** fogalmát egymástól, amikor pont a (T)-axiómához való eltérő viszonyulás alapján definiáljuk a vélekedés fogalmát, s ezzel magát a doxasztikus logikát (↬79.o.). A modális (T)-séma episztemikus átértelmezése eredményeként az alábbi formulához jutunk.

(EL.6) $(\mathrm{T}) \stackrel{\text{def}}{=} \boldsymbol{K}_i\phi \to \phi$

tudás axióma (HUN) – knowledge axiom (ENG) – igazság axióma (HUN) – truth axiom (ENG) – episztemikus (T)-séma (HUN) – epistemic (T)-schema (ENG)

Nagyon fontos minőséget jelent ez tétel a tudás fogalmának értelmezésekor. Ha azt mondjuk, hogy tudunk valamit, akkor annak igaznak kell lennie. Nagyon erős állítás ez – különösen a vélekedés, hit definíciójával szembeállítva.

Tovább bővítjük az episztemikus kalkulust, ha felvesszük az axiómák közé a $\boldsymbol{K}$ tudásoperátorra vonatkozó (S4) modális sémát (↬72.o.). Ennek a lépésnek az lehet az értelme, hogy ezáltal az ágens saját tudásának tudásáról, vagyis introspekcióról, egyfajta reflektáltságról beszélhetünk. A formulában természetesen ebben az esetben is ki kell cserélnünk a modális operátorok szimbólumait.

(EL.7) $(\text{S4-PIK}) \stackrel{\text{def}}{=} \boldsymbol{K}_i\phi \to \boldsymbol{K}_i\boldsymbol{K}_i\phi$

pozitív introspekciós axióma (HUN) – positive introspection axiom (ENG) – (S4) (LOG) – (KK) (LOG) – reflektált pozitív tudás (HUN) – reflected positive knowledge (ENG)

Az axiómát úgy interpretálhatjuk, hogy ha az i ágens tud valamit, akkor tudja azt is, hogy tudja azt a valamit. Ez abban az értelemben erős tudás, hogy reflektált. A fejezetet elején mottóként közölt versrészlet szereplői közül az – utolsóként megénekelt – bölcsre illik ez a jellemzés. Az ősi versben dramaturgiai szerepe van, ahogy követik egymást az egyre inkább reflektált tudással bíró szereplők, így lehet eljutni a teljesen (reflektálatlanul) tudatlan embertől a reflektáltan tudó bölcsig. Az episztemikus logika másként "működik". Ha az episztemikus kalkulusba felvesszük az (S5) modális sémát is (↬73.o.), akkor olyan tudásoperátorhoz (tudástípushoz) jutunk, amit a köznapi gondolkodás még nem sokra tart. Ennek az új episztemikus sémának a jelentését és értelmét csak akkor érthetjük meg igazán, ha átalakítjuk az (S5)-sémát az $A = \neg\phi$ egyenlőség, illetve a (Dual)-séma (↬66.o.) alapján.

1) $\Diamond A \to \Box \Diamond A$
2) $\neg \Box \neg A \to \Box \neg \Box \neg A$
3) $\neg \Box \phi \to \Box \neg \Box \phi$

Az átalakítás első lépésben a (Dual)-szabályt alkalmazva kicseréljük a lehetséges modális operátort a szükségszerűre. A második lépésben behelyettesítjük az új változót a negált A-formulákra vonatkozó egyenlőség alapján, a következő lépésként már csak a $\boldsymbol{K}$ tudásoperátor szimbólumát kell átvezetni, és mindezek után az (S5)-axiómát a következő alakban ábrázolhatjuk.

(EL.8) $\text{(S5-NIK)} \stackrel{\text{def}}{=} \neg \boldsymbol{K}_i \phi \to \boldsymbol{K}_i \neg \boldsymbol{K}_i \phi$

negatív introspekciós axióma (HUN) – negative introspection axiom (ENG) – (S5) (LOG) – reflektált negatív tudás (HUN) – reflected negative knowledge (ENG)

Az axióma értelmezése az lehet, hogy ha az i ágens nem tud valamit, akkor tudja, hogy nem tudja. Logikai szempontból ez az – átalakított (S5)-sémával kiegészített – kalkulus a legerősebb. Eszerint az igazi bölcs az, aki egyaránt tudja azt, hogy mit tud, és azt is, hogy mit nem.

Definiálhatunk még további levezetési szabályokat is, például a **monotonitás** vagy a **kongruencia** szabályait. Ezeket azért mondhatjuk levezetési szabályoknak, mert a nyelv zártságát definiálják a kongruencia és monotonitás tulajdonságaira nézve.

(EL.9) $\text{(Mon)} \stackrel{\text{def}}{=} \vdash (\phi \to \psi) \Rightarrow \vdash (\boldsymbol{K}_i \phi \to \boldsymbol{K}_i \psi)$

monotonitási szabály (HUN) – monotony rule (ENG) – (Mon) (LOG)

(EL.10) $\text{(Cgr)} \stackrel{\text{def}}{=} \vdash (\phi \leftrightarrow \psi) \Rightarrow \vdash (\boldsymbol{K}_i \phi \leftrightarrow \boldsymbol{K}_i \psi)$

kongruencia szabály (HUN) – congruence rule (ENG) – (Cgr) (LOG)

Ezeket a szabályokat úgy is értelmezhetjük, hogy ezek azt fejezik ki, hogy a tudással rendelkező ágensek képesek tudássá formálni a logikai állításokat, vagyis ha egy adott episztemikus modellben érvényes egy kondicionális vagy egy bikondicionális művelet, akkor az összetevőkre vonatkozó tudások köztött is fennáll ugyanaz a művelet. Lehetne még további szabályokat definálni, amelyek mind az ágensek logikai következtető képességét fejeznék ki más és más logikai műveletekre fókuszálva, de ezt a képességet kifejezhetjük egy még általánosabb formában is. Rögzíthetünk ugyanis egy nagyon erős, és ezért sokak által vitatott következtetési szabályt is, ami azt írja le, hogy ha a logika

szabályai szerint egy adott formális elméletből (mondhatjuk így is: formulák valamilyen konjunkciójából) levezethető egy tétel, akkor az elmélet minden formulája (vagyis a konjunkciók elemi komponensei és a konjunkcióból következő "új" formula egyaránt) ismert az adott ágens számára. Formálisan kifejezve ezt a következő képletet kapjuk.

(EL.11) $(\mathrm{RK}) \stackrel{\mathrm{def}}{=} \vdash (\phi_1 \wedge \phi_2 \wedge \ldots \wedge \phi_n \rightarrow \psi) \Rightarrow \vdash (\boldsymbol{K}_i\phi_1 \wedge \boldsymbol{K}_i\phi_2 \wedge \ldots \wedge \boldsymbol{K}_i\phi_n \rightarrow \boldsymbol{K}_i\psi), n \in \mathbb{N}$

logikai mindentudás szabálya (HUN) – knowledge logical omniscience rule (ENG) – (RK) (LOG)

Akik ezt a szabályt "bevezették", érvényesnek tartották, azt abból a megfontolásból tették, hogy ennek segítségével lehet az episztemikus képességekkel rendelkező ágensek viselkedését minél jobban lemodelezni, számítógépes szimulációját minél jobban megvalósítani. Tették ezt abban a – jogos – meggyőződésben, hogy az ágenseik mindig megbízhatóan képesek a logikai szabályok figyelembe vételével helyes következtetéseket levonni. Ez a számítógépes modellezés során teljesen elfogadható. Akik viszont ezt a szabályt támadják, azok arra hivatkoznak, hogy az emberek esetében sokszor nem teljesen megalapozott ez az elvárás, az emberek nem képesek minden és mindenfajta logikai feladatot megoldani. A fenti szabályra valóban rá lehet húzni a "logikai mindentudás" kifejezést, de az ember esetében ennek teljes körű érvényességét nem szabad feltételeznünk. Ez nyilván nem logikai és – bizonyos értelemben – nem elméleti probléma. Egy cselekvéselmélet építése során azonban mégiscsak figyelembe kell venni az ember korlátait, ezért úgy gondolom, hogy ezt a szabályt fenntartásokkal kell kezelnünk. Nem eldobni kell, csak mindvégig figyelni kell az emberek valóságos logikai következtető képességeinek tényleges korlátaira. Ezt azonban nem ebben a fejezetben fejtem ki, hanem majd később.[29]

Superhuman logical capacities stalnaker

> Laci tudja, hogy Kati nem tudja, hogy Béla nem hiszi, hogy Klára tudja, hogy Péter azt hiszi, hogy a németek mind tudják, hogy az amerikaiak hisznek abban, hogy *A*.

Egy ilyen állítást tényleges tartalmát már nem igazán vagyunk képesek befogadni, noha az egymába ágyazott elemi állítások mindegyike könnyen értelmezhető. A hétköznapi gyakorlatban azonban sosem fordulnak elő ilyen hosszúságú (és összetettségű) mondatok, ezért nem tudjuk értelmezni. Egy számítógép számára viszont – persze csak akkor, ha az általa olvasható formában áll rendelkezésre – ugyanolyan könnyedén feldolgozható ez a mondat is, mint azok az egyszerűek, amelyeket még mi is kezelni tudunk. A fenti példamondat a közösségi helyzeteket jellemzi, amikor nem az összetett mondat bonyolultsága jelenti a nehézséget, hanem az egy csoportba tartozó, tudással rendelkező ágensek egyéni tudásainak egymásra vonatkozása. A másikfajta feldolgozási nehézséget az összetett mondatok logikai szerkezetének bonyolultsága okozhatja. Ez egyetlen tudásképes ágens esetében is jelentkezhet.

korlátozott racionalitás (bounded rationality)

R. Stalnaker.1996-os cikke [705]

(S4)-séma, a tranzitivitás kérdése sokan nem fogadják el

kritika1: Fred Dretske [217]

kritika2: R.Nozick.1981 [541]

[29] amikor már meglesz, ugró hivatkozást betenni ide

kritika3 T.Williamson1999 [821]

kritika4 T. Williamson 2013 Gettier Edmund1963: [287] [823]

Timothy Williamson 'Knowledge and Its Limits' című könyvében 8.2.paragrafus [822]

az episztemikus logika egyik komoly feltételezése, hogy az ágensek teljesen racionálisak abban az értelemben, hogy tökéletes következtetési képességgel rendelkeznek. Ez egyben nyilván komoly korlátja is az elméletnek, ha a tényleges társadalmi gyakorlat tényleges ágenseire akarjuk mindezt vonatkoztatni.

1.11.2 Doxasztikus logika

A (T)-séma azért lehet alkalmas a tudás és vélekedés fogalmainak elválasztására, mert amíg a tudás fogalmával szemben fenntarthattuk a séma érvényességét (↬76.o.), addig a vélekedésre vonatkozóan pont az ellenkezőjét kell mondanunk: ebben az esetben érvénytelennek kell tartanunk az igazság axiómát. Azt ugyanis mondhatjuk, hogy valaki igaznak vél valamit, ami hamis, csak ilyenkor hozzá kell tennünk, hogy tévesen véli, tévesen hiszi azt, ami nem igaz. Azt viszont nem mondhatjuk, hogy valaki tud valamit, ami hamis. Ha valami hamis, akkor az nem tudás, csak – hamis – vélekedés.[30] Ezen a ponton lehet elágaztatni az episztemikus és doxasztikus logikát egymástól azon az alapon, hogy amíg az előbbi a tudás, addig az utóbbi a vélekedés fogalmának formalizálásával foglalkozik. Mielőtt belekezdenék a doxasztikus logika vagyis a vélekedés (hit) logikájának bemutatásába, jeleznem kell, hogy ezt a logikát annyira közelállónak tekintem az episztemikus logikához, hogy nem definiálok számára új nyelvet, a tételeket ugyanúgy az episztemikus logikára utaló 'EL' szimbólummal azonosítom, a tételek számozását folytatom. Annak ellenére tehát, hogy a doxasztikus logikát külön fejezetben tárgyalom, sok szempontból .

A doxasztikus logika központi kategóriája a **vélekedés**, amit a tudás fogalmával szemben kell értelmeznünk. Ha a tudásra érvényesnek, a vélekedésre érvénytelennek tartjuk a (T) tudás axiómát, akkor kérdés, hogy milyen tulajdonsággal jellemezhetjük a vélekedés fogalmát. Ehhez a modális logika (D)-sémáját kell elővenni, amit persze újfajta módon kell értelmeznünk. Ahhoz, hogy megértsük, mi lehet az megfelelő új interpretáció, át kell alakítani kicsit a modális logika (D)-sémáját.

1) $(\mathrm{D}) \equiv \Box\phi \rightarrow \Diamond\phi$
2) $\Box\phi \rightarrow \neg\Box\neg\phi$
3) $\boldsymbol{B}_i\phi \rightarrow \neg\boldsymbol{B}_i\neg\phi$

Az 1) és 2) között kicseréltük a $\Diamond$ lehetséges operátort a (Dual)-szabály alapján a $\Box$ szükségszerű operátorra (külső és belső tagadással együtt), majd a 2) és 3) között a modális szükségszerű operátort a $\boldsymbol{B}_i$ vélekedésoperátorral helyettesítettük. Ezek alapján a következő formulával határozhatjuk meg a vélekedés egy nagyon fontos tulajdonságát.

(EL.12) $(\text{D-CB}) \stackrel{\text{def}}{=} \boldsymbol{B}_i\phi \rightarrow \neg\boldsymbol{B}_i\neg\phi$

konzisztencia axiómája (HUN) – Consistency of Belief axiom (ENG) – doxasztikus (D)-séma (HUN) – doxastic (D)-schema (ENG)

[30] A köznyelv természetesen nem határolja el egymástól ezt a két fogalmat ennyire pontosan és egyértelműen.

Ez a formula a vélekedés (hit) konzisztenciájáról szól, ugyanis azt rögzíti, hogy ha valaki hisz valamiben (ϕ-ben), akkor nem hihet annak az ellenkezőjében ($\neg\phi$-ben). Ez a tudásra nem lenne igaz, csak a vélekedésre. Ezért ez a (T)-axiómánál gyengéb episztemikus (D)-séma a vélekedést, a hitet jellemző doxasztikus logika specifikumát adja.

A vélekedéshez a $\boldsymbol{B}$ szimólumot és a következő axiómát rendelhetjük.

(EL.13) $\text{BELIEF} \stackrel{\text{def}}{=} \neg \boldsymbol{B}_i \bot$

vélekedés (HUN) – belief (ENG)

A formulát úgy értelmezhetük, hogy valaki (az i ágens) nem hiheti azt, hogy egy ellentmondás igaz lehet (nagyon fontos nem "elfelejteni", hogy ebben az esetben az episztemikus (T)-sémát nem tartjuk érvényesnek). A (T)-séma érvénytelensége miatt a doxasztikus logika nem olyan erős, mint az episztemikus logika (T)-sémával megerősített változata.

(EL.14) $\text{(S4-PIB)} \stackrel{\text{def}}{=} \boldsymbol{B}_i\phi \rightarrow \boldsymbol{K}_i\boldsymbol{B}_i\phi$

a hit pozitív introspekciós axiómája (HUN) – positive introspection of belief (ENG) – doxastic (S4) (LOG) – (BB) (LOG)

(EL.15) $\text{(S5-NIB)} \stackrel{\text{def}}{=} \neg\boldsymbol{B}_i\phi \rightarrow \boldsymbol{K}_i\neg\boldsymbol{B}_i\phi$

a hit negatív introspekciós axiómája (HUN) – negative introspection of belief (ENG) – (S5) (LOG)

Should Knowledge Entail Belief? [320]

(EL.16) $\text{KB} \stackrel{\text{def}}{=} \boldsymbol{K}_i\phi \rightarrow \boldsymbol{B}_i\phi$

a hit tudás alapú axióma (HUN) – knowledge imply belief axiom (ENG)

doxasztikus logikában is felvehető a (RK)-szabály "megfelelője", annyit kell csak tennünk, hogy a tudásoperátort kicseréljük a vélekedésoperátorra.

(EL.17) $\text{(RB)} \stackrel{\text{def}}{=} \vdash (\phi_1 \wedge \phi_2 \wedge \ldots \wedge \phi_n \rightarrow \psi) \Rightarrow \vdash (\boldsymbol{B}_i\phi_1 \wedge \boldsymbol{B}_i\phi_2 \wedge \ldots \wedge \boldsymbol{B}_i\phi_n \rightarrow \boldsymbol{B}_i\psi), n \in \mathbb{N}$

a logikai mindenhit szabálya (HUN) – belief logical omniscience rule (ENG) – (RB) (LOG)

[321] [708] [361]
vietnami phd [219]
Dynamic-Epistemic Spatial Logic [469]
(B)-séma a szimmetricitást tételezi

Fitch- és Moore-probléma

mondani, de nem hinni (saying but disbelieving)

G.E. Moore által felvetett hitproblémát a következő mondattal reprezentálhatjuk:

> Az a helyzet, hogy *A*, de én nem hiszem, hogy *A*. (*A* but I do not believe that *A*.)

Jaako Hintikka elemezte először az episztemikus logika eszköztárára támaszkodva Everything is Knowable ? How to Get to Know Whether a Proposition is True [781]
episztemikus logika könyv [505]

1.11.3 Kollektív tudás

Sok szó esett már eddig a tudás, a hit fogalmáról, és bár mindvégig nyilvánvaló volt, hogy az ezeket reprezentáló operátorokat ($\boldsymbol{K}_i$, illetve $\boldsymbol{B}_i$) össze kell kapcsolnunk azokkal az ágensekkel, akit "hordozzák" a tudás- és hittartalmakat, viszont eddig csak jeleztük ezt a kapcsolatot, de nem igazán használtuk ki ezt az információt. A társadalmi cselekvések és társadalmi kapcsolatok értelmezése során rendkívüli jelentősége van annak a ténynek, hogy a tudás, a hit lényege szerint kollektív jelenség, az ágensek a saját egyéni tudásaikat kicserélhetik, megoszthatják egymás között, az egyéni tudásaikat mások tudásaira "vonatkoztathatják", és ezáltal az egyéni tudások mellett másfajta, kollektív tudások is megjelennek a társadalmi térben. A kollektív tudás jelenségének fontosságát önmagában jelzi az a tény, milyen sokféle terminust használunk ebben a fogalomkörben. A későbbiekben a **közös tudás**, a **közismert tudás**, illetve a **elosztott tudás** fogalmait fogom majd a társadalmi cselekvés modelljébe beemelni, de ezekkel együtt fordulnak elő a terepen az olyan kategóriák, mint az együttes tudás, kölcsönös tudás, köztudott tudás, magától értetődő tudás, nyilvános tudás, kumulatív tudás, kollektív bölcsesség, szétszórt tudás, csoporttudás. Mindezeket pedig együttesen lefedi a kollektív tudás felettes fogalma.

A három kiemelten kezelni kívánt fogalom közül igazából csak a közös tudás és az elosztott tudás fogalma a fontos, a közismert tudás kategóriájára csak azért van szükségünk, hogy a közös tudás jelenségét értelmezni tudjuk a segítségével. A következőkben ezért előbb a közös tudás fogalmát mutatom be, majd ezt követi az elosztott tudás fogalmának rövid ismertetése. A közös tudás jelenségét először David Lewis vizsgálta monografikus igényességgel a 'Convention' című könyvében [444], majd még sokan követték őt a jelenség értelmezésében. A jelenség megragadására John McCarthy, a mesterséges intelligencia egyik úttörő kutatója javasolta használatba venni a "minden bolond tudja" ("any fool knows") kifejezést [237], de ez – számomra – inkább a közismert tudás jelentését sugallja. Nehéz ezekkel a kifejezésekkel pontosan operálni. nem is tehetünk mást, mint hogy egy-egy kifejezést kiragadunk a szokásos diskurzusból, és pontos, rögzített jelentéssel látjuk el őket. Ezáltal persze elszakadunk attól a lehetőségtől is, hogy a hétköznapi gyakorlatból eredő intuícióinkra hivatkozzunk.

Ahhoz, hogy a kollektív tudás és más hasonló fogalmak elemzésébe kezdjünk, először be kell vezetnünk a társadalmi csoport fogalmát. A csoportba tartozó ágensek tudásait, hiteit vizsgálva, a csoport egyes tagjait indexekkel jelölve, a csoport fogalmára hivatkozva definiálhatjuk majd a kollektív tudás általános kategóriája alá sorolható további tudásfogalmakat. A tudást hordozó ágens, az ágenseket összefogó csoport fogalmával itt még nagyon kell foglalkoznunk, ezekre a fogalmakra később még úgyis kitérek még. Most elég csak rögzíteni, hogy vannak ágensek, akik csoportokba rendeződ(het)nek. A csoportot G-vel jelöljük, az ágenseket, ha ez nem zavaró, akkor az i természetes számokkal azonosítjuk, ha szükséges, akkor az a_i szimbólumot alkalmazzuk rájuk. A közösség ekkor a tagjainak uniója.[31]

(EL.18) $$G \stackrel{\text{def}}{=} \bigcup_{i=1}^{n} a_i$$

---

csoport (HUN) – group (ENG) – közösség (HUN)

[31] Amíg nem zavaró, a közösség és a csoport terminusait felváltva, itt még azonos jelentés mentén alkalmazom.

Az i. ágenshez kapcsolható $\boldsymbol{K}_i$ egyéni tudás operátorát már sokat használtuk korábban, most kell kidolgoznunk azokat a tudásfogalmakat, amelyek segítségével már meg tudjuk ragadni az egyéni tudások kölcsönösségét, egymásra vonatkozatottságát, illetve le tudjuk írni a csoport egészéhez köthető tudásokat is. Ehhez azonban először meg kell tudnunk adni azt, hogy az ágensek egyénileg milyen tudásokkal rendelkeznek. Ehhez nem kell mást tennünk, mint összegezni az egyes egyénekre az elemi tudásaikat. Ekkor is használhatjuk a már bevezetett $\boldsymbol{K}_i$ tudásoperátort, és ha csak magában használjuk, akkor az i. egyén össze tudását érttjük alatta, ha pedig a tudásoperátor után teszünk még az adott tudáselemre utaló jelet, akkor értelemszerűen az arra a konkrét tudáselemre vonatkozó tudás tényét állítjuk. Most tehát az egyén teljes tudását kell definiálnunk.

(EL.19) $$\boldsymbol{K}_i \stackrel{\text{def}}{=} \bigcup_{j=1}^{m} \boldsymbol{K}_i\phi_j$$

---

egyén teljes tudása (HUN) – agent's full knowledge (ENG)

Az egyén teljes tudása (tudattartalma) az egyéni tudáselemek uniója. Ez a fogalom segíteni fog minket a kollektív tudás körébe tartozó fogalmak értelmezésében. Lássuk, hogyan!

Közös tudás

A közös tudás fogalmának értelmezéséhez először a **közismert tudás** kategóriáját kell bevezetnünk, ami nem egy bonyolult fogalom. Azt a jelenséget ragadhatjuk meg vele, amikor egy csoporton belül egyénileg mindenki rendelkezik ugyanazzal a ϕ tudással, vagyis amit "minden bolond tud". Ha a G csoport bármelyik i tagja tudja ϕ-t, akkor azt a következő formulával írhatjuk le.

(EL.20) $$\boldsymbol{K}_G^E\phi \stackrel{\text{def}}{=} \bigwedge_{i\in G} \boldsymbol{K}_i\phi$$

---

közismert tudás (HUN) – köztudott tudás (HUN) – együttes tudás (HUN) –magától értetődő tudás (HUN) – közös tudás (HUN) – mutual knowledge (ENG)

Nehéz széles körben elfogadott, a fenti formulával definiált, pontos jelentés mentén értelmezett kifejezést rendelni ehhez a fogalomhoz. A definíció alatt felsorolt nyelvi alakzatok mindegyike használható lenne, csak a választási kényszer miatt emeltem ki közülük egyet (a 'közismert tudás' kifejezést), de ha nem lesz zavaró, a későbbiekben is megengedhetőnek tartom, hogy szinonímaként használjuk az alternatív kifejezéseket a fogalomra vonatkozóan. [32] A közismert tudás definíciója a G csoport egészére, de egyetlen ϕ tudáselemre vonatkozik. Továbbléphetünk egyet, ha erre a fogalomra támaszkodva definiáljuk a csoport teljes közismert tudását. Ehhez a $\boldsymbol{K}_i$ egyéni teljes tudások (tudáskészletek) közös metszetét kell előállítanunk.

(EL.21) $$\boldsymbol{K}_G^E \stackrel{\text{def}}{=} \bigcap_{i=1}^{n} \bigcup_{j=1}^{m} \boldsymbol{K}_i\phi_j$$

---

csoport teljes közismert tudása (HUN) – group's full mutual knowledge (ENG)

[32] Csak a kölcsönös tudás terminust zárnám ki (még akkor is, ha az angolban az ennek megfelelő 'mutual knowledge' alakot használják a leggyakrabban). A kölcsönösség számomra sokkal inkább sugallja a szimmetricitást és reflektáltságot, ami majd a közös tudás fogalmához lesz elengedhetetlen.

A $\boldsymbol{K}_G^E$ közismert tudáskészlet minden $\boldsymbol{K}_i$ egyéni tudáskészletben benne van, vagyis a $\boldsymbol{K}_i^E$ egyéni tudatok össztartalma mindig szélesebb vagy azonos terjedelmű, mint a $\boldsymbol{E}_G$ közismert tudat. Lehetséges (bár nem valószínű), hogy a $\boldsymbol{K}_G^E$ közismert tudat megegyezik valamelyik i személy $\boldsymbol{K}_i$ egyéni tudattartalmával. Amennyiben nem ez a helyzet áll fent (és ez a valószínűbb kimenet), akkor azt mondhatjuk, hogy $\boldsymbol{K}_G^E$ közismert tudat különbözik minden egyes $\boldsymbol{K}_i$ egyéni tudattól. Ilyen esetben hiába van jelen a $\boldsymbol{K}_G^E$ kollektív tudat minden eleme az összes $\boldsymbol{K}_i$ egyéni tudatban, mégis különbözik mindegyiktől, vagyis a kollektív tudat mindegyik egyéni tudathoz képest más. Ez a fontos meglátás az alapja Emilé Durkheim társadalomelméletének is.

A tudást közismertnek mondhatjuk abban az értelemben, hogy mindenki részesül belőle. Azonban ez nem jelenti azt, hogy annak is tudatában lennének a csoport tagjai, hogy ezt a valamennyiük által egyénileg birtokolt tudást valamennyien ismerik. Sőt, elvileg az is elképzelhető, hogy mindenki azt feltételezi, hogy csak ő rendelkezik az adott tudással, senki más. Andersen meséje a meztelen királyról szép példája annak, hogy ami közismert tudás egy közösségen belül, az még nem feltétlen jelent közös tudást is egyben. Mindenki látta, tudta, hogy a királyon nincs ruha, de amíg valaki el nem kiáltotta magát, hogy "a király meztelen", addig senki sem tudhatta, hogy amit egyénileg mind tudnak, azt mindenki más is tudja, vagyis a tudás közös. Az Andersen-mesének persze létezik más olvasata is, de ez a tény nem teszi érvénytelenné a rá való hivatkozást ezen a ponton (és erre később még visszatérek).

Ha azt akarjuk megragadni, ami az Andersen-mese tanulsága, akkor a **közös tudás** (vagy kölcsönös tudás) fogalmát kell definiálnunk. Ehhez új elemként fel kell vennünk a nyelvünkbe a közös tudást reprezentáló $\boldsymbol{K}_G^C$ modális operátort, amit a következőképpen definiálhatunk.

(EL.22) $\boldsymbol{K}_G^C$

közös tudás (HUN) – kölcsönös tudás (HUN) – common knowledge (ENG) – mutual knowledge (ENG)

A fogalom értelmezése: a G csoport minden tagja rendelkezik a közös tudással.

———————-

képletek, értelmezések, összehasonlítások
modeling common belief: [291]
közös tudásról újra: [37]

———————-

Fentebb említettem, hogy Andersen 'A császár új ruhája' című meséjéhez többféle értelmezést is hozzárendelhetünk. Legalább három lehetőség adódik. Azért interpretálhatunk többféleképpen is, mert van egy nagy adag bizonytalanság abban, hogy miként lehet előállítani azt a helyzetet, ami a közös tudáshoz eléréséhez szükséges "tudom, hogy tudja" állapotot eredményezi. Az első lehetőség az, hogy valóban nem tudják a csoporttagok, hogy a többiek tudják, ami maguk egyénileg mind tudnak, és a nyilvános bejelentés mindenkit meglep. Ilyen esetekben az a bejelentés, aminek eredményeként a közös tudás létrejön, csodálkozást, meghökkentést, meglepetést vált ki a csoport tagjaiban. A második értelmezés szerint a csoporttagok mind sejthetik, de nem tudhatják biztosan, hogy a többiek is a közös tudás birtokában vannak. Ekkor a mindenki által tudott tudás közzététele a korábbi bizonytalanságból adódó feszültség feloldódásához, egyfajta megkönnyebbüléshez, megnyugváshoz vezet. A harmadik értelmezési lehetőséget nem

lehet teljes mértékben elválasztani a második interpretációtól, de érdemes elkülöníteni tőle. Előfordulhat ugyanis, hogy a egyetemesen birtokolt egyéni tudások azért nem válhatnak sokáig közös tudássá, mert a közzététel útjában társadalmi akadályok vannak, valamiért nem lehet kimondani a közösség előtt a mindenki által tudott tudást. Ez is feszültséget okozhat a csoport tagjaiban, ám az ilyen helyzetekhez másfajta érzelmek kapcsolódnak, és ha mégis megszületik a nyilvános bejelentés, akkor azt felszabadultság követi (az Andersen-mesében ez utóbbival állunk szemben, amit a tömeg reakciója bizonyít: a királyt leleplező felkiáltás után a tömeg nevetésben robban ki).

---

valahol később a nyilvánosságról írni és ide hivatkozni vissza: nyilvánosság egymásba ágyazódó, a kétszemélyes köröktől kezdődő és az egyre nagyobb számosságú, végül a teljes csoportot magába foglaló teljes nyilvánosság

---

Dan Sperber és Deirdre Wilson ‘Kölcsönös tudás és relevancia a megértés elméletében’ cikk [702]

1.11.4 Közös tudás

lásd még a karizmatkus uralomnál a Bourdieu idézetet

ez a fogalom a konvencióhoz is kell

Lewis konvenciófelfogásában kiemelt helye van a közös tudásnak

kritika: Margaret Gilbert + magyar tanulmány a konvencióról

Szolcsányi Tibor, Davidson háromszögeléséről, [736]

Chwe, rational ritual

Binmore, short introduction to game theory könyvében szó esik a közös tudásról. [86]

Auman: lehetetlen a közös tudás [38]

p-belief: DovMondererandDovSamet, Approximating Common Knowledge with Common Belief, [512]

Elosztott tudás

A közös tudás fogalmának tárgyalásakor az volt a fontos, hogy mi az a tudáselem, amit a közösség tagjai mind közösen birtokolnak, vagyis hogy mi az, ami ott van a csoport mindegyik tagjának a fejében. Az **elosztott tudás** fogalmával nem az ebben az értelemben vett közös részt akarjuk megragadni, hanem valami mást. A kollektív bölcsesség elve, a ‘sok kicsi sokra megy’ mondása és még sok minden más ősrégi megfigyelés fejezte és fejezi ki azt a társadalmi tényt, hogy a közösség olykor több lehet, mint a részei, vagyis a tagjai összessége. Azt a tudásösszességet, ami a közösség egésze rendelkezésére áll nevezhetjük elosztott tudásnak. A korábban bevezetett egyéni teljes tudás fogalmára támaszkodva már könnyen definiálhatjuk ezt a fogalmat.

(EL.23) $$\boldsymbol{K}_G^D \stackrel{\text{def}}{=} \bigcup_{i=1}^{n} \bigcup_{j=1}^{m} \boldsymbol{K}_i \phi_j$$

---

csoport teljes elosztott tudása (HUN) – group’s full distributed knowledge (ENG)

Ez a tudásfogalom nem más, mint az egyéni tudattartalmak (tudáselemek) uniója, vagyis a csoport tagjainak fejében levő tudás valahogyan képzett összessége.

which reads "it is distributed knowledge to every agent in G.

Francis Galton 1907-ben publikált egy cikket a Nature-ben egy Plymout városában tartott állatvásár tapasztalatairól [279]. A látogatókat versenyre invitálták: meg kellett becsülni egy levágandó bika vágás utáni súlyát. A versenynek jelentkezési díja volt, ami miatt gyakorlatilag ki lehetett zárni a "vicces becsléseket", viszont komolynak mondható díjat ajánlottak fel a nyertes számára, ami miatt sokan beszálltak a játékba. A verseny után a közel nyolcszáz játékos által leadott papírokat Galton megkapta, és elemezte a rajtuk található adatokat. Meglepő eredmény született: az 1198 font súlyú állatra leadott becslések mediánja 1207 font lett, ami az egy százalékos hibahatáron belül maradt.

Galton a "kísérlet" eredményét a demokratikus szavazás, a Vox Populi elvének igazolásának is szánta, bár ebben nem volt teljesen igaza (mivel itt tényre vonatkozó átlagolásról, a demokratikus szavazás során értékre vonatkozó átlagolásról van szó).

> „This result is, I think, more important creditable to the trustworthiness of a democratic judgment than might have been expected." [279]

"egy ember egy szavazat" elve

Luckmann, Thomas; Berger, Peter L., A valóság társadalmi felépítése. Tudásszociológiai értekezés Jószöveg Műhely Könyvkiadó, 1998

RATIONAL DYNAMICS AND EPISTEMIC LOGIC IN GAMES [769] Binmore 1992 justifies the above steps in the strictly dominated (SD) algorithm by means of iterated mutual knowledge of rationality:

„E.g., A can be sure that E will disregard the right-hand column, as A knows that E is rational. And E can later remove the second column since she knows that A knows her rationality leading to discarding action c, and will therefore remove the bottom-most row. And so on."

Muddy Children' (Fagin et al. 1995).

---

közösségi nyomás, hatás, Benoit Mandelbrot, Máté-hatás: „a kinek van, annak adatik" (Máté 25:29).

Karinthy Jézus-Barabás meséje

1.11.5 Dinamikus episztemikus logika

A tudásról, vélekedésről eddig úgy beszéltünk, hogy adottnak és változatlannak vettük tudásaink, vélekedéseink tartalmát. Ezek a tartalmak nyilván nem statikusak, módosulnak az időben, régóta fennálló tudások, hitek tűnnek el, új tudások, új hitek bukkannak fel, terjednek el széles körben, állandó változásban élünk. Nagyon jó lenne, ha meg tudnánk ragadni ezeket a változási folyamatokat. Ebben segíthet minket a dinamikus episztemikus (doxasztikus) logika.

tudás, vélekedés és változás, update, ugrade, revizió

dinamikus episztemikus logika [778]

Public announcements, public assignments and the complexity of their logic [780]

tudás, tudatosság, cselekvés Action models for knowledge and awareness [779]

fontos írás: Carlos E. Alchourrón and Peter Gärdenfors és David Makinson On the Logic of Theory Change: Partial Meet Contraction and Revision Functions [7]

Lindström, St. and Wl. Rabinowicz: DDL Unlimited. Dynamic Doxastic Logic for Introspective Agents. In: Erkenntnis 51, 1999, p. 353-385.

Linski, L.: On Interpreting Doxastic Logic. In: The Journal of Philosophy 65, 1968, p. 500-502.

Segerberg, Kr.: Default Logic as Dynamic Doxastic Logic. In: Erkenntnis 51, 1999, p. 333-352.

Wansing,H.: A Reduction of Doxastic Logic to Action Logic. In: Erkenntnis 53, 2000, p. 267-283.

Stalnaker, On Logics of Knowledge and Belief [708]

belief revision [706] Iterated Belief Revision [709]

Social Norms, Rational Choice and Belief Change: [29]

1.11.6 Tudástipológiák

A tudás fogalmát nemcsak az episztemikus és doxasztikus logika művelői elemezték, sok más tudományterület foglalkozott még a tudás kérdéskörével.

rejtett tudás hallgatólagos tudás, tacit tudás implicit vs. explicit tudás személyes tudás Ryle-tipológia

1.12 Preferencialogika

Ha azt mondjuk, hogy *A*-t preferáljuk *B*-vel szemben, akkor nem teszünk mást, mint valamiféle sorrendet alakítunk ki a két dolog között. A sorrendállítás pedig rendezést jelent. Rendezésből persze többféle van, erről bővebben írtam a könyv relációelméleti részében (↬38.o.). A matematikai rendezésfogalmak azonban túl általánosak, ezért sok jelenségre "ráhúzhatók". Ha speciálisabb rendezési fogalmakat szeretnénk definiálni, szükségünk lehet további tulajdonságokat kikötni. A reflexív, antiszimmetrikus és tranzitív tulajdonságok egyidejű feltételezésével többféle rendezési relációt is jellemezhetünk. A **része** (partitív), a **tartalmazási**, a **nem öregebb** vagy a **kisebb vagy egyenlő** reláció esetén egyaránt fennállnak a jelzett tulajdonságok, és érezzük, hogy szükség lenne további minőségeket hozzájuk rendelni, hogy jól elkülöníthessük őket egymástól. A partitív reláció esetében megtette ezt Peter Simon vagy Archello Varzi. Előbbi könyvet írt a rész fogalmáról [685], utóbbi axiómarendszert dolgozott ki a partitív reláció fogalmára [786], amelyben a három alaptulajdonságon túl még 4-5 további tulajdonságot definiált.[33]

(a_1) $\forall x(\mathbf{P}(x,x))$

(a_2) $\forall x\forall y(\mathbf{P}(x,y)\wedge\mathbf{P}(y,x)\rightarrow x=y)$

(a_3) $\forall x\forall y\forall z(\mathbf{P}(x,y)\wedge\mathbf{P}(y,z)\rightarrow\mathbf{P}(x,z))$

Az első három axióma rögzíti, hogy a **P** partitív reláció reflexív (a_1), antiszimmetrikus (a_2) és tranzitív (a_3). A következő lépésben a **P** relációval néhány új mereológiai relációt definiálunk, majd az újakkal még újabbakat hozunk létre:

(1) $\mathbf{PP}(x,y)\stackrel{\text{def}}{=}\forall x\forall y(\mathbf{P}(x,y)\wedge\neg\mathbf{P}(y,x))$

[33] A partitív reláció Varzi-féle axiomatizálását bemutattam magyarul [730].

(2) $\mathbf{O}(x,y) \stackrel{\text{def}}{=} \forall x \forall y \exists z(\mathbf{P}(z,x) \wedge \mathbf{P}(z,y))$

(3) $\mathbf{U}(x,y) \stackrel{\text{def}}{=} \forall x \forall y \exists z(\mathbf{P}(x,z) \wedge \mathbf{P}(y,z))$

(4) $\mathbf{OX}(x,y) \stackrel{\text{def}}{=} \forall x \forall y(\mathbf{O}(x,y) \wedge \neg\mathbf{P}(x,y))$

(5) $\mathbf{UX}(x,y) \stackrel{\text{def}}{=} \forall x \forall y(\mathbf{U}(x,y) \wedge \neg\mathbf{P}(x,y))$

(6) $\mathbf{PO}(x,y) \stackrel{\text{def}}{=} \forall x \forall y(\mathbf{OX}(x,y) \wedge \mathbf{OX}(y,x))$

(7) $\mathbf{PU}(x,y) \stackrel{\text{def}}{=} \forall x \forall y(\mathbf{UX}(x,y) \wedge \mathbf{UX}(y,x))$

Sorba definiáljuk a következő relációkat: valódi része (1), átfedése (2), kívüllevősége (3), részhatáros átfedése (4), részhatáros kívüllevősége (5), valódi lefedése (6), valódi kívüllevősége (7). Ezekután a **P** része, **O** átfedése és **U** kívüllevősége relációk segítségével először meghatározhatjuk a kiterjeszthetőség (erős) elvét (a_4), majd az (a_5)-(a_7) mereológiai axiómák elfogadásával definiálhatjuk a zárt (extenzionális) mereológia elméletét.

(a_4) $\forall x \forall y(\neg\mathbf{P}(x,y) \exists z(\mathbf{P}(z,x) \wedge \neg\mathbf{O}(z,y)))$

(a_5) $\forall x \forall y(\mathbf{U}(x,y) \rightarrow \exists z \forall w(\mathbf{O}(w,z) \leftrightarrow (\mathbf{O}(w,x) \vee O(w,y))))$

(a_6) $\forall x \forall y(\mathbf{O}(x,y) \rightarrow \exists z \forall w(\mathbf{P}(w,z) \leftrightarrow \mathbf{P}(w,x) \wedge P(w,y))))$

(a_7) $\forall x \forall y \exists z((\mathbf{P}(z,x) \wedge \neg\mathbf{O}(z,y)) \rightarrow \exists z \forall w(\mathbf{P}(w,z) \leftrightarrow (\mathbf{P}(w,x) \wedge \neg\mathbf{O}(w,y)))))$

További axiómák felvételével egyrészt definiálhatjuk az általános (extenzionális) mereológia elméletét (a_8) másrészt meghatározhatjuk az atomos (a_9), illetve atom nélküli mereológiákat (a_{10}).

(a_8) $\exists x\phi \rightarrow \exists z \forall y(\mathbf{O}(y,z) \leftrightarrow (\exists x(\phi \wedge \mathbf{O}(y,x))$

(a_9) $\forall x \exists y(\mathbf{P}(y,x) \wedge \neg\exists z\mathbf{PP}(z,y))$

(a_{10}) $\forall x \exists y(\mathbf{PP}(y,x))$

A mereológia elméletének bemutatásához, megérttetéséhez nyilván nem elégséges, hogy formulákkal definiáljuk az alaprelációkat, meg az axiómákat, és kész. Én mégis ezt tettem meg itt. Azért csak ennyit foglalkoztam a kérdéssel, mert csak azt akartam megmutatni, hogy a RÉSZE reláció meghatározásához milyen logikai apparátust kell megmozgatni. Nem elég rögzíteni az első három relációtulajdonság meglétét, mert azzal csak egy "szimpla" rendezési relációt határozhatunk meg, ami túl általános, még nagyon sokféle dolgot képes lefedni. Láthattuk, hogy további 5-6 axióma kell ahhoz, hogy a partitív reláció specifikumát megtaláljuk.

Ez az egész pedig csak azért érdekes e könyv mondanivalója szempontjából, mert a preferenciarelációra is ki lehet jelenteni, hogy az aszimmetrikus és tranzitív egyben, ám ezzel még biztos nem tudjuk megragadni a preferenciareláció különlegességét, hiszen több más ismert relációt jellemezhetnek ezek a minőségek, ezért keresni kell még további tulajdonságokat hozzá. Ezt végzik el azok a szerzők, akik a preferencialogika különböző elméleteit dolgozzák ki. Georg Henrik Von Wright is, a követői is szokták a preferencia terminusa helyett a – magyarra lefordíthatatlan – **betterness** terminust használni. A kategória tartalmát talán a meglehetősen kicsavart **jobbnak tartottság** kifejezéssel adhatnánk vissza, de e jelzésen túl, hogy létezik, ezt a terminust a természetes nyelvi szövegben

inkább nem használom.[34] Hansson szerint amikor másokról beszélünk, akkor inkább a "preferálja" terminust használjuk, és amikor magunkra utalunk vagy imperszonális módon nyilatkozunk, akkor inkább a "jobb" terminust alkalmazzuk [336].

A preferenciarendezés valamilyen halmaz elemeinek kívánatosság szerinti sorbarendezését jelenti. A preferenciarendezés reláció. Mint minden relációnak, úgy a preferenciarendezésnek is vannak argumentumai, relátumai. A preferencia jelenségéről beszélve könnyen átválthatunk, átváltunk reifikált tárgyalásmódba. Amikor a 'preferencia' terminust használjuk (például feltesszük a kérdést: "Neked mi a preferenciád?"), akkor észrevétlenül a preferenciarendezés eredményére irányul a figyelmünk. Ilyenkor a preferenciarelációt tárgyiasítjuk. A preferenciarendezés fogalma az összehasonlítandó alternatívák viszonyát vonatkozik, nem az összehasonlítás eredményére. Ez a beállítódás arra hasonlít, amikor a "vágy" terminusát használjuk. A vágyunk tárgya egy preferenciareláció reifikátuma. Amikor a vágyunkra gondolunk, már nem a valamivel való összehasonlításra, hanem csak a rendezés eredménye jár a fejünkben. Az angol nyelvben ritkán használt, de létező terminus a 'preferent', magyarban nincs megfelelője, ezért a **preferált alternatíva** kifejezést használjuk a preferenciarendezés előnyben részesített relátumára vonatkozóan.

A preferenciarendezést el kell választani a **preferenciális választástól** (preferential choice) is. A választásainkban valamilyen módon mindig megnyilvánulnak a preferenciáink, de nem minden preferenciánk jelent választást is egyben. Von Wright két példát is ad erre vonatkozóan. Az egyik példa szerint preferálhatjuk ugyan a különféle időjárás jelenségek közül (napos idő, viharos idő, kánikula stb.) egy egyiket a másikkal szemben, de nem választhatunk belőlük. De ugyanez igaz az egészségi állapotunkra is. Hiába tarjuk jobbnak az egészséget a betegségnél, nem áll módunkban választani ezen állapotok közül (legfeljebb csak áttételesen, az életmódválasztásokon keresztül).

A preferencia fogalmilag feltételezi a **diszkrimináció** létezését: ha m alternatívát preferáljuk n-vel szemben, akkor ezennel n-t diszkriminálunk is egyben. A diszkrimináció (vagy a diszkriminálás) a preferencia (vagy a preferálás) konverze. Természetesen a tárgyiasítás a másik irányban is működhet, vagyis a reifikálás a rendezési reláció másik relátumára a **diszkriminált alternatívára** is vonatkozhat. Ezzel azonban kevesebbet foglalkoznak a szakirodalomban. Amikor a hasznosság (utility) fogalmát elemzik a közgazdászok, akkor olykor felbukkan az a nézet, hogy a hasznosság dipólusos fogalom, tehát a szokásos, pozitív értelméhez képest létezik a másik véglete is: ez a negatív hasznosság, kellemetlenség, károsság (disutility), de ez ritkán történik meg. Pedig a döntéseink során figyelünk mindkét pólusra: a munkába álláskor a munkás hasznosságának számít a munkabére, a kellemetlensége pedig a munkavégzés okozta fáradság, kimerülés.

A preferenciarelációt, a szemantikailag ide kapcsolható további relációkat, valamint az ezek közti összefüggéseket a preferencialogika (BL) segítségével lehet definálni. Mielőtt bemutatnám, hogyan tehetjük meg ezt, előtte még el kell választani a preferenciák két típusát. Ez a megkülönböztetés azért nagyon fontos, mert ez alapján lehet majd a döntéseinkben, választásainkban, mérlegeléseinkben tetten érhető kétféle értékválasztási, preferenciarendezési mód, helyzet között különbséget tenni. Elsőként az intrinzikus prefe-

[34] Praktikus megfontolások miatt azonban a preferenciarelációt leíró formális logika azonosítására, az egyes tételek jelzésére a 'BL' szimbólumot fogom használni, ami a 'Betterness Logic' kifejezés rövidítése. Ezt a megoldást csak azért választottam, mert a logikák azonosítójaként használt jelek között a BL még szabad volt, míg a PL jelzést más logika már korábban "lefoglalta".

rencia fogalmát elemzem, és majd másodjára térek rá az extrinzikus preferenciák tárgyalására.

1.12.1 Intrinzikus preferencia

Von Wright saját esszéjében nem használta ki, de a társadalmi cselekvés modellezése során komoly jelentősége lesz annak a megkülönböztetésnek, amelynek segítségével von Wright definiálta a **intrinzikus** (intrinsic), illetve a **extrinzikus** (extrinsic) preferencia fogalmát.

> „. . . egy személy például preferálhatja a bordói vörös bort a rajnai fehérrel szemben azért, mert az orvosa azt ajánlotta neki, vagy azért, mert sajátmaga kitapasztalta, hogy az első bor jobbat tesz a gyomrának vagy általában az egészségének. Ebben az esetben az 'ez jobb, mint az' ítélet alapjául vagy indokául szolgál a preferenciának. Azt a preferenciát nevezem extrinzikusnak.
>
> Előfordulhat azonban az is, hogy ez a személy nem azért preferálja a bordói bort a rajnaival szemben, mert azt gondola, hogy az jobbat tesz neki, hanem pusztán csak azért, mert az elsőt jobban szereti. Ilyenkor az egyik bor szeretete nem indoka, értelme a preferenciájának. . . . " [794, 14]

Ez utóbbit nevezi von Wright intrinzikus preferencának. Mondhatnám, hogy ez a megkülönböztetés azért fontos, mert von Wright a preferenciákról szóló esszéjében csak az intrinzikus preferenciákkal foglalkozik. De nyilván nem véletlenül. Von Wright azt akarja megfogni, hogy mi a lényege azoknak a kedveléseknek, preferenciakinyilatkozásoknak, amelyek a legbensőnkből jönnek, amelyek a legkevésbé és a leglassabban változnak, amelyek kinyilvánításakor nem mérlegelünk, "nem gondolkozunk", amelyek az értékelköteleződéseink legmélyebb rétegét jelentik. Egy későbbi tanulmányában így ír erről.

> „Az *x* és *y* közötti preferencia extrinzikus, ha létezik egy olyan (nem cirkuláris) érv, amely megadja, *x*-t *miért* preferáltuk *y*-nal szemben. Minden más esetben a preferencia intrinzikus. Azt, hogy *x* intrinzikusan preferált *y*-nal szemben, úgy is kifejezhetjük, hogy azt mondjuk, *x* 'önmagában' vagy 'önmagáért' preferált *y*-nal szemben. Az intrinzikus preferencia kinyilvánítása, a legtöbbb esetben, kifejezi azt, amit *szeretünk*." [795]

Ezzel a megkülönböztetéssel persze sok esetben nagyon nehéz mit kezdeni. Az, hogy mikor számít valami extrinzikus vagy intrinzikus preferenciának, mikor jön valami belülről, mikor támaszkodunk külső észérvekre, sok esetben nagyon nehezen dönthető el, ha eldönthető egyáltalán. De egyelőre nem a nehézségekkel kell foglalkozni, hanem a jelenségek jobb megértésével, amihez nyilván hasznos eszköz az ideáltípusokra támaszkodó megközelítés alkalmazása.

Igazából minden preferenciát négy-argumentumú relációként kellene kifejeznünk. Az összehasonlítandó alternatívák mellett ugyanis mindig meg kellene adnunk, hogy az adott preferencia kinek az értékítéletét fejezi és mikor. Hiszen, tudjuk jól, ugyanazon dolgokra vonatkozó preferenciarendezés más lehet attól függően, hogy kiknek a preferenciájáról van szó, illetve egy személy preferenciája is változhat az időben. Ennek ellenére a preferenciával foglalkozó elméletek nem veszik figyelembe az ágencia és az idő dimenzióit.

Leginkább azért nem, mert a preferencia jelenségének megértéséhez érdemben nem (vagy csak keveset) ad hozzá e két szempont figyelembe vétele. Ezért a továbbiakban személytelenített és időtlenített (imperszonális és atemporális) módon fogom én is elemezni a preferencia fogalmát. A teljeskörű leírást biztosító mondat a következő lenne:

> *A* preferálja m-t n-nel szemben t időpontban

Ebből az állításból azonban – elhagyva a személyi és időbeli vonatkozásokat – csak annyit hagyunk meg a továbbiakban, hogy:

> valaki preferálja m-t n-nel szemben

Ez a megoldás egyszerűbbé teszi a jelenség tárgyalását. Von Wright is erre hivatkozva fogadta el ezt a szűkítést. Ezen túl azonban tett még egy olyan további leegyszerűsítést is, amit a preferenciával foglalkozó későbbi kutatók már nem fogadtak el. Von Wright ugyanis nemcsak az extrinzikus preferenciákra zárta ki a tárgyalásból, de lemondott az intrinzikus preferenciák változásainak vizsgálatáról is:

> „Csak azokat a preferenciákat tanulmányozzuk, amelyek egy személy intrinzikus preferenciái egy adott pillanatban. Ezzel a tárgyalásból kizárjuk mind a preferenciákat támogató érveket, mind a preferenciák változásának lehetőségét." [794, 23]

A fenti szűkítések miatt von Wright preferencialogikája statikus maradt, mivel nem tudta leírni a preferenciák változásait. Sem a mérlegelés nélküli, "belülről jövő" preferenciák változásait, sem a preferenciák mérlegelés, befolyásolás hatására bekövetkező, "külsődleges" változásait. A von wright-i preferencialogikát ezért ki kell egészíteni az extrinzikus preferenciákat leíró logikával, illetve bővítenünk kell olyan módon, hogy dinamikus preferencialogikához jussunk. Ezt azonban csak később végezzük el. Előtte definálnunk kell egy nyelvet, ami az intrinzikus preferenciák leírására alkalmas lehet.

(BL.0) $\mathscr{L}_{BL}(\mathscr{L}_{QL}, \text{PREF}, \mathbf{P})$, ahol

$\mathscr{L}_{QL}$	predikátumlogika
$\mathbf{P}$	preferenciareláció

Bár a preferencia fogalmának első formális elméletét Sören Halldén dolgozta ki még 1957-ben [319], a preferencia fogalmának tárgyalásakor nem az ő könyve lett a legtöbbet hivatkozott forrás. A mai napig sokkal inkább Georg Henrik von Wright 'The Logic of Preference' című esszéje számít hivatkozási alapnak e tárgyban [794]. Halldén is, von Wright is szintaktikai alapon definálta a preferencia fogalmát, ahogy a kortársak közül még sokan mások, aztán a hetvenes években már megjelentek azok az elméletek, amelyek – döntő módon – a lehetséges világ szemantika keretein belül igyekeztek modellt adni ennek a fogalomnak [333]. A következőkben előbb bemutatom, hogy milyen tulajdonságokkal próbálták meg definiálni a preferenciarelációt a legfontosabb szerzők, majd egyenként elemzem, hogy az egyes tulajdonságok elfogadásának milyen következményei lehetnek, végül áttérek arra, hogy a modális logika segítségével hogyan lehet teljesebbé tenni a preferencia fogalmának leírását.

1.12.2 Az intrinzikus preferencia tulajdonságai

Von Wright abból indult ki, hogy bár a preferenciarelációt sok minden "dolog" között alkalmazhatjuk, elég csak a világállapotok közti preferenciarelációt vizsgálni, mert minden más eset visszavezethető erre.[35] Ezután megadott öt axiómát, amivel definiálni lehet a preferencia fogalmát.

Von Wright első lépésben két relációalgebrai tulajdonság (az aszimetrikusság és a tranzitivitás) tételezésével a preferenciarelációt a SZIGORÚ RENDEZÉSI RELÁCIÓ egyik típusának minősítette. Ezután pedig megállapított három további ekvivalenciát, amelyek segítségével transzformációs (konjunkciós, disztribúciós, amplifikációs) műveletek végezhetők el. Az öt kritérium formulája a következő:

(BL.1) $\text{STRICT_PREFERENCE}(m,n) \equiv \mathbf{P}(m,n) \equiv \mathbf{P} \equiv \mathbf{R}^{>} \equiv m > n \equiv >$

szigorú preferenciarendezés (HUN) – erős preferenciarendezés (HUN) – preferenciarendezés (HUN) – preferencia (HUN) – strict preference (ENG) – strong preference (ENG)

(BL.2) $\text{ASYMMETRIC}(\mathbf{P}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n(\mathbf{P}(m,n) \rightarrow \neg\mathbf{P}(n,m))$

P reláció **aszimmetrikus**

(BL.3) $\text{TRANSITIVE}(\mathbf{P}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n \forall o(\mathbf{P}(m,n) \wedge \mathbf{P}(n,o) \rightarrow \mathbf{P}(m,o))$

P reláció **tranzitív**

(BL.4) $\forall m \forall n(\mathbf{P}(m,n) \leftrightarrow \mathbf{P}(m \wedge \neg n, \neg m \wedge n))$

P relációra érvényes a **konjunktív expanzió** elve

(BL.5) $\forall m \forall n \forall o \forall p(\mathbf{P}(m \vee n, o \vee p) \leftrightarrow (\mathbf{P}(m \wedge \neg o \wedge \neg p, \neg m \wedge \neg n \wedge o) \wedge \mathbf{P}(m \wedge \neg o \wedge \neg p, \neg m \wedge \neg n \wedge p) \wedge \mathbf{P}(n \wedge \neg o \wedge \neg p, \neg m \wedge \neg n \wedge o) \wedge \mathbf{P}(n \wedge \neg o \wedge \neg p, \neg m \wedge \neg n \wedge p)))$

P relációra érvényes a **disztribúció** elve

(BL.6) $\forall m \forall n \forall o(\mathbf{P}(m,n) \leftrightarrow (\mathbf{P}(m \wedge o, n \wedge o) \wedge \mathbf{P}(m \wedge \neg o, n \wedge \neg o)))$

P relációra érvényes az **amplifikáció** elve

A preferenciareláció fenti tulajdonságaira támaszkodva von Wright további két fontos diadikus preferenciarelációt határozott meg, ami azonban – tágabb kontextusba helyezve a dolgokat – három relációval hozható kapcsolatba. Az egyik az **indifferencia** reláció, amit von Wright azzal a minőséggel definál, hogy két állapot egyike sem preferált a másikkal szemben [794, 52]. Ezen a ponton vigyázni kell, mert könnyen zavart okozhat az a tény, hogy ugyanezt a relációt más szerzők egészen más szemantikával értelmezik, és hol összehasonlíthatatlansági relációnak [20], hol 'ugyanolyan, mint' relációnak nevezik [154]. Von Wright is tisztában van azzal, hogy kétfajta relációról van szó, ezért a most tárgyalt relációnak kicsit később a **gyenge indifferencia** nevet adja [794]. A későbbi félreértések elkerülése végett von Wright névadási javaslatát csak a jelzős szerkezetben viszem tovább. Mindez persze csak terminológiai kérdés, jelentését tekintve az új relációt egyféleképpen vezethetjük be a formális modellünkbe.

[35] Von Wright a preferenciák lehetséges irányultságaiként az állapotok mellett a dolgokat, eszközöket, illetve a cselekvéseket említi még. Abban nem nagyon van vita a preferenciával foglalkozó szerzők között, hogy elég az állapotokkal foglalkozni, mert minden másra irányuló preferencia kifejezhető ezen az alapon. Ennek a döntésnek komoly előnye, hogy az állapotokra vonatkozóan könnyű kijelentéseket tenni, és ezekre azonnal alkalmazható a propozícionális logika teljes eszköztára.

(BL.7) $\text{INCOMPARABILITY}(m,n) \equiv \mathbf{R}^{\#}(m,n) \equiv \mathbf{R}^{\#} \equiv \mathbf{J}(m,n) \equiv \mathbf{J} \stackrel{\text{def}}{=} \forall m \forall n (\neg \mathbf{P}(m,n) \wedge \neg \mathbf{P}(n,m))$

összehasonlíthatatlanság (HUN) – gyenge közömbösség (HUN) – gyenge indifferencia (HUN) – incomparability (ENG) – weak indifference (ENG) – ugyanolyan értékes, mint (HUN) – same value as (ENG)

A közömbösségi reláció terminusra vonatkozóan pedig, amit von Wright az erős indifferenciának nevezett el, meghagyom a sokkal inkább elterjedt – jelző nélküli – indifferencia vagy közömbösség megnevezést. Ebben az esetben tagadás nélkül állítjuk az egyidejűleg két irányú preferálás tényét.

(BL.8) $\text{INDIFFERENCE}(m,n) \equiv \mathbf{I}(m,n) \equiv \mathbf{I} \equiv \mathbf{R}^{\equiv} \equiv \equiv\!\!\!| (m,n) \equiv m \equiv\!\!\!| n \equiv \equiv\!\!\!| \stackrel{\text{def}}{=} \forall m \forall n (\mathbf{P}(m,n) \wedge \mathbf{P}(n,m))$

indifferencia (HUN) – erős indifferencia (HUN) – erős közömbösség (HUN) – strong indifference (ENG) – indifference (ENG) – értékegyenlőség (HUN) – value equality (ENG) – ugyanolyan, mint (HUN) – ugyanolyan jó, mint (HUN) – equally good as (ENG) – same as (ENG)

Utóbbi relációra vonatkozóan von Wright bevezetett még egy újabb megnevezést is, és **értékegyenlőség** relációnak nevezte el, de ezt a terminust nem fogom használni, csak a történeti érdekesség miatt említem meg. Von Wright a közömbösségi relációt nem a fent megadott képlet alapján definiálta, hanem ahelyett ugyanolyan axiómákat rögzített, mint amilyeneket a preferenciareláció bevezetésekor kötött ki.

(BL.9) $\text{SYMMETRIC}(\mathbf{I}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n (\mathbf{I}(m,n) \rightarrow \mathbf{I}(n,m))$

I reláció **szimmetrikus**

(BL.10) $\text{TRANSITIVE}(\mathbf{I}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n \forall o (\mathbf{I}(m,n) \wedge \mathbf{I}(n,o) \rightarrow \mathbf{I}(m,o))$

I reláció **tranzitív**

(BL.11) $\forall m \forall n (\mathbf{I}(m,n) \leftrightarrow \mathbf{I}(m \wedge \neg n, \neg m \wedge n))$

I relációra érvényes a **konjunktív expanzió** elve

(BL.12) $\forall m \forall n \forall o \forall p (\mathbf{I}(m \vee n, o \vee p) \leftrightarrow \mathbf{I}(m \wedge \neg o \wedge \neg p, \neg m \wedge \neg n \wedge o) \wedge \mathbf{I}(m \wedge \neg o \wedge \neg p, \neg m \wedge \neg n \wedge p) \wedge \mathbf{I}(n \wedge \neg o \wedge \neg p, \neg m \wedge \neg n \wedge o) \wedge \mathbf{I}(n \wedge \neg o \wedge \neg p, \neg m \wedge \neg n \wedge p))$

I relációra érvényes a **disztribúció** elve

(BL.13) $\forall m \forall n \forall o (\mathbf{I}(m,n) \leftrightarrow \mathbf{I}(m \wedge o, n \wedge o) \wedge \mathbf{I}(m \wedge \neg o, n \wedge \neg o))$

I relációra érvényes az **amplifikáció** elve

A fenti formulák együttes igazsága – von Wright intuíciója szerint – szükséges és elégséges feltétel az értékegyenlőségi reláció meghatározásához. A preferenciareláció axiómáival összevetve látható, hogy az ötből tétel közül csak az elsőben különbözik a két reláció: a preferencia aszimmetrikus, az értékegyenlőség szimmetrikus. Ez azt is jelenti egyben, hogy a von Wright értelmezésben vett preferenciareláció és értékegyenlőség reláció segítségével előállíthatunk egy olyan preferenciarendezési fogalmat is, amelyben a preferencia fogalma nem a 'szigorú értelemben vett jobb', hanem a 'legalább olyan jó' jelentés mellett értelmezhető.

(BL.14) $\text{WEAK_PREFERENCE}(m,n) \equiv \mathbf{R}(m,n) \equiv \mathbf{R} \equiv \mathbf{R}^{\geq}(m,n) \equiv \mathbf{R}^{\geq} \equiv m \geq n \equiv \geq \stackrel{\text{def}}{=} \forall m \forall n (\mathbf{P}(m,n) \vee \mathbf{I}(m,n))$

gyenge preferencia (HUN) – preferencia (HUN) – legalább olyan jó, mint (HUN) – jobb vagy ugyanolyan (HUN) – weak preference (ENG) – preference (ENG)

A legalább olyan jó értelmezésű $\mathbf{R}$ gyenge preferenciarelációra ugyanazokat a tulajdonságokat rá lehet vetíteni, mint amiket a szigorú preferenciarendezés, illetve közömbösségi relációval szemben von Wright kikötött. A három preferenciarelációt, a szigorú és gyenge preferenciarendezést, valamint a közömbösségi relációt (és olykor még az összehasonlíthatatlansági relációt is) együttesen ($\mathbf{P}$, $\mathbf{I}$, $\mathbf{J}$) **preferenciastruktúrának** is szokták nevezni, és bizonyos tételeket a struktúra egészére vonatkozóan szoktak kimondani [618].

A preferenciareláció eddigi axiómái alapján von Wright már meghatározhatónak gondolta két fontos monadikus (vagy unáris) preferenciareláció, a **jó** (GOOD) és a **rossz** (BAD) fogalmát is. Mivel von Wright a preferenciát a világállapotokra vonatkoztatva definiálta, ezért ebben az elméletben a jó és rossz kategóriája is (világ)állapotokra vonatkozóan érvényes. Szerinte a jó (állapot) így tehát az az állapot, amely feltétel nélkül preferált saját ellentétével szemben, míg a rossz (állapot) az, amelynek az ellentéte feltétel nélkül preferált magával szemben.[36] A meghatározások megfelelő formulái [794].

(BL.15) $\quad \text{GOOD}(m) \equiv \mathbf{G}'(m) \equiv \mathbf{G}' \stackrel{\text{def}}{=} \mathbf{P}(m, \neg m)$

(BL.16) $\quad \text{BAD}(m) \equiv \mathbf{B}'(m) \equiv \mathbf{B}' \stackrel{\text{def}}{=} \mathbf{P}(\neg m, m)$

Voltak, akik elfogadták ezt a meghatározást, és voltak, akik elutasították (erre később még visszatérek). Ez persze általában is igaz. Halldén és Von Wright preferencialogikájára sokan reagáltak, bizonyos pontokon támogató, máshol kritikus módon. Voltak szerzők nem osztották Von Wright intuícióit, és részben vagy teljesen más axiómákat tartottak meghatározónak. Roderick M. Chisholm és Ernest Sosa egy egészen eltérő formalizmust dolgozott ki von Wright elméletéhez képest [154]. Az aszimmetrikusság tulajdonságait fenntartották ők is, de ezután von Wright megoldásához képest teljesen más axiómákat javasoltak. Másodiknak a tranzitivitás duális fogalmát, a negatív vették fel az axiómák közé.[37]

(BL.17) $\quad \text{CONTRA-TRANSITIVE}(\mathbf{P}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n \forall o((\neg\mathbf{P}(m,n) \wedge \neg\mathbf{P}(n,o)) \rightarrow \neg\mathbf{P}(m,o))$

negatívan tranzitív (HUN) – negatively transitive (ENG)

Ezután definiáltak még további három axiómát, de mielőtt bemutatnám ezeket, előtte érdemesebbnek tűnik felsorolni azt, hogy milyen további monadikus és diadikus preferenciarelációkat vezettek még be Chisholmék az első két axiómájukra támaszkodva. Azért érdemes előbb megismerni ezeket, mert az ismeretükben könnyebb lesz majd értelmezni a további axiómákat. Lássuk tehát az öt új Chisholm-Sosa-féle preferenciarelációt.[38]

(BL.18) $\quad \text{SAME_VALUE_AS}(\mathbf{S}(m,n)) \equiv \mathbf{S}(m,n) \stackrel{\text{def}}{=} \forall m \forall n(\neg\mathbf{P}(m,n) \wedge \neg\mathbf{P}(n,m))$

ugyanolyan (HUN) – ugyanolyan értékes, mint (HUN) – same value as (ENG)

(BL.19) $\quad \text{INDIFFERENT}(\mathbf{I}''(m)) \equiv \mathbf{I}''(m) \stackrel{\text{def}}{=} \forall m(\neg\mathbf{P}(m,\neg m) \wedge \neg\mathbf{P}(\neg m, m))$

indifferens (HUN) – közömbös (HUN) – indifferent (ENG)

[36] Von Wright egy másik tanulmányában megjegyezte, hogy ezt a meghatározást először A.P. Brogan dolgozta ki még 1919-ben [795].

[37] Chisholmék nem használtak semmilyen megnevezést a szóbanforgó tulajdonságra. A formulájukat azonban ekvivalens alakra lehet hozni a **negatívan tranzitív** relációtulajdonság képletével (↬35.o.), ahogy azt M. Roubens és Ph. Vincke megmutatta [618, 3]. Ezért használom itt ezt a megnevezést.

[38] Chisholm és Sosa az öt preferenciareláció megnevezésébe mindig beleírta, beleértette az 'intrinzikusan' jelzőt, amit én itt – kényelmi okok miatt – elhagyok.

(BL.20) $\text{NEUTRAL}(\mathbf{N}''(m)) \equiv \mathbf{N}''(m) \stackrel{\text{def}}{=} \forall m \exists n(\mathbf{I}(m) \wedge \mathbf{S}(m,n))$

neutrális (HUN) – semleges (HUN) – neutral (ENG)

(BL.21) $\text{GOOD}(\mathbf{G}''(m)) \equiv \mathbf{G}''(m) \stackrel{\text{def}}{=} \forall m \exists n(\mathbf{I}(m) \wedge \mathbf{P}(m,n))$

jó (HUN) – good (ENG)

(BL.22) $\text{BAD}(\mathbf{B}''(m)) \equiv \mathbf{B}''(m) \stackrel{\text{def}}{=} \forall m \exists n(\mathbf{I}(m) \wedge \mathbf{P}(n,m))$

rossz (HUN) – bad (ENG)

A továbblépés előtt érdemes itt pár megjegyzést tenni. Az ugyanolyan, mint reláció megegyezik a korábban bemutatott von wrighti gyenge indifferencia relációval (↬92.o.). A fenti öt relációból csak az első diadikus, a többi monadikus, vagyis csak mindegyiknek egy argumentuma van. Arra is figyelni kell nagyon, hogy amíg Chisholm és Sosa itt mindegyikhez egy-egy pontos és egyértelmű jelentés rendel, addig a köznyelv ezeket a terminusokat sokféle értelemben használja.

Az $\mathbf{S}(m,n)$ ugyanolyan és az $\mathbf{I}(m)$ indifferens reláció csak abban különbözik egymástól hogy az első két-, a második egyargumentumú. Két állapot (alternatíva) ugyanolyan, ha egyik sem jobb, mint a másik, egy állapot pedig önmagában indifferens, ha nem jobb a tagadásánál, és a tagadása sem jobb az állapotnál. Az $\mathbf{S}(m,n)$ és az $\mathbf{I}(m)$ segítségével lehet definiálni a neutrális állapotot. E fogalom alatt Chisholm és Sosa egy olyan m indifferens állapotot értenek ($\mathbf{I}(m)$), amely legalább egy másik n állapothoz képest ugyanolyan értékű ($\mathbf{S}(m,n)$). Hasonló logika mentén a $\mathbf{G}(m)$ jó az az $\mathbf{I}(m)$ indifferens állapot, amely legalább egy másik állapothoz képest jobb ($\mathbf{P}(m,n)$), a $\mathbf{B}(m)$ rossz állapot pedig az az $\mathbf{I}(m)$ indifferens állapot, amelyhez képest legalább egy másik állapot jobb ($\mathbf{P}(n,m)$).[39] Ezek alapján könnyebben meg lehet majd érteni, milyen minőségeket akartak Chisholmék az alábbi három axiómájukkal megragadni.

(BL.23) $\forall m \forall n((\neg\mathbf{P}(m,\neg m) \wedge \neg\mathbf{P}(\neg m,m) \wedge \neg\mathbf{P}(n,\neg n) \wedge \neg\mathbf{P}(\neg n,n)) \rightarrow (\neg\mathbf{P}(m,n) \wedge \neg\mathbf{P}(n,m)))$

(BL.24) $\forall m \forall n((\neg\mathbf{P}(n,\neg n) \wedge \neg\mathbf{P}(\neg n,n) \rightarrow \mathbf{P}(m,n)) \rightarrow \mathbf{P}(m,\neg m))$

(BL.25) $\forall m \forall n((\neg\mathbf{P}(n,\neg n) \wedge \neg\mathbf{P}(\neg n,n)) \rightarrow \mathbf{P}(n,\neg m) \rightarrow \mathbf{P}(m,\neg m))$

Ha fenti három formula közül az elsőbe behelyettesítjük az $\mathbf{I}(m)$ és $\mathbf{S}(m,n)$ relációk képleteit, akkor egy jóval egyszerűbb, áttekinthetőbb és így könnyebben értelemezhető formulához jutunk.

(BL.26) $\forall m \forall n((\mathbf{I}(m) \wedge \mathbf{I}(n)) \rightarrow \mathbf{S}(m,n))$

A képletet úgy értelmezhetjük, hogy ha két állapot indifferens önmagában, akkor egymáshoz képest ugyanolyan értékűek lesznek. A második és harmadik axióma képletébe is behelyettesítve az $\mathbf{I}(m)$ reláció formuláját, újra csak egyszerűbb eredményt kapunk.[40]

[39] Újra hangsúlyozom, hogy a köznyelvben ezeket a fogalmakat sokkal pongyolább módon használjuk, nem teszünk ilyen éles és pontos különbséget e terminusok között.

[40] Egy másik tanulmányban Chisholm és Sosa ezt a két axiómát egy közös diszjunktív formulába összevonva fejezte ki [153].

(BL.27) $\forall m \forall n((\mathbf{I}(n) \to \mathbf{P}(m,n)) \to \mathbf{P}(m,\neg m))$

(BL.28) $\forall m \forall n((\mathbf{I}(m) \to \mathbf{P}(n,\neg m)) \to \mathbf{P}(m,\neg m))$

Az első formulát úgy interpretálhatjuk, hogy ha van egy n indifferens állapot, és ebből az következik, hogy egy másik m állapot preferált n-hez képest, akkor az m állapot preferált a saját $\neg m$ tagadásához képest. A második formula szerint ha az n indifferens állapot létezéséből következik, hogy ez az n állapot preferált egy másik m állapot tagadásához, tehát $\neg m$-hez képest, akkor az m állapot preferált a saját $\neg m$ tagadásához képest.

Láthattuk tehát, hogy Chisholm és Sosa öt axiómája – egy kivételével – teljesen eltér von Wright elméletétől, de hogy ennek mi az oka, a vitában kinek milyen érvei voltak, erre a kérdésre nem egyszerre és nem is itt válaszolok, hanem az elméletek által kezelt egyes problémákra különö fejezeteket szentelve fogom összehasonlítani a különböző nézőpontokat.

Itt érdemes röviden áttekinteni, hogy Bengt Hansson hogyan kezelte a preferenciareláció fogalmát [329]. Hansson először megmutatta, hogy mind a von Wright, mind a Chisholm-Sosa szerzőpár által lefektetett elmélet túl erős, és a preferenciareláció mindennapi használatával kapcsolatos intuíciónknak több ponton ellentmondó következményei vannak. Ezért Bengt Hansson gyengébb axiómákat vezetett be, bár tanulmánya végén ezekre vonatkozóan is érvényesnek tartotta azt az elvárását, hogy a preferenciák logikájának kidolgozásához talán erősebb logikai nyelvre lehet szükség. Bengt Hansson a preferenciarelációt a 'jobb vagy ugyanolyan jó' értelmeben fogta fel, tehát az $\mathbf{R}(m,n)$ gyenge preferenciarendezést elemezte, és erre a relációra kikötött két rendezési, illetve két disztribúciós axiómát. Az első bengt hanssoni tétel a tranzitivitás, a második a **totalitás** tulajdonságát rögzítette. A tranzitivitási feltételt már felvettük a listába, a totalitás formulája így néz ki:

(BL.29) $\forall m \forall n(\mathbf{R}(m,n) \vee \mathbf{R}(n,m))$

R reláció **erősen összefüggő** (totális)

Hansson totalitási kritériuma megfelel a relációk tárgyalása során bemutatott erősen összefüggő tulajdonságnak (↬33.o.), a tranzitivitás és totalitás alapján pedig azt mondhatjuk, hogy a bengt hanssoni **R** reláció totális (összefüggő) előrendezés (↬39.o.). A következő lépésben Hansson először megmutatta, hogy milyen továbbbi tételek vezethetők le a két axiómájából, majd – miután több, túl erős, elfogadhatatlan következményekkel járó változatot kipróbált és elvetett – javasolt két gyengített disztribúciós axiómát.

(BL.30) $\forall m \forall n(\mathbf{R}(m,n) \wedge \mathbf{R}(m,o)] \to \mathbf{R}(m,n \vee o))$

(BL.31) $\forall m \forall n(\mathbf{R}(m,o) \wedge \mathbf{R}(n,o)] \to \mathbf{R}(m \vee n,o))$

A két tétel egy balról és egy jobbról érvényes disztribúciós kritériumot fogalmaz meg. Az első formula azt írja le, hogy ha egy m állapot preferáltabb mint két másik (n vagy o),

akkor az m állapot a két másik (n és o) állapot diszjunkciójánál is preferáltabb. A második tétel ennek a fordítottját követeli meg, vagyis azt, hogy ha egy o állapotnál két másik (m vagy n) is preferáltabb, akkor azok konjunkciója is preferáltabb az o állapothoz képest. Beng Hansson hosszan érvelt amellett, hogy végül miért ezeket a tételeket fogadta el, erre én is visszatérek a disztribúciós elv tárgyalásakor.

———————-

Nicholas Rescher 1966-ban [602]

———————-

A preferencialogika fejlődéstörténetéhez hozzátartozik még, hogy von Wright egy évtizeddel később visszatért a kérdés tárgyalására. Az 1963-as könyvével kapcsolatban megjelent kritikák, alternatív formalizálási kísérletek után újra elővette a preferencia problémakörét, és 1972-ben részben újraformálta elméletét [795]. Saját értékelése szerint nem is annyira a kritikákat fogadta meg, mint saját maga törekedett egyfelől a kérdések egyértelműsítésére, másfelől a válaszok egyszerűsítésére.

Tanulmányában von Wright először felsorolta azokat a kérdéseket, amelyeket minél pontosabban tisztázni kell egy preferencialogika kidolgozásakor, majd megállapította, hogy nem igazán arra lát esélyt, hogy ki lehessen dolgozni egy széles körben elfogadott elméletet. Inkább csak azt az állapotot tartotta elérhetőnek, hogy a különböző intuíciók mentén bár szabatos, de párhuzamos értelmezések alakulhassanak ki. Von Wright újra megpróbálkozott a preferencia fogalmának formalizálásával.

Von Wright az új elmélet központjába is a $\mathbf{P}$ szigorú preferencirendezést tette. Első axiómaként felvette az aszimetrikusságot, másodikként pedig egy új elvet rögzített, amit az (érték)összehasonlíthatóság elvének nevezett el. Ez az összehasonlítható relációtulajdonsággal egyezik meg (↬34.o.).

(BL.32) $\text{COMPARABLE}(\mathbf{P}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n \forall o (\mathbf{P}(m,n) \rightarrow (\mathbf{P}(m,o) \vee \mathbf{P}(o,n)))$

$\mathbf{P}(m,n)$ preferenciareláció **összehasonlítható**

Az összehasonlíthatóság elvének felvétele egyben azt is jelentette, hogy von Wright nem írta elő a – sokat vitatott – tranzitivitás tulajdonságának fennállását, ami viszont a két axiómából már levezethető vált. Ez az alábbi módon látható be:

1) $\mathbf{P}(n,o) \rightarrow \neg\mathbf{P}(o,n)$
2) $(\mathbf{P}(m,n) \wedge \mathbf{P}(n,o)) \rightarrow (\mathbf{P}(m,n) \wedge \neg\mathbf{P}(o,n))$
3) $(\mathbf{P}(m,n) \wedge \mathbf{P}(n,o)) \rightarrow ((\mathbf{P}(o,n) \vee \mathbf{P}(m,o)) \wedge \neg\mathbf{P}(o,n))$
4) $(\mathbf{P}(m,n) \wedge \mathbf{P}(n,o)) \rightarrow (\mathbf{P}(m,o) \wedge \neg\mathbf{P}(o,n))$
5) $(\mathbf{P}(m,n) \wedge \mathbf{P}(n,o)) \rightarrow \mathbf{P}(m,o)$

Első lépésben felvesszük a $\mathbf{P}(n,o)$ aszimmetrikusságát (1. axióma). Második lépésben a kondicionális mindkét oldalát kibővítjük ugyanazzal a $\mathbf{P}(m,n)$ taggal (PL: konjunkció bevezetése). Harmadik lépésben a kondicionális jobb oldalán levő $\mathbf{P}(m,n)$ helyébe behelyettesítjük az összehasonlíthatóság képletének jobb oldalát (2. axióma). Negyedik lépésben elhagyjuk a kondicionális következményéből a jobb oldali $\neg\mathbf{P}(o,n)$ összetevőt (PL: konjunkció eliminálása). Ezzel a tranzitivitás képletéhez jutunk.

Megállapítható tehát, hogy a $\mathbf{P}(m,n)$ szigorú preferenciarendezés aszimmetrikus, irreflexív és tranzitív. Ezután von Wright bevezette a $\mathbf{J}(m,n)$ (gyenge) indifferenciarelációt (valójában az összehasonlíthatatlansági relációt (↬92.o.), és az eddig axiómák és tételek

alapján levezette, hogy $\mathbf{J}(m,n)$ reflexív, tranzitív és szimmetrikus reláció. A preferenciarendezés és a gyenge indifferencia reláció segítségével pedig kimondta az (érték)-ekvivalencia elvét, amely az alábbi háromféle állapot lehetőségét tételezi:

(BL.33) $\mathbf{J}(m,o) \wedge \mathbf{J}(o,n) \rightarrow \mathbf{J}(m,n)$
$\mathbf{P}(m,n) \wedge \mathbf{J}(m, o \rightarrow \mathbf{P}(o,n))$
$\mathbf{P}(m,n) \wedge \mathbf{J}(n,o) \rightarrow \mathbf{P}(m,o)$

P-re és **J**-re érvényes az **(érték)-ekvivalencia elve** (HUN) – principle of (value)-equivalence (ENG)

Az ekvivalencia elve az axiómákból levezethető következmény, de önmaga nem axióma. Von Wright az aszimmetricitáson és az összehasonlíthatóságon túl több axiómát már nem definiált, tehát az 1963-as modelljéhez képest jóval gyengébb axiómarendszert javasolt. Úgy tűnik, hogy leginkább Bengt Hansson kritikája hatott leginkább von Wrightre, mert az átalakított modellje sokkal jobban hasonlít Hansson javaslatához, mint akár a saját korábbi, akár Chisholmék elképzeléseihez.

Bár von Wright sosem tartotta önálló axiómának a kontrapozíció elvét (↬44.o.), de azért megemlítette azt az új tanulmányában. Ezt az elvet Halldén javasolta felvenni a preferencia modelljébe [319]. Von Wright [795] nem építette be ugyan az új axiómarendszerébe, de az a tény, hogy alkalmazta, arra utal, hogy úgy gondolta, a preferenciára vonatkozóan érvényesnek lehet tartani. Ezt az elvet Chisholm és Sosa támadta (akkor elsősorban mint Halldén javaslatát), de hogy miért és hogyan kritizálták a kontrapozíció elvét, azt később mutatom be (↬103.o.). Egyelőre álljon itt magának az elvnek a formulája.

(BL.34) $\forall m \forall n(\mathbf{P}(m,n) \leftrightarrow ((\mathbf{P}(\neg n, \neg m))))$

P preferenciarelációra érvényes a **kontrapozíció elve**

Az 1972-es tanulmányában von Wright nagy teret szentelt annak, hogy megmutassa, miként lehet leírni a lehetséges világ szemantika alapján a ceteris paribus preferencia fogalmát. Erre a kérdésre külön fejezetben térek ki (↬120.o.), ezért itt nm tárgyalom. A cikke végén nekifutott újra annak a feladatnak, hogy hogyan lehet az olyan monadikus fogalmat, mint a 'jó' vagy a 'rossz' kategóriája definiálni a 'jobb' fogalom segítségével. Von Wright az előbbi kettőt érték-abszolút, az utóbbit összehasonlító-érték fogalomnak tartotta, és másként definálta őket az 1963-as megoldásához képest [795].

Von Wright felvetette, hogy a jó és a rossz fogalmának meghatározásához G.E. Moore üres világ (empty world) fogalmára lehetne támaszkodni, csak ehhez tisztázni kéne, mit is jelent pontosan ez a fogalom. Von Wright javaslata az volt, hogy az üres világ fogalmát az **ellentmondással** (↬18.o.) lehetne definiálni, mert az logikailag lehetetlen világ, ahol semmi sem igaz és semmi sem hamis. Ezt a világot a lehetséges világok sorában a nulladiknak tekintve (0-világ) és w_0-val jelölve olyan fogalomhoz jutunk von Wright szerint, amely alapján mondhatjuk, hogy jó az a w_x világ, amelyre igaz, hogy $\mathbf{P}(x,0)$, rossz az a w_y világ, amelyre igaz, hogy $\mathbf{P}(0,y)$. Mindezt a totális állapotokra írhatnánk így. Egy tetszőleges egyedi m állapotra vonatkoztatva a jót és a rosszat a következőképpen határozhatnánk meg: $\mathbf{P}_C(m,0)$ a jó, $\mathbf{P}_C(0,m)$ pedig a rossz definíciója. A jó és rossz fogalmára vonatkozó új von wrighti javaslatokat a monadikus relációkról szóló fejezetben értékelem (↬119.o.).

Lennart Åqvist, Chisholm-Sosa Logics of Intrinsic Betterness and Value [25]

Preferenciastruktúra

Az eddigiekből már kiderülhetett, de tézisszerűen még nem mondtam ki, hogy a preferencia jelenségét nem egyetlen, hanem inkább több preferenciareláció segítségével írhatjuk le igazán. Első nekifutásra tán még tűnhetne úgy, hogy a preferenciarendezés fogalma jelenti a jelenség lényegét, de ez nem így van. A különböző preferencialogikái már abba is különböznek, hogy a szigorú vagy a gyenge preferenciarendezést teszik-e az elmélet középpontjába. De kiderült az is, hogy az indifferencia vagy az összehasonlíthatatlnasági relációknak is olykor komoly szerepe van. Az elméletek ezért mozdultak el abba az irányba, hogy ne egyik vagy másik preferenciarelációról beszéljünk, hanem preferenciastruktúráról. Ez annyit jelent, hogy egyfelől általában vett **preferenciarelációról** beszélünk, és megkülönböztetjük egymástól ennek típusait (**R** gyenge és **P** erős rendezés, **I** indifferencia, **J** összehasonlíthatatlanság)[41], másfelől ezeket a preferenciareláció-típusokat, valamint a köztük értelmezett műveleteket egy struktúrába összefogva beszélünk **preferenciastruktúráról**. Kiindulásként vegyünk egy A (tartó)halmazon értelmezett $(\mathbf{P}, \mathbf{I}, \mathbf{J})$ preferenciastruktúrát, és nézzük meg, milyen tulajdonságokat definiálhatunk ezen a struktúrán [618].

(BL.35) $\textsc{irreflexive}(\mathbf{P}) \wedge \textsc{reflexive}(\mathbf{I}) \wedge \textsc{irreflexive}(\mathbf{J})$

P és **J** irreflexív, **I** reflexív

(BL.36) $\textsc{asymmetrical}(\mathbf{P}) \wedge \textsc{symmetrical}(\mathbf{I}) \wedge \textsc{symmetrical}(\mathbf{J})$

P aszimmetrikus, **I** és **J** szimmetrikus

(BL.37) $\mathbf{P} \cap \mathbf{I} = \varnothing \wedge \mathbf{P} \cap \mathbf{J} = \varnothing \wedge \mathbf{I} \cap \mathbf{J} = \varnothing$

P, **I** és **J** kölcsönösen kizárják egymást

(BL.38) $\mathbf{P} \cup \mathbf{P}^{-1} \cup \mathbf{I} \cup \mathbf{J} = A \times A$

P és a konverze ($\mathbf{P}^{-1}$), **I** és **J** együttesen kiadják a tartóhalmazon vett egész relációt

Az első két tétel azokat a relációtulajdonságokat rendeli a relációkhoz, amelyeket már korábban megismertünk. Fontos összefüggést fejez ki a harmadik tétel, miszerint a három alapreláció diszjunktnak tekinthető egymáshoz képest. Végül a negyedik tételből kiderül az is, hogy ha a három alaprelációhoz hozzávesszük még a szigorú rendezés konverzét, akkor a négy reláció teljesen lefedi az tartóhalmazon értelmezhető teljes relációt. Mivel bizonyítható az is, hogy az **P** rendezés konverze is diszjunkt az indifferencia és összehasonlíthatatlansági relációkkal, így kijelenthető, hogy a a $(\mathbf{P}, \mathbf{P}^{-1}, \mathbf{I}, \mathbf{J})$ preferenciastruktúrában a négy reláció teljes partíciót alkot.[42] Sem a hármas, sem a négyes relációstruktúrában nem szerepel az **R** gyenge rendezési reláció. Ezt azonban könnyen pótolhatjuk az alábbi összefüggés alapján.

(BL.39) $\mathbf{R} = \mathbf{P} \cup \mathbf{I}$

[41] Tehát – figyelem – ezeket mind preferenciarelációnak nevezzük.

[42] A teljes partíció megállapításához szükséges, hogy a komponensek diszjunktak legyenek egymáshoz képest. A négyes preferenciastruktúra negyedik eleme itt a $\mathbf{P}^{-1}$ konverzreláció. Általában véve nem igaz, hogy egy reláció konverze diszjunkt lenne az alaprelációjával, de mivel itt **P** aszimmetrikus, ezért ebben az esetben fenáll ez a feltétel.

A új rendezési reláció segítségével további összefüggéseket is megállapíthatunk a hármas struktúra elemei és a gyenge rendezés között. Ehhez használnunk kell a konverzreláció (↬43.o.), a duálreláció (↬44.o.) és a komplementer reláció (↬44.o.), valamint a teljesség vagy erős összefüggőség (↬33.o.) és az üres reláció (↬31.o.) fogalmát. Először a hármas struktúra elemeit írhajuk fel a gyenge rendezési reláció segítségével, majd a gyengerendezés teljességére adhatunk meg egy új feltételt.

(BL.40) $(\mathbf{P},\mathbf{I},\mathbf{J}) = (\mathbf{R}\cap\mathbf{R}^{-d},\mathbf{R}\cap\mathbf{R}^{-1},\overline{\mathbf{R}}\cap\mathbf{R}^{-d})$

az (**I**, **J**, **R**) relációstruktúra elemei az **R** segítségével kifejezve

(BL.41) $\textsc{strongly_connected}(\mathbf{R}) \leftrightarrow \mathbf{J} = \underset{\cdot}{\mathbf{E}}$

az **R** gyenge preferenciarendezés erősen összefüggő, ha **J** üres

---

A Ferrers-tulajdonság és a szemirendezés fogalmait is meg kellene itt hivatkozni? vagy inkább a tranzitivitás fejezetnél????????

Preferencia és reflexivitás

A preferenciaelméleteken belül nem igazán szokás reflektálni a reflexivitás kérdésére. Ez talán magyarázható azzal, hogy a reflexivitás részben összefügg olyan más tulajdonságkörökkel (például a szimmetricitással, az összekapcsoltsággal), amelyeket viszont mindig sokkal nagyobb figyelemmel kísértek. Az irreflexivitás több más relációtulajdonságból is következik. Például az aszimmetrikus reláció vagy az összefüggő reláció mindig irreflexív (↬38.o.). Von Wright sohasem vette fel az axiómarendszerébe a reflexivitás tulajdonságát, de több helyen is kategórikusan kijelentette, hogy az intrinzikus preferenciának irreflexívnek kell lennie mondván: „egy állapot sem lehet jobb önmagánál” [795, 147]. A reflexivitás direkt tárgyalását tekintve az egyetlen kivételt talán az jelenti, hogy a preferenciastruktúrák tárgyalásakor felveszik az axiómák közé a preferenciarelációk reflexivitási tulajdonságait.

Preferencia és szimmetricitás

Az aszimmetrikusság feltételezése nagyon erős, sokan nem is fogadják el. A von wrighti axióma elfogadása azt jelenti, hogy két világállapot összehasonlításakor mindig valamelyiket szigorúan jobbnak találjuk a másiknál. Azok, akik ezt a tulajdonság tételezik, azok az **szigorú preferenciát** teszik az elméletük középpontjába. Ezzel szemben azok, akik a **gyenge preferencia** fogalmát tartják alapvetőnek, és arra alapozzák elméletüket, azok a gyenge preferenciához szükséges **antiszimmetrikus** tulajdonság (↬33.o.) létezését írják elő. Bármelyiket is vesszük fel a modellünkbe kiindulásként, könnyen levezethető az egyik fogalom a másikból, ezért nagy jelentősége nincs választásnak. Mivel ezt korábban még nem rögzítettük, érdemes itt felvenni a tételek közé a gyenge preferencia antiszimmetrikusságát.

(BL.42) $\textsc{antisymmetric}(\mathbf{R})$

R gyenge preferenciarendezés **antiszimmetrikus**

A szimmetrikusság egyik preferenciarendezésnél sem jöhet szóba, viszont – ahogy azt már korábban rögzítettem – mind a közömbösségi, mind az összehasonlíthatatlansági reláció szimmetrikus.

Preferencia és összefüggőség

Von Wright nem foglalkozott a teljesség tulajdonságával, a Chisholm és Sosa páros sem, Bengt Hansson viszont a totalitás tulajdonságát felvette az axiómái közé (↬95.o.). Bár Bengt Hansson elméletében fontos szerepe van a totalitás tételezésének,[43] ettől még Sven Ove Hansson igaza van, amikor azt állítja, hogy preferencialogikában általában nem követelik meg ezt a tulajdonságot. Meg is magyarázta, miért. Azért nem, mert ez a feltétel nagyon erős elvárást teremt, ami sokszor nem vagy csak nagyon nehezen teljesíthető. De miért is kéne teljesülnie? Ha valaki preferálja m-t mind n-nel, mind o-val szemben, akkor tud választani, miért kéne foglalkoznia azzal, hogy n és o közül melyiket preferálja inkább. Még konkrétabban: ha autót akarunk vásárolni, és két lehetőség között ingadozunk, akkor ez ezek közti döntéshez nincs szükségünk arra, hogy állást foglaljunk a további húsz kocsivásárlási lehetőséggel kapcsolatban. Elméleti szempontokból azonban sokat jelent, ha valahol kimondhatjuk a teljesség fennállását, ezért a formális elméletek szeretik ezt a tulajdonságot (lásd Bengt Hansson).

Sven Ove Hansson röviden kitér arra a kérdésre, hogy mit lehet tenni, ha egy halmazon nem teljes a rendezés (incomplete), és teljesé akarjuk azt tenni. S.O. Hansson szerint három kimenet lehetséges: vagy egyértelműen megoldható (unique resolvable), amikor egyetlen elrendezést kapunk, vagy többféleképp megoldható (multiply resolvable), amikor többféle rendezést kapunk, vagy megoldhatatlan (irresolvable), amikor nem jutunk teljes rendezéshez. Utóbbira példa: választani a gázai piramis lerombolása vagy a pandák kihalása között. Az első két eset S.O. Hansson összehasonlíthatatlannak (incomparable), a harmadikat inkommenzurábilisnak (incommensurable) nevezi [336].

Preferencia és a disztribúció elve

A negyedik von wrighti axióma, a disztribúció elve (↬91.o.) azzal a kérdéssel kapcsolatos, hogy miként működik a preferencia a diszjunkció esetén, vagyis akkor, amikor úgy kell a dolgokat összehasonlítani, hogy az állapotok diszjunktív kapcsolatban lehetnek egymással. Von Wright erre azt a példát adta, hogy a munkahelyén valakinek döntenie kell arról, hogy mit preferál jobban: a fizetésemelést (m) vagy hosszabb éves szabadságot (n) a napi munkaidő csökkentéshez (o) képest. A kérdés, hogy mit lehet mondani az ilyen összetett állapotokra vonatkozóan ($(m \vee n)$ az o-val szemben). Von Wright ilyen-olyan feltételezések és levezetések után hozta ki azt formulát, amit végül axiómának minősített.

BL.xx $\forall m \forall n \forall o \forall p (\mathbf{P}(m \vee n, o \vee p) \leftrightarrow (\mathbf{P}(m \wedge \neg o \wedge \neg p, \neg m \wedge \neg n \wedge o) \wedge \mathbf{P}(m \wedge \neg o \wedge \neg p, \neg m \wedge \neg n \wedge p) \wedge \mathbf{P}(n \wedge \neg o \wedge \neg p, \neg m \wedge \neg n \wedge o) \wedge \mathbf{P}(n \wedge \neg o \wedge \neg p, \neg m \wedge \neg n \wedge p)))$

Nyugodtan mondhatjuk, hogy sokadik ránézesre sem tűnik intuitívnek a formula. Ebben a formájában nehéz befogadható értelmezést adni neki. Pedig van. Több más szerző mellett

[43] Amikor nem teljesül az összefüggőség (totalitás) tulajdonsága, akkor a preferenciarendezést Bengt Hansson csak részlegesnek nevezi, és ezen túl is igaz, hogy a totalitás tulajdonságára több fontos tételében is erősen támaszkodik [329].

Patrick Girrard is újraértelmezte, leegyszerűsítette az eredeti von wrighti formulát és megmutatta, hogyan azt le lehet vezetni a konjunktív expanzió elvének és a jóval egyszerűbb disztributivitási tulajdonságoknak a felhasználásával [297]. Az átalakítás eredményeként a von wrighti formula két egyszerűbb és érthetőbb formulára visszavezethető.

(BL.43) $\forall m \forall n \forall o(\mathbf{P}(m \vee n, o) \leftrightarrow (\mathbf{P}(m, o) \wedge \mathbf{P}(n, o)))$

P preferenciareláció a diszjunkcióra **jobbról konjunktíve disztributív**

(BL.44) $\forall m \forall n \forall o(\mathbf{P}(m, n \vee o) \leftrightarrow (\mathbf{P}(m, n) \wedge \mathbf{P}(m, o)))$

P preferenciareláció a diszjunkcióra **balról konjunktíve disztributív**

A fenti képletek alapján már az is megállapítható, hogy a von Wright által megfogalmazott disztribúciós axiómák meglepően közel vannak Bengt Hansson axiómáihoz (↬95.o.). Az egyszerűsített von wrighti formulákhoz képest a hanssoni disztribuciós feltételek csak annyiban gyengébbek, hogy amíg az utóbbiak csak az egyik irányban érvényesek, tehát csak kondicionális művelet kapcsolja össze a két tagmondatot, addig az előbbi esetben kétirányú kapcsolatot ír le a bikondicionális logikai kapcsoló.

BL.xx $\forall m \forall n \forall o((\mathbf{R}(m, n) \wedge \mathbf{R}(m, o)) \rightarrow \mathbf{R}(m, n \vee o)))$

BL.xx $\forall m \forall n \forall o((\mathbf{R}(m, o) \wedge \mathbf{R}(n, o)) \rightarrow (\mathbf{R}(m \vee n, o)))$

Persze azért van egy másik különbség is közöttük, csak azt talán kicsit nehezebb észrevenni. Különböznek ugyanis abban is, hogy amíg von Wright a $\mathbf{P}(m, n)$ szigorú preferenciára, addig Bengt Hansson az $\mathbf{R}(m, n)$ gyenge preferenciára mondta ki a disztributivitási axiómáit. Tanulságos egyébként az a mód is, ahogy Bengt Hansson eljutott az axiómáihoz. Először feltételezte és értékelte az erős disztribúciós axiómákat, majd meggyőző példákkal igazolta, hogy ezek elfogadása intuícióinkkal ellentétes következményekkel járna [329]. Első lépésben Hansson feltette, hogy a következő formulát kellene elfogadnunk.

(BL.45) $\forall m \forall n \forall o \forall p(\mathbf{P}(m \vee n, o \vee p) \leftrightarrow (\mathbf{F}(m, n, o, p) \veebar \mathbf{F}(n, m, o, p) \veebar \mathbf{F}(m, n, p, o) \veebar \mathbf{F}(n, m, p, o)))$

A formula jobb oldalán látható négy argumentumú **F**-operátor azt fejezi ki, hogy az argumentumokba írt négy alternatíva az adott preferenciasorrendbe rendezhető. A formula intuitívnak tűnik, hiszen annyit fejez ki, hogy az első két alternatív diszjunkcióját jobbnak értékeljük a második két alternatíva diszjunkciójához képest, akkor az egyenkénti összehasonlítás során az első két alternatíva előbbre kell, hogy kerüljön a második két alternatívánál (illetve fordítva). Másként: ha a bal oldalon kifejezett preferenciarendezés fennáll, akkor a jobb oldali képlet négy komponense közül az egyiknek kell teljesülnie. Valóban úgy tűnik, hogy ez a feltételezés elfogadható. B. Hansson azonban a fenti formulából levezet egy abból következő tételt, amiről viszont egy ellenpéldával bebizonyítja, hogy vállalhatatlan értelmezést is lehetővé tesz. A levezetett tétel így néz ki:

(BL.46) $\forall m \forall n \forall o(\mathbf{I}(m, n) \rightarrow \mathbf{I}(m \vee o, n \vee o))$

A képletben az $\mathbf{I}(m,n)$ indifferenciareláció szerepel (amit Bengt Hansson $\mathbf{S}(m,n)$ 'same as' relációnak nevez). Az alapfeltételezésből logikailag következő tétel azonban túl erős, és vállalhatatlan következményekkel jár. Ezt Hansson a következő példával szemlélteti. Tegyük fel az alábbiakat.

m	valaki megnyeri a fődíjat az ötös lottón
n	valaki megnyeri a fődíjat a hatos lottón
o	valaki nyer egy kettest az ötös lottón
$m \vee o$	valaki megnyeri a fődíjat és nyer egy kettest az I. lottón
$n \vee o$	valaki megnyeri a fődíjakat mind az ötös, mind a hatos lottón

Nyilván nem kell sokáig magyarázni, hogy amíg az m és n állapotokra mondhatjuk, hogy ugyanolyan jók (vagyis indifferensek egymással szemben), addig az $(m \vee o)$ és az $(n \vee o)$ "összetett alternatívák" messze nem tekinthetők egyenrangúaknak. Ezért Hansson a kezdeti feltételezést elvetette. A következő lépésben keresett egy gyengébb formulát abban reménykedve, hogy az már ki fogja elégíteni az intuícióinkat. Az addigi axiómák és tételek alapján levezetett egy következő tételt, ami így nézett ki:

(BL.47) $\forall m \forall n \forall o \forall p((\mathbf{I}(m,n) \wedge \mathbf{I}(o,p)) \rightarrow \mathbf{I}(m \vee \neg n, o \vee \neg p))$

A formula úgy értelmezhető, hogy ha van két-két állapot (m és n, illetve o és p), amelyek páronként egyforma értékesek ($\mathbf{I}(m,n)$ és $\mathbf{I}(o,p)$), akkor a két állapotpárból képzett két-két speciális diszjunktív állapot ($(m \vee \neg n)$ és $(o \vee \neg p)$) is ugyanolyan értékűek. Azonban ez a tétel sem maradhatott meg, mert Hanssonnak erre is akadt ellenpéldája. A történet részben megint a lottózásról szól, de most más alternatívák kerülnek a lottónyeremények mellé. Ezeket mutatom be a következő táblában.

m	valaki 100 millió forintot nyer az ötös lottón
n	valaki 100 millió nyer a hatos lottón
o	valaki meghal egy balesetben, mert elüti egy zöld busz
p	valaki meghal egy balesetben, mert elüti egy kék busz

Az első két alternatíva nyilván ugyanannyit ér az ember számára, és igaz ez a második kettő alternatívára is. De ezek konjunkcióit egymással összevetve már egyáltalán nem áll fent az értékegyezés.

Ezek után vezette be Hansson a már ismertetett axiómáit, amelyeket "hivatalosan" is axiómaként minősített. Ezután azonban, miután a disztribúciós axiómákból kiindulva további tételeket vezetett le, – némiképp meglepő módon – maga hozott példát arra, hogy még ezek a legyöngített formulák is képesek az elvárásainkkal ellentétes eredményeket produkálni.[44] A példa szerint valaki távol az otthonától a családjától kap leveleket. Tegyük fel, hogy a következő hírek érkeznek hozzá.
A kérdés az, vajon hogyan rendezi el ezeket a híreket akkor, ha diszjunkciót is megengedünk közöttük. Egy furcsa, de logikailag nem támadható kiindulópontot választva Hansson bemutat egy logkailag korrekt levezetést, amelynek a végén az intuíciónkkal ellentétes eredményhez jutunk. A formulákkal kifejezett lépések a következők.

[44] Hansson egy lábjegyzetben megjegyezte, hogy Sven Danielssonnak tulajdonítja bemutatott példáját.

m	a kutyája jól van
n	a fia jól van
o	a lánya jól van
p	a felesége meghalt autóbalesetben

1) $\mathbf{R}(m, n \vee (o \wedge p))$
2) $\mathbf{R}(m, (n \wedge p) \vee o))$
3) $\mathbf{R}(m, (n \vee (o \wedge p)) \vee ((n \wedge p) \vee o))$
4) $\mathbf{R}(m, (n \vee (o)) \vee ((n) \vee o))$
5) $\mathbf{R}(m, (n \vee o))$

Kiindulásként két preferendezést állított fel Hansson diszjunktív állapotok között. A fenti első sor ($\mathbf{R}(m, n \vee (o \wedge p))$) azt fejezi ki, hogy jobb az a hír, mely szerint 'a kutyája jól van' ahhoz az összetett hírhez képest, hogy 'a fia jól van (n) vagy a lánya jól van, de a felesége meghalt' ($o \wedge p$). A rossz hír jelenléte a diszjunkció jobb oldali tagjában az egész jobb oldali relátumot ($n \vee (o \wedge p)$) diszpreferálta. Ugyanez igaz a második sorra is a megfelelő módosításokkal. A harmadik sor formulájához az első disztribúciós axióma alkalmazásával juthatunk, amikoris az axióma képletébe behelyettesítjük a fenti két sort, mint az axióma bal oldali konjunkciójának két elemét. A harmadik formulából úgy kapjuk meg a negyediket, hogy kétszer alkalmazzuk a belső konjunkciókra a konjunkció eliminálásának műveletét (↬25.o.). Ennek eredményeként elhagyhatjuk a belső konjunkciók p-t tartalmazó tagjait. A negyedik sor formuláját sima propozíciólogikai egyszerűsítésekkel "kitisztíthatjuk", így jutunk az ötödik formulához. Ez viszont olyat állít, ami intuícióellenes, hiszen a $\mathbf{R}(m, (n \vee o))$ formula annyit jelent, hogy az illető jobban örülne annak a hírnek, hogy 'jól van a kutyája' ahhoz képest, hogy 'jól van a fia vagy a lánya'. Feltételezhetjük ugyan, hogy lehetnek ilyen beállítódású személyek is, de semmiképpen sem ezt az attitűdöt tartjuk általánosnak, ezért ezt a következtetést mint kontraintuitív végeredményt nyugodtan "elutasíthatjuk". Bengt Hansson rezignált végkövetkeztetése ezekután csak az, hogy a két disztribúciós axiómát nem lehet feltétel nélkül alkalmazni.

Preferencia és a kontrapozíció elve

A témakör felvezetésekor említettem, hogy Sören Halldén felvette a posztulátumai közé a kontrapozíció elvét (↬97.o.). Von Wright az 1963-as esszéjében még nem beszélt róla, de az 1972-es cikkében már igen, és noha az elvet nem vette fel az új axiómarendszerébe, az intuícióinkkal megegyezőnek és ezért – legalább részben – elfogadhatónak tartotta. Tette ezt annak ellenére, hogy előbb Chisholm és Sosa szerzőpár[154], majd Bengt Hansson is meggyőző példákat hozott arra, hogy feltétel nélkül nem szabad elfogadni ezt az elvet [329]. Chisholm és Sosa ellenpéldája nem annyira a hétköznapi élet problémateréből merített, de ettől még a belőle kibontható kellemetlen logikai következmény fel- és elismerhető. Két állítást mutattak be azzal, hogy hasonlítsuk össze őket, illetve a tagadásaikat egymással.
Chisholmék szerint a $\mathbf{P}(m, n)$ ítélet elfogadható, vagyis jobb az az állapot, amiben boldog kócsagok vannak ahhoz képest, amiben "csak" kövek vannak, viszont az állítás kontraponáltja, vagyis a $\mathbf{P}(\neg n, \neg m)$ ítélet szerintük már nem igazán fogadható el kérdezvén: "miért lenne jobb állapot az, amikor nincsenek kövek azzal szemben, amikor nincsenek boldog kócsagok". A két semmi se nem jobb, se nem rosszabb egymással szemben.

m	vannak boldog kócsagok
n	vannak kövek
$\neg m$	nincsenek boldog kócsagok
$\neg n$	nincsenek kövek

Jobban megérthetjük a Chisholmék példája mélyén rejlő problémát, ha felidézem a Bengt Hansson által bemutatott esetet is [329]. Képzeljük el, hogy valaki valamilyen versenyen indul, ahol lehet első és lehet további díjakat nyerni. Tekintsük a következő alternatívákat.

m	valaki első díjat nyer
n	valaki valamilyen díjat nyer
$\neg m$	valaki nem nyer első díjat
$\neg n$	valaki semmilyen díjat sem nyer

Első lépésként nyugodtan feltételezhetjük, hogy $\mathbf{P}(m,n)$ fennáll, hiszen az első díj többet ér bármilyen más díjnál. Ennek a konkrét relációnak a $\mathbf{P}(\neg n,\neg m)$ kontraponáltja viszont már nem igaz, hiszen ahhoz képest, hogy az ember nem nyeri meg az első díjat (ami itt együtt jár azzal, hogy semmilyen díjat nem szerez), bármilyen díj elnyerése jobb lehet, tehát a semmilyen díj elnyerésének alternatívája nem jobb az első díj el nem nyerésének alternatívájához képest. Ebben a történetben már egyértelműen látszik, hogy az ellentmondást az okozza, hogy a két alternatíva (m és n) nem független egymástól. Ez azt jelenti, hogy sem maguk, sem a tagadásaik nem teljesen diszjunktak, ami azért okoz problémát, mert ez a helyzet lehetetlenné teszi köztük az egyértelmű összehasonlítást.

Mielőtt továbblépnék a megoldás irányába, előtte még fontosnak tartom jelezni, hogy ez a probléma nemcsak a preferenciarendezés sajátja. S.V. Hansson arra is mutatott példát, amikor a kontrapozíció az indifferenciareláció esetében is gondot okoz [336, 349]. Tekintsük a következő helyzetet (illetve azok tagadásait).

m	A-nak két példánya van Rousseau Társadalmi szerződéséből
n	A-nak egy példánya van Rousseau Társadalmi szerződéséből
$\neg m$	A-nak nincs két példánya Rousseau Társadalmi szerződéséből
$\neg n$	A-nak nincs példánya Rousseau Társadalmi szerződéséből

Egy ilyen helyzetben – mondja S.V. Hansson – a két alternatíva m és n indifferens lehet A számára, hiszen egyszerre úgyis csak egy könyvet olvashat az ember.[45] Mondhatjuk tehát, hogy $\mathbf{I}(m,n)$ fennáll. De nem mondatjuk ezt az alternatívák kontraponáltjára, hiszen $\neg n$ esetén A nem olvashatja a könyvet, míg $\neg m$ esetén van rá esélye, hiszen attól, hogy nincs két példány a könyvből, egy még lehet.

Már jeleztem, hogy mi okozza itt a problémát: ezek az ellentmondások akkor keletkeznek, amikor az alternatívák nem teljesen függetlenek egymástól. Sven Ove Hansson ezért azzal javasolta kiegészíteni a kontrapozíció elvét rögzítő axiómát, hogy abból ki kell zárni az alternatívák közötti logikai inkonzisztencia lehetőségét. Ezt a kizáró feltételt az alábbi módon fejezhetjük ki.

[45] Itt fontos, hogy csak erre a szempontra figyeljünk a példában, mert ha beemelnénk azt a szempontot is, hogy a két könyvből az egyiket el lehetne adni és új könyvet lehetne érte venni, akkor a "két-könyves" állapot nyilván jobbnak számítana.

(BL.48) $\forall m \forall n (\mathbf{P}(m,n) \leftrightarrow (((m \wedge \neg n) \neq \bot \wedge (\neg m \wedge n) \neq \bot) \rightarrow \mathbf{P}(\neg n, \neg m)))$

ellentmondásmentes kontrapozíció elve (HUN) – principle of contradiction-free contraposition (HUN)

Ugyan a fenti tétel a szigorú preferenciarendezésre vonatkozik, de ugyanúgy ki lehet mondani az indifferenciarelációra nézve is. Ezzel a szűkítéssel már elkerülhetjük a kontrapozíció elvének ellentmondásait – gondolhatnák, de nem így van. Annak bizonyságára, hogy semmi sem olyan egyszerű, mint ahogy látszik, a Sven Ove Hansson által javasolt, fenti szűkítve-pontosító definíció után rögtön bemutathatunk egy olyan példát, amely ellentmondani látszik az ellentmondásmentes kontrapozíció elvének. A példa – ki mástól, mint – Sven Ove Hanssontól származik [333]. Képzeljük el az alábbiakat.

m	valaki felajánl 100 dollárt jótékony célokra
n	valaki felajánl 50 dollárt jótékony célokra
$\neg m$	valaki nem ajánl fel 100 dollárt jótékony célokra
$\neg n$	valaki nem ajánl fel 50 dollárt jótékony célokra

Ebben a helyzetben nyilván fennáll a $\mathbf{P}(m,n)$, de nehéznek tűnik elfogadni a $\mathbf{P}(\neg n, \neg m)$ kontrapontált igazságát. Miért lenne jobb az egyik semmi a másik semminél? Hansson logikai érveket is hoz amellett [333], hogy bár furcsa, de mégiscsak el kéne fogadnunk a kontrapozíció érvényességét ebben az esetben is. Én csak egy másik példát hoznék, hogy pszichológiailag jobban elfogadhatóvá váljék mindez. Képzeljük el, hogy lottózunk, és két alternatívát kell értékelnünk, valamint képzeljük még azt is ide, hogy mindkét alternatívát megálmodtuk, és felébredve konstatáltuk, hogy a nyeremény csak álom volt.

m	nyerünk 2 milliárd 5 forintot a lottón
n	nyerünk 1429 forintot a lottón
$\neg m$	nem nyerünk 2 milliárd 5 forintot a lottón
$\neg n$	nem nyerünk 1429 forintot a lottón

Nyilvánvaló, hogy fenntarthatjuk a $\mathbf{P}(m,n)$ igazságát. Ezek után gondoljunk bele abba, melyik ébredés utáni felismerés lehet fájdalmasabb a két negált alternatíva közül. A példa alapján talán már jobban elfogadható, hogy az álmodott (képzelt) hatalmas nyeremény nem teljesülése iránt nagyobb fájdalmat érzünk, mint a kis nyeremény "elvesztése" miatt – még akkor is, ha mindkét esetben ugyanazt a semmit "nyertünk", vagyis jobban preferáljuk a kisebb fájdalmat (a kisebb nyeremény elmaradása miatt), mint a nagyobb fájdalmat (a nagyobb nyeremény elmaradása miatt).

Talán kicsit nehezebb elfogadni az igazságát, de úgy tűnik, hogy az előző két példa csak látszólag mond ellent az ellentmondásmentes kontrapozíció törvényének. De utóbbi elv bevezetése még nem ad magyarázatot Chisholm és Sosa ellenpéldájára. Chisholmék boldog kócsagos történetében benne rejlő látszólagos ellentmondást akkor oldhatjuk fel, ha már többet tudunk a következő fontos preferencialogiai axiómáról, a konjunktív expanzió elvéről. Sören Halldén, Bengt Hansson, Sven Ove Hansson is hivatkozott arra, hogy a kontrapozíció és a konjunktív expanzió elvei között logikai azonosság van [329]. Ezt egyébként könnyű belátni. A konjunktív expanzió elvéből kiindulva lépésről lépésre el lehet jutni a kontrapozíció elvét kimondó formuláig.

(BL.49) a konjuktív expanzió és a kontrapozíció elve logikailag azonos

1) $\mathbf{P}(m,n) = \mathbf{P}(m \wedge \neg n, \neg m \wedge n)$
2) $\mathbf{P}(m,n) = \mathbf{P}(\neg n \wedge m, n \wedge \neg m)$
3) $\mathbf{P}(m,n) = \mathbf{P}(\neg n \wedge \neg\neg m, \neg\neg n \wedge \neg m)$
4) $\mathbf{P}(m,n) = \mathbf{P}((\neg n) \wedge \neg(\neg m), \neg(\neg n) \wedge (\neg m))$
5) $\mathbf{P}(m,n) = \mathbf{P}((\neg n), (\neg m))$
6) $\mathbf{P}(m,n) = \mathbf{P}(\neg n, \neg m)$

Az első sor mutatja a konjunktív expanzió elvét. A második sorban csak megcseréltük a konjunkciók összetevőinek sorrendjét. A harmadik lépésben betettünk egy-egy kettős tagadást, ami – mint tudjuk – nem változtat az igazságértékeken. A negyedik sorban zárójellel határoltuk el mindegyik változó negáltját, és az így képzett "új" változókra (tehát a $\neg m$-re és az $\neg n$-re) nézve újra alkalmaztuk a konjuktív expanzió elvét ("visszafele"). Így kaptuk az ötödik sor formuláját, amiből már csak a fölösleges zárójeleket kellett elhagyni, hogy a hatodik sorban láthassuk a kontrapozíció törvényének képletét.

Ez alapján mindazt, amit a kontrapozícióval kapcsolatban bemutattam, elmondhatnám a konjunktív expanzió elvével kapcsolatban is. Még azt is mondhatnánk, hogy szükségtelen a konjunktív expanzió elvének önálló tárgyalása, de ez nem így van. Látni fogjuk, hogy további, eddig még nem tárgyalt problémákkal kerülünk szembe a következő fejezetben. Logikai szempontból valóban nem lenne szükség a konjuktív expanzió elvének elemzésére, de gyakorlati megfontolások miatt mégis érdemes megtenni ezt. Egyszerűen azért, mert a logikailag ekvivalensnek tekinthető "jelenségekre" másfajta nézőpontból közelítünk, és ez – legalább részben – másfajta kép kialakításához, másfajta összefüggések felismeréséhez segíthet minket.

Preferencia és a konjunktív expanzió elve

Von Wright harmadik axiómája, a konjunktív expanzió elve (↬91.o.) az (állapot)változások és a preferencia között állapít meg összefüggést. Von Wright abból indult ki, hogy a két állapot (m és n) együttes figyelembe vételével négyféle, egymást kölcsönösen kizáró világállapot lehetséges:

$$(m \wedge n) \underline{\vee} (m \wedge \neg n) \underline{\vee} (\neg m \wedge n) \underline{\vee} (\neg m \wedge \neg n)$$

és ezek bármelyikének fennállását feltételezzük, mindegyik esetben négyféle változás lehetséges (minden marad a régiben, az egyik állapot marad, a másik változik, az egyik marad, a másik változik, valamint mindkettő változik). Bármelyik ilyen kettős-állapotban vagyunk is, mindegyik esetben az a változás lehet kívánatos a $\mathbf{P}(m,n)$ preferenciával rendelkező ágens számára, amelyik az $(m \wedge \neg n)$ állapotot eredményezi a $(\neg m \wedge n)$ helyett. Ez mind a négy esetben más és más transzformációt kíván meg, de a változás tartalma mindig az előbbi kell legyen.

Ugyanezt a problémát azonban más irányból is meg lehet közelíteni. Fel lehet ugyanis tenni azt a kérdést, hogy mi is történik valójában a preferenciáink kinyilvánításakor. Azt tudjuk, hogy ilyenkor állapotokat hasonlítunk össze. Ez mit jelent? Képzeljünk el egy gyereket [336], akinek azt mondjuk, hogy:

> Jobb az, ha van egy macskánk, mint egy kutyánk.

Ha erre a kijelentésre a kisgyerek atr válaszolja, hogy:

> Nem! Az a jobb, ha van egy kutyánk, ha van már egy macskánk is.

Ennek hallatán azt feltételezzük, hogy a gyerek nem érti még az összehasonlítás (és a választás) lényegét. Az összehasonlítás (választás) ugyanis feltételez valamiféle diszjunktságot az alternatívák között abban az értelemben, hogy kizárja a két állapot együttes kívánását. Ilyenkor – az egymást kölcsönösen kizáró – a 'kutya-de-nem-macska' vagy a 'macska-de-nem-kutya' állapotokat (alternatívákat) kell összehasonlítani. Nos, pontosan ez fejezi ki a konjunktív expanzió elve. Ezért javasolta már Halldén is, aztán von Wright is az elv elfogadását. Sokan hivatkoztak Héctor-Neri Castañeda véleményére, miszerint ezt az elvet fel lehet fedezni már Szent Pálnak a házasságról szóló mondatában is.

> „De ha magukat meg nem tartóztathatják, házasságban éljenek: mert jobb házasságban élni, mint égni." (Pál Apostolnak Korinthusbeliekhez írt első levele, 7.9)

Vagyis – ahogy Castañeda értelmezi – „jobb házasságban élni és nem égni, mint nem házasságban élni és égni." Ha az összehasonlítást (választás) így értelmezzük, akkor szükségesnek látszik a konjunktív expanzió elve. Ezzel azonban nem mindenki van így. Chisholm és Sosa például elutasította ezt az elvet [154]. A kócsagos példájukat kicsit átalakítva a következő alternatívákat ajánlották összehasonlításra.

m	vannak kövek
n	vannak boldogtalan kócsagok
$m \wedge \neg n$	vannak kövek és vannak boldog kócsagok
$\neg m \wedge n$	nincsenek kövek és nincsenek boldog kócsagok

Chisholmék szerint kijelenthetjük, hogy az $m \wedge \neg n$ alternatíva jobbnak mondható, mint a $\neg m \wedge n$ állapot, ellenben az m-t inkább közömbösnek érezzük az n-hez képest. Még inkább érezhetjük az ellentmondást abban a példában, amit Chisholmék egy másik cikkükben adtak meg [153]. Ebben a következő állapotokat kell összehasonlítani.

m	nincs három boldog görög
n	van két boldogtalan római
$m \wedge \neg n$	nincs három boldog görög és nincs két boldogtalan római
$\neg m \wedge n$	van három boldog görög és van két boldogtalan római

Az egyszerű alternatívák összehasonlításából az jön ki, hogy **P**(m,n) igaz, hiszen a két boldogtalan római léte rosszabb, mint a boldog görög nem létezése. Ezzel szemben a boldog görögök és boldogtalan rómaiak nem létezésével szemben jobb a három boldog görög létezése még a két boldogtalan rómaival együtt is, mivel az boldogságok mérlege még így is pozitív.[46]

Sven Ove Hansson ennek a problémának a megoldására is ugyanazt javasolja, mint a kontrapozíció esetében: pontosítani, szűkíteni kell a konjunktív expanzió formuláját. Azért kell ezt tenni, mert itt is ugyanaz a probléma oka, az ellentmondásos állapotok kialakulása. Szemléltetésként bemutat egy olyan példát, amiben a következő két alternatíva között kell dönteni.

[46] röviden említeni Chisholmék hedonisztikus szemléletét, és ami ebből fakad.

m	nyerek 100 millió forintot a lottón, és tartozom valakinek 2000 forinttal
n	tartozom valakinek 2000 forinttal
$m \wedge \neg n$	nyerek 100 millió forintot a lottón, és tartozom valakinek 2000 forinttal, és nem tartozom valakinek 2000 forinttal

Ebben a helyzetleírásban jól látszik az a probléma, hogy $m \wedge \neg n$ állapotok között ellentmondás van. A logikai ellentmondás, az elméleti gondok ellenére természetesen meg tudjuk oldani az ilyen helyzeteket, igenis tudunk dönteni ilyenkor is, mégpedig úgy, hogy az összehasonlítás ebben az esetben más lesz: nem $m \wedge \neg n$ és $\neg m \wedge n$, hanem m és $\neg m \wedge n$ alternatívák közül választunk. Éppen ezért javasolja azt Hansson, hogy ezt a "pontosítást" emeljük be a preferencia modelljébe. Ekkor a következő formulához jutunk.

(BL.50) $\forall m \forall n(\mathbf{P}(m,n) \leftrightarrow (m \wedge \neg n = \bot \rightarrow \mathbf{P}(m, \neg m \wedge n)) \underline{\vee} (m \wedge \neg n \neq \bot \rightarrow \mathbf{P}(m \wedge \neg n, \neg m \wedge n)))$

ellentmondásmentes konjunktív expanzió elve (HUN) – principle of contradiction-proof conjunctive expansion (ENG)

Ez az ellentmondásmentes szűkítés hasonlít a kontrapozíció elve esetében alkalmazott megoldáshoz, amit már nyilván nem tarthatunk véletlennek annak ismeretében, hogy a konjunktív expanzió és a kontrapozíció elvei logikailag ekvivalensnek tekinthetőek. Kiderült azonban, hogy ez a szűkítés még mindig nem elég. Sven Ove Hansson mutatott egy olyan példát, ami miatt az ellenmondásmentes konjunktív expanzió elvét még tovább kellett pontosítani [336]. Lássuk a példát!

m	valaki utazik a Holdra
n	valaki űrhajóval utazik
$m \wedge \neg n$	valaki utazik a Holdra és nem űrhajóval utazik
$\neg m \wedge n$	valaki nem utazik a Holdra és űrhajóval utazik

Itt más típusú probléma jelentkezik, mint a korábbiakban. A táblázatban feltüntetett mind a négy állapot lehetséges logikailag. Jelenlegi tudásunk és felkészültségünk szerint azonban az $m \wedge \neg n$ állapot technikailag nem elképzelhető, így minden olyan összehasonlítás, amelyben ez az alternatív szerepel, értelmetlennek tűnik, nem megengedhető. De ez a probléma orvosolható, mégpedig úgy, hogy az ellentmondásmentesség követelményét ki kell terjeszteni erre az esetre is. Nemcsak a logikai ellentmondások lehetőségét kell tehát kizárni, hanem azt is, hogy szemantikailag értelmetlen alternatívákat kelljen összehasonlítani valami mással. Ezt a további szűkítést **ellentmondásbiztos** megoldásnak nevezem, szemben a korábbi ellentmondásmentes szűkítéssel, amely "csak" a logikai ellentmondásokat zárta ki a formulákból.

(BL.51) $\forall m \forall n(\mathbf{P}(m,n) \leftrightarrow (m \wedge \neg n = ??????? \rightarrow \mathbf{P}(m, \neg m \wedge n)) \underline{\vee} (m \wedge \neg n ?????? \rightarrow \mathbf{P}(m \wedge \neg n, \neg m \wedge n)))$

ellentmondásbiztos konjunktív expanzió elve (HUN) – principle of contradiction-proof conjunctive expansion (ENG)

Ha ezt a kiegészítést megtettük, akkor természetesen ezt a második szűkítést érdemes elvégeznünk a kontrapozíció elvére vonatkozóan is. Így az alábbi módosított formulához jutunk.

(BL.52) $\forall m \forall n(\mathbf{P}(m,n) \leftrightarrow (((m \wedge \neg n) \neq ???? \wedge (\neg m \wedge n) \neq ????) \rightarrow \mathbf{P}(\neg n, \neg m)))$

ellentmondásbiztos kontrapozíció elve (HUN) – principle of contradiction-proof contraposition (HUN)

A konjunktív expanzió és a kontrapozíció elveinek kétszeres pontosítása után visszatérhetünk Chisholm és Sosa legelső kócsagos példájához (↬104.o.). Sven Ove Hansson szerint ha a példára kétféleképpen is alkalmazzuk a konjunktív expanzió elvét (azaz a kontrapozíció bal- és jobb oldalára egyaránt rávetítjük a konjunktív expanzió képletét), akkor belátható, hogy Chisholmék példája nem tartalmaz ellentmondást – akkor sem, ha a "hétköznapi intuíciónk" számára ez nehezebben befogadható [336].

Preferencia és választás

Bár a von wrighti konjuntív expanzió elvével kapcsolatban írt tanulmányt R.E. Jennings [379], mégsem az elvről szóló fejezetben foglalkozom Jennings véleményével, hanem a választás és preferencia kapcsolatát taglaló részben. Ennek az az elsődleges oka, hogy Jennings nem a korábban kifejtettek értelemben szólt hozzá a konjunktív expanzió elvéhez, hanem csak azt állította, ez tulajdonság nem a preferenciarelációt, hanem a **választás** (choice) jelenségét jellemzi. Jennings szerint ha *m*-t preferálom *n*-nel szemben, akkor az azt is kell, hogy jelentse, hogy van valamilyen tapasztalatom mind az *m*-mel, mind az *n*-nel kapcsolatban. Valaminek a megtapasztalása azonban még nem jelenti automatikusan azt is, hogy tudnánk, milyen lehet annak a valaminek a hiánya, ezért erre nem is hivatkozhatunk az összehasonlítások során.

Az explikáció érdekében Jennings elkülönített kétféle preferenciafogalmat, és ezeket $\mathbf{P}^j$-preferencia és $\mathbf{R}^j$-preferencia kifejezésekkel jelölte.[47] Mindkét esetben ugyanazt a helyzetet értelmezzük:

> valaki az *m* alternatívát preferálja *n*-nel szemben

Azonban amíg a $\mathbf{P}^j$-preferenciával azt az értelmezést fejezzük ki, hogy:

> valaki jobban tetszik az *m*, mint az *n*-t

addig $\mathbf{R}^j$-preferencia segítségével azt reprezentáljuk, hogy:

> valaki inkább szeretné *m*-t, mint az *n*-t

Jennings szerint a két értelemzés abban különbözik egymástól, hogy amíg az előbbiben nincs semmilyen akarati tényező, addig az utóbbiban van, és így az előbbi valóban a preferenciát mint tetszést, az utóbbi pedig inkább a preferenciát mint választást (choice) reprezentálja. Ez a különbség megnyilvánul abban (is), hogy a preferencia esetében nem érvényes a konjunktív kiterjeszthetőség elve, míg a választás esetében igen. A kétféle reláció további két fontos szempontból is eltér egymástól. Egyfelől különböznek a magyarázatokban betöltött szerepük tekintetében, másfelől más előfeltételekkel élünk akkor, amikor a kétféle relációt tulajdonítjuk valakinek. A $\mathbf{P}^j$-preferencia és $\mathbf{R}^j$-preferencia aszimmetrikus viszonyban állnak egymással abban az értelemben, hogy az utóbbiakat mindig megmagyarázhatjuk az előbbiekre hivatkozva, de ez a lehetőség fordítva már nem áll fenn. Jennings példájával élve arra a kérdésre, hogy:

[47] Jennings saját relációi jelölésében természetesen nem használ felső indexet, azt csak én teszem itt ki, hogy nem legyen keveredés a már bevezetett **P** és **R** relációk és Jennings relációi között.

— Miért kéne neked inkább a narancs, mint mandarin?

adhatunk olyan magyarázatot, hogy egyszerűen hivatkozunk arra:

— Mert jobban szeretem a narancsot a mandarinnál.

de arra kérdésre, hogy:

— Miért szereted jobban a narancsot a mandarinnál?

nem tudunk értelmesen azzal válaszolni, hogy utalunk a vonatkozó $\mathbf{R}^j$-preferenciánkra.

A másik különbség a jenningsi $\mathbf{P}^j$- és $\mathbf{R}^j$-preferenciák között abban van, hogy milyen fokú és milyen típusú magyarázatot fogadunk el, amikor meg akarjuk indokolni őket. A $\mathbf{P}^j$-preferencia esetén az előző kérdésre elfogadjuk az alábbi választ:

— Hát, csak.

Ebben az esetben a magyarázat lényege az, hogy egyszerűen hivatkozhatunk a létező $\mathbf{P}^j$-preferenciánkra: "ezt szeretjük jobban". Itt nincs racionális magyarázat, nincs érvelés, nincs megokolás, nincs másra hivatkozás, csak a preferencia "puszta" deklarálása. Ezzel szemben a $\mathbf{R}^j$-preferencia esetén mindig kell valamilyen indoklás, valamilyen okra való hivatkozás. Ekkor a 'Miért?' kérdésre nem fogadjuk el az "önmagára hivatkozó" indoklást. Az alábbi kérdésre adott választ nem érezzük "teljesnek", elfogadhatónak.

— Miért kéne neked jobban fahéj, mint gyömbér?
— Hát, csak kéne.

Ilyenkor nem elég a 'csak' válasz. Ha mégis ezt feleljük, akkor azt úgy fogják értékelni mások, hogy nem akarjuk felfedni a preferenciáinkat, esetleg egyfajta deklarációként. Ha ugyanis azt mondjuk:

— Nekem jobban kéne most fahéj, mint gyömbér.

ezt inkább úgy lehet érteni, hogy nem a preferenciánkat fejeztük ki, hanem "kihirdettük", közöltük a választásunkat, de nem értelmezhetjük úgy, hogy ezzel magyarázatot adtunk volna a 'miért' kérdésre. A választásainkank mindig (de legalábbis az esetek többségében) van indoka, sőt, nagyon sok esetben ezek az indokok pont saját belső $\mathbf{P}^j$-preferenciánk.

A kétféle preferencia közti harmadik különbségről már volt szó: $\mathbf{P}^j$-hez kell tapasztalat az alternatívákról, $\mathbf{R}^j$-hez nem feltétlen szükséges. Jennings szerint mondhatjuk, hogy:

— Inkább szeretnék a pokolra kerülni, mint a mennyországba.

de – tényleges tapasztalat hiányában – nem mondhatjuk, hogy:

— Nekem jobban tetszik a mennyország, mint a pokol.

Jennings ezek után egy formalizálási javaslatot is bemutat az újonnan bevezetett fogalomra, a $\mathbf{R}^j$-preferenciára. Kiindulásként rögzíti, hogy a $\mathbf{P}^j$ reláció argumentumai állapotok, az $\mathbf{R}^j$ reláció argumentumai cselekvések (és csak kényelmi, rövidítési okok miatt használjuk ugyanazokat a jeleket az argumentumok változóinak). Abból indulhatunk ki, hogy ha $\mathbf{R}^j(m,n)$ igaz, vagyis valaki inkább tenné m-t, mint a n-t, akkor erre mondhatjuk azt is, hogy választania kell az alternatívák közül. Az m és n közti választás az exkluzív

diszjunkció (biszubtrakció) műveletével fejezhető ki (↬20.o.), vagyis írhatjuk, hogy $m \underline{\vee} n$. Az exkluzív diszjunkció képlete: $(m \vee n) \wedge \neg (m \wedge n)$, ahol a formula második tagjával zárjuk ki a két alternatíva együttes fennállását, és ezzel biztosítjuk az alternatívák diszjunktságát. Ezt a formulát viszont – a disztribúciós szabály segítségével – átalakíthatjuk a következő alakra: $(m \wedge n) \vee (\neg m \wedge n)$. Ha ezek után az m és n közötti választás tevékenységét $\#(m,n)$ szimbólummal jelöljük, akkor vonatkoztathatjuk rá az előző összefüggést.

(PL.53) $\text{CHOICE}(m,n) \equiv \#(m,n)$

(PL.54) $\#(m,n) \leftrightarrow \#((m \vee n) \wedge \neg (m \wedge n)) \equiv \#(m \wedge \neg n, \neg m \wedge n)$

a választás itt azt jelenti, hogy m-t választja

A választás mint exkluzív diszjunkció tételezésének van egy fontos következménye: ha a két alternatíva valamelyikének létezése maga után vonja a másik létezését, akkor nem beszélhetünk választásról, hiszen ekkor a két alternatíva nem, hogy nem zárja ki egymást, de az egyik feltételezi a másikat.

érvényes a konjunktív expanzió elve. konjunktívan kiterjeszthető

unális (unal) és duális (dual) leírás, illetve magyarázat.

A $\mathbf{P}^j$- és $\mathbf{R}^j$-preferenciák a cselekvéseink duális magyarázatában adhatnak segítséget.

K_1 — Miért csináltad x-t?
V_{11} — Azért, mert x-t akartam csinálni.
V_{12} — Azért, mert szeretem x-t csinálni.
K_2 — Miért csináltad x-t az y helyett?
V_{21} — Azért, mert inkább tenném x-t, mint y-t.
V_{22} — Azért, mert jobban szeretem csinálni x-t, mint y-t.

A kérdésekre adott válaszok magyarázóereje eltérő. A V_{11} és V_{21} válaszok üresek, nem érezzük elfogadható indoklásoknak őket. Ezzel szemben érvényesnek, elfogadhatónak mondható a V_{12} és V_{22} válasz, amikor egyszerűen hivatkozunk a preferenciáinkra (unális vagy duális módon). Fontos észrevennünk, hogy a V_{12}, illetve a V_{22} válaszok a $\mathbf{P}^j$-, illetve $\mathbf{R}^j$-preferenciának felelnek meg. Az egyik fontos különbség pedig az közöttük, hogy a $\mathbf{P}^j$-preferencia "önmagát magyarázza", nincs szüksége "külső" indoklásra, érvelésre, az $\mathbf{R}^j$-preferencia viszont önmagában nem áll meg, mindig szüksége van "külső" érvekre, indokokra.

> „My reasons for wanting to do x are my reasons for doing x. Similarly, the reasons why I would rather do x than y are my reasons for doing x rather than y. ... Thus the relation between liking to do x and doing x is different from the relation between wanting to do x and doing x. There is a parallel difference between the relation between having a P-preference of doing x to doing y and doing x rather than y, and the relation between having the R-preference (do x) R (do y) and doing x rather than y. Our likes provide reasons for and explanations of both our wants and our actions. Our P-preferences provide explanations of both our R-preferences and our choices." [379, 565]

Jennings értékelése átkötés a választáshoz akarati, intencionális többlet van itt, ezt kell pontosan megragadni

———————-

A preferenciákat vizsgálva von Wright felvetett még egy további tipizálási lehetőséget is, amikor elkülönítette azokat helyzeteket, amelyekben van **kockázat** (risk) azoktól, amelyekben nincs. Még azt is javasolta (igaz, nem indokolja meg, miért), hogy a kockázatmentes (risk-free) preferenciakinyilvánításokat nevezhetnénk **tiszta preferenciáknak** (pure preferences), míg a kockázatot is magába foglaló preferenciarendezésekre használhatnánk a **preferenciális választások** (preferential choices) kifejezést. Ha egy ponton módosítunk von Wright korábbi példáján azt feltételezve, hogy a munkavállalónak úgy kell döntenie a munkahelyén, hogy a fizetésemelést (m) vagy hosszabb éves szabadságot (n) választja inkább a napi munkaidő csökkentéshez (o) képest, hogy az első alternatíva diszjunkciójában (m vagy n) szereplő két összetevő közül csak az egyik teljesülhet, viszont nem előre lehet tudni, hogy melyik. Ha eközben igaz az is, hogy az alternatívák páronkénti összehasonlítás alapján az illető jobbnak tartja a fizetésemelést (m-t) a rövidebb munkaidőnél (o-nál), viszont utóbbit jobbnak ítéli a hosszabb szabadságnál (n-nél), akkor a döntési helyzetbe kockázat "kerül". Ha a munkavállaló az ($m \vee n$) opciót választja, akkor – a döntése után – kiderülhet, hogy a legpreferáltabb állapothoz jut, de az is lehet, hogy a legkevésbé kedvelt alternatívához. Ha viszont a második alternatívát választja, akkor biztosan nem érheti el a legkívánatosabb állapotot, amire adott valószínűsége pedig lett volna az első alternatíva választása esetén. Már ebből a példából is látszik, hogy a preferenciális választásokra nem igazak a disztributivitási kritériumok, ami miatt von Wright rögzítette is, hogy esszéjében csak a tiszta preferenciákkal foglalkozik.

Ha fel akarjuk venni ezt a megkülönböztetést az ontológiánkba, akkor definiálni kell a kockázat fogalmát, ami alapján már könnyen meg lehet határozni a most bemutatott kétféle preferenciafogalmat is.

(BL.55) $\text{RISK} \stackrel{\text{def}}{=} m|p_m, n|p_n$

(BL.56) $\text{PURE_PREFERENCE} \equiv \text{PREFERENCE} \stackrel{\text{def}}{=} \mathbf{P}(m|p_m, n|p_n) \wedge p_m = p_n = 0$

(BL.57) $\text{PREFERENTIAL_CHOICE} \equiv \text{CHOICE} \stackrel{\text{def}}{=} \mathbf{P}(m|p_m, n|p_n) \wedge p_m \neq 0 \wedge p_n \neq 0$

Kockázatmentes helyzetekben a tiszta preferenciáinkat nyilváníthatjuk ki, ezeket tarthatjuk az igazi, külső befolyástól (valószínűségekre alapozott megfontolásoktól) mentes értékkifejeződéseknek. Az intrinzikuskus preferencia kritériumrendszerét erre a típusra dolgozta ki von Wright. A **kockázatterhes helyzetekben**, amikor a különböző valószínűségekkel ellátott állapotok között kell választanunk, más logika mentén működik az ember. Ezt később fogom elemezni.

Ceteris paribus preferencia

Az ötödik axiómát von Wright még az amplifikáció elveként vezette be (↬91.o.), de később (és mindenhol máshol) ezzel a tulajdonsággal már a ceteris paribus preferenciát írta le (hogy aztán az 1972-es tanulmányában már holisztikus preferenciának hívta). A lehetséges terminusok közül végül a ceteris paribus preferencia kifejezés vált elterjedtté, ezt használom én is.

Bengt Hansson kritizálta az amplifikáció elvét mondván, hogy ellentmondáshoz vezethet, ha az alternatívák nem függetlenek egymástól. Könnyen szemléltethető mindez, ha feltételezzük, hogy $o = m \wedge n$. Ezt az egyenlőséget behelyettesítve az amplifikáció elvének formulájába, majd a fölösleges belső konjunkciókat kiiktatva, illetve a belső konjunkciók szimmetriáját kihasználva a következő képletet kapjuk:

$$1)\ \forall m \forall n \forall o(\mathbf{P}(m,n) \leftrightarrow (\mathbf{P}(m \wedge o, n \wedge o) \wedge \mathbf{P}(m \wedge \neg o, n \wedge \neg o)))$$
$$2)\ \forall m \forall n \forall o(\mathbf{P}(m,n) \leftrightarrow (\mathbf{P}(m \wedge m \wedge n, n \wedge m \wedge n) \wedge \mathbf{P}(m \wedge \neg(m \wedge n), n \wedge \neg(m \wedge n))))$$
$$3)\ \forall m \forall n \forall o(\mathbf{P}(m,n) \leftrightarrow (\mathbf{P}(m \wedge n, n \wedge m) \wedge \mathbf{P}(m \wedge \neg(m \wedge n), n \wedge \neg(m \wedge n))))$$
$$4)\ \forall m \forall n \forall o(\mathbf{P}(m,n) \leftrightarrow (\mathbf{P}(m \wedge n, m \wedge n) \wedge \mathbf{P}(m \wedge \neg(m \wedge n), n \wedge \neg(m \wedge n))))$$

A negyedik sorban már látszik az ellentmondás, hiszen a bikondicionális jobb oldalán álló konjunkció első tagja ($\mathbf{P}(m \wedge n, m \wedge n)$) ellentmondásban van von Wright **P**-relációjával, ami – lévén szigorú preferenciarendezés – aszimmetrikus. Ezt persze korrigálni lehet a korábban már többször alkalmazott ellentmondásbiztos klauzula beépítésével az elv formulájába. Ez tehát "kezelhető" problémát jelent. És nem is lényeges ponton "támadja" a ceteris paribus preferencia fogalmát. Von Wright számára nem ezért volt fontos ez a fogalom.

Amikor preferenciáról beszélünk, akkor két állapotot hasonlítunk össze. Feltehető a kérdés, hogy amikor így döntünk, akkor vajon figyelembe vesszünk-e más állapotokat, illetve állapotváltozásaokat, és ha igen, akkor mennyit? Mi van akkor, ha figyelünk arra is, hogy nem csak két, de több állapot is fennállhat akkor, mikor éppen azt a két állapotot hasonlítunk össze, és nemcsak az éppen összehasonlított állapotok változhatnak, hanem bármi más is. Akkor érthetjük meg, hogy miért is fontos ez a kérdés, ha von Wright harmadik axiómájára gondolunk, miszerint preferálni m-t n-nel szemben ekvivalens azzal, hogy a jelen helyzethez képest preferáljuk azt a változást, ami az $(m \wedge \neg n)$ helyzetbe visz minket ahhoz a változáshoz képest, ami az $(\neg m \wedge n)$ helyzetet eredményezi. Ebben az interpretációban már sokkal érthetőbb, miért lehet feltenni azt a kérdést, hogy ilyen esetekben miért nem veszünk figyelembe más változásokat, más helyzeteket is.

A kérdésre válaszul von Wright háromféle lehetőséget különít el egymástól: az abszolút, a feltétel-nélküli és feltételes preferencia kinyilvánításának helyzetét. Amikor az egyén az $(m \wedge \neg n)$ állapotot eredményező változást nemcsak a $(\neg m \wedge n)$, de bármely más állapotot eredményező változáshoz képest is preferálja, akkor **abszolút preferenciáról** (absolute preference) beszélhetünk. Ilyen esetekben ha egy világállapot tartalmazza a $(\neg m \wedge n)$ állapotot, akkor azt ez az egyén mindenképpen jobban tartja az összes többi világállapotnál, amelyben $(\neg m \wedge n)$ is benne van. Az ilyen abszolút preferenciából azonban egy ember (egy adott pillanatban) csak eggyel rendelkezhet (ezt a tézist von Wright röviden le is vezeti, amit itt nem mutatok be), ezért ezt a preferenciatípust von Wright elhanyagolhatónak véli (mivel a gyakorlatban érdektelen).

A **feltételes preferencia** (conditional preference) fogalmát úgy értelmezi von Wright, hogy ilyen esetekben két állapot összehasonlítása függ egy harmadik állapottól. Azt a példát adja erre, hogy ha valakinek pénzre van szüksége és el kell adnia egy ruhadarabját, akkor három ruhadarabot (trikót, cipőt és nadrágot) feltételezve másként értékelheti azt a kérdést, hogy melyiket adja el a trikó (m) és a cipő (n) közül attól függően, hogy mit gondol a nadrág (o) eladásáról. Ha nem kell (nem akarja) eladni a nadrágot, akkor preferálja a cípőt a trikóval szemben (mert az "megy" a nadrághoz), vagyis az utóbbit fogja eladni. Viszont ha el kell (el akarja) adni a nadrágot (mert, mondjuk, kinőtte), akkor már inkább a trikó megtartását választja a cipővel szemben (mert nadrág nélkül úgysem használhatná).

A feltétlen vagy **feltétel-nélküli preferenciáról** (unconditional preference) akkor beszélhetünk, amikor feltételezhetjük, hogy a világ nem változik más állapotait tekintve, tehát csak a *m* és *n* állapotokat érintő változások történnek. E helyzet leírására szokták használni a **ceteris paribus preferencia** kifejezést.[48] A von wrighti preferencialogika a feltétel-nélküli, ceteris paribus preferenciákról szól. A feltétel-nélküliség mozzanatát von Wright az ötödik axiómájával definiálta. Ez a tétel egy elsőrendű logikai szintaxison alapuló formula igazságát tételezi. Azóta kiderült, hogy a ceteris paribus elvet igazából a modális logika segítségével lehet jól megragadni. Von Wright megoldási javaslata után sokáig az volt a kérdés, mi lesz a jó megoldás erre. Sven Ove Hansson már egy 1995-ös cikkében is tett egy javaslatot a ceteris paribus preferencia formalizálására [335], amin aztán a 2001-es nagy preferenciatanulmányában módosított [336], de azt a megoldást, amit egyre szélesebb körben fogadnak el a szakirodalomban, mások dolgozták ki. Ezt viszont csak egy későbbi fejezetben mutatom be (↬120.o.).

Preferencia és tranzitivitás

A tranzitivitás feltételét is sokan vitatják arra hivatkozva, hogy a tényleges társadalmi gyakorlatban sok olyan példát találhatunk, amikor nem érvényesül a tranzitivtás. Von Wright is tudatában van a tranzitivitással szembeni kételyeknek, de azt válaszolja, hogy a tranzitivitás ellenében felhozott ellenpéldák szerinte mindig az extrinzikus preferenciákra vonatkoznak, amikor valami hatására időközben megváltozik valakinek a preferenciarendezése (hiszen, mondja von Wright, a tranzitivitás mérése időbeli folyamat, ami alatt változhat az extrinzikus preferencia). Szerinte az intrinzikus preferenciákra fenntartható a tranzitivitás kritériuma.

Még ha sokan vannak, akik túl erősnek tartják, ezért nem fogadják el a preferencia tranzitivitási feltételét, a tranzitivitás tulajdonságának fontosságát senki sem tagadja. Nélkülözhetetlenségét jelzi az a tény is, hogy sokféle módon próbáltak gyöngíteni a tranzitivitás kritériumán, és új tranzitivitás-közeli fogalmakat vezettek be. Ilyenek előfordultak a relációtulajdonságok között is (↬32.o.), emlékezzünk csak például a kvázitranzitivitás vagy az aciklusosság fogalmára. A relációtulajdonságokkal való minősítésen túl úgy nyerhetünk további tranzitivitás-közeli fogalmakat, hogy több relációt vonunk be a definíciókba. Ilyeneket definiáltuk a relációelméleti fejezetben, például az IP- vagy a PI-tranzitivitás fogalmait (↬37.o.). Mivel akkor még nem vezettük be a különböző preferenciarelációkat, így egyetlen relációra támaszkodva kellett megadnunk a definíciókat. Most már egyszerűsíthetünk a korábbi formulákon, ha az **R** gyenge preferencia fogalmából képzett **P** szigorú preferencia, **I** indifferenciarelációk meghatározásait is figyelembe vesszük.

(PL.58) $\text{PP-TRANSITIVE}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n \forall o((\mathbf{P}(m,n) \wedge \mathbf{P}(n,o)) \rightarrow \mathbf{P}(m,o))$

(PL.59) $\text{IP-TRANSITIVE}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n \forall o((\mathbf{I}(m,n) \wedge \mathbf{P}(n,o)) \rightarrow \mathbf{P}(m,o))$

(PL.60) $\text{PI-TRANSITIVE}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n \forall o((\mathbf{P}(m,n) \wedge \mathbf{I}(n,o)) \rightarrow \mathbf{P}(m,o))$

(PL.61) $\text{II-TRANSITIVE}(\mathbf{R}(m,n)) \stackrel{\text{def}}{=} \forall m \forall n \forall o((\mathbf{I}(m,n) \wedge \mathbf{I}(n,o)) \rightarrow \mathbf{I}(m,o))$

A későbbiekben még használjuk ezeket a tranzitivitási tulajdonságokat, illetve a különböző tézisekben, levezetésekben nem egyszer lehet hivatkozni azokra az összefüggésekre,

[48] A 'ceteris paribus' latin kifejezés jelentése: 'minden más változatlan'.

amelyek a tranzitivitás fogalomkörébe tartozó tulajdonságok között állapíthatók meg. Fel lehet tenni itt a kérdést, miért van szükség ilyen sokféle fogalomra a tranzitivitás fogalombokrán belül. Azért, mert a racionális emberi cselekvések magyarázatában igyekszünk kizárni az egymás után kimondott preferenciaállítások sorából a ciklikusság megjelenését. De miért is baj a ciklikusság? Erre a kérdésre Amos Tversky pénznyelő (money-pump) példájával szokás válaszolni. Ha nem követelnénk meg a tranzitivitást, akkor a viselkedésünket ugyanúgy irracionálisnak lehetne minősíteni, mint Tversky példaemberének cselekvését.

> „A tranzitivitás a racionális visekedés egyik alapvető és talán legszigorúbb elve. Mondhatjuk azt is, hogy aki megsérti a tranzitivitás elvét, az voltaképpen 'pénznyelőként' működik. Tegyük fel, hogy valaki preferálja y-t x-hez, z-t y-hoz, valamint x-t z-hez képest. Ekkor megalapozott lehet az a feltételezésünk, hogy az illető hajlandó fizetni azért, hogy kicserélhesse először x-t y-ra, aztán y-t z-re, végül z-t x-re. Ám ha így tesz, a végére ugyanoda jut, ahonnan indult, csak épp kevesebb lesz a pénze." [759, 45]

dominancia elv: [314]

A preferenciarelációkat (szigorú rendezés, gyenge rendezés, indifferencia) egymás után többször is lehet alkalmazni, vagyis sorba lehet őket kötni. Ilyen esetekben egyszerűsíthetünk a jelölésen, ha az egymást követő diadikus viszonyokat lineaárisan "összekötjük":

$$\text{(BL.62)} \qquad \mathbf{P}(m,n) \wedge \mathbf{P}(n,o) \wedge \mathbf{I}(o,p) \equiv PPI \equiv (m > n) \ \wedge (n > o) \wedge (o \shortparallel p) \equiv m > n > o \shortparallel p$$

Bármennyire is fontosnak és szükségesnek érezzük a tranzitivitást mind a preferencia, mind a racionális vilsekedés meghatározásához, az intranzitivitás, a ciklikusság gyakran tetten érhető a cselekvéseinkben. Sven Ove Hansson szerint négyféle tipikus tranzitivitási kör létezik, és az elemzésekben felhozott legtöbb példa ezek valamelyikébe tartozik [336]. A továbblépés előtt érdemes bizonyos jelölési konvenciókat bevezetni.

$\mathbf{I}^n\mathbf{P}$-CYCLE	egymás után n-szer **I**-, majd egy **P**-reláció követi egymást $(n \geq 2)$
IPIP-CYCLE	egy **I**-, egy **P**-, **I**-, **P**-reláció követi egymást
IPP-CYCLE	egy **I**-, majd két **P**-reláció követi egymást
$\mathbf{P}^n$-CYCLE	egymás után n-szer **P**-reláció követi egymást $(n \geq 3)$

Az $\mathbf{I}^n\mathbf{P}$-ciklus szemléltetésére a R. Duncan Luce példáját hívom segítségül [460]. Tegyük fel, hogy valaki jobban szereti a kávét egy kockacukorral inni, mint öttel (ami azt is jelenti, hogy képes megkülönböztetni az egy-cukros kávét az öt-cukrostól). Ezután képzeljünk el egy kísérletet, amelyben 401 csésze kávét kell összehasonlítania úgy, hogy az egymást követő i. kávékba rendre $\left(1 + \frac{i}{100}\right) x$ gramm cukrot tesznek (ahol x a kockacukor súlya). Nyilvánvaló, hogy ilyen kicsit mennyiségi különbségek mellett senki sem képes megkülönböztetni az egymást követő kávékban levő cukor mennyiségét, és így a kávé ízét ugyanolyannak érzi, ezért minden egymást követő kávépárost egyformának ítél (közömbös velük szemben). Ezt kell mondania az első összehasonlítástól kezdve az utolsóig. A közömbösségi reláció tranzitivitását feltételezve viszont bármelyik két kávéra is igaz kell legyen az, hogy ugyanolyannak tartja őket a kísérleti alany. Ezzel azonban az indifferens ítéletek sorának végére jutva ellentmondás keletkezik, hiszen az alany az első összehasonlítási körben az egy-cukros kávéhoz képest rosszabbnak tartotta az öt-cukros kávét,

míg a második – szomszédos páros – összehasonlítási körben az 1. és a 401. csészében levő kávékat ugyanolyannak ítélte meg, holott bennük ugyanannyi cukor (egy, illetve öt kockányi) van, mint az első körös összehasonlításban.

Az $\mathbf{I}^n\mathbf{P}$-ciklus legegyszerűbb változata az **IIP**-ciklus, amelyre Sven Ove Hansson Frank Restle 1961-es könyvéből idéz egy példát [336]. Restle olyan esetet ír le, amelyben valakinek három alternatíva közül kell választania úgy, hogy a három összehasonlításból kettő nehéz, egy pedig könnyű döntést jelent. A alternatívák a következők:

m	utazás Kaliforniába
n	utazás Floridába
o	utazás Kaliforniába + egy alma

Ebben a helyzetben mind az első és második, mind a második és harmadik opció közötti választás nehéz döntés (a plusz alma elhanyagolható előnyt jelent ebben az összehasonlításban, a döntés lényege ettől nem változik). Bármilyen kicsi és az alma értéke az utazások értékéhez képest, az első és harmadik opció öszehasonlításakor mégiscsak többletet jelent az utóbbi számára, tehát itt könnyen döntést lehet hozni. Igaz tehát, hogy **I**(m,n), **I**(n,o), illetve **P**(m,o), vagyis teljesül az **IIP**-ciklus. Lényegében ugyanezt a helyzetet vizsgálta Dan Ariely is abban az empirikus kísérletében, amelyben a résztvevőknek egy római és egy párizsi utazás között kellett választani, miközben harmadik opcióként felajánlották számukra a római utat egy ingyenreggelivel együtt [28], bár Ariely kérdésfeltevése és megközelítésmódja nem igazán az intranzitív ciklusok vizsgálatához kötődik. Az általa levont tanulság azonban itt is hasznos. A kísérletben a résztvevők gyakran választották az ingyenreggelivel járó római utat, és Ariely úgy értelmezi ezt a helyzetet, hogy mindez azért történt így, mert a harmadiknak kínált csali ajánlat segített "kimenekülni" abból a helyzetből, amelyben nem tudott igazából dönteni az ember két egyenragú opció között. Erre a kérdésre később, a racionalitás tárgyalásakor még visszatérek.

Folytatva az intranzitivitási ciklusok tárgyalását érdemes Restle korábbi példáján módosítanunk kicsit. Ha az utazással kapcsaolatos választási lehetőségeket kiegészítjük egy további alternatívával, akkor olyan helyzetethez jutunk, amellyel az **IPIP**-ciklust szemléltethetjük. Vegyük a következő négy alternatívát:

m	utazás Kaliforniába
n	utazás Floridába
o	utazás Floridába + egy alma
p	utazás Kaliforniába + egy alma

A négy választási lehetőségre nézve igaz lehet az, hogy az egyén nem tud dönteni Kalifornia és Florida között, de a plusz alma miatt a harmadik opciót választja a másodikkal szemben, aztán ha egy-egy almával "növelik" két város vonzerejét, akkor továbbra is döntetlen marad köztük, viszont az elsőhöz képest a negyedik alternatíva többet ér az ajándék alma miatt. Tehát fennállnak a következő rendezések: **I**(m,n), **P**(n,o), **I**(o,p), illetve **P**(m,p). Ez azt jelenti, hogy ez a példa már az **IPIP**-ciklus jelenségébe tartozik.

Az **IPP**-ciklusra Hansson Yew-kwang Ng példáját idézi [534]. Tegyük fel, hogy a döntési feladat abból áll, hogy papírlapcsomagok közül kell választani kétféle szempont

alapján. A vastagságot tekintve az első csomag papírja vastagabb, mint a másodiké, de az első és harmadik, valamint a második és harmadik csomag lapvastagságai között nincs érezhető különbség. Ekkor igaz a következő: $\mathbf{P}^v(m,n), \mathbf{I}^v(n,o), \mathbf{I}^v(o,p)$. A második szempont szerint, a papír fehérsége alapján a második csomag papír fehérebb a harmadiknál, de az első és harmadik, illetve az első és második között nem lehet különbséget tenni. Ebben a dimenzióban a következő rendezések élnek: $\mathbf{I}^f(m,n), \mathbf{P}^f(n,o), \mathbf{I}^f(o,p)$. Ha a két szempontot együttesen figyelembe véve akarunk döntést hozni, és azt a papírt keresnénk, amelyik a legvastagabb és legfehérebb, akkor gondba kerülnénk, mert a preferenciáink ciklusosak. Az vastagság dimenziójában az m-t, a fehérség szempontja szerint az y-t kéne választanunk, de a kettős szempont szerint nem lenne "győztes".

A negyedik ciklustípus, a $\mathbf{P}^n$-ciklus, amire Warren S. Quinn önkínzó önkéntes példáját hozhatjuk fel [585]. Quinn gondolatkísérletében olyan speciális szerkezetet szerelnek egy önkéntesre, amellyel fokozatosa, kis lépésekben lehet növelni az áramütés nagyságát. Az eszköz 1000 fokozatot képes kezelni. Az egyes fokozatokban okozott áramütés, illetve az így kiváltott fájdalom nagysága közötti különbséget nem lehet érezni, de minél távolabb van egymástól két fokozat, annál inkább érzékelhetővé válik a különbség, egyre nő a fájdalomérzet. A skála végén már elviselhetetlen mértékű fájdalmat képes kiváltani az eszköz. A kísérlet szereplője minden héten dönthet arról, hogy egy fokozattal növeli-e a szintet, vagy kiszáll a játékból. Minden növelésért jutalmat kap. A kísérletben minden egyes lépésre igaz, hogy az önkéntes jobbnak tartja az adott fokozatot a megelőzőnél, hiszen a két szomszédos fokozatban nem érzékelhető különbség az áramütések, illetve az okozott fájdalmak nagysága között. Ám minél messzebbre jut az első lépéstől, annál biztosabban mondhatjuk, hogy a "sokadik" fokozatot már nem találja jobbnak a legelsőhöz képest (hiszen ekkor már érzékelhetően nagy különbség van a fájdalmak mértékében). Azt kell mondanunk tehát, hogy a lépésenkénti összehasonlítás alapján ez a preferenciarendezés intranzitív. Ha viszont az egyes fokozatokat "összevonjuk", és nagyobb intervallumokat veszünk figyelembe, akkor megváltozik a helyzet. Ha például a 0., az 500. és az 1000. lépést hasonlítjuk össze, akkor egyértelmű, hogy az egyén azt fogja mondani, hogy preferálja az 500. fokozatot az 1000.-hez képest, a 0.-at az 500.-hoz képest, valamint a 0.-at az 1000.-hez képest, vagyis a rendezése ekkor és így tranzitívvá válik. Quinn kísérletete $\mathbf{P}^{1001}$-ciklusú intranzivitást mutat be, hiszen 1001 lépésen keresztül, minden egymást követő fokozatban $\mathbf{P}$-rendezésről van szó, míg a sorozat két végpontja között már nem áll fent ez a kapcsolat. Quinn példája persze nem igazán életszerű, de hogy ezzel ő is tisztában van, azt szépen jelzi az a tény, hogy a fenti gondolatkísérletén túl említést tesz teljesen hétköznapi jelenségekről is, amelyekben ugyanazt a jelenséget tapasztalhatjuk, mint az önkínzó önkéntes példájában. Minden egyes evésünket felfoghatunk úgy, mint egy kielégüléssel, élvezettel járó apró lépést, és egyik evéstől sem nő meg látható módon a súlyunk, de ha kicsit "messzebbről" vagyis nagyobb időszeletben nézünk a folyamatra, akkor már észrevehetjük a hízásunkat. De ugyanez igaz minden egyes cigi elszívása okozta élvezet és az esetleg megjelenő tüdőrák fájdalma vagy a mindennapi lustálkodásaink öröme és az elpocsékolt élet feletti kesergés között feszülő ellentmondásra is [585].

A bemutatott ciklusos jelenségekből két mozzanatot emelnék ki. Egyfelől a példákban láthatóan gyakran van fontos szerepe annak a ténynek, hogy az egyes lépések (alternatívák) között nm tudunk igazán különbséget tenni. Ilyen esetekben a **diszkriminációképességünk** alacsony foka (vagy másként: a diszkriminációképtelenségünk) magyarázhatja meg a jelenséget. Másfelől a fenti példák között "megjelent" egy olyan új je-

lenség is, amilyennel még nem foglalkoztunk eddig. Yew-kwang Ng példájában ugyanis egyértelműen az okozta a "problémát", hogy először két rendezési szempont szerint kellett (lehetett) ítélkezni, majd ezek után és ezek alapján kellett a végső eredményt kihozni. A két említett mozzanat nem teljesen független egymástól.[49] Az elméleti kapcsolat talán nem szükségszerű, de történetileg biztosan összefonódott a két dolog. Amikor Amos Tversky elemezte az intranzitivitás problémáját [759], és a cikkében felvetette a diszkrimináció problémakörét mint fontos magyarázó tényezőt, akkor – többek között – R. Duncan Luce-ra is hivatkozott. Utóbbi volt az, aki bevezette a szemirendezés (semiorder) fogalmát [460], ami egy preferenciarelációból és egy indifferencia-relációból áll az alábbi feltételek mentén.

(BL.63) SEMIORDER $\stackrel{\text{def}}{=}$

1) $\forall m \forall n (\mathbf{P}(m,n) \veebar \mathbf{P}(n,m) \veebar \mathbf{I}(m,n))$
1) $\forall m (\mathbf{I}(m,m))$
3) $\forall m \forall n \forall o \forall p ((\mathbf{P}(m,n) \wedge \mathbf{I}(n,o) \wedge \mathbf{P}(o,p)) \rightarrow \mathbf{P}(m,p))$
4) $\forall m \forall n \forall o \forall p ((\mathbf{P}(m,n) \wedge \mathbf{P}(n,o) \wedge \mathbf{I}(n,p)) \rightarrow \neg(\mathbf{I}(m,p) \wedge \mathbf{I}(o,p)))$

szemirendezés **P** a preferenciareláció, **I** az indifferenciareláció

Tversky megjegyezte, hogy szemirendezés a lexikografikus rendezésnek egyik típusát is jelenti egyben. Észre kell azonban vennünk, hogy ez a fogalom egy nagyon fontos ponton eltér az eddig rendezési fogalmainktól. Itt ugyanis nem egy, hanem két rendezési fogalomról (egy preferencia- és egy indifferenciarelációról) van szó, ami az "egyszerű" rendezésekhez képest egy eddig nem látott problémához vezet: ahhoz, hogy a kétféle rendezési relációtól végleges (valamiféle együttes) eredményt várhassunk el, az kell, hogy a "belső" relációk között is megállapítsunk valamilyen viszonyrendszert. Csak ennek létezésekor (és ennek ismeretében) lehetséges ugyanis az egyesített végeredmény elérése. A probléma kezelése, megoldása azonban már kivezet minket a preferencialogika szűkebb értelemben vett területéről, ezért itt most nem foglalkozom vele. Később természetesen még visszatérek erre a kérdésre.

Williamson hasonlósági relációja [820].

a négy-argumentumú $\mathbf{S}(m,n,o,p)$ hasonlósági relációt úgy értelmezhetjük, hogy 'm legalább annyira hasonlít n-re, mint amennyire o hasonlít p-re'.

(BL.64) SIMILARITY $\equiv \mathbf{S}(m,n,o,p) \equiv \mathbf{S} \stackrel{\text{def}}{=}$

S1) $\forall m \forall n \forall o \forall p (\mathbf{S}(m,n,o,p) \wedge \mathbf{S}(o,p,m,n))$
S2) $\forall m \forall n \forall o \forall p \forall s \forall t (\mathbf{S}(m,n,o,p) \wedge \mathbf{S}(o,p,s,t) \rightarrow \mathbf{S}(m,n,s,p))$
S3) $\forall m \forall n \forall o (\mathbf{S}(m,m,n,o))$
S4) $\forall m \forall n (\mathbf{S}(m,n,n,n) \rightarrow m = n)$
S5) $\forall m \forall n (\mathbf{S}(m,n,n,m))$

———————-

folyt.

———————-

[49] A kétféle szempont, és a köztük teremtendő "összehangolási kényszer" egyébként megjelenik Quinn példájában is.

1.12.3 Monadikus preferenciarelációk

von Wright megközelítéséhez képest Roderick M. Chisholm és Ernest Sosa hedonisztikus kiindulópontot választ magának:

> „ . . . we shall also assume that pleasure is the *only* thing that is intrinsically good and displeasure the *only* thing that is intrinsically bad." [154]

mit jelent a jobbnak tartani?

Saját hedonisztikus nézőpontjukon keresztül visszautasítják von Wright jóság-felfogását mondván, hogy a jóság minőségéhez szükséges a gyönyör, boldogság érzete, és nem elegendő szenvedés, boldogtalanság hiánya. Von Wright definíciójával szemben hozott ellenpéldájuk a következő:

m	nincsenek boldogtalan kócsagok
$\neg m$	vannak boldogtalan kócsagok

A fentiek alapján a von wrigthi rendezés nyilvánvalóan fennáll, de ettől még – Chisholmék szerint – az m állapotot nem tartjuk intrinzikusan jónak. Hasonló érvelés mentén utasítják el von Wright **B** rosszaság definícióját is.

Chisholmék elutasítják az a particionálást is, amely szerint $\mathbf{G} \underline{\vee} \mathbf{B} \underline{\vee} \neg\mathbf{G} \underline{\vee} \neg\mathbf{B}$ vagy másként $\mathbf{G} \underline{\vee} \mathbf{B} \underline{\vee} \mathbf{N}$, ahol **N** a monadikus indifferencia reláció

megintcsak a kócsagos példájukat hozzák megkülönböztetik az indifferenciát és a neutralitást és utóbbira már elfogadják ők is a hármas partíció létezését.

monadikus vagy diadikus, illetve abszolút vagy relatív preferencia utóbbi esetben az abszolút minősítés természetesen mást jelent, mint von Wright preferencialogikájában.

Sven Ove Hansson több cikkében is foglalkozott a monadikus preferenciafogalmakkal [332], [336].

legjobb, legrosszabb,

1.12.4 Preferencia és modális logika

von Wright 1972-es cikkében már két lehetséges világ értelmezést kínál az intrinzikus preferencia fogalmára [795, 146-147]. Lehetéges világ szemantika alapján: m-világ (illetve n-világ) azt jelenti, hogy abban a világban m (illetve n) fennáll.

C_i-világok (amikben nem áll fent mindkettő??????

Def$_1$: m $\forall$-preferált n-nel szemben C_i-világ feltételei közepette, ha minden olyan C_i-világ, amely egyben m-világ, de nem n-világ, preferált minden olyan C_i-világgal szemben, amely egyben n-világ, de nem t-világ

Def$_2$: m $\exists$-preferált n-nel szemben C_i-világ feltételei közepette, ha néhány olyan C_i-világ, amely egyben m-világ is preferált minden olyan C_i-világgal szemben, amely nem n-világ, egyben n-világ, de nem létezik olyan C_i-világ, amely t-világ is egyben, és preferált minden olyan C_i-világgal szemben, amely nem m-világ.

az előzmény Craig Boutilier preferencia [110]

Semantically, a goal can be viewed as a set of possible worlds, those states of affairs that satisfy the goal proposition. Intuitively, if we ignore considerations of ability, the set of goal worlds should be those considered most desirable by an agent (or its designer). To achieve all goals is to ensure that the actual world lies within this desirable set.

Ceteris paribus preferencia modális logikai alapon

Korábban már bevezettem a ceteris paribus preferencia fogalmát (↬112.o.), és jeleztem azt is, hogy a fogalom formális leírását modális logikai alapon lehet jól megoldani. Most érkeztünk el arra a pontra, amikor bemutathatom, hogyan lehet ezt a formalizálást elvégezni. Kiindulásként idézzük fel az amplifikáció elvének formuláját, amit von Wright összekapcsolt a ceteris paribus preferenciával (a formulából itt elhagyjuk a kvantorokat).

$$\mathbf{P}(m,n) \leftrightarrow (\mathbf{P}(m \wedge o, n \wedge o) \wedge \mathbf{P}(m \wedge \neg o, n \wedge \neg o))$$

A ceteris paribus elv azt akarja megragadni, hogy a preferenciáinkat úgy fejezzük ki két állapot között, hogy közben minden más állapotot változatlannak ítélünk meg. Először érdemes röviden értelmezni, mit is jelent ez a kijelentés. Az összehasonlítandó állapotok összességét nevezhetjük a tárgyalási univerzumnak (universe of discourse). Ebben a modellben a ceteris paribus elvet úgy kell értelmeznünk, hogy ha két állapotot (m-t és n-t) vetünk össze, akkor minden o állapotot, alternatívát a tárgyalási univerzumban, amelynek nincs direkt relevanciája az értékelés során, konstansnak tekintünk az összehasonlítási folyamat egészében. Vegyünk egy példát! Tegyük fel, hogy az esőköpenyt (r) preferáljuk az esernyővel (u) szemben uPr miközben harmadik potenciális alternatívaként ott van az az állapot is, hogy csizmát viselünk (b). Ekkor a tényleges, aktuális preferenciánkat így írhatjuk le: $r \wedge \neg bPu \wedge b$. Megváltozna az értékelésünk, ha elvesztenénk a csizmánkat, mert ez a tény átfordítaná a preferenciánkat, és az esőkabátot preferálnánk az esernyővel szemben, de az amplifikációs szabály elfogadása kizárja ezt a lehetőséget. A ceteris paribus elv érvényessége azt jelenti, hogy a csizma (r alternatíva) léte (vagy nem léte) nem befolyásolhatja a döntést, tehát mindvégig konstansnak kell tartanunk a hatását, és ezért el kell tekintenünk tőle.

Doyle-Wellman 1991 Desire = ceteris paribus preference? Hic et nun itt es most preferencia????

Ceteris paribus elv a kontextualis érvényességet, a preferenciák relativitását fejezi ki

Preferencia: állapotoktól propozíciókig

Fontos előzményként kell tekintenünk Jon Doyle és Michael Wellman cikkére, amelyben a szerzők először bontották fel két összetevőre a ceteris paribus preferenciák reprezentálásának feladatát [215]. Doyle és Wellman abból a praktikus szempontból fontos kérdésre keresste a választ, hogy miként lehetne a propozíciók között összehasonlításokat (és ezáltal preferenciarendezéseket) tenni. Doyle-ék számára azért volt fontos ez a kérdés, mert von Wright állapotok közti preferenciákról beszélt, amiket individuális propozíciókkal mindig le lehetett írni, de amikor a mérnökök a számítógépes döntéstámogató rendszerek fejlesztése során a céltételezésekben, eszközkiválasztásokban megbúvó preferenciákat is modellezni akarták, akkor hamar kiderült, hogy nem járható út az, amikor a preferenciákat az individuumok szintjén adják meg a rendszer számára, mert végtelen sok adat összegyűjtésére, betáplálására és számítására lenne szükség. Ehelyett persze kínálta magát az a lehetőség, hogy az individuális szintű adatok helyett általánosabb szintű adatokat kezeljenek, de itt újfajta gondok jelentkeztek.[50]

[50] Doyle és Wellman nem igazán választotta el a preferenciarendezéseket a másfajta rendezésektől, ezért sokszor egyszerűen csak összehasonlításokról beszéltek a cikkükben. Ez a mondanivalójuk érvényességén

Az individuumok összehasonlítása “egyszerű”. Ha Béla magasabb, mint Kati, vagy Hans jobban szereti a magasabb fizetést az alacsonynál, akkor erre könnyű képleteket, számítógépes programokat írni. De mi van akkor, ha individuumok valamilyen osztályát kell összehasonlítani. Például egy ruházati cég számára fontos lehet az olyan megállapítás, hogy a dánok magasabbak és soványabbak, mint a szicíliaiak, mert ennek ismeretében más méretű ruhákkal megy be a két piacra. Amikor ilyenfajta döntéseket kell hozzunk, akkor persze mindig találhatunk kivételeket, de ezt tudva is törekszünk arra, hogy ne individuumokra, hanem individuumok egész osztályára tudjunk összehasonlító állításokat tenni. Vannak tehát az individuumok (ahol könnyen tudunk összehasonlításokat tenni), és vannak az individuumok osztályát leíró propozíciók, ahol – egyelőre – még nem megoldott az összehasonlítás.

Első lépésként megpróbálhatjuk az individuumokra vonatkozó összehasonlítást (preferenciarendezést) az individuumok valamely osztályára vonatkozó összehasonlítássá alakítani a **sima felemelés** (simple lifting) módszerével. Tegyük fel, hogy az alábbi egyedeink vannak egy adott Ω halmazban: $m, n \in \Omega$, és ezek a következő osztályokba tartoznak: $m \in \phi, n \in \psi$. Ekkor nyilván igaz, hogy: $\phi, \psi \subseteq \Omega$. Ekkor az egyedek közti a már jólismert preferenciarendezési műveleteket értelmezhetjük: a $\succeq$ legalább olyan jó, a $\succ$ szigorúan jobb és az $\sim$ indifferens relációt. Ezek alapján – a sima felemelés módszerével – definiálhatjuk az osztályok (propozíciók) közti rendezéseket is.

(BL.65)
$$1)\ \phi \succeq \psi \stackrel{\text{def}}{=} \forall m \in \phi \forall n \in \psi (m \succeq n)$$
$$2)\ \phi \succ \psi \stackrel{\text{def}}{=} \forall m \in \phi \forall n \in \psi (m \succ n)$$
$$3)\ \phi \sim \psi \stackrel{\text{def}}{=} \forall m \in \phi \forall n \in \psi (m \sim n)$$

Ezzel a megoldással azonban komoly gondok adódnak, amit akkor érthetünk meg igazán, ha egy konkrét példán keresztül szembesülünk az elképzelés nem kívánt következményeivel. Ehhez először definiálnunk kell pár további feltételt. Elsőként rögzítenünk kell azt, hogy az összehasonlítandó szempontokat (propozíciókat) logikailag függetlennek kell tartani egymástól. Ha ezt nem követelnénk meg, akkor az összehasonlítások során logikai ellentmondásokhoz juthatnánk (erről már volt szó korábban).[51] A logikai függetlenséghez az szükséges, hogy a propozíciók és komplementereik összes lehetséges metszetében legyen előfordulás, a dolgok szemléltetéséhez elégséges a szemifüggetlenség – némileg gyengébb – fogalma is, amikor nem kell az összes lehetséges kombinációra megkövetelni a nem-ürességet.

(BL.66) $$\forall i \forall j (\phi_i \cap \phi_j \neq \varnothing \wedge \phi_i \cap \overline{\phi}_j \neq \varnothing \wedge \overline{\phi}_i \cap \phi_j \neq \varnothing \wedge \overline{\phi}_i \cap \overline{\phi}_j \neq \varnothing)$$

logikai függetlenség (HUN) – logical independence (ENG)

(BL.67) $$\forall i \forall j (\phi_i \cap \overline{\phi}_j \neq \varnothing \wedge \overline{\phi}_i \cap \phi_j \neq \varnothing)$$

logikai szemifüggetlenség (HUN) – logical semi-independence (ENG)

A logikai szemifüggetlenség alapján definiálhatunk háromféle legnagyobb rendezési fogalmat, amiket a következő formulákkal fejezhetünk ki.

semmit sem változtat, és ha nem zavaró, olykor én is átveszem a nyelvezetüket. Ebben a kontextusban tehát az összehasonlítás műveletét a preferenciarendezés értelmében használjuk, még akkor is, ha nyilvánvalóan másfajta – például magasság szerinti – összehasonlításról van szó.

[51] Megkeresni és hivatkozni a megfelelő helyre !!!!!!!!!!

(BL.68) 1) $\forall i((\phi_i \geq \overline{\phi}_i \wedge \forall k \forall l(\phi_k \overline{\phi}_l \neq \varnothing)) \rightarrow \forall m \in \phi_i \forall n \in \overline{\phi}_i (m \geq n))$
2) $\exists i(\phi_i \geq \overline{\phi}_i \rightarrow \forall m \in \phi_i \forall n \in \overline{\phi}_i (m \geq n))$
3) $\exists i \exists j(\phi_i \geq \overline{\phi}_i \rightarrow \forall m \in \phi_i \forall n \in \overline{\phi}_j (m \sim n))$

(no comparatives theorem)

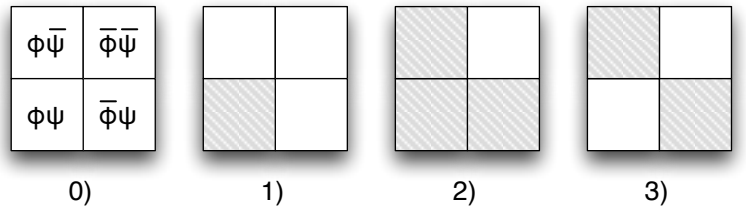
A legnagyobb minősítésnek fent meghatározott három fokozatának Doyle-ék a következő nveket adták:

- maximálisan legnagyobb (maximal greatness)
- közepesen legnagyobb (intermediate greatness)
- minimálisan legnagyobb (minimal greatness)

Jobban megérthetjük mind a fenti definíciókat, mind a bemutatni kívánt problémát, ha konkrét Doyle és Wellman a következő példát hozta. Vegyünk két szempont, a fizetést és a napi munkaidő hosszát, és ezek alapján ítéljük meg a választható munkahelyeket. Külön-külön mondhatjuk, hogy a magas fizetés jobb, mint az alacsony, illetve a rövidebb munkaidő jobb, mint a hosszabb, ezek alapján egyértelmű összehasonlításokat tudunk tenni. De gondba kerülünk, ha a két szempontot egyszerre akarjuk figyelembe venni. Egy ábra segítségével megvilágíthatjuk, miért. Mivel egy szemponton belül kizáró vagylagos választható lehetőségek vannak, az alternatívák így egymás ellentettjeinek, komplementereinek számítanak. Ezt figyelembe véve jelöljük a választási lehetőségeket a következő módon.

ϕ	– magasabb fizetés	$\overline{\phi}$	– alacsonyabb fizetés
ψ	– rövidebb munkaidő	$\overline{\psi}$	– hosszabb munkaidő

Mivel a két szempont független egymástól, ezért a szemifüggetlenség feltétele fennáll. Ekkor a $(\phi\psi, \phi\overline{\psi}, \overline{\phi}\psi, \overline{\phi}\overline{\psi})$ kompozíciósorozat teljes partíciót alkot, vagyis egymást kölcsönösen kizárják. Ezt a következő módon ábrázolhatjuk.



A választható konkrét munkahelyeket mindig besorolhatjuk az ábra 0) jelű négyzetének valamelyik negyedébe, az 1)-2)-3) jelű négyzetek pedig azt mutatják, hogy a korábban definiált három rendezési reláció melyik negyede(ke)t jelöli ki a legjobbnak. Az 1) négyzet jelöli a maximálisan legnagyobb rendezés eredményét, amikor azokat a konkrét egyedeket soroljuk ide, amelyek mindkét szempont szerint jobbak voltak. Ezek azok a munkahelyek, amelyek magas fizetést adnak és rövid munkaidejűek (vagyis mindkét szempont szerint jobbak). A 2) négyzet a közepesen legnagyobb rendezésnek felel meg. Ez az összehasonlítás azokat az individuumokat minősíti jobbnak, amelyek legalább egy szempont szerint jobbak a többinél. Ilyenek lesznek mindazok a munkahelyek, amelyek vagy magasabb fizetést, vagy rövidebb munkaidőt kínálnak, vagy mindkettőt egyszerre nyújtják. Ez jóval gyengébb feltétel az előzőhöz képest, így ennek a kritériumnak sokkal

több munkahely képes megfelelni. Az ábra 3) jelű négyzete mutatja, hogy mit adna vissza a minimálisan legnagyobb rendezés. Ekkor a két egyenlően legnagyobb állapotot jelent az, amikor az egyik szempont szerint jobbak, a másik szerint rosszabbak az alternatívák, illetve fordítva.

Ezzel a rendezési hármas javaslattal az a legnagyobb probléma, hogy nem ad módot mindenféle más összehasonlításra. Nem kerülhetünk olyan helyzetbe például, amikor a magasabb fizetésű és hosszabb munkaidejű munkahelyet preferálnánk az alacsonyabb fizetéssel és rövidebb munkaidővel szemben. Márpedig olykor kívánatos lenne. Például akkor, amikor a ceteris paribus elvet szeretnénk modellezni. Valahogyan gyöngíteni kell a feltételeken, hogy be tudjuk vezetni a ceteris paribus elvet. Doyle és Wellman újítása az volt ezen a ponton, hogy bevezették a kontextuális összehasonlítás (contextual comparison) fogalmát, amit elvezethet minket a ceteris paribus összehasonlítás lehetőségéhez. Ha azt mondjuk, hogy két individuum 'ugyanaz, miközben minden más változatlan' (same, other things equal), akkor ez két lépésben definiálható. Elsőként bizotsítani kell azt, hogy valamik ugyanazok, amit az individuumok közötti ekvivalenciával ragadhatunk meg, majd azt kell valahogy megfogni, hogy minden más változatlan. Ehhez kell a kontextuális ekvivalencia fogalma, amire már "ráülhet" a ceteris paribus összehasonlítás (comparison ceteris paribus). A kontextuális ekvivalenciát Doyle-ék az alábbi módon definiálták.

(BL.69) $\text{CONTEXTUAL_EQUIVALENCE}(m,n) \equiv m \overset{\text{c}}{\equiv} n \bmod \phi, \psi, \ldots \overset{\text{def}}{=} \wp(\wp(\Omega)) \mapsto \varepsilon(\Omega)$

kontextuális ekvivalencia (HUN) – contextual equivalence (ENG)

Az individuumok közti ekvivalenciarelációk segítségével ragadhatjuk meg a propozíciókat, hiszen a propozíciók (vagy azok komplementerei) alá sorolható egyedek mind ekvivalensek egymással ebben a minőségükben. A kontectuális ekvivalencia fogalmának segítségével már definiálni lehet a gyengén nagyobb ceeteris paribus rendezési relációt.

(BL.70) $\text{WEAKLY_GREATER_CP}(\phi,\psi) \equiv \phi \trianglerighteq \psi \overset{\text{def}}{=} \forall m \in \phi\overline{\psi} \forall n \in \overline{\phi}\psi(m \succeq n \wedge (m \overset{\text{c}}{\equiv} n \bmod \psi\overline{\psi}, \overline{\phi}\psi))$

gyengén nagyobb ceteris paribus (HUN) – weakly greater ceteris paribus (ENG) – gyengén jobb ceteris paribus (HUN)

Miért is érdekes ez? A definíció azokat az individuumokat hasonlítja össze, amelyekre fennáll, hogy $\phi\overline{\psi}$-be, illetve a $\overline{\phi}\psi$-be tartoznak, nem pedig azokat, amelyekre az lenne igaz, hogy a ϕ-be, illetve a ψ-be tartoznának. Azért értelmes ez így (és nem értelmes másként), mert az az összehasonlítás, amely során mindkét individuum bene lenne ϕ-ben, nem tudna köztük releváns módon különbséget tenni csak az alapján, hogy kielégítik-e ψ-t (hiszen ϕ-t mindketten kielégítik). Mivel fordítva is hasonlót mondhatnánk, vagyis ha az összehasonlítandó egyedek benne lennének a ψ-ben, akkor azokat csak a ϕ alapján lehetne összehasonlítani. Ezért csak a definícióban alkalmazott összehasoníltási mód végezheti el a két propozíció közti összehasonlítást.

(BL.71) $\text{STRICTLY_GREATER_CP}(\phi,\psi) \equiv \phi \triangleright \psi \overset{\text{def}}{=} \phi \trianglerighteq \psi \wedge \neg(\psi \trianglerighteq \phi)$

erősen nagyobb ceteris paribus (HUN) – strictly greater ceteris paribus (ENG) – szigorúan jobb ceteris paribus (HUN)

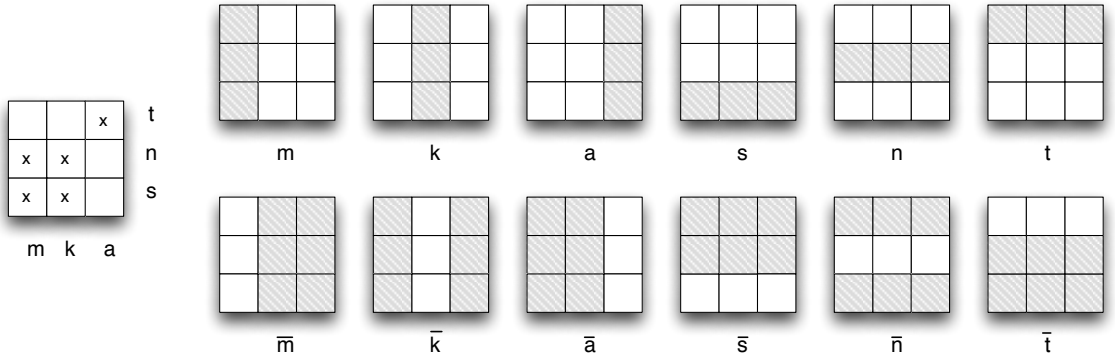
(BL.72) $\text{EQUALLY_GREAT_CP}(\phi,\psi) \equiv \phi \triangleq \psi \overset{\text{def}}{=} \phi \trianglerighteq \psi \wedge \psi \trianglerighteq \phi$

egyenlően nagy ceteris paribus (HUN) – equally great ceteris paribus (ENG) – ugyanolyan jó ceteris paribus (HUN)

Nézzünk meg egy újabb példát: hasonlítsuk össze az embereket a testmagasságuk, illetve testsúlyuk alapján. Soroljuk be úgy az embereket, hogy mindkét dimenzióban legyen három-három fokozatunk. A tulajdonságfokozatokat, illetve ezek komplementerét mind egy-egy proporzíciónak tarthatjuk, és az alábbi módon jelölhetjük.

m = magas	$\overline{m}$ = közepes + alacsony	s = sovány	$\overline{s}$ = normál + túlsúlyos
k = közepes	$\overline{k}$ = magas + alacsony	n = normál	$\overline{n}$ = sovány + túlsúlyos
a = alacsony	$\overline{a}$ = magas + közepes	t = túlsúlyos	$\overline{t}$ = sovány + normál

A két tulajdonság fokozatait preferenciasorrendbe rakhatjuk, a magasság szerint a magastól az alacsony felé, a súly szerint a soványtól a túlsúlyos felé csökkenő rendezéssel. A bevezetett jelölésünket használva legyen érvényes a következő két sorrend: $m \trianglerighteq k \trianglerighteq a$, illetve $s \trianglerighteq n \trianglerighteq t$. Végül tételezzük fel még azt, hogy van 5 személyünk, akikre vonatkozóan a következők igazak: van egy magas és sovány, egy közepesen magas és sovány, egy magas és normál súlyú, egy közepesen magas és normál súlyú, végül egy alacsony és túlsúlyos emberünk. Mindezeket az információkat a következő ábra segítségével szemléltethetjük vizuálisan.



A kérdés, hogy mit mondhatunk a ceteris paribus preferenciákról ebben az esetben. Nos, két állítást tehetünk. Mondhatjuk, hogy egyfelől a magas + sovány gyengén jobb ceteris paribus a magas + normál súlyúhoz képest, másfelől a közepes + sovány gyengén jobb ceteris paribus a közepes + normál súlyúhoz képest, vagyis: $m \wedge s \geq m \wedge n$, illetve $k \wedge s \geq k \wedge n$.

Benthemék

Johan van Benthem, Patrick Girard és Olivier Roy szerint a feltétel-nélküliséget modális logikai alapon kell inkább megragadni, ezért kidolgozták a ceteris paribus preferencia $\mathscr{L}_{\wp}$ modális nyelvét és modeljét [771].

vannak állapotok: s, t

és vannak propozíciók: φ, ψ

A $\preceq$ jobb reláció **előrendezés** (↬39.o.)

erősen összefüggő (totális) előrendezés

a propozíciók közötti preferenciaoperátorok definiálása az állapotok közt érvényes jobb reláció segítségével:[52]

[52] Itt áttérek a infix-jelölésmódra .....

(BL.73) $\leq_{\exists\exists} \stackrel{\text{def}}{=} \mathfrak{M}, u \vDash \varphi \leq_{\exists\exists} \psi \stackrel{\text{def}}{\Leftrightarrow} \exists s, \exists t : \mathfrak{M}, s \vDash \varphi \wedge \mathfrak{M}, t \vDash \psi \wedge s \preceq t$

létezik olyan ψ-állapot, amely legalább olyan jó, mint néhány φ-állapot (HUN) – there is a ψ-state that is at least as good as some φ-state (ENG)

(BL.74) $\leq_{\forall\exists} \stackrel{\text{def}}{=} \mathfrak{M}, u \vDash \varphi \leq_{\forall\exists} \psi \stackrel{\text{def}}{\Leftrightarrow} \forall s, \exists t : \mathfrak{M}, s \vDash \varphi \rightarrow \mathfrak{M}, t \vDash \psi \wedge s \preceq t$

minden φ-állapotra létezik legalább egy legalább olyan jó ψ-állapot (HUN) – for every φ-state, there is a ψ-state that is at least as good (ENG)

(BL.75) $<_{\exists\exists} \stackrel{\text{def}}{=} \mathfrak{M}, u \vDash \varphi <_{\exists\exists} \psi \stackrel{\text{def}}{\Leftrightarrow} \exists s, \exists t : \mathfrak{M}, s \vDash \varphi \wedge \mathfrak{M}, t \vDash \psi \wedge s \prec t$

létezik olyan ψ-állapot, amely szigorúan jobb, mint néhány φ-állapot (HUN) – there is a ψ-state that is strictly better than some φ-state (ENG)

(BL.76) $<_{\forall\exists} \stackrel{\text{def}}{=} \mathfrak{M}, u \vDash \varphi <_{\forall\exists} \psi \stackrel{\text{def}}{\Leftrightarrow} \forall s, \exists t : \mathfrak{M}, s \vDash \varphi \rightarrow \mathfrak{M}, t \vDash \psi \wedge s \prec t$

minden φ-állapotra létezik legalább egy szigorúan jobb ψ-állapot (HUN) – for every φ-state, there is a strictly better ψ-state (ENG)

(BL.77) $<_{\forall\forall} \stackrel{\text{def}}{=} \mathfrak{M}, u \vDash \varphi <_{\forall\forall} \psi \stackrel{\text{def}}{\Leftrightarrow} \forall s, \forall t : \mathfrak{M}, s \vDash \varphi \wedge \mathfrak{M}, t \vDash \psi \wedge s \prec t$

minden φ-állapotra minden ψ-állapot szigorúan jobb (HUN) – for every φ-state, every ψ-state is strictly better (ENG)

(BL.78) $>_{\exists\forall} \stackrel{\text{def}}{=} \mathfrak{M}, u \vDash \varphi >_{\exists\forall} \psi \stackrel{\text{def}}{\Leftrightarrow} \exists s, \forall t : \mathfrak{M}, s \vDash \varphi \wedge \mathfrak{M}, t \vDash \psi \wedge t \prec s$

létezik olyan φ-állapot, amely szigorúan jobb minden ψ-állapotnál (HUN) – there is a φ-state that is strictly better than every ψ-state (ENG)

(BL.79) $\leq_{\forall\forall} \stackrel{\text{def}}{=} \mathfrak{M}, u \vDash \varphi \leq_{\forall\forall} \psi \stackrel{\text{def}}{\Leftrightarrow} \forall s, \forall t : \mathfrak{M}, s \vDash \varphi \wedge \mathfrak{M}, t \vDash \psi \wedge s \preceq t$

minden ψ-állapot szigorúan jobb, mint minden φ-állapot (HUN) – every ψ-state is strictly better than every φ-state (ENG)

(BL.80) $\geq_{\exists\forall} \stackrel{\text{def}}{=} \mathfrak{M}, u \vDash \varphi \geq_{\exists\forall} \psi \stackrel{\text{def}}{\Leftrightarrow} \exists s, \forall t : \mathfrak{M}, s \vDash \varphi \wedge \mathfrak{M}, t \vDash \psi \wedge t \preceq s$

létezik olyan φ-állapot, amely legalább olyan jó, mint minden ψ-állapot (HUN) – there is a φ-state that is at least as good as every *psi*-state (ENG)

definálni az első négy operátort $\mathscr{L}_{\wp}$ nyelven

(BL.81) $\varphi \leq_{\exists\exists} \psi \stackrel{\text{def}}{=} E(\varphi \wedge \Diamond^{\leq}\psi)$

(BL.82) $\varphi \leq_{\forall\exists} \psi \stackrel{\text{def}}{=} A(\varphi \rightarrow \Diamond^{\leq}\psi)$

(BL.83) $\varphi <_{\exists\exists} \psi \stackrel{\text{def}}{=} E(\varphi \wedge \Diamond^{<}\psi)$

(BL.84) $\varphi <_{\forall\exists} \psi \stackrel{\text{def}}{=} A(\varphi \rightarrow \Diamond^{<}\psi)$

Amennyiben feltételezzük, hogy a $\preceq$ jobb reláció erősen összefüggő (totális), akkor a másik négy preferenciaoperátort is definiálhatjuk az alábbi formulák segítségével:

(BL.85) $\text{STRONGLY_CONNECTED}(\preceq) \rightarrow (\varphi <_{\forall\forall} \psi \stackrel{\text{def}}{=} E(\varphi \wedge \Box^{\leq}\neg\psi))$

(BL.86) $\text{STRONGLY_CONNECTED}(\preceq) \rightarrow (\varphi >_{\exists\forall} \psi \stackrel{\text{def}}{=} A(\varphi \rightarrow \Box^{\preceq}\neg\psi))$

(BL.87) $\text{STRONGLY_CONNECTED}(\preceq) \rightarrow (\varphi \leq_{\forall\forall} \psi \stackrel{\text{def}}{=} E(\varphi \wedge \Box^{\prec}\neg\psi))$

(BL.88) $\text{STRONGLY_CONNECTED}(\preceq) \rightarrow (\varphi \geq_{\exists\forall} \psi \stackrel{\text{def}}{=} A(\varphi \rightarrow \Box^{\prec}\neg\psi))$

Von Wright preferencialogikája a $<_{\forall\forall}$ modális operátor segítségével írható le.

Ha formalizálni akarjuk az $\Lambda^{\mathscr{L}_{\wp}}$ logikát a $\mathscr{L}_{\wp}$ nyelven, akkor a következő öt axiómát kell rögzítenünk.

S4 logika kell a $\Diamond^{\preceq}$, K logika kell a $\Diamond^{\prec}$ és az S5 logika kell a E globális egzisztenciális modalitáshoz.

(BL.89) $\text{INCLUSION}_1 \vdash \Diamond^{\prec}\varphi \rightarrow \Diamond^{\preceq}\varphi$

(BL.90) $\text{INTERACTION}_1 \vdash \Diamond^{\preceq}\Diamond^{\prec}\varphi \rightarrow \Diamond^{\prec}\varphi$

(BL.91) $\text{INTERACTION}_2 \vdash \varphi \wedge \Diamond^{\preceq}\psi \rightarrow (\Diamond^{\prec}\psi \vee \Diamond^{\preceq}(\psi \wedge \Diamond \preceq <\varphi))$

(BL.92) $\text{INTERACTION}_3 \vdash \Diamond^{\prec}\Diamond^{\preceq}\varphi \rightarrow \Diamond^{\prec}\varphi$

(BL.93) $\text{INCLUSION}_2 \vdash \Diamond^{\preceq}\varphi \rightarrow E\varphi$

levezetési szabályok:

(BL.94) MP

(BL.95) NECESSITATION

(BL.96) UNIFORM SUBSTITUTION

The logic $\Lambda^{\mathscr{L}_{\wp}}$ is sound and complete with respect to the class of preference models .....

ceteris paribus preferencia két értelemzési lehetősége:
1) normalitási értelem: minden más normális (all other things being normal)
2) változatlansági értelem: minden más változatlan (all other things being equal)

A normalitási értelemben vett CP-elvre példa: valaki preferálja a vörös bort a fehérrel szemben, kivéve, ha halat eszik, mert akkor a fehéret preferálja. A vörös bor a tipikus preferencia, a fehér bor az atipikus preferencia, amely adott feltételek mellett a tipikus helyébe lép (a legyőző – defeater

későbbiek

Meghyn Bienvenu, Jérôme Lang és Nic Wilson kidolgoztak egy olyan általános modális nyelvet és modellt a preferenciára, amely a filozófusok által kidolgozott preferencialogikákat, és mesterséges intelligencia világán belül kifejlesztett preferenciareprezentációs nyelveket egyaránt képes integrálni [84]. Ennek – más egyéb okok mellett – azért érezték szükségét, mert figyelemmel akartak lenni az adott nyelven történő gépi következtetések számításigényeire is.

> „However ... adopting the full expressivity of $\mathscr{L}_{CP}$ also means accepting an increase in the computational complexity of reasoning.
>
> In this paper we make a trade-off between complexity and expressivity, and choose to work with our simple, yet still fairly expressive, preference logic PL, which covers and generalizes many AI preference languages as well as preference logics (including von Wright's), but also (we believe) will be sufficiently expressive in most practical situations." [771]

Preferencia és változás

A preferencia általános fogalmára nézve különösen igaz, de még az intrinzikus preferenciára vonatkoztatva is kijelenthető, hogy a preferenciáink nem statikusak. Az extrinzikus preferencia együtt változhat azzal, ahogy valamilyen hitünk, ismeretünk megváltozik a világról. Ez a hatás az intrinzikus preferencia esetében is jelentkezhet, bár nyilván lassabban, kevésbé reflektált módon. De a preferenciáink másfajta külső hatásnak is ki vannak téve. Kaphatunk parancsokat, megfogalmazhatnak morális elvárásokat velünk szemben, új információhoz juthatunk, új összefüggéseket ismerhetünk fel, javaslatokat, tanácsokat adhatnak nekünk, melyek eredményeként változhatnak a világra vonatkozó értékeléseink.

conditionals

Johan van Benthem, Sieuwert van Otterloo and Olivier Roy tanulmánya [768]

$\mathfrak{M} = \langle W, N, \{\preceq_i\}_{i \in N}, V \rangle$

$\preceq_i$ előrendezés (↬39.o.)

igazságdefiníció:

$\mathfrak{M}, w \vDash \Diamond_i \varphi \overset{\text{def}}{\Leftrightarrow} \exists w' \left(w \preceq_i w' \right) \wedge \mathfrak{M}, w' \vDash \varphi$

preferencia upgrade

Johan van Benthem és Fenrong Liu ‘Dynamic logic of preference upgrade’ című cikke [775].

javaslat (suggestion) parancs (order)

más a helyzet az indifferencia és az össszehasonlíthatatlanság esetén

update upgrade

$\mathbf{K}_i \varphi$-vel jelöljük, hogy az i ágens tudja ($\mathbf{K}$), hogy φ.

Liu

Fenrong Liu több cikke is a preferenciák változásáról szól könyve revisited cikke de Jongh és Liu 2009-es cikke

1) dynamic betterness language

suggestion upgrade

betterness change programs

2) priority language, priority dynamics

bevezeti a prioritás fogalmát, tehát ez már másodrendű preferencia fogalmat használ és az aggregációs témakörbe tartozik

> „ . . . extrinsic preference as given by priority orders of propositions that encode relevant criteria for comparing worlds or objects.”

támaszkodnak ARS prioritás modellre: két alapoperátor:

sequential composition G;G? adds the graph G on top of G? in the order: all nodes in the first come before all those in the second,

the parallel composition G||G? is the disjoint union of the graphs G and G?, without any order links between them.

megadják a két ARS-operátor segítségével hogyan lehet a prioritás-gráfot módosítani (update). The set ?(G , A) of basic graph updates is defined by: ?(G,A) := $A|G_1; G_2|G1||G_2$.

Sven Ove Hansson négy operátort definált [335]

———-

For better or for worse: Dynamic logics of preference: [770]

megjegyzések:

állapotokra vonatkoztatja a preferenciarelációt, amelyek nem állítások, tehát a tagadás helyett inkább komplementálást kell használnunk.

Johan van Benthem és szerzőtársai erre azt mondják, hogy állapotokra propozícionális változókat használnak, ami a [771]

———-

1.12.5 További elméletek

Warren S. Quinn [584]

modern elméletek, mérnökök közt: [613]

Kanger is írt egy rövid cikket a preferencialogikáról, amiben Wright elméletét is kritizálta

Dragan Cvetković áttekintése, programjai,

Rescher

extrinzikus preferencia, preferenciaváltozás: Reasoning about Preference Dynamics [455]

A von wrighti preferencialogikáról Fenrong Liu: [454]

Preference Change: Approaches from Philosophy, Economics and Psychology [310]

utility function

hasznosságfüggvény

valós értéket lehet a preferátumokhoz rendelni folytonos és monoton (?) függvényt lehet képezni

1.13 Deontikus logika

A deontikus logika (vagy imperatív logika) az előírások, kötelezések, tiltások, megengedések logikai formalizálásával foglalkozik. Nem célom bemutatni itt a deontikus logika kialakulásának, fejlődésének történetét, nem akarok foglalkozni az elmélet legfontosabb eredményeivel sem. A témáról szóló könyvek, tanulmányok közül íme egy elfogult és teljesnek messze nem mondható lista: [611], [791], [437], [15], [331], [330], [143], [144], [393], [654], [467], magyarul olvasható Solt Kornél munkája [700].

A cselekvéselmélet számára a deontikus logika mint a modális logika egyik ága a leginkább használható. Azért van ez így, mert az előírások mindig valamilyen cselekvésre vonatkoznak, tehát a cselekvéselmélet központi fogalma, a cselekvés van a deontikus logika fókuszában is. A modális logika különböző kalkulusait meghatározó sémák közül a (D) séma az, amit a deontikus logikához kapcsolhatunk.

$$\text{(D)} \quad \Box A \rightarrow \Diamond A$$

A deontikus logikában a két modális operátor, a ‘szükségszerű’ és a ‘lehetséges’ átértelmezhető úgy, hogy ‘kötelező’, illetve ‘megengedett’ legyen a jelentésük. Ekkor a (D) sémát úgy lehet interpretálni, hogy ‘ha valami kötelező, akkor az megengedett is egyben’. A deontikus logika mint a kötelezések logikája ‘arról szól’, hogy mit tudunk mondani a jövőről, pontosabban – a jövőre vonatkozó – kötelezések és megengedések világáról. Minden kötelezés, tiltás, megengedés a “jövőre irányul”, hiszen az ilyen helyzetek ágensei (vagy legalábbis egy részük) azzal a várakozással fordul az előírások felé, hogy azok a jövőben teljesülni fognak. Ez a jövőre vonatkozás adja a deontikus logika egyik specialitását, hiszen ez azt jelenti, hogy ez a logikai szakterület nem tényekkel foglalkozik. Ettől még sok minden fontos felismerést, tudást köszönhetünk neki, de azt a hiányosságát, hátrányát mindenképpen rögzíteni érdemes, miszerint ez a diszciplina nem tud azonban semmi értelmeset mondani arról, hogy mikor, miért valósulnak meg azok a cselekvések, amelyekre az előírások irányulnak.[53]

Deontikus operátorok

A deontikus logika első elméletében von Wright négy deontikus operátorról írt [796], és az utána következő szerzők is kitartottak amellett, hogy a deontikus operátorok a logikai négyszögre “húzhatók rá”. Persze, ebben a diskurzusban is “felbukkantak” ugyanazok a problémák, amelyek a logikai négyszög más területeken való alkalmazásakor jelentkeztek (↬52.o.). A **kötelező** és a **tiltott** univerzális kategóriákkal nem nagyon voltak problémák, de bizonytalanság, pontatlan fogalmi határok jellemezték a partikuláris kategóriákat. A **megengedett** értelmezhető volt mind az állító, mind a tagadó partikuláris kategóriákra, bár intuitívabbnak tűnt az állító jellegű használata, és olykor “nehézséget okozott” a négyszög (O)-sarokjának megnevezése, illetve pontos értelmezése. A fontosságuk miatt sok nyelvi forma termelődött ki a kategóriák jelölésére. Íme egy – biztos nem teljes – lista magyarul:

kötelező	megparancsolt, kellő, támogatott, obligát, ajánlott, muszáj, szükséges, nem mellőzhető, nem lehet nem megtenni
tilos	tiltott, nem lehet, nem lehetséges, nem szabad, nem megengedett, megengedhetetlen, tűrhetetlen, nem tűrt, elfogadhatatlan, kötelező tartózkodni
megengedett	szabad, lehet, lehetséges, jogos, nem tilos, nem tiltott, tűrt, elfogadott, tolerált
mellőzhető	szabad, lehet, lehetséges, jogos, nem kötelező, elhagyható, elmulasztható, eltekinthető, eltekintett, felmentett, tolerált

A használt kifejezések listájából látszik, hogy a természetes nyelvi szintre (is) jellemző a terminológiai túlhasználat, a pongyolaság. A nyelv rugalmassága miatt egyes kategóriák több fogalomhoz köthetően is feltűnnek, a tagadás különböző nyelvi formáin keresztül többszörösen újrahasznosítjuk az alapterminusainkat. A terminológiai gazdagság (és pongyolaság) vélhetőleg minden nyelven tetten érhető. Angolul is “sokszínűséget” tapasztalhatunk (és ez a lista is nyilván bővíthető lenne):

obligatory	commanded, duty, compulsory, required
forbidden	impermissible, banned, prohibited, prohibition, disallowed
permitted	permissible, allowed, right, free, claim
omissible	non-obligatory, no-duty, duty-free, gratuitous, exemption

[53] Erről a kérdésről bővebben írtam: [721].

Csakúgy, mint a kategoriális logika bármely más területén, a legnagyobb problémát itt is a partikuláris kategóriák jelentetik. A **megengedett** fogalmát ugyanis a hétköznapi/jogi kommunikáció során többféle értelemben is használjuk. Induljunk ki Jan C. Joerden szemléletes példájából [381, 204]. Két mondatban is ugyanazt a 'megengedett' terminust használjuk, miközben más és más jelentés tulajdonítunk neki.

(1) Ha valaki egy közúti baleset helyszínére ér, megengedett számára, hogy segítsen.

(2) Ha valaki megszerezte a vezetői engedélyét, megengedett számára, hogy gépjárművet vezessen.

Az első mondatra azonnal rá lehet vágni, hogy a baleset helyszínén nemcsak megengedett, de kötelező a segítség. Ez persze akkor azt is jelenti, hogy ebben az esetben a kötelezettség maga után vonja a megengedés aktusát, amiből következik, hogy a két fogalom között semmiképpen sincs ellentétes viszony. A második mondatban azonban más a helyzet. Ekkor ugyanis a kötelezettséget épp', hogy szembenállónak érezzük a mondatba foglalt megengedéssel szemben. Arról pedig nyilvánvalóan szó sincs, hogy a jogosítvány megszerzésével párhuzamosan kötelességünk lenne autót vezetni. Ebben az esetben a kötelezés már kontradiktórikus viszonyban áll a megengedéssel. Joerden két példája mellé odatehetek még egy harmadikat is (ami az első kettőhöz képest megintcsak – részben – eltérő jelentéssel rendelkezik).

(3) Megengedett, hogy ne szavazzon valaki a parlamenti választásokon.

Ez a megengedés azt fejezi ki, hogy valami (itt: a szavazás cselekedete) nem kötelező, ami viszont logikailag "megengedi" azt az értelmezést, hogy ez a valami, vagyis a szavazás akár tilos is lehet.

Van tehát három példamondatunk, amelyekben három – egymást részben átfedő, de nem teljesen azonos – jelentés mellett használjuk ugyanazt a nyelvi jelet, a 'megengedés' terminusát. A probléma tisztázásának elvi lehetőségét a logikai hatszög eredményeinek alkalmazása nyújthatja ezen a területen is. A logikai hatszög "felfedezése" után G. Kalinowski vezette be a deontikus hatszög fogalmát.[54] Az univerzális deontikus kategóriák, a **kötelező** (OBLIGATORY) és a **tiltott** (FORBIDDEN) fölöttes kategóriájaként be kell vezetnünk a **preskriptív** (PRESCRIPTIVE) fogalmát, ami a deontikus hatszög (U)-sarkába "kerül". A két partikuláris deontikus kategóriára használhatjuk a **megengedett** (PERMITTED), illetve az **mellőzhető** (OMISSIBLE) fogalmakat, és ezek fölöttes fogalmaként az (Y)-sarokba betehetjük az **opcionális** (OPTIONAL) kategóriát.

A deontikus logikai hatszög használata ugyanazért kívánatos, mint amiért a logikai hatszöget "szeretjük": egyértelmű kapcsolatok teremthetők így a hatszög kategóriái között (**kontradiktórius**, **kontrárius**, **szubkontrárius**), illetve olyan kategoriális felosztáshoz juthatunk, amelyre már érvényesül a JEPD-elv (ami a logikai négyzet elemeire még nem teljesült). Utóbbi szempont (elvárás) nagyon fontos mozzanat. Már Jeremy Bentham is törekedett arra, hogy a JEPD-elvnek megfelelő felosztást használjon a jogi relációk között [449, 63], ezért alkalmazta azt a kategóriahármast (**kötelezés**, a **tiltás** és a **nem tiltás és nem kötelezés**), ami teljesen egybevág a deontikus logikai hatszög kontrárius viszonyban levő fogalmainak csoportjával. Ugyanezt az összefüggést szemléltethetjük a deontikus logikai operátorok egymáshoz való viszonyát bemutató ábrán is:

[54] G. Kalinowski, La logique des normes. PUF, Paris (1972), idézi: [518]. Jörg Hansen szerint a deontikus logikai négyszöget Arthur N. Prior alkalmazta először 1955-ben [327].

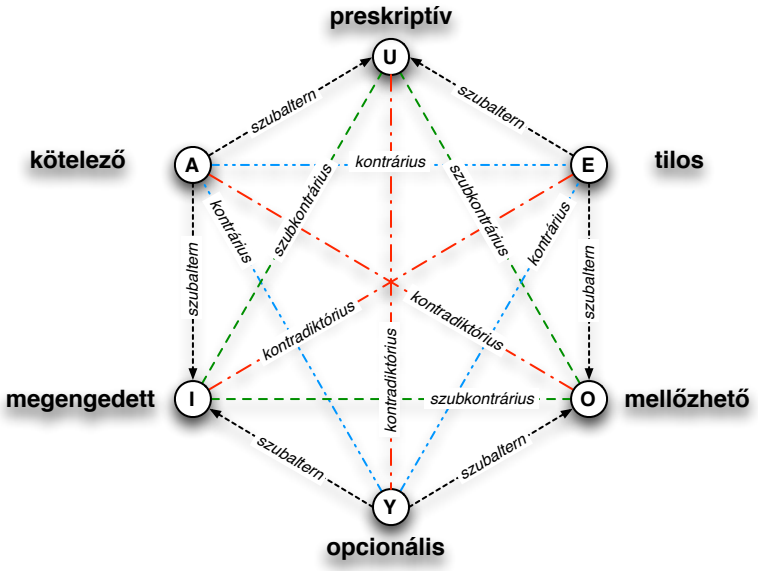


Figure 1.9: deontikus logikai hatszög

Az ábra szépen mutatja, hogy bár a két univerzális fogalom, a **kötelező** (OBLIGATORY, O) és a **tiltott** (FORBIDDEN, F) egymást kizáró kategóriák, a két partikuláris kategória, a **megengedett** (PERMITTED, P) és a **mellőzhető** (OMISSIBLE) egyszerre vannak átfedésben egymással, illetve valamelyik preskriptív fogalommal. Ezek az átfedések okozzák a deontikus operátorok kezelésének nehézségeit, pongyolaságát. A JPED-elv viszont fennáll a két preskriptív kategória (kötelezés és tiltás), valamint a két partikuláris kategória konjunkciója, tehát az **opcionális** (OPTIONAL) kategória között.

Természetesen a deontikus logikai hatszög elfogadása, alkalmazása nem ígér csodát. Bevezetése "mindössze" azt az előnyt ígéri, hogy sok esetben azonosítani tudjuk, hogy valamely konkrét nyelvi megnyilatkozáshoz melyik fogalmat (logikai kategóriát) kell hozzárendelnünk. A fejezet első felében bemutatott konkrét mondatainkra visszautalva: az első példa a deontikus logikai hatszög (I)-sarkába, a második az (Y) pontjába, a harmadik pedig az (O) esetbe tartozik.

McNamara: gratuitous=non-obligatoy [491]

Sztenderd deontikus logika

A korábban bemutatott modális axiómák (↬128.o.) közül a (D) deontikus axióma jelenti a kulcsot a **sztenderd deontikus logika** (Standard Deontic Logic, SDL) felépítéséhez. A deontikus logikai séma definiálásakor a szükségszerű, illetve a lehetséges ($\Box$, illetve $\Diamond$) modális operátorok helyett ezek deontikus megfelelőit ($\underline{\mathbf{O}}$, illetve $\underline{\mathbf{P}}$) használhatjuk. A deontikus operátorok hatókörében levő *A* szimbólum nem egy cselekvés neve (ahogy azt von Wright még az első modelljében tételezte), hanem egy olyan propozíció jele, amely propozíció azt írja le, hogy az adott cselekvést végrehajtották (ezt az értelmezést Prior javasolta először). A deontikus logikai szakirodalomban elterjedt felfogás szerint tehát $\underline{\mathbf{O}}(A)$ interpretációja: 'kötelező, hogy az a helyzet megvalósuljon, amelyet az *A* leíró mon-

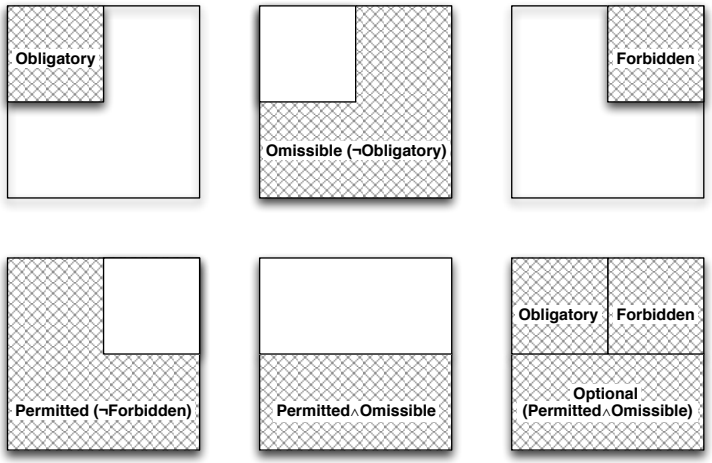


Figure 1.10: deontikus operátorok a JEPD elv szerint

dattal fejezünk ki' [327]. A továbblépéshez rögzítenünk lell a deontikus logika nyelvét.

(DL.0) $\mathscr{L}_{DL}(\mathscr{L}_{QL}, \underline{\mathbf{O}})$, ahol

$\mathscr{L}_{QL}$	propozícionális logika
$\underline{\mathbf{O}}$	kötelező deontikus logikai operátor

A deontikus logika modelljét – Jörg Hansen leírását követve [327] – a következő axiómasémák segítségével építhetjük fel, amihez hozzávesszük még az $\underline{\mathbf{F}}$ tiltás és a $\underline{\mathbf{P}}$ megengedés operátorának definícióját.

(DL.1) $\text{PERMITTED}(A) \equiv \underline{\mathbf{P}}(A) \stackrel{\text{def}}{=} \neg\underline{\mathbf{O}}(\neg A)$

megengedett az A cselekvés akkor és csak akkor, ha nem kötelező tartózkodni a végrehajtásától.

(DL.2) $\text{FORBIDDEN}(A) \equiv \underline{\mathbf{F}}(A) \stackrel{\text{def}}{=} \underline{\mathbf{O}}(\neg A)$

tiltott az A cselekvés akkor és csak akkor, ha kötelező tartózkodni a végrehajtásától.

(DL.3) Ha $\vdash A \leftrightarrow B$, akkor $\vdash \underline{\mathbf{O}}(A) \leftrightarrow \underline{\mathbf{O}}(B)$

(Ext) extenzionalitás elve (principle of extensionality): ha A kötelező ($\underline{\mathbf{O}}(A)$), akkor minden vele ekvivalens B propozíció is kötelező ($\underline{\mathbf{O}}(B)$).

(DL.4) $\underline{\mathbf{O}}(A \wedge B) \rightarrow (\underline{\mathbf{O}}(A) \wedge \underline{\mathbf{O}}(B))$

(M) monotonitás elve (principle of monotony): ha A és B együttesen kötelező, akkor egyenként is az.

(DL.5) $(\underline{\mathbf{O}}(A) \wedge \underline{\mathbf{O}}(B)) \rightarrow \underline{\mathbf{O}}(A \wedge B)$

(C) kumulativitás elve (principle of cumulativity): ha A és B kötelező, akkor ketten együtt is azok (ez a monotonitás elvének megfordított fennállását fejezi ki).

(DL.6) $\underline{\mathbf{O}}(A) \rightarrow \underline{\mathbf{P}}(A)$

(D') deontikus axióma (deontic axiom): ha az A cselekvés kötelező ($\underline{\mathbf{O}}(A)$), akkor megengedett is egyben ($\underline{\mathbf{P}}(A)$).

(DL.7) $\underline{\mathbf{O}}(A \vee \neg A)$

(N) deontikus szükségességi szabály (deontic necessitation rule): ha valami tautológia, akkor az kötelező.

Hansen megadta azt is, hogy miként teremthető meg a kapcsolat von Wright deontikus logikájával. Az (M) és (C) axiómák alapján levezethető von Wright **disztribúció elve** (principle of distribution), miszerint: $(\underline{\mathbf{P}}A \vee \underline{\mathbf{P}}B) \rightarrow \underline{\mathbf{P}}(A \vee B)$. A (D') séma ekvivalens a von wright-i **megengedés elvével** (principle of permission), ami így néz ki: $\underline{\mathbf{P}}A \vee \neg\underline{\mathbf{P}}A$. Mindezek alapján definiálhatunk egy deontikus nyelvet, illetve egy arra épülő deontikus rendszert.

DEONTIKUS LOGIKAI NYELV (THE LANGUAGE OF DEONTIC LOGIC)

A deontikus logika $\mathscr{L}_{DL}$ ábécéje megegyezik a propozícionális logika $\mathscr{L}_{PL}$ nyelvével, amely tartalmazza még az '$\underline{\mathbf{O}}$' deontikus operátort. Az $\mathscr{L}_{DL}$ nyelvet ugyanúgy definiáljuk, mint a $\mathscr{L}_{PL}$ propozícionális logika nyelvét annyi többletszabály tételezésével, hogy ha $A \in \mathscr{L}_{DL}$, akkor $\underline{\mathbf{O}}(A) \in \mathscr{L}_{DL}$ is igaz. A $\underline{\mathbf{P}}(A)$ megengedés operátorával rövidíthetjük a $\neg\underline{\mathbf{O}}(\neg A)$, illetve a $\underline{\mathbf{F}}(A)$ tiltás operátorával $\underline{\mathbf{O}}(\neg A)$ kifejezéseket.

SZENDERD DEONTIKUS LOGIKAI RENDSZER (THE SYSTEM SDL)

Az *SDL* az a legkisebb halmaz, amely tartalmazza a $\mathscr{L}_{DL}$ nyelv összes tautológiapéldányát, az (M), (C), (D) és (N) axiómasémákat, valamint zárt az (Ext) és a modus ponens szabályaira nézve.

---

ez nem-iteratív SDL, vagyis nincs megengedve a deontikus operátorok iterációja, például: O(O(A))

tolerálás: von Wright [792, 90]

gyenge és erős megengedés von Wright: [792, 86]

Arend Soeteman, tartózkodni valamitől vs. eltekinteni hasonlósága és különbsége [699, 151]

a negyedik deontikus operátor lehetősége [699, 168]

Angelika Kratzer [421]

Megemlítem, de mivel nem használom a modellépítésben, nem mutatom be azt a továbbfejlesztési irányt, amely a deontikus logikai hatszög további kiterjesztéseként tekinthető: Alessio Moretti, The Geometry of Standard Deontic Logic [518].

A deontikus logika redukciója az alethikus logikára

Egymástól függetlenül, szinte teljesen egyidőben, lényegét tekintve ugyanazt a redukciós programot fogalmazta meg Stig Kanger [392] és Alan Ross Anderson [12], [14], [17]ENG)-[18]HUN). Anderson H.B. Bohnert 1945-ből származó javaslatát formalizálta, amely szerint a deontikus logikát vissza lehet vezetni az altehikus logikára úgy, hogy a deontikus operátorokat a szankció, szankcionálás fogalmára hivatkozva definiáljuk. Ehhez

a normál alethikus logika elméletére (kalkulusára), az alethikus logika **szükségszerű**, illetve **lehetséges** ($\Box$, illetve $\Diamond$) operátorára, valamint egy új kategóriára, az előírt cselekvés végre nem hajtásakor kiszabandó S **szankció** fogalmára van szükség (Anderson olykor szankció-logikaként hivatkozott az új logikára). Az új elemként bevezetett fogalmat később elnevezték **prohairetikus konstansnak** (prohairetic constant), leginkább azért, mert Kanger is hasonló megfontolásokat követve szintén egy új fogalom felvételét javasolta az alethikus logika sémája mellé, de ő nem a szankciót, hanem a morál által előírt Q cselekvést (pontosabban az azt leíró propozíciót) emelte be a modelljébe. A két javaslat teljesen ekvivalens egymással [327], mégha a kötelezés deontikus operátorának definíciója el is tér egymástól a két esetben:

(DL.8) $\underline{\mathbf{O}}_Q(A) \equiv \underline{\mathbf{O}}(A) \equiv \text{OBLIGATORY}(A) \equiv \underline{\mathbf{O}}A \stackrel{\text{def}}{=} \Box(Q \to A)$

(DL.9) $\underline{\mathbf{O}}_S(A) \equiv \underline{\mathbf{O}}(A) \equiv \text{OBLIGATORY}(A) \equiv \underline{\mathbf{O}}A \stackrel{\text{def}}{=} \Box(\neg A \to S)$

Ezt a redukciót természetesen alá lehet támasztani megfelelő axiómákkal is. A két esethez az alábbi formulákat lehet hozzárendelni:

(DL.10) $\Diamond Q \qquad (\mathrm{D}_Q)$

($\mathbf{D}_S$) a formula interpretációja az lehet, hogy 'lehetséges, hogy S hamis', vagy másként: 'a szankcionálás kontingens, nem szükségszerű',

(DL.11) $\Diamond\neg S$

($\mathbf{D}_Q$) a formulát pedig értelmezhetjük úgy, miszerint 'lehetséges, hogy Q igaz'.

(DL.12) $\Box(A \to B) \to (\Box A \to \Box B)$

(K) a modális logikai séma azt követeli meg, hogy ha egy kondicionális szükségszerű, akkor a szükségszerű előtag materiálisan implikálja a szükségszerű utótagot.

(DL.13) $\vdash A \to \vdash \Box A$

(RN) ha az A a nyelv eleme, akkor a $\Box A$ kifejezés is az.

Az eddigiek alapján természetesen már sorban definiálhatjuk a deontikus operátorokat:

(DL.14) $\underline{\mathbf{O}}_S(A) \equiv \underline{\mathbf{O}}(A) \equiv \text{OBLIGATORY}(A) \equiv \underline{\mathbf{O}}A \stackrel{\text{def}}{=} \Box(\neg A \to S)$

(DL.15) $\underline{\mathbf{P}}_S(A) \equiv \underline{\mathbf{P}}(A) \equiv \text{PERMITTED}(A) \equiv \underline{\mathbf{P}}A \stackrel{\text{def}}{=} \neg\Box\neg(A \wedge \neg S)$

(DL.16) $\underline{\mathbf{F}}_S(A) \equiv \underline{\mathbf{F}}(A) \equiv \text{FORBIDDEN}(A) \equiv \underline{\mathbf{F}}A \stackrel{\text{def}}{=} \Box\neg(A \wedge \neg S)$

(DL.17) $\underline{\mathbf{N}}_S(A) \equiv \underline{\mathbf{N}}(A) \equiv \text{OMISSIBLE}(A) \equiv \underline{\mathbf{N}}A \stackrel{\text{def}}{=} \neg\Box(\neg A \to S)$

(DL.18) $\underline{\mathbf{D}}_S(A) \equiv \underline{\mathbf{D}}(A) \equiv \text{PRESCRIPTIVE}(A) \equiv \underline{\mathbf{D}}A \stackrel{\text{def}}{=} \neg\Box(\neg A \to S)$

(DL.19) $\underline{\mathbf{A}}_S(A) \equiv \underline{\mathbf{A}}(A) \equiv \text{OPTIONAL}(A) \equiv \underline{\mathbf{A}}A \stackrel{\text{def}}{=} \neg\Box(\neg A \to S)$

magyarázat a rövidítésekre: optional = (A)lternative prescriptive = (D)ominated omissible = (N)on-obligatory

ideális világ (Kanger), szubideális világ (Anderson) mindketten

○

Bár itt nem jogelmélettel foglalkozunk, mégis érdemes felidézni a híres jogtudós, Hans Kelsen jogfelfogását – egy másik híres jogtudós, H.L.A. Hart interpretálásában:

> „A jog az az elsődleges norma, ami szankciót állapít meg. Nincs olyan törvény, ami tiltja az emberölést, csak olyan törvény van, ami arra utasítja a hivatalos személyeket, hogy bizonyos feltételek esetén bizonyos szankciókat alkalmazzanak azokra, akik megölnek valakit." [346, 50]

Az idézetből ugyan nem derül ki egyértelműen, de tény, hogy Kelsen a jog fogalmát a szankció fogalmából eredeztette, amit más jogtudósok vitattak ugyan, de ez számunkra itt nem lényeges. Az viszont fontos, hogy pontosan ezt a kelseni megközelítést reprodukálta a logika nyelvén Anderson és Kanger deontikus logikai redukciója. Peng-Hsiang Wang – Kangerék után jó ötven évvel – arra vállalkozott, hogy tisztázza, összeegyeztethető-e Kelsen jogfelfogása és Anderson deontikus logikai elmélete [808]. A válasza az lett, hogy igen, de ehhez részlegesen át kell alakítani az andersoni formalizmust az alapján, ahogy von Wright különbséget tesz a kétfajta **Ought** kategória, a deontikus és technikai kötelezőség között.

○

Makinson áttekintése a deontikus logika fejlődésén
[467]

Kellellentétes parancs paradoxona

Arthur N. Prior fedezte fel elsőként a materiális implikáció analógiájaként megteremtett deontikus logikai implikációval kapcsolatos paradoxont [582]. Tanulmányában – von Wright pár évvel korábban kidolgozott deontikus rendszerére utalva – Prior először rögzíti azt a nyilvánvaló tényt, hogy egyfelől az implikáció átírható konjunktív alakba, másfelől a kötelezés traszformálható megengedésbe, így az alábbi átalakítások érvényesek:

$$\underset{\sim}{\mathbf{O}}(A \to B) = \underset{\sim}{\mathbf{O}}(\neg A \wedge B) = \neg \underset{\sim}{\mathbf{P}} \neg (\neg A \wedge B) = \neg \underset{\sim}{\mathbf{P}} (A \wedge \neg B)$$

Ezután Prior felidézi von Wright deontikus rendszerének egy törvényét, miszerint ha egy cselekvés tiltott, akkor az tiltva van bármely más cselekvéssel konjunkcióba állítva, de tiltva van bármely más cselekvéstől való tartózkodással együtt is. Ezt formálisan így lehet kifejezni:

$$(\neg \underset{\sim}{\mathbf{P}} A \to \neg \underset{\sim}{\mathbf{P}} (A \wedge B)) \wedge (\neg \underset{\sim}{\mathbf{P}} A \to \neg \underset{\sim}{\mathbf{P}} (A \wedge \neg B))$$

A továbbiakban a formula második összetevője az érdekes Prior számára, mivel abban el lehet végezni egy helyettesítést a deontikus implikáció fenti összefüggése alapján. Ekkor kapjuk a következő formulát:

$$\neg \underset{\sim}{\mathbf{P}} A \to \underset{\sim}{\mathbf{O}} (A \to B)$$

Ennek az értelmezése során meglepő következtetéshez juthatunk. A formula ugyanis úgy interpretálható, hogy ha valami tiltva van ($\neg \underline{\mathbf{P}}A$), akkor ha valaki mégis megteszi azt (A-t), akkor meg kell tenni bármi mást is (B-t). Prior ezt azzal példázza, hogy ha tiltott lopni, de valaki mégis lop, akkor – például – a házasságtörést is el kell követnie. Ez ellenkezik a normákkal kapcsolatos intuíciónkkal. Ezt a problémát nevezték el Prior-paradoxonnak.

Chisholm tovább gondola a problémát, és 1963-ban írt cikkében új fogalmat vezetett be [148]. A logikában akkor már régóta használt **tényellentétes** vagy kontrafaktuális (contrafactual) kifejezés mintájára megalkotta a **kellellentétes** vagy kontranormált (contrary-to-duty) kötelesség fogalmát.[55] Ezen azokat a helyzeteket értette, amelyekbe akkor kerülünk, miután megszegtük valamely kötelesség előírását, egy újabb – feltételes – kötelességhez kell igazodnunk. Vétségeink újabb kötelmeket szülnek – mondja Chisholm. Az igazi problémát azonban az a tény jelenti, hogy az ilyen helyzetek paradoxont rejtenek magukban. Chisholm a következő példával szemlélteti mindezt.

1) Kötelező, hogy valaki átmenjen segíteni a szomszédjának.
2) Kötelező, hogy ha valaki átmegy segíteni, akkor előtte szólnia kell, hogy megy.
3) Kötelező, hogy ha valaki nem megy segíteni, akkor tilos előtte szólnia, hogy megy.
4) Valaki nem megy át.

Az 1) és 4) mondatok egyértelműen leírhatók, de a 2) és 3) propozíciókat kétféleképpen is megragadhatjuk. Ebből következik, hogy ezt a négy mondatot négyféleképpen is formalizálhatjuk, bár egyik sem lesz tökéletes. A négyféle leírási mód így néz ki:

i)	$\underline{\mathbf{O}}(A)$	$\underline{\mathbf{O}}(A \to B)$	$\neg A \to \underline{\mathbf{O}}(\neg B)$	$\neg A$
ii)	$\underline{\mathbf{O}}(A)$	$\underline{\mathbf{O}}(A \to B)$	$\underline{\mathbf{O}}(\neg A \to \neg B)$	$\neg A$
iii)	$\underline{\mathbf{O}}(A)$	$A \to \underline{\mathbf{O}}(B)$	$\underline{\mathbf{O}}(\neg A \to \neg B)$	$\neg A$
iv)	$\underline{\mathbf{O}}(A)$	$A \to \underline{\mathbf{O}}(B)$	$\neg A \to \underline{\mathbf{O}}(\neg B)$	$\neg A$

Akármelyik leírási kísérletet vizsgáljuk meg, mindegyik esetről kiderül, hogy vagy inkonzisztens a négy formula, vagy az egyik kijelentés levezethető valemelyik másból, tehát nem függetlenek a mondatok egymástól [748], [327]. Márpedig az intuíciónk szerint a fenti négy mondat egymástól független és konzisztens.

Korábban már szó esett arról, hogy a deontikus logikára nem érvényes a (T) modális kalkulus (↬71.o.). Ebben az esetben viszont mondhatjuk azt is, hogy a formula kontraponáltja sem lehet igaz, vagyis:

$$\nvdash \neg A \to \neg \underline{\mathbf{O}}A$$

Természetes nyelven ezt úgy fejezhetjük ki, hogy ha nem áll fent egy A propozícióval leírt cselekvés, abból még nem következik az, hogy ne lenne az a cselekvés kötelező. Ez teljesen egybevág azzal a mindennapos tapasztalattal, hogy a normákat nem egyszer megsértik az emberek, de ettől a normák még nem vesztik el az érvényességüket.

———-

R.P. McArthur, Anderson's deontic logic and relevant implication [481]

[55] Magyar nyelven egyetlen forrást találtam, amely hivatkozott a contrary-to-duty kategóriájára. Szolcsányi Tibor a 'kötelezettség-ellentétes norma' kifejezést vezette be a 'contrary-to-duty' fordításaként [748]. Tartalmilag megfelelő ez a kifejezés is, de túl hosszúnak találom, ezért kitartok saját javaslatom mellett.

Diadikus deontika

monadikus deontikával szemben jött létre a feltételes kötelességet (conditional obligation) kell kezelni Prior paradoxon of derived obligation Chisholm-paradoxon contrary-to-duty kontranormált helyzete

Mindezek hatására von Wright bevezeti a diadikus deontikus operátor (dyadic deontic operator) fogalmát [789]. Az új diadikus formula a következő alakú:

(DL.20) $\underset{\sim}{\mathbf{O}}_D(A/B) \equiv \underset{\sim}{\mathbf{O}}(A/B)$

means that A is obligatory under the condition B. The old, ?monadic? O- operator is then defined as OA = O(A/?),

Bengt Hansson az 1969-es 'An Analysis of some Deontic Logics' című tanulmányában alapozza meg a diadikus deontikus logika (Dyadic Deontic Logic, DDL) új rendszerét [330].

az axiomatizálást elvégezte Wolfgang Spohn [703] Lennart Åqvist [26]

he introduced a semantics based on a betterness relation for conditional obligations (it was axiomatized only six years later)

> „The problem (...) is what happens if somebody nevertheless performs a forbidden act. Ideal worlds are excluded. But it may be the case that among the still achievable words some are better than others. There should then be an obligation to make the best out of the sad circumstances." [330]

The pregnant woman example can be represented by an ideal situation from the perspective of traffic law ? in which the car goes to the right and signals that it will go to the right, but the situation in which the car goes to the left and signals that it will go to the left is better to the one in which the car goes to the left but signals that it will go to the right. As mentioned in the previous section, it is precisely the possibility of violation, that led Meyer and Wieringa to introduce the use of norms and deontic logic in computer science. Moreover, the formalism has become popular also in other areas of computer science too, such as non-monotonic logic and qualitative decision theory. [99]

a normák nem mindig vezethetők vissza egyértelműen a jó és a rossz szembeállítására

le lehet-e lőni a terroristák által eltérített, utasokkal teli repülőgépet, ha tudni lehet, hogy még nagyobb kárt akarnak okozni vele a terroristák?

A diadikus deontikus logikai operátorok segítségével természetesen ki lehet fejezni a monadikus operátorokat is. Ehhez a feltételként egy tautológiát kell megadnunk, aminek eredményeként az operátorunk feltétel nélkülivé "változik":

(DL.21) $\underset{\sim}{\mathbf{O}}_m(A) \equiv \underset{\sim}{\mathbf{O}}(A) \equiv \underset{\sim}{\mathbf{O}}A \stackrel{\text{def}}{=} \underset{\sim}{\mathbf{O}}(A/\top)$

ahol $\underset{\sim}{\mathbf{O}}_m$ a monadikus operátort, $\top$ a tautológiát jelöli.

David Lewis, Semantic analyses for dyadic deontic logic [443]

leválaszthatóság

leválaszthatóság (detachment)

Andrew J.I. Jones és Ingmar Pörn, 1985, Ideality, sub-ideality and deontic logic [387]

, Dyadic Deontic Detachment [456]

P.S. Greenspan alapján elkülöníthetjük a diadikus logikák két csoportját aszerint, hogy a **faktuális leválasztás** (factual detachment) vagy a **deontikus leválasztás** (deontic detachment) elvének érvényesülését engedik meg [308]. A két leválasztási elvet így fejezhetjük ki:

(DL.22) $B \wedge \underline{\mathbf{O}}(A/B) \rightarrow \underline{\mathbf{O}}(A)$

(FD) a faktuális leválasztás

(DL.23) $\underline{\mathbf{O}}(B) \wedge \underline{\mathbf{O}}(A/B) \rightarrow \underline{\mathbf{O}}(A)$

(DD) a deontikus leválasztás formulája

jelöl

támadhatóság

támadhatóság (defeasability)

Carlos E. Alchourron, Detachment and Defeasibility in Deontic Logic 1996 [6]

normák támadhatóságának elemzése Van der Torre és Tan, [777]

Prakken és Sergot, 1997) Prakken, H; Sergot, M. J. (1997). Dyadic deontic logic and contrary-to-duty obligations. . Defeasible Deontic Logic, edited by Donald Nute, Kluwer Academic Publishers, Dordrecht/Boston/London.

További deontikus logikai paradoxonok, dilemmák, bibik

Hansson: revenger's paradox [334]

vannak paradoxonok, amelyeket a klasszikus deontikus logikában jelentkeznek, bizonyos további (nem-klasszikus) deontikák pont arra jöttek létre, hogy ezek közül párat kiküszöböljenek

Computer scientists are usually surprised when they read the philosophical literature, because the posed problems seem to have a trivial solution. For ex- ample, the most famous deontic paradox of all, is often posed as the problem to give a consistent representation such that none of the sentences can be derived from the others:

1. A certain man should go to the assistance of his neighbors, 2. If he goes, he should tell them he is coming 3. If he does not go, he should not tell them that he is coming 4. He does not go.

In SDL the set Oa, O(a ? t), ?a ? O(?t), ?a is inconsistent, and in Oa, a ? O(t),?a ? O(?t),?a the sentences are not logically independent. However, this problem is trivially solved by replacing the material implication by a strict implication, or a relevant implication, or a defeasible implication. This has been known since the early days of deontic logic, as may be expected since both deontic logic and conditional logic were major branches of philosophical logic. In general, any paradoxical consequence can be solved by simply weakening the logic (e.g., solve Ross? paradox by replacing standard deontic logic by a non-normal modal logic). The misconception is simply due to the fact that the deontic paradoxes do not work the same way as experiments in engineering or the sciences. They are just used to illustrate the formal system, not to guide research in the area. A similar phenomena and misconception has been present in the field of defeasible reasoning and deontic logic, where the use of the infamous Tweety example has

been criticized for similar reasons. The reason why contrary-to-duty paradoxes have been discussed for fifty years in deontic logic is that a lot of normative reasoning is directly or indirectly related to violations, just like in defeasible reasoning a lot of reasoning is directly or indirectly related to exceptions.

free choice permission

Hans Kamp Zimmermann Katrin Schulz

[465]

Jennings kiterjeszti a problémát, nem csak az engedély hatályába tartozó cselekvésekről lehet szó, de a cselekvőkre vonatkozó diszjunkció is okozhat gondot. [378, 263]

> Fred vagy Bill jöhet.

ekvivalens azzal a mondattal, hogy:

> Fred jöhet, és Bill jöhet.

De Jennings szerint ha elhagyjuk a deontikus logika világát, akkor is jelentkezhetnek hasonló problémák. Bizonyságul erre további két példát is bemutat. Először az összehasonlítások világából hoz egy mondatot:

> Pierre fontosabb mint William vagy Réné.

aminek kiterjesztő átalakítása meghagyja az eredeti mondat jelentését.

> Pierre fontosabb, mint William, és Pierre fontosabb, mint Réné.

Jennings másik példája minden eddiginél egyszerűbb, leírása nem igényel semmi különös logikát, de a probléma ebben az esetben is jelentkezik. Vegyük az alábbi mondatot:

> Ha Laurie vagy Alison távozik, akkor távozom én is.

Ez a mondat ugyanúgy egészíthető ki, mint az eddigiek: a kiegészítés közben a konjunkciót kell használjuk. Az átalakított mondat így néz ki:

> Ha Laurie távozik, akkor távozom én is, és Alison távozik, akkor távozom én is.

A példák alapján elfogadható Jennings érve, miszerint igazából nem deontikus logika problémájáról, hanem valami másról van itt szó. Kétféle megoldási lehetőséget is bemutat, hogy aztán az egyik mellett letegye a voksát. Az egyik értelmezés szerint ezekben a mondatokban azért alkalmazunk ‘vagy’-ot az ‘és’ helyett, mert az utóbbi használata egyértelműen a konjunktív jelentést írná elő. Ha ugyanis ezt a mondatot olvassuk:

> Ha Laurie és Alison távozik, akkor távozom én is.

akkor egyértelműen arra gondolunk, hogy a feltétel két tagja egyszerre fent kell, hogy álljon a következmény létrejöttéhez, vagyis akkor távozom én is, ha már mindketten távoztak. Az eredeti mondatban pedig ez nem volt benne.

a természetes nyelv nem oldja meg egyérteéműen ezt a problémát ezért minden formalizálási kísérlet gondban van

A jó szamaritánus paradoxona

paradox of the good Samaritan Prior [582] vagy inkább? [583]

„You ought to help the assaulted person. Therefore, there ought to be an assaulted person."

A valamilyen jó cselekedetet jelent, amit akkor kell csak megtenni, ha B-t is megtette valaki B valamilyen rossz cselekedetet jelent

ha $\vdash A \to B$ akkor a neccessiation elvéből következik: $O(A) \to O(B)$

Lennart Åqvist 1967-ben: Good Samaritans, Contrary-to-Duty Imperatives, and Epistemic Obligations [24]

megoldása a DDL segítségével: [327, 66]

A szelíd gyilkos paradoxona

Forrester's gentle murderer' paradox

[262]

megoldása a DDL segítségével: [327, 66]

Jörgensen dilemma

Jörgen Jörgensen (Jørgen Jørgensen) [388]

Ross paradox

[611]

counterintuitive derivation of ?you ought to mail or burn the letter? from ?you ought to mail the letter?, is typically viewed as a side effect of the interpretation of ?or? in natural language.

'Post the letter or burn it!'

Plato's Dilemma

Plato's Dilemma (Lemmon 1962) [436]

paradox of commitment

[582]

[336]

$\underline{\mathbf{O}}(\neg A) \to \underline{\mathbf{O}}(A \to B)$ ha megteszed azt, ami tiltva van, akkor meg kell tenned bármi mást is ha ellopsz egy autót, akkor hajts át egy gyalogoson

Knower

[336] [750]

Deontikus logika a preferencialogikára alapozva

Hansson [336]
van Benthem ás társai [773]
[772]

A ΔEON-közösség fejlődése

áttekintés: Hansen, Imperatives and Deontic Logic. On the Semantic Foundations of Deontic Logic [327]

a legfontosabb problémák az idők során [99]:

period	tradition	main issue
50s	monadic modal logic	relation $\underline{\mathbf{O}}$ and $\underline{\mathbf{P}}$
60s	dyadic modal logic	relation $\underline{\mathbf{O}}$ and facts, violations, sub-ideality and optimality, CTD
70s	temporal deontic logic	relation $\underline{\mathbf{O}}$ and time
80s	action deontic logic	relation $\underline{\mathbf{O}}$ and actions
90s	defeasible deontic logic	dilemmas, CTD
00s	imperatives, normative systems	Jorgensen's dilemma

Ten philosophical problems [326]

1. How can deontic logic be reconstructed in accord with the philosophical position that norms are neither true nor false? 2. When is a set of norms to be termed ?coherent?? 3. How can deontic logic accommodate possible conflicts of norms? How can the resolution of apparent conflicts be semantically modeled? 4. How do we reason with contrary-to-duty obligations which are in force only in case of norm violations? 5. How to define dyadic deontic operators with regard to given sets of norms and facts? 6. How to distinguish various kinds of permissions and relate them to obligations? 7. How can meaning postulates and intermediate terms be modeled in semantics for deontic logic reasoning? 8. How to define counts-as conditionals and relate them to obligations and permis- sions? 9. How to revise a set of regulations or obligations? Does belief revision offer a satis- factory framework for norm revision? Can the belief merging framework deal with the problem of merging sets of norms? 10 ?????

Imperszonális deontika

A deontikus logikával kapcsolatban indulásától fogva sokakban megfogalmazódott az a remény, hogy ez a diszciplina segítheti majd a normatív, jogi szövegek gépi elemzésében, az ilyen szövegeken belüli inkonzisztenciák megtalálásában. Ezt az ígéretét azonban az évtizedeken keresztül tartó folyamatos fejlődése ellenére sem tudta beteljesíteni ez a tudományterület. Nyilván nem szabad, nem lehet egyetlen tényezővel megmagyarázni mindezt, de egy magyarázó lehetőséget azért kiemelnék a sok közül.

A deontikus logika eddig bemutatott elméletének komoly hiányossága a személytelensége, imperszonális jellege. Kötelezésekről, tiltásokról, megengedésekről van szó, de úgy, hogy közben nem tudjuk, hogy ezek kire, kikre irányulnak. Egész pontosan nem is igazán az a baj, hogy nem tudjuk, kikre vonatkoznak az előírások, vagy kik adják ki ezeket, bár

azért ezek is fontos kérdések, hanem sokkal inkább a megközelítés mondanikus jellege a probléma.

A személytelenséget az imperszonális jelző alkalmazásával ragadhatjuk meg. Az **imperszonális deontikában** "nincsenek" személyek, így nem lehet kapcsolatot teremteni a deontikus operátorok és az ágensek között, emiatt nem lehet az ágencia egész problémakörét sem kezelni. De még nagyobb baj, hiány az, hogy ez a deontika nem képes megragadni azt a tényt, hogy a társadalmi cselekvésekben megnyilvánuló normativitás lényegét tekintve relációs jellegű, aminek kezelésére relációs fogalmakra, és ezért **perszonális deontikára** lenne szükség. A társadalmi cselekvés formális modelljének felépítésében kulcsszerepet szánok a perszonális deontika elméletének, amit a későbbi fejezetekben mutatok majd be.

supererogation

Urmson egy 1958-as cikkében vetette fel azt a problémát, hogy a deontikus logikában bevezetett modalitásokkal leírhatjuk ugyan a kötelességhez, a megengedéshez és a tiltáshoz igazodó cselekvéseket, ám ezzel a fogalomhármassal nem lehet minden morális aktust megragadni. A *szent* és a *hős* cselekvése kilóg ebből a klasszifikációból [765](eng)-[766](hun). Bizonyos emberek akkor is teljesítik kötelességeiket, amikor a többiek nem. A többség gyengeségből vagy gyávaságból nem teszi meg, ami kötelező lenne számára, csak kevesek tudnak erősek maradni. A szent ellenáll a kísértésnek, a hős legyőzi a félelmét, de mindkettő megcsinálja, amit morális kötelességének tart (saját maga és a környezete). A különbség annyi csak köztük, hogy miben győzik le magukat, minek állnak ellen: a szent a vágyainak, a hős a félelmeinek.

kétféle

supererogation - túlteljesítés: actions that are good to do but not bad not to do

presept (előírás) tanács (counsel)

which is the object: actor vs. action

suberogation: actions that are bad to do but not good not to do

a másik irányban: ???? (offence)

az első elméletformálási kísérlet: Chisholm, supererogation and offence [149]

1964-ben Chisholm kicsit módosít az elméletén [150]

Roderick M. Chisholm és Ernest Sosa: [153]

JoelFeinberg:1961 (annyira nem érdekes) [245] magyarul: krokovay felelősség kötetében

Paul McNamara: a deontikus hatszög bővítése, a supererogation modellbe emelése [491], [490] axiomatizálás: [470]

Jan C. Joerden a túlbuzgóság (Supererogation) kezelhetősége, modellbe emelése érdekében arra tesz javaslatot, hogy a deontikus hatszög helyett inkább deontikus tízszöget (Deontological Decagon) kellene használnunk [381].

kritika, még általánosabb keretbe illesztés: [516], [518]

A supererogation fogalma segíthet tisztázni a jézusi etika legfontosabb tézisének lényegét.

> „ . . . ha ellenséged éhezik, adj neki enni, ha szomjazik, adj neki inni. Ha ezt teszed, izzó parazsat raksz a fejére.
> Ne engedd, hogy legyőzzön a rossz, inkább te győzd le a rosszat jóval.
> Ne tartozzatok senkinek semmivel, csak kölcsönös szeretettel, mert aki embertársát szereti, a többi törvényt is megtartja.

> Hiszen a parancs: ne törj házasságot, ne ölj, ne lopj, a másét ne kívánd, s ami egyéb parancs még van, mind ebben az egyben tetőződik: Szeresd embertársadat, mint saját magadat.” (Róm. 12.10-?????)

Az iszlám szankcionálási gyakorlat

A deontikus logikában nem nagyon esik róla szó, ezért önmagában is érdekes lehet az iszlám kultúra szankcionálási gyakorlatának bemutatása, de különös jelentőséget adhat, egyfajta ellenpontot jelenthet az a tény, hogy az elméletileg létező, de az európai kultúrkör szankcionálási gyakorlatából hiányzó elemek itt megjelennek.

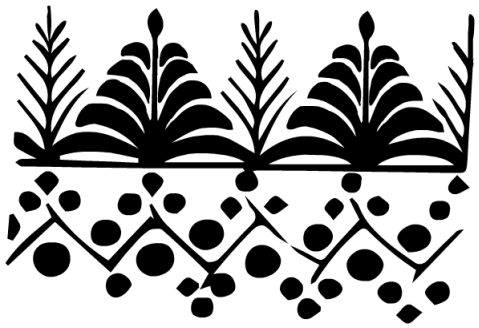
modalitás	szankció elvégzésekor	szankció mellőzésekor
megparancsolt	nincs semmi	büntetés
ajánlott	jutalom	nincs büntetés
tiltott	büntetés	nincs semmi
helytelenített	nincs büntetés	jutalom
közömbös	nincs semmi	nincs semmi

A problem about permission

David Lewis, A problem about permission [440]
mégnézni, hátul nem akartam-e foglalkozni ezzel

1.13.1 Dinamikus logika és deontika

J.J.C. Meyer az andersoni alapokat integrálta a dinamikus logikával és így más felépítésű deontikus logikát javasolt. [504]



Chapter 2

A társadalom modelljei

Ágensek

2.1 Ego

2.1.1 Szelf

The Emergence of Self Terrence W. Deacon, James W. Haag and Jay Ogilvy
szelf

Feinberg

Todd E. Feinberg, Altered Egos. How the Brain Creates the Self: [247]
agysérülések utáni mentális betegségek:
Ego nem érzékeli magát betegnek Ego nem tartja magáénak egyes testrészeit, elidegeníti magától Ego nem érzékeli a világ egy részét (például a bal oldalt)
a tárgyakhoz való viszony: Ego-close vs. Ego-distant (saját cipő vs. más cipője)
„the feeling of how one relates to things in a personal sense. Our identities are built around this sense of relatedness. Personal relatedness provides the structure within which the self is anchored in the world. The self is a continuum of relationships." 30.
állandó változás a kétféle állapot között: elidegenedés vs. ??? jamais vu (where the familiar appears strange) vs. deja vu (where the unfamiliar appears familiar)
Capgras szindróma (kicserélt dolgokban való hit) arcfelismerő képesség elvesztése 36 de el lehet veszteni a hely- vagy tárgyfelismerés képességét is 37 Freud családi románc Otto Rank: 15 Hero-birth myth A Capgras szindrómában az érzelmi viszony változott meg alapvetően, nem az észlelési képesség
Frégoli szindróma + környezeti megkettőződés
When a patient misidentifies a stranger as a relation, we call it Fregoli syndrome. When the patient mislocates the hospital in his or her own backyard and doubles the place, we call it environmental reduplication. However, both syndromes represent a change in the self, in the sense that one's relatedness to the world undergoes a transformation. The bound- aries of the self are altered, and one's personal world draws nearer. 49.
Capgras represents withdrawal of personal relatedness Fregoli-like disturbance is an insertion of personal relatedness the self has withdrawn its involvement

confabulation

„Many areas of the brain make a contribution to the self. It appears that the problem is not in deciding where in the brain the self is: Rather, the question is, if the self, the ego, is widely distributed across the brain, how is it possible for a single self to exist? This problem needs to be addressed if we are to understand how the brain creates a unified person, the 'I.' "71.

Patients with personal confabulation confuse the margins of the self, figuratively speaking. By this I mean, in the context of the confabulation, that the self makes a personal appearance in the guise of a character within a story of the patient's own making. Patients with asomatognosia, on the other hand, seem to lose a part of the self in a more literal sense. The arm in the patient with severe asomatognosia may be totally rejected and personified as if it belonged to someone else. I would now like to discuss another condition?mirror misidentification?in which the patient misidentifies not the actual body, but its mirror image. 72

mirror misidentification Susan used sign language in front of the mir- ror in her bedroom. When asked what she was doing, she reported that she was communicating with the "other" Susan. She believed that there was another Susan who was identical to her in appearance, age, back- ground, education, and so on.

Susan was able to accurately identify other family members, physicians, and neighbors. She never misidentified anyone but herself in the mirror.

Susan's condition is a Capgras syndrome for her mirror image. 74

the misidentification occurred in any reflective surface, from glass in store windows to car mirrors. She did not, however, misidentify her reflection in the mirror of her cosmetic compact. It occurred to me that perhaps its small size made it less likely for the symptom to occur. So I first instructed her to take out the compact and asked her to identify her image, which she did correctly. Using a series of mirrors of increasing size, I was able to convince her as the reflection got larger, despite her initial protestations, that they represented her own reflection. I repeated this process over several days until her husband reported that the behavior had vanished. 78.

There is a separation of knowledge in a psycho- logical sense: one domain of knowledge-perception of a person in the mirror-is not integrated with another-a self-concept. Susan and Rosamond have a pathological separation of the self as knower-the "I" that is the subjective aspect of the self that looks into the mirror and the self as known-the "me" that is the objective aspect of the self that is seen in the mirror. 78.

Gordon G. Gallup made a rather remarkable discovery. ... only in humans, chimps, and orangutans. Interestingly, gorillas show no ability to recognize themselves in the mirror. 79.

Double Indemnity Doppelganger autoscopia (heutoscopia or hallucination speculaire) A central characteristic of autoscopic hallucinations is the close connection between the image and subject. There is a feeling of belonging to or of connection with the autoscopic hallucination that serves an important distinguishing feature. Unlike the Capgras patient, who denies or disavows the relationship between the image and relation, the hallucinating person feels a part of the autoscopic hallucination. There is a feeling of oneness with the hallucination. 81

The vision of the self represents a projection of the self. The remarkable feature about autoscopy is that the self, the ego, the standpoint from which the world is psychologically

perceived, coheres as an entity that can be experienced outside of the body. Autoscopy teaches us that the self can be split, and one can have a dual sense of awareness or viewpoint both as experiencing and observing entities. The widespread belief in "guardian angels," the experience of the depersonalized self in out-of-body experiences in religious settings, the common occurrence of the idea of a presence under conditions of stress in normal persons, all suggest that fragmentation of the self represents a common human behavior. 84.

Imaginary Companions and Guardian Angels The imaginary companions are spoken to, played with, and have an almost physical space in the child's world. 84. the imaginary companion is often a projection of the child's own ideals. 85. the fantasy had restored to some degree Miriam's lost relationship with her then-absent mother. 86. Miriam was expressing her own difficulties by attributing them to Susan 86.

The child who employs his imagination and the people of his imagination to solve his problems is a child who is working for his own mental health. He can maintain his human ties and his good contact with reality while he maintains his imaginary world. Moreover it can be demonstrated that the child's contact with the real world is strengthened by his periodic excursions into fantasy. It becomes easier to tolerate the frustrations of the real world and to accede to the demands of reality if one can restore himself at intervals in a world where the deepest wishes can achieve imaginary gratification. 86-87.

the related adult phenomenon is ... the idea of a presence. This was described as a feeling or impression, sometimes amounting to a delusion, that one is "not alone." There is a sense of a presence of someone beyond the self. It is not a visual hallucination, nor is it a misidentification; it is rather a feeling or sense of company. Both the imaginary companion and the idea of a presence appear under conditions of stress. 87.

The imagined, the alter ego, the imaginary companion, or nemesis rep- resent personifications, reifications of hopes or fears. The mind creates concrete representations of inner, sometimes hidden, emotions and gives them a life of their own. 89.

Alien Hand Syndrome 94. corpus callosum sérül, megszűnik a kapcsolat a két agyfélteke között The hand performs actions that are seemingly purposeful and voluntary, yet the patient claims that the movements are unwanted and involuntary and are not consistent with their conscious intentions. 94

intermanual conflict: a páciens a jobbkezével kanállal akart volna levest enni, miközben a balkeze egy kenyérdarabot rakott volna a szájba 98

hogy lehetséges egységes működés "kettévágott agy setében? 99

hogyan látunk? Torsten Hubel and David Wiesel how a single neuron in the brain could respond to a line in the world. 1- simple cortical cells? pontokat látnak 2- complex cortical cell: a pontokat látó sejtekt látják, így már vonalat érzékelnek 3- hypercomplex cell:

visual neurobiologist Semir Zeki terms "topical convergence"

Higher order visual cells respond preferentially to highly specific and complex stimuli such as hands or faces. Some of these neurons respond best to the frontal view of a face and others to a side view. Neurons of this type led to the somewhat fanciful notion of the "grandmother" cell, a cell so specific that it fires only to the vision of one's grandmother! Cells of that degree of specificity do not exist, but it is true that the farther along a sensory processing stream that one looks, the more specific a cell's response characteristics become. 113

While the cells early in the visual stream have small receptive fields and "know" where each line of the face is, these early cells do not "know" that a given line is part of a face. They do not have knowledge of the "big picture" of the face that will emerge later on in the processing stream. The face cells, on the other hand, "know" there is a face, but due to the process of topical convergence, these cells don't know where the face is located in space. Cells of the brain project higher and higher in a hierarchical fashion in order to code for increasingly specific complex and abstract properties, but information coded by cells earlier in the process is not and cannot be lost in awareness. As Zeki observes, cells comprising both the early and the late stages of the visual processing stream must make a unique contribution to consciousness. In this way the conscious mind is "spread out" across the activity of the many neurons that make up different regions of the brain. 113-114

a karteziánus színpad gondolata nem működőképes, mert egyrész nincs egyetlen hely az agyban, ahol minden agyi aktivitás egyesülne, másrészt a gondolathoz kellene egy inner homunculus, ami nicsn, de nem is oldaná meg, csak doébb tolnáeggyel a problémát (ki az a homunculus, kinek jelent?) 115

binding problem: különböző feladatokat ellátó területeket hogy kapcsolja össze az agy? az autó színét, formáját, hangját stb. más területek észlelik, dolgozzák fel, és ezek hogyan kapcsolódnak egymáshoz? 116

Francis Crick: synchronized neuronal firing in the 35 to 75 Hertz range "might be the neural correlate of visual awareness." 117

Gerald Edelman: re-entrant cortical integration (RCI)

If there is no area in the brain where all the regions that con- tribute to the self physically "come together.," then how do we exist as whole, integrated, unified persons?

just as the unification of the visual system contributes to our sense that we are unified beings, our ability to produce unified actions is essential to our feeling of wholeness as a person.

Tinberger: cselekvésvezérlés

The instinct originates from the highest brain regions of the fish, and surely some particular areas of the fish brain are more necessary for instinctual behaviors than others. But this instinct has no unified spatial locus in the brain. Even though the instinct is unified in the behavior of the fish, it is not materially unified in the fish brain. The brain locus of an instinct, and hence the source of the unified action of the fish, is distributed across large regions of the fish brain. 119

vast areas of the brain are involved in the complicated act of speech production, yet somehow when we "will" to speak, the entire act is unified into a whole and coordinated behavior. 120

when we search for the inner source of the center of the "will," we find ourselves at impasse. Just as there is no "pontifical neuron" sitting atop the perceptual hierarchy, there is no single brain region creating and controlling our actions and intentions. 120

emergence and brain One popular example of an emergent system is water, for which it is noted that the properties of liquidity, wetness, and transparency do not apply to a single water molecule, but to the aggregate "water." 121

the emergent feature is said to be "greater than the sum of its parts." 122

A second key concept of emergence theory is the notion of constraint, 122

hierarchical control is the essential and distinguishing characteristic of life. 122

It is through constraint that hierarchical systems operate. There is often a reciprocal relationship between emergence and constraint in hierarchical systems. ... When the system is considered as a whole, the mitochondria contribute to the emergence of the lung, and the lung in turn constrains the mitochondria. 122-123

The third key concept of emergence theory that I wish to discuss is perhaps its most controversial. This concept is the notion of nonreducibility, the claim that the wholes created by an emergent system can- not be explained simply by or reduced to the properties of its constituent parts.

the nonreducibility of a higher level to a lower level in a hierarchical system is a key characteristic of an emergent property. .... if a property is emergent, it is not reducible to those parts. It follows that if a property in a hierarchy is reducible to the properties of things at lower levels of the hierarchy, it is not emergent.

When we view with one eye open, we lack depth perception. When we view our environment with both eyes open, due to the differences between the images generated by each eye, the cyclopean eye possesses depth perception for the integrated image. Depth perception is an emergent property of vision not present when viewing with either eye alone. 125.

... pótolni a kimaradt részt

My purpose in elaborating at some length on this topic is to point out that purposeful action can be distinguished from teleonomic (but non-purposeful) behavior by its greater degree of constraint over a multitude of nested parts. 134

nested vs. non-nested hierarchy (compositional hierarchy)

Allen and Starr (1982) provided a definition of a nested hierarchy: "A nested hierarchy is one where the holon at the apex of the hierarchy contains and is composed of all lower holons. The apical holon consists of the sum of the substance and the interactions of all its daughter holons and is, in that sense, derivable from them. Individuals are nested within populations, organs within organisms, tissues within organs, and tissues are composed of cells" (p. 38). 162.o

Allen, T. F. H., and T. B. Starr. Hierarchy. Perspectives for Ecological Complexity. Chicago: University of Chicago Press, 1982.

Salthe, S. N. Evolving Hierarchical Systems: Their Structure and Representation. New York: Columbia University Press, 1985. ———————

Giulio Tononi: An information integration theory of consciousness [751] Tudatosság nélkül nincsen céllal végzett cselekvés, azonban az utóbbiak is szoros kapcsolatban állnak az érzelmi és affektív folyamatokkal

———————

corpus callosum

A corpus callosum, azaz kérgestest, azaz CC a két agyféltekét összekötő rostköteg. Több, mint 200 millió idegszálból áll. A CC segítségével "kommunikál" a két oldal egymással, másrészt a CC koordinálja a két oldal "munkáját".

hemisphere - agyfélteke

frontal lobe - frontális lebeny

spinal cord - gerincvelő

cerebral cortex - agykéreg

cerebrospinal fluid, a condition called hydrocepbalus.

ipsilateral vs contralateral

anterior commissure
chimerics, chimera - kiméra
pineal gland - tobozmirigy
cortical - kortikális
inferior temporal cortex

Daniel Stern: az önzérzékelés, a szelf kialakulása

Daniel N. Stern *A csecsemő személyközi világa* című könyvében [711]

bontakozó szelf érzése	0-2 hónapos korban
magszelf érzése	2-6 hónapos korban
szubjektív szelférzés	7-15 hónapos korban
verbális szelférzés	15 hónapos kortól

Ezek nem egymást váltó szakaszok, megjelenésük után mindegyik továbbra is fennmarad, és az egész élet folyamán “funkcionál”.

kategorikus affektus (öröm, bánat) vitális affektus (dinamika, időbeli, intenzitásbeli mintázat, az intenzitás és az idő amodális kategoráibóél tevődik össze

Önismeret

[9] de se hitek szelftörténelem önazonosság kijelentések tartalmának Szelfhez kötése
Tőzsér János, az önismeret határa [752]
George Herbert Mead: én és az általánosított másik

2.2 A világ

világállapotok, alternatívák
társadalmi kapcsolat, társadalmi cselekvés
szándék, vágy, törekvés, intenció, döntés

Chapter 3

Egyéni döntések

mentális állapot tudat intencionalitás fenomenális-tudatos öntudat test-elme viszonya

négyféle mentális állapot: észlelés érzet érzelem gondolat

három reláció: azonosság szuperveniancia konstitúció

3.1 Intencionalitás

tudat jelensége

valami reprezentációja reprezentációs tartalom tartalomtípus reprezentációra irányultság

a tudat tárgyra irányultsága (tárgyrairányultsága)

intecionalitás = valamire vonatkozás (aboutness) (Fodor cikke az elmefilozófia könyvben)

irányultság beállítódás motiváltság attitűd diszpozíció

Huoranszky: „Wittgenstein, Anscombe és Ryle érveinek hatására egészen a hatvanas évek végéig az volt az általánosan elismert álláspont, hogy, mivel a szándéktulajdonítás és a cselekvés leírása konceptuálisan nem független egymástól, az intenciók nem lehetnek a cselekvés (hume-i) okai. A konceptuális kapcsolat érvét azonban D. Davidson Actions, Reasons and Causes cím? írásában (1963) meggy?z?en cáfolta. Másfel?l kételyek merültek fel a tekintetben is, hogy vajon a kauzalitásról kialakított, els?sorban a hume-i filozófián nyugvó elképzelésünk módosításra szorul-e vagy sem (lásd pl. J. Searle Intentionality (Cambridge University Press 1983.) cím? könyvének 3. és 4. fejezetét”

3.1.1 Brentano

> „Minden pszichikai közös jellemvonása abban található, amit sajnálatos módon gyakran oly félreérthető kifejezéssel tudatnak neveznek, vagyis egyfajta szubjektív magatartásban, egyfajta, mint mondják, *intencionális* viszonyban, amely valamire irányul, ami talán valójában nem létezik, ám belsőleg tárgyként adott. Nincs hallás valami meghallott, hívés valami elhitt, remény valami reménylett nélkül, nincs törekvés a törekvés célja nélkül, és nincs öröm anélkül, hogy valaminek ne örülnénk, s még sorolhatnánk.” [124, 55]

Brentano szerint az intencionlitás három típusa, osztálya

> „Az első alaposztály a legtágabb értelemben vett képzeteké ...
> A második alaposztály az ítéleteké ... az ítélet alapja, miként a vágyakozásé is, mindig valamilyen képzet; ... Mi tehát a megkülönböztető jegye annak, amikor nem egyszerűen képzetet alkotok, hanem ítélek is? – Ebben az esetben a képzetalkotáshoz egy második, a képzet tárgyára irányuló intencionális viszony is kapcsolódik, a képzet tárgyának elismerése vagy elvetése. Aki Istent említi, Isten képzetét fejezi ki, aki pedig azt mondja: van Isten, arról beszél, hogy hisz benne.
> A harmadik alaposztály a kifejezés legszélesebb értelmében vett érzelmi mozgásoké, kezdve attól, hogy valamely gondolat egyszerűen kellemesen vagy kellemetlenül hat rám, egészen a meggyőződésen alapuló örömig vagy szomorúságig, illetve a cél- és eszközválasztás legbonyolultabb jelenségéig. Arisztotelész mindezt az *órexisz* fogalmában foglalta össze." [124, 56-57]

Arisztotelész 'órexisz' kategóriája a vágyakozás, törekvés fogalmával azonosítható.

> „A második alaposztályban az intencionális viszony tartalma elismerés vagy elutasítás volt; a harmadik alaposztályban ugyanez szeretet vagy gyűlölet, avagy (ami éppoly helyes kifejezés) tetszés vagy nemtetszés." [124, 57]

A pszichikai és fizikai jelenségek közti különbséget keresve Brentano az intencionalitást jelöli meg

> „Minden pszichikai jelenség azáltal jellemzett, amit a középkor skolasztikusai intencionális (nyilvánvalóan mentális is) inegzisztenciának (valamire irányuló belső létezésnek) neveztek, s amit mi, habár nem is egészen kétértelműség nélküli kifejezéssel, egy tartalomra való vonatkozásnak, egy tárgyra való irányulásnak (ami alatt itt nem realitást kell érteni) vagy pedig az immanens tárgyiasultságnak nevezhetnénk. Mindegyik magába foglal valamit a tárgyakból, habár nem mindegyik egyenlő módon. A képzetben valamit megjelenítünk, az ítéletben valamit elismerünk vagy elvetünk, a szeretetben valamit szeretünk, a gyűlöletben valamit gyűlölünk, a vágyakozásban valamit kívánunk stb." [123, 161]

Frederick Chisholm egy lexikonbejegyzésében elemzi Brentano intencionalitás fogalmát. [151] hogyan lehet az intencionális

mondat prefixumokat keres

Mások által már javasolt, de Chisholm által nem megfelelőnek talált kritériumok:

existentional generalization

nonexistentional occurence

referential opacity

Megfelelő kritérium:

M intencionális, ha minden p propozícióra igaz, hogy $M(p)$ logikailag kontingens.

nem intencionális

'lehetetlen' prefix, mert a lehetetlen (egyes négyzetek körök) mondat mindig igaz, tehát nem lehet kontingens

horváth gizella, érték-cikkében ír brentano könyvéről, a jó fogalmáról

Héctor-Neri Castañeda [529]

3.1.2 Husserl

Husserl cogito cogitatio

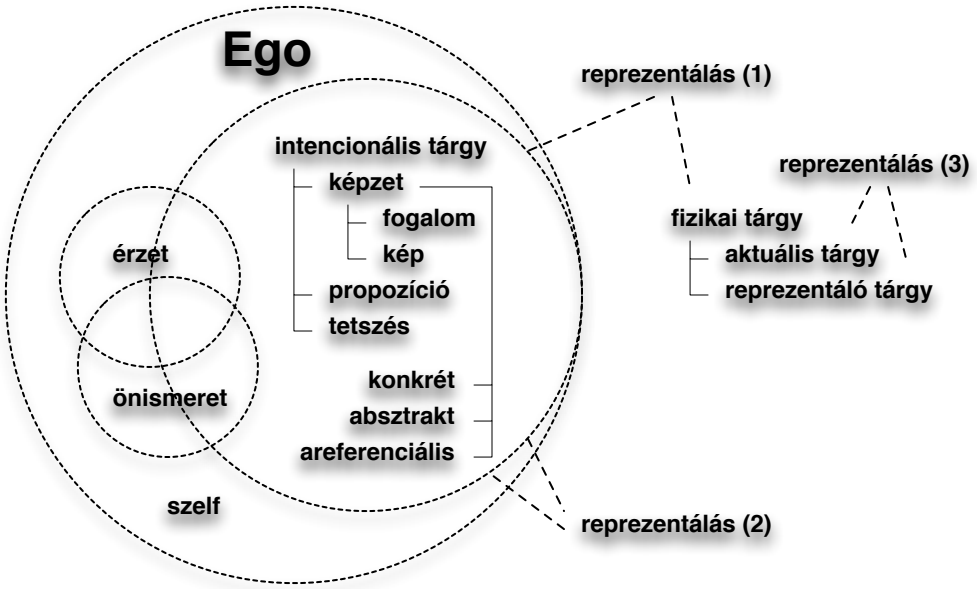


Figure 3.1: Brentano intencionalitás fogalma

3.1.3 Anscombe, Dennet, Fodor, Searle

Anscombe, Dennet, Fodor, Searle

Dennet: az intencionalitás filozófiája, [202]

Searle

kollektív intencionalitás mi-intencionalitás 121.o. a módszertani individualizmus talaján maradva lehet beszélni róla

ez a McClintock-féle motivációs gyűrű (meg a játékelmélet) értelmezésekor is fontos ez a szempont. Amikor a gyűrű tengelyei körül vagyunk, ott egyéni intencionalitásról beszélhetünk, de amikor interdependens helyzetekbe kerülnek a résztvevők, akkor használható a mi-tudat, kollektív intencionalitás

mindkét helyről hivatkozni ide

A *p* propozícionális tartalom minden esetben ugyanaz

értékelés, ítélkezés, figyelmeztetés, könyörgés, imádkozás, mentegetőzés, panaszkodás, becslés, vélekedés, előrejelzés átok, eskü, áldás vallomás, hazugság/hazudás, titkolódzás/titkolás diplomakihirdetés, fokozatszerzés, kinevezés

Searle [652, 92]

Ruth Millikan: miben más az emberi és nem-emberi reprezentáció? hat különbség van

felszólító reprezentációk, szándék-doboz [509]

ezt lehet, hogy Bratman-elméletnél kéne említeni?????????

az állatoknál a kijelentő és felszólító reprezentációk egybemosódnak, az embereknél nem

„Azok a reprezentációk, amelyek esetében nem tehető különbség kijelentés és fel-

beszédaktustípus	asszertívum	direktívum	komisszívum	expresszívum	deklaráció
megfeleltetési irány	szó-világ	világ-szó	világ-szó	nincs	szó-világ és világ-szó
ágens		Alter	Ego		
őszinteségi feltétel	vélekedés	vágy	vágy		
érvényességtípus	igaz-e	teljesült-e	teljesült-e		
példák	kijelentés, leírás, osztályozás, magyarázat	utasítás, parancs, kérés	ígéret, fogadalom, felajánlás, megállapodás, jótállás, fenyegetés	bocsánatkérés, köszönetnyilvánítás, gratuláció, felköszöntés, részvétnyilvánítás	hadüzenet, kinevezés, házastárssá nyilvánítás

szólítás között, a tényállásokat közvetlenül kötik össze a cselekvésekkel, tehát azzal, amit a tényállás esetén tenni kell. Az emberi hitek nincsenek közvetlenül a cselekvésekhez kötve. Ha nem kombinálják őket megfelelő vágyakkal, az emberi hitek tehetetlenek. És az emberi vágyak szintén tehetetlenek, ha nincsenek megfelelő hitekkel kombinálva." [510, 267]

következtetési képesség (és szükség)

„Mivel az emberi lények belső reprezentációs rendszerében a kijelentő és a felszólító funkciók szét vannak választva, szükség van újbóli integrálásukra. Ezért az emberek gyakorlati következtetéseket végeznek, a hiteket és a vágyakat újszerű módokon kombinálják, hogy azok először szándékokat, majd cselekvést eredményezzenek." [510, 267]

„Ha az igaz hitek funkcionálisak, a hamis hitek pedig legtöbbször nem rosszabbak az üres elménél, akkor még a tévedésre rendkívül hajlamos hitrgözítő eszközök is jobbak lehetnek a hitrögzítő eszközök teljes hiányánál ... Mivel rendkívül pontos hitrögzítő eszközök tervezése nehéz, talán valójában előnyösebb túl sok hitet rögzíteni, megengedve, hogy közülük egyesek hamisak lesznek. Ennek megfelelően hitfelhasználó mechanizmusainkat talán arra tervezték gondosan, hogy képesek legyenek tolerálni a hamis hitek magas arányát." [510, 259-260]

ez nagyon kell a weberi hatszöghöz ia
[9]

Altrichter Ferenc az intencionalitás fogalmát úgy értelmezi azon belül a propozícionális attitűdöket

A propozícionális attitűdök összefüggő rendszert, sőt kauzális rendszert alkotnak. Ha Ego úgy vélekedik, hogy kevés benzin van a kocsijában, és tudja, hogy a tervezett útra kevés lenne, akkor tudja azt is, hogy egy benzinkútnál tankolnia kell. A vélekedéseink, hiteink, propozícionális attitűdjeink vagyis az intencionális tárgyak rendszere kauzális.

A propozícionális attitűdökre (azok kialakulására, változására) hatnak a külvilág tárgyai, az intencionális állapotaink, vélekedéseink „kauzálisan kapcsolódnak a külvilág dolgaihoz, eseményeihez és sajátosságaihoz". Ez már egy második kauzális viszonyrendszert jelent, ami a belső és külső világ "tárgyai" közötti kapcsolatokat jellemzi. A kétfajta kauzális rendszer egy egységet (Altrichter kifejezésével élve: "kauzális teret") alkotva folyamatosan "kauzális befolyással" van az Egora.

A kauzális hatásrendszer létezésén túl azonban az attitűdjeink további “kényszereknek” vannak alávetve. A propozícionális attitűdjeink világában ugyanis szemantikai, logikai kényszerek is érvényesülnek.

> „Már maga az a tény, hogy a mentális tartalom igazként vagy hamisként értékelhető, hiteinket, gondolatainkat beemeli a logikai térbe, és aláveti mentális állapotainkat a logika szabályainak, amelyek az emberi elme vagy tudat konstitutív szabályai.” [9, 139]

A logikai-szemantikai kényszerek létezése és “befolyásolási képessége” azért különösen fontos, mert feltételeznünk kell azt is, hogy a racionalitásnak akár az ideális, akár a reális formáját is vesszük figyelembe, a racionalitásnak erős attitűdbefolyásolási képességet kell tulajdonítanunk.

zártság elve....

> „A reális racionalitás regulatív szabálya a logika konstitutív szabályaival összhangban – hogy úgy mondjam – felülbírálja a kauzális tér elemeit, és bizonyos pszichikus állapotokat egyszerűen megsemmisít és kiküszöböl a kauzálisan releváns állapotok köréből.” [9, 140]

A zártság elve megköveteli az önismeret létét. Akkor változtathat valaki az attitűdjein, ha tud róla, hogy korábban milyen intencionális állapotban volt. Ez viszont már másodrendű intencionalitást jelent, hiszen olyan Egonak például tudnia kell, hogy tudott valamit korábban.

olyan fogalmakkal kell tehát bővíteni mind az Ego egyes szám első személyű, mind a megfigyelő harmadik személyű ontológiáját, mint a hit, a vélekedés, vágy, a gondolat.

Az önismeret fogalmialg szükséges tehát, ehhez viszont már másodrendű intencionalitást kell a modellbe emelnünk. Az önismeret fogalma pedig nyilván megköveteli az öntudatnak, a szelfnek a fogalmát.

> „Az önismeretnek viszont logikailag szükséges feltétele az öntudat: az tehát, hogy képesek legyünk önmagunkat megkülönböztetni minden mástól, hogy képesek legyünk önmagunkra vonatkozó, *de se* hiteket fenntartani, vagy gondolatokat gondolni, hogy képesek legyünk önmagunkra az ‘én’ szóval vagy ennek valamilyen nyelvi ekvivalensével referálni, és hogy képesek legyünk önmagunkat a propozcionális attitűdök szubjektumaként felfogni.” [9, 140]

3.1.4 egyéb

három emberi képesség, fakultás: kognitív (vélekedés) affektív (vágy, érzelem, érzés, öröm, fájdalom) konatív (intenció, akarat, elköteleződés)

cognition, affection, conation belief, desire, intention

conation, as goal-directed striving or purposive activity, involved two meanings of the goal or end of the striving. "One is the obtaining of means and the other making affective [sic] use of the means

mentális állapot, melynek tartalma egy cselekvési terv ?????

———————-

3.1.5 Neurológiai alapok

Damasio, A. (1985). Understanding the mind's will. [182] „. . . SMA performs a key role in the cerebral organization of action. More specifically, I agree with the notion that the SMA plays a crucial role in volitional processes." SMA – supplementary motor area

anterior cingulate is fontos, az SMA mellett van

———————————-

3.1.6 Értelem, tudat

szubjektív értelem self

tudat

"Like Spinoza and Damasio, Hazlitt begins with the body. Imagine a child putting a hand in a flame. The bodily sensation of pain teaches the brain about danger. "I will not touch fire again or it will hurt," thinks the child. But wait a minute, says Hazlitt: when the child learns the lesson, it is imagining a being that does not yet exist: its own future self.

To adopt Damasio's terms, the brain is mapping a body that is still only imaginary. From feeling comes the capacity for imagination and hence for empathy. If we can imagine our future self, we can also imagine other selves. The human mind thus has a natural capacity not only for self-interest, as Hobbes had proposed, but also for disinterest (in the proper sense of the word). For Hazlitt, this insight was the starting-point for a lifetime's commitment to both liberal politics and the empathetic power of the arts. I have a hunch that he would have considered Damasio's findings a cause for joy, not despair."

3.2 Reprezentáció

tartalom propozícionális tartalom képi tartalom

3.3 Teleológia vs. teleonómia

The Teleological Conception of Practical Reasons Douglas W. Portmore Mind, Vol. 120, No. 477 (January 2011), pp. 117-153

Arisztotelész teleológia értelmezése

teleologikus gondolkodás, beállítódás, de Sausa

oksági állítás formulája:

$A \rightarrow B$

a teleologikus állítás formulája:

$(B \rightarrow G) \rightarrow B$

ahol G valamilyen célt, jövőbeni állapotot jelent

Ernest Nagel két cikke (egy előadás két része) a teleológia és a biológia kapcsolatáról 1977-ből [525] [526]

Mayr 1961: két megközelítés a biológiában: hogyan működik (funkcionális biológia) miért úgy működik (evolócióbiológia)

A teleonómia fogalmát Colin Pittendrigh vezette be 1958-ban. Az ő meghatározását azonban túl tág értelműnek találta Ernst Mayr, aki a 'Cause and Effect in Biology' c.

cikkében [478] a biológia által vizsgált kétfajta cél-követő jelenséget (end direction) elkülönítve egymástól a teleonómia és teleológia fogalmainak párhuzamos használatát javasolta. Mayr azért kritizálta Pittendrigh értelmezését, mert az nem tett különbséget aközött, hogy egyfelő az individuumok (rendszerek) célkövető módon (szándékosan - purposefully) képesek viselkedni, másfelől a genetikai programok evolúciójában vélelmezett célra törekvés között [478, 1504]. Mayr a teleonómia fogalmát így valamely program, információs kód alapján viselkedő individuum, rendszer célkövető minőségére javasolta leszűkíteni. Ebben az értelemben Mayr szerint a biológia mint tudomány a teleonómiára tud csak példával szolgálni.

Ernst Mayr: Teleological and Teleonomic, a New Analysis [479]

„It would seem useful to rigidly restrict the term teleonomic to systems operating on the basis of a program of coded information." (p.42)

Ernst Mayr: The Idea of Teleology [480]

emergencia, Todd Feinberg könyve

Hayne Reese: teleology vs. teleonomy - purposefulness, having an internal determination - pusposiveness, serving or effecting a useful function [598]

teleológiai funkció fogalma, Larry Wright [830]

f akkor és csak akkor x funkciója, ha
a) x egy korábbi y entitás reprodukciója, amely korábban f-t csinálta
b) x azért reprodukálódott, mert y korában f-t csinálta

feladat fogalma, John R. Searle [652]

plato.stanford.edu:

intention intentionality teleological theory of mental content emotion intrinsic vs. extrinsic values reasons for actions: agent-neutral vs. agent-relative value theory

3.4 Érvényesség

beszédaktus, Austin, Searle, Habermas, Hare

argumentáció, bizonyítás: áttekintés [429]

"mintha" [203]

érvényesség, jog, kényszerítő erő, követés [271] 152.oldal

3.5 Intenció

Az intenció mint szándék az intencionáltságnak csak egy típusa, és vigyázni kell arra, hogy a két terminus alaki közelsége nehogy áthallásokat, fölösleges és szándékolatlan konnotációkat okozzon.

> „Az *intencionalitás* meglehetősen balszerencsés szóalkotás, és mint annyi más szerencsétlen szót a filozófiában, ezt is a németül beszélő filozófusoknak köszönhetjük. A szó arra utal, hogy az intencionalitás az irányítottság értelmében valamiféle kapcsolatban áll a 'szándékkal', olyan értelemben például, hogy ma este moziba szándékozom menni. (A német nyelvben ez nem okoz problémát, mert az *Intentionalität* másképp hangzik, mint az *Absicht*, a mozilátogatás szándékának hétköznapi értelmében vett szándék.) Úgyhogy ne feledjük, az angolban a szándék csupán az intencionalitás számos formájának az egyike." [652, 89]

Searle figyelmeztetése természetesen nemcsak az angol nyelv esetében érvényes, a magyarban (és még nyilván sok más nyelvben) ugyanúgy igaz, hogy az intenció az intencionáltság egyik formája csupán. Nagyon oda kell tehát figyelni akkor, amikor az intencionalitásról beszélünk.

!!!!!!!!!!!!!!!!!!!

Ajzen, intencióelmélete

[3] [4]

négy cikk az intencionalitásról, logikáról: [462]

Szécsi Gábor, intencionalitás és okság, [733]

von Wright, Anscombe és mások elvetik a cselekvések oksági elméletét

> „Fölfogásukban a cselekvések intencióira hivatkozó leírások nem tekinthetők kauzális magyarázatoknak, mert a leírások tárgyait jelentő cselekvések logikailag nem különböztethetők meg a föltételezett intencióktól. Azaz egy cselekvés azonosítását megelo?zően csupán egyetlen intencionális eseményt észlelünk, amit viszont egyszer cselekvésként, máskor pedig e cselekvést eredményező intencióként írunk le. A föltételezett intenció okként való leírása így nem jelenti azt, hogy ez az intenció valóban a cselekvés oka is lenne, hiszen a leírásban szereplő események egyetemes törvényszerűséget nem példázó, belső viszonya nem felel meg az okság hume-i kritériumainak." [733, 530]

Davidson, Searle, Bratman elfogadják, persze módosítva

> „Az az okság-értelmezés ugyanis, amely az intencionális események és állapotok kettős, logikai-fiziológiai természetét egyetlen kauzalitás-fogalommal képes megragadni, áthidalhatja a test-tudat probléma dualista és monista megközelítései között húzódó szakadékot, s ezáltal új tartalommal töltheti meg magát a problémát is. Ennek a sajátos, dualista-monista konszenzusnak a lényege pedig az, hogy bár minden egyes intencionális esemény és állapot azonos egy fizikaival, logikai jellemzőik meghatározása mégsem természettudományos perspektívából történik. Azaz a kauzalitás-problematikára vetítve ez azt jelenti, hogy jóllehet minden egyes intencionális oknak és okozatnak megfelel egy fizikai jelenség, a közöttük fönnálló viszonyok nem testesítenek meg egyetemes törvényszerűségeket. Az intencionális események és állapotok közötti oksági relációk logikai jellemzői ugyanis mindig valamilyen individuális kognitív háttérre vezethetők vissza. ...
> Az intencionális eseményeknek és állapotoknak ezt a kettős vonását, a kauzális függést és a nomologikus függetlenséget együttesen megvilágító és Davidson által 'szabálytalan' monizmusnak nevezett fölfogás tehát egy sajátos, 'intencionalizált' okságfogalomra épül. " [733, 531]

534. teleje: Davidson meghatározása a szabálytalan monizmus (anomalous monism) fogalmára

Searle Intentionality: észlelési intencionlitás cselekvési intencionlaitás reflexió

542 Az észlelés abban az értelemben tekinthető 'kauzálisan ön-referens' intencionális eseménynek, hogy az 'észlelési tapasztalat' mint okozat saját okát, az észlelt jelenséget prezentálja. Az észlelés intencionális tartalma, az 'észlelési tapasztalat' tehát nem más, mint saját okának, vagy ahogy Searle nevezi, 'teljesülési föltételének' (condition of satisfaction) a mentális prezentációja. Ugyanilyen értelemben 'ön-referens' a cselekvés intencionális tartalma, a 'cselekvési tapasztalat' (experience of acting) is, amit Searle a

cselekvés közvetlen intencionális okával azonosít, azaz egyfajta intencióként fog föl. Az ilyen belső tapasztalatként értelmezett intenció ‘kauzális ön-referencialitásának’ lényege tehát az, hogy az intenció egyúttal oka is annak a cselekvésnek, amit a cselekvő számára a cselekvés folyamatában reprezentál. Azaz, ahogy Searle fogalmaz, az intencio?k ‘teljesülési föltételei’ az ‘észlelési tapasztalatokhoz’ hasonlóan ‘megkövetelik azt, hogy az intencionális állapotok maguk meghatározott kauzális viszonyban álljanak teljesülési föltételeik fönnmaradó részeivel.’

ilyenkor a kognitív háttér limitatív problémákat és cselekvéseket reprezentáló fogalmaira irányuló reflexió mintegy megvilágítja a cselekvő számára az elvégzendő elemi aktusok intencionális jellemzőit. 548.

549: elemi cselekvési szituáció limitatív cselekvési szituáció

fontos kérdés: az ok (cause) és az indok, indíték (reason) megkülönböztetése

egy kis etimológia: ind-ok, ind-ít, ind-ít-ék, ind-ul, ind-ul-at, ind-ul-ás

emotion, e-motion,

A cselekvést magyarázva, az ok és indok megkülönböztetésével foglalkozva Donald Davidson ‘Actions, Reasons, and Causes’ című tanulmányában (angolul: [191], magyarul: [190]) az indokot az ok egyik fajtájaként értelmezi. Ebben szembe megy G. Ryle, G.E.M. Anscombe, G.H. von Wright nézetével, akik megkülönböztetik a cselekvések gyakorlati indokait az események okaitól – nagyjából azon az alapon, hogy egyfelől a cselekvésekhez szükséges intencionális állapotok nem események, és az okságot események között állapítjatjuk meg, másfelől az indokokra nem találhatjuk meg az okokra jellemző általános törvényszerűséget, nomologikus minőséget.

pro-attitűd (pro-attitude) és vélekedés (belief)

elsődleges indok (primary reason)

két tézis az elsődleges indokokra:

> „1. Ahhoz, hogy megértsük, miként racionalizál egy tetszőleges fajtájú indok egy cselekvést, szükséges és elégséges azt látnunk (legalább körvonalaiban), hogyan építhető fel a cselekvéshez rendelt elsődleges indok.
> 2. A cselekvés elsődleges indoka nem más, mint a cselekvés oka.” [190, 8])

> „C1. *R* egy cselekvés végrehajtásának elsődleges indoka (ahol *A* cselekvés *d* leírásban szerepel), akkor és csak akkor, ha *R* egy, a cselekvőnek egy bizonyos vonást hordozó cselekvések iránti pro-attitűdjéből, valamint a cselekvő egy olyan meggyőződéséből áll, miszerint az *A* cselekvés *d* leírásban hordozza ezt a vonást.” [190, 9])

> „Nehány filozófus felismerte, hogy a nem-teleologikus magyatázatok képtelenek felutatni indok- alapú igazolást.” [190, 12])

> „ . . . amikor indok által magyarázunk egy cselekvést, akkor valójában újraleírjuk azt; az újraleírás egy bizonyos mintába helyezi a cselekvést, és pontosan ezt nevezzük a cselekvés magyarázatának.” [190, 13])

> „C2. Egy cselekvés elsődleges indoka a cselekvés oka.” [190, 14])

a szándék, amellyel a sofőr felemeli a karját,szintén nem esemény. A szándék: semmi – sem esemény, sem attitűd, se diszpozíció, se tárgy. 15

18. teteje fontos gondolat

3.5.1 Bratman előtti korszak: vágy-vélekedés modellek

Az intencionáltság, sőt, általában az emberi cselekvés elemzésében fontos mérföldkő volt G.E.M. Anscombe *Intention* című könyve, amely 1957-ben jelent meg [22]. Anscombe elemzésében az intenció fogalma alatt mindvégig szándékot ért. Azt viszont nem vizsgálja, hogy mi a különbség a vágy és a szándék között.

> „A man says 'I am going to walk' and we say 'that is an expression of intention, not a prediction'. But how do we know?" [22, 6]

Megijedni, reszketni, összerezzenni, zihálni – ezek nem intencionális cselekvések, nem cselekvések, elmosogatni, olvasni, taxit hívni – ezek cselekvések.

mindennek lehet oka, de csak az intenciónak lehet indoka

az előrejelzésekhez, predikciókhoz tényeket, bizonyítékokat (evidence) rendelünk, az intenciókhoz, szándékokhoz cselekvésre vonatkozó indokokat (reason)

karmol, tol, nyom, bevizez olyan cselekvések, amelyek valamilyen hatásokkal járnak (okoznak), kauzatív igék

Anscombe szerint szándékolt cselekvések azok, amelyekre értelmesen alkalmazható a – bizonyos értelemben vett – 'Miért?' kérdés. Mit jelent a bizonyos értelemben? Anscombe a választ több lépésben fejti ki. Először rögzíti, hogy milyen értelemben nem használható a 'Miért?' kérdés.

1) nem tudtam, hogy ezt teszem

Anscombe szerint egy cselekvést mindig többféle módon írhatunk le

A –

B –

C –

D –

ha a cselekvő A és B leírásokat adja, a kérdező pedig a C vagy D leírásokra kérdez rá, akkor a cselekvő válaszolhatja a fenti kérdést

2) nem akartam ezt tenni

mit is jelent ez?

1) akaratlagos (voluntary)

2) nem akaratlagos (involuntary)

2.a) megfigyelés által szerzett tudás, melynek okát csak megfigyelés révén tudhatjuk meg például: perisztaltikus bélmozgás

2.b) megfigyelés nélküli tudás (knowledge without observation, non-observational knowledge) ilyenek a saját testi mozgások cselekvései/történései = nincsen semmiféle érzékelés (történés), amely alapján tudunk a dolgokról például: saját lábam, kezem, fejem, testrészem, annak helyének, állapotának tudata

> „ . . . the class of things known without observation is of general interest to our enquiry because the class of intentional actions is a sub-class of it. I have already said that 'I was not aware I was doing that' is a rejection of the question 'Why?' whose sense we are trying to get at; here I can further say 'I knew I was doing that, but only because I observed it' would also be a rejection of it." [22, 14]

2.b.i) megfigyelés nélkül tudjuk a jelenséget, de az okát csak megfigyeléssel tudhatjuk meg például izomrángás elalvás során, reflexrúgás orvosi vizsgálat során,

2.b.ii) megfigyelés nélkül tudunk a jelenségről és az okáról is megfigyelés nélkül tudunk például: a krokodil hangjára hirtelen hátraugrunk

9par-ban a ellentétről beszél: reason for acting vs. reason for thinking something is true

10,par mentális ok, mentális okozás (mental çause) más, mint a motiváció

más, mint az érzelmek tárgya:

Egy kisgyereknek egy vörös ruhadarabra azt mondják, 'az csak egy darab szatén' (bit of satin), de a gyerek úgy érti, hogy 'egy darab sátán' (bit of Satan), ami miatt rémálmai lesznek. A félelem tárgya a ruhadarab volt, a félelem oka a félreértett mondat. Természetesen az érzelem oka és tárgya lehet ugyanaz. Ha éjszaka megpillantunk egy maszkos embert az ablakunk alatt, akkor az ijedtségünknek mind a tárgya, mint az oka az adott látvány (a félelmetes maszk) lesz.

sokszor mondjuk, hogy azért csináltam meg valamit, mert egy vágy erre késztetett, de az intenciót nem lehet mindig a vágyra visszavezetni (például kopogtatnak az ajtón, és odemegyek kinyitni)

nem tudom, ezt idézi-e vagy így gondolja:

> „In philosophy a distinction has sometimes been drawn between our motives and our intentions in acting as if they were quite different things. A man's intention is what he aims at or chooses; his motive is what determines the aim or choice; and I suppose that ' determines ' must here be another word for 'causes '." [22, 18]

motívumok

- visszautaló motívumok (backward-looking motives) például: bosszú, csodálat, szánalom
- interpretatív motívumok (motive-in-general or interpretative motive)
- előretekintő motívumok (forward-looking motives) = intention

> „Why is it that in revenge and gratitude, pity and remorse, the past event (or present situation) is a reason for acting, not just a mental cause?" [22, 21]

what the agent reports in answer to the question ' Why ?' is a reason for acting if in treating it as a reason he conceives it as something good or bad, and his own action as doing good or harm. 22

Whether in general good and harm play an essential part in the concept of intention it still remains to find out. 22

When the question ' Why ? ' about a present action is answered by a description of a future state of affairs, this is already distinguished from a mental cause just by being future. 23

a megfigyelések érzékszerveinken keresztül szerzett észlelések, tapasztalatok

= rendelkezünk egy elkülönítve leírható érzékeléssel és ezt az érzéki benyomást függetleníteni tudjuk az azt kiváltó dologtól (Anscombe példája: liftben érzett "süllyedő" érzés, amit elkülöníthetünk a lifttől magától, hiszen olyan érzést máskor is érezhetünk, mondjuk a vidámparkban hullámvasúton ülve)

elválasztható, elkülöníthető (separable) vagy nem elválasztható, nem elkülöníthető (non separable) kettősség

ez az elkülöníthetőség szemben áll más érzetekkel, például a reflexrúgás érzetét nem tudjuk elkülöníteni magától a kiváltó ütéstől

elkülönítve leírható érzések a megfigyelés által tudott dolgokat jellemzik a nem elkülönítve leírható érzések a megfigyelés nélkül tudott dolgokat

2.a.i) eset nem elkülöníthető érzés, viszont az okát megfigyeléssel tudjuk csak meg, ezért kizárhatjuk a szándékos cselekvések köréből az ilyen eseteket

a megfigyelés bizonyítékot "eredményez"

a cselekvésünkre vonatkozó, megfigyelés nélküli tudás a gyakorlati tudás (practical knowledge)

gyakorlati tudás vs. kontemplatív, spekulatív (contemplative, speculative)

perception-independent practical knowledge of what one is doing vs. perceptual knowledge of what is happening at a given moment to the material one is working on (Anscombe 89)

Anscombe vetette fel először azt az elképzelést, miszerint az igazodás irányultsága (direction of fit) szerint érdemes különbséget tenni olyan kognitív állapotok között, mint a vélekedés (belief) és konnatív állapot (conative state) mint a vágy (desire).

> „ . . . is there any description which is *the* description of an intentional action, given that an intentional action occurs?" [22, 37]

egy másik gondolati szálat jelent Anscombe azon meglátása, hogy az intencionális cselekvés mindig valamely leírás szerinti cselekvést jelent (acting under a description)

acting under a description

ez a narratív tudomány fele való lépés is egyben

Ian Hacking, Rewriting the Soul [317] [318]

Jean Beer Blumenfeld, Goldman's account of intentional action, [95]

jövővezérelt intenció (future-directed intention) és *jelenvezérelt intenció* (present-directed intention)

3.5.2 Bratman: BDI-elmélet

Michael Bratman, Intention, Plans, and Practical Reason [117]

a vélekedés és a vágy mellé fel kell venni harmadik mentális tényezőnek az intenció, szándék (intention) fogalmát is.

jövőre irányuló szándék és jelenre irányuló szándék

elköteleződés (commitment) fogalma

szándék támadható, megvűltoztatható, módsítható, visszavonható (defiesable, revocable)

a szándék és elköteleződés szükséges a koordinációhoz

szándék: 1) tehetetlensége (intertia) van 2) önmagában érveket szolgáltat a további mérlegelések során 3) viselkedésbefolyásoló (conduct controlling) proattitűd (pro-attitide)

terv (plan) szándékhoz hasonló részleges hierarchikus(an szervezett) (hierarchically structured)

> „The strategy of settling in advance on such partial, hierarchically structured plans, leaving more specific decisions till later, has a deep pragmatic rationale. On the one hand, we need to coordinate our activities both within our own lives and socially, between lives. And we need to do this in ways compatible with our limited capacities to deliberate and process information. Further, given these same limitations we

> need a way to allow prior deliberation to shape later conduct. This argues for being planning creatures. On the other hand, the world changes in ways we are not in a positions to anticipate; so highly detailed plans about the far future will often be little use and not worth bother. Partial, hierarchically structured plans for the future provide our compromise solution." [117, 29-30]

31-32.o. két fontos követelmény a tervvel kapcsolatban a tervnek az ember vélekedéseihez képest belsőleg konzisztensnek kell lennie (internally consistent)

a tervnek cél-eszköz koherensnek kell lennie (means-end coherence)

ez a két minőség adja a tervek, szándékok, illetve a vágyak, értékelések közötti megkülönböztetés lehetőségét a vágyakra nem érvényesítjük egyik kritériumot sem lehet inkonzisztens két vágy bennem, de egy szándék vagy egy terv nem. Vágyhatok egyszerre meccsre menni és vacsorázni egyet este, de nem tervezhetem őket egyszerre végrehajtani.

Másfelől pedig a vágyakhoz (mint lehetéges célokhoz) nem kell eszközöket rendelnünk, szemben a szándékokkal és tervekkel, amelykhez igen.

> „. . . intentions do not provide reasons for actions to be placed on the scale with desire-belief reasons for actions, rather than to provide reasons to be weighed in favor of one considered alternative over another. The reasons to be weighed in deliberation remain desire-belief reasons." [117, 34]

> „Practical reasoning, then, has two levels: prior intentions and plans pose problems and provide a filter on options that are potential solutions to those problems; desire-belief reasons enter as considerations to be weighed in deliberating between relevant and admissible options. This two-level structure is an essential part of the way in which intentions and plans play their coordination-facilitating role, and so part of the way in which intentions enable us to avoid being merely time-slice agent – agents who are constantly starting from scratch in their deliberations.." [117, 35]

36 flat-out desire (all-or-nothing)

aszimmetria tézis:

> „. . . intention-belief inconsistency is closer to criticizable irrationality than is intention-belief incompleteness. This is the asymmetry thesis." [117, 38]

internal or external assessment of action a belső értékelés figyelembe veszi az elköteleződéseket és szándékokat, míg a külső eltekint ezektől. Emiatt eltér, eltérhet a kétfajta értékelés.

áz ágens racionalitása

háromfajta intenció deliberatív, mérlegelő nem deliberatív, nem-mérlegelő policy-based,

az újragondolás (reconsideration) és racionalitás kérdése

háromféle újragondolás nem-reflektív mérlegelő (deliberative) poliy-based

> „. . . some of an agent's belief about the future will depend in a critical way on his intentions. My belief that I will be at Tanner Library this afternoon is based on my knowledge that I intend to go there." [117, 63]

> „. . . given our limits and the important role of plans in reliably extending the influence of present deliberation to future action we may expect that reasonable habits of (non)reconsideration will involve a tendency not to reconsider a prior plan except when faced with some problem for that plan." [117, 66-67]

Bratman two faces on intention

Jing Zhu Intention and Volitiion [842]

Knowledge of Intention [675]

Essays on Anscombe's Intention [260]

az intenció elkötelezettséget jelent a versengő vágyak közepette, ami 1) legalább időleges állandóságot jelent és 2)más terveket lehet ráépíteni

Bratman [117]

For Bratman (1987), intention is a distinctive practical attitude marked by its pivotal role in planning for the future. Intention involves desire, but even predominant desire is insufficient for intention, since it need not involve a commitment to act: intentions are ?conduct-controlling pro-attitudes, ones which we are disposed to retain without reconsideration, and which play a significant role as inputs to [means-end] reasoning? (Bratman 1987, p. 20). The plans for action contained in our intentions are typically partial and must be filled out in accordance with changing circumstances as the future comes.

Among the advantages of being able to commit ourselves to action in advance, albeit defeasibly, are: *(i)* the capacity to make rational decisions in circumstances that leave no time to deliberate, or lend themselves to deliberative distortion; *(ii)* the capacity to engage in complex, temporally extended projects that require coordination with our future selves; and *(iii)* the capacity for similar coordination with others.

wikip: It could be argued that the BDI model is now becoming somewhat dated: the principles of the architecture were established in the mid-1980s, and have remained essentially unchanged since then. With the explosion of interest in intelligent agents and multi-agent systems that has occurred since then, a great many other architectures have been developed, which, it could be argued, address some issues that the BDI model fundamentally fails to. Furthermore, the focus of agent research (and AI in general) has shifted significantly since the BDI model was originally developed. New advances in understanding (such as Russell and Subramanian?s model of ?bounded- optimal agents?) have led to radical changes in how the agents community (and more generally, the artificial intelligence community) views its enterprise.

[625]

Cohen és Levesque formalizmust dolgozott ki Bratman intencióelméletére [164]

P.R. Cohen and H.J. Levesque, Rational interaction as the basis for communication, [166]

Casteñada míg a vélekedés tartalma propozíció, addig az intencióé egy cselekvés, amire Casteñada új fogalmat vezetett be: 'practition' praktozíció

Többek között azért érdemes ezt az új fogalmat alkalmazni, mert ezáltal elkerülhetővé válik a deontikus logikában jól ismert probléma: ha a cselekvés logikáját akarjuk kidolgozni, akkor azt kell mondanunk, hogy ha valakinek szándékában áll megtenni A-t, akkor abból logikailag következik az, hogy szándékában áll megtenni $(A \vee B)$-t, ami a cselekvések világában nem plauziblis következtetés.

3.5.3 más elméletek

Searle, Intentionality, [650]

Douglas Ehring, "Normal" Intentional Action, [227]

Bentzen intencionalitás logikája [66]

Roderick M. Chisholm [152]

Wayne A. Davis [193]

Ullmann, Tamás (2010): A láthatatlan forma. Sematizmus és intencionalitás. Budapest, L'Harmattan

Daniel C. Dennett, Az intencionalitás filozófiája [202] D.C: Dennett: tartalom és tudatosság [201]

Dennet intencionalitás koncepciója [261]

Tomasello, M. (1999). Having intentions, understanding intentions, and understanding communicative intentions. In : Zelazo, P.D., Astington, J.W., and Olson, D.R. (Eds.), Developing theories of intentions (pp. 63-75). Mahwah, NJ: Lawrence Erlbaum Associates.

3.5.4 Az intencionalitás értékelése

Intentional Action and Side Effects in Ordinary Language [403] [249]

3.6 Vágy

three faces of desire [644] desire kontra aversion

a vágy/averzió kettőshöz mindig kapcsolódik a kielégülés vs. kielégületlenség (frusztráció) jelensége/állapora

alapvető szükségletből minden embernek kettő van. Az egyik a tartozni akarok vhová. A másik a teljesíteni akarok vmit. Az első szükségletben a kötődés, a másodikban a szabadság iránti vágy nyilvánul meg

két típus:

"maintenance goals", i.e. goals about keeping things as they are "achievement goals", i.e. goals about state of things not yet realized

3.6.1 Castelfranchi: vélekedés és vágy

Cristiano Castelfranchi, Fabio Paglieri, The Role of Beliefs in Goal Dynamics: Prolegomena to a Constructive Theory of Intentions [141]

the main point is that goal processing and intention revision are largely determined by belief revision – that is, the process by which we come to intend something crucially depends on our assumptions concerning how we consider the world to be, i.e. our beliefs. s on our assumptions concerning how we consider the world to be, i.e. our beliefs. More precisely, our aim here will be to detail a model in which specific types of beliefs act as filters and triggers from goal activation to goal execution. According to this view, in order to activate, promote, drop, or suspend a goal, an intention, or an intentional action, one has to provide or modify the appropriate beliefs.

belief and goal are mental representations, two cognitive items funkcióikban különböznek egymástól (nem formájukban, tartalmukban vagy más belső jellemzőikben)

a cél fogalma kapcsolódik: several other concepts familiar both in Artificial Intelligence (e.g. desires and intentions), in cognitive psychology (e.g. motivations, wishes, aims), and in social sciences (e.g. preferences).

“a goal is defined as an anticipatory internal representation of a state of the world that has the potential for and the function of (eventually) constraining/governing the behaviour of an agent towards its realization. The defining function of goals is to shape, to direct in a teleological sense the actual behaviour of the system.” 77.o.

goals prefigure a certain state of affairs, as opposed to representing a state of the world which is believed to be the case.

“An agent is said to be rational if it chooses to perform actions that are in its own best interests, given the beliefs it has about the world. For example, if I have a goal of staying dry, and I believe it is raining, then it is rational of me to take an umbrella when I leave my house. (...) You would still be inclined to refer to my behaviour as rational even if I was mistaken in my belief that it was raining: the point is that I made a decision that, if my beliefs were correct, would have achieved one of my goals.” (Wooldridge, 2000, p. 1)

vélekedéstipizálás

a célvonatkozású vélekedéseken (goal-related belief) belül céltámogató vélekedések (goal-supporting belief)

céltámogató vélekedések (goal-supporting beliefs)

there is a specific structure of beliefs that is necessary to maintain and justify goal without such beliefs, the related goal would be dropped from its current state, and would change its nature and functional properties.

vélekedéstámogató vélekedések (belief-supporting beliefs)

- motiváló vélekedések (motivating beliefs): ezek mellett léteznek nem-doxasztikus befolyások is (pl. physical arousal, emotional triggering) - ??? vélekedések (triggering beliefs) - feltételes vélekedések (conditional beliefs)

- előszűrő/előértékelési vélekedések (assessment beliefs): ezek előszűrik a vélekedéseket, illetve kialakítják a követhető célok körét - önmegvalósulási vélekedések (self-realization beliefs) - megvalósultsági vélekedések (satisfaction beliefs) - lehetetlenségi vélekedések (impossibility beliefs)

- költségbecslő vélekedések (cost beliefs)

- inkompatibilitási vélekedések (incompatibility beliefs) - végcél inkompatibilitási vélekedések (terminal incompatibility beliefs) - instrumentális inkompatibilitási vélekedések (instrumental incompatibility beliefs) - feleslegességi imkompatibilitási vélekedések (superfluity)

- preferencia vélekedések (preference beliefs) - érték vélekedések (value beliefs) - sürgősségi vélekedések (urgency beliefs)

- előfeltevési vélekedések (precondition beliefs) - inkompetencia vélekedések (incompetence beliefs) - hiányzó feltélek vélekedések (lack of conditions beliefs)

- cél-eszköz vélekedések (means-end beliefs)

ezek a vélekedések a szerzők szándéka szerint kimerítőek, ami elsődorban az általuk “közvetített” funkciókra vonatkozik

a vélekedések listája helyett célszerű inkább : given goal is always supported by a structure of beliefs, rather than a mere list.

folyamatok:

- aktiválás (activation)

- értékelés (evaluation)

- mérlegelés (deliberation)

-ellenőrzés (checking)

- cselekvés (action)
aktív célok (active goals) = vágyak (desires)
követhető célok (pursuable goals)
választott célok (chosen goals)
végleges/végehajtandó célok (executive goals)

3.6.2 Cross: diszkrepancia alapú vágyelmélet

Cross, The Modal Logic of Discrepancy [177]
incompatibility and incompleteness
state of affairs (A,B,C,D ...)
két típus: G – goal, az Ego célállapotja
B – belief, az Ego hiedelmi állapota[1]
vágy: $G = (A \wedge B \wedge C \wedge D)$
$B_0 = (A \wedge \sim B \wedge C \wedge D)$ – incompatible
$B_1 = (A \wedge B \wedge D)$ – incomplete
"In claiming that we desire only what we lack, Socrates affirms a discrepancy between what one wants and what one has. Socrates refers only to what the agent in fact has or lacks and not (as I have) to what the agent believes she has or lacks, but the role of the notion of discrepancy in Socrates? claim is evident." Platon, Symposium 42-43.
$\bigtriangledown \bigtriangleup \otimes \odot$
Ego kielégült vagy elégedett
If $\oplus A$ is true, then the agent has no reason to act to bring it about that A because, as far as she knows, no action is needed.
inkompatibilitási diszkrepancia:
$\bigtriangleup A$ is true iff the agent's goal entails A, but her belief state entails $\neg A$
$\bigtriangleup A$ means that the agent has a reason to bring it about that A or the agent needs to bring it about that A.
inkomplettségi diszkrepancia:
Ego whose goal requires A neither accepts A nor accepts $\neg A$
$\bigtriangledown A$ is true iff the agent?s goal entails A, but neither A nor $\neg A$ is entailed by her belief state. Informally,
$\bigtriangledown A$ means that the agent has a reason to make sure that A or the agent needs to make sure that A.
az első diszkrepancia a vágy megragadására szolgál, a második nem, mert az első késztetést ad cselekvésre (to bring about the truth of A), míg a második késztetést csak ahhoz ad késztetést, hogy a cselekvő megbizonyosodjék A állapotáról
indifferencia:
$\odot A$ is true iff the agent accepts A, but her goal requires neither A nor $\neg A$.
Informally, $\odot A$ means that the agent indifferently accepts that A

3.6.3 magasabbrendű vágyak

Wayne A. Davis [194]
Kris McDaniel and Ben Bradley, Desires [483]

[1] Előfordul a bizonyossági állapot kifejezés is. mit.bme.hu mesterséges intelligencia könyvre hivatkozni.

Eleonore Stump, Sanctification, Hardening of the Heart, and Frankfurt's Concept of Free Will[712]

David Lewis, Desire as Belief, [441] Desire as Belief II, [442]

Timothy Schroeder, Reflection, Reason, and Free Will, [645]

Timothy Schroeder, Three Faces of Desire [644]

Contours of Agency: Essays on Themes from Harry Frankfurt, [134]

John Martin Fischer (ed.), Free Will: Free agency, moral responsibility, and skepticism, Routledge, 2005

Stephen Schiffer, A Paradox of Desire, [641]

Scott Sturgeon, Normative Judgement, [713]

Kenneth A. Taylor, De Re and De Dicto: Against the Conventional Wisdom, [742]

A. L. Herman, A Solution to the Paradox of Desire in Buddhism, [352]

Christine Swanton, Weakness of Will as a Species of Executive Cowardice, [720]

Harry G. Frankfurt, The Importance of What We Care About, [267]

3.7 Másodrendű vágy: érték

> Egyszer, amikor Diogenész Kraneioszban sütkérezett a napon, Nagy Sándor ezt mondta neki: "Kérj tőlem, amit akarsz!" Erre ő: "Ne álld el előlem a napot!"

Copp, Reasons and Needs [172]

David Lewis koncepciója: az érték mint másodrendű vágy [697]

Gilbert Harman áttekinti Frankfurt, Lewis és Copp másodrendű vágy mint érték értelmezését és elveti őket

[341]

Lewis koncepcióját összevetni Fitch javaslatával. Frederic B. Fitch 'érték' definíciója [258]

$(value\ p) \equiv \exists q \exists r[q \wedge [(knows\ q)\ C\ (strives\ for\ [p \wedge r])]]$

ahol C - partial causation

Charles R. Pigden, Desiring to Desire: Russell, Moore and David Lewis: [572]

Weber, Cs. Kiss Lajos érték hivatkozása

magyar jogász

Harry Frankfurt, másodlagos vágy, másodlagos akarat

Charles Taylor sources of the self [741] Frankfurt másorendű vágy fogalmára támaszkodik

Reasons from Within. Desires and Values [301]

Why don't you want to be rich? Preference explanation on the basis of causal [311]

Ahhoz, hogy megértsük, mik az értékek, a szervezet viselkedéséből kell kiindulnunk. Már ezen a szinten találkozunk preferenciákkal, azaz pre-reflektív megegyezésekkel a szervezet szükségletei és a környezet adottságai között. A természeti állapot nem jelent apátiát. A preferenciák nem tudatos döntésekből erednek. A tudatos döntések viszont elkerülhetetlenek, amikor a preferenciák konfliktusba kerülnek egymással. Emiatt, Dewey számára az organikus impulzusok csupán szükséges, de nem elégséges feltételei annak, hogy a vágyak cselekvésbe lendüljenek. Maguk a vágyak is módosulhatnak a reflexió hatására. Az értékek nem egyszerűen a vitális háttér elemei, hanem a reflexió

által alátámasztott vágyak. Értékekről akkor beszélhetünk, amikor felmerül egy konfliktus. Csak amikor problémák jelentkeznek, csak akkor szükséges tudatosan értékelni, hogy az általunk reflektálatlanul követett értékorientációk közül melyik élvez elsőbbséget. Különbséget kell tennünk aközött, amit egy cselekvő kíván, és aközött, amit reflexió következtében ?kívánatosnak? tart.

Dewey Experience and Nature könyvében [204] Úgy tekint a kommunikációra, mint egy eseményre, amelyben a személy megnyílik mások előtt. Mivel a kommunikáció közös, megosztott tapasztalathoz vezet, úgy tekinthetjük, mint egy út az egocentrizmus meghaladása felé. Az altruizmust Dewey úgy írja le, mint azt a radikális attitűdöt, amelyben készen vagyunk arra, hogy a többiek ?megrázzanak?, ?megmozgassanak? , annak érdekében hogy megvalósítsuk önmagunkat a többiekkel együtt és a többieken keresztül ? mint ?megrázó? interszubjektivitás. (shattering intersubjectivity). Mivel a kommunikációra, interszubjektivitásra és a megosztott tapasztalatra fókuszál, kiutat talál a individuális és a kollektív tapasztalat inadekvát dualizmusából. Ugyanakkor azt állítja, hogy az interszubjektivitás tapasztalata nem csupán a különleges - vallásos - élményekben adott, hanem a mindennapi kommunikációs helyzetekben is. Ezen az úton jut el ahhoz a gondolathoz, hogy maga a demokratikus tapasztalat is értékteremtő. A demokrácia Dewey szekularizált vallásává válik.

horváth gizella, érték-tanulmánya, 32-34: Taylor értékelmélete Frankfurt másodlagos vágy koncepciójára épül gyenge és erős értékelések

önkifejezés, identitásfenntartás mint alapérték (és mint magyarázó tényező):

> „ . . . a politikai cselekvés különbözni látszik a magánfogyasztók és termelők piaci viselkedésétől. Amellett is érvelhetünk, hogy a politikai cselekvés természeténél fogva publikus, abban az alapvető értelemben, hogy nemcsak egyéni nyereményeken alapul, hanem a világ lehetséges jövőbeli állapotaira értelmezett preferenciákon, amelyek magukban foglalhatják mások nyereményeit, valamint szabályokra és intézményekre vonatkozó preferenciákon, amelyek garantálhatják ezeket a nyereményeket. Bizonyos politikai cselekvések, beleértve azt a döntést, hogy a választó elmegy szavazni, lehetnek egyszerűen nyilvános önkifejezési formák, amelyek egyetlen szerepe az, hogy nyilvánosan kifejezésre juttathassák a kérdéses preferenciákat.” [180, 567]

a politikában a szenvedélyes kisebbségek erősebb hatással bírnak [411]

3.7.1 Osgood

az elemzés szempontjából releváns dimenzió is - értékelés (evaluation): a helyes kategorizálás

- az üzenetek értékelésének gyenge-erős (potency) minősége
- a passzív-aktív (activity) vetülete

evaluation, potency, activity

3.7.2 Axiológiai hatszög

3.8 Akarat

akarat és érdek szembeállítása,

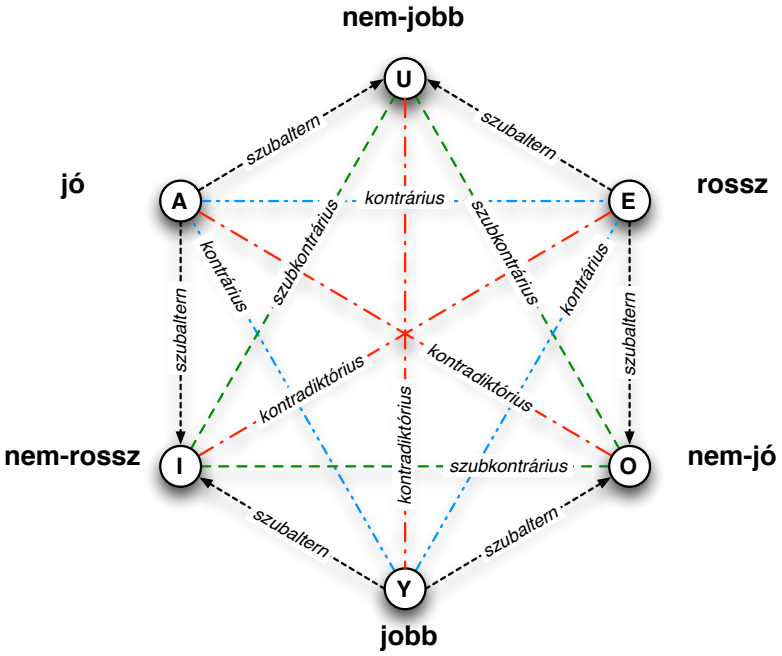


Figure 3.2: értékfogalmak hexagonja

az érdek fogalmának homályossága
gazdasági érdek és nem-gazdasági érdek konfliktusa
rövidtávú és hosszútávú érdek konfliktusa
egyéni és csoportérdek konlfiktusa
Schoppenhauer az akarat szabadságáról szóló könyvében deklarálta:

> „Tehetsz, amit akarsz: de élted minden adott pillanatában csak egy határozott dolgot és szükségképen csakis ezt az egyet akarhatod, semmi mást. [643, 25]

> „Az öntudatnak egyetlen tárgy az akarati tény és annak feltétlen uralma a test tagjai felett, ezt értjük voltaképpen azon, hogy tehetek, »amit akarok«. Csupán ennek az uralomnak a gyakorlása, vagyis a *tett*, bélyegzi (az akaratot), még az öntudat számára is, akarati ténynek. Mert amíg keletkezőben van, addig *kívánság* a neve, ha kész, akkor *elhatározás*. De hogy valóban elhatározás, azt még az öntudatnak is csak a *tett* bizonyítja be, mert mindaddig változékony. . . .
> *Kívánni* lehet ellentétest, de *akarni* csak az egyiket közülük: még pedig hogy melyiket, azt az öntudatnak is csak a tett teszi nyilvánvalóvá.” [643, 19]

elsőrendű akarat, másodrendű akarat
akaratkonfliktus
kényszer
erőszak
Harry Frankfurt az állatok szándékolt cselekvéséről [266]
“Tehát részben akarunk és részben nem akarunk, nem szörnyűséges valami, hanem a lélek betegsége ez. Nem ugrik egészen talpra, midőn az igazság emeli, mert lenyűgözi a

megszokások terhe. Ezért két akarat létezik, az egyik közülük nincs meg a maga teljes valóságában, s ami hiányzik belőle, a másikban lapul ... ” [36, 234]

Benjamin Libet készenléti potenciál (RP - readiness potential), nincs szabad akarat, már a tudatosulás előtt dönt az agy

Libet modelljének újraértelmezése, a szabad akarat lehetősége Jeff Miller és Judy Trevena új kísérlete

3.9 Egyéni döntések

A preferencialogika bemutatásakor már elkülönítettem egymástól a tiszta preferencia és a preferenciális választás fogalmait egymástól (↬109.o.), és ezt a megkülönböztetést érdemes szem előtt tartani az egyéni döntések jelenségének tárgyalása során.

a preferenciák konstrukciója: sok helyzetbe úgy megyünk be, hogy nincsenek igazán rögzült előzetesen preferenciánk, hanem maga helyzet alakítja ki azt bennünk [446]

How to decide what to do? [187]

3.9.1 Racionalitásértelmezések

Amennyire fontos fogalom a racionalitás, annyira sokféleképpen próbálják meg a fogalom jelentését megragadni, a definícióját megadni. Ez nyilván nem véletlen. Az eltérő szakmai nézőpontok máshová teszik a hangsúlyt. Mielőtt áttekintenék párat a legfontosabb racionalitásértelmezések közül, szükségesnek tartok egy előzetes elkülönítést megtenni. Az általam továbbra is alapvetőnek tartott – és a későbbiekben még elemezni kívánt (↬280.o.) – max weberi cselekvéstipizálásában kiemelt szerepet van a racionalitás fogalmának, és ezt a weberi felfogást fontosnak tartom ütköztetni a későbbi racionalitásértelmezésekkel. Weber négy cselekvéstípust különített el egymástól: a cél- és értékracionális, valamint az érzelmi és a habituális cselekvéseket [810]. Bár Weber pontosan definiálta az egyes cselekvéstípusokat, de arra nem reflektált, hogy lehet-e, és ha igen, akkor milyen dimenziók alapján lehet a négy cselekvéstípust elrendezni egymáshoz képest egy fogalmi térben. Erre többen is kísérletet tettek, én Bertalan László javaslatát emelném be a társadalmi cselekvés modelljébe [77].

Úgy gondolom, hogy nincs miért meghaladni Max Weber híres cselekvéstipológiáját, ami az alábbi négy cselekvéstípust különítette el egymástól [810]:

- értékracionális
- célracionális
- érzelmi
- habituális

racionális = cél- és értékracionális aracionális = érzelmi és habituális)

mérlegelés és a racionalitás különbsége

Bertalan figyelmeztet, hogy a weberi racionalitásfogalomnak van egy szűkebb és egy tágabb értelmezése. A szűkebb értelmezés szerint racionálisnak csak a célracionálist tarthatjuk, és a további három típus , tehát az értékracionális, az érzelmi és a habituális valamilyen mértékig irracionális. A tágabb értelmezés szerint racionális a célracionális és az értékracionális, míg irracionális az érzelmi és a habituális.

A cselekvő beállítódásának jellemzésekor az egyik fontos szempont az, hogy a cselekvő milyen szempontokra figyel, egyáltalán: van-e, amire reflektál, ha igen, mik ezek,

és mit tesz akkor, ha több szempontot is figyelembe kell, figyelembe akar venni. A mérlegelő beállítódás fogalma szükségszerűen feltételezi, hogy a cselekvő több szempontra figyel. A mérlegelésnek enélkül nincs értelme. Eléggé nyilvánvalónak látszik, hogy Weber is ezért nevezte racionálisnak az értékracionálist, mert a mérlegelés mozzanata hozzátartozik.

Mérlegelő magatartás a tágabb értelemben vett racionális celekvést jelent, amivel szembe állítható a nem-mérlegelő vagy aracionális magatartás. A szűkebb értelemben vett racionalitásfelfogás alapján csak a célracionális cselekvés racionális, amihez képest minden más beállítódás irracionális.

A habituális cselekvés során nincs mérlegelés, mert a cselekvő nem is igazán reflektál saját magára, a cselekvés rutinszerű, automatikus. Az érzelmi cselekvésben már tetten érhető valamiféle tudatosság, reflektáltság, de ekkor sincs még "szükség" mérlegelésre, mivel a cselekvő irányultságában az egyetlen érzelmi motívum van a fókuszban, semmi más.

"elborult az agyam", "semmit sem mérlegeltem", "elveszettem az ítélőképességemet", "kontroll nélkül", "mentem a farkam után", "elragadott az indulat", "elveszettem a fejem"

a mérlegelés ítéletet (judgement) hoz, előtte "dönt", a nem-mérlegelés nem hoz ítéletet

Mondhatjuk úgy is, hogy az aracionális cselekvések azért nem-mérlegelő cselekvések, mert a semmit és az egyet nem lehet mérlegelni.

később a véleményaggregálás (judgement aggregation) a mérlegelő magatartás körébe fog tartozni

Szűkebb értelemben vett racionalitás

teleológikus

$(B \to G) \to B$

$(M \to G) \to M$

$(B \to ((M \to G) \to M)) \to B$

a racionális vélekedési képesség fogalmához a rejtett, személyes tudás fogalmát axiomatizálni kell

John Broome három racionalitásfeltételt kötött ki [125].

(CL.0) $(\mathbf{B}_i(\phi) \wedge \mathbf{B}_i(\phi \to \psi)) \to \mathbf{B}_i(\psi)$

modus ponens (HUN) – Modus Ponens (ENG)

(CL.1) $(\mathbf{B}_i(M \to G) \wedge \mathbf{I}_i(G) \wedge \mathbf{B}_i(\mathbf{I}_i(M) \to M)) \to \mathbf{I}_i(M)$

szükséges eszközök (HUN) – necessary means (ENG)

(CL.2) $(\mathbf{B}_i(\mathbf{O}F) \wedge \mathbf{B}_i(\mathbf{I}_i(F) \to F)) \to \mathbf{I}_i(F)$

krasia (HUN) – krasia (ENG)

Az első broome-i feltétel értelmezése: ha az i személy hiszi, hogy ϕ, és hiszi, hogy $\phi \to \psi$, akkor i hiszi, hogy ψ. Ez nem jelent mást, mint hogy alkalmazza a modus ponens következtetési szabályt, vagyis képes a következetések levonására, a logikus gondolkozásra. Vannak, akik szerint ez kevés, a racionalitáshoz a logikai helyesség önmagában még nem elég. Evans-Pritchard a zandék törzsénél végzett vizsgálatai alapján azért különböztette meg a tudományos és a logikus gondolkozást, hogy rámutathasson, a törzsi társadalmak tagjai is logikusan (de még nem tudományosan) gondolkodnak.

> „A tudományos képzetek azok, amelyek összhangban vannak az objektív valósággal mind a premisszáik érvényességét, mind az állításaikból levont következtetéseket tekintve . . . A logikai képzetek azok, amelyekben a gondolkodás szabályainak megfelelően az igaz premisszákból levont következtetések is igazak, a premisszák igazsága lényegtelen. . . . Egy fazék eltört égetés közben, valószínűleg a homoknak köszönhetően. Vizsgáljuk meg a fazekat, nézzük meg, tényleg ez volt-e az ok. Ez logikus és tudományos gondolkodás. A betegséget a boszorkányok okozzák. Egy ember megbetegedett. Kérdezzük meg a jósdát, melyik boszorkány a felelős érte. Ez logikus és tudománytalan gondolkodás.” [235]

Evans-Pritchard megkülönböztetésével egyet lehet érteni, de ez az egyetértés még nem mond semmit arról, mit tartsunk racionálisnak, a logikus vagy a tudományos gondolkozást. Ha a tudományos gondolkodást követelnénk meg, azzal túl erős feltételt teremtenénk. Ráadásul a tudománytörténetből rengeteg olyan példát hozhatnánk, amikor egy tudományosan megalapozott premissza igazságáról kiderült annak hamissága, és ilyen esetekben egy darabig racionálisnak, majd utána nem-racionálisnak kellene tartanunk ugyanazt a hitet. Ezért én azt az értelmezést fogadom el, amely szerint – a most vizsgált kontextusban – a racionális feltételéül a logikus gondolkodás képességét kell megszabnunk.

Broome második feltétele annyit jelent, hogy ha az i személy úgy véli, hogy az általa megvalósulni kívánt G cél akkor teljesül a jövőben, ha az M eszköz is megvalósul, és i úgy gondolja, hogy M teljesülése az ő szándéktól függ, akkor i akarni fogja M megvalósulását. Magyarán: ha valamilyen célt el akarunk érni, és tudjuk, hogy a cél elérésének eszközeként valamit meg kell tenni, ami a mi szándékunkon múlik, akkor az utóbbit akarnunk kell megtenni. Ez a feltétel összeköti a célra irányuló cselekvések megvalósításának szándékát és a célok eléréséhez szükséges eszköz jellegű cselekvésekre irányuló szándékokkal.

A harmadik Broome-féle kritérium szerint ha i úgy véli, hogy meg kell tennie F-t, miközben úgy gondolja, hogy F-t csak akkor teheti meg, ha akarja megtenni, akkor i-nek akarnia kell megtenni F-t. Ez a feltétel a cselekvés ??????

Stalnaker kétféle racionalitásfogalma [707]

Balkányi Péter, racionalitás és cselekvés

[45]

Robert J. Aumann, Rule-Rationality versus Act-Rationality [41] Rationality and Bounded Rationality [39]

P.R. Cohen and H.J. Levesque, Speech acts and rationality, [163]

közgazdaságtanban:

objective function - célfüggvény = preferencia

Logic as General Rationality: [631]

Harsányi János szerint a racionalitás szabályokhoz való következetes, koherens, konzisztens igazodást jelent

Advances in Understanding Rational Behavior [343]

a cél-eszköz racionalitás leszűkítő értelmű, mert nem foglalja magába a kritériumhoz igazodó viselkedést (felmászni a legmagasabb hegyre példa) A másik probléma ezzel a modellel az, hogy csak az adott cél eléréséhez szükség alternatív eszközök mérlegelését teszi lehetővé, és ezáltal nem fedi le az alternatív célok közti racionális választás jelenségét (tehát nem lehet megmagyarázni, ha valaki egy adott cél helyett valami mást választ,

tehát a változás, a dinamika hiányzik a rendszerből)

közgazdaságtan hozadéka: ehelyett jobb az a modell, amelyikben vannak preferenciák (preferences), lehetőségek (opportunities), és az ágens utóbbiakból választ az előbbiek alapján. Amennyiben az alternatív lehetőségek halmazából valaki kiválasztja az egyiket, akkor ez egyben azt is jelenti, hogy a többi lehetőséget nem választja, és a nemválaszott lehetőségekben rejlő kívánatosság-potenciált értékelhetjük a választási egyfajta költségeként is. A dolgokat így értékelve egyszerűen megmagyarázhatjuk azokat a jelenségeket, amikor valaki akkor is célt vált, ha az alternatív célokra vonatkozó preferenciái változatlanok maradtak, egyszerűen csak azért, mert az alternatívák költségei (még pontosabban: az effajta költségekről való ismeretei) megváltoztak

a preferencia-lehetőség modell magába foglalja mind a cél-eszköz modellt, mind a kritérium-teljesítő modellt ha az ágens preferenciái megfelelnek bizonyos konzisztencia és folytonossági feltételeknek, akkor ez a modell lehetővé teszi, hogy a preferenciáinkat egy **U hasznosságfüggvénnyel** reprezentáljuk, és ebben az esetben a racionális cselekvés a **haszonmaximalizáló** viselkedést jelent.

döntéselmélet hozadéka nem mindig, sőt, inkább ritkán cselekszünk biztos tudás birtokában, az életünket sokkal jobban jellemzi az a tény, hogy **bizonytalanság** (uncertainty) és **kockázat** (risk) mellett kell döntéseket hoznunk.

[458]

(CL.3) CERTAINTY $\stackrel{\text{def}}{=}$

bizonyosság

(CL.4) UNCERTAINTY $\stackrel{\text{def}}{=}$

bizonytalanság

(CL.5) RISK $\stackrel{\text{def}}{=}$

kockázat pl. lottózás várható valószínűsége pontosan ismert, tehát ezt kockázatnak kell minősítenünk

A döntéselmélet lényege – Harsányi szerint– az, hogy ha elfogadunk pár konzisztenia és folytonossági axiómát, akkor mondhatjuk, hogy a racionálisan cselekvő ágens a saját szubjektív várható hasznossága maximalizására törekszik ennek modellezéséhez a bizonyosság menti racionális cselekvés axómái mellé fel kel még venni egy újat, a (sure thing principle) 95.o.

de ez a modell még mindig nem elég az össze racionális cselekvés leírásához, hiszen a játékelméleti helyzetekre nem alkalmazható a játékokban bizonytalanság van, hiszen nem tudhatjuk, hogy a másik játékos hogyan lép a játék során

és a modellnek le kell még fednie az etikai cselekvések világát is az etika világát Harsányi úgy definálja, hogy racionális morális értékítéletek elmélete, ami alatt azt kell érteni, hogy az ágensek pártatlan (impartial) és személytelen (impersonal) kritériumokra alapozott preferenciák közül való racionális választást (ítélkezést) ért

amíg a játékelmélet a potenciálisan konlfiktusos egyéni érdekek (amelyek lehetnek egoista és altruista irányultságúak!) elmélete, addig az etika a társadalom, a közösség mint egész közös érdekeinek elmélete (98.o.)

játékelmélet

komplett vs. inkomplett információ

perfekt vs. imperfect információ

az inkomplett játékokból komplett, bár imperfekt játékokat lehet transzformálni, ha a kölcsönösen konzisztens elvárások (mutually consistent expectations) feltételét fenntartjuk

kooperatív és nem-kooperatív játékok megkülönböztetése (J.F. Nash két cikkében) kooperatív játékokban egyrészt a játékosok tehetnek kikényszeríthető elköteleződéseket, egyezségeket, másrészt szabad a kommunikáció köztük.[343] Harsányi szerint az előbbi szempont az erősebb, bár az utóbbit szokták inkább hangsúlyozni (a fogolydilemma csak akkor lesz igazából kooperatív játék, ha kikényszeríthető egyezséget lehet kötni a játékosok között, a kommunikáció, a megegyezés lehetősége még kevés).

nem-kooperatív játékokban ezért csak az önérvényesítő megegyezéseknek (self-enforcing aggreements) lehet esélye a sikerre. Ezt nevezik egyensúlyi pontnak (equilibrium point) (Nash 1950, 1951) legjobb válasz (best reply): az a lépés (stratégia), amelyik a legjobb kifizetést eredményezi a játékosnak akkor ha feltételezzük, hogy a másik játékos stratégiája konstans (nem változik). egyensúlyi pont = a játékos legjobb válaszaiból álló lépéskombináció

Nash egyensúlyi pont fogalmának bevezetése után azt gondolták, hogy ez elégséges és megfelelő feltétel lehet a racionalitás meghatározásához [343] Selten 1965-ben olyan példát mutatott, ami irracionális egyensúlyi pont

	$Alter_1$	$Alter_2$
Ego_1	1,3	1,3
Ego_2	0,0	2,2

Két egyensúlyi pont van: $EqP_1 = (Ego_1, Alter_1)$ és $EqP_2 = (Ego_2, Alter_2)$, de csak az utóbbi **perfekt egyensúlyi pont**, és csak is ez reprezentálhatja a racionális cselekvést.

ezt a játék extenzív formájú reprezentációja képes megmutatni

a magyarázat: a második (B) játékosnak az egyik döntési ágon nincs valódi döntési lehetősége (döntetlen számára a helyzet, hiszen a saját döntésétől független a saját és a társa nyereménye

perturbed game reszkező kéz: nem úgy lép, ahogy tervezte, akarta volna az ilyen játékban a nem-perfekt egyensúlyi pont elveszti ezt a státuszát, csak a perfekt egyensúlyi pont képes megtartani

bargaining game [343] nem mindegy, hogy milyen szabályok vannak a tárgyalási folyamatra - ki beszélhet, kinek - ki kezdheti a tárgyalást - milyen a tárgyalások nyilvánossági foka (nyilvános vagy privát) - visszavonhatók-e a korábbi egyezségek, van-e mód újratárgyalásra vagy végleges, visszavonhatatlan ajánlatokat engednek csak meg - lehet-e egyoldalúan kötelező ígéreteket vagy fenyegetéseket tenni vagy sem

ha van három játékos, és három kifizetési feltétel: 1) egyedül játszó játékos nem nyer semmi 2) a kétszemélyes koalíció 100 dollárt nyer 3) háromszemélyes koalíció 100 dollárt nyer

milyen szabályok érvényesek ezen felül? a) ha a játékosok elkülönülten páronként is találkozhatnak, akkor az első – mindegy, hogy kik között – létrejött találkozó szereplői vélhetőleg azonnal koalíciót formálnak és egyenlően elosztják maguk között a nyereményt b) ha nincs mód szeparált találkozókra, és nincs elég idő a hármas találkozón arra, hogy kettős koalíció alakulhasson ki, akkor a három fél megegyezehet a hármas

koalícióban, és harmadolhatják az összeget c) más feltételek mellett kialakulhat bármilyen koalíció, de ezt nem lehet előre megjósolni

Harsányi 1969-es cikke [342]

racionalitásértelmezések - adott cél(ok)hoz kiválasztani a legmegfelelőbb eszközt az (eszköz)alternatívák közül - célok közül kiválasztani a legmegfelelőbbet egy konzisztens preferenciaskála segítségével (az egyes lehetséges célokhoz rendelt relatív fontosság vagy hasznosság szerint) - ezáltal dinamika kerül a modellbe

a közgazdaságtan azért volt sikeres, mert egy komoly magyarázóerővel rendelkező motivációs elméletet húzott rá a gazdasági cselekvések magyarázatára ez egyszálú motivációs magyarázat volt, mert az gazdasági önérdek mint egyetlen létező motiváció fogalmára támaszkodott, ami ugyan sok területen működőképesnek bizonyult, de még a gazdaság területén sem lehetett ezzel a túlegyszerűsített modellel mindent megmagyarázni. vannak nem-gazdasági és nem-egoista motívumok is. A gazdaságon túli világban, például a politikában még inkább igaz ez

ha bővíteni akarjuk a modellt azzal, hogy más motívumokat is figyelembe veszünk, akkor felmerül a kérdés, hogy honnan tudhatjuk meg az emberekről, hogy milyen relatív súlyokat, hasznosságokat rendelnek az egyes célokhoz, eszközökhöz, gazdasági és nem-gazdasági érdekekikhez, normáikhoz stb. meg kell figyelni az emberek tényleges viselkedését, és az így szerzett empirikus tényekből fel lehet állítani hasznosságfüggvényeket, majd ezek alapján már meg lehet magyarázni az emberek további cselekvéseit. ez akkor nem körbenforgó érvelés, ha az empirikus megfigyelésből szerzett paraméterek száma hangsúlyosan kevesebb, mint a megmagyarázni kívánt jelenségek, paraméterek száma

olyan motivációs elméletre van szükség, amly nem egydimenziós, tehát nem csak a gaudasági önérdek motivációjára épít, de nincs is túl sok új motivációra sem

négy posztulátum (az első három az önző és önzetlen motivációk kérdésével foglalkozik, a negyedik)

1) pártatlanság, a közjó követése alacsony költségek esetén (postulate of ‘low-cost’ impartiality and public spirit)
   amikor olyan helyzetekben vagyunk, amikor közvetlenül nem vagyunk érdekeltek, akkor hajlandók vagyunk pártatlanok lenni, a köz érdekében, altruista módon fellépni, ha nem túl megerőltető (a bírót kivonják az ügyintézésből, ha érdekeltsége van egy ügyben, de ha nincs, akkor képes részrehajlás nélkül működni – enélkül nem lenne működőképes közigazgatásunk), vagy elmegyünk szavazni, hogy racionális módon gondolkodva nem éri meg nekünk (de csak azért, mert kevés fáradságba kerül)
2) konfliktusos helyzetben érdekaggregálás pártatlan szimpatizánsok segítségével (postulate of interest-aggregation by impartial sympathetic third-parties) érdekellentétes helyzetekben a felek közötti kompromisszumot segíthet elérni a pártatlan harmadik fél ez magyarázza meg a csoportérdek, osztályérdek, nemzetérdek fogalmát, jelenségét ez adja morál alapját is: olyan megoldásokat keresni a konfliktusok feloldására, amely minden ember érdekére tekintettel van
3) mások iránti személyes elköteleződések elsajátítása racionális választások révén (postulate of acquiring personal commitments by rational choice) az ember személyes elköteleződései (a család, a munkatársak, abarátok stb. iránt) altruizmust eredményez(het)nek. Ezek az elköteleződések valamilyen racionális választás eredményeiként jöttek létre korábban. Miután az ember így elköteleződött, utána ezek

a mindenkori racionális döntéseket, de ezek az elköteleződések korábban racionális döntési helyzetekben születtek

4) gazdasági és társas elfogadási motívumok (postulate of economic and social-acceptance motivations) az emberi cselekvések alapvetően két domináns érdekkel magyarázhatóak: gazdasági nyereségek és társas elfogadás (aconomic gain and social acceptance). Ez a két fő céltárgyunk.

irracionalitások

a társadalom statikus és dinamikus magyarázata, utóbbi a fontos talán legfontosabb feladat megmagyarázni a társadalmi értékek aktuális eloszlását, amit csak a múlt ismerete alapján tehetünk meg

az értékváltás legtöbbször a feltételek változása miatt következik be (a nagycsalád adta előnyökért sokáig megéri vállalni a sok gyereket, de idővel már többet jelent az embereknek a kényelmetlenségektől megszabadulni, tehát kevés gyereket vállalni

az értékváltás sokszor puszta tudatlanságból fakad, mert nem tudunk eleget az új alternatíváról, és a régi ismertségéből fakadó biztonságérzet erősebb

mik az értékváltások legnagyobb akadályai (disincentives) mások szankciói (leginákbb azokl, akik érdekeltek a változatlanságban)

érdek lehet egista és altruista ez szemben áll sokak felfogásával (pl. Hischmann: az érdek ésszerű önzés)

az intézmények társadalmi funkciói nem igazán a kollektívista, hanem inkább az individualista álláspontról magyarázhatók meg

Jack Hirshleifer azt hangsúlyozza, hogy a racionalitás elsősorban módszert, nem pedig eredményt jelent.

> "A racionalitás instrumentális fogalom. Ha valaki saját céljait (preferenciáit) figyelembe véve megfelelő eszközöket (cselekvéseket) választ, akkor racionális; ha nem, akkor irracionális. A 'megfelelő' itt inkább a módszerre, mint az eredményre vonatkozik. A racionális viselkedés olyan cselekvés, amit a logika és az érvényesség más normái szerint mérlegelve hajtanak végre. A változásoknak köszönhetően pedig egy jó megoldás nem mindig vezet jó eredményhez." [358]

sok tanulmány van benne a racionalitásról, elköteleződésről [570]

Horton, Robin, African Traditional Thought and Western Science, in: Wilson (ed.): Rationality, pp.131-171.

Habermas szerint egy "megnyilatkozás racionalitása a bírálhatóságára és a megindoklásképességére vezethető vissza" [316, 16]. Ez a definíció új szálat hoz be a racionalitásértelmezések világába. Legalábbis akkor, ha kicsit továbblépünk egy egyénhez kötött racionalitás fogalmán. A habermasi javaslat (ami persze visszavezethető lenne legalább Weberig) ugyanis a racionális diskurzus feltételeként megadott bírálhatóságot és megindoklásképességet egy új dimenzióban is értelmezhetjük. A kommunikációban megnyilvánuló racionalitás, a nyilvánosság racionalitása azt (is) jelenti, hogy nincs semmilyen tilalom a megnyilatkozásokra vonatkozóan.

mit jelent egy vélemény? privát vagy nyilvánoaan kinyilvánított (vállalt) vélemény utóbbiba már beleszámít a közösség véleménye is

backward induction csapás a racionalitásra penny piching [362]

H. Simon **korlátozott racionalitás** tézise [681] az ember saját korlátozott információfeldolgozó képesség miatt nem tud teljesen racionális lenni

Sen racionalitás és társadalmi választás: [667]

Thomas Schwartz Rationality and the Myth of the Maximum [648]

Suzumura racionalitásfelfogása: 19. o. környéke [718]

[105]

List-Pettit racionalitásértelmezése: A Φ csoportos véleménykészlet (collective set of judgement) racionális, ha teljes, konzisztens és deduktíve zárt (↬256.o.).

Málik és Mészáros racionalitásfelfogására is hivatkozni.

Suzumura:

racionlitás feltétele purposive, azt szándékozzon tenni, ami neki a legjobb nem tranzitivitás, mert az túl erős, de valamilyen optimalitás a tranzitivitással kapcsolatban

választás

Ahhoz, hogy a választás fogalmát definiálni tudjuk, szükségünk van pár kisegítő fogalomra. A választás alternatívák, lehetőségek összehasonlítását, majd az összehasonlítási szempont szerinti legjobb elem(ek) kijelölését jelenti. A választás fogalmának meghatározása előtt így meg kell tudnunk mondani, mit jelent a legjobb elem fogalma. Ha adva van egy halmazunk, amin értelmezve van egy bináris reláció, akkor mindig megvizsgálhatjuk, hogy az adott reláció szerinti rendezés eredményeként lesz(nek)-e legnagyobb elem(ek). Ha igen, akkor ezek segíthetnek minket a választás fogalmának meghatározásához. Látni kell azonban, hogy ezek a tulajdonságok elsődlegesen nem magára a relációra, hanem a tartóhalmaz egy vagy több elemére vonatkoznak, de a legjobb, legnagyobb elem fogalma megkívánja a rendezési reláció létét a definícióban, vagyis ezek az extremitást kifejező fogalmak a halmaz és reláció fogalmára egyaránt támaszkodik. Mégpedig a következő módon [718].

(RL.6) $\text{GREATEST} \equiv G(S,\mathbf{R}) \stackrel{\text{def}}{=} (S \subset \mathbb{X} \wedge \mathbf{R} \subset \mathbb{X} \times \mathbb{X}) \rightarrow \{m \in S | \forall n \in S : (m,n) \in \mathbf{R}\}$

(R-)legnagyobb (elem) (HUN) – (**R**-)greatest (element) (ENG) – legjobb (elem) (HUN) – best (element) (ENG) – G-optimális (elem) (HUN) – G-optimal (element) (ENG) – (felülről) domináns elem (HUN)

(RL.7) $\text{MAXIMAL} \equiv M(S,\mathbf{R}) \stackrel{\text{def}}{=} () \rightarrow \{m \in S | \forall n \in S : (n,m) \notin \mathbf{P}(\mathbf{R})\}$

(R-)maximális (elem) (HUN) – (**R**-)maximal (element) (ENG) – M-optimális (elem) (HUN) – M-optimal (element) (ENG) – (felülről) nem domináns elem (HUN)

Mind a legnagyobb, mind a maximális elem valamiféle extremitást jelent. Ha meg akarjuk érteni azt, hogy miben különböznek egymástól, akkor ebben leginkább az segíthet, ha felidézzük a preferenciastruktúrára vonatkozó tételeket, azok közül is leginkább a szigorú preferenciarendezés, valamint annak konverze, az indifferencia és az összehasonlíthatatlansági reláció mint négyes partíció létezésére vonatkozó tételt (↬98.o.). Utóbbi szerint az említett négy reláció egymást kölcsönösen kizárja, és együttesen pedig kiteszik az univerzális relációt. Ebből viszont az következik, hogy amíg a legnagyobb elemek halmazába azok az elemek tartoznak, amelyek kielégítik a szigorú preferenciarendezés, illetve az indifferencia reláció feltételeit, addig a maximális elemek halmaza azokból az elemekből áll, amelyek nem felelnek meg a konverz szigorú preferenciarendezésnek, vagyis amelyek kielégítik vagy a szigorú preferenciarendezés vagy az indifferencia vagy az összehasonlíthatósági reláció feltételeit. Ebből perdig nyilvánvalóan az következik,

hogy a legnagyobb, illetve a maximális eleme halmazok között az összehasonlíthatatlansági reláció jelenti a különbséget. Ezt úgy is megfogalmazhatjuk, hogy a legnagyobb eleme halmazban azok az elemek vannak, amelyek minden más elemnél "nagyobbak" és egymás között indifferencia reláció áll fent, a maximális eleme halmaz pedig azokat az elemeket tartalmazza, amelyek elemei között vagy az indifferencia vagy az összehasonlíthatatlansági reláció érvényesül. A legnagyobb elemek tehát mind indifferensek egymáshoz képest, a maximális elemek vagy indifferensek vagy összehasonlíthatatlanok. Az igazi különbség tehát az összehasonlíthatatlansági relációban van.

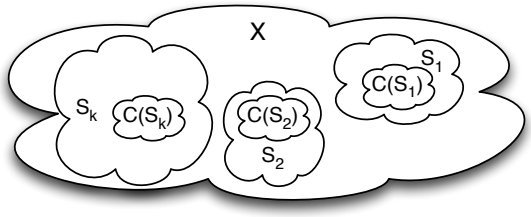
A két extremitás között két további fontos összefüggést is megállapíthatunk [105]. Egyfelől igaz az, hogy $G(S,\mathbf{R}) \subset M(S,\mathbf{R})$, vagyis a legnagyobb elemek halmaza mindig része a maximális elemek halmazának, másfelől bizonyítható az a tétel is, hogy ha egy reláció teljes és reflexív, akkor a maximális és legnagyobb elemek halmaza megegyezik.

Az extremitások, azaz legnagyobb (legjobb) és maximális elem fogalmaira támaszkodva már meg lehet határozni a **választási függvény** fogalmát, de előtte még tisztázni kell pár dolgot. Első lépésként be kell vezetni a $(\mathbb{X},\mathscr{S})$ **választási tér** fogalmát, ahol $S \subset \mathbb{X}$ $(S \in \mathscr{S})$ az alternatívák teljes $\mathbb{X}$ halmazán képzett részhalmazokat, vagyis az ágens számára a választás során elérhető állapotok halmazait jelentik, $\mathscr{S}$ pedig az S részhalmazok hatványhalmazát. A **választási halmaz** a választási tér valamelyik konkrét S halmazának egy nem-üres részhalmaza, és a **választási függvény** a választási tér választási halmazra való leképezését végzi el.

(CL.8) $\text{CHOICE_FUNCTION} \stackrel{\text{def}}{=} \forall S \in \mathscr{S}(C : \mathscr{S} \mapsto C(S) \mid C(S) \neq \varnothing)$

választási függvény (HUN) – choice function (ENG) – döntési függvény (HUN) – decision function (ENG)

A választási függvény alapján előállt választási halmazok elemei még nincsenek rendezve. Vannak a kiválaszottak, azok az elemek, amelyek $C(S)$-ekbe tartoznak, és vannak a választhatók, azok az elemek, amelyek S-ekbe tartoznak.



Itt még nem feltétel a választási függvény egyeleműsége, lehet benne több kiválaszott elem is. Ha ez fordul elő, akkor döntetlentörő szabályok szerint lehet eldönteni a végleges (egyetlen) választást, de ez nem jelent igazi probléma, nem is nagyon fogom elemezni. Érdemes viszont pár példát bemutatni a választási függvényre, mégpedig olyanokat, amelyek majd szembeállíthatók a később bevezetendő racionális választási függvény fogalmával. A lottósorsolás esetében az alternatívák halmazát a leadott szelvények jelentik, és a választás a véletlen alapján történik meg. A résztvevők megtippelik a számokat, és a gép véletlenszerűen választ a számok közül, de ezzel voltaképp a leadott szelvények (és áttételesen az azokat birtokló személyek) közül választ ki egyet vagy többet. Ez a példa mutatja azt is, hogy a választási halmaz fogalma még megengedi, hogy a procedúra több

kimenetet adjon. A döntetlentörő szabály itt annyi, hogy a nyereményt annyi részre kell osztani, ahányan bekerültek a választási halmazba. A lottóhúzáshoz képest csak egy ponton van különbség a tombolajáték esetében, amikor a leadott szelvények közül – megintcsak véletlenszerűen – egyetlen győztest húznak ki a "jelöltek" közül.

A véletlen választás elve mellett létezik más nem-racionális módszer is. Például az általános iskolai osztályokban a hetes szerepébe a névsor szerinti kerülnek be a tanulók. Ilyenkor is választásról beszélünk, hiszen egy alternatívahalmazból egy (vagy több) elemet kiválasztunk. Meg lehetne szavaztatni mindenkit arról, hogy ki legyen a hetes, lehetne pénzfeldobással dönteni, és lehet a nevek kezdőbetűi alapján. Ezt a mechanikus elvet egyébként alkalmazzák, alkalmazták a politika világában is. Amikor még létezett Jugoszlávia, akkor a tagországok vezetőiből álló államelnökség tagjai közti úgy döntötték el, hogy ki legyen az államelnök egy adott időszakra, hogy az államok neve szerinti sorrendben váltották egymást a tagországok vezetői. De a görög városállamokban is előfordult, hogy bizonyos tisztségek betöltésére a névsor szerint jelölték ki a polgárokat egymás után.

Ezek a példák szépen mutatják azt is egyben, hogy a választási függvény létezése önmagában még nem "garantálja" annak racionális jellegét. Pedig a racionalitás feltételének érvényesülését jó lenne elvárni. Ezt a minőséget valamilyen gyenge rendezési preferenciarendezési reláció, illetve az ennek alapján képzett legnagyobb elem halmaz választási függvényhez rendelésével biztosíthatjuk.

(CL.9) $\text{G-RATIONAL_CHOICE_FUNCTION} \stackrel{\text{def}}{=} \exists \mathbf{R} \subset \mathbb{X} \times \mathbb{X}\ \forall S \in \mathscr{S}(C(S) = G(S, \mathbf{R}))$

(G-)racionális választási függvény (HUN) – (G-)rational choice function (ENG)

Akkor racionális egy választási függvény, ha az általa kiválasztott választási halmaz megegyezik egy adott gyenge preferenciarendezés szerinti legnagyobb elemek halmazával. A racionalizálást tehát a preferenciarendezés "végzi el" (mondhatjuk is azt, hogy az adott preferenciareláció a választás/döntés racionalizálása). Ez azért is tűnik szerencsés megoldásnak, mert így jobban kezelni tudjuk azt a vitatott kérdést, hogy ti. milyen kritériumai is vannak a racionális választásnak. Ha megengedjük, hogy a preferenciareláció többféle lehessen, akkor nem kell kizárólagos módon állást foglalnunk ebben a kérdésben, és a fenti meghatározáshoz képest elfogadhatunk másfajta racionalitásértelmezéseket is. Például ha a másik extremitásfogalmat, a maximális elem kategóriáját alkalmazzuk itt, akkor máris más racionális választási függvényhez jutunk.

(CL.10) $\text{M-RATIONAL_CHOICE_FUNCTION} \stackrel{\text{def}}{=} \exists \mathbf{R} \subset \mathbb{X} \times \mathbb{X}\ \forall S \in \mathscr{S}(C(S) = M(S, \mathbf{R}))$

(M-)racionális választási függvény (HUN) – (M-)rational choice function (ENG)

Van tehát egy választási függvényünk, van egy hozzárendelhető preferenciarelációnk, amire vonatkozóan különböző minőségű és erősségű tulajdonságokat tételezhetünk abban a reményben, hogy ezáltal jellemezni, tipizálni tudjuk a választói döntéseket, beállítódásokat.

> „ . . . akkor tekintünk egy aggregált gazdasági szereplőt – tehát a döntési függvényét – racionálisnak, ha bevezethető olyan reláció az alternatívák halmazán, amely szerinti optimalizáló választás ugyanazt eredményezi, mint a döntési függvény szerinti választás. Világos, hogy ebben a keretben annyiféle racionlitás vethető fel, amennyi tulajdonsága lehet a racionalitást biztosító relációnak, így a tranzitivitás ennek csak egy fontos speciális esete." [463]

Ez a hozzáállás azt is lehetővé teszi, hogy a preferenciarelációra vonatkozó feltételek mentén különböző erősségű racionalitáskritériumokat állítsunk fel a választási függvényre vonatkozóan. Hogy ezt megtehessük, szükségünk van még egy fontos fogalomra, amivel kapcsolatot teremthetünk a választás és racionalitáskritériumot "szállító" reláció között. Ehhez fel kell idéznünk a közgazdaságtan egyik központi kategóriáját, a **kereslet** fogalmát. A közgazdászok ezt a fogalmat nagyon mélyen beágyazták a mikroökonómia szövetébe, ami – sok minden mellett – azért is különösen fontos volt a közgazdaságtan számára, mert "külsőleg" lehetett mérni – szemben a preferencia (vagy a hasznosság) fogalmával, amit meg nem. A kereslet változását a jószágok fogyasztásán keresztül "objektíven" követhetjük, mérhetjük, a jószágokra vonatkozó "belső" hasznosság szubjektíven érezhető, a megfigyelőnek nincs direkt hozzáférése. A közgazdászok körében sokáig vitatott volt, hogy a hasznosságot (preferenciát) és keresletet hogyan lehet összekapcsolni egymással. Az áttörést a – Paul A. Samuelson és Hendrik Hoathakker nevéhez köthető – **kinyilvánított preferencia** fogalmának megjelenése hozta el [414]. Samuelsonék újítása az volt, hogy a – keresletet generáló – fogyasztók vásárlásait, választásait értelmezhetjük úgy, mint saját belső preferenciáik kinyilvánításait, és az egyének választásainak teljes halmaza voltaképp az egyének preferenciáit adja meg. Erre már fel lehet írni egy összefüggést, és akkor a preferencia fogalma "kezelhetővé" válik. A választás és a kinyilvánított preferencia közti kapcsolat persze többféleképpen is értelmezhető, adhatunk egy erős és egy gyenge meghatározást, az alábbiak szerint.

(CL.11) $\text{WEAKLY_REVELALED_PREFERENCE} \equiv \mathbf{R}_c \stackrel{\text{def}}{=} \bigcup_{S \in \mathscr{S}} [C(S) \times S]$

---

gyengén kinyilvánított preferencia(reláció) (HUN) – weakly revealed preference (relation)(ENG) – Richter-reláció (HUN) – Richter-relation (ENG)

(CL.12) $\text{STRONGLY_REVELALED_PREFERENCE} \equiv \mathbf{R}_c^* \stackrel{\text{def}}{=} \bigcup_{S \in \mathscr{S}} [C(S) \times \{S \setminus C(S)\}]$

---

erősen kinyilvánított preferencia(reláció) (HUN) – strongly revealed preference relation (ENG) – Samuelson-reláció (HUN) – Samuelson-relation (ENG)

Mindkét relációnak ugyanaz a lényege: létezik az alternatívák teljes készletének egy olyan halmaza, amelynek m is, n is eleme, és ezek közül az egyén az előbbit preferálja az utóbbival szemben, vagy másként: az egyén mindkét elemet választhatta volna, de ténylegesen az m-t választotta. A kétféle értelmezés csak abban tér el egymástól, hogy az egyikben megszorítás van a preferenciareláció második argumentumára vonatkozóan (a formulában $S \setminus C(S)$ szerepel a S helyett), míg a másikban nincs. Ezt úgy interpretálhatjuk, hogy az egyén n-t választhatta volna, de nemcsak, hogy nem választotta, de aktuálisan még vissza is utasította (ezt fejezhetjük ki azzal a szűkítéssel, hogy n mindig csak az S-nek azon részhalmazában szerepelhet, amely az S alaphalmaz és a $C(S)$ választási részhalmaz különbsége).

A továbbiakban azonban a választási függvénnyel, a kinyilvánított preferenciarelációval kapcsolatos racionalitásértelmezéseket Kotaro Suzumura megközelítését elfogadva mutatom be, aki könyvében precízen és kimerítően tárgyalja ezt a témakört [718]. Történelmileg Samuelsonék axiómái voltak az elsők, amelyek kényszereket fogalmaztak meg a kinyilvánított preferencia fogalmára vonatkozóan. Samuelson feltételét el is nevezték gyenge axiómának, míg Houthakkerét erősnek. Ezek mellé Suzumura definiált egy har-

madik kritériumot is, amit Houthakkerről nevezett el, és a három fogalmat az alábbi módon definiálta.

(CL.13) $(\text{WARP}) \stackrel{\text{def}}{=} \forall m \forall n \in \mathbb{X}\left(\mathbf{R}_c^*(m,n) \rightarrow \neg \mathbf{R}_c(m,n)\right)$

kinyilvánított preferencia gyenge axiómája (HUN) – Weak Axiom of Revealed Preference (ENG) – Weak Axiom of Consumer Behaviour (ENG)

(CL.14) $(\text{SARP}) \stackrel{\text{def}}{=} \forall m \forall n \in \mathbb{X}\left(\left[\mathbf{R}_c^*\right]^+(m,n) \rightarrow \neg \mathbf{R}_c(m,n)\right)$

kinyilvánított preferencia erős axiómája (HUN) – Strong Axiom of Revealed Preference (ENG) – Strong Axiom of Consumer Behaviour (ENG)

(CL.15) $(\text{HARP}) \stackrel{\text{def}}{=} \forall m \forall n \in \mathbb{X}\left(\left[\mathbf{R}_c\right]^+(m,n) \rightarrow \neg \mathbf{R}_c^*(m,n)\right)$

kinyilvánított preferencia Houthakker axiómája (HUN) – Houthakker's Axiom of Revelaed Preference (ENG)

Mindhárom tulajdonság esetén szerepel a gyenge és erős értelemben kinyilvánított preferencia a definícióban, és kétszer alkalmazni kell az adott relációra az $\mathbf{R}^+$ tranzitív lezárás műveletét is (↬45.o.).

Amartya K. Sen a 1970-es években írt cikkeiben több racionalitáskritériumot is megfogalmazott, és ezeket α-, β, γ-tulajdonságoknak nevezte el. Maga Sen is megjegyezte, hogy az α-tulajdonság megegyezik Herman Chernoff által definiált tulajdonsággal, amit a tétel megjelenése óta Chernoff-axiómának neveznek [146]. Elsőre bemutatom, ahogy Sen megadta az α-tulajdonság definícióját, majd másodjára melléteszem azt a formulát, ahogy Suzumura fejezte ki Chernoff axiómáját.

(CL.16) $(\text{SEN-}\alpha) \stackrel{\text{def}}{=} \forall m\left(\left(m \in S_1 \wedge S_1 \subset S_2\right) \rightarrow \left(m \in C(S_2) \rightarrow m \in C(S_1)\right)\right)$

Sen α-tulajdonsága (HUN) – Sen's α-property (ENG) – irreleváns alternatívák függetlensége (HUN) – independence of irrelevant alternatives (ENG) – (basic) contraction consistency (ENG)

$(\text{CA}) \stackrel{\text{def}}{=} \forall S_1 \forall S_2 \in \mathbb{S}\left(S_1 \subset S_2 \rightarrow \left(S_1 \cap C(S_2) = \varnothing \vee S_1 \cap C(S_2) \subset C(S_1)\right)\right)$

Chernoff axiómája (HUN) – Chernoff's axiom (ENG) – örökítő (HUN) – öröklődő (HUN) – megőrző (HUN) – preserving (ENG)

Sen azt javasolta, értelmezzük úgy ezt a tulajdonságot, hogy ha a világbajnok egy pakisztáni, akkor ennek a világbajnoknak Pakisztán bajnokának is kell lennie [662]. Ez a kritérium az örökítésről, örökletességről szól. Ha egy nagyobb halmazból kiválasztunk valamit "győztesnek", akkor annak bármely olyan részhalmazából is ki kell választanunk ugyanazt a győztes, amely részhalmazok tartalmazzák azt a legjobb elemet. Sen megjegyezte még, hogy erre feltételre használják az irreleváns alternatívák függetlensége kifejezést is, de fontos tudni, hogy ilyenkor ezt más értelemben teszik, mint ahogy az a társadalmi választások elméletében szokás (↬210.o.). Itt az egyén racionalitásának jellemzésére, ott a kollektív preferencia kialakítását végző eljárás egyik kritériumaként használják ezt a terminust. Kenneth J. Arrow egy 1959-es cikkében definiált egy tulajdonságot [32], ami egyetlen apró mozzanatban tért csak el a Chernoff axiómától. Arrow tulajdonságát később Suzumura elnevezte Arrow-axiómának, és a következő formában mutatta be.

(CL.17) $(\text{AA}) \stackrel{\text{def}}{=} \forall S_1 \forall S_2 \in \mathbb{S}(S_1 \subset S_2 \rightarrow (S_1 \cap C(S_2) = \varnothing \vee S_1 \cap C(S_2) = C(S_1)))$

(választás konzisztenciájának) Arrow axiómája (HUN) – Arrow's axiom (of choice consistency) (ENG) – Arrow's independence condition (ENG) – erősen örökítő (HUN) – erősen megőrző (HUN) – strongly preserving (ENG)

Az a pici kis eltérés, ami megkülönbözteti a Chernoff- és az Arrow-axiómákat egymástól, a formulák utolsó tagjában látható: eltérő műveleteket írnak elő a választási függvények között. A különbség miatt nyilván nem vonhatjuk össze a két tulajdonságot, de a hasonlóság miatt az Arrow-feltételt is megőrzési tulajdonságnak tarthatjuk, olyannak, ami az örökletességre erősebb feltételt támaszt a Chernoff-axiómához képest. Más nézőpontból tekintve a dolgokra kicsit másfajta interpretációt is adhatunk az örökítés problémájára. Az egyszerűség kedvéért egyelemű halmazokra váltva: ha van egy m alternatívánk, amit előnyben részesítünk n-nel és o-val szemben, akkor ez a beállítódás nem szabad, hogy megváltozzon attól, hogy még egy elemet "beveszünk" az összehasonlításba. Vagyis nem lehetséges az, hogy egy p elem felvétele után már nem m-t preferáljuk n-nel vagy o-val szemben. Mindezt egy gyakran idézett gyakorlati példával szemléltetve: ha a csoki- és vaníliafagyik közül a csokit választjuk, majd ezután közlik velünk, hogy van még málna is, akkor – ha konzisztensek akarunk maradni – nem mondhatjuk másodjára, hogy akkor mégis vaníliát kérünk.[2]

Sen az általa definiált β-tulajdonságot a következő formában határozta meg.

(CL.18) $(\text{SEN-}\beta) \stackrel{\text{def}}{=} \forall m \forall n((m, n \in C(S_1) \wedge S_1 \subset S_2) \rightarrow (m \in C(S_2) \leftrightarrow n \in C(S_2)))$

Sen β-tulajdonsága (HUN) – Sen's β-property (ENG)

Sen javaslata értelmében ezt a minőséget úgy interpretálhatjuk, hogy ha a pakisztáni bajnokok közül legalább egy világbajnok is egyben, akkor az összes pakisztáni bajnoknak világbajnoknak kell lennie [662]. Ha bővítjük az alternatívák halmazát, és az eredeti – szűkebb terjedelmű – halmaz legjobb (maximális) elemei közül egy a bővítés után is bekerül a legjobbak (maximálisak) közé, akkor a "korábbi" társaival is ennek kell megtörténnie.

———

Sen könyv, 18., 50. oldalon a β-tulajdonságról

———-

Sen definiált egy harmadik fontos tulajdonságot is, amit először γ-tulajdonságnak [664], majd később a **kiterjesztés konzisztenciájának** nevezett el [658]. Sen az új kritériummal azt akarta biztosítani, hogy ha az alternatívák tetszőleges halmazaiból ugyanazt az elemet mindig kiválasztják, akkor a halmazok uniójából is ki kell választani.

(CL.19) $(\text{EC}) \stackrel{\text{def}}{=} \forall i(S_i \in \mathscr{S} \rightarrow (\bigcap_i C(S_i) \subseteq C(\bigcup_i C(S_i))))$

kiterjesztési konzisztencia (HUN) – (basic) expansion consistency (ENG) – Sen γ-tulajdonsága (HUN) – kiterjedő (HUN) – Sen's γ-propertiy (ENG) – összehangoltság (HUN)

[2]Természetesen az új alternatívát (példánkban a málnát) preferálhatjuk a régi opciókkal szemben.

Igazodva a Sen által korábban megadott interpretációk "műfajához", azt adhatjuk meg a γ-tulajdonság értelmezéseként, hogy ha valaki minden országban megnyeri az országos versenyt, akkor a végső összesítésben is nyernie kell, tehát a világbajnoknak is neki kell lennie. Sen megjegyezte azt is [664], hogy ha csak a választási részhalmazok véges halmazaira fókuszálunk, akkor érdekes komplementaritást vehetünk észre az α- és γ-tulajdonságok között. Mindez akkor válik különösen szembetűnővé, ha az α-tulajdonság formuláját kicsit átalakítjuk.

(CL.20) $(\text{SEN-}\alpha) \stackrel{\text{def}}{=} \forall i(S_i \in \mathscr{S} \rightarrow (\bigcup_i C(S_i) \subseteq C(\bigcap_i C(S_i))))$

A Sen által definiált tulajdonságok segíthetnek a racionális választással, választási függvénnyel szemben támasztott elvárások értelmezésében. A racionalitáskritériumok értelmezésében és meghatározásában azonban Marcel Richter hozzájárulása fontosabbnak mondható [604]. Ő vezette be a **kongruencia** fogalmát, ami gyorsan az egyik legelfogadottabb racionalitáskritériummá vált. A kongruencia fogalmának kétféle (erős és gyenge) értelmezést is lehet adni – az alábbi módon:

(CL.21) $(\text{WCA}) \stackrel{\text{def}}{=} \forall S \in \mathscr{S}(\forall m \in S \wedge (\exists n \in C(S) : \mathbf{R}_c)) \rightarrow m \in C(S)$

---

gyenge kongruencia axiómája (HUN) – Weak Congruence Axiom (ENG)

(CL.22) $(\text{SCA}) \stackrel{\text{def}}{=} \forall S \in \mathscr{S}(\forall m \in S \wedge (\exists n \in C(S) : [\mathbf{R}_c]^+)) \rightarrow m \in C(S)$

---

erős kongruencia axiómája (HUN) – Strong Congruence Axiom (ENG)

A kétféle tulajdonság meghatározása csak abban tér el egymástól, hogy amíg a gyenge kongruencia a Richter-reláció, addig az erős kongruencia a Richter-reláció tranzitív lezártjának való megfelelést írja elő a formulában.

(CL.23) $(\text{FR}) \stackrel{\text{def}}{=}$

---

teljes racionalitás (HUN) – Full Rationality (ENG)

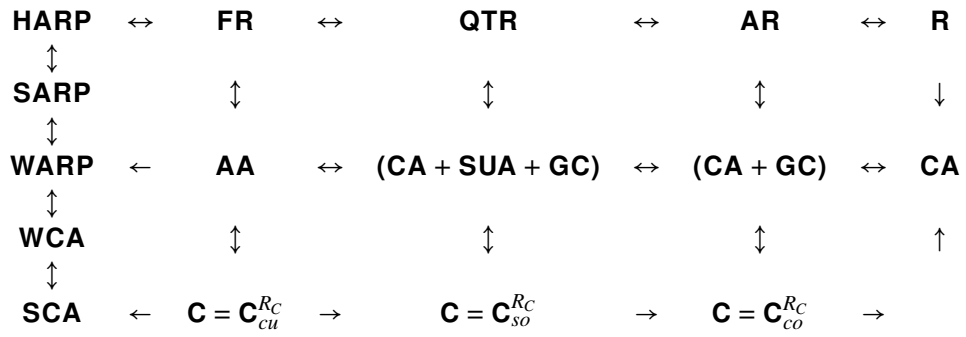
útfüggetlenség axiómái
Sen átfogó cikke, 1973: [665]
Magyarkúti Gyula 2001 [463]

A racionalitáskritériumok között érvényes összefüggéseket tételszerűen nem mutatom itt be, Suzumura társadalmi választásokról szóló könyvében minden tételt kimond és be is bizonyít [718]. A könyvből viszont átemelek egy grafikus ábrázolást, amely szemléletesen mutatja a kritériumok közti kapcsolatrendszert.

bodó beáta

döntési mechanizmus tulajdonságai: döntésképes (decisive) Richter 1971 stabil - stable valódi - real injektív - injective tökéletes - pefect

10. oldal: reláció megszorítása, esetleg felvenni ezt is? 11. aciklikusság definiciója 11 reláció nem összehasonlítható része dominánsan racionalizálható D-rac nem dominánsan racionalizálható ND-rac 27. rac. kritériumok öröklődő (preserving) összehangolt (coordinated?)



választás és norma

Amartya K. Sen az 1990-as években fenntartásokat fogalmazott meg a racionális választás addigra egyeduralkodóvá vált paradigmájával szemben [658], [668], és a társadalmi normák figyelembevételének fontossága mellett érvelt. Szemléletes példája a következő volt. A racionális választás paradigmája alapján – akármelyik racionalitáskritériumot is fogadjuk el – az alábbi két választást leíró választási függvényt inkonzisztensnek kell minősíteni:

$$C(S_1) \equiv C(\{m,n\}) = \{m\}$$
$$C(S_2) \equiv C(\{m,n,o\}) = \{n\}$$

Pedig könnyen kerülhetünk olyan helyzetbe, amikor teljesen racionális magyarázatot adhatunk arra, hogy miért választunk látszólag nem konzisztens módon, ha bővítjük a választható alternatívák halmazát. Gondoljunk csak arra – javasolta Sen –, amikor egy gyümölcsöstálban egyetlen alma van. Ekkor a választást értelmezhetjük úgy, hogy m jelenti azt, hogy nem vesszük el az utolsó almát a tálból, míg n-nel jelölhetjük azt az opciót, amikor elvesszük az almát. A tényleges gyakorlatban sokszor azt tapasztalhatjuk, hogy az emberek nem veszik el az almát, tehát m-t választják. És ha nem veszik el az almát, akkor azt azért teszik, mert érvényesnek tartják azt a társadalmi normát, amely tiltja az utolsó gyümölcs elvételét a tálból. Amennyiben viszont van egy másik alma is a tálban, amit a harmadik o alternatíva "felvételével" fejezhetünk ki, akkor ez a norma már nem tiltja valamelyik alma elvételét a kettőből. Elvehetjük akár azt is, ami előzőleg még tiltva volt, tehát ekkor választhatjuk az n alternatívát.

Sen 1977-es cikkében már arról írt, hogy az eredeti konzisztenciaelvárásokat kétféleképpen is megsérthetjük, noha ilyenkor is racionálisnak érezhetjük a cselekvéseinket [668]. Az egyik lehetőséget **választófüggőségnek** (chooser dependence), a másikat **menüfüggésnek** (menu dependence).

Kotaro Suzumura és Walter Bossert a normativitás jelenségének megragadására, a **normafüggő választás** (norm-conditional choice) modellezésére azt javasolta, hogy az eredeti választási modellt ki kell egészíteni a normák tiltó hatását kifejező kivételhalmazokkal [104].[3] Ezt úgy tehetjük meg, hogy a normák által tiltott m_k alternatívákat összekapcsoljuk adott választási halmazokkal (S_k, m_k), és ezekből létrehozzuk az $\mathbf{N} = \{\bigcup_k (S_k, m_k)\}$ norma alapú kivételhalmazunkat.

8.o. nem érdemes bővíteni a halmaz-alapú kivételkezelésre

[3] Lehetne a 'normakondicionált választás' kifejezését is használni itt, ez pontos fordítása lenne az angol terminusnak, de én jobbnak tartom a normavezérelt minősítést. Jobban kifejezi a norma választást befolyásoló "jelenlétét".

normafüggő racionalitás (norm-conditional rationality)

Sen: commitment + preference [657]

később ide alkalmazni Adrékáék elméletét? amikor a norma jelen van, akkor a kivételt a BUT operátort kell alkalmaznunk, egyébként az ON_THE_OTHER_HAND

[105]

Társadalmi választások elmélete

Sen 77-es áttekintése: [666]

3.9.2 döntési szabályok

Amikor valaki egyetlen szempont szerint akar választani az alternatívák halmazából, akkor ezt megteheti a szemponthoz "illeszkedő" preferenciareláció segítségével. Amikor azonban több szempont is felmerül a váalsztás során, akkor a választás döntési helyzetté alakul, ami annyiban lesz több és más, hogy az eltérő szempontok más és más rendezéseket eredményezhetnek. Ezt nyilván kezelni kell valahogy, amit azonban többféleképpen tehetünk meg. A döntéselmélet – legalább részben – ennek a problémának a tárgyalására jött létre. A sokféle szempont egyeztetése során lehetséges, hogy a szempontokat egyenrangúaknak, nem rendezetteknek tekintik, és lehetséges, hogy különböző szintekre sorolják, vagyis rendezik őket. Nem célom itt az egyéni döntési eljárások, módszerek, szabályok ismertetése.[4] A probléma érzékeltetése végett röviden bemutatok három jól ismert, a későbbiekben is használni kívánt döntési stratégiát, a maximin, a maximax és a minimax szabályt.

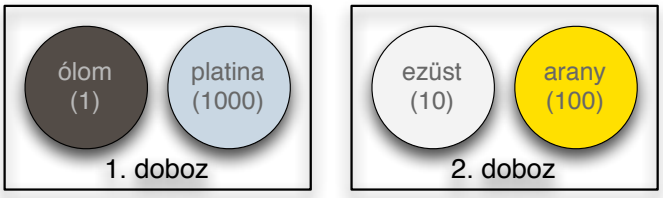
A **maximin szabály**, amit neveznek még pesszimista, biztonságra törekvő, kockázatkerülő módszernek, Wald-féle Wald kritériumnak, a legrosszabbak legjobbikát választja ki az alternatívák közül. A szabály lényege, hogy első körben kiválasztja az alternatívák egyes szempontok szerinti leggyengébb értékelését, majd második körben ezen leggyengébb elemek közül kiválasztja a legjobbat. A **maximax szabály** vagy másként: az optimista módszer első körben kiválasztja az alternatívák egyes szempontok szerinti legerősebb értékelését, majd második körben ezek közül veszi a legjobbat. Ez módszer tehát a legjobbak legjobbikát keresi meg. A **minimax szabály** megkeresi az egyes szempontok szerint várható legrosszabb eredményeket, és ezek közül kiválasztja a legkisebb veszteséget ígérő alternatívát. Ez a szabály kockázatsemleges, az elmaradó haszon minimalizálására törekszik, arra, hogy a döntéshozónak a végén a legkisebb legyen a megbánása. A módszert nevezik még Savage vagy Savage-Niehans-kritériumnak is. A maximin és a maximax szabályok közti különbség szemléltetésére érdemes megidéznem Carl Hempel példáját ide [350].

A maximin szabály szerint első lépésben megkeressük a két doboz minimumát, ami az ólom (1) és az ezüst (10), második lépésben pedig vesszük a minimumok közül a legnagyobbat, ami az ezüst. A maximax szabály először megkeresi a két doboz maximumát, ami itt a platina (1000) és az arany (100), majd másodjára kiválasztja utóbbiak maximumát, vagyis a platinát.

---

mentális keret, framing elmélet

[4] A témakört jól bemutatja ez a magyar nyelvű tankönyv: [744].



az emberek döntő többsége kockázatvállalóként viselkedik, ha rizikós döntésében lehetséges veszteségekkel szembesül, de kockázatkerülővé válik, amikor lehetséges nyereségek között választ. ebben van manipulációs potenciál, ezt átvinni, vagy meghivatkozni a manipulációs részbe

3.9.3 Viselkedésgazdaságtan: “döntési irracionalitások”

Ariely, viselkedésgazdaságtan mindig szűk keresztmetszetek vannak (lehetnek) a figyelem (attention) nagyon fontos, mert abból csak egy van, tehát nagyon szűkös erőforrás

Dan Ariely, Kiszámíthatóan irracionális [28] a nehéz döntéseknél azt választjuk, amit elénk tesznek (donor felajánlások Európában)

előzetes döntés van a beteg műtétjéről, de “kiderül”, hogy egy vagy két gyógyszert még nem próbált ki az orvos. Mit tesz? Ha egy gyógyszert kell kell kipróbálnia, akkor visszahívja a műtőből, ha kettőt, akkor nem. AKkor az újabb döntéskényszer terhe miatt az alapértelmezett döntést hagyja érvényesülni

Economist előfizetési lehetőségek: print 85 dollár online 125 dollár online & print 125 dollár nagy arányban a kettős ajánlatot választották

de amikor a középső ajánlat elmaradt, akkor sokkal inkább a print opciót kérték. Amikor nem tudunk dönteni, gyenge a preferenciarendezésünk, akkor az elénk tett, sugallt javaslatot fogadjuk el.

Schelling fogalma: precommitment

a döntés során az egyik alternatíva miatti veszteség jobban fáj, mint a másik miatti nyereség, a veszteségre érzékenyebbek vagyunk

előfordul, hogy nem mérjük fel a nem-döntés káros következményeit Buridán szamara helyzetbe kerülünk

előfeszítés, előhangolás jelensége: a vélekedéseink, elfogultságaink befolyásolják az érzéseinket, észleléseinket,döntéseinket, választásainkat prekondíció, előítélet, sztereotípia hasznos is tud lenni, mert nem lehet mindig, mindent a nulláról indítva megítélni, dönteni

Megfigyelt “különös” döntési mechanizmusok (valószínűségek becslése) - Az emberek intuitív értékelései gyakran szisztematikusan eltérnek a Bayes elvtől - Az emberek gyakran túlértékelik a “konkrét” adatokat - Az embereknek általában hibás elképzelésük van a “véletlenről” - Az emberek túlságosan a saját tapasztalatukra támaszkodnak - Az emberek sokszor kiválasztanak egy információt és ahhoz ragaszkodnak, ahhoz igazítják értékelésüket - Az emberek gyakran túlértékelik saját "képességüket" bizonytalan események “bejóslására”

Herbert Simon, korlátozott racionalitás

„Egy olyan világban, amelyben viszonylag szűkösek az információk, és amelyben csupán néhány, egyszerű döntési problémát kell megoldani, az információ majdnem mindig

pozitív jószágnak számít. Egy olyan világban, amelyben a figyelem az egyik legszűkösebb erőforrás, az információ költséges luxuscikk is lehet, mivel figyelmünket a fontos dolgokról kevésbé fontosakra terelheti." (Simon, 1982, 81. o.)

SIMON, H. (1986). Rationality in Psychology and Economics, Journal of Business, 1986, Vol. 59, No. 4, pp.2.

SIMON, H. (1955). A Behavioral Model of Rational Choice, Quarterly Journal of Economics, 69 (February), pp. 99?118.

keretezés TVERSKY, A. ? KAHNEMAN, D. (1981). The framing of decisions and the psychology of choice. Science, 211 (4481), pp. 453?458.

RABIN, M (2008). Pszichológia és közgazdaságtan, Alinea Kiadó ? Rajk László Szakkollégium.

preferenciarendezéssel kapcsoaltos kísérletekről, a tranzitivitáshoz való viszonyról [744]

Chapter 4

Közösségi döntések

4.1 Aggregálás mint közösségi döntési technika

Brian Barry [55]

harc (combat) alku(dozás) (bargaining) eszmecsere, tárgyalás (discussion on merits) szavazás (voting) véletlen sorshúzás (chance) versengés, verseny (contest) autoritatív eljárás (authoritative determination)

A társadalmi választások elméletének fókuszában az a folyamat áll, amelynek keretében (és eredményeként) az egyéni választásokból (döntésekből) közösségi választás (döntés) lesz valahogy. Közösségi döntést egyéni döntésekből többféleképpen lehet "csinálni". A történelmileg leggyakoribb megoldás a **döntési privilégium** kialakítása és fenntartása, amikor egy társadalmi norma kijelöli az emberek valamely körét, akiknek a kezébe teszi a közösség egészét érintő döntés jogát. Ezt a lehetőséget - bármennyire is fontos a történelemben - itt nem vizsgáljuk. A másik közösségi döntési mód az érintettek közötti **konszenzus** megteremtése, vagyis olyan mechanizmus kidolgozása és működtetése, amelynek révén mindenki ugyanazt a döntést fogadja el egységesen. Mivel ez a döntési mechanizmus nagyon erős feltételen nyugszik, hisz a konszenzus megteremtése sokszor nagyon idő- és energiaigényes folyamat, a módszer fontos eszköze a **vétó** technikájának alkalmazása. A teljes pozitív konszenzus eléréséhez képest gyengébb a negatív konszenzus, amikor ha nem is teljes mértékben értenek egyet a felek, de legalább nem ellenzik a vita során javasolt dolgokat. Hogy ilyen alkalmak során ne lehessen eltérni a konszenzus követelményétől, a vétójog az utolsó eszköz az egyet nem értők kezében. A konszenzusos döntések fontos típusát jelentik a **kompromisszum** alapú döntési eljárások, amelyekről Max Weber mint kompromisszumtestületekről ír [810]. A kompromisszum elve "gyengített" konszenzus elvként minősíthető, és akkor alkalmazzák, amikor azonos társadami szerepek mentén csoportosuló emberek meg kell, hogy egyezzenek egymással úgy, hogy a csoportok (vagy azok képviselői) között fenntartják a konszenzus elvét, de a csoport tagjai egyénileg valamennyien már nem nem vesznek részt a vitában. A kompromisszumtestületek működésére az érdekképviseleti szervek (állami szervek, munkaadók, illetve munkavállalók képviseleti szervei) közti tárgyalások szolgáltatják a legjob példát. Ilyenkor a döntéshozó testületnek olyan konszenzusos döntést kell hoznia, ami minden oldal számára elfogadható, ami szinte mindig valamilyen kompromisszum eredménye. A kompromisszum jelenségét találjuk meg a stratégiai szavazás, azon belül is a szavazatcsere intézménye mögött (↬226.o.). Amikor mindenki egyenrangú félként beleszól(hat)

a közös ügyek megvitatásába, akkor már szavazótestületekről, **szavazás** alapú döntéshozatalról beszélhetünk. Itt mindig valamilyen többségi elv alapján születik a döntés úgy, hogy valamilyen szavazási eljárással összegzik az egyéni véleményeket egyetlen közösségi döntéssé. A továbbiakban erről a közösségi döntési módszerről lesz szó.

A szavazás elve

A szavazás tehát a közösségi döntéshozatal egyik fontos, bár nem kizárólagos módszere. A szavazásra jogosult szavazók leadják a szavazataikat és egy összegző eljárás segítségével közös döntést hoznak ki – ezt jelenti a szavazás fogalma. Az egész döntési folyamatban sok elemzendő kérdés van, amelyek közül párat később még tárgyalni fogok. Bármennyire is fontos kérdés, hogy pontosan milyen is a szavazási eljárás menete, hogyan bonthatjuk fel összetevőkre a szavazás folyamatát, itt nem foglalkozom ezzel. Létezik egy "bonyolult", sok összetevős és sok lépésből álló folyamat, amely az emberek fejében levő sok-sok egyéni véleményt valahogyan egyetlen közösségi véleménnyé aggregálja. Ehhez először a véleményt szavazattá kell alakítani, aztán – olykor, ha szükséges – a véleményt polarizálni kell (eközben preferencia esetén azt kell figyelni, hogy az egyén mit preferál, faktuális vélemény esetén hogyan vélekedik). Előfordulhat az is, hogy a (binárisan polarizált) véleményeket súlyozni szükséges. A folyamat végén valamilyen algoritmus szerint összegezni kell az egyéni szintű adatokat. Sok technikai kérdést lehetne vizsgálni, de nem teszem. Egyetlen formulát mutatok be (pusztán csak a lehetőségek érzékeltetése végett), a többségi eljárás képletét.

(AL.0) $$\mathbf{F}^{AP}(m,n) \equiv \mathbf{F}(\langle\Phi_i\rangle)_{i\in\mathbb{N}}(m,n) = \sum_{i=1}^{n} \delta(\Phi_i^{\pm}(m,n)) > |\mathbb{N}|/2, (\delta : \phi_i^{\pm} \mapsto \mathbb{B}, \mathbb{B} = \{1,0\})$$

ahol $\mathbf{F}^{AP}(m,n)$ **(abszolút) többségi szavazás**, $\Phi_i^{\pm}$ a polarizálófüggvény, δ az indikátorfüggvény, $|\mathbb{X}| = |\mathbb{P}| = 2$ az alternatívák száma, $|\mathbb{N}|$ a szavazók száma.

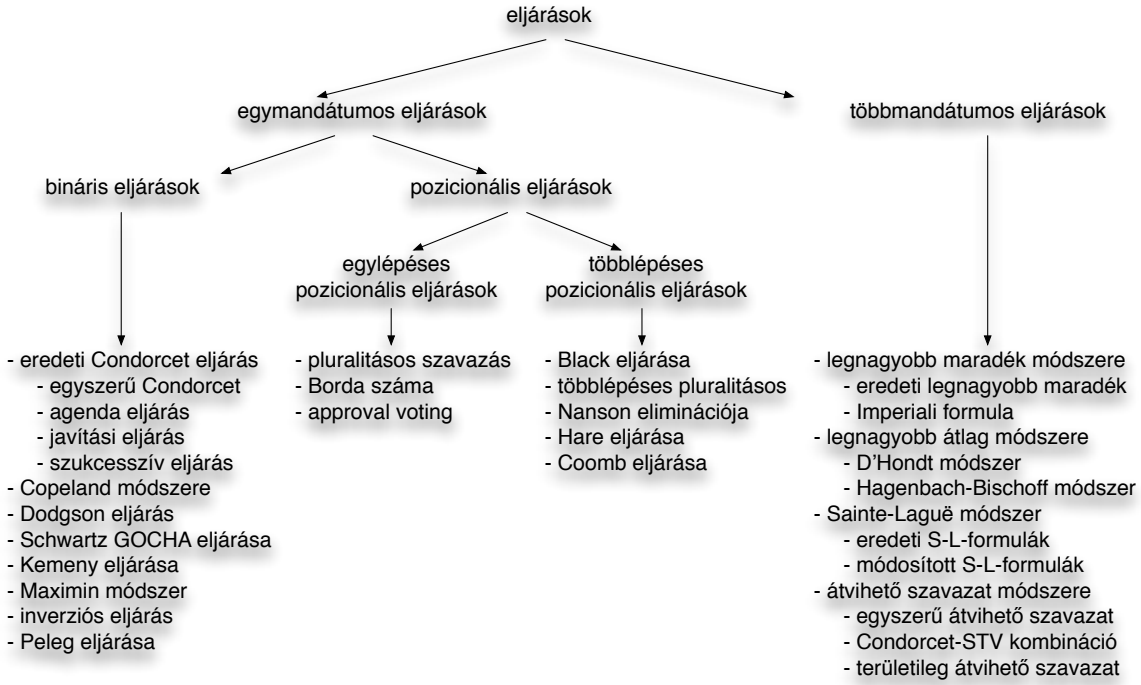
A teljes szavazási folyamatot le lehet írni pontos képletekkel, formulákkal. A választások elméletéről matematikailag nagyon jól megalapozott szakkönyvek születtek,[1] de én kihagyom ezt a kérdéskört. Ami egy cselekvéselmélet számára fontos lehet, azt megemlítem itt-ott, a részletek pedig már nem az emberről szólnak, hanem a technikáról. Ebben a könyvben ez másodlagos. Azzal sem foglalkozom, hogy milyen szavazási eljárások, szabályok, algoritmusok léteznek. Alan D. Taylor vagy húsz szavazási szabályt mutat be alaposan a könyvében [740]. Korábban Mészáros József kollégámmal együtt mi is összeállítottunk egy könyvecskét erről, ott sok részlet olvasható e tárgyban [494]. Azzal zárom ezt a fejezetet, hogy bemutatok egy ábrát a legismertebb választási eljárásokról. Nem állítom, hogy az ábra kimerítően bemutatná az összes fontos algoritmust. Vannak, akik másként csoportosítanák az eljárásokat. Az egészből csak azt tartom lényegesnek, hogy az ábra azt mutatja, nincs is olyan sok szavazási eljárás.

---

[234]
véleményaggregálás két ága:

- egy személy több, különböző értékelési szempontok szerint képzett véleményhalmaza alapján kell egyetlen aggregált döntést hozni
- több személy véleményhalmaza alapján kell egyetlen aggregált döntést hozni

[1] Hivatkozom itt egy 2005-ben megjelent könyvre: [740].



Természtesen az is lehetséges, hogy a két probléma egyszerre jelentkezik, amikor több személynek, több szempont szerint kell egyetlen döntést hoznia. Ezzel a lehetőséggel a továbbiakban nem foglalkozom. A két aggregálási feladat egységes modellben kezelhető. A különböző elméletek, modellkomponensek általában valamelyik problémakörhöz kapcsolódnak, és értelemszerűen magunkon viselik a terület terminológiáját. Ettől nem lehet teljesen elszakadni, ezért előfordulhat, hogy egyes fejezetekben egyik vagy a másik terület terminológiája, példái kerülnek túlsúlyba. Ez azonban sosem változtat az egységes modell lehetőségén, létezésén és az egységes kezelésmód iránti elköteleződésen.

———-

csoportrendezés, közösségi rendezés (social ordering), közösségi vagy kollektív preferencia (social preference)[2]

ide még írni kell

———-

Nem könnyű megmondani, mi is a különbség az **egyéni döntés** és a **közösségi döntés** között. Az azonosságot könnyebb megállapítani köztük. Mindkét esetben a döntés a lehetséges világállapotokból való választást jelenti. A döntés feltételezi, hogy van miből választani (vannak alternatívák), a döntés pedig ezt az alternativitást szünteti meg. Ha nem foglalkozunk azzal a kérdéssel, hogyan keletkezik (illetve hogyan változik) maga a döntés az emberben, akkor a legnyilvánvalóbb eltérés az egyéni és közösségi döntés között az, hogy az előbbiben egy, az utóbbi során több, mint egy ember hoz döntést.[3] Ha

[2] A 'social preference' kifejezésnek van egy teljesen más értelmezése is, amit inkább társas preferenciának fordítottam (↬292.o.).

[3] Ezen értelmezés mentén gondok adódhatnak a diktatórikus döntés jelenségének kezelésekor, de ezzel a kérdéssel itt most nem foglalkozom.

viszont több embernek kell döntést hoznia, akkor felmerül a kérdés, hogy milyen módon lehet a több egyéni döntésből egyetlen közösségi döntést előállítani, vagyis hogy lehet az egyéni döntéseket összegezni. Sokféle kifejezést használnak erre az összegzési folyamatra: **aggregálás**, kombináció, összehasonlítás (aggregation, combination, comparison). Az 'összegzés' teljesen megfelelne, de jóval elterjedtebbnek érzem az 'aggregáció' terminusát, ezért inkább ez utóbbit fogom használni. A közösségi döntés definitív mozzanata az egyéni döntések vagy **komponensdöntések** aggregálása.

Az egyéni döntésekhez képest a közösségi döntések elemzése nemcsak azért nehezebb, mert több ember döntésének vizsgálata értelemszerűen több feladatot kíván. A közösségi döntések vizsgálata teljesen újfajta problémákat hoz a felszínre. Később megpróbálom majd tisztázni, mi a különbség a vélemény- és a preferenciakinyilvánítás között. De amíg el nem jutok oda, addig is szükségesnek tartom jelezni, hogy a közösségi döntések világában el kell különíteni egymástól, hogy:

- milyen problémák, paradoxonok keletkez(het)nek a **véleményaggregálás** (judgement aggregation) folyamata során
- milyen társadalmi elvárásoknak, feltételeknek lehet (illetve nem lehet) egyszerre megfelelni a **preferenciaaggregálás** (preference aggregation) folyamata során
- milyen hatása van az **aggregálási eljárásoknak** a folyamat végeredményre
- milyen módon lehet biztosítani az egyéni megnyilatkozások őszinteségét (sincerely voting) vagy másként a csalásbiztosságot, illetve hogyan lehet kezelni (ha kell, elkerülni) a stratégiai vagy taktikai szavazást (strategic or tactical voting)

Kétféle véleményt érdemes megkülönböztetni. Vannak a preferenciakinyilvánítások, amelyek **összehasonlító ítéletek** (comparative judgements).[4] Ezek valamilyen értékelési szempont (evaluative point of view) szerint a lehetséges világállapotok mint alternatívák közül választanak ki egyet [513]. Ezt a problémahelyzetet lehet értelmezni úgy is, hogy az összehasonlítandó világállapotok mint alternatívák helyett a világállapotokat leíró propozíciók az alternatívák, így az ítélet ilyen esetekben is propozíciókból való választást jelent. A preferenciakifejezésekkel szembeállíthatóak a tényszerű ítéletek vagy **tényítéletek** (factual judgement), amelyek világra vonatkozó propozíciók közül való választást jelent. Utóbbi esetben a propozíciók igazsága vagy hamissága a kérdés, a választás az igaznak tartott állítások választását jelenti, a preferenciakinyilvánítás során pedig a jobb/rosszabb alternatívát leíró propozíciók közül választunk. Mindkét helyzetben közös viszont az, hogy az egyéni ítéleteken mellett gyakran igény merül fel közösségi vagy kollektív ítéletek (collective judgement) meghozatalára is. Ebben a fejezetben ennek problémáit tárgyaljuk.

Mit jelent az egyéni vélemény, pontosabban: egyéni kinyilvánított vélemény? Annak kinyilvánítását, hogy az adott propozícióval (alternatívával) egyetért, nem ért egyet vagy tartózkodik vele szemben (mert nem akar vagy nem tud véleményt mondani). A kollektív vagy közösségi ítélet (vélemény) a közösség tagjainak egyéni véleménye alapján előállított, aggregált származtatott vélemény. Az aggregálás folyamatát nevezhetjük szavazásnak is. Nevezzük a véleményt (ítéletet, szavazatot) nyilvánítókat szavazóknak (döntéshozóknak, megnyilatkozóknak), akik kifejezik a véleményüket (ítélkeznek, szavaznak) az alternatívák halmazával kapcsolatban.

---

felvezetni az aggregálási logika (AL) definiálását

---

[4] Az összehasonlító ítéletek az értékítéletek alá sorolhatók, de ezt a tényt itt nem használjuk ki.

(AL.1) $\mathscr{L}_{AL}(\mathscr{L}_{PL}, \mathbb{S}, \mathbb{A}, \Phi, \mathbb{C}, \{x,y,z\}, \{\alpha,\beta,\gamma\}, \mathbf{F})$, ahol

$\mathscr{L}_{PL}$	propozícionális logika
$\mathbb{S}$	világállapotok halmaza
$\mathbb{A}$	ágensek halmaza
Φ	vélemények halmaza
$\mathbb{C}$	választások halmaza
$\{x,y,z\}$	ágensek
$\{\alpha,\beta,\gamma\}$	világállapotok
$\mathbf{F}$	választási függvény

Az új nyelvünkben több új predikátumot fel kell vennünk. Vezessük be a következő jelöléseket:

(AL.2) $\mathbb{N} = \{1, \ldots, i, \ldots, |N|\}$

ahol N a szavazók halmaza, i az i. szavazó, $|\mathbb{N}|$ a szavazók száma.[5]

(AL.3) $\mathbb{P} = \{m, n, \ldots, o\}, |\mathbb{P}| = p$

ahol m, n, p lehetséges világállapotok, amelyekről a szavazók kinyilváníthatják a véleményüket, P a világállapotok halmaza, p a lehetséges alternatívák száma; az ítélek lehetnek propozíciók (a propozícionális logika jól-formált formulái), illetve preferenciarendezések.

(AL.4) ϕ_k

ahol ϕ_k egy lehetséges vélemény(alternatíva), amelyet a szavazók választhatnak, amikor kinyilvánítják a véleményüket.

(AL.5) $\mathbb{X} = \{\phi_1, \phi_2, \ldots, \phi_k\}$

ahol ϕ_k egy lehetséges véleményalternatíva, X a véleményalternatívák halmaza,

(AL.6) $\Phi_i = \{\phi_k, \ldots, \phi_m\} \subseteq \mathbb{X}$

ahol Φ_i az i. szavazó véleményprofilja (az alternatívákra vonatkozó véleményeinek halmaza).

(AL.7) $\Phi = \langle \Phi_1, \ldots, \Phi_i, \ldots, \Phi_n \rangle = \langle \Phi_i \rangle_{i \in \mathbb{N}}$

ahol Φ az egyes szavazók Φ_i véleményprofiljainak *rendezett n-ese* (ha nem fontos a rendezettség, akkor elegendő a halmazjel: $\{\Phi_i\}_{i \in N}$).

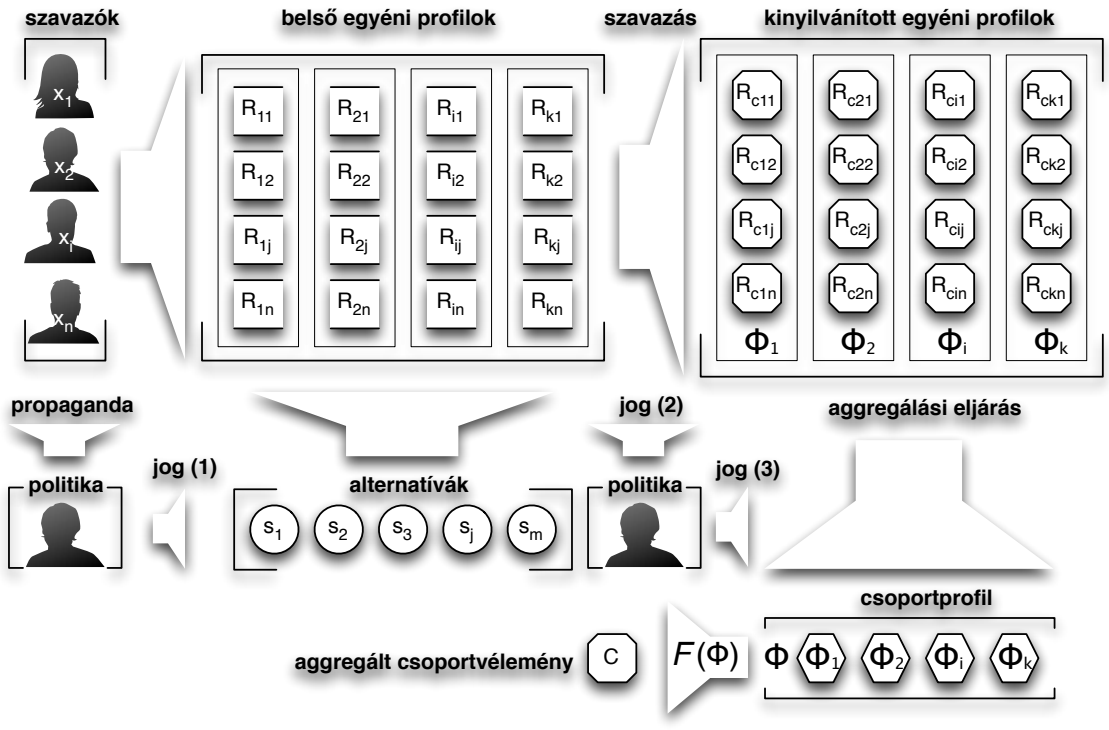
(AL.8) $\mathbf{F}(\langle \Phi_i \rangle_{i \in \mathbb{N}}) \equiv \mathbf{F}(\Phi) : \mathbb{D}_F \mapsto \mathbb{X}, \mathbb{D}_F \subseteq \mathbb{X}^n$

közösségi vélemény függvény (HUN) – csoportdöntési fügvény (HUN) – group decision function (ENG) – társadalmi jólét függvény (HUN) – social welfare function (ENG) – preferenciaaggregálási függvény (HUN) – preference aggregation function (ENG) – constitution (ENG) – közösségi preferencia (HUN) – social preference (ENG) – kollektív választás(i szabály) (HUN) – collective choice (rule) (ENG) – kollektív döntés(i szabály) (HUN) – collective decision (rule) (ENG) – allokációs mechanizmus (HUN) – allocation mechanism (ENG) – (társadalmi) választási függvény (HUN) – (social) choice function (ENG) – társadalmi döntési módszer/függvény (HUN) – social

[5] Az itt bevezetett új halmazt, az ágensek halmazát a természetes számokkal jelzem, és a későbbiekben még az is előfordulhat, hogy az ágensekre a természetes csak a természetes számokkal mint azonosítókkal fogok utalni.

decision method/function (ENG) – szavazási rendszer (HUN) – voting system (ENG) – voting scheme (ENG) – voting mechanism (ENG) – social choice procedure (ENG)

Érdemes egy ábrában is feltüntetni a most jellemezni kívánt aggregálási folyamat legfontosabb összetevőit, az ezek közti kapcsolatokat. A helyzet legfontosabb kapcsolata az egyének és a világállapotok közti kapcsolat, amelyről itt semmit sem tudunk mondani. Tényként vesszük itt tudomásul, hogy egy adott csoport tagjai kifejezik a véleményüket a világállapotok valamely halmazáról. Attól függően, hogy melyik véleményaggregálási helyzetben vannak, más lesz az egyéni vélemények típusa. Az egyéneknek véleményük lehet a világállapotok kívánatosságáról, amit preferenciarendezés formájában fejezhetnek ki, de véleményük lehet a világállapotokra vonatkozó faktuális kijelentés is. Ezek az egyéni vélemények – alapértelmezés szerint – még őszinték, viszont nem nyilvánosak abban az értelemben, hogy az ágensek nem nyilvánították ki őket valamilyen szavazási aktus keretében. Ha ez a szavazás megtörténik, akkor beszélhetünk az ágensek kinyilvánított véleményeiről, és ezeket az ágensekhez rendelve megkapjuk a külső egyéni profilokat. A szavazás aktusában lehet megragadni a stratégiai szavazás jelenségét. A külső egyéni profilok alapján lehet létrehozni a csoportprofilt, majd ez alapján a közösségi döntés végeredményét. Az aggregálási eljárás – szűkebb értelemben vett – fogalma ezt a véleményössszesítési folyamatot ragadja meg.



Ez az ábra az aggregálási folyamat legáltalánosabb értelmezését jelenti. Magába foglalja az aggregálási feladatok mindkét ágát, a vélemény- és a preferenciaaggregálás jelenségét. A következő két nagyobb fejezetben bemutatom mindkét területet, előbb a preferenciák, majd a vélemények aggregálásának miként jét.

4.1.1 Preferenciaprofilok tipológiája

Ha az egyéni preferenciák helyett társas vagy közösségi preferenciákat vizsgálunk, akkor az egyéni preferenciák valamely összességét preferenciaprofilnak hívjuk. Ezek a preferenciaprofilok rendelkezhetnek olyan tulajdonságokkal, amelyekre támaszkodva fontos és érdekes tételek vezethetők le. Ebben a fejezetben ezeket a speciális mintázattal rendelkező profilokat mutatom be, de itt nem foglalkozom a belőlük levezethető tételekkel. Amikor preferenciaprofilokat jellemzünk, akkor nemcsak az összahasonlítandó alternatívákra vonatkoztatjuk az állításainkat, de a minősítésbe bevonjuk a preferenciarendezéseket végrehajtó személyeket is. Ezt persze interpretálhatjuk úgyis, hogy ilyenkor magukra az egyéni preferenciarendezésekre mondunk ki állításokat, ami logikai szempontból azt jelenti, hogy másodrendű logikai konstrukciókat kell kezelnünk (↬60.o.), hiszen relációk (predikátumok) tulajdonságaira, kapcsolataira kell a kvantorainkat vonatkoztatni.

Az első elemzett preferenciaprofil az egycsúcsú preferencia volt, amivel kapcsolatban fontos, ún. lehetségességi tételeket mondtak ki a társadalmi választások elméletén belül (↬215.o.). Ken-ichu Inada volt az, aki először tett kísérletet egy – bizonyos szemponból – kimerítőnek minősíthető preferenciaprofil-tipológia kidolgozására [369], [370]. Inada 1969-es cikkében véglegesített tipológiáját befolyásolta Amartya K. Sen javaslata, amely eredményeként egy általánosabb kategória alá be lehetett sorolni Inada legfontosabb típusait, az ún. értékkorlátozott preferenciákat [659], [669]. Inada profiltipológiáját később – egy új szempont figyelembe vételével – egy további típussal kibővítették, viszont az inadai típusok közül többeket elhagytak idővel, és inkább csak a Sen által is fókuszált értékkorlátozott profilokkal foglalkoztak alaposabban. Az egyes preferenciaprofilok ismertetésekor kezdetben igazodom Inada logikájához, de később eltérek az általa követett gondolatmenettől, és a legfontosabb, a mai napig elemzett típusok bemutatásakor már a későbbi szerzők megközelítését veszem át.

Inada preferenciaprofiljai

Ken-ichu Inada először három olyan preferenciaprofilt mutatott be, amelyek tetszőleges számú szavazó esetén tranzitív közösségi preferenciát képesek adni az egyszerű többségi döntési eljárással. A preferenciaprofil-típusok leírásához Inada abból a kérdésből indult ki, hogy elméletileg milyen egyéni preferenciarendezések lehetségesek. Ha adva van három alternatívánk (m, n, o), akkor azokat tizenháromféleképpen lehet preferenciasorba állítani. Hatféleképpen rangsorolhatjuk őket úgy, hogy szigorú rendezést feltételezünk mindhárom alternatíva között, három-háromféle sorrend lehetséges akkor, ha egy alternatívát jobbnak, illetve rosszabbnak tartunk a másik kettő, egymással közömbös alternatívához képest, valamint egyféleképpen fordulhat elő az a helyzet, amikor mindhárom alternatíva közömbös egymással szemben. A szemléletesebb megjelenítés miatt kivételesen eltérve az eddigi gyakorlattól, a preferenciarelációk argumentumait infix-jelölésmódban kezelve, az alábbi táblázatban mutatom be a fent kifejtett lehetőségeket (csak a példa kedvéért jelzem, hogy az infix jelölésmód alkalmazásával a következő két formula ekvivalens egymással: $\mathbf{P}(m,n) \wedge \mathbf{P}(n,o) \equiv m\mathbf{P}n\mathbf{P}o$.
Mivel az indifferencia reláció szimmetrikus, így az argumantumai felcserélhetőek egymással, a második és harmadik oszlopban szereplő konfigurációkban az indifferenciareláció két oldalán levő argumentumok szerepelhetnének fordítottt sorrendben is. A táblázat az egyéni rendezések lehetséges mintázatait mutatja. A szavazás során a szavazók ezen ren-

$m\mathbf{P}n\mathbf{P}o$	$m\mathbf{P}n\mathbf{I}o$	$m\mathbf{I}n\mathbf{P}o$	$m\mathbf{I}n\mathbf{I}o$
$m\mathbf{P}o\mathbf{P}n$	$n\mathbf{P}o\mathbf{I}m$	$n\mathbf{I}o\mathbf{P}m$	
$n\mathbf{P}m\mathbf{P}o$	$o\mathbf{P}m\mathbf{I}n$	$o\mathbf{I}m\mathbf{P}n$	
$n\mathbf{P}o\mathbf{P}m$			
$o\mathbf{P}m\mathbf{P}n$			
$o\mathbf{P}n\mathbf{P}m$			

dezésminták bármelyikét választhatják, így előfordulhat az az eset, amikor mind a tizenhárom különböző rendezésminta szerint szavazott legalább egy szavazó. Amikor tehát az egyéni preferenciák aggregálását kell modelleznünk, akkor azt kell feltételeznünk, hogy a fenti tizenhárom mintázat mindegyike szerint szavaznak valamennyien, és ezek alapján kell az összesítést elvégeznünk. A kérdés az, hogy vannak-e olyan konstellációk a rendezésminták összessége között, amelyekre képesek lehetnek a többségi szavazat konzisztenciáját, tranzitivitását boztosítani. Ken-ichi Inada ezen lehetőségeket vizsgálva különített el különböző típusú preferenciákat [370]. Az inadai preferenciaprofil-típusok mind olyanok, amelyek bizonyos feltételek mentén biztosítani képesek a többségi eljárás konzisztenciáját. Fontos hangsúlyozni, hogy itt nem egyéni preferenciákról van szó, hanem egyéni preferenciatípusok adott konfigurációjáról, vagyis preferenciaprofilokról.

Amennyiben feltételezzük, hogy a szavazói populáció olyan, hogy az egyéni preferenciák között nem fordulnak elő szigorú rendezések (vagyis a fenti táblázat első oszlopában látható rendezésminták szerint senki sem szavaz), akkor Inada szerint **dichotóm preferenciáról** beszélhetünk. Ilyenkor a szavazók mindegyikére igaz, hogy az alternatívákat két csoportba osztva rendezik el úgy, hogy az egyes csoportokon belüli alternatívák között közömbösek. Az alternatívák így két csoportra oszlanak el és a szavazók így valóban dichotóm módon választ(hat)anak egyik vagy másik csoportból maguknak. Mielőtt megmutatnám, hogy milyen formula segítségével lehet leírni a dichotóm preferenciaprofil fogalmát, szükséges egy megjegyzést tenni Inada jelölési megoldására vonatkozóan. A következő formulákban át fogok térni az alternatívákat jelző m, n, o változók helyett az x, y, z változók használatára, hogy egyértelműen jelezhessem, a szavazási minták eltérhetnek egymástól a konkrét értékeiket tekintve, miközben a mintázatukat tekintve azonos rendezési logikát mutatnak. Az $m\mathbf{P}n\mathbf{P}o$ formula így a fenti táblázat első oszlopának mind a hat elemét reprezentálhatja. Ezek alapján a dichotóm preferencia formuláját a következőképpen írhatjuk fel.

(AL.9) $\quad (\text{DIP}) \stackrel{\text{def}}{=} \forall x \forall y \forall i (\neg (\mathbf{P}_i(x,y) \wedge \mathbf{P}_i(y,z)))$

dichotóm (preferencia) (HUN) – dichotomous preference (ENG)

Inada kimondott egy tételt az ilyen preferenciaprofil esetére vonatkozóan, de ezt később mutatom majd be (↬216.o.). A következő preferenciaprofilokat már úgy kapjuk meg az inada építkezés során, hogy feltételezzük azt, hogy előfordulhatnak szigorú egyéni preferenciarendezések (vagyis a preferenciaminták között szerepelnek a fenti táblázat első oszlopából). Ez a feltétel önmagában azonban még nem szolgáltat érdekes mintázatot számunkra, tehát további feltételeket kell megfogalmaznunk. Az egyik ilyen pluszfeltétel lehet, hogy kikötjük azt, hogy amikor fennáll egy adott szigorú hármas preferenciarendezés, akkor kizárjuk a rendezés két végén álló alternatíva közti fordított irányú szigorú rendezés létezését. A fenti táblázat alapján ez azt jelenti, hogy az első oszlop mintái közül

megengedjük azokat, amelyek az egyik alternatíva (x) szigorúan jobb, mint a másik (z), de fordítva nem áll fent a kapcsolat köztük, vagyis az első három sorban levő típusok létezését engedjük csak meg. Inada ezt a típust nevezte el **visszhangzó preferenciának**. A kifejezéssel arra akart utalni, hogy ha van egy szavazó, aki tranzitív preferenciákkal rendelkezik és a többiek is követik (visszhangozzák) őt ebben, akkor nyilvánvalóan teljesen a közösségi szavazat tranzitivitása is. Ezt a triviális esetet lehet általánosítani a most megfogalmazott feltétellel, amit a következő módon formalizálhatunk.

(AL.10) $\text{(ECP)} \stackrel{\text{def}}{=} \forall x \forall y \forall z \forall i(\mathbf{P}(x,y) \wedge \mathbf{P}_i(y,z) \rightarrow \neg \mathbf{P}_i(z,x))$

visszhangzó (preferencia) (HUN) – echoic preference (ENG)

A következő preferenciaprofil felvezetéséhez Inada újracsak egy triviális példát hozott értelmezési mintául. Ha a szavazó közösség két diszjunkt csoportra osztható a szavazók "ízlése" szerint, azaz amíg az egyik csoport tagjai $x\mathbf{P}y\mathbf{P}z$ sorrend szerint szavaznak, addig a másik csoportba tartozók a $z\mathbf{P}y\mathbf{P}x$ mintát követik, akkor a többségi eljárás konzisztens közösségi végeredményt képes kihozni. Ezt a triviális esetet úgy lehet általánosítani, hogy amennyiben vannak, aki az $x\mathbf{P}y\mathbf{R}z$ sorrend szerint preferálják az alternatívákat, akkor vagy vannak olyanok is, akik a fordított rendezési sort preferálják ($z\mathbf{P}y\mathbf{R}x$), vagy vannak, akik a két szélső értéket tekintve közömbösek ($x\mathbf{I}z$). Ha ez így van, akkor – Inada javaslata alapján – **antagonisztikus preferenciáról** beszélhetünk, ami a következő formulával fejezhetünk ki.

(AL.11) $\text{(ANP)} \stackrel{\text{def}}{=} \forall x \forall y \forall z \forall i((\mathbf{P}_i(x,y) \wedge \mathbf{P}_i(y,z)) \rightarrow (\mathbf{P}_i(z,y) \wedge \mathbf{P}_i(y,x)) \vee \mathbf{I}_i(x,z))$

antagonisztikus (preferencia) (HUN) – antagonistic preference (ENG)

A dichotóm, a visszhangzó és az antagonisztikus preferenciaprofilok fogalmával szemben Inada úgy vezette be a **tabu preferencia** fogalmát, hogy jelezte, hogy utóbbi létezése esetén csak páratlan számú szavazó esetén remélhetünk konzisztens végeredményt a többségi eljárás alkalmazása során. A tabu preferencia lényege az, hogy egyfajta "preferenciaminta" nem fordulhat elő. Ezt azzal a kettős feltétellel vélte megragadhatónak Inada, hogy kizárta az alternatívák közti teljes közömbösség, valamint a kizárandó preferenciarendezés lehetőségét az egyéni preferenciarendezések köréből. Ezt az alábbi módon írhatjuk le.

(AL.12) $\text{(TAP)} \stackrel{\text{def}}{=} \forall x \forall y \forall z \forall i(\neg(\mathbf{I}_i(x,y) \wedge \mathbf{I}_i(y,z)) \wedge \neg \mathbf{P}_i(y,x))$

tabu (preferencia) (HUN) – taboo preference (ENG)

Az első három bemutatott inadai preferenciaprofil tetszőleges számú, a negyedjére elemzett mintázat csak páratlan számú szavazó esetén képes konzisztens végeredményt produkálni. Az első három, illetve a negyedik típus között elemezte Inada azokat a preferenciamintákat, amelykre későbbb Sen az értékkorlátozott minősítést alkalmazta. Mivel az idők során ezek a preferenciarendezések kerültek igazán az elemzői figyelem fókuszába, ezeket önállóan, a következő fejezetben mutatom be.

Értékkorlátozott preferenciaprofilok

A legelőször vizsgálatba vont és azóta is a legtöbbet hivatkozott preferenciaprofil-típus az egycsúcsú preferencia. A fogalmat Duncan Black vezette be 1948-ban [91], Arrow

is hivatkozott rá 1951-es könyvében [33], majd utánuk még sokan mások is használták a kategóriát. Ha a szigorúan rendezett alternatívák halmazában van olyan elem, amely minden más elemhez képest jobbnak minősül, és a többi alternatíva értékelése (rendezése) speciális mintázat szerint, montonon módon csökkenve változik a maximális elemtől két irányban távolodva adott skálán mérve, akkor beszélünk **egycsúcsú preferenciáról**. Ekkor a lineárisan rendezett alternatívákra vonatkozóan a szavazók preferenciái vagy mindig szigorúan növekvő vagy szigorúan csökkenő vagy egy pontig először szigorúan növekvő, majd a ponttól szigorúan csökkenő függvényképet mutatnak. Az egycsúcsú preferenciát az alábbi képlettel írhatjuk le.

A fogalom értelmezéséhez segítséget jelenthet, ha bemutatok néhány példát az egycsúcsú preferenciákkal kapcsolatban. Az egyik példa hétköznapi. Ha egy nagy teremben több ágy van sorban egymás mellé rakva (mint egy laktanya körletében), és csak egy melegítőtest van, akkor minden egyes ágy birtokosának egycsúcsú a preferenciája, vagyis mindenki meg tudja mondani, hogy hol lenne számára a hőforrás ideális helye. Ez a hely természetesen mindenki esetében más és más pont lehet a teremben, ami részben a terem hossztengelye mentén elfoglalt pozíciótól, részben az ágytulajdonos "hőigényétől", fázósságától függ. Akik jobban szeretik a meleget, azok közelebb szeretnének a kályhához lenni, de lehet egy olyan pont, aminél közelebb már nem lenne annyira jó nekeik, mert túlzottan sok meleg jönne rájuk. Ilyen helyzetben mindenki meg tudja valahogyan határozni azt, hogy neki mi lenne a hőforrás ideális elhelyezése, és az mindenki számára egycsúcsú preferencia elrendezést jelentene a lehetséges helyek halmazán.

Gyakran hivatkoznak az egycsúcsú preferenciákra a közgazdaságtanban, illetve a politikai szociológia világában. Az érdeklődést Harold Hotelling híres cikke váltotta ki először [365], amiben a szerző egy tengerparti sétányon vagy egy kisváros főutcáján működő boltok elhelyezkedéséről állított fel egy meglepő és igencsak "erős" tézist. Hotelling modellezte az a jólismert tényt, hogy amikor két konkurrens boltos helyet keres magának a város főutcáján (vagy az üdülőövezet sétányán), akkor mindketten a főutca (sétány) közepén fogják megnyitni a boltjukat. A magyarázat egyszerű. Ha feltételezzük, hogy a verseny miatt a boltosok minden minőségükben (árban, kínálatban, kiszolgálásban stb.) megegyeznek, akkor a potenciális vevőknek csak az marad meg egyetlen döntési szempontként, hogy melyik bolt van közelebb hozzájuk (a lakásukhoz, nyaralójukhoz). A Hotelling tétel szerint a boltosok akkor remélhetik a legnagyobb forgalmat, ha a lehető legtöbb vevőnek a lehető legkisebb sétálnivalója lesz, ezt pedig akkor érhetik el, ha valamennyien az utca, a sétány közepére települnek. Az tétel azért érdekes számunkra, mert a magyarázat erősen támaszkodik az egycsúcsú preferencia létezésére (illetve annak tulajdonságaira). A szavazói preferenciák egycsúcsúságával magyarázzák a politikai pártok medián szavazókra optimalizáló viselkedését, és hasonló okok miatt kap fontos szerepet ez a fogalom a társadalmi választások elméletén belül a lehetségességi tételek között (↬216.o.).

Az egycsúcsú preferenciák jellegzetessége az, hogy van egy maxinális elem és a többi alternatíva ehhez képest egyre rosszabbnak minősül, bármelyik irányban is távolodunk a maximális elemtől. Ennek a speciális mintázatnak elképzelhetjük egyfajta tükörképét is, amikor van egy minimális elem, és minden más elem ehhez képes egyre jobbra értékelt, minél jobban távolodunk a minimális elemtől mindkét irányban. Az ilyen preferenciát Ken-ichi Inada **egyvölgyű preferenciának** nevezte el [369].

Az egyvölgyű preferencia fogalmával kifejezhetjük például a "közrosszhoz" való vi-

szonyunkat. Amikor olyan döntést kell hozni, amit nem nagyon támogatnak az érintettek (hol épüljön börtön, menekülttábor, hulladéktároló telep, szemétfeldolgozó üzem stb.), akkor a preferenciáinkat nem a jobb, hanem inkább a rosszabb alternatívák jelzésével szerint fejezzük ki. William Vickrey már Inada cikke előtt is utalt erre a preferenciatípusra, bár ő egymélypontú (single troughed) preferenciának nevezte el [788]. Úgy tűnik, hogy az inadai 'single-caved' kifejezés helyett végül inkább a 'single-dipped' nyelvi szerkezet terjedt el igazán a szakirodalomban [399], [400], [51].

Inada egy harmadik profiltípust is definiált, amit ugyanazzal a módszerrel határozott meg, mint amivel az egyvölgyű preferencia fogalmát is "elválasztotta" az egycsúcsú preferencia fogalmától. Ekkor azonban a legnagyobb (legjobb) elem helyett nem a legkisebb (legrosszabb) elem cseréjét javasolta, hanem ehelyett inkább a hármas összehasonlítás harmadik elemére fókuszált az új fogalom definiálásakor. A **szeparáló preferenciának** elnevezett új típus lényege az, hogy az alternatívák közül a szavazók mindig a két szélsőértéket hasonlítják össze vagyis minden esetben figyelmen kívül hagyják a "középső értéket".

Inada preferenciafogalmainak vagyis az egycsúcsú, az egyvölgyű és a szeparált preferencia általánosításaként Amartya Sen javasolta használatba venni az **értékkorlátozott preferencia** (value-restricted preference) fogalmát [659], [669]. Sen abból indult ki, amit már Inada is észrevett (sőt, a tételeiben ki is használt), hogy Arrow a lehetségességi tételeinek bizonyításakor (↬216.o.) valójában csak hármas összehasonlításokkal operált, vagyis nem használta ki azt, hogy a tétel feltételei az összes alternatívára "egyszerre" vonatkozzanak. Sen az inadai három preferenciatípust arra a közös vonásra hivatkozva vonta egy felettes kategória alá, hogy azok az alternatívák hármas összehasonlításakor mindig kizárják valamelyik elemet. Három elem (m,n,o) viszonylatában ki lehet hagyni a legjobb, a legrosszabb és a középső elemet, és ennek megfelelően beszélhetünk **felső-kizáró** (best-restricted), **alsó-kizáró** (worst-restricted) és **közép-kizáró** (medium-restricted) preferenciáról.

Az egycsúcsú preferenciák Sen jellemzésében az alsó-kizáró típusba tartoznak, hiszen ebben az esetben nincsen legrosszabb elem egyik szavazó összehasonlításában sem. Az egyvölgyű preferenciáknak a felső-kizáró preferenciaprofil felel meg, hiszen senki preferenciarendezésében sem szerepel a legjobb elem. A szeparáló preferencia esetében pedig közép-kizáróságról beszélhetünk, amikoris a középső alternatívát hagyja ki az összes szavazó a rendezéséből. Mindezek alapján a három preferenciaprofilt a következőképpen definálhatjuk.

(AL.13) $(\text{SPP}) \stackrel{\text{def}}{=} \forall m \forall n \exists! o(((m < n \wedge n \leq o) \rightarrow \mathbf{P}(n,m)) \wedge ((m > n \wedge n \geq o) \rightarrow \mathbf{P}(n,m)))$

egycsúcsú (preferencia) (HUN) – single-peaked (preference) (ENG)

(AL.14) $(\text{SDP}) \equiv \succ^s \stackrel{\text{def}}{=} \forall m \forall n \exists! o(((m < n \wedge n \leq o) \rightarrow n \succ^s m) \wedge (m > n \wedge n \geq o) \rightarrow n \succ^s m))$meg kell fordítani még

egyvölgyű (preferencia) (HUN) – egyvermű (preferencia) (HUN) – single-caved (preference) (ENG) – single-dipped (preference) (ENG) – single troughed (preference) (ENG)

(AL.15) $(\text{GSP}) \stackrel{\text{def}}{=}$

(szeparáló preferencia) (HUN) – (two-)group-separated preference (ENG) – group-separable preference (ENG)

Ballester és Haeringer tétele értelmében [46]

Miguel A. Ballester és Guillaume Haeringer A characterization of the single-peaked domain

Sen jellemzésében: medium-restricted Every group-separable profile is also medium-restricted.

Ballester és Haeringer tétele értelmében [46]

(AL.16) $(\text{SSP}) = \text{WORST-RESTRICTED} \wedge \neg\alpha - config$

(AL.17) $(\text{SDP}) = \text{BEST-RESTRICTED} \wedge \neg\overline{\alpha} - config$

(AL.18) $(\text{GSP}) = \text{MEDIUM-RESTRICTED} \wedge \neg\beta - config?????$

Az itt bemutatott preferenciatípusokra vonatkozó matematikai definíciókat Robert Bredereck, Jiehua Chen és Gerhard J. Woeginger cikke alapján adom meg [119].

Single-crossing preference

egyszer kereszteződő preferencia előjelet egyszer váltó , egyszeres előjelváltó

Kevin W.S. Roberts vezette be a **hierarchikus tapadás feltétel** (hierarchical adherence condition) fogalmát [607]

jellemzése: [120]

Joshua S. Gans és Michael Smart a többségi szavazás és a .... kapcsolatát elemezte [280]

(PL.19) $(\text{SCP}) \equiv \succ^{s} \stackrel{\text{def}}{=}$

??? (preferencia) (HUN) – single-crossing preference (ENG)

Single-crossing, strategic voting and the median choice rule [636]

Általánosítás

Indifferences and Domain Restrictions [49]

Salvador Barberà és Bernardo Moreno a topmonotonitás fogalmát bevezetve az egycsúcsú és a Single-crossing preferencia [53]

4.2 Preferenciaaggregálás

A preferenciáról szóló fejezetben bemutattam, hogy e jelenségről szólva mindig négyféle preferencia-vonatkozású relációt kell figyelembe vennünk (↬86.o.). A társadalmi választások elméletében - csakúgy, mint más preferencia-vonatkozűsú elméletekben – elsősorban a gyenge preferenciarendezés (↬92.o.) és az erős preferenciarendezés (↬91.o.) fogalmait használják, de olykor-olykor felbukkan a preferenciarelációk körébe sorolható közömbösség (↬92.o.), illetve az összehasonlíthatatlanság (↬92.o.) relációja is.

Mielőtt a preferenciaggregálás sajátosságainak tárgyalásába kezdenék, röviden felvázolok egy olyan értelmezési keretet, amely végig segítségemre lehet a jelenségek bemutatásában azáltal, hogy összeköti az egyes elemzendő kérdésköröket.

4.2.1 Manipuláció

ügyrend (eljárás) napirend (alternatívák) agenda szavazás alternatívák halmazának változtatása (1) új dimenzió, szempont felvetése és ezáltal új alternatíva lehasadása valamelyik régiről, 2) vagy teljesen új alternatíva megjelenése 3) az alternatívák felszeletelése részalternatívákra és ezek szerint szavazás - ez egyfajta kognitív gerrymandering) [6]

a stratégiai szavazásról, a szavazói tanulás jelenségéről [502]

A társadalmi választások elméletén belül a szavazás jelenségének talán legfontosabb vonása a manipuláció lehetősége. Az egyéni preferenciák aggregálása egyetlen közösségi véleménnyé nem kizárólag az egyéni véleményektől függ, hanem sok minden mástól, és ezeket mind valamilyen szinten meg lehet változtatni, amivel a végeredményt is lehet befolyásolni. A manipulációnak azonban sok értelme, sok formája, sokféle "színtere" lehetséges. A társadalmi választások elméletében legelterjedtebb manipulációértelmezés az egyéni preferenciák kinyilvánításával, a leadott szavazatok őszinteségével kapcsolatos (↬224.o.). A stratégiai szavazásnak nevezett jelenség a szavazók által direkt módon megvalósított szavazatmanipulációt jelenti, amikor a saját egyéni preferenciarendezésüket megváltoztatják a szavazatgazdák, vagyis a szavazók. Ezzel szemben áll az az indirekt szavazatmanipuláció, amikor ugyanúgy az egyéni preferenciarendezések változnak meg, de itt nem a szavazók maguk változtatják meg tudatosan saját véleményüket, hanem mások hatására, a külső véleménybefolyásolások eredményeként lesz más a vélemény. Ekkor a szavazók őszintén szavaznak, szemben a stratégiai szavazás őszintétlenségével. A külsődleges véleménybefolyásolás hatására bekövetkező vélményváltoztatást azonban, bár fontossága nem kérdőjelezhető meg, itt mégsem vizsgálom, mivel ez más tudományterület elemzési körébe tartozik.

Másfajta manipulációról beszélhetünk akkor, amikor az alternatívák halmaza változik meg valamiért. Az nagyon fontos kérdés, hogy miért is változ(hat)nak meg az értékelendő alternatívák. A dolgokat itt is leegyszerűsítve azt mondhatjuk, hogy szavazási kontextusban az alternatívák maguktól, a szavazók felől tekintve nem változnak meg. Ha változik valami e téren, akkor mindez azoknak a hatására történik, akik felelnek a szavazásért, akik kontroll alatt tartják a választási folyamatot. Ha ezeket elnevezzük aggregációgazdáknak, akkor azt mondhatjuk, hogy az **alternatívamanipuláció**, amikor az alternatívák, a jelöltek listáját megváltoztatják az az aggregációgazdák, kívül esik az aggregációs folyamatok logikáján. Ez a kérdés a választás normatív kontextusában elemezhető, ezért ennek tárgyalását is elhagyom itt. Fontos kérdés, de nem az aggregálási folyamat jelenségébe tartozik.

Hasonlókat mondhatunk a **szavazójog-manipuláció** kérdéséről is. A szavazások, választások végeredményét befolyásolja az a tény, hogy kik a szavazók, kiknek a véleménye számít. Ez is fontos történelmi, társadalmi kérdés, gondoljunk csak a nők választójogának megjelenésére, a cenzus alapú választójog eltűnésére, a 'ki a szabad', 'ki a felnőtt' kérdésekre adott válaszok történetére, a szavazati jog megvonása mint egyéni vagy kollektív büntetés jelenségére, vagy – még konkrétabban – a Magyarország határain kívül élő magyarok, magyar állampolgárok szavazati jogának kezelésére. Ebben a jelenségben is van manipulációs potenciál, nyilván ez is a választáshoz tartozik, de ezt sem elemzem itt. Ugyanazokért nem, mint ahogy a korábbi témaköröket is elhagytam: más tu-

[6] A manipuláció bizonyos lehetőségeiről, területeiről magyar nyelven érdekes áttekintést ad Meszerics Tamács tanulmánya [503].

dományterület relevanciájába tartozik.

Van azonban két olyan manipulációs lehetőség a választásokkal kapcsolatban, amit végre már "elemezhetek". Az egyik az **algoritmusmanipuláció**,[7] amikor az aggregálási eljárást megváltoztatják az aggregációgazdák, hogy ezzel befolyásolják a végeredményt, a másik a **körzethatár-manipuláció**, amikor a körzethatárokat változatják meg politikai, gazdasági előnyök reményében. Az első kérdéskört alaposabban körbe járom (↬204.o.), a második problémát csak röviden említem (↬243.o.).

4.2.2 Algoritmusmanipuláció

Azt, hogy a szavazási eljárás választása önmagában befolyásolhatja a csoportszavazás végeredményét, már legalább a rómaiak óta tudjuk. Nem lehet nem megidézni azt a mesét, amiről ifjabb Plinius egyik levelében számolt be [576, 8. könyv, 14. levél], és ami a társadalmi választások elméletének talán leggyakrabban idézett példájává vált [240], [606], [552]. A történetet Plinius eredeti meséjéhez képest kicsit megváltoztatva mutatom be, hogy szemléletesebb, karakteresebb legyen a tanulsága.

Ifjabb Plinius beszámolója szerint Afranius Dexter szenátor holtan találják a villájában, és a gyanú a felszabadított rabszolgáira terelődik. Tegyük fel, hogy a szenátorok egy 12 tagú csoportjának kell meghoznia kell a végső döntést, akik három alternatíva közül választhatnak:

F	felmentés
K	kivégzés
S	száműzés

Tegyük fel még azt is, hogy a szenátorok három csoportokba sorolhatók. Az 1. csoport tagjai hisznek a vádlott ártatlanságában, tehát szeretnék felmenteni, vagyis az F alternatívát teszik saját preferenciasoruk elejére. Továbbá úgy gondolkodnak még, hogy ha valamilyen oknál fogva nem lehet felmenteni a vádlottat, akkor a száműzés (S alternatíva) mellett voksolnának, s utolsó helyre tennék a kivégzés lehetőségét (a K alternatívát). A referenciarendezésük tehát az alábbi formában írható le röviden: FSK. A második csoport rangsora legyen KSF (vagyis: kivégzés-börtön-felmentés), míg a harmadik csoportot jellemezzük a SFK száműzés-felmentés-kivégzés) rendezéssel (e csoport tagjai olyan érvekre támaszkodhatnának, hogy bár bűnösnek tartják a vádlottat, az első helyen ezért szerepel a börtönbüntetés kiszabása, de a halálbüntetést ellenezve második helyre inkább a felmentési alternatívát tennék). Tegyük fel, hogy a három csoport létszáma a következő:

1. csoport	FSK	5
2. csoport	KSF	4
3. csoport	SFK	3

[7] Direkt módon nem így nevezi meg, de tartalmilag mint napirend-manipulációról (agenda manipulation) beszél erről a jelenségről Peter C. Ordeshook. Az eljárási algoritmusokat a szavazások lebonyolításának ügyrendjeként felfogva lehetne érvelni a napirend, de még inkább az ügyrend fogalmának használata mellett, de a politikai elemzések világában később teret nyert napirendelemzési technikák miatti terminológiai ütközés miatt én nem használom Ordeshook kifejezését.

A szavazási összesítést kétféle módon is elvégezzük, nézzük meg, hogyan. Vegyük előbb a plurális eljárást!

Plurális eljárás. Vizsgáljuk meg, melyek alternatívára hányan szavaznak külön-külön, és a legtöbb szavazatot kapott alternatíva legyen a nyertes. Ebben az esetben az eredmény a következő:

F	5 szavazat	az első csoport mindegyik tagja ide szavaz
B	3 szavazat	a harmadik csoport mindegyik tagja ide szavaz
K	4 szavazat	a második csoport mindegyik tagja ide szavaz

Tehát a nyertes az F alternatíva, vagyis ebben az esetben a vádlottat felmentenék. Most nézzük meg a másik, az ún. bináris eljárás eredményét!

Bináris eljárás. Az eljárás két lépésből áll. *(i)* Először vizsgáljuk meg, a szenátorok bűnösnek tartják-e a vádlottat vagy nem? Ekkor az eredmény a következő:

bűnös:	4+3 szavazat	a második és harmadik csoport tagjai ide szavaznak
nem bűnös:	5 szavazat	az első csoport mindegyik tagja ide szavaz

Tehát a szenátorok szerint a vádlott bűnös! *(ii)* Másodjára vizsgáljuk meg, hogy milyen büntetést kell kiszabni a vádlottra. Ekkor már csak két döntési alternatívánk marad, mert a felmentés lehetőségét ki kell zárnunk. A döntési menet tehát:

K	4 szavazat	a 2. csoport mindegyik tagja ide szavaz
B	3+5 szavazat	a harmadik csoport az elsődleges preferenciája, míg az első csoport a másodlagos preferenciája szerint szavaz ide

A nyertes a B alternatíva, a vádlottat börtönbe csuknák. A bemutatott eset azt szemlélteti, hogy a szavazások során nemcsak a leadott szavazatok tényleges eloszlása, de az alkalmazott aggregálási eljárás is befolyásolhatja a végeredményt. Sokféle paradoxont beazonosítottak, megneveztek e jelenségkörön belül, párat röviden bemutatok közülük. Az egyiket Ordeshook dominált győztes paradoxonának (dominated-winner paradox) nevezte el [552], és a példában mutatott eljárás azt az elvárásunkat sérti meg, hogy ha van egy alternatíva, amely egy másikkal szembeni összehasonlításban mindig alulmarad (vagyis mindenki preferálja x-t y-nal szemben), akkor a közösségi döntésben is megnyilvánuljon ez a sorrend. Hogy lássuk ezt, feltételezzünk három szavazót, négy alternatívát és a következő preferenciarendezéseket:

1. szavazó	$x\ y\ b\ a$
2. szavazó	$a\ x\ y\ b$
3. szavazó	$b\ a\ x\ y$

Az aggregálási eljárás pedig legyen olyan, hogy először az x és a alternatívák közötti páros összehasonlítások alapján hoz ki részgyőztes, majd ezt követően a győztest előbb a b-vel, majd az újabb győztest az y-nal veti össze, és ezzel véget is ér az algoritmus. Ebben az összehasonlítási folyamatban x és a versenyéből utóbbi jön ki győztesen (hiszen a 2. és 3. szavazónál is előbbre van a sorban). A második menetben a-t és b-t kell

összevetni, aminek *b* lesz a nyertese (az 1. és 3. szavazó miatt). Harmadjára *b* és *y* kerül egymással szembe, és ezt a fordulót *y* nyeri (mert az 1. és 2. szavazónál előbb van *b*-nél). Ez az algoritmus tehát a kollektív döntés győzteseként az *y* alternatívát hozza ki, ami szembeállítható azzal a ténnyel, hogy egyénileg mindhárom szavazó jobbnak tartotta *x*-t *y*-nál.

Másféle paradoxon adódik akkor, ha a Borda-pontok módszerével aggregálunk közösségi döntést. Ordeshook két paradoxont is megnevezett [552], itt csak azt mutatom be, aminek az "elsőkből utolsók" (winner-turns-loser) paradoxon nevet adta. Tételezzük fel az alábbi preferenciaeloszlásokat hét szavazó és négy alternatíva között.

1.	2.	3.	4.	5.	6.	7.	alternatíva	Borda-pont
a	*b*	*c*	*a*	*b*	*c*	*a*	*a*	$18 = (3\times4)+(2\times2)+(2\times1)$
b	*c*	*x*	*b*	*c*	*x*	*b*	*b*	$19 = (2\times4)+(3\times3)+(2\times1)$
c	*x*	*a*	*c*	*x*	*a*	*c*	*c*	$20 = (2\times4)+(2\times3)+(3\times2)$
x	*a*	*b*	*x*	*a*	*b*	*x*	*x*	$13 = (2\times3)+(2\times2)+(3\times1)$

A Borda-módszer szerint a legjobb helytől indulva a gyengébbek felé egyre kevesebb (itt: 4-3-2-1) pontot kell adni minden egyéni rangsorban elért helyezésért, és a közösségi döntéshez ezeket kell összegezni. Így eljárva a fenti példában a *c* alternatíva jön ki győztesnek. A paradoxon akkor keletkezik (válik láthatóvá), amikor ugyanezen szavazók és alternatívák közül kiveszünk egyet, az *x* alternatívát (mondjuk azért, mert – ha az alternatívák politikai jelölteket reprezentálnak – az *x* azonosítójú jelölt meghal rögtön a szavazás után). Ekkor újra kell számolni a Borda-pontokat úgy, hogy a preferenciarendezésekből kivesszük az *x* alternatívát.

1.	2.	3.	4.	5.	6.	7.	alternatíva	Borda-pont
a	*b*	*c*	*a*	*b*	*c*	*a*	*a*	$15 = (3\times3)+(2\times2)+(2\times1)$
b	*c*	*a*	*b*	*c*	*a*	*b*	*b*	$14 = (2\times3)+(3\times2)+(2\times1)$
c	*a*	*b*	*c*	*a*	*b*	*c*	*c*	$13 = (2\times3)+(2\times2)+(3\times1)$

A tanulság egyből levonható a két táblázat összevetéséből (most az *a* nyert a *b*-vel és a korábbi nyertes *c*-vel szemben). Az a tény, hogy a vesztesek közül valaki kiesik a versenyből, megváltoztathatja az elsőség kérdését. Ez a paradoxon egyben fontos példaként tekinthető arra vonatkozóan, hogy nem mindig érvényesül az irreleváns alternatívák függetlenségének elve.[8] A Borda-pontozás további paradoxonokat is okozhat. Ha a pontozási módszeren változtatunk, és az egyéni győzteseknek nem annyi pontot adunk, mint az alternatívák száma, hanem kevesebbet (és ekkor értelemszerűen az egyéni listák végefelé már nem fognak pontot kapni az alternatívák), akkor más és más győztest hozhatunk ki [552]. Ennek ismerete és alkalmazása olykor komoly manipulációs potenciált jelenthet a Borda-pontozást alkalmazó tendereztetések, versenyeztetések lebonyolítói számára.

A különböző szavazási, választási eljárások leírásával, következményeik feltárásával több tanulmányunkban is foglalkoztunk Mészáros József kollégámmal közösen [494], [500]. Ezek tanulságait nem mutatom be itt részletesebben, csak egyetlen példát idézek fel arra vonatkozóan, hogy milyen hatása lehet a különböző többmandátumos eljárásoknak

[8] Ezt az elvet a későbbiekben még részletesebben tárgyalni fogom (↬210.o.).

mandátumok elosztására vonatkozóan. Az 1994-es parlamenti választás adatai alapján a megyei listán a parlamenti pártokra leadott tényleges szavazatok alapján kiszámoltuk, hogy milyen eredmények születtek volna, ha más és más eljárással számolták volna ki az eredményeket [500]. A táblázatban feltüntetett eljárásokat itt nem mutatom be, most elég csak annyit tudni róluk, hogy mindegyiküket alkalmazni lehetett volna az éles számítások során (a hivatkozott korábbi tanulmányainkban bővebben lehet olvasni róluk).

	MSZP	SZDSZ	MDF	FKGP	FIDESZ	KDNP
legnagyobb maradék módszere	54,1	17,9	9,8	6,7	5,2	5,7
Imperiali módszer	55,4	17,9	9,6	6,7	4,7	5,2
d'Hondt módszer	57,5	19,4	9,1	6,0	3,3	4,1
Hagenbach-Bischoff módszer	54,9	17,1	9,6	7,0	5,2	5,7
Sainte-Laguë módszer (osztó: 1)	53,9	17,1	10,4	7,0	5,2	6,0
Sainte-Laguë módszer (osztó: 1,4)	55,7	18,4	9,8	5,2	4,4	4,9
Sainte-Laguë módszer (osztó: 1,2)	54,1	17,9	9,8	6,7	5,2	5,7

Az adatokból szépen kiolvasható, hogy ha nem is radikális, de nem is elhanyagolható változásokat okozhat az a tény, hogy kicseréljük a szavazatok összegzését végző algoritmusokat. Tudni kell azonban, hogy ennél sokkal komolyabb változást is tapasztalhatnánk, ha a választási rendszerben komolyabb átalakítást végeznénk. A magyar választási rendszer ugyanis több, egymástól jól elkülöníthető, és más logikát megtestesítő részből áll, és ha valamelyik összetevő súlyát megnövelnénk a többi rovására (például a listás rendszerkomponenshez rendelt mandátumok számát megnövelnék az egymandátumos választókerületi részrendszer rovására), akkor az sokkal jelentősebb következményekkel járna. Ennek a kérdéskörnek azonban ezen a ponton nem szentelek több figyelmet.

A bemutatott példák azt szemléltetik, hogy az aggregálási algoritmus maga befolyással lehet a közösségi döntés végeredményére, vagyis amikor közösségi döntést képezünk valamilyen eljárással, akkor az egyéni vélemények összessége mellett az összegzés technológiája is befolyással lehet a kimenetre. Ez felveti annak szükségességét, hogy megvizsgáljuk és értékeljük az összegző eljárásokat. Mielőtt ezt megtennénk, rá kell még mutatnom egy problémára, amivel a kutatók az aggregálási folyamatok nagyon korai vizsgálata során szembesültek. Ez a többszáz éve ismert paradoxon folyamatosan arra sarkallta a kutatókat, hogy keressék az ellentmondás feloldásának lehetőségét.

4.2.3 Condorcet-paradoxon

Nicolas de Condorcet francia matematikus és filozófus már a 18. században leírta azt a paradoxont, amit később róla neveztek el. Legalább Condorcet óta tudjuk tehát, hogy amikor az egyéni preferenciák alapján egy közösségi sorrendet kell kialakítani valamilyen alternatívahalmazon, akkor a többségi eljárással nem remélhetünk minden esetben végeredményt. A Condorcet-féle módszer a Condorcet-győztes elemet úgy hozza ki, hogy az egyéni rendezések alapján összehasonlítja az összes alternatívát egymással, és a páros összehasonlítások részeredményeit összesítve a többségi elv alapján hirdeti ki a győztest. A Condorcet-paradoxon lényege az, hogy könnyen preferenciakör, **tranzitív kör** alakulhat ki az ilyen eljárás során. A körkörös többség jelenségét szemléltethetjük az alábbi táblázat segítségével (szigorú preferenciarendezést feltételezve).

	feltétel	következmény
1. egyéni vélemény	$\mathbf{P}(m,n) \wedge \mathbf{P}(n,o)$	$\mathbf{P}(m,o)$
2. egyéni vélemény	$\mathbf{P}(n,o) \wedge \mathbf{P}(o,m)$	$\mathbf{P}(n,m)$
3. egyéni vélemény	$\mathbf{P}(o,m) \wedge \mathbf{P}(m,n)$	$\mathbf{P}(o,n)$
többségi vélemény	$\mathbf{P}(m,n) \wedge \mathbf{P}(n,o) \wedge \mathbf{P}(o,m)$	?

Mindegyik egyén racionálisnak mondható (mert tranzitív) preferenciákkal rendelkezik, de a csoport egészére, ha többségi szavazással kell dönteni az egyes alternatívák közti rendezésekről, akkor a tranzitivitásnak ellentmondó eredményhez jutunk. A Condorcet-paradoxon feloldására történtek mindenféle kísérletek, de ezek számunkra itt, most nem annyira érdekesek. Sokkal fontosabb az az – időben jóval későbbi – felismerés, hogy a paradoxon mélyében alapvetőbb ellentmondások rejlenek. Ennek megértéséhez meg kell néznünk, milyen tulajdonságokkal jellemezhetjük az aggregálási eljárásokat, és ezek alapján milyen összefüggéseket, tételeket állapíthatunk meg rájuk vonatkozóan.

Tartalmát tekintve ugyan nincs köze a választás problematikájához, mégis érdemes megemlíteni, hogy a hétköznapi életből jól ismert *kő-papír-olló* játék is a Condorcet-paradoxon lényegét adó "körbeverésen" alapul.

4.2.4 Arrow lehetetlenségi tétele

Nem nagyon vitatkoznék azzal, ha valaki azt állítaná, hogy a preferenciaaggregálás egész világa Kenneth J. Arrow lehetetlenségi tétele körül forog. Arrow 1950 fogalmazta meg először híres tételét [31], majd 1951-ben egy kis könyvet jelentetett meg ugyanerről a kérdésről [33].[9] Az egyéni preferenciákat aggregáló függvényt **társadalmi jóléti függvénynek** (social welfare function) nevezve, az aggregálási folyamatra vonatkozóan különböző elvárásokat megfogalmazva, Arrow kimondott egy tételt, miszerint azok az elvárások, amelyeket elég széles körű egyetértés alapján a véleményösszegzési eljárásokkal szemben támasztunk, nem elégíthetők ki egyszerre. Vegyük sorba egymás után, milyen elvárásaink lehetnek az összegző függvénnyel szemben.[10]

Nyilván meg kell követelnünk az aggregálási szabálytól azt, hogy mindenféle konstellációban képes legyen eredményt produkálni – akárhogyan is alakulnak az alternatívák és az egyéni preferenciarendezések. Ez az elvárás azonban kétféle dolgot is jelent, ami alapján a bemenetre és a kimenetre is megfogalmazhatunk egy-egy részfeltételt. A bemenetre vonatkozó feltételünk az lehet, hogy mindenféle bemenetet be tudjon fogadni a rendszer, ebben az értelemben legyen egyetemes, univerzális. A kimenetre vonatkozóan pedig azt az elvárást rögzíthetjük, hogy egyetlen, tehát egyértelmű végeredményt adjon ki a rendszer az összesítés után. Az persze előfordulhat – szinte bármelyik aggregáló szabály esetében –, hogy a szavazatok összegzése döntetlent eredményez. Ilyenkor szükség lehet egy másodlagos **döntetlentörő szabályra** (tie-break rule), ami segít a döntetlent megszüntetni. Ez azonban a főszabályra vonatkozó elvárásunkat nem érinti lényegi módon. A két tulajdonság, amit a fentiek alapján definiálhatunk a következő.

[9] Az 1951-es könyvet kisebb módosításokkal úra kiadták 1963-ban, én is erre hivatkozom itt.

[10] Megintcsak sokféle terminust használhatunk akkor, amikor elvárásokról beszélünk. Az összegzőfüggvény **tulajdonságaira** (property) vonatkozóan **feltételeket** (condition), **kényszereket** (constraint), **kritériumokat** (criterion) írhatunk elő. A későbbiekben ezeket a terminusokat mind használni fogom – ugyanabban az értelemben.

(AL.20) $(\text{UNI}) \equiv \text{UNIVERSAL}(\mathbf{F}) \stackrel{\text{def}}{=} \mathbb{D}_F = \mathbb{X}^n$

univerzális (értelmezési tartomány) (HUN) – korlátozatlan értelmezési tartomány (HUN) – universal (domain) (ENG) – unrestricted domain (ENG)

(AL.21) $(\text{DEC}) \equiv \text{DECISIVE}(\mathbf{F}) \stackrel{\text{def}}{=} |F(\Phi)| = 1$

döntésképes (HUN) – ügydöntő (HUN) – decisive (ENG)

Az első követelmény az **univerzalitás** tulajdonsága, ami tehát azt jelenti, hogy az összegző függvény értelmezve van minden lehetséges bemenetre. Matematikailag ezt úgy ragadhatjuk meg, hogy az $\mathbf{F}$ aggregáló függvény $\mathbb{D}_F$ értelmezési tartományát egyenlővé tesszük a teljes állapotteret jelző $\mathbb{X}^n$ halmazzal. Az univerzalitás felétele azt jelenti tartalmilag, hogy az aggregáló függvénytől elvárjuk azt, hogy bármilyen (bárki által megfogalmazott) lehetséges preferenciarendezést figyelembe vegyen a közösségi vélemény kialakításakor, ami egyet jelent azzal, hogy az egyének szabadon alakíthassák ki saját egyéni sorrendjeiket a világ dolgait illetően. Ez az elv tehát az szabad véleménykinyilvánítás egy fontos esetét írja le. Ha elhagyjuk az univerzalitás követelményét, akkor valamilyen módon korlátozott tartományról kell beszélnünk. Erre példa az egycsúcsú preferenciák esete (↬201.o.), amelyre vonatkozóan Duncan Black fontos tétel fogalmazott meg, amit később fogok bemutatni (↬216.o.). Bár a jelenség lényegét tekintve nem ide tartozik, de azért érdemes megjegyeznem, hogy amikor a diktatórikus rendszerek megakadályozzák, hogy a szavazáson tetszőleges jelöltek indulhassanak, akkor voltaképpen az itt tárgyalt univerzalizási kritériumnak nem tesznek eleget.

Amíg az első elvárás a függvény bemenetére rögzített egy kritériumot, addig az aggregáló függvény **döntésképes** jellegére irányuló feltétel a függvény kimenetére irányul. A fentebb már jelzett elvárás, miszerint az összegző eljárásnak minden lehetséges bemenetre egyértelmű végeredményt kell adnia, azt a szimpla és nagyon praktikus elvet akarja érvényesíteni, hogy értelme legyen az aggregálási eljárás lefolytatásának. Arra ugyanis pont azért van szükség, mert a sok egyéni vélemény nem egyezik meg egymással, tehát nincs konszenzus az egyéni szinten, az érintettek nem tudnak kollektíve dönteni. Ilyen helyzetekben azért kell az aggregálási eljárást elindítani, hogy legyen kollektív döntés. De csak akkor beszélhetünk döntésről, ha egyetlen kimenetünk van. Ez biztosítja az eljárás döntésképes tulajdonsága.

Taylor terminusa: valódi (resolute), döntetlenmentes, egyelemű halmazt hoz ki visszavezeti Duggan and Schwartz (1993, 2000) and goes back at least to Gärdenfors

———-

Sen CCR is Decisive = Complete Closure of R

———-

A lehetetlenségi tétel első megfogalmazásakor Arrow még öt tulajdonságot különített el egymástól [31], és csak később alakította át úgy az elméletét, hogy négy tulajdonságról legyen szó. A tétel első megfogalmazásában még szerepelt a **monotonitás** fogalma is, amit a relációelméleti fejezetben mi már definiáltunk, mint fontos függvénytulajdonságot (↬42.o.). A monotonitás kritériumát itt nyilván az $\mathbf{F}$ összegzőfüggvényre kell vonatkoztatnunk.

(AL.22) $(\text{MON}) \equiv \text{MONOTONE}(\mathbf{F}) \stackrel{\text{def}}{=} (\mathbf{F}(\langle\Phi_i\rangle_{i\in\mathbb{N}}) = m \wedge (\Phi_i = m \rightarrow \Phi_i' = m)) \rightarrow \mathbf{F}(\langle\Phi_i'\rangle_{i\in\mathbb{N}}) = m$

monoton (HUN) – monotone (ENG)

Arrow ezt az egyéni és csoportos értékek közti pozitív kapcsolat (positive association of social and individual values) elvárásaként is értelmezi [31]. Ha az egyéni rendezésekben az egyik alternatíva mindig egy másik előtt szerepel, akkor ugyanennek a rendezésnek kell megjelennie a kollektív preferencia szintjén is. A monotinitás helyett azonban nagyon hamar egy másik tulajdonság került be a lehetetlenségi tétel(ek)ről folytatott diskurzusba, az **irreleváns alternatíváktól való függetlenség** elvárása.[11] A fogalom értelmezését kezdhetjük azzal, amit Arrow maga mondott erről.[12]

> „Az ordinális hasznossághoz való újkeletű ragaszkodást a megkülönböztethetetlenek azonosságának leibniz-i elvével magyarázhatjuk. Csak a megfigyelhető különbségeket tehetjük meg a magyarázataink alapjául. ... Az Irreleváns Alternatívák Függetlenségének Feltétele egy lépéssel tovább megy a megfigyelhetőség követelményén. Ha adott a társadalom számára választható alternatívák halmaza, akkor – elméletileg – elvárható, hogy valaki képes lehet az összes alternatíva mérlegelésére, de nem várhatjuk el azt, hogy figyelembe vegyen a társadalom számára nem elérhető (nem létező) alternatívák közti preferenciákat." [33, 109-110]

Azt az elvárást kell itt megfogalmazni valahogyan, hogy a közösségi preferenciarendezések eredményei csak az összehasonlításba bevont alternatívákra vonatkozó egyéni preferenciáktól függjenek, semmi mástól. Ha az egyéni preferenciák két alternatíva között nem változtak, de – esetleg – más alternatívák között igen, akkor a közösségi preferencia sem változhat a szóbanforgó alternatívákra.

(AL.23) $(\text{IIA}) \equiv \text{IIA}(\mathbf{F}) \stackrel{\text{def}}{=} \forall i \forall m \forall n((\Phi_i(m,n) \leftrightarrow \Phi_i'(m,n)) \rightarrow (\mathbf{F}(\Phi)(m,n) \leftrightarrow \mathbf{F}(\Phi_i')(m,n)))$

független az irreleváns altenatíváktól (HUN) – independent from the irrelevant alternatives (ENG) – binary independence (ENG)

Az elv szemléltetésére nézzünk meg egy egyszerű példát. Tegyük fel, hogy van három szavazónk, v_1, v_2 és v_3, valmint négy alternatívánk, m, n, o és p. A következő táblázatban tüntessük fel azt, hogy a szavazók hogyan rangsorolják az alternatívákat, illetve milyen sorrendet hoz ki az **F** szavazási függvény két alternatíva, m és n között. Az irreleváns alternatívák függetlensége azt jelenti, hogy ha az első sorban feltüntetett egyéni rendezésekből az összegző függvény a m alternatívát hozza ki a n elé, akkor minden más profilegyüttesre is tartania kell ezt a sorrendet.

v_1 szavazó	v_2 szavazó	v_3 szavazó	**F**	
(m,n,o,p)	(n,p,o,m)	(o,m,p,n)	$m > n)$	
(m,o,p,n)	(o,n,p,m)	(m,n,o,p)	$m > n)$	jó
(p,m,n,o)	(n,m,o,p)	(o,p,m,n)	$m > n)$	jó
(n,p,o,m)	(o,p,n,m)	(p,m,o,n)	$m > n)$	rossz
(m,n,o,p)	(n,p,o,m)	(o,m,p,n)	$m < n)$	rossz

[11] Az irreleváns alternatívák függetlensége kifejezés korábban már előkerült, az egyéni racionalitás jellemzésekor (↬184.o.). A két fogalom különbözik egymástól. Itt a kollektív preferencia kialakítását végző eljárás egyik fontos feltételét, korábban a racionális egyén egyik tulajdonságát jelenti

[12] Az idézet első mondatának értelmezéséhez tudni kell azt, hogy itt ordinális rendezésről beszélünk mindvégig.

A táblázatban két-két példát mutatok be az elvnek megfelelő (2-3. sor), illetve az elvet megsértő (4-5. sor) konstellációra. A negyedik sor példája azzal sérti meg a vizsgált elvet, hogy a v_1 szavazó esetében nem áll fent két alternatíva közti sorrend változatlansága, az ötödik sor pedig azért mutat "rossz eloszlást", mert az összegzőfüggvény megváltoztatja az alternatívák viszonyát egymáshoz. Azok a választási szabályok tehát, amelyek az első három sor alapján működnek, kielégítik az irreleváns alternatívák függetlenségének kritétiumát, a két utolsó sor pedig olyan választási függvényeket reprezentál, amelyek megsértik az (IIA) elvét. Korábban egyébként már bemutattam egy algoritmust, ami ugyancsak nem felelt meg az irreleváns alternatívák függetlensége elvének (↰206.o.).

———————

irreleváns alternatívákról [596]

Bengt Hansson a legjobb alternatíva függvényre támaszkodva alternatív meghatározást adott az irreleváns alternatívák függetlenségének elvére Bengt Hansson, Choice Structures and Preference Relations [328]

———————-

A következő szempont arra vonatkozik, hogy az összegző függvény hogyan "viselkedjen" akkor, ha az egyéni profilok nagyon hasonlítanak egymásra, netán ugyanolyanok, vagyis mi legyen a végeredmény az egyéni vélemények egyhangúsága esetén. Ezt a közösségi véleményt kifejező függvény Pareto-hatékonyságával fejezhetjük ki. A Pareto-hatékonyságot azonban háromféle értelemben is megadhatjuk, és a háromféle értelmezést más és más tételek során lehet felhasználni. Nézzük meg először a definíciókat!

(AL.24) $(\text{wPE}) \equiv \text{WEAKLY_PARETO_EFFICIENT}(\mathbf{F}) \stackrel{\text{def}}{=} \forall S \forall m \forall n((a = (\mathbf{R}_1^a, \mathbf{R}_2^a, \ldots, \mathbf{R}_n^a,) \wedge (m,n) \in \mathbb{X} \wedge C^a = \mathbf{F}(a)) \rightarrow ((m,n) \in \bigcap_{i \in N} \mathbf{P}(\mathbf{R}_i^a) \rightarrow (m \in C^a(\{m,n\}))))$

gyengén Pareto hatékony (HUN) – weakly Pareto efficient (ENG) – gyengén Pareto optimális (HUN) – weakly Pareto optimal (ENG) – wPo (ENG)

(AL.25) $(\text{PE}) \equiv \text{PARETO_EFFICIENT}(\mathbf{F}) \stackrel{\text{def}}{=} (a = (\mathbf{R}_1^a, \mathbf{R}_2^a, \ldots, \mathbf{R}_n^a,) \wedge (m,n) \in \mathbb{X}) \rightarrow ((m,n) \in \bigcap_{i \in N} \mathbf{P}(\mathbf{R}_i^a) \rightarrow (m \in S \rightarrow n \notin C^a(S)))$

Pareto hatékony (HUN) – Pareto efficient ((ENG)) – Pareto optimális (HUN) – Pareto optimal (ENG) – Po (ENG)

(AL.26) $(\text{sPE}) \equiv \text{STRONGLY_PARETO_EFFICIENT}(\mathbf{F}) \stackrel{\text{def}}{=} (a = (\mathbf{R}_1^a, \mathbf{R}_2^a, \ldots, \mathbf{R}_n^a,) \wedge (m,n) \in \mathbb{X} \wedge C^a = \mathbf{F}(a)) \rightarrow ((m,n) \in \bigcap_{i \in N} \mathbf{P}(\mathbf{R}_i^a) \rightarrow (m \in S \rightarrow n \notin C^a(S)))$

erősen Pareto hatékony (HUN) – strongly Pareto efficient (ENG) – erősen Pareto optimális (HUN) – strongly Pareto optimal (ENG) – sPo (ENG)

(AL.27) $(\text{UNA}) \equiv \text{UNANIMOUS}(\mathbf{F}) \stackrel{\text{def}}{=}$

egyhangú (HUN) – egyöntetű (HUN) – unanimous (ENG)

egyhangú, ha az m alternatívát minden szavazó prefereálja az n-nel szemben, akkor n nem lehet győztes.

———————

Ha az egyöntetűséget mint a konszenzus elvének megvalósítását értelmezzük, vagyis az egyhangú szavazás elvárásával a konszenzust várunk el, akkor a szavazók a szavazási folyamat során könnyen stratégiai szavazókká válhatnak, és elkezdhetik "működtetni" a szavazatcsere intézményét (↰226.o.).

konszenzus és egyetértés (aggreement) közti különbségtétel

Az is fontos elvárás lehet a választási függvénnyel szemben, hogy ne egyetlen személy (vagy csoport) véleménye érvényesüljön az aggregálás során, vagyis azt várjuk el, hogy mindenki véleménye számítson az összegzés során, és ne fordulhassanak elő oligarchikus vagy diktatórikus megoldások. Ha arra a szempontra figyelünk, hogy milyen az egyéni és csoportos vélemények közti viszony, kétféle tulajdonságot is definiálhatunk, a **kikényszerítettséget** és a **diktatórikusságot** (illetve ezek komplementerét). Ezeket azonban akkor érthetjük meg jól, ha előtte még más fogalmakat is bevezetünk a modellünkbe. Először is képeznünk kell a szavazók tetszőleges csoportját, koalícióját, amelyek közül kiemeljük majd azokat, amelyeknek valamilyen módon meghatározó befolyása van a szavazás kimenetére.

(AL.28) $(\text{DEC}) \equiv \text{DECISIVE}(\mathbf{V}) \stackrel{\text{def}}{=} \exists\mathbb{V}\exists m\exists n(((m,n) \in \mathbb{X} \wedge \mathbb{V} \subset \mathbb{N} \wedge \mathbb{V} \neq \varnothing) \rightarrow ((m,n) \in \bigcap_{i\in\mathbb{V}} \mathbf{P}(\mathbf{R}_i^a) \rightarrow \{m\} = C^a(\{m,n\})))$

(lokálisan) döntő (HUN) – (locally) decisive (ENG)

(AL.29) $(\text{ADE}) \equiv \text{ALMOST_DECISIVE}(\mathbf{V}) \stackrel{\text{def}}{=} \exists\mathbb{V}\exists m\exists n(((m,n) \in \mathbb{X} \wedge \mathbb{V} \subset \mathbb{N} \wedge \mathbb{V} \neq \varnothing) \rightarrow (((m,n) \in \bigcap_{i\in\mathbb{V}} \mathbf{P}(\mathbf{R}_i^a) \wedge (n,m) \in \bigcap_{i\in\mathbb{N}\setminus\mathbb{V}} \mathbf{P}(\mathbf{R}_i^a)) \rightarrow \{m\} = C^a(\{m,n\})))$

(lokálisan) majdnem döntő (HUN) – (locally) almost decisive (ENG)

(AL.30) $(\text{BLO}) \equiv \text{BLOCKING}(\mathbf{V}) \stackrel{\text{def}}{=} \exists\mathbb{V}\exists m\exists n(((m,n) \in \mathbb{X} \wedge \mathbb{V} \subset \mathbb{N} \wedge \mathbb{V} \neq \varnothing) \rightarrow ((m,n) \in \bigcap_{i\in\mathbb{V}} \mathbf{P}(\mathbf{R}_i^a) \rightarrow m \in C^a(\{m,n\})))$

(lokálisan) blokkoló (HUN) – (locally) blocking (ENG)

(AL.31) $(\text{ABL}) \equiv \text{ALMOST_BLOCKING}(\mathbf{V}) \stackrel{\text{def}}{=} \exists\mathbb{V}\exists m\exists n(((m,n) \in \mathbb{X} \wedge \mathbb{V} \subset \mathbb{N} \wedge \mathbb{V} \neq \varnothing) \rightarrow (((m,n) \in \bigcap_{i\in\mathbb{V}} \mathbf{P}(\mathbf{R}_i^a) \wedge (n,m) \in \bigcap_{i\in\mathbb{N}\setminus\mathbb{V}} \mathbf{P}(\mathbf{R}_i^a)) \rightarrow m \in C^a(\{m,n\})))$

(lokálisan) majdnem blokkoló (HUN) – (locally) almost blocking (ENG)

Az itt megadott, a szavazók koalíciójára vonatkozó ‘döntő’ (decisive) értelmezés eltér a korábban definiált ‘ügydöntő’ (decisive) tulajdonságtól (↬212.o.), amit majd később, a May-tételhez kapcsolódva használunk fel (↬219.o.). A majdnem döntő minőség formulája itt azt fejezi ki, hogy a majdnem döntő koalíció tagjai egyhangú véleménye határozza meg a közösségi döntés végeredményét úgy, hogy eközben mindenki más preferenciája pont a koalíció tagjainak preferenciájával ellentétes. Ha tehát a majdnem döntő csoport tagjai mind m-t preferálják n-nel szemben, és rajtuk kívül mindenki más az n-t az m ellenében, legyenek utóbbiak akármennyien, akkor is a koalíció véleménye számít, tehát az m alternatíva lesz a győztes. A majdnem döntő koalíció definíciójában szerepel a koalíción kívüliek ellenkező véleménye, míg a döntő csoport meghatározásából ez a mozzanat hiányzik. Éppen ezért a döntő csoport fogalma általánosabb, mert a definíciójában nincs semmilyen további feltétel kikötve. Kicsit félrevezetőnek is tűnhet a ‘majdnem döntő’ kifejezés, mert ha a meghatározása szerint a majdnem döntő csoport tagjai a csoporton kívüli összes szavazó egyhangú ellenvéleménye ellenére képesek meghatározni a közösségi végeredményt, akkor felmerülhet a kérdés, hogy a teljes ellenkezés dacára elért győzelemnél mi lehet még erősebb (hogy ne legyen szükség a ‘majdnem’ jelző alkalmazására). A válasz pedig az, hogy logikailag általánosabb, tehát erősebb az a fogalom (a ‘döntő koalíció’ fogalma), amire semmilyen feltétel sem vonatkozik.

A (majdnem) döntő, illetve a (majdnem) blokkoló koalíció fogalmak még lokálisak abban az értelemben, hogy két adott alternatíva (m és n) összehasonlításán alapulnak, de a definíciók nem követelték meg, hogy minden alterntatívára igazak legyenek a megadott feltételek. Globálissá tehetjük a fogalmakat akkor, ha minden alternatíva összehasonlítására kiterjesztjük a feltételek érvényességét (vagyis az alternatívákra vonatkozóan egzisztenciális kvantorok helyett univerzális kvantorokat alkalmazunk).

(AL.32) $(\text{GDEC}) \equiv \text{GLOBALLY_DECISIVE}(\mathbf{F}) \stackrel{\text{def}}{=} \exists\mathbb{V}\forall m\forall n(((m,n)\in\mathbb{X} \wedge \mathbb{V}\subset\mathbb{N}\wedge\mathbb{V}\neq\varnothing)\rightarrow((m,n)\in \bigcap_{i\in\mathbb{V}}\mathbf{P}(\mathbf{R}_i^a) \rightarrow \{m\}=C^a(\{m,n\})))$

(globálisan) döntő (HUN) – (globally) decisive (ENG)

(AL.33) $(\text{GADE}) \equiv \text{GLOBALLY_ALMOST_DECISIVE}(\mathbf{F}) \stackrel{\text{def}}{=} \exists\mathbb{V}\forall m\forall n(((m,n)\in\mathbb{X} \wedge \mathbb{V}\subset\mathbb{N}\wedge\mathbb{V}\neq\varnothing)\rightarrow (((m,n)\in \bigcap_{i\in\mathbb{V}}\mathbf{P}(\mathbf{R}_i^a)\wedge(n,m)\in \bigcap_{i\in\mathbb{N}\setminus\mathbb{V}}\mathbf{P}(\mathbf{R}_i^a)) \rightarrow \{m\}=C^a(\{m,n\})))$

(globálisan) majdnem döntő (HUN) – (globally) almost decisive (ENG)

(AL.34) $(\text{GBLO}) \equiv \text{GLOBALLY_BLOCKING}(\mathbf{F}) \stackrel{\text{def}}{=} \exists\mathbb{V}\forall m\forall n(((m,n)\in\mathbb{X} \wedge \mathbb{V}\subset\mathbb{N}\wedge\mathbb{V}\neq\varnothing)\rightarrow((m,n)\in \bigcap_{i\in\mathbb{V}}\mathbf{P}(\mathbf{R}_i^a) \rightarrow m\in C^a(\{m,n\})))$

globálisan) blokkoló (HUN) – (globally) blocking (ENG)

(AL.35) $(\text{GABL}) \equiv \text{GLOBALLY_ALMOST_BLOCKING}(\mathbf{F}) \stackrel{\text{def}}{=} \exists\mathbb{V}\forall m\forall n(((m,n)\in\mathbb{X} \wedge \mathbb{V}\subset\mathbb{N}\wedge\mathbb{V}\neq\varnothing)\rightarrow (((m,n)\in \bigcap_{i\in\mathbb{V}}\mathbf{P}(\mathbf{R}_i^a)\wedge(n,m)\in \bigcap_{i\in\mathbb{N}\setminus\mathbb{V}}\mathbf{P}(\mathbf{R}_i^a)) \rightarrow m\in C^a(\{m,n\})))$

(globálisan) majdnem blokkoló (HUN) – (globally) almost blocking (ENG)

A döntő-, illetve blokkolóképesség fogalmaira támaszkodva "végre" definiálhatjuk az oligarcha (oligarchikus szavazó) és a diktátor (diktatorikus szavazó) fogalmát, amelyekre aztán a lehetetlenségi tételek kifejtése során sokat fogunk hivatkozni.

(AL.36) $(\text{OLI}) \equiv \text{OLIGARCHIC}(\mathbf{F}) \stackrel{\text{def}}{=}$

oligarchikus (HUN) – oligarchic (ENG)

(AL.37) $(\text{NOLI}) \equiv \text{NON-OLIGARCHIC}(\mathbf{F}) \stackrel{\text{def}}{=} \neg\text{OLIGARCHIC}(\mathbf{F})$

nem oligarchikus (HUN) – non-oligarchic (ENG)

(AL.38) $(\text{DIC}) \equiv \text{DICTATORIAL}(\mathbf{F}) \stackrel{\text{def}}{=} \exists!\, i\in\mathbb{N}(\mathbf{F}(\mathbf{R}^a(m,n)) \leftrightarrow \mathbf{R}_i^a(m,n))$

diktatórikus (HUN) – dictatorial (ENG)

(AL.39) $(\text{NODIC}) \equiv \text{NON-DICTATORIAL}(\mathbf{F}) \stackrel{\text{def}}{=} \neg\text{DICTATORIAL}(\mathbf{F})$

diktátormentes (HUN) – nem-diktatórikus (HUN) – non-dictatorial (ENG)

Az $\mathbf{F}$ véleményösszegző függvény diktatórikus, ha az általa aggregált kollektív vélemény minden esetben megegyezik egyetlen (az i.) szavazó $\mathbf{R}_i^a(m,n)$ véleményével.[13] Ha ilyen nincs, akkor a szavazási rendszer nem diktatórikus. Meg kell jegyezni, hogy a diktátormentesség követelménye – a dolgokat a tényleges társadalmi praxis felől szemlélve – nem igazán erős. Az egyszemélyi akarat (a diktatúra) érvényesülését ugyanis ki lehet zárni már

[13] Az általam adott meghatározása tartalmazza, de általában nem szokták beletenni a definíciókba az egyetlen individuumra való szűkítést.

azzal is, ha két személy (egy kisebb csoport) véleménye érvényesül. A diktátormentesség megkövetelése a véleménymonopólium kizárását biztosítja, de a véleményoligopóliumét már nem.

kérdés: itt és fentebb kell-e a két-alternatívás szűkítés a definícióban????? Annak ellenére, hogy Arrow már a tétele első megfogalmazásakor definálta és használta a kikényszerítettség fogalmát, a témakör diskurzusában kevesebb figyelmet kapott ez a kategória.

(AL.40) $(\text{IMP}) \equiv \text{IMPOSED}(\mathbf{F}) \stackrel{\text{def}}{=}$

kikényszerített ????? (HUN) – imposed (ENG)

(AL.41) $(\text{NOIMP}) \equiv \text{NON-IMPOSED}(\mathbf{F}) \stackrel{\text{def}}{=} \neg\text{IMPOSED}(\mathbf{F})$

nem kikényszerített ????? (HUN) – non-imposed (ENG)

A nem-kikényszerítettség tulajdonságát Arrow az állampolgári szuverenitás (citizen's sovereignty) feltételének is nevezi [31].

A társadalmi jóléti függvényre vonatkoztatható feltételek bemutatása után következhet Arrow lehetetlenségi tételének kimondása: nem létezik olyan társadalmi jóléti függvény, amely minden preferenciaprofilra helyes rendezést eredményez (tehát reflexív, tranzitív és anti-szimmetrikus relációt) és teljesíti az (UNI), (PE), (IIA) és (NOD) feltételeket.

(AL.42) $\neg\exists\mathbf{F}\big((\text{UNI})(\mathbf{F}) \wedge (\text{PE})(\mathbf{F}) \wedge (\text{IIA})(\mathbf{F}) \wedge (\text{NODIC})(\mathbf{F})\big)$

Arrow lehetetlenségi tétele (HUN) – Arrow's impossibility theorem (ENG) – General Possibility Theorem (ENG) – Arrow's paradox (ENG)

A tételt az idők során sokféleképpen is bizonyították. Sen két lépésben adta meg a bizonyítást [662, Chapter 3*]. Először egy segédtételben kimondta, hogy ha egy aggregálási metódusra teljesül a (UNI), a (PE) és a (IIA) feltétel, illetve van legalább egy olyan szavazó, aki az összes alternatívapárra vonatkozóan majdnem döntő, akkor annak a szavazónak diktátornak kell lennie. Második lépésben pedig megmutatta, hogy ha az első három feltétel teljesül, akkor mindig van ilyen majdnem döntő szavazó, aki – a segédtétel értelmében – diktátor is egyben.

A jól ismert, széles körben alkalmazott kollektív döntési eljárások közül, amelyek megsértik valamelyik feltételt a fenti négyből, a többségi eljárás kielégíti a (PE), a (IIA) és a (NOD) feltételeket, de az univerzalitási kritériumot már nem [662, 38]. Sen bemutatott egy konkrét példát arra, hogy a Borda-pontozási rendszer hogyan sérti meg az irreleváns alternatívák függetlensége elvét.

1. szavazás				2. szavazás			
preferenciák		Borda-pontok		preferenciák		Borda-pontok	
1. szavazó:	$m\,n\,o$	m:	$7 = 1\times3+2\times2$	1. szavazó:	$m\,o\,n$	m:	$7 = 1\times3+2\times2$
2. szavazó:	$o\,m\,n$	n:	$4 = 1\times2+2\times1$	2. szavazó:	$o\,m\,n$	n:	$3 = 3\times1$
3. szavazó:	$o\,m\,n$	o:	$7 = 2\times3+1\times1$	3. szavazó:	$o\,m\,n$	o:	$8 = 2\times3+1\times2$

Sen a példájában két szavazás során kiosztott Borda-pontozást mutat be, amikor a két szavazás csak annyiban tér el egymástól, hogy az 1. szavazó másként szavaz a két fordulóban, mert menet közben átértékeli a preferenciasorrendjét, megcseréli o és n sorrendjét. Az első fordulóban m és o lesz a Borda-győztes, a második fordulót viszont o

nyeri meg egyedüliként, miközben m és o egymáshoz való viszonya nem változott meg. Mivel a végeredmény mégis változott, így kimondhatjuk, hogy az eljárás megsérti az (IIA) feltételt.

———————-

Arrow tétele után sokféle átértelmezés, új fogalmak, új bizonyítások jelentek meg a szakmai diskurzusban.

az univerzalitás kritérium néllkül is igaz marad a tétel, tehát már három feltétel együttes fennállását is ki lehet zárni.

Arrow bizonyítása a diktátor szavazóra épül leírja Amartya K. Sen [662] elemzi még Kotaro Suzumura [718]

Másfajta bizonyítást mutatott be Salvador Barberà [47]. Bizonyítását egy új fogalomra támaszkodva adta meg, amikor bevezette a **pivotális szavazó** fogalmát. E fogalommal azt a szavazót jelölte meg, akinek a szavazata fordulópontot jelent a szavazás során, vagyis ha ő megváltoztatja a saját szavazatát, akkor a közösségi döntés maga is megváltozik. Barberà bizonyításának lényege úgy összegezhető, hogy ha társadalmi jóléti függvény kielégíti az irreleváns alternatívák függetlensége és a Pareto-hatékonyság elvét, akkor

(AL.43) $(\text{PIV})(\mathbf{F}) \equiv \text{PIVOTAL_VOTER} \stackrel{\text{def}}{=}$

pivotális szavazó (HUN) – kulcsszavazó (HUN) – pivotal voter (ENG)

pivotális szavazó [471] [472]

$\mathbf{P}_i$ – a szavazók preferenciarendezései $\langle \mathbf{P}_i^a \rangle_{i \in \mathbb{N}}$ – preferenciaprofil = az összes szavazó preferenciarendezésének rendezett n-ese

megint más bizonyítást mutat be: John Geanakoplos [283]

Peter Bernholz, A General Constitutional Possibility Theorem [76]

két könyv, matematikai alapú áttekintés, kritika [276] [740]

elsőrendű logikával leírva [305] [306]

[288], Suzumura, Sen, [468, 16]

(AL.44) $|A| \geq 3 \wedge (D = WL(A)^m \vee D = L(A)^m) \wedge (?\text{PP}) \rightarrow (\text{DICT})$

Gibbard oligarcha tétele (HUN) – Gibbard's oligarchy theorem (ENG)

Tegyük fel, hogy jAj ? 3; és legyen az F egy kvázi Arrow-típusú összjóléti függvény, melyre teljesül i) D = WL(A)m vagy D = L(A)m; ii) a Pareto-feltétel. Ekkor az F oligarchikus.

4.2.5 Lehetségességi tételek

Az Arrow tétel és az azt követő további lehetetlenségi tételek azt mutaták meg, hogy mi lehetetlen, vagyis mit nem tehetünk meg, de nyitva hagyták a kérdést, hogy mi lehetséges, vagyis mit tehetünk meg. Tudunk-e pozitív állításokat tenni a társadalmi választási függvényekre vonatkozóan. A válasz természetesen az, hogy igen. A következőkben ezek közül bemutatok pár fontos tételt.

már maga Arrow is rámutatott két esetra, amikor létezhet megoldás

más feltételeket definiált Ken-Ichi Inada 1955-ben [371]

illetve 1964-ben: [369] ekkor különböző preferenciatípusokra mondta ki a konzisztens társas preferencia létezését többségi szabály mellett

Sen: possibility theorem [659]
A Simple Characterization of Majority Rule: [135]
Blau tétele: [93]

Black tétel

Duncan Black 1948-ban, tehát még Arrow lehetetlenségi tételének "felbukkanása" előtt megfogalmazott egy tézist, ami egy speciális preferencia esetén az Arrow-féle kritériumok együttes kielégíthetőségét állította [91]. Black tézise nem az Arrow-tétel cáfolata, hiszen a preferenciák egy szűkebb tartományára, mégpedig az (SPP)(**P**)) egycsúcsú preferenciákra (↬201.o.) vonatkozóan áll csak fent az igazsága. Az egycsúcsú preferenciák létét csak úgy feltételezhetjük, ha a választási függvény értelmezési tartományára nem áll fent az univerzalitás kritériuma, vagyis korlátozott tartományról beszélünk. A Black tétel éppen ezért "létezhet" az Arrow-féle lehetetlenségi tétel mellett, mert az előbbi érvénytelenként kezeli az utóbbi egyik alapfeltételét. A Black tétel az alábbi kondicionálissal írható le.

(AL.45) $$(\mathbf{F} = \mathbf{F}^{AP} \wedge (\text{SPP})(\mathbf{P})) \rightarrow ((\text{NODIC})(\mathbf{F}) \wedge (\text{IIA})(\mathbf{F}))$$

Black tétel (HUN) – Black theorem (ENG)

A Black-tétel értelmében tehát amennyiben az egyéni preferenciák egycsúcsúak, akkor a(z egyszerű) többségi döntési szabály nem diktatórikus és kielégíti az irreleváns alternatívák függetlenségének elvét.

---

elmondani, hogy miért korlátozott tartományról van itt szó
Black's median voter theorem

Arrow két tétele

Ritkábban szokták idézni, de ettől még tény, hogy Arrow a lehetetlenségi tételéről szóló könyvében foglalkozott azzal a kérdéssel is, hogy vannak-e lehetségeségi tételek, és be is mutatott belőlük párat.

Inada tétel

A preferencia fogalmáról bővebben értekezve korábban bemutattam, hogy vannak preferenciák, amelyek jól meghatározott profillal rendelkeznek (↬197.o.). Ezek közül néhány esetben kimondhatunk lehetségességi tételeket a társadalmi választási függvényre vonatkozóan.

- dichotóm preferencia
- visszhangszerű preferencia
- antagonisztikus preferencia
- értékkorlátos preferencia
  - egycsúcsú preferencia
  - egyvölgyű preferencia
  - szeparált preferencia

- tabu preferencia

Az első három típus esetében nincs megkötés a szavazók számosságára vonatkozóan, az utolsó kettőre annyi elvárás van, hogy a szavazók számra páratlan kell legyen.

Ken-ichi Inada tétele [370] többségi elárás tranzitív

(AL.46) $(\mathbf{F} =$

Inada tétel (HUN) – Inada theorem (ENG)

May tétel

Az Arrow-féle lehetetlenségi tételre azonnal reagált Kenneth O. May, aki egy 1952-es cikkében fejtette ki azt a tételét, amit később róla neveztek el [476]. May kétalternatívás helyzetet (m és n választási alternatívával), benne minden egyén számára három döntési lehetőséget $(\mathbf{P}(m,n), \mathbf{I}(m,n), \mathbf{P}(n,m))$ feltételezett, és a közösségi döntések jellemzésére bevezette a **csoportdöntési függvény** (group decision function) kifejezést, ami ugyanaz a fogalmat jelöli, mint a véleményösszegző függvény terminusa. Ezek után megfogalmazott négy feltételt, amelyek teljesülését az F csoportdöntési függvénytől "el lehet várni".

May elsőként a döntésképesség (decisive) tulajdonságát vezette be. Mivel nem adott pontos formulákat a fogalmaira, ezért a szóbeli leírásait nekünk kell interpretálni és formalizálni. A szakirodalomban ezt a may-i kritériumot leggyakrabban az univerzalitás fogalmával azonosítják [662], [178]. Ezt a fogalmat korábban már bemutattam (↬209.o.). Én kicsit másként értelmezem May felfogását. Szerintem ő azt akarta rögzíteni, hogy a csoportdöntési függvény legyen minden lehetséges bemenetre értelmezve, és minden bemenő értékre adjon egyértelmű kimenetet. Nem véletlen, hogy May maga döntésképesnek (decisive) nevezte ezt a feltételt: a tétel bizonyítása során az egyértelmű végeredmény elvárására hivatkozott. Igaz, hogy ettől még a feltételek közé kell sorolnunk a tetszőleges bemenet "kezelésének" elvárását, tehát az univerzalitást is. Ebből az következik, hogy – megítélésem szerint – May a döntésképes fogalma alatt az $\mathbf{F}$ választási függvény döntésképes és univerzális jellegét értette.[14]s

May következő két tulajdonsága valamifajta szimmetriát fejez ki – egyfelől a szavazók, másfelől az alternatívák között. A szavazók közti szimmetriát az **anonimitás**, az alternatívák köztit a **semlegesség** tulajdonságaival ragadhatjuk meg. A szimmetricitást itt úgy kell érteni, hogy mind a szavazók, mind a választható alternatívák egyenrangúak egymással, nincs egyik sem kivételezett helyzetben a többihez képest, és ebben az értelemben nincs semmiféle aszimmetria közöttük. Az anonimitás definiálásához szükség van egy σ segédfüggvényre, amely a szavazók sorrendjét képes tetszőleges módon előállítani (matematikailag ez annyit jelent, hogy a $\sigma : \mathbb{N} \mapsto \mathbb{N}$ függvény a szavazók $\mathbb{N}$ halmazának tetszőleges permutációját képes létrehozni). A segédfüggvény segítségével már meghatározhatjuk az **anonimitás** tulajdonságát.

(AL.47) $(\text{ANYM}) \equiv \text{ANONYM}(\mathbf{F}) \stackrel{\text{def}}{=} \forall m \forall n \forall i \forall \sigma(\mathbf{F}(\langle \mathbf{R}_i^a(m,n)\rangle_{i\in\mathbb{N}}) = \mathbf{F}(\langle \mathbf{R}_{\sigma(i)}^a(m,n)\rangle_{i\in\mathbb{N}}))$

anomin (HUN) – egalitáriánus (HUN) – egalitarian (ENG) – anonym (ENG)

[14] Az egyértelműség kedvéért azt jeleznem érdemes itt, hogy ez a may-i 'döntésképes' minősítés más, mint a korábban bevezetett, a szavazók koalíciójára vonatkozó 'döntő' tulajdonság (↬212.o.).

Az anonimitás feltétele azt jelenti, hogy a csoportdöntés nem változik, ha a bemenő egyéni szavazatok sorrendjét tetszőleges módon megváltoztatjuk.[15] Ebből azonnal következik, hogy egyfelől mindenki szavazata egyenlőnek számít, másfelől ezáltal teljesíteni lehet az anonimitás követelményét, hiszen a végeredményt tekintve nem számít, hogy ki, mikor, melyik szavazólapon adta le szavazatát. May meg is jegyezte, hogy lehetne ezt a tulajdonságot egalitáriánusnak is nevezni, amely kategória egyértelműen utal a szavazók körében elvárt "egyenlőség" szimmetriájára.

A másik szimmetricitási tulajdonság az alternatívákra, jelöltekre vonatkozik. Ahogy a neutralitásnál azt kellett biztosítani, hogy egyik szavazó se legyen se jobb, se rosszabb helyzetben a többihez képest, úgy itt valami hasonlót kell az alternatívákra megkövetelni. Elvárjuk, hogy minden jelölt helyzete legyen ugyanolyan, vagyis az alternatívák egyenlő eséllyel rendelkezzenek, ne legyen köztük bármiféle aszimmetria. Az anonimitás kritériumával azt fejezhetjük ki, hogy a csoportdöntés érzéketlen az alternatívákra, nem kedvez egyik számára sem.

———-

anonimitásról, semlegességről: [61]

———————-

(AL.48) $$(\text{NEUT}) \equiv \text{NEUTRAL}(\mathbf{F}) \stackrel{\text{def}}{=} \forall m \forall n \forall o \forall p((\forall i \forall \sigma(\mathbf{R}_i^a)(m,n) \leftrightarrow \mathbf{R}_{\sigma(i)}^a)(o,p)) \wedge (\forall i \forall \sigma(\mathbf{R}_i^a)(n,m) \leftrightarrow \mathbf{R}_{\sigma(i)}^a)(p,o))) \rightarrow ((\mathbf{F}(\langle \mathbf{R}_i^a(m,n)\rangle_{i\in\mathbb{N}})) \leftrightarrow \mathbf{F}(\langle \mathbf{R}_{\sigma(i)}^a(o,p)\rangle_{i\in\mathbb{N}})) \wedge ((\mathbf{F}(\langle \mathbf{R}_i^a(n,m)\rangle_{i\in\mathbb{N}})) \leftrightarrow \mathbf{F}(\langle \mathbf{R}_{\sigma(i)}^a(p,o)\rangle_{i\in\mathbb{N}})))$$

semleges (HUN) – neutral (ENG)

A formula értelmezéséhez figyelembe kell venni azt a tényt, hogy May olyan helyzetet feltételezett, amelyben három döntés, tehát háromféle szavazat képzelhető csak el, amit leegyszerűsíthetünk arra, hogy adott alternatívára vagy *igen*-nel vagy *nem*-mel szavaz-(hat)tak a szavazók. Ekkor viszont az egy jelöltre leadott igenlő szavazatok komplementer viszonyban vannak a jelöltre nem adott szavazatokkal, és ha ilyenkor minden szavazat esetében megcseréljük a szavazatok minőségét (*igen*-ről *nem*-re és fordítva), majd így végezzük el az összegzést, akkor ugyanazt a végeredményt kell kapjuk. Ha ez a feltétel teljesül, akkor valóban biztosíthatjuk az alternatívák semlegességét, hiszen ebben az esetben nem tudunk különbséget tenni a jelöltek között. Az alternatívák iránti semlegességet csak a döntetlen esetén kell feladni, amikor egy másodlagos, döntetlentörő szabály segítségével kell eldönteni, ki legyen a győztes. Ez a szabály azonban kivételszabálynak minősül, nem tartozik a csoportdöntési függvény magjába.

a neutralitásról Luis Ubeda [760]

May negyedik kritériumként bevezette a **pozitív érzékenység** tulajdonságát, amiről ő maga megjegyezte, hogy erősebb Arrow második feltételéhez képest, és amiről mi megjegyezhetjük, hogy ez az új kategória nagyon hasonlít a korábban már bemutatott monotonitás fogalmára (↬209.o.).

(AL.49) $$(\text{POSR}) \equiv \text{POSITIV_RESPONSIVENESS}(\mathbf{F}) \stackrel{\text{def}}{=} \forall m \forall n (\forall i \forall \sigma())$$

[15] Az anonimitás és a sorrend felcserélhetőségének összefüggőségét szépen érzékelteti az a videó, amelyen az látható, hogy egy zárthelyi dolgozatot késve leadó diák dolgozatát a tanár nem akarja elfogadni. Erre a diák megkérdezi a tanártól, ismeri-e őt, és a tagadó válasz után becsúsztatja a dolgozatát az asztal szélén álló többi dolgozat közé. Lásd: http://www.youtube.com/watch?v=wRpJM4-PoT8

pozitív érzékenység (HUN) – pozitív reakcióképesség (HUN) – positive responsiveness (ENG)

May azt az interpretációt adta a fogalomra, hogy ha a csoportdöntés m alternatívára nézve kedvező (vagy legalább indifferens), és az egyéni szavazók közül valaki megváltoztatja a korábbi döntését n-ról m-re, akkor a csoportdöntési függvény továbbra is kedvező kell maradjon m-re nézve.

——

ez csak két-alternatívás helyzetek esetén értelemezhető?????

——-

Az aggregáló függvények fenti tulajdonságaira, valamint az **egyszerű/többségi többségi szavazás** (simple majority voting)[16] eljárásra támaszkodva May a következő tételt fogalmazta meg:

(AL.50) $$(|\mathbb{X}| = 2 \to \mathbf{F} = \mathbf{F}^{AP}) \leftrightarrow \text{DEC}(\mathbf{F}) \wedge \text{ANYM}(\mathbf{F}) \wedge \text{NEUT}(\mathbf{F}) \wedge \text{POSR}(\mathbf{F})$$

May tétel (HUN) – May theorem (ENG)

A két-alternatívás $\mathbf{F}$ csoportdöntési függvény akkor és csak akkor (abszolút) többségi szavazás, ha $\mathbf{F}$ ügydöntő, anonim, semleges és monoton.[17]

A May-tételben az ekvivalenciának mindkét iránya fontos. Egyfelől igaz, hogy ha egyszerű többségi szavazási eljárást alkalmazunk, akkor teljesítjük a fenti négy igazságossági feltételt, másfelől igaz az is, hogy ha a jelzett négy – a méltányosságérzetünket kielégítő – tulajdonság teljesül egy véleményaggregáló függvényre, akkor az egyszerű többségi eljárás lesz. A tétel jelentőségének és helyének érzékeltetéséhez hozzátartozik, hogy tudjuk: csak kétalternatívás közösségi döntési helyzetekben érvényes.

——————

A többségi szavazással hozható kapcsolatba a logrolling, a szavazatcsere jelensége, ami a többségi és kisebbségi véleményeket valló emberek csoportjai közti különös viszonyból fakad, a jelenséget azonban csak azután tudom bemutatni (↬226.o.), miután már bevezettem a stratégiai szavazás fogalmát.

A többség(i szavazat) zsarnoksága ....

——————-

Franz Dietrich ls Christian List: Majority voting on restricted domains [208]

Sen-Pattanaik tétel

A többségi döntésre vonatkozóan Amartya K. Sen és Prasanta K. Pattanaik foglalta össze, hogy milyen szükséges és elégséges feltételek mentén racionális választási üggvényt remélni.

Necessary and Sufficient Conditions for Rational Choice under Majority Decision [669]

[16] May egyszerű többségi szavazásnak nevezte a vizsgált módszert, ami kétalternatívás helyzetekben megegyezik az abszolút többségi eljárással, de más esetekben nem feltétlenül. Az egyszerű jelzővel utalni lehet a relatív többségi eljárásra is, ezért jobbnak gondolom itt (is) az abszolút jelző használatát.

[17] A tételt leíró formula első kondiconálisában megadott feltételre igaz a következő: $|\mathbb{X}| = 2 \leftrightarrow |\mathbb{P}| = 2$.

Pareto-elv kiterjesztése

Sen 5.fejezet
pozitív responsiveness vs.
a többségi szavazás és a Pareto-kiterjesztés egy mozzanatban különbözik

Fishburn tétel

Fishburn reprezentatív rendszert jellemezte, P. Young hivatkozik rá: The Theory of Social Choice, Princeton University Press, 1973

Young tétel

A többségi szavazás mellett gyakran alkalmazott aggregálási technika a pontozásos szavazás, a Borda-pontok módszere, úgyhogy adta magát a kérdés, hogy vajon milyen feltételek mellett alkalmazható ez a szavazási technika. J. Smith cikke volt az első 1973-ban, ami ezzel a kérdéssel foglalkozott [695], de (el)ismertebbé H.P. Young tétele vált, amit Young egy 1975-ös cikkében publikált [831]. Young a pontozásos módszer alkalmazhatósági feltételeire vonatkozóan újabb tulajdonságokat definiált, és ezekre támaszkodva mondta ki tételét. Két új tulajdonságot vezetett be, és további kettőt használt a korábban definiált tulajdonságok közül. A pontozásos módszer axiomatizálását Young egy korábbi cikkében már elvégezte [833], de ott még - részben – más tulajdonságokat használt. Mivel ekkor már közismert volt, hogy a pontozásos módszer ellentmond az irreleváns alternatívák függetlensége elvének, ezért Young javasolta elhagyni ezt a feltételt. Rögtön javasolt viszont bevezetni két új fogalmat, az egyik egy konzisztenciafeltételt, a másik egy folytonossági elvet fogalmazott meg.

(AL.51) $(\text{RIF}) \equiv \text{REINFORCED}(\mathbf{SDM}) \stackrel{\text{def}}{=}$

megerősített (HUN) – reinforced (ENG) – konzisztens (HUN) – consistent (ENG)

Ez a konzisztenciakritérium fontos, intuitívnak mondható elvárást fogalmaz meg. Ha a szavazók két körzetben (N_1 és N_2 csoportban), egymástól teljesen elkülönítve szavaznak ugyanazokra a jelöltekre, az N_1 csoportban a W_1, az N_2-ben W_2 a győztesek halmaza, és van legalább egy olyan jelölt, aki a győztesek mindkét csoportjában benne van, akkor a két szavazói csoport ($N_1 \cup N_2$) szavazatainak összesítése után a győztesek halmaza a két részgyőztes halmazának metszete lesz ($W_1 \cap W_2$). Peyton Young a 'konzisztens' terminust használta erre a tulajdonságra, Roger B. Myerson szerint H. Moulin volt az, aki kicserélte ezt a 'reinforcement' terminussal megtartva a fogalom eredeti jelentését [524], és a javaslatot Myerson is elfogadta, én pedig a 'reinforcement' terminusát más szófaji alakkal helyettesítettem.

A következő tulajdonság értelmezéséhez újra azt kell feltételeznünk, hogy két teljesen elkülönült csoportban zajlik a szavazás (N_1 és N_2). A két csoportban legyen A és B jelölt a nyertes külön-külön. Ilyen helyzetben fel lehet tenni a kérdést, hogy van-e olyan feltétel, amely mentén ki lehet hozni győztesnek az egyik jelöltet a másik ellenében, ha a két körzetet (csoportot) egyesítjük. Erre az egyik igenlő válasz az lehet, hogy elvárjuk az archimédeszi tulajdonság érvényesülését, amit a következőképpen definiálhatunk.

(AL.52) $(\text{CONT}) \equiv \text{CONTINUOUS}(\mathbf{SDM}) \stackrel{\text{def}}{=}$

folytonos (HUN) – continuous (ENG) – archimédeszi (HUN) – Archimedean (ENG) – ellenállhatatlan többség (HUN) – overwhelming majority (ENG)

Az újonnan definiált archimédeszi tulajdonság akkor áll fenn, ha az egyik csoport részgyőztesét úgy hozhatjuk ki az egyesítés utáni győztesnek, hogy egy kellően nagy számmal beszorozzuk a csoport létszámát, és ezáltal a győztesre leadott szavazatok számát, amivel már le lehet győzni a másik csoport részgyőztesét. Ugyanezt a minőséget fejezi ki az archimédeszi preferencia fogalma is, amit később fogok elemezni (↬259.o.). Az 'ellenállhatatlan többség' kifejezést Roger B. Myerson javasolta,[18] mint ami tartalmilag legpontosabb fejezi ki a szóban forgó tulajdonságot [524].

(AL.53) $(\text{SCR}) \equiv \text{SCORING_RULE} \stackrel{\text{def}}{=}$

pontozásos szavazás (HUN) – scoring rule (ENG)

(AL.54) $(\text{SCR}) \leftrightarrow (\text{ANYM}) \wedge (\text{NEUT}) \wedge (\text{RIF}) \wedge (\text{CONT})$

Young tétel (HUN) – Young theorem (ENG)

A social decision method satisfies anonymity, neutrality, reinforcement and continuity if and only if the method is a scoring rule.

---

a reinforcement tulajdonság biztosítja, hogy a szavazási módszer esetén nem érvényesül a többkörzetes paradoxon (multiple district paradox).

ezt a paradoxont meg kell korábban említeni

IMPOSSIBILITY THEOREMS IN THE ARROVIAN FRAMEWORK DONALD E. CAMPBELL: 36- Relaxing the Pareto criterion 52 Relaxing transitivity 57 Relaxing the domain condition 64 Relaxing independence of irrelevant alternatives 70 Modifications of the Arrovian framework 75

98.o. The main condition used by Arrow was called Independence of Irrelevant Alternatives. This condition pre-defines the 'local' treatment of alternatives (or pairs of alternatives, or sets of alternatives, etc.) in voting procedures. The counterpart of his condition is called below a 'Locality' condition.

DOMAIN RESTRICTIONS* WULF GAERTNER, 132.o.

CLAUDE d'ASPREMONT and LOUIS GEVERS, SOCIAL WELFARE FUNCTIONALS AND INTERPERSONAL COMPARABILITY, 461.o-tól

invariance axioms pareto indifference, independence, neutrality separability

Pattanaik 366 social ranking rule social decision rule

Peleg, 398. voting paradox 402-404, veto player, dictator, choice problem 404-405.o. voting procedure, social choice correspondence - SCC anonymous, neutral, monotonic, has the SPA property (strong positive association . stronger concept as monotonicity)

[18] A tulajdonságokat jelző terminusokra vonatkozóan nincs megegyezés, hogy azok milyen szófaji alakban szerepeljenek. Az egész könyvben igyekszem melléknév szófajú terminusokkal hivatkozni a tulajdonságokra, de ez nem mindig sikerül. Myerson javaslatánál is ez a helyzet: sem magyarul, sem angolul nem találtam értelmesen használható melléknévi alakot.

„A következő, összoptimum függvényekről szóló tétel kiutat jelenthet a lehetetlenségi tételek köréből abban a speciális esetben, amikor a megengedett halmazok közös része nemüres, azaz létezik egy olyan alternatíva (nevezhetjük status quo-nak), amely minden helyzetben – mondjuk döntésképtelenség esetében – adoptálható.” [468, 26]

Allan Gibbard, Aanund Hylland és John A. Weymark [290]

(AL.55) $|A| \geq 3 \wedge |V| \geq 2 \wedge S \subset \wp P \wedge \bigcap S \neq \varnothing \wedge \#S > 1 \rightarrow$

Gibbard-Hylland-Weymark tétel (HUN) –

Legyen jAj ? 3, jV j ? 2, S 1?2 P(A). Tegyük fel, hogy S =6 ;, és tegyük fel, hogy S > 1. Tetszőleges D 1?2 WL(A)m esetén található olyan Arrow-típusú összoptimum függvény, mely teljesíti a Pareto-feltételt és mégsem diktatórikus.

4.2.6 No-show paradox

A monotonitás feltétele magától értetődőnek tűnik, hiszen ha egy alternatívára több szavazat érkezik, akkor azt érezzük intuitívnek, hogy ez a többletszavazat ne rontson az alternatíva helyzetén. Mégsem mindig ez a helyzet. Vannak olyan összegzőeljárások, amelyek megsértik a monotonitási tulajdonságot. Fishburn és Brams cikkéből idézhetünk ide egy példát szemléltetésképpen [256].

Egy kisvárosban három jelölt indul a polgámesteri címért: Mrs. Bitt, Mr. Huff és dr. Wogg. A szavazás során a Hare-féle preferenciális szavazás eljárását alkalmazzák. Ennek a lényege röviden annyi, hogy ha az első preferenciaként megadott jelöltek közül valaki megszerzi a szavazatok felét, akkor ő nyer. Ha nem ez a helyzet, akkor az első helyes szavazatok alapján kihúzzák a legkevesebb szavazatot szerző jelöltet, és a többi alternatívára szűkítve folytatják tovább az eljárást – az első fordulóhoz hasonló logika mentén.[19] A szavazás során mindenkinek meg kell adnia a szavazólapon, hogy milyen sorrendben támogatja a három jelöltet. A szavazás napján a Smith házaspár kocsija lerobban útban a szavazóhelység fele, és lekésik a szavazást. Mindkettejük preferenciája megegyezett az alábbi minta szerint: Bitt > Huff > Wogg, tehát a szavazólapra a BHW sorrendet írtak volna. A másnapi újságban megjelent a hír, miszerint a választást Mr. Huff nyerte. Közölték a szavazási eredményeket is.

szavazatok	sorrendek	H > W	W > H
417	BHW	417	0
82	BWH	0	82
143	HBW	143	0
357	HWB	357	0
285	WBH	0	285
324	WHB	0	324
1608		917	691

Az újságbeszámolóból kiderült az is, hogy Mrs. Bitt (a Smith házaspár kedvence) esett ki az első forduló után, mivel az első helyes szavazatok az alábbi módon alakultak:
Mivel az első számlálás után Mrs. Bitt esett ki, a bent maradt két jelöltre már könnyen ki lehetett számolni, hogy kire mennyi szavazat jutott a második számlálási körben (ezek

[19] A módszer leírása itt olvasható magyarul: [494].

Bitt	499 (417 + 82)
Huff	500 (143 + 357)
Wogg	609 (285 + 324)

a számok vannak a táblázat utolsó két oszlopában), és ezek alapján jött ki dr. Wogg győzelme. Ezt látva a – korábban aktuáriusként dolgozó – Mrs. Smith kiszámolta, hogy mi lett volna, ha ők is időben odaérnek, és mindketten leadják a Mrs. Bitt számára kedvező szavazataikat. Ez a plusz két szavazat ugyanis megfordította volna Mrs. Bitt és Mr. Huff közti első helyes szavazatok sorrendjét, és ebben az esetben utóbbinak kellett volna kiesni az első fordulós összegzés után. Ekkor persze a megmaradt két jelöltre adott szavazatok is másként oszlottak volna el.

szavazatok	sorrendek	B > W	W > B
419	BHW	419	0
82	BWH	82	0
143	HBW	143	0
357	HWB	0	357
285	WBH	0	285
324	WHB	0	324
1610		644	966

Az újraszámolásból az derült ki, hogy a Mrs. Smithék számára “adott” plusz két szavazat nemhogy javított volna a helyzetükön, de inkább rontott, hiszen a legkevésbé kedvelt jelöltet segítette volna a szavazatuk. Ezt a paradox helyzetet Fishburn és Brams ‘No-Show paradoxon’-nak nevezték el.

Az okos és kitartó Mrs. Smith azonban további felfedezéseket is tett.[20] Észrevette azt is, hogy ha abban a hipotetikus helyzetben, amikor ők is szavaztak volna, ketten másként adták volna le a szavazatukat azok közül, akik a valóságban a BWH sorrendet jelölték be (82-en), és ehelyett a WBH-re szavaznak, akkor a táblázat kicsit, a végeredmény nagyon megváltozott volna. Ekkor ugyanis az első számlálás után megint csak Mrs. Bitt esett volna ki, és a bent maradt két jelölt, Mr. Huff és dr. Wogg között nagyjából az első esethez hasonló módon lakult volna a második körös eloszlás.

szavazatok	sorrendek	H > W	W > H
419	BHW	419	0
80	BWH	0	80
143	HBW	143	0
357	HWB	357	0
287	WBH	0	287
324	WHB	0	324
1610		919	691

A képzelt eset lehetett volna persze teljesen reális is, így a példa azt mutatja, hogy ezen aggregáló eljárás esetében a monotonitás feltétele nem teljesült, hiszen a második esethez

[20] Fishburnék a cikkben leírnak még további két paradoxont, de ezeket itt nem mutatom be.

képest csak annyi változás történt a harmadik esetben, hogy a második eset nyertese a harmadik esetben "kapott" két plusz szavazatot, a többletszavatatok azonban nem javítottak, hanem rontottak a helyzetén.

———————-

A no-show paradoxon az olyan helyzeteket jellemzi, amelyekben a szavazók bizonyos csoportja számára előnyösebb lehet nem szavazni, mint a saját preferenciáikat kinyilvánítani.

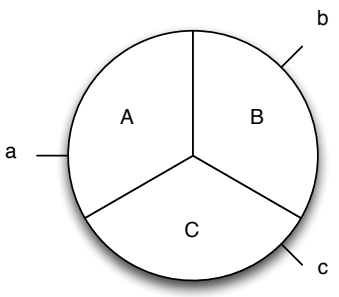
Moulin: [523]

4.2.7 Reprezentációs paradoxon

Az előző fejezetben tárgyalt problémához sok szepontból hasonlít a reprezentáció paradoxona, amit Thomas Schwartz írt le először 1985-ben [649]. Schwartz azt a közkeletű nézetet támadta meg a cikkében, amely szerint közösségi döntési helyzetekben valamilyen vélemény képviseletének sikeressége mindig arányos a véleményt képviselő reprezentációs erő nagyságával. Másként, kicsit konkrétabban mondva: ha egy szavazótestületben (például parlamentben) egy véleménycsoport (például egy párt) érvényesíteni szeretné a véleményét, akkor annál biztosabban teheti ezt meg, minél több szavazattal rendelkezik. Schwartz szerint nem mindig van ez így. Előfordulhat, hogy a kevesebb szavazat, pontosabban az alacsonyabb reprezentációs szint előnyösebb kimenetet eredményezhet.

elmélet

Schwartz három példát is bemutatott, amik alapvetően abban különböztek egymástól, hogy milyen magatartást (kooperatív, rafinált és naív hozzáállást) feltételezett a szavazóktól. Itt csak a naív, őszinte szavazói viselkedéshez tartozó esetet ismertetem.



Ebben a történetben három körzethez (*A*-hoz, *B*-hez, *C*-hez) tartozó 100 képviselőnek kell dönteni abban a kérdésben, hogy építsenek-e új haditengerészeti bázist vagy ne, és ha igen, akkor hova. A döntés pillanatában már van egy bázis, az *a* pontban, és lehetséges döntés az, hogy minden marad a régiben, nem épül új hadikikötő (status quo). A három körzet és a három potenciális bázis helyét, ezek egymáshoz való viszonyát mutatja az ábra. Az *A* csoportba 23, a *B*-be 28, a *C*-be 49 mandátum (szavazási lehetőség) tartozik. A három körzet képviselői rendre a következő preferenciákkal rendelkeznek.

körzet	szavazatszám	preferencia
A	23	*a b c*
B	28	*b c a*
C	49	*c a b*

Hunno Nurmi szerint ez a paradoxon nagyon közel áll a no-show paradoxonhoz.

4.2.8 Szavazatmanipuláció

A társadalmi választások kimenetét elsődlegesen az egyéni szavazók által leadott szavazatok határozzák meg. Már az is egy kérdés, hogy miért úgy szavaznak, ahogy szavaznak,

és tudjuk jól, hogy a végeredményt azáltal is lehet módosítani, megváltoztatni, hogy az egyéni szavazók véleményére próbálnak meg hatást gyakorolni még a szavazás előtt, esetleg a szavazás során. Az előzetes véleménybefolyásolás, a lojalitáskialakítás és lojalitásfenntartás, a kampányolás indirekt technikái többnyire legális módokon hatnak, míg a szavazás közben alkalmazott direkt befolyásolási technikák (például nyílt szavazás során kifejtett nyomásgyakorlás) döntően illegálisak, de ettől még létez(het)nek. Bármennyire is fontosak ezek a befolyásolási jelenségek, velük itt – egyetlen kivételtől eltekintve – nem foglalkozom. Ezen a ponton azt mondom, vegyük tudomásul, hogy a szavazók valamilyen módon, korábban kialakult véleménnyel, preferenciarendezéssel érkeznek el a szavazás pillanatához. Ideális esetben ennek megfelelően adják le a szavazatukat, amikoris azt mondhatjuk, hogy őszintén szavaztak, vagyis mindenféle stratégiai megfontolás nélkül pontosan a saját preferenciáiknak megfelelően adják le a szavazataikat. Elképzelhető azonban az is, hogy mégsem ez a helyzet, vagyis vannak szavazók, akik valamiért az elsődleges/eredeti preferenciáktól eltérnek a szavazás során. Érdemes megfontolni, hogyan is nevezzük az ilyen helyzeteket, mert bár sokféle kifejezés verseng e tárgyban, ezek némelyike rossz konnotációt hozna be a diskurzusba. A stratégiai szavazás vagy taktikai szavazás még értéksemleges, a hazug szavazó (liar) vagy manipulatív szavazó (manipulation of voter) kifejezéseknek már egyértelműen negatív felhangja van. Akármelyik terminust is használjuk, a lényeg persze nem változik. Azt is látni kell azonban, hogy az a definíciós mozzanat, miszerint minden szavazás stratégiai, amennyiben nem a tényleges preferenciarendezése alapján szavaz valaki, nagyon hétköznapi, nagyon gyakori jelenségekre is ráhúzható. Eszerint ugyanis már azt is stratégiai szavazásnak kell tartanunk, ha egy politikai választáson a két nagy párt és a az egy kis párt jelöltjei közül egy szavazó inkább valamelyik nagy párt politikusára szavaz, noha szíve szerint a kis párt politikáját követné. Ha viszont tudja, hogy a két nagy versengése mellett a kicsinek esélye sincs megnyerni a választást, és így a kis pártra adott szavazata "elveszne", nem értékesülne, és ezért inkább a neki szimpatikusabb nagy párt képviselőjére teszi le a voksát, akkor már stratégiai, hazug, manipulatív szavazónak minősülne. Lehet azon vitatkozni, ki mennyire érzi kritizálhatónak, hazugnak, manipulatívnak az ilyen magatartást, de itt és most nem az ilyen helyzetek gyakorlati értékelését, hanem csak az elméleti leírását szeretném elvégezni.[21]

A **stratégiai szavazás** során tehát a szavazó azért ad le az igazi preferenciájához képest eltérő, tehát "színlelt" szavazatot, mert azt előnyösebbnek véli az eljárás végeredményét tekintve. Nézzük meg egy egyszerű példát erre a jelenségre [239]. Legyen három alternatíva (U, B, L) és a szavazólapokon lehesen mindegyik alternatívára szavazni úgy, hogy a teljes rendezés során az első alternatíva kapjon 2 pontot, a második 1 pontot, a harmadik ne kapjon semmit (ez a Borda-szavazásnak felel meg). A táblázat baloldali része mutatja a tiszta szavazás eredményét, amikor mindenki a valósi preferenciáinak megfelelően szavazott. A jobb oldalon látható a stratégiai (hazug) szavazás eredménye, ami olyan módon állt elő, hogy mindhárom szavazótábor úgy szavazott, hogy megcserélte a második és

[21] A manipulálhatóság problémáját Allan Gibbard elemezte először formális módon a lehetetlenségi tételről szóló cikkében [289]. Ebben az 1973-as tanulmányában Gibbard a kérdés előzményeként Michael Dummet és Robin Farquharson 1961-es cikkében megfogalmazott sejtésére hivatkozott: „This would be to assume that every voter votes 'sincerely,' whereas it seems unlikely that there is any voting procedure in which it can never be advantageous for any voter to vote 'strategically,' i.e., non-sincerely." [223, 34]. Gibbard másik előzményként említi William Vickrey egy 1960-as cikkét is [788], amelybenVickrey egyértelműen a preferenciák stratégiai álcázásáról (strategic misrepresentations of preferences) értekezik.

harmadik helyen álló alternatíváit.

	igaz szavazás			hazug szavazás		
	LBU	BLU	UBL	LUB	BUL	ULB
szavazat	104	98	7	104	98	7
pontszám	306	307	14	208	203	216

A táblázatban látható eredmények magukért beszélnek. A stratégiai szavazás következményeként a legkisebb saját táborral rendelkező (U) alternatíva győzött. Nagyon életszerű magyarázatot lehet adni ennek a jelenségnek. Az első két, nagy táborral rendelkező alternatíva sokkal közelebb állt egymáshoz a harmadikhoz képest, de a győzelem reménye miatt mindkét tábor tagjai úgy gondolkodhattak a stratégiai szavazás során, hogy úgy gyöngíthetik a riválisukat, ha az esélytelennek gondolt (és mindkettőjüktől távol eső) alternatívára szavaznak második helyen, így a konkurrencia nem kapja meg a második helyekért járó egy-egy pontot. Ám amikor mindkét tábor tagjai fegyelmezetten így is tettek, akkor annak az lett a következménye, hogy a nevető harmadik – egyébként teljesen esélytelen – alternatíva győzedelmeskedett. Az is látszik a táblázatból, hogy ha csak páran lettek volna fegyelmezetlenek a nagyobb táborokból, és az igaz szavazataikat adták volna le, akkor bármelyik nagy győzhetett volna. A bemutatott példa láthatóan "mesterséges", de a stratégiai szavazásra sok valós példát is találhatnánk. Az ausztráliai parlamenti választásokon (ahol a Single Transferable Vote eljárást használják) 2004-ben egyértelműen kimutatható volt, hogy a zöld-, a demokrata- és munkáspárti szavazók a Family First konzervatív párt jelöltjére szavaztak, amit a szakértők a stratégiai szavazás jelenségével magyaráztak [167].

A stratégiai szavazás első – eddig tárgyalt – szintjén a szavazók "magukban" szavaznak, egyedül változtatják meg eredeti preferenciáikat, nincs stratégiai kapcsolat közöttük. A szavazatok megváltoztatásának azonban vannak más lehetőségei is. A szavazók (bizonyos csoportjai) megállapodhatnak abban, hogy kölcsönös előnyt ígérő helyzetekben együttesen megváltoztatják a szavazataikat. Ezt a jelenséget nevezik szavazatcserének, amikor valamilyen "politikai koalíció" alapján szavaznak stratégiailag az emberek.

4.2.9 Szavazatcsere

A **szavazatcsere** (logrolling) vagy **szavazatkereskedelem** (voting trade) jelenségét ugyan a stratégiai szavazás körébe kell sorolnunk, mégis érdemes önálló fejezetben tárgyalni. A jelenségnek vannak ugyanis olyan jellemzői, amelyek túlmutatnak a szavazás, azon belül is a stratégiai szavazás intézményén. Minél inkább a szavazatcsere interakciós vonásaira koncentrálunk, vagyis minél inkább figyelembe vesszük a jelenségben rejlő dinamikát, annál inkább távolodunk el az egyszerű szavazás jelenségétől, és közeledünk másfajta, dinamikus módon kezelhető társadalmi kapcsolatokhoz.

Na, de nézzük meg, mit is jelent a logrolling jelensége. A megértéséhez abból a megállapításból kell kiindulnunk, hogy a (többségi) szavazás nem veszi figyelembe az egyéni preferenciák intenzitásában meglevő különbségeket [129]. Egy adott ügy iránt nagyon elkötelezett ember egyetlen támogató szavazata ugyanannyit ér, mint egy majdnem teljesen közömbös szavazóé, aki valamiért ellenző szavazatot ad le. Amikor ilyen érzelmi intenzitáskülönbségek tapasztalhatók a szavazók körében, akkor lehet értelme a szavazatcsere intézményének. Ha ugyanis az erősen elkötelezett támogató hajlandó valamilyen

előnyt kínálni az addig a másik oldalon álló fél támogató szavazatának megszerzéséért, és az az egy szavazat megváltoztatása épp elég ahhoz, hogy a szavazás végeredménye megforduljon, akkor mindkét résztvevő jobban járhat a logrolling révén. Az elkötelezett ember nyereménye a szavazási eredmény megváltozása, a közömbös szavazó pedig a szavazatáért kapott előny miatt jobban járhat, mint anélkül. A logrolling alkalmazása így Pareto-optimális helyzetet teremthet. Nézzünk meg egy példát [382, 163]. Legyen két kérdés (ϕ és ψ), amiben három szavazónak (*A*-nak, *B*-nek és *C*-nek) kell döntenie, és a két kérdés elfogadásához rendeljük a következő nyereményeket (illetve veszteségeket). Ezekből a két kérdés elfogadása esetén származó nettó társadalmi nyereséget is ki lehet számolni.

	ϕ esetén	ψ esetén		ϕ-re	ψ-re
A nyereménye	-100	-100	*A* szavazata	nem	nem
B nyereménye	-200	500	*B* szavazata	nem	igen
C nyereménye	600	-300	*C* szavazata	igen	nem
nettó társadalmi nyereség	300	100	többségi szavazat	nem	nem

Ha nincs logrolling, akkor a többségi szavazat alapján mindkét kérdésben elutasító döntés születik, még akkor is, ha mindkét kérdésnek van egy-egy lelkes támogatója, akik nagy nyereményhez jutnának valamelyik kérdés elfogadásával. Mivel a többségi vélemény mindkét esetben elutasító lenne, senki sem nyerne és senki sem vesztene a szavazáson. Ha viszont a szavazás végeredményének mgváltoztatására erős motivációval rendelkező két szavazó, *B* és *C* megegyezne egymással, hogy saját ellenző szavazataikat kölcsönösen megváltoztatják a másik kedvéért, akkor megváltozna a végeredmény, mindkét kérdés többségi véleménnyé válhatna. Egyénileg mindketten, *A* is, *B* is, jobban járnának, hiszen a két kérdés együtteséből pozitív nyereséggel jöhetnének ki, sőt, még realizálódhatna a nettó társadalmi nyereség is, vagyis a csoport egésze is nyerne az ügyön. Mindez persze azon az áron valósulhatna meg, hogy egy szavazónak (*A*-nak) így veszteséget kellene elszenvednie. A tanulság tehát az, hogy ha nem kerülnek kapcsolatba, csak saját egyéni szavazataikra "támaszkodnak" az emberek, akkor elutasítóak lesznek olyan kérdésekben is, ahol egyébként a többségnek és a közösség egészének is jobb lenne ellentétes módon dönteni. Ha mód van a szavazatcserére, akkor ez az optimálisabb helyzet elérhető.

A logrolling jelensége az egyéni preferenciák intenzitáskülönbségeiből táplálkozik. Ezt a technológiát ezért csak akkor lehet eredményesen alkalmazni, ha valamely véleménykisebbségnek jóval erősebb a véleményintenzitása a többséghez képest [129, 134]. A logrolling jelensége kis létszámú és nyilvános döntési eljárással dolgozó testületekben fordulhat elő könnyebben, ezt nevezik explicit logrollingnak. A kis létszám a szavazat értékessége miatt fontos (ha nagyon sok szavazó van, akkor egy szavazat nem sokat ér), a nyilvános szavazás pedig az ellenőrizhetőséget "biztosíthatja" (titkos szavazás esetén nem lehetne tudni, hogy megvalósult-e a szavazatcsere). Buchanan és Tullock szerint létezik szavazatcsere nagylétszámú szavazás során is, amit ők indirekt (máskor implicit) logrollingnak neveztek el. Olyan eseteket soroltak ide, amikor a kampányszervezők úgy állítják össze a szavazási napirendet (pártjelölteket), hogy legyen olyan kérdés (jelölt) benne, amely (aki) iránt erősen elkötelezett szavazókat lehet odavonzani az adott párthoz, csoporthoz még akkor is, ha a csomag más elemei esetleg már nem vonzóak az elkötelezett szavazók számára.

Az állami kiadások folyamatos emelkedését nyilván sok összetevőre bontva magyarázhatjuk csak meg, de a lehetséges magyarázó tényezők között figyelembe vehetjük azt is, hogy a logrolling jelensége sok esetben növeli a kollektív költségeket. Ezt a felismerést már 1959-ben megfogalmazta Gordon Tullock egy cikkében [754]. Ezt a tézist elmondhatjuk más formában is: a logrolling révén alkalmi többségi koalícióba lépve másokra (a kisebbségre) lehet terhelni a többség számára fontos célok költségeinek egy – akár jelentős – részét. Egyszerű példával lehet szemléltetni mindezt. Legyen egy kisvárosunk, ahol három út (A, B, C) felújítására lenne szükség, de a város költségvetésében csak kettőre van meg a fedezet (természetesen az előző évben befizetett adókból). Ekkor szavazást lehet kiírni arról, melyik kettő legyen felújítva a három közül. Az egyszerűség kedvéért tegyük fel, hogy a három út három (egyenlő lélekszámú) városrészben van, és az ott élők mind a saját városrészükben levő út rendbetételét szeretnék elérni. Ha nincs előzetes "megegyezés" az érintettek egy része között, akkor minden városrész lakói *igen*-t mondanának saját útjuk és *nem*-et a többi út felújítására, amiből a többségi vélemény mindhárom esetben az elutasítás lenne. Ha azonban két városrész lakói megegyeznének a szavazatcserében, akkor egymás útjaira is támogatólag szavazva már többségi véleményt szerezhetnének saját útjaik felújítására.

	szavazatcsere nélkül			szavazatcserével		
	A-i út	*B-i út*	*C-i út*	*A-i út*	*B-i út*	*C-i út*
A városrész lakóinak szavazata	igen	nem	nem	igen	igen	nem
B városrész lakóinak szavazata	nem	igen	nem	igen	igen	nem
C városrész lakóinak szavazata	nem	nem	igen	nem	nem	igen
többségi szavazat	nem	nem	nem	igen	igen	nem

A szavazatcsere eredménye az, hogy két út megépülhet a rendelkezésre álló pénzből. Ezt az előnyt úgy nyerhették el a két városrész lakói, hogy az általuk befizetett adóhoz még hozzájárult a harmadik városrészből befolyt adók összege is, vagyis az nyertesek útfelújítását részben a vesztesek adóiból finanszírozták. Ez az időleges koalícióképzési lehetőség gyakran és könnyen egy olyan spirált indíthat be bárhol a világon, ami – a koalícióba lépő szereplőket állandóan cserélgetve – folyamatosan költi a mások adóját, amit a következő körben más ügyekben "kell" újabb adóelvonásokkal kompenzálni. Ez nyilvánvalóan arra készteti az ilyen koalícióképzési, szavazatkereskedelmi folyamatokat irányító politikusokat, hogy állandóan növeljék az adóbevételeket, hiszen csak ilyen módon lehet új és új időleges koalíciókat képezni.

A logrolling jelenségére mutat be egy konkrét politikai példát William H. Riker az USA történetéből [606]. Magyar példaként hivatkozhatunk az 1990-es MDF-SZDSZ paktumra, amelyben a stabil kormányozhatóság érdekében megegyezett egymással a két legnagyobb kormány-, illetve ellenzéki párt a kétharmados törvények körének kiterjesztéséről, valamint a köztársasági elnök személyéről.

A szavazatcsere, különösen az explicit logrolling intézménye annyiban speciális a stratégiai szavazás jelenségéhez képest, hogy az érintettek itt egyeztetett módon, tudatosan, kölcsönös engedményeket téve térnek el a szavazataik kinyilvánítása során a valós véleményüktől. Ezt igazán jól csak kis létszámú, a nyilvánosságtól elzárt módon egyezkedni képes szavazótestületekben lehet megtenni, másfelől ez gyakorlati eredményét tekintve már sokkal inkább a kompromisszumtestületek működésére hasonlít (↬191.o.). A szavazatcsere a kölcsönös haszon reményében tett kölcsönös engedmények eredménye.

Minél inkább a kompromisszum megkeresésének, megteremtésének technikái válnak dominánssá egy adott társadalmi kapcsolatban, annál kevésbé beszélhetünk stratégiai szavazási problémáról, és annál inkább csúszunk át az alkudozáselmélet, a koalícióelmélet, a politikai hatalomtechnikákat leíró elméletek világába.

Megemlítem itt, de egyáltalán nem elemzem a korrupció kérdését is, ami – legalább részben – a stratégiai szavazás területére tartozik. A direkt szavazatvásárlás vagyis a megvesztegetés fontos esetei eredményüket tekintve azt jelentik, hogy fontos döntéshozó emberek valamely szavazás során a kapott előnyökért cserébe megváltoztatják a véleményüket. Ezen az alapon van ebben a fejezetben lenne a helye ennek a jelenségnek is, de ahogy a kompromisszumkötés világát sem elemzem itt, úgy a korrupció egész jelenségét sem tárgyalom. Ha jobban belegondolunk a logrolling és adózás kapcsolatára hozott korábbi példába, akkor észrevehetjük, hogy ide lehetne még kötni a járadékvadászat egész kérdéskörét is, de sem ennek bővebb kifejtését, sem a témakör elemzését nem vállalom itt.

Bernholz [72] [73] [74]
voting trade [417] [418] [485]
Robert Wilson axiomatikus leírása [827]
[407]

4.2.10 Csalásbiztos eljárás

Miután "kiderült,", hogy egyfelől a tényleges társadalmi gyakorlatban létezik a stratégiai szavazás intézménye, másfelől az erre reflektáló elméletekben is meg lehet ragadni, jellemezni lehet a jelenséget, akkor felmerül a kérdés, vajon vannak-e olyan társadalomtechnikák, amelyek segítségével meg lehet akadályozni vagy legalább meg lehet nehezíteni azt, hogy emberek "csaljanak", hazudjanak a preferenciáinkat illetően. Van-e – szavazatmanipulációs szempontból vett – csalásbiztos aggregálási technika? A válasz az, hogy bizonyos helyzetekben elképzelhető.

Nézzünk meg egy példát Csontos László egyik cikkéből [180], amelyben a következő történet szerepelt a Clarke-adó mint csalásbiztos technika hasznával kapcsolatban. Képzeljünk el egy szigeten elterülő kisvárost, ahol egyszercsak felmerül egy híd építésének az ötlete a sziget és a szárazföld között. Ilyenkor hihetőnek tűnik azt is feltételezni, hogy a városka lakói között lesznek, akik ellenzik, és lesznek, akik támogatják az ötletet. Ha nem akarják, hogy egy ilyen fontos kérdésben a többségi akarat alapján döntsenek, mivel az a vesztes kisebbségre túl nagy "terheket" róna, ezért más döntési technikát kell keresni. A következő megfontolásból lehet kiindulni. Minél erősebben támogatja vagy ellenzi valaki a tervet, annál inkább hajlandó lehet pénzzel is támogatni a híd felépítését, ha támogatja a tervet, vagy – épp ellenkezőleg – hajlandó lehet kárpótlást fizetni azért az ötlet támogatóinak, hogy ne épüljön meg a híd. Ezt a – pro és kontra – fizetési hajlandóságot kellene valahogyan figyelembe venni a döntés során. A probléma csak az, hogy egy egyszerű szavazás során nem könnyű biztosítani azt, hogy a szavazók megvallják a többiek számra ezeket az elköteleződéseiket. Az ellenzők igyekeznének úgy szavazni, hogy ne derüljön ki, hogy ők hajlandóak lennének kárpótlásra, míg a támogaták azt nemfednék fel szívesen, mennyire érné meg nekik az ötlet megvalósítása. Az egyik megoldása lehet ennek a problémának az ún. Clarke-adó kivetése. Ezt akkor érthetjük meg jól, ha egy táblázatban feltüntetjük, hogy az érintetteknek (háztartásoknak) milyen fizetési haj-

landóságaik vannak (amikor is feltételezhetjük, hogy ezek a felajánlások nyilvánosak és egyben igazak).

	fizetési hajlandóság ($-ban)		Clark-adó
háztartás	híd mellett	híd ellen	
1. háztartás	0	9000	0
2. háztartás	0	6000	0
3. háztartás	0	1000	0
4. háztartás	4000	0	0
5. háztartás	17000	0	12000
összes háztartás	21000	16000	0

Az tényleges egyéni preferenciákat az egyének természetesen nem fedik fel egy sima szavazás során. Ennek feltárását célozza meg a Clarke-adó alkalmazása, aminek lényeg az, hogy csak azokra a szavazókra vetik ki, akik – értékben is kifejezhető – szavazata döntő a szavazás végeredménye szempontjából. Ez úgy derül ki, hogy először is mindenkinek meg kell vallania a fizetési hajlandóságát, és ezek alapján ki lehet számolni, melyik opció mellett lennének hajlandóak több anyagi áldozatra a szavazók. A jelen példában a támogatók ajánlottak fel többet (21 ezer dollárt a 16 ezer ellenében). Ebben az esetben döntő szerepe van az ötödik háztartást képviselő szavazónak, hiszen ha ő nem ajánlana fel 16 ezer dollárt a híd felépítésére, akkor az építkezés ellenzői nyernének. A híd felépítésének kérdését tehát ő dönti el. A Clarke-adó mértéke úgy állapítják meg, hogy a választási lehetőségekhez rendelt összesített hozzájárulások különbségéből levonják a döntő szavazó felajánlását. Ez a jelen példában $|(21000-16000)-17000| = 12000$ dollár. Ennyit kellene tehát fizetnie az 5. szavazónak, miközben mást nem terhelnének adóval. A fizetendő adó összege kevesebb, mint amennyit a legelkötelezettebb támogató hajlandó lenne fizetni a tényleges preferenciái alapján, de a nagy kérdés itt az, hogy miért vallaná be ezt, ha a végén csak neki kell fizetnie adót. Nos, ha csak pár ezer dollárral kevesebbet jelölne meg (bárhol 12 és 17 ezer dollár között), akkor a Clarke-adó ugyanannyi maradna és neki kellene továbbra is megfizetnie. Ha például csak 13 ezret vallana be, akkor a következőképpen számolnák ki a Clarke-adót: $|(17000-16000)-13000| = 12000$ dollár. Ez ugyanannyi, mint korábban volt. Amennyiben 12 ezer dollárt vagy annál kevesebbet vallana meg, akkor viszont megfordulna a szavazás kimenete, és a híd ellenzői nyernének (döntetlen esetén a státus quo vagyis a híd nélküli állapot maradna fent). Ebben az esetben a döntő szavazó elesne a híd felépítéséből származó előnyöktől, ezért neki ekkor sem érné meg erre a szintre lemenni.

Belátható, hogy a másik oldalon, az ellenzők között sem éri meg a preferenciákat hamisan megadni. Amikor az ellenzők feltüntetik a fizetési hajlandóságaikat, akkor ezzel azt jelzik, hogy mennyi veszteséget jelentne számukra a híd felépítése, és ezt mekkora összeggel lennének hajlandóak megakadályozni. Ha az 1. szavazó azzal akarná megakadályozni a híd felépítését, hogy a ténylegesnél nagyobb veszteséget jelezne előre (amivel nagyobb fel-nem-építési kárpótlást vállalna magára), mondjuk 14 ezer dollárt, akkor megnyerhetnék a választást (21-21 ezer dolláros döntetlen mellett a status quo maradna fent), de neki magának ekkor 14 ezer dollárnyi Clarke-adót kéne fizetnie ($|(21000-21000)-14000|$) dollár. A fizetendő adó viszont 5 ezer dollárral több lenne, mint a tényleges fizetési hajlandósága, emiatt értelmetlen lennne hazudnia.

A Clarke-adó tehát mindkét oldal résztvevőit arra ösztönzi, hogy a valós fizetési ha-

jlandóságaikat vallják meg, és ezáltal a tényleges egyéni preferenciák (tehervállalási hajlandóságok) alapján szülessen meg a közösségi döntés. Ez egyfajta igazságosság megvalósulását képes támogatni azzal, hogy megakadályozza a valós preferenciák eltitkolásának lehetőségét. A módszernek azonban gyakorlati szempontokból tekintve komoly hiányosságai vannak. A preferenciák feltárása fontos lehet ugyan a kollektív közérzet szempontjából, de a preferenciák, a vélt nyereségek vagy veszteségek kinyilvánításának semmi köze nincs az utánuk következő valós folyamatokhoz (itt a híd felépítéséhez). A módszer alapján nem lehet semmit sem kezdeni azzal, ha nem elég a pénz a tervek megvalósításához, de arra sem ad semmi útmutatót, hogy mit kell tenni abban az esetben, ha netán maradna a projekt végén a beszedett Clarke-adóból. Ez a megoldás nem kezeli az érintettek preferenciái, vágyai, tudásai és a megvalósult folyamatok valós költségei közti diszkrepanciát.

Edward H. Clarke 1971 [160] Edward H. Clarke 1977 [161]

4.2.11 Gibbard-Satterthwaite tétel

Richard Zeckhauser nem-manipulálhatósági feltétel (non-manipulability condition) [841]

determinisztikus vagy sztochasztikus

egyedi (unique) vagy nem egyedi (non-unique)

taktikai szavazás (tactical voting)

(AL.56) $(\text{MAN}) \equiv \text{MANIPULABLE}(\mathbf{F}) \stackrel{\text{def}}{=}$

manipulálható (HUN) – manipulable (ENG)

(AL.57) $(\text{STRP}) \equiv \text{STRATEGY-PROOF}(\mathbf{F}) \stackrel{\text{def}}{=} \neg\text{MANIPULABLE}(\mathbf{F})$

színlelésbiztos (HUN) – nem manipulálható (HUN) – hazugságbiztos (HUN) – strategy-proof (ENG) – non-manipulable (ENG)

A voting rule is a mapping f from in U to A:3 A voting rule f is manipulable precisely when there is an individual i 2 N; preferences u0i 2 U, and a preference profile u 2 U; such that ui(f(u0i;u?i)) > ui(f(u)):

A. Feldman, nem-manipulálhatóság [248]

Kelly 1987 Ellenmanipulálhatóság Mj 163.

a közösségi preferenciát képző eljárás a preferenciarendezés teljes ismeretén alapul. A teljes preferenciarendezés feltételezése helyett gyakran meg kell elégendünk (elég lehet) a legjobb elemeket kiválasztani az alternatívák halmazából.

A (szavazat)manipulációval kapcsolatos leghíresebb tételt ketten is kimondták nagyjából egyidőben, Allan Gibbard [289], illetve Mark Allen Satterthwaite [639], ezért kettejük nevét összekötve hivatkoznak a “közös” Gibbard-Satterthwaite tételre.

Gibbard

nincs szerepe a véletlennek

Gibbard azzal indítja az elemzését, hogy a preferenciaaggregálási függvények (vagy ahogy Gibbard nevezi őket: a szavazási sémák) a – játékelméleti értelemben vett – játékok speciális eseteit jelentik. Ahogyan megadunk egy játékot, ugyanúgy kell leírni egy szavazási sémát is. Ami a játékokban a stratégiát jelenti, az a szavazásban megfelel egy adott preferenciarendezésnek (egy leadott szavazatnak). Ez a hasonlóság azonban még kevés ahhoz, hogy megragadhassuk a stratégiai szavazás jelenségét. Az igazi, illetve a hazug

szavazatok “kezeléséhez” még tudnunk kell azt is, hogy az adott játékos (szavazó) melyik preferenciarendezése az “igazi”, az őszinte és melyik nem (tehát stratégiai). Ezt megadhatjuk egy függvénnyel, ami minden szavazóhoz hozzárendeli a lehetséges preferenciarendezések halmazából az egyetlen őszinte preferenciarendezést. A manipulálhatóság ezek alapján viszont már nem a játékok, hanem csak a szavazások sajátosságaként értelmezhető.

a szavazási séma nem enged meg döntetlent (hiszen függvény), az Arrow-féle konstitúció viszont megengedi ezt.

GS: ha az alternatívák száma több kettőnél, akkor a nem diktatórikus, racionális választás ki van téve a stratégiai manipuláció lehetőségének ugyanezt megfordítva: csak a diktatórikus közösségi választási szabály

manipulálható (↬231.o.)

(AL.58) $(\text{STRATEGIC_VOTING}) \stackrel{\text{def}}{=}$

színlelt szavazás (HUN) – stratégiai szavazás (HUN) – taktikai szavazás (HUN) – strategic voting (ENG) – tactical voting (ENG) – sophisticated voting (ENG) – insincere voting (ENG)

taktikázásbiztos (strategy-proof) értsd: stratégiai, nem őszinte, számító szavazással nem lehet befolyásolni a végeredményt csalásbiztos, csalásgátló, helyezkedésgátló, taktikaérzéketlen

(AL.59) $\big((\text{STRP}) \wedge (\text{UNI})\big) \rightarrow (\text{DIC})$

Gibbard-Satterthwaite tétel (HUN) – Gibbard-Satterthwaite theorem (ENG)

Gibbard-Satterthwaite tétele: Nincs olyan racionális döntési mechanizmus egy olyan legalább háromszemélyes helyzetben, amely egyszerre kielégítené az (UNI) és (STRP), valamint (NODIC) feltételeket.

ha azt akarjuk elérni, hogy a racionális döntés őszinte legyen (megfeleljen az (STRP) feltételnek), azaz mindenki a legpreferáltabb alternatíváját nyilvánítsa ki, akkor szükségképpen a racionális választás elitista jellegű lesz (lesz egy olyan diktátor, aki a döntési helyzetben belül ‘túl nagy’ hatalommal rendelkezik, és preferenciája automatikusan kollektív preferenciává válik).

Salvador Barberà és J.S. Kelly azon a ponton próbált single valued

Salvador Barberà alternatív bizonyítást adott a Gibbard-Satterthwaite tételre a pivotális szavazó fogalmára támaszkodva [48].

Barberà és Peleg 1990-es alternatív bizonyítása [54]

A korábbi bizonyítások támaszkodtak az univerzális értelmezési tartomány feltételére (↬209.o.), amit új bizonyításukban Barberàék elhagytak, viszont felvették a feltételek közé azt, hogy az egyéni preferenciák folytonosak. A feltételeknek e két együttes módosítását azzal indokolták, hogy bár sok esetben valóban tartható az univerzális tartomány elvárása, vagyis hogy a szavazási séma hozzáfér az összes lehetséges egyéni preferenciához, de vannak esetek, amikor – a végtelenség miatt – ez a feltétel nem teljesülhet. Amikor a a preferenciákat a valós számok többdimenziós terében folyamatos módon lehet kifejezni, és ilyenkor a kifejezhető preferenciák számossága végtelenné válik.

> „Gibbard’s (1973) original proof relies on Arrow’s impossible theorem, and it is known that the latter is still valid when preferences are restricted to be continuous, and even under further restrictions over preferences.” [54, 32]

Lars-Gunnar Svensson alternatív bizonyítása [719]

PhilipJ.Reny:2001 az Arrow-paradoxon és a GS-tétel közti kapcsolatot mutatja be [600]

Alan D. Taylor: az Arrow-tétel és a GS-tétel ekvivalens egymással [740]

Berberàék: Individual versus group strategy-proofness: when do they coincide? [50]

McKelvey tétel

A lehetetlenségi tételek egymást kizáró feltételeket állapítanak meg a közösségi vélemény-aggregálások világában. A preferenciákkal kapcsolatos lehetetlenségi tételek sorában az Arrow-tétel mondta ki a legfontosabbat: amikor az egyéni véleményeket összegezzük, nem úszhatjuk meg a diktatórikus mozzanatot, a közösségi vélemények kialakításakor valaki(k) mindig döntő szerepbe kerül(nek). A Gibbard-Satterthwaite tétel azzal pontosította mindezt, hogy ráirányította a figyelmet a stratégiai szavazás, a szavazatmanipuláció jelenségére. A tétel középpontjában két egymást kizáró feltétel került, amelyek közül választanunk kell: a szavazási rendszerünk vagy az őszinte szavazást vagy a diktátormentességet biztosíthatja, a kettőt együtt nem. A most bemutatandó, harmadik lehetetlenségi tétel szerint a szavazás során minden megtörténhet (anything might happen), mert a szavazás lebonyolítása, a **szavazási napirend** vagy menetrend (agenda setting) kialakítása olyan erőt ad annak kezébe, aki e fölött rendelkezik, amivel bármilyen végeredményt ki tud hozni. És ez a végeredmény nyilvánvalóan az az alternatíva lesz, amit ő (a napirendkezelő) legjobban preferál.

Richard Drummond McKelvey volt az, aki 1976-ban írt egy cikket, amiben megfogalmazta a – később róla elnevezett – tételét [485], amit aztán némely ponton kiegészített, módosított egy három évvel későbbi tanulmányában [486]. Érdekes kapcsolatot jelez az a tény, hogy nagyon gyakran káosz-tételként hivatkoznak McKelvey tézisére. Azért van ez így, mert McKelvey gondolatmenetének kiindulópontjában egy olyan lehetségességi tétel áll, amely kimondja ugyan, hogy a többségi szabály biztosíthatja a Condorcet-győztest bizonyos feltételek mellett, ám McKelvey ezek után épp azt mutatja meg, hogy ez az elvi lehetőség – az alternatívák terében tapasztalható kaotikus viszonyok miatt – gyakorlatilag sosem valósulhat meg.

a preferenciák kezelésekor a téri metafora használata, térbeli

(AL.60) *a*

nyerőhalmaz (HUN) – winset (ENG)

(AL.61) *a*

magpont (HUN) – üres nyerőhalmaz pont (HUN) – core point (ENG) – empty-winset point (ENG) – Plott equilibrium (ENG) – majority-rule equilibrium (ENG) – Condorcet winner (ENG)

Duncan Black tétele (↬216.o.) az egycsúcsú preferenciák (↬201.o.) létezése esetén áll fent, és ez a feltétel természetesen nem mindig teljesül. Nem is teljesülhet minden esetben, hiszen az Arrow-tétel pont ennek lehetőségét tagadja. A Black-tételt így lehetségességi tételként is értelmezhetjük, és feltehetjük a kérdést, vajon vannak-e másfajta feltételek, amelyek alapján ki lehetne másfajta lehetségességi tételeket is mondani. Természetesen léteznek további ilyen feltételek és tételek. 1966-ban Plott cikkében megírta,

hogy a preferenciák sugaras vagy radiális szimmetriája (radial symmetry) esetén a többségi eljárás

1972-ben Otto A. Davis, Morris H. DeGroot és Melvin J. Hinich bebizonyította az euklideszi preferenciákra, hogy $W(x^*) = \varnothing$ nyerőhalmaz eleme(i) akkor és csak akkor többségi Condorcet-győztes(ek), ha megegyeznek a teljes mediánnal (total-median) [192].

A lehetségessségi tételek feltételei összefüggenek egymással: a Black-tételben kihasznált feltétel, a preferencia egycsúcsúsága speciális esete a sugaras szimmetriának, utóbbi pedig a teljes medián egyik altípusa [175]. Ezek a feltételek azonban nem mindig teljesülnek a tényleges társadalmi gyakorlatban. McKelvey meg is jegyezte egy helyen, hogy a teljes medián követelménye olyan erős, hogy gyakorlatilag lehetetlen kielégíteni. Minél több szavazó van, annál könnyebb megsérteni az elvet, hiszen elegendő egyetlen szavazó preferenciájának miniális megváltoz(tat)ása, és borul a feltételekben megfogalmazott kényes egyensúly [486, 1085].

McKelvey az 1976-os cikkében még feltételezte a preferenciákra vonatkozóan az euklideszi tulajdonság létezését, vagyis azt, hogy az alternatívák tere euklideszi, de az 1979-es tanulmányában ezt tovább általánosította, amikor az euklideszi tér feltételezése helyett elegendőnek tartotta tetszőleges topológiai tér létezését. David Austen-Smith és Jeffrey Banks 1999-ben pedig azt bizonyítota be, hogy az euklédeszi preferencia feltételezése helyett elég lehet a folytonossági feltétel rögzíteni [42].

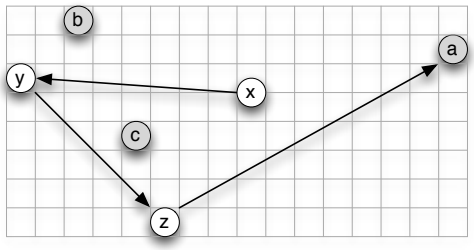
Ronald Rogowski szerint az alternatívák többdimenziós térében az egyéni preferenciák tranzitivitása és egycsúcsúsága mellett is előfordulhat körkörösség [609]

(AL.62) $(|\mathbb{X}| > 2 \wedge |\mathscr{N}| > 2) \rightarrow$ (DICT)

McKelvey (káosz) tétel (HUN) – McKelvey (chaos) theorem (ENG)

Ha több, mint három szavazó van, legalább három politikai kérdés van napirenden, akkor a döntési folyamat nem diktátormentes, vagyis a napirendet a ‘diktátor’, az ‘agenda setter’ fogja irányítani.

további fontos feltétel: a diktátor ismeri a többiek preferenciáit nincsenek tranzakciós költségek



—— A napirend-manipuláció híres példajaként hivatkoznak a ‘Powell Amendment’ név alatt arra az Adam Clayton Powell amerikai képviselő által gyakorolt módszerre, amelynek a tartalmi célja az volt, hogy a szövetségi forrásokból ne támogassanak olyan állami szintű intézményeket, alapítványokat, iskolákat, amelyeken belül valamilyen szegregációs politika érvényesült, és mindezt olyan formai eszközökkel érte el, hogy a “szegregációs tartalmú” törvényekhez mindig benyújtott egy olyan kiegészítést, amely a tiltotta a szegregáció [606]

HR 7535. School Construction Aid Bill. Amendment Denying Federal Funds to Any State that Fails to Comply With the Decisions of the Supreme Court.

	igen	nem
Northern Democrats:	77	42
Southern Democrats:	0	104
Republicans:	148	46
Total:	230	197

HR 7535. School Construction Aid Bill. Final Passage (Rejected)

	igen	nem
Northern Democrats:	116	3
Southern Democrats:	3	102
Republicans:	75	119
Total:	199	229

magyar példa a napirend-manipulációra Körösényi András tanulmányában olvasható [411]

az útfüggőség és napirend-manipuláció kapcsolata????

szavazási paradoxon = körörös többség

két híres kísérlet:

[439] [577]

Richard D. McKelvey és Richard G. Niemi reprezentációs megoldása [487]

Richard D. McKelvey és Peter C. Ordeshook áttekintése: [488]

magyar kísérletek: Meszerics Tamas, Stratégiai viselkedés és bizottsági döntés [502]

Korábban említettem, hogy a McKelvey-tétel a lehetségességi tételekkel szemben megfogalmazott lehetetlenségi tétel, ami gyakorlati szempontokra hivatkozva "utasítja el" a diktátormentesség lehetőségét. De azt is említettem, hogy mindez bizonyos típusú preferenciákra vonatkozóan áll fent. Fel lehet itt tenni azt a kérdést, hogy más preferenciatípusok esetében lehet-e még Condorcet-értelemben stabil a többségi szavazás. Stephen Ansolabehere és James M. Snyder Jr. szerint elképzelhető ilyen eset [23].

(AL.63) → ,????

Ansolabehere-Snyder tétel (HUN) – Ansolabehere-Snyder theorem (ENG)

Barberà és társai Two necessary conditions for strategy-proofness: On what domains are they also sufficient? [52]

másik fontos cikk az egyvölgyű preferenciákra (♀**??**.o.) vonatkozik

Domains, ranges and strategy-proofness: the case of single-dipped preferences [51]

Rothstein: Order restricted preferences and majority rule [616]

a medián szavazó elméletével szemben a reprezentatív szavazó elmélete Rothstein: Representative voter theorems [617]

Herve Moulin bizonyítása szerint az egycsúcsú preferenciák esetén a társadalmi választási függvény nem-manipulálható [522].

(AL.64) ????

Moulin tétel (HUN) – Moulin theorem (ENG)

4.2.12 napirend

McKelvey (káosz) tétele a napirendkezelés manipulatív erejéről (is) szól. Fontos itt megjegyezni, hogy a társadalmi választások elméletében kezelt napirend-tematika eltér a politológusok által elemzett politikai napirend kérdéskörétől, noha kapcsolat nyilván van a két fogalomhasználat és megközelítés között.

első felismerés az 1920-as években Jay G. Blumler és Elihu Katz

Pippa Norris 1999, agenda setting

Bernard C. Cohen: a napirenddel ki lehet jelölni, hogy miről gondolkodjunk, de nem lehet meghatározni azt, hogy hogyan gondolkodjunk

a média sajátos hatalma szelekciós erejében rejlik Gianpietro Mazzoleni és Winfried Schultz: különbséget kell tennia média politikai hatása és a média saját politikája között

Duggan-Schwartz tétel

John Duggan és Thomas Schwartz végezte el a Gibbard-Satterthwaite tétel általánosítását 1993-ban egy publikálatlan preprintben [740, 81], és ennek egy változata jelent meg 2000-ben [221].

(AL.65) ????

Duggan-Schwartz tétel (HUN) – Duggan-Schwartz theorem for Linear Ballots (ENG)

In the context of linear ballots, if n is a positive integer and A is a set of three or more alternatives, then any voting rule for (A, n) that is non-imposed and cannot be manipulated by an optimist or a pessimist has a nominator.

[740, 87]

(AL.66) ????

Duggan-Schwartz tétel nem lineáris szavazatok esetén (HUN) – Duggan-Schwartz theorem for non-Linear Ballots (ENG)

In the context of non-linear ballots, if n is a positive integer and A is a set of three or more alternatives, then any voting rule V for (A, n) that is non-imposed and cannot be manipulated by an optimist or a pessimist is a dictatorship in the sense that then there is a voter such that, for every profile P, the set V(P) contains at least one of the alternatives that this voter has tied for top.

4.2.13 AI-válasz

A társadalmi választások elméletének egyik kulcseleme az aggregálási, választási eljárás fogalma. A közösségi döntések meghozatalához szükséges választási eljárások algoritmusok, így nem kellene, hogy meglepő legyen az a tény, hogy az 1990-es évektől kezdve megjelentek matematikusok, mérnökök, informatikusok is ezen a kutatási terepen.

A formális eszközök iránti igény korábban is vonzotta az ilyen felkészültségű szakembereket ide. Miután egyre másra jelentek meg olyan új feladatok, amikor számítástechnikai környezetben kellett szimulálni ilyen-olyan választási helyzeteket,[22] érthető, hogy a fejlesztések, fejlesztők oldalán is erősödött az igény a szavazáselméletek tanulságainak becsatlakoztatására. Az azonban mégiscsak meglepő, hogy milyen technikákat hoztak be ebbe az új diskurzusba az informatika, matematika oldaláról.

Ahhoz, hogy érteni lehessen, milyen fogalmakkal, elméletekkel operálnak a mérnökök, informatikusok ezen a terepen, tudnunk kell valamit a témánk szempontjából nagyon fontos bonyolultságelmélet világából. A számítástechnika fejlődése során a matematikusok, informatikusok sokat foglalkoztak az algoritmusok problémakörével. Az egyik kiemelten fontos kérdés az volt számukra, hogy milyen számítási feladatot jelentenek a különféle algoritmusok a gépek számára, vagyis mennyi idő szükséges valamely probléma megoldására, a hozzá tartozó algoritmus lefuttatására. Anélkül, hogy a fogalmak lényegi tartalmát, s főleg a mögöttük levő elméleteket megértsük, szükségünk van arra, hogy tudjuk, milyen fogalmak mit jelentenek, mire és miért használják őket a mérnökök. P- és NP-problémákról, illetve utóbbiakon belül NP-nehéz és NP-teljes problémákról szoktak beszélni. Nekünk csak annyit kell tudni, hogy ezek a megoldandó feladatok bonyolultságát, és a megoldáshoz szükséges számításigény nagyságrendjét fejezik ki. A P osztály a polinom időben megoldható problémákból áll, azaz olyan problémákból, melyekhez létezik olyan k konstans, hogy a probléma n hosszú bemenet esetén $O(n^k)$ idő alatt megoldható. Ez annyit jelent, hogy az ide tartozó algoritmusok vállalható idő alatt lefutnak a számítógépeken. Ezzel szemben az NP-nehéz (NP-teljes) osztályba tartozó problémák olyanok, hogy nagy elemszám mellett gyakorlatilag megoldhatatlanok. A megoldhatatlanságot itt csak abban az értelemben kell érteni, hogy kiszámíthatatlanok, vagyis ha elindítanánk tetszőleges képességű és méretű számítógépen a problémát elvileg megoldani képes algoritmust, akkor az sosem jutna a futási idő végére. Elvileg tehát megoldható lenne, csak gyakorlatilag nem számítható ki.

Ezek a problémaosztályok "normál ügymenetben" kihívást jelentenek a matematikusok számára, amikor a cél az ilyen bonyolult problémák megoldása valamilyen eszköz, trükk segítségével. Kiderült azonban, hogy nagyon praktikus haszna is lehet az ilyen problémáknak. Praktikus védelmet jelenthetnek ugyanis olyan helyzetekben, amikor valamilyen informatikai erőforráshoz való hozzájutást effajta problémák megoldásának közbeiktatásával lehet megfelelő módon kezelni ("csak akkor léphetsz be ezen a ponton, ha megoldod ezt a problémát"). A NP-osztályba tartozó feladatok tehát egyfajta kapuőrként is használhatók. És ez a kifordított gondolat segített abban, hogy a mérnökök megjelenjenek a szavazáselméletek világában és használni kezdjék ezeket a bonyolultságelméleti fogalmakat. Ha ugyanis az a feladat, hogy megakadályozzuk az – egyéni vagy koalíciós – stratégiai szavazás lehetőségét, akkor erre meg kell keresni és hasznosítani kell azokat a számítástechnikai eszközöket, amelyeket más területeken már sikeresen alkalmaztak.

Az 1990-es évek kezdetén Bartholdi szerzőtársaival olyan megoldásokat dolgozott ki, amelyek a fenti elveket ültették át a gyakorlatba. Kiderült, hogy a társadalmi választási eljárásokat valóban annál kevésbé lehet manipulálni, minél bonyolultabbak ezek az eljárá-

[22] Egyre bővül az a lista, amely a számítástechnika és a választási probléma összekapcsolódásával kapcsolatos jelenségekről szól – a multiágens alapú döntési, tervező rendszerektől a kollaboratív szűrés technikáját alkalmazó rendszereken át a weboldalakon elhelyezett linkeket szavazásokként felfogó és a keresések hatékonyságát ezáltal növelő keresőrendszerekig.

sok algoritmuselméleti szempontból. Ez praktikusan azt jelenti, hogy az aggregálási mechanizmusok integritásának a szükséges számításigény értelmében vett komplexitás az igazi biztosítéka [57], [56]. Bartholdiék megoldási javaslatai mind a jelöltek, mind a szavazók számára nagy számokat írtak elő, így tudták biztosítani a megfelelő számítási bonyolultságot.

Vincent Conitzer, Tuomas Sandholm és Jérôme Lang azzal érveltek, hogy a jelöltekre vonatkozó számossági elvárások nem mindig életszerűek, hiszen a választási helyzetek egy részében csak kevés alternatíva közül kell választani, ezért olyan számításokat végeztek, és ezek alapján olyan új javaslatokat tettek le az asztalra, amelyek úgy biztosítottak megfelelő számítási bonyolultságot az egyes választási eljárésokkal kombinálva, hogy a jelöltek számára alacsony számú alternatívát is fent lehetett tartani [170]. Tanulmányukban először röviden bemutatták, milyen dimenziókban, milyen kérdések mentén vizsgálják a jelenséget, Kik a manipulátorok? Lehetnek egyéni vagy csoportos, koalícióba szerveződő szavazók. Mit tudnak a manipulátorok a tiszta szavazók szavazatairól, preferenciáiról? Lehet teljes vagy tökéletlen a tudásuk. Súlyozottak-e vagy sem a szavazók? Ezáltal lehet különbséget tenni különböző szavazótípusok, szavazók között (például részvényesi szavazás során). Mi a manipuláció célja? Vagy az, hogy egy adott jelölt nyerjen (ekkor beszélhetünk konstruktív szavazásról), vagy az, hogy ne nyerjen (ez a destruktív szavazás). A vizsgálatba bevontak 9-10 szavazási módszert, és a fenti szempontok alapján sorra mondtak ki tételeket arra vonatkozóan, hogy a különböző feltételek mentén adott eljárások esetében adott (2-7 fős) jelöltszámok mellett milyen számítási problémát jelent a manipulációs szavazás végrehajtása (megakadályozása). Csak a szemléltetés kedvéért bemutatok itt egy táblázatot a cikkből, ami a súlyozott, koalíciós, konstruktív manipulációkkal kapcsolatos adatokat tartalmazza.

jelöltek száma	2	3	4, 5, 6	≥7
Borda	P	NP-teljes	NP-teljes	NP-teljes
véto	P	NP-teljes	NP-teljes	NP-teljes
STV	P	NP-teljes	NP-teljes	NP-teljes
kétfordulós többségi	P	NP-teljes	NP-teljes	NP-teljes
Copeland	P	P	NP-teljes	NP-teljes
maximin	P	P	NP-teljes	NP-teljes
randomizált kupa	P	P	P	NP-teljes
reguláris kupa	P	P	P	P
többségi	P	P	P	P

A táblázat legfontosabb "üzenete" az, hogy bizonyos eljárások esetében 3 jelöltnél is, másoknál pedig 4-7 jelöltnél már meg lehet akadályozni a manipuláció lehetőségét, mert olyan feltételek esetén a számítási bonyolultság már kellően naggyá válik.

A szavazatmanipuláció speciális esete a szavazatvásárlás vagyis a megvesztegetés (bribery). Kicsit furcsa, szokatlan, egyeseknek talán elfogadhatatlan a szavazatvásárlást megvesztegetésnek minősíteni, pedig valóban az. Valaki (itt a szavazó) pénzért (vagy más ellenszolgáltatásért, valamilyen előnyért) olyan döntést hoz, ami másoknak kedvező, és az illető ezt a döntést a pénz (előny) felajánlása nélkül nem hozta volna meg, akkor beszélünk megvesztegetésről, csalásról, megkenésről, és ez a meghatározás illik a szavazatvásárlás esetére is. Piotr Faliszewski társaival együtt amellett érvelt, hogy a megvesztegetési problémát érdemes a manipulációs problémától elkülönítve vizsgálni, mert speciális tulajdon-

ságai vannak [238]. Nem is annyira a szűkebb értelemben vett választás jelenségén belül érhető tetten ez a speciális jelleg, hiszen az "önkéntes", külső befolyástól mentes stratégiai szavazás eljárástechnikai értelemben nem különböztethető meg a szavazatvásárlás útján megváltoztatott véleményektől. A megvesztegetés a eljárástechnikai részleteken túl válik érdekessé (és elemezhetővé). Mennyi szavazatot kell megvenni, hogy biztosítani lehessen az adott jelölt győzelmét, ugyanannyit kell-e fizetni minden megvásárolt szavazónak, vagy vannak minősített szavazók, összesen mennyibe kerül a szavazatvásárlás, lehet-e pénzen kívüli eszközökkel fizetni – ilyen kérdések merülnek fel, amelyekkel kapcsolatban ugyanúgy és ugyanolyan indokok miatt kell a bonyolultságelméletet segítségül hívni, ha meg akarjuk akadályozni az ilyen manipulációk lehetőségét. Ha el akarjuk lehetetleníteni az iylen próbálkozásokat, akkor valóban ismerünk kell minden olyan technikai lehetőséget, ami módot adhat arra, hogy a manipulációs kísérletek sikeresek lehessenek. Érdekes tanulságokat nyerhetünk az ilyen vizsgálatokból. A megvesztegetés modellezésekor például kiderült, hogy a jelenség egyik alfajaként minősített "hátul cserélős" megvesztegetés (swap bribery) során indirekt módon is lehet sikert elérni [239]. A megvesztegetés során ugyanis lehetséges, hogy valaki vagy nagyon drágán hajlandó a saját preferencialistája első helyén levő jelöltet elmozdítani a csúcsról, vagy erre a egyáltalán nem vehető meg. A preferencialistáján hátrébb levő jelöltek között viszont már sokkal könnyebben (sokkal olcsóbban) meg változtatja a véleményét. A választás befolyásolása szempontjából azonban sok esetben sikeres taktika lehet, ha az ilyen szavazók esetében a gyenge preferenciák sávjában érik el a szavazatcseréket, és ez a sok apró rangsorváltoztatás a hátsó preferenciarégiókban egymással összeadódva mégiscsak meghozhatja a kívánt eredményt.

4.2.14 Párosításelmélet

A kollektív preferenciák kezelésének egy speciális területét jelentik azok a problémák, amelyeket a párosításelmélet vizsgál a hatvanas évek elejétől fogva. David Gale és Lloyd S. Shapley írta le először azt a problémát, amit párosítási feladatnak, majd később a problémát leíró elméletet **párosításelméletnek** neveztek el [277]. A probléma megoldásra Shapleyék egy olyan algoritmus javasoltak, ami a mai napig Gale-Shapley algoritmus néven ismert és hivatkozott szerte a világban. A stabil házasítási vagy párosítási probléma (stable matching problem) arról szól, hogy össze kell párosítani lányokat és fiúkat egymással azok preferenciarendezései alapján a lehető legoptimálisabb módon, vagyis úgy, hogy mindeki a lehető leginkább kedvelt partnerével kerülhessen egy párba. A stabil megoldást itt úgy értelmezhetjük, hogy a kapcsolatok között nincs egyetlen **blokkoló pár** sem, vagyis nem létezik olyan fiú-lány kettős, akik nem házastársak, de mindketten boldogabbak lennének egymással a jelenlegi párjukhoz képest. Ezt úgy is fogalmazhatjuk, hogy ha egy fiú és egy lány nem házas, akkor legalább az egyikük jobban szereti a jelenlegi házastársát, ezért nem lehet őt elcsábítani, vagyis az algoritmus válásbiztos (divorce-proof). Azt természetesen nem várhatjuk el, hogy minden esetben mindenkit össze lehessen kapcsolni a legjobban akart partnerével, de azt mindenképpen biztosítani kell, hogy mindenki olyan jelöltet kapjon a végén, aki rajta van a preferencialistáján (vagyis mindenki találjon magának valakit a preferencialistájáról).[23] Shapleyék erre

[23] A probléma legáltalánosabb formájában meg szokták engedni azt, hogy valaki ne találjon párt magának. Az ilyen esetekre azt mondják, hogy valaki "önmagával is párba kerülhet".

a feladatra dolgozták ki azt az algoritmust, ami végül is képes volt minden esetben stabil megoldást szállítani.

A **késleltetett elfogadási algoritmus** (deferred acceptance algorithm) lényege az, hogy a fiúk minden körben felkérik a lányokat a preferencialistájuk alapján,[24] hogy legyenek a párjuk (táncpartnerük stb.), a lányok pedig a kérők közül a nekik legjobban tetsző fiúknak feltételes igen-t (talán-t) mondanak (ha van ilyen), míg a többieknek nem-et. Az elküldött fiúk a következő körben felkérik a preferencialistájukon szereplő következő lányt. A lányok az új kérőket elutasítják, ha azok hátrébb vannak, mint a már esetleg talán-nal talonba tett fiú vagy ha annál jobb helyen áll az új jelölt, akkor lecserélik a régit az újra. A lecserélt fiú újra beáll a kérők közé, és a következő fordulóban – a többi visszautasított társával együtt – megy a preferencialistája következő helyén szereplő lányhoz. Ekkor a lányok újra választanak. Ez így megy addig, amíg minden fiú ki nem köt valahol, vagyis amíg talán választ nem kap. Mivel a jelenség szemléltetésére Gale és Shapley nehezebben követhető példát adott, ezért itt inkább egy Biró Péter és társai által bemutatott példát ismertetek [89]. A példában szereplő négy lány és négy fiú az alábbi sorrendben szeretne magának társat választani.

	lányok rangsorai		fiúk rangsorai
Anna:	Gábor > Henrik	Endre:	Dóra > Anna > Bea > Csilla
Bea:	Gábor > Ferenc > Endre > Henrik	Ferenc:	Dóra > Csilla > Anna > Bea
Csilla:	Gábor > Henrik > Endre > Ferenc	Gábor:	Dóra > Csilla > Bea > Anna
Dóra:	Henrik > Gábor	Henrik:	Bea > Anna > Csilla

Az első körben Endre, Ferenc és Gábor felkéri Dórát, Henrik Beát. Dóra Gábornak, Bea Henriknek mondja, hogy ‘talán’. Mivel Endre és Ferenc kosarat kaptak, tovább próbálkoznak a második menetben. Ekkor Endre Annát kéri fel, de most is elutasítják. Ferenc Csillát kéri fel, aki ‘talán’ választ kap. A harmadik körre már csak Endre maradt pár nélkül. Ekkor megkéri a következő lányt a listáján, Beát. Beának Endre jobban tetszik, mint Henrik, ezért Henriktől elveszi és átadja Endrének a ‘talán’ jelzést. A negyedik fordulóban Henrik felkéri a preferencialistáján szereplő következő lányt, Annát és ezzel véget is ér az algoritmus, hiszen az utolsó fiú is párra lelt. Mivel minden fiú megállapodott, a lányok elfogadják a jelenlegi partnerüket, a párok “véglegesítődnek”. Ezt a folyamatot bemutatom egy táblázat segítségével is, hogy látszódjanak az egyes fordulók eseményei. Az egyes fordulókban az adott lányhoz bejelentkező fiú vannak feltüntetve, és aláhúzással jelölöm azt, ha az adott fordulóban egy lány talonba tett vagy ott tartott valakit (a harmadik körben előfordult, hogy a talonban kicseréltek valakit).

	1. kör	2. kör	3. kör	4. kör
Anna		Endre		Henrik
Bea	Henrik	Henrik	Endre, Henrik	Endre
Csilla		Ferenc	Ferenc	Ferenc
Dóra	Endre, Ferenc, Gábor	Gábor	Gábor	Gábor

Az algoritmus mindenkit összerendelt azok közül, akiket el tudtak képzelni párjuknak a maguk számára. Természetesen nem adta ki mindenkinek a legjobb megoldást, de min-

[24] A preferenciarendezéseknek nem kell teljesnek lennie, vagyis megengedhető, hogy valaki ne mindenkit vegyen fel a listájára, vagy másként: bárkit ki lehet zárni a potenciális jelöltek közül.

denkinek "vállalható" pár jutott, olyan, aki rajta volt az elképzelhető partnerek listáján. Az algoritmus a fiúk számára taktikázásbiztos (strategy-proof), de a lányok számára nem az. Alvin Roth egy 1984-es cikkében bebizonyította, hogy nincs olyan stabil párosítási mechanizmus, amelyben minden szereplő számára az igazmondás a domináns stratégia [615]. Ez azonban nem akkora baj, ugyanis ha a szereplők nem ismerik a többi szereplő preferenciáit, akkor nem tudják előre megmondani, hogyan kéne módosítani a sorrendjükön, hogy számukra kedvező módon befolyásolhassák a végeredményt. Ez azt jelenti, hogy az elméleti lehetőség nem igazán váltható át gyakorlati haszonná: sokszereplős piacon gyakorlatilag nem lehet megjósolni, mi lehet a kiment, és emiatt nem igazán lehet manipulálni a dolgok lefolyását.

A párválasztási probléma algoritmusa valóságidegennek tűnhet ahhoz képest, ahogy a házasság, a párválasztás ténylegesen megvalósulás a valós társadalmi gyakorlatban. Vannak azonban olyan társadalmi jelenségek, amelyek a párválasztási problémához nagyon hasonlítanak, és a megoldásuk az itt bemutatott mintát követi. Shapleyék a "programadó cikkükben" is hivatkoztak a kollégiumi felvételik során a kollégiumok és az oda jelentkezők összekapcsolására vagy csak egyszerűen a szobatársak összerendelésének kérdésére, de ez nem nagyon tér nagyon a házassítási problémától. Talán a legfontosabb (és a legtöbbet hivatkozott) példa a párosítási feladatok közül az amerikai rezidensek központosított kórházválasztási procedúrája. A probléma lényegét tekintve itt is ugyanaz , mint a párválasztás esetében, hiszen a lehető legoptimálisabb módon kell összerendelni a kórházakat és a pályakezdő orvosgyakornokokat. Ennek a példának a gyakorlati jelentősége azonban már egy pillanatig sem kérdőjelezhető meg, és az a tény, hogy ezt a feladatot már sok évtizede a Gale-Shapley algoritmussal oldják meg folyamatosan, a megoldás erejét és jelentőségét egyértelműen bizonyítja. Kiderült, hogy ezt a módszert alkalmazták már az ötvenes évek elejétől fogva (tehát még a Gale-Shapley-algoritmus "felfedezése", leírása előtt, és a kórházak számára optimális módon). Az algoritmus módosítására csak évtizedekkel később volt szükség, akkor is csak azért, mert a hatvanas években már egyre több fiatal orvos házaspár keresett magának közös munkahelyet, vagyis nekik az a legfontosabb, hogy ugyanabban a városban kapjanak állást. Ezt feltételt az eredeti algoritmus nem tudta figyelembe venni, ezért kellett kicsit módosítani rajta [88]. Ezt a módszert természetesen nemcsak orvosok esetében lehetett alkalmazni. Az egyetemi jelentkezési rendszerekben pont ugyanez a feladat, csak ilyenkor a diákok és az iskolák, egyetemek közti párosításokat kell egyensúlyba hozni, és bizony ennek a problémának nem lehet vitatni a gyakorlati jelentőségét.[25] Ugyancsak vitathatatlan gyakorlati haszonnal jár annak azoknak a vesecsereprogramok a vezérlése, amelyet pont a párosításelmélet fenti sikerei nyomán indítottak el a világ egyre több pontján [88]. Ebben az esetben a vesére váró betegek, illetve a vesedonorok közti összerendeléseket kellett és kell minél optimálisabb módon megteremteni.

Van tehát egy algoritmus, amivel a fent bemutatott, illetve az azokhoz hasonló, további párosítási poblémákat sikeresen meg lehet oldani, a módszer bizonyított a munkaerőpiacon (rezidensek elosztása), az oktatásban (felvételik lebonyolítása), az egészségügyben (vesedonorok megszervezése). A procedúra értékelésekor fontos látnunk, hogy a módszer alkalmazása – a látszattal szemben – egyáltalán nem a központi tervezés megvalósulását jelenti. A problémakör Nobel-díjas kutatója, Alvin Roth szerint másról van itt szó.

[25] A magyar felvételei rendszer és a párosításelmélet kapcsolatáról bővebben lásd: [88], [405], [406].

> „Vegyük észre, hogy az itt vizsgált központosított piacok nem jelentenek központi tervezést, ahogyan a legtöbben ezt értelmezik, mivel ezek a piacok úgy lettek megalkotva, hogy érzékenyek legyenek a szereplők kinyilvánított preferenciáira, ahelyett, hogy egy tervező ezektől független céljainak elérését szolgálnák. Ami itt központosítva van, az nem az elérendő cél, hanem a piaci mechanizmus maga." idézi: [88]

Mielőtt foglalkoznék még a késleltetett elfogadási algoritmus néhány fontos következményével, bemutatom, hogy lehet a tételt formális módon leírni. Amennyiben $\mathbb{M}$-vel jelöljük a férfiak halmazát, $\mathbb{W}$-vel a nőkét, $>_a$-vel valamely a ágens preferenciarendezését és μ-vel ($\mu : \mathbb{M} \cup \mathbb{W} \mapsto \mathbb{M} \cup \mathbb{W}$) a párosítási függvényt, akkor néhány tulajdonság definiálása után formálisan is megadhatjuk a Gale-Shapley tételt: a házassági problémának mindig létezik stabil megoldása.

(AL.67) $\exists m \exists w (\mu(m) = w \leftrightarrow \mu(w) = m) \wedge \forall w \forall m ((\mu(w) \in \mathbb{M} \vee \mu(w) = w) \wedge (\mu(m) \in \mathbb{W} \vee \mu(m) = m))$

párosítás(i függvény) (HUN) – matching (function) (ENG)

(AL.68) $\text{BLOCKED_INDIVIDUALLY}(\mu) \stackrel{\text{def}}{=} \exists a \in \mathbb{M} \cup \mathbb{W} (a >_a \mu(a))$

egyénileg blokkolt párosítás (HUN) – matching blocked individually (ENG) – matching blocked by an individual (ENG)

(AL.69) $\text{BLOCKED_COLLECTIVELY}(\mu) \stackrel{\text{def}}{=} \exists w \in \mathbb{W} \exists m \in \mathbb{M} (w >_m \mu(m) \wedge m >_w \mu(w))$

párosan blokkolt párosítás (HUN) – matching blocked collectively (ENG) – matching blocked by a pair of agents (ENG)

(AL.70) $\text{STABLE}(\mu) \stackrel{\text{def}}{=} \neg \text{BLOCKED_INDIVIDUALLY}(\mu) \wedge \neg \text{BLOCKED_COLLECTIVELY}(\mu)$

stabil párosítás (HUN) – stable matching (ENG)

(AL.71) $\exists \mu (\text{STABLE}(\mu))$

Gale-Shapley tétel (HUN) – Gale-Shapley theorem (ENG)

A tételre Gale és Shapley konstruktív bizonyítást adtak, amikor megmutatták, hogy az algoritmust, aminek alkalmazásával mindig el lehet jutni a feladat stabil megoldásához [277]. Annak ellenére, hogy a házasítási problémának mindig van stabil megoldása, azért az is fontos tény, hogy ez az algoritmus nem mindig és nem mindenkinek biztosítja a legjobb párosítást. A Gale-Shapley algoritmus ismeretében Lester E. Dubins és David A. Freedman egy 1981-es cikkben feltette a kérdést, hogy abban az esetben, ha a játékosok között vannak olyanok, akik ismerik a többiek preferenciáit, akkor van-e esélyük arra, hogy a saját preferenciakinyilvánításaik megváltoztatásával befolyásolják az algoritmus menetét, és ezzel a végerdményt – nyilván úgy, hogy számukra kedvezőbb lehessen. Lehet-e manipulálni, érdemes-e hazudni, vagyis lehet-e itt machiavellizmusról beszélni – tette fel a kérdést Dubins és Freedman [218]. A válasz az volt, hogy az egyetemek esetében elképzelhető, hogy némelyiküknek előnyösebb lehet hazudni, azonban a hallgatók esetében ez a lehetőség már nem áll fent.[26] Minden hallgatóra (és azok minden koalíciójára) igaz, hogy az algoritmus biztosítja a legjobb stabil megoldást számukra, és ezen nem tudnak semmilyen taktikázással javítani. Dubinsék cikkére válaszul David Gale és Marilda Sotomayor példát és bizonyítást hozott arra, hogy több stabil illeszkedés esetén

[26] Dubins és Freedman az egyetemek és a diákok összerendelésének problémáját vizsgálta.

a másik irányban mindig lehetséges hazugsággal előnyt szerezni legalább egy szereplő számára [278].[27]

Sok más tétel és alkalmazási lehetőség lenne még bemutatható a késleltetett elfogadási algoritmussal kapcsolatban, de ezek elemzésével itt nem foglalkozom tovább. Az elmélet további tételeiről, az elméleti eredmények konkrét gyakorlati alkalmazásairól értő összefoglalást nyújt Alvin Roth 2008-as áttekintő tanulmánya [614].

4.2.15 Körzethatár-manipuláció

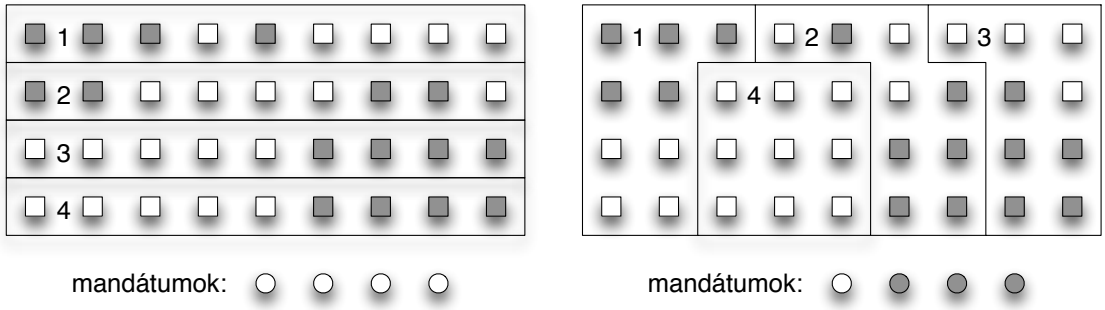
Az egyéni vélemények közösségi véleményekké transzformálásának következő befolyásolási lehetősége a **körzetesítés** vagyis a körzethatárok megváltoztatása (gerrymandering). Itt azokról a körzethatárokról van szó, amelyek meghatározzák az aggregálási számítások alapjául szolgáló területi egységeket. Még pontosabban: a földrajzi térben szétszórtan elhelyezkedő szavazópolgárokat egy-egy szavazataggregációs közösségbe egybefogó határokról van szó, amelyeket természetesen – mindig és szükségszerűen önkényes – politikai döntések határoznak meg. Ez a körzetesítés így mindig rendelkezik valamilyen mértékű manipulációs potenciállal.

Mielőtt bemutatnék egy egyszerű példát a körzethatárok átszabásában rejlő manipulációs lehetőségre, fel kell hívnom a figyelmet arra, hogy két fontos körzetfogalmat el kell különítenünk egymástól, nem szabad őket összekevernünk egymással. A szavazatokat szavazókörökben gyűjtik össze. Ezeket a körzeteket kicsire kell szabni, mert igazodni kell ahhoz az elváráshoz, hogy a szavazópolgárok kis fáradsággal fel tudják keresni. A szavazókörökhöz így viszont csak kevés embert lehet hozzárendelni, amiből következően a mandátumokat sem lehet egy-egy szavazókörben kiosztani. Ezért van szükség a választókerület fogalmára, ami a mandátumok kiosztásának területi egysége. Természetesen szükség van arra is, hogy kapcsolatot teremtsünk a kétféle körzetfogalom között, vagyis összerendeljük a szavazatok összegyűjtésének színtereit, a szavazóköröket és a mandátumok kiosztásának egységeit, a választókerületeket egymással. Akármilyen elveket is rögzítünk azonban, amikhez igazodni kell a szavazókörök és választókerületek párosításakor, mindig marad hely az eseti (önkényes) döntésnek, és ez adja a körzethatár-átszabások elvi lehetőségét. Ezek alapján nézzük meg, hogyan is működhet ez.

Képzeljük el, hogy van két párt, harminchat szavazókörzet. Feltételezzük fel még, hogy minden körzetben ugyanannyi ember szavaz, és az egyszerűség kedvéért tegyük fel azt is, hogy minden szavazókörzetben az egyik vagy másik pártra adják le az összes szavazatot (ez ugyan életszerűtlen feltételezés, de a lényeget ez nem érinti). Tegyük fel még fel azt is, hogy egy-egy mandátumot lehet kiosztani négy választókerületben. A kérdés már csak az, hogy a harminchat szavazókörből hogyan csináljunk négy választókerületet, ahol egy-egy mandátumot lehet nyerni. A következő ábrán a harminchat szavazókör kétféle választókerületi besorolását mutatom be.

A szürke és fehér négyzetek jelölik az egyik vagy másik pártra egységesen szavazó szavazóköröket (nevezzük a két pártot fehér és szürke pártnak). A választókerületi felosztást jelzik a szavazókörök köré rajzolt körzethatárok. Ha egyszerű többségi szavazással dön-

[27] Mivel Gale és Sotomayor a házasítási problémát használták fel az érvelésük során, ezért szellemes módon utaltak Dubins és Freedman 'Machiavelli and the Gale-Shapley algorithm' című cikkére, amikor saját tanulmányuknak a 'Ms. Machiavelli and the stable matching problem' címet adták.



tünk, akkor nagyon eltérő végeredményt ad a két felosztás. A baloldali besorolás alapján a négy választókerület mindegyikében a fehér párt nyeri el a mandátumot, míg a jobboldali felosztásban a szürke párt három, a fehér párt egy mandátumot kap.

Bár a gerrymandering valóban létező, "művelhető" manipulációs lehetőség és technika, ahhoz képest, hogy mennyire sokat hivatkoznak rá, nincs akkora potenciál benne. Sok országból, sok történelmi helyzetből lehet példákat hozni arra, hogy alkalmazták, sőt, sokszor sikerrel alkalmazták ezt a technikát, de ettől még tudni kell a helyén kezelni a dolgokat. Mivel az 1990-ben megállapított választókerületi határok sok esetben nem feleltek meg a törvényi előírásoknak (az időközben lezajlott népmozgalmi folyamatok eredményeként), 2008-ban készítettünk egy olyan alkalmazást, ami a 2006-os parlamenti választások adataira támaszkodva, bizonyos paramétereket szabadon változtatva átszabta a választókerületek határait.[28] Bár az alkalmazás használatakor különböző paramétereket beadva egészen eltérő választókerületi határokat és ezzel eltérő szimulált végeredményeket lehetett kihozni, az is kiderült az ilyen próbálkozásokból, hogy nagyon radikálisan eltérő végeredményeket nem lehetett elérni, akárhogy is változtak a körzethatárok.[29]

4.2.16 Többkörzetes paradoxon

Sorolhatnánk a körzetmanipuláció, a gerrymandering fogalma alá azt a körzethatárokkal kapcsolatos jelenséget, amit Peter C. Fishburn és Steven J. Brams szerzőpáros **többkörzetes paradoxonnak** nevezett el, és amire egy 1983-as cikkében mutatott be egy szemléltes példát [256]. A helyzet azonban pont fordítva van. A többkörzetes paradoxon létezése biztosítja a körzethatár-manipuláció elvi lehetőségét, vagyis az előbbi jelenség logikailag megelőzi az utóbbit. A többkörzetes paradoxon ugyanis azt a helyzetet írja le, amikor a szavazatokat – a szavazók térhez való kötöttsége miatt, tehát pusztán gyakorlati okokból – körzetekbe rendezve gyűjtik össze, és a szavazási adatok összegzése során paradoxon adódik. Ha nem lennének körzetek, akkor nem adódna semmi gond, de ha vannak körzetek, akkor paradoxon keletkezhet. És ha már van paradoxon, akkor az ilyen helyzeteket már lehet manipulálni, vagyis akkor már helye van a gerrymanderingnek is. A többkörzetes paradoxon thát az alap, de még nem tudjuk, miről is szól pontosan ez a

[28] Az EOVK-határoló alkalmazásról beszámoló cikk: Pártos térképészek manipulálhatják a választásokat, at: http://index.hu/belfold/eovk4876/

[29] Nem kifejezetten a gerrymandering szempontjából, hanem inkább a szavazatok területi megoszlásainak vizsgálata miatt kollegáimmal korábban több tanulmányban, könyvben is foglalkoztunk a választások térbeliségével [497], [496], [499], [495], [498], [501].

paradoxon. A no-show paradoxon kapcsán korábban már bemutatott kisváros (↬222.o.) szavazási anomáliái közé tartozik ez az új eset is. Bramsék cikkének főszereplője, Mrs. Smith csinált egy újabb kimutatást a kisvárosi választás adatai alapján. A szavazást két szavazókerületre bontva, egy keleti és egy nyugati körzeten belül végezték el a városka lakói. Az adatok így a két körzetre bontva is rendelkezésre álltak Mrs. Smith számára. Az alábbi táblázat mutatja a különböző preferenciasorrendekre leadott szavazatokat a két körzet szerinti bontásban.

összes szavazat	sorrendek	keleti körzet	nyugati körzet
417	BHW	160	257
82	BWH	0	82
143	HBW	143	0
357	HWB	0	357
285	WBH	0	285
324	WHB	285	39
1608		588	1020

Az adatok elemzéséből meglepő következtetést lehet levonnni. Korábban már bemutattam, hogy ezen a választáson a három jelölt, Mrs. Bitt, Mr. Huff és dr. Wogg közül Mr. Huff lett a nyertes (↬222.o.). Ha azonban a kerületekre bontható adatok alapján kiszámolnánk, hogy ugyanazon eljárást alkalmazva ki nyerte volna meg a választást a két körzeten belül külön-külön, meglepő eredményhez jutnánk. A két körzetben az alábbi szavazatösszesítést kapnánk.

jelöltek	keleti körzet	nyugati körzet
Bitt	160 (160+0)	339 (257+82)
Huff	143 (143+0)	357 (0+357)
Wogg	285 (0+285)	324 (285+39)

Az első körben kiesne a legkevesebb szavazatot kapott jelölt, azaz a keleti körzetben Mr. Huff, a nyugatiban dr. Wogg. A második körös számítás során tehát Mrs. Bittnek a keleti körzetben dr. Wogg, a nyugatiban Mr. Huff lenne az ellenfele, és az alábbi adatok alapján számolnák ki a körzetnyerteseket.

jelöltek	keleti körzet	nyugati körzet
Bitt	303 (160+143)	624 (257+82 +285)
Huff		396 (357+39)
Wogg	285 (0+285)	

A két körzetben külön-külön számítás esetén tehát mindkét esetben Mrs. Bitt nyerte volna meg a versenyt, miközben tudjuk, hogy az összesített eredmények alapján más lett a győztes. A paradoxon nyilvánvaló.

Ez a jelenség a Simpson-paradoxon speciális esetének tekinthető [628].

továbbbi paradoxonok: [542]

4.2.17 Többesetes paradoxon

Bizonyos értelemben a választási paradoxonok vég nélkül szaporíthatóak, ezért fontos kérdés, hogy ezek között vannak-e olyanok, amelyek visszavezethetők valamely más paradoxonra. A többkérdéses vagy **többesetes választás paradoxona** (multiple-elections paradox) jelenségét egy 1998-ban megjelent cikkben Steven J. Brams, D. Marc Kilgour és William S. Zwicker nevezte meg először [114], és bár a cikk végén megmutatták azt is, hogy a jelenség értékelhető a szavazási paradoxon egyik változataként, mégis önálló névvel, újfajta minőséggel hivatkoztak erre a problematikára. A paradoxon olyan választási helyzetekben jelentkezhet, amelyekben egyfelől egyszerre több témában is szavazni kell, másfelől mindegyik témában *igen* vagy *nem* lehet a választási lehetőség. Bramsék az egyes témákat, szavazási kérdéseket propozícióknak, míg az összes szavazási kérdés együttesét kombinációnak nevezték el. Az ilyen szavazások során a szavazókat annyiféle szavazói profilba sorolhatjuk be, ahány szavazatkombináció lehetséges elméletileg. Ha valóban csak *igen*/*nem* értékek közül lehet választani, akkor k szavazási kérdés esetén a szavazatkombinációk száma összesen k^2 lehet. Az egyes szavazók által leadott tényleges szavazatkombinációkat mindig besorolhatjuk valamelyik szavazatkombináció-profilba, és összeszámolhatjuk, mennyi szavazat érkezett összesen ezekre. A szavazatkombinációkat az adott sorszámú kérdésre adható igen vagy nem szavazatok szeint I-vel vagy N-nel jelölhetjük, tehát egy szavazatkombináció mindig annyi I vagy N betűből áll, ahány kérdést tesznek fel a választás során. Hogy miként lesz ebből paradoxon, jobban érthetővé válik, ha megnézünk egy konkrét példát. Tegyük fel, hogy három kérdésről kell szavazni, csak *igen*/*nem* válasz lehetséges, 13 szavazónk van, és az alábbi szavazatok érkeztek az egyes szavazatkombinációkra.

profilok	III	IIN	INI	INN	NII	NIN	NNI	NNN
szavazatok	1	1	1	3	1	3	3	0

Az adatokat úgy értelmezhetjük, hogy négy szavazatkombinációra 1-1, háromra 3-3, egyre pedig 0 szavazat érkezet. A legtöbb támogatás tehát azok a “kimenetek” kapták, amelyekben két *nem* és egy *igen* szavazat érkezett. Ezekre a leggyakoribb szavazatkombinációkra mondhatjuk azt is, hogy ezek a leggyakoribb választótípusok is egyben, de ez az eloszlás nem a szavazásra feltett kérdésekre vonatkozik, hanem a kérdésekre adott válaszok adott kombinációjára. Ha az egyes szavazási kérdésekre akarunk végeredmény kihozni, akkor ha egyszerű többségi szavazással döntünk, akkor mindhárom kérdésben (az elsőben, másodikban és harmadikban) 7:6 arányban a nem szavazatok kerülnének többségbe, vagyis a közösségi döntés a három ‘nem’ (NNN) lenne. És pont ez a mozzanat, amit paradox jellegűnek tarthatunk. Az NNN szavazatkombinációra ugyanis senki sem szavazott, vagyis senki sem érezheti úgy, hogy az ő kedvére való végeredmény született a szavazás során. A közösségi döntés áll a legmesszebb az egyéni szavazói profiloktól, vagy másként: a győztes (szavazat)kombináció kapta a legkevesebb szavazatot. Kétféle aggregálási mód létezik tehát, amelyek eltérő végeredményeket hozhatnak ki olykor. Amikor az összes kérdésre leadott szavazatprofilok gyakoriságát hasonlítjuk össze egymással, akkor kombinációaggregálásról (combination aggregation) beszélhetünk. Amikor az egyes kérdésekre (propozíciókra) adott – igenlő – szavazatokat egyenként összegezzük, tehát először kérdésenként döntjük el, hogy a támogató vagy ellenző vélemények lettek-e a győztesek, majd ez alapján képezhetjük a győztes véleményprofilt az összes

kérdésre együttesen, akkor Bramsék szerint propozícióaggregálásról (proposition aggregation) van szó. A szembeállítás tartalma nyilvánvaló, a javasolt kifejezéspárt azonban nem tartom szerencsésnek, mert nem szemantikailag egynemű fogalmakat használ, amikor a kétféle aggregálási lehetőséget a propozíció és kombináció eltérő minőségeivel akarja megragadni. Ezen a ponton jobbnak tartom az **esetenkénti aggregálás** (vagy: egyenkénti aggregálás) és a **kombinált aggregálás** terminusait alkalmazni.

A Brams-Kilgour-Zwicker szerzőtrió még sokféle formában bemutatta a paradoxont, elemezték kettő- vagy négy-kérdéses változatban, a tartózkodási lehetőség megengedésével vagy olyan paraméterek mentén, amikor a győztes kombinációra adott legkisebb szavazatszámban döntetlen adódik, vagyis más kombinációk is ugyanannyi szavazatot kapnak, mint a végső győztes kimenet. Bemutattak egy olyan konstellációt, amikor négy kérdésre adott 31 szavazat az alábbi megoszlást mutatta.

profilok	IIII	IIIN	IINI	INII	NIII	IINN	ININ	INNI
szavazatok	0	4	4	4	4	1	1	1
profilok	NIIN	NINI	NNII	INNN	NINN	NNIN	NNNI	NNNN
szavazatok	1	1	1	1	1	1	1	5

Ez az eset azért különösen érdekes paradoxon, mert az egyenkénti aggregálás során 16:15 arányban minden kérdésben az *igen* szavazatok győznek, vagyis a IIII profil a győztes, miközben egyetlen szavazó sincs, aki a négy kérdésben ilyen mintázat szerint szavazott volna. Ezzel szemben a kombinált aggregálás végeredményeként az NNNN kombináció jönne ki győztesként, hiszen a legtöbben (öten) erre a kombinációra szavaztak. A kétféle aggregálási technika tehát a két szélsőséges kombinációt, az IIII- és az NNNN-profilt hozná ki győztesnek, vagyis az eljárás felcserélése teljesen megfordítaná a szavazás végeredményét. Ezért nevezték el Bramsék ezt az esetet teljes megfordítás paradoxonának (complete-reversal paradox).

A paradoxon lehetőségét szemléltető elméleti példák bemutatása után a szerzőtrió bemutatott pár konkrét példát az amerikai politikatörténetből, amikor a többkérdéses választás paradoxona valósult meg. 1990-ben California államban 28 kérdésből álló népszavazást tartottak. A Los Angeles városában és a környékén leadott szavazatokat elemezve kiderült, hogy az egyenkénti aggregálás szerinti nyerő kombinációra egyetlen szavazat sem érkezett. Ez persze volt nem akkora meglepetés, hiszen 28 kérdés esetén az összes lehetséges szavazatkombináció száma 2^28, azaz majdnem 270 millió, ami a leadott durván két milliós szavazathoz képest nagyságrendekkel nagyobb, vagyis a lehetséges kombinációk döntő részére nem érkezett egyetlen szavazat sem. Ennek ellenére azért voltak olyan kombinációk, amelyek az egyenletes eloszláshoz képest jóval nagyobb szavazatszámot kaptak. A legtöbb szavazatot (1,75%-ot) a minden kérdésre tartózkodást adó profil volt,[30] a második helyen végzett a *teljes-nem* kombináció (1,72%), ötödik lett a Los Angeles Times újság által sugallt szavazási javaslat (0,29%), kilencedik pedig a *teljes-igen* mintázat (0,20%).

A történelmi példák között Brams és társai olyan jelenségre is hivatkoztak, ami ugyan nem a népszavazás témakörébe tartozik, mert nem *igen*/*nem*-mel megválaszolható kérdésekre kell válaszolni, mégis "produkálja" a többesetes választási paradoxont. Az elnökválasztás évében háromfajta politikai posztra is szavaznak az emberek az USA-ban, az el-

[30] A szavazás során három válasz volt lehetséges, a támogatáson és ellenzésen túl lehetett tartózkodni is.

nöki pozícióra, valamint a szenátusi és a képviselőházi tagságra. Ha a szavazási lehetőségként csak a demokrata (D), illetve republikánus párti (R) jelöltekre adott szavazást tekintjük (ez Amerikában teljesen helytálló feltételezés), valamint a szenátusi és képviselőházi tagságot egy-egy pozíciónak tekintjük (ez is elfogadható leegyszerűsítés, mert politikai szempontból sokszor az az igazán lényeges, hogy melyik pártnak van többsége az adott testületben), akkor három leadható szavazat (elnök, képviselőház, szentáus) és kétféle szavazási lehetőség (R vagy D) létezik. Ilyen körülmények között a többkérdéses paradoxon ugyanúgy előállhat, csak ekkor – Bramsék szerint – ekkor nem propozícióaggregálásról, hanem pozícióaggregálásról (office aggregation) beszélhetünk. Ekkor a három pozícióról külön-külön szavaznak, és esetenkénti aggregálással döntenek a poszt betöltéséről, ám a politikai életet leginkább a három különálló kérdésben kialakult eredmény kombinációja határozza meg. Ennek a kérdésnek, vagyis az egységes vagy osztott kormányzás (unified or divided government) helyzetnek az értékelése az amerikai politikai elemzések fontos témája régóta. 1976-ban egységes, 1980-ban osztott kormányzás jött létre, és a szavazatoknak az egyes kombinációkra vonatkozó eloszlását a Bramsék által feldolgozott formában (a szavazatokat 315-316 körzetre vetítve) mutatja a következő táblázat.

kombináció	DDD	DDR	DRD	DRR	RDD	RDR	RRD	RRR	
1976	*40,8*	5,7	2,6	1,9	11,7	8,2	8,5	20,6	100% (n=316)
1980	22,2	2,5	2,2	0,6	16,8	12,7	*14,3*	28,6	100% (n=315)

Az adatokból könnyen kiolvasható, hogy 1976-ban nagy volt a nyomás az egységes kormányzásra, hiszen a DDD és az RRR kombináció együtt megkapta a szavazatok majdnem kétharmadát. Ekkor a teljes demokrata párti győzelem született, tehát nem alakult ki paradoxon. 1980-ra viszont nagyot változott a szavazók véleménye. Ekkorra már a szavazók fele elfogadta az osztott kormányzás lehetőségét, sőt, az esetenkénti aggregálás eredményeként ténylegesen is létrejött az osztott kormányzás. Viszont a kialakult végeredmény, az RRD állapot csak a negyedik legnépszerűbb kombináció lett, tehát az egyenkénti és a kombinált aggregálás által kihozott eredmények elég messze kerültek egymástól. Mivel nem a legrosszabb kombinált eredmény lett az egyéni győztes, ezért a szerzők szemiparadoxonnak nevezték el ezt a konkrét eredményt. A valóságos történelmi példa tehát nem az elmélet által előírt, “tiszta"" módon valósult meg, de a dolgok paradox jellegét azért még így is felismerhettük a gyakorlatban. Bramsék hoztak még egy harmadik példát a törvénykezési gyakorlatból, de ezt már nem mutatom itt be.

A többszörös választási paradoxonhoz hasonlít az Ascombe-paradoxon, de mivel a szakirodalomban önálló néven hivatkoznak rá, én is külön fejezetben mutatom be.

4.2.18 Anscombe-paradoxon

A G.E.M. Anscombe nevével jelzett paradoxon a többségi elv egy furcsa deficitjére mutat rá [21]. A jelenség szemléltetésére Anscombe adott egy példát, amelyben tizenegy szavazó (A,B, . . . , K) tizenegy kérdésről szavazott. Az eredményt mutató táblázatban az *igen* szavazatokat 1, a *nem*-eket 0 jelöli. Mivel a többségi szavazással hozott csoportdöntés minden esetben *igen* (1) volt (6:5 arányban), ezért azt mondhatjuk, hogy az egyéni *igen* (1) szavazatok sikeres voltak, mivel a csoportdöntéssel megegyeztek, az egyéni *nem*-ek pedig sikertelenek.

	A	B	C	D	E	F	G	H	I	J	K	csoportdöntés
1. kérdés	0	1	1	1	1	0	0	1	1	0	0	1
2. kérdés	0	1	0	1	1	0	0	1	0	1	1	1
3. kérdés	0	1	0	0	1	0	0	1	1	1	1	1
4. kérdés	0	1	0	0	0	1	1	0	1	1	1	1
5. kérdés	0	0	0	0	1	1	1	0	1	1	1	1
6. kérdés	0	0	0	0	0	1	1	1	1	1	1	1
7. kérdés	0	0	0	0	0	1	1	1	1	1	1	1
8. kérdés	1	0	0	0	0	0	1	1	1	1	1	1
9. kérdés	1	0	1	0	0	0	0	1	1	1	1	1
10. kérdés	1	0	1	1	0	0	0	1	1	0	1	1
11. kérdés	1	0	1	1	0	0	0	1	0	1	1	1
sikeres szavazat	4	4	4	4	4	4	5	9	9	9	10	

A táblázat legalsó sorából kiolvasható, hogy heten voltak olyanok (A-G), akiknek a szavazata az esetek többségében sikertelen volt. Ez azt jelenti, hogy a többségi szavazás végeredményeként kisebbségbe kerülhetnek azok, akik a többségi vélemények többségével egyetértenének, vagy másként: többen lehetnek olyanok, akiknek a szavazata az esetek többségében sikertelen lesz. Ez az Anscombe-paradoxon lényege.

Jonathan Lamb Gorman válaszcikkében egy jóval egyszerűbb példát javasolt [303], ami áttekinthetőbbé tette az esetet leíró táblázatot is, ezért inkább azt használom a továbbiakban (még pontosabban: Gorman példáját a Carl Wagner által átrendezett formában veszem át [803], mivel az alakilag jobban illeszkedik az itt használt megoldásokhoz). Gorman példájában három alternatíváról (A, B, C) kell szavazni $(igen/nem)$ öt szavazónak, többségi eljárással. A szavazás eredményét mutató táblázatban a ':-)' szimbólum azt jelzi, hogy az adott szavazó adott szavazata megegyezik-e az adott kérdésben hozott többségi döntéssel, vagyis sikeres volt-e a szavazat.

kérdések:	*A*	*B*	*C*
1. szavazó véleménye	igen :-)	igen	nem
2. szavazó véleménye	nem	nem :-)	nem
3. szavazó véleménye	nem	igen	igen :-)
4. szavazó véleménye	igen :-)	nem :-)	igen :-)
5. szavazó véleménye	igen :-)	nem :-)	igen :-)
többségi vélemény	igen	nem	igen

Ebben a táblázatban is a többséggel való egyetértést mutató jel eloszlása zavarbaejtő. Az első, második és harmadik szavazó csak egy-egy esetben került be a többségi véleményt osztók csoportjába, és ők, hárman, ebben a "kisebbségi minőségükben" többségben vannak a maradék kettő "igazi többségi" szavazóval szemben, akik több sikeres szavazatot adtak le, mint sikertelent.

> „When the majority get what they want, they do so at the expense of the minority's frustration, but at least nobody is privileged ... a method which derives the upshot from an equal weight attached to each person's desire must achieve the satisfaction of the desires of the greatest possible number and so be both the best and fairest." [21, 163-164]

Carl Wagner megoldása: 3/4 [803], [804]

Még durvább a helyzet, ha az előző táblázatban bemutatott szavazói vélemények létezését feltételezve és képviseleti demokráciát "szimulálva" olyan jelöltekre kell szavazni, akik a három kérdésben előre kinyilvánítják a preferenciáikat, és ez alapján szavaznak rájuk a szavazók. Tegyük fel, hogy két jelölt van, az alábbi "ígéretekkel".

kérdések:	*A* kérdés	*B* kérdés	*C* kérdés
1. jelölt meghirdetett profilja	igen	nem	igen
2. jelölt meghirdetett profilja	nem	igen	nem

A táblázatból az olvasható ki, hogy az első jelölt a csoporttagok véleményei alapján megállapítható többségi eredmények képviseletét ígéri, míg a második jelölt mindenben pont az ellenkezőjét. Azt várhatnánk, hogy ilyen helyzetben a többségi véleményt képviselő jelöltre érkezik több szavazat, de nem ez a helyzet. A csoportból ugyanis az első három tag két kérdésben inkább egyetért a második jelölt nézeteivel, ezért rá szavaz, és ez már többséget biztosít annak a képviselőjelöltnek, aki minden kérdésben a többségi véleménnyel ellentétes álláspontot hirdet.

Lehetne talán azt gondolni, hogy az osztott kormányzás paradoxona mint a többkérdéses választás paradoxonának egyik altípusa (↬245.o.) és az Anscombe-paradoxon megegyezik egymással, de ez nem így van. Hannu Nurmi mutatott be egy olyan esetet, ami példázza az előbbit, de nem igaz az utóbbira [542].

	1. kérdés	2. kérdés	3. kérdés
1. szavazó véleménye	Y :-)	Y	Y :-)
2. szavazó véleménye	X	X :-)	Y :-)
3. szavazó véleménye	X	Y	Y :-)
4. szavazó véleménye	Y :-)	X :-)	X
5. szavazó véleménye	Y :-)	X :-)	X
többségi vélemény	Y	X	Y

Látható, hogy a többesetes választás paradoxona fennáll, hiszen nyertes kombináció (YXY) egyszer sem fordul elő az egyéni vélemények között, viszont az Anscombe-paradoxon nem teljesül, mivel majdnem mindenki (négyen az öt közül) két kérdésben (tehát többségében) a többségi véleményt vallja.

4.2.19 Osztrogorszkij paradoxon

Noha Mojszej Jakovlevics Osztrogorszkij[31] már az 1900-as évek elején írt arról a paradoxonról, amit később róla neveztek el, a társadalmi választások elméletével foglalkozó kutatók csak a hetvenes években figyeltek fel erre a problémára [542]. A paradoxon lényege a következő. Legyen három kérdéskör, és négy választói csoport (A, B, C, D), amelyek a szavazók 20-20-20-40 százalékát reprezentálják valamint két párt (X és Y). Tegyük fel, hogy a három kérdésben a választói csoportok az alábbi táblázat szerint szavaznak a két párt (programjai) közül valamelyikre.

[31] Az angolul többféleképpen is hivatkoznak Osztrogorszkij nevére, például így: Moisey Yakovlevich Ostrogorsky/Ostrogorski

csoport	1. kérdés	2. kérdés	3. kérdés	támogatott párt
A (20%)	*X*	*X*	*Y*	*X*
B (20%)	*X*	*Y*	*X*	*X*
C (20%)	*Y*	*X*	*X*	*X*
D (40%)	*Y*	*Y*	*Y*	*Y*
győztes párt (1)	-	-	-	*X*
győztes párt (2)	*Y*	*Y*	*Y*	*Y*

Az ellentmondás akkor és abban érhető tetten, amikor dönteni kell arról, hogy melyik párt a győztes az összes kérdés figyelembe vétele mellett. Ha az egyes csoportokra nézzük meg, melyik pártot választanák, akkor az első három csoport esetében az *X*, a *D* csoport esetében az *Y* párt nyerne, és ezek alapján 60:40 arányban az előbbi lenne a végső győztes. Ha a három kérdésben egyenként hoznánk ki részgyőzteseket, akkor mindegyik kérdésben az *Y* párt jönne ki győztesként, s így természetesen a végső győztes is az lenne. Th. Bezembinder és P. Van Acker a kétféle aggregálási módszer megnevezésére javasolta bevezetni az IP-, illetve PI-egyesítés (IP/PI-amalgamation) kifejezéseket ahhoz igazodóan, hogy előbb a kérdések (Issue), majd a személyek (Person) szerint aggregálunk-e vagy fordítva [79]. Az nem kérdés, hogy a kétféle aggregálási módszert érdemes elkülöníteni és egyértelműen megnevezni, de a Bezembinderék által javasolt rövidítések egyfelől már "foglaltak" a tranzitivitással, indifferenciával kapcsolatos elméletek világában, másfelől a diszkurzív paradoxon jelenségével való párhuzamosság miatt inkább az ott bevezetett megkülönböztetést használom itt is (↬253.o.). Az egyes kérdések szerinti összegzés összegzését IBI-módszernek nevezem (az 'issue-by-issue' kifejezésre utalva), míg azt a módszert, amikor előbb az emberek, csoportok, esetek szerint összegeznek, majd az így kialakult részeredményeket vetik össze a második körben, azt nevezem CBC-metódusnak (a 'case-by-case' kifejezés kezőbetűire hivatkozva ezzel).

Hannu Nurmi mutatott egy olyan esetet, ami mind az Osztrogorszkij-paradoxonra, mind a többkérdéses választás paradoxonára (↬245.o.) például szolgálhat [542]. Az osztott kormányzás korábban bemutatott problémáját szemléltetik az alábbi adatok. Két párt, a demokraták (D) és a republikánusok (R) jelöltjeire szavaz 13 szavazó három pozíció betöltése kapcsán (képviselőházi és szenátusi tagság, valamit a kormányzói poszt elnyerése a tét).

szavazók	képviselőház	szenátus	kormányzóság
1-3 szavazó	D	D	D
4 szavazó	D	D	R
5 szavazó	D	R	D
6 szavazó	D	R	R
7-9 szavazó	R	D	R
10-12 szavazó	R	R	D
13 szavazó	R	R	R

Az adatokból kitűnik, hogy a három poszt közül egyet a republikánus párt szerez meg (képviselőház), illetve kettőt a demokrata párt, vagyis RDD kombináció születik. Erre egyénileg senki sem szavazott, tehát fennáll a többesetes választás paradoxona. Viszont meg lehet állapítani az Osztrogorszkij-paradoxont is, hiszen a kétféle aggregálási módszerrel egyszer a demokrata, másszor a republikánus párt jön ki "abszolút" győztesnek.

Ennek ellenére a két paradoxon nem azonos egymással, amit önmagában szemléltet az a tény, hogy az Osztrogorszkij-paradoxonra hozott első esetre nem igaz az osztott kormányzás paradoxona, hiszen az egyenkénti összegzéssel kihozott nyertesek (YYY) kombinációja egyben a legtöbb egyéni szavazatot kapta (40%).

[573]

[533]

Douglas W. Rae és Hans Daudt szerint minden Osztrogorszkij-paradoxon háttérében van egy Condorcet-paradoxon [588], amit közel sem tartottak véletlennek. Szerintük a két jelenségnek közös gyökere van, amire vonatkozóan a következő sejtést fogalmazták meg.

> „. . . in two properties shared by each paradox: (i) each has as its "input" some binary relation comparing pairs of objects, yet (ii) each has its "output" some list or set of sets containing three or more objects. Each paradox concerns a relation of pairwise decision or observation functioning in an n-wise world. It is possible to prove that any such pattern must produce surprises? . . .
> Something of the sort is done in Arrow's celebrated "general possibility theorem", since the condition he calls "independence of irrelevant alternatives" restricts attention to pairwise composition, yet the assumption of "universal domain" requires the consideration of n-wise decisions." [588]

Már az Anscombe-paradoxonra is igaz volt, az Osztrogorszkij-paradoxonra még inkább elmondható, hogy ezek a problémák már nem csak, vagy talán nem is elsősorban a preferenciaaggregálásról szólnak, hanem valami másról. Itt már nem a világállapotok elrendezéseiről, értékeléseiről van szó, hanem inkább a vélemények összesítéséről, amit már az aggregálási problémák egy másik típusaként lehet elemezni, és amit én is külön fejezetben mutatok be.

4.3 Véleményaggregálás

kijelentésaggregálás véleményaggregálás nézetaggregálás

4.3.1 Doktrinális paradoxon

Miközben Wijnand J. Schoenmakers egy 1986-os kétoldalas cikkében az adatbázisok egyesítésének veszélyeire akarta felhívni a figyelmet, a probléma forrását egy olyan képzeletbeli példával szemléltette, ami teljes mértékben alkalmas volt arra, hogy a véleményaggregálás problámájára is rá lehessen mutatni vele [642]. Schoenmakers történetében egy bíró két tanút kérdezett ki különböző időpontokban. Az első tanú azt állította, hogy az A állítás igaz, a második pedig azt, hogy $A \to B$. A két kijelentés alapján a bíró levonta azt a következtetést, hogy B is igaz (és erre hivatkozva halálra ítélte a vádlottat). Később kiderült, hogy mindkét tanú hamisnak tartotta a B állítást, bár más indokok alapján. Az első tanú nem gondolta sem B-t, sem $(A \to B)$-t igaznak, de mivel a bíró ezt nem kérdezte tőle, nem szólt semmit erről. A második tanú pedig úgy vélte, hogy az $A \to B$ állítás igaz ugyan, de tudta azt is, hogy a B nem. Ebből levonhatta magának a következtetést, hogy akkor az A nem lehet igaz, de ezt magától értetődőnek tartotta, ezért ezt nem is mondta a

bírónak. A két tanútól átvett állítások alapján a bíró helyes logikai következtetést vont le, ám mégis rossz ítéletet hozott.

Példájával Schoenmakers azt akarta szemléltetni, hogy a különböző forrásokból származó adatok egyesítése könnyen ellentmondásokhoz vezethet. A mérnökök, informatikusok számára azért volt fontos mindez, mert az ilyen paradoxon lehetősége komoly fenyegetést jelenthet az adatbázisok egyesítése során, mivel az egyébként önmagukban konzisztens adatbázisok az egyesítés után (eredményeként) inkonzisztenssé válhatnak. A történet tanulsága azonban messze túlmutat az adatbázisok egyesítésének problémáján. Ugyanez a probléma rejlik például a doktrinális paradoxon mélyében.

A **doktrinális paradoxon** (doctrinal paradox) vagy más néven bírói paradoxon problémáját Kornhauser és Sager fedezte fel [410]. Kornhauserék felismerése az volt, hogy ha egy ügyben nem egyetlen, hanem több bíró dönt valamiért (ahogy Amerikában a fellebbviteli bírósági tárgyalások esetében ez így van), és többségi eljárással összegzik a bírók véleményeit, akkor könnyen paradoxonhoz juthatunk. Nem mindegy ugyanis, hogy a döntéshozatal melyik fázisában összegzik az egyéni véleményeket. Az egyik megoldás szerint előbb közösen tisztázni kell a tényállást. Ehhez a büntethetőséget implikáló egyes feltételekről kell közös véleményt kialakítani, és minden tényállási kérdésben többségi szavazást kell alkalmazni. Az egyes feltételekre így megállapított közös vélemény alapján lehet aztán a büntetésre vonatkozó konklúziót levonni. Létezik azonban egy másik módszer is, ami szerint mindenkinek magának kell döntenie a tényállási feltételeről, és a közösségi döntésre csak – az egyéni döntések alapján – a bűnösség kimondásakor, az ítélethozatal során van szükség.

Egy egyszerű példa segítségével szemléltethető, hogy miként juthatunk így ellentétes eredményekhez [410]. Tegyük fel, hogy egy olyan ügyben kell dönteni, amelyben két szempont (A és B) alapján kell megállapítani, hogy bűncselekmény történt-e, amit szankcionálni kell (S) vagy sem ($\neg S$). A két feltétel és a szanció legyen a következő:

A: x és y között érvényes a szerződés
B: x megszegte a szerződést
S: x kártérítésre kötelezett

A döntés logikai szerkezete egyértelmű. Ha A és B feltételek fennállnak, akkor meg kell állapítani a bűncselekmény bekövetkeztét és ki kell szabni a S büntetést. Nézzük meg, mi történik akkor, ha három bírónak kell közösségi döntést hoznia ebben az ügyben.

döntési alternatívák:	A	B	$A \wedge B \rightarrow S$
1. bíró véleménye	igen	nem	nem
2. bíró véleménye	nem	igen	nem
3. bíró véleménye	igen	igen	igen
közösségi vélemény (1)	igen	igen	–
közösségi vélemény (2)	–	–	nem

A paradoxon léte szembeszökő. Ha a premisszák esetében alkalmazzuk a többségi szavazást (1), akkor mindkét feltételre igenlő választ kapunk, ami alapján kötelezően ki kell mondani a bűnösséget és elmarasztaló ítéletet kell hozni (S). Ha az ítélethozatal során támaszkodunk a közösségi döntésre (2), akkor nemleges válaszhoz jutunk a bűnösséget

illetően, ami felmentő ítélet eredményez ($\neg S$). Kornhauser és Sager az első esetet **kérdésről kérdésre szavazásnak** (issue-by-issue voting) nevezi. Ekkor minden egyes eldöntendő kérdésről szavazást tartanak, majd az így kialakított közös álláspont alapján vonják le (akkor már a logikai következtetés segítségével) a végleges döntést. A másodikat esetet (ami a konklúzió alapú döntési helyzetnek felel meg) **ügyről ügyre szavazás** (case-by-case voting) címkével illetik, mivel ekkor minden bíró az ügy egészéről mond ítéletet, és aaztán ezeket összegzik többségi szavazással [410]. Azért nevezték el doktrinális paradoxonnak, mert az $A \wedge B \rightarrow S$ formulával reprezentálható jogi doktrinát (legal doctrine) tartották a bírói döntéshozatal alapzatának.

Franz Dietrich is bemutat egy paradoxont, de az ő példájában más a döntésbe vont, egymáshoz kapcsolódó kijelentések viszonya [206]. Képzeljük el azt a helyzetet, amikor egy kormányülésen kell az alábbi ügyben dönteni (amikor feltétezhetjük, hogy a vélemények megoszlását tekintve a kormány egésze három, egyenlő létszámú csoportra osztható).

A: megengedhető a költségvetési hiány
B: az oktatásra többet kell költeni
$A \rightarrow B$: *ha* megengedhető a költségvetési hiány, *akkor* az oktatásra többet kell költeni

A döntés logikai szerkezete itt más lesz, mint az előző példában, ha úgy viszik be a döntéshozó testület elé az ügyet, hogy az A és a $A \rightarrow B$ kérdéseket jelölik meg premisszaként és a B kérdést konklúzióként. Tegyük fel, hogy az alábbi szavazatokat kapjuk:

döntési alternatívák:	A	$A \rightarrow B$	B
1. csoport véleménye	igen	igen	igen
2. csoport véleménye	igen	nem	nem
3. csoport véleménye	nem	igen	nem
közösségi vélemény (1)	igen	igen	–
közösségi vélemény (2)	–	–	nem

A táblázat alapján látszik, hogy ugyanúgy paradoxonhoz jutottunk, mint az előző példában: a kétféle aggregálási technika alkalmazásával eltérő végeredményhez juthatunk. Mindhárom véleménycsoport konzisztens a következtetéseiben, a véleményösszegzés különböző útjai mégis ellentmondáshoz vezetnek. Ez a példa egyébként azt a jelenséget is szemlélteti, hogy a szavazási napirend kialakításával befolyásolhatjuk a végeredményt.

A doktrinális paradoxon lényege az, hogy ha a döntéshozók (politikusok, bírók, szakértők) egy csoportjának logikailag összefüggő állításokról kell közös döntést hoznia, akkor előfordulhat, hogy ha az ügybe vont kijelentésekről (premisszákból) egyenként képeznek közös (többségi elvű) döntéseket, majd ezek alapján vonják le a végkövetkeztetésüket, illetve a premissza-kijelentések alapján levont egyéni konklúziókból közös (többségi elvű) véleményt alkotnak, akkor a kétféle módon képzett közös döntések ellentmondásba kerülhetnek egymással. E rövid jellemzés során azonban már olyan terminusokat használtam, amelyeket a későbbi – a jelenség általánosítását elvégző és a diszkurzív paradoxon fogalmát középpontba helyező – elméletek termeltek ki. Nézzük meg, hogyan épült tovább a “paradoxonok elmélete”.

4.3.2 Diszkurzív paradoxon, List-Pettit lehetetlenségi tétele

2002-ben Christian List és Philip Pettit a doktrinális paradoxon "belső logikáját" elemezve, és a problémát általánosítva bevezette a diszkurzív paradoxon fogalmát, majd megfogalmazott egy lehetetlenségi tételt, amire azóta List-Pettit tételként hivatkoznak a szakirodalomban [452]. A vizsgált szakterület is saját nevet "kapott": egyre inkább úgy hivatkoztak erre a vizsgálati irányra, mint a **véleményaggregálás** (judgement aggregation) elemzésével foglalkozó tudományterületre.

Egy 2002-es cikkben List és Pettit Kornhauserék doktrinális paradoxonát értékelve rögtön bemutattak egy másik példát, amely a doktrinális paradoxont "kitermelő" élethelyzethez képest teljesen eltérő világból, a gazdaság területéről származott. Ezzel a kiterjesztéssel azt akarták szemléltetni, hogy az élet minden területén felbukkanhat a doktrinális paradoxonhoz hasonló jelenség. A példában szereplő kisvállalat dolgozóinak, akik egyben tulajdonosai is a cégnek, közösen kell dönteniük arról, lemondjanak-e fizetésük egy részéről annak érdekében, hogy egy biztonsági berendezést felszereljenek a cég telephelyén. Három szempontot mérlegelhet mindenki egyéni döntése során: *(i)* mennyire komoly az elhárítandó veszély, *(ii)* mennyire ígérkezik hatékonynak a berendezés, illetve *(iii)* személy szerint képes-e viselni a lemondás pénzügyi terheit. Ezek alapján – feltételezve, hogy hárman vesznek részt a döntésben – az alábbi táblázat szerint összegezhetjük az egyéni és közösségi döntéseket (itt is elég az *igen* vagy *nem* döntéseket feltüntetni).

szempontok:	komoly veszély	hatékony berendezés	viselhető anyagilag	lemond-e
1. személy véleménye	igen	nem	igen	nem
2. személy véleménye	nem	igen	igen	nem
3. személy véleménye	igen	igen	nem	nem
közösség véleménye (1)	igen	igen	igen	–
közösség véleménye (2)	–	–	–	nem

A paradoxon egyértelműen kiolvasható a táblázatból. Ha a tagok egyénileg döntenek, és ez alapján aggregálnak közösségi döntést, akkor mindegyik döntéshozó számára létezhet olyan szempont, amely szerint nem támogatja a tehervállalást, és emiatt a három – a beruházást ellenző – egyéni döntésből származtatott közösségi döntés is csak elutasító tartalmú lehet. Miközben minden egyes szempont alapján többségi vélemény volt arról köztük, hogy az adott szempont szerint támogatni kellene a beruházást.

premissza alapú megközelítés (premise-based approach)

konklúzió alapú megközelítés (conclusion-based approach)

A véleményaggregálási folyamattal, techikával szemben két szinten lehet követelményeket megfogalmazni. Egyfelől rögzíteni lehet racionalitási feltételeket az egyéni szavazókkal szemben, másfelől elvárásokat lehet megfogalmazni a közösségi aggregálással szemben. Az egyéni racionalitásra vonatkozó a következő racionalitáskritériumokat fejezhetünk ki.

X tartalmazza minden ϕ_i propozíció negáltját is: $\phi_i \in X \leftrightarrow \neg\phi_i \in X$

X tartalmaz legalább kettő elemi propozíciót (A, B), illetve ezek konjunkcióját $(A \wedge B)$.[32]

Φ_i véleményprofil tartalmazza mindazon propozíciókat, amelyeket igaznak fogad el az i. szavazó

[32] A szerzők megjegyzik, hogy bármely más logikai kapcsoló is szerepelhetne a konjunkció helyén.

(AL.72) $\text{COMPLETE}(\Phi_i) \stackrel{\text{def}}{=} \forall i \forall k (\phi_k \in \Phi_i \vee \neg\phi_k \in \Phi_i)$

az i. szavazó Φ_i véleményprofilja **teljes**, ha tartalmaz minden propozíciót vagy annak negáltját (akár mindkettőt).

(AL.73) $\text{CONSISTENT}(\Phi_i) \stackrel{\text{def}}{=} \forall i \forall k \neg(\phi_k \in \Phi_i \wedge \neg\phi_k \in \Phi_i)$

az i. szavazó Φ_i véleményprofilja **konzisztens**, ha egy propozíció és annak negáltja közül csak az egyiket tartalmazza.

(AL.74) $\text{DEDUCTIVELY_CLOSED}(\Phi_i) \stackrel{\text{def}}{=} \forall i \forall k (\Phi_i \vdash \phi_k \rightarrow \phi_k \in \Phi_i)$

ahol $\vdash$ a szintaktikai következmény jele; az i. szavazó Φ_i véleményprofilja **deduktíve zárt**, ha valamely ϕ_k propozíció logikailag levezethető a véleményprofilba tartozó propozíciókból, akkor a propozíció is beletartozik a véleményprofilba.

A teljességi feltétel azt biztosítja, hogy mindenki állást foglaljon valahogy az adott véleménytér minden propozíciókettősével kapcsolatban, de ez az elvárás önmagában még nem eredményez racionális véleményprofilt, hiszen ez a kritérium "megengedi", hogy valaki egyszerre elfogadon egy állítást és annak tagadását. Ezt "korrigálja" a konzisztenciakritérium, amely a racionális véleménynek csak az ellentmondó állítások lehetőségét kizáró profilt tartja. A deduktív zártság feltétele pedig azt "garantálja", hogy a szavazók egyénileg mind képesek levonni a logikailag szükséges következtetéseket. A három fenti tulajdonság alapján már megfogalmazhatjuk mind az egyéni, mind a csoport véleményprofilokra vonatkozó racionalitáskritériumot.

(AL.75) $\text{RATIONAL}(\Phi_i) \stackrel{\text{def}}{=} \text{COMPLETE}(\Phi_i) \wedge \text{CONSISTENT}(\Phi_i) \wedge \text{DEDUCTIVELY_CLOSED}(\Phi_i)$

egyénileg racionális (HUN) – individually rational (ENG)

A Φ_i egyéni véleménykészlet (personal set of judgement) racionális, ha teljes, konzisztens és deduktíve zárt.

Az aggregált (csoportos, közösségi, kollektív) véleményprofilra vonatkozóan is ugyanezt a racionalitáskritériumot követelhetjük meg:

(AL.76) $\text{RATIONAL}(\Phi) \stackrel{\text{def}}{=} \text{COMPLETE}(\Phi) \wedge \text{CONSISTENT}(\Phi) \wedge \text{DEDUCTIVELY_CLOSED}(\Phi)$

csoportosan racionális (HUN) – group ??? rational (ENG)

A Φ csoportos véleménykészlet (collective set of judgement) racionális, ha teljes, konzisztens és deduktíve zárt.

Ez azonban kevés, meg kell fogalmaznunk azt az elvárást is, miszerint az aggregáló függvénynek is racionálisnak kell lennie. Ehhez az univerzalitás és anonimitás tulajdonságainak (↬**??**.o.) érvényességén túl még egy további feltételt ki kell kötnünk.[33]

(AL.77) $\text{SYSTEMATIC}(\mathbf{F}(\Phi)) \stackrel{\text{def}}{=} \forall \Phi, \Phi' \in \mathbb{D}_F : \forall \phi, \psi \in X : (\forall i : \ \phi \in \Phi_i \leftrightarrow \psi \in \Phi_i') \rightarrow (\phi \in F(\Phi) \leftrightarrow \psi \in F(\Phi_i'))$

szisztematikus (HUN) – systematic (ENG)

Az $\mathbf{F}(\Phi)$ függvény szisztematikus, ha Ezen tulajdonságok segítségével már megfogalmazható a List-Pettit tétel, amely a fenti racionalitáskritériumok egyidejű kielégíthetetlenségét állítja [452].

[33] List és Pettit az 'univerzalitás' terminusa helyett az 'univerzális értelmezési tartomány' (universal domain) kifejezést használja, de ugyanabban az értelemben, ahogy azt korábban már bemutattam.

(AL.78) $\exists i(\neg(\text{RATIONAL}(\Phi)\wedge\text{RATIONAL}(\Phi_i)\wedge\text{UNIVERSAL}(\mathbf{F})\wedge\text{ANONYM}(\mathbf{F})\wedge\text{SYSTEMATIC}(\mathbf{F})))$

List-Pettit tétel (HUN) – List-Pettit theorem (ENG)

Nem létezik olyan **F** véleményaggregáló függvény, amely univerzális, anonim és szisztematikus, valamint minden racionális egyéni véleményprofilt racionális kollektív véleményprofillá összegez.

A tétel nagyon erőset állít, és úgy tűnhet, mintha a társadalmi gyakorlat nem is igazolná vissza az igazát. List és Pettit válasza erre az, hogy a ténylegesen létező, ebben az értelemben működőképes véleményösszegzési eljárások során mindig megsértünk valamilyen racionalitási elvet a fent definiáltak közül. Ha így van, akkor a következő fontos kérdés az lehet, hogy milyen módonkon lehet enyhíteni az elvárásokon, hogy még "vállalható" racinonalitású eljárásokat kapjunk.

Pauly és van Hess [564] Pauly-van Hees tétel

(AL.79) a

Pauly-van Hees tétel (HUN) – Pauly-van Hees theorem (ENG)

A Pauly-van Hees tétel tétel szerint tehát egy véleményösszegző függvény, amely kielégíti az univerzális értelmezési tartomány és a szisztematicitás feltételeket, akkor és csak akkor összegez racionális véleményprofilokat racionális kollektív véleményhalmazzá, ha diktatúra valamely $i \in \mathbb{N}$ individuumra.

mi a List-Petit tétel és az Arrow-tétel viszonya [453]

Majority voting on restricted domains [208]

4.3.3 Aggregálás általában

Általánosítás

Robert Wilson: „It is natural to ask whether procedures for aggregating attributes other than preferences are subject to similar restrictions. This paper answers the question by proving the counterpart of Arrow?s result in the context of a general theory of aggregation." [828]

Rubinsteint és Fishburn Algebraic aggregation theory [623]

Elad Dokow and Ron Holzman Aggregation of binary evaluations [211]

Klaus Nehringa és Clemens Puppe Abstract Arrowian aggregation [531]

———————————————————-

4.3.4 Liberális paradoxon

Amartya Sen liberális paradoxon [663]

James M. Buchanan kritikája [128]

Gibbardnak is van példája a liberális paradoxonra, Suzumura is elemzi a könyvében

Málik: Sen + perverz effektusok Bernholz-tétel - Externalities as a Necessary Conditions for Cyclical Social Preferences [75]

áttekintés: List, Synthese [451]

List: véleményaggregálás általánosabb, mint a preferenciaaggregálás [207]

4.3.5 Multiple elections

The paradox of multiple elections [114]

4.4 Másodrendű preferenciák

intranzitív indifferencia kezelésére Wagnerék másodrendű megoldást javasolnak,

[435]

matematikában van két részben rendezett halmaz, A és B. A két halmazon értelmezett $A \times B$ Descartes-szorzat ????, illetve lexikografikus rendezés a következőképpen definiálható.

(AL.0) $\text{PRODUCT_ORDER}((a_1, b_1), (a_2, b_2)) \equiv (a_1, b_1) \leq (a_2, b_2) \stackrel{\text{def}}{=} a_1 \leq^* a_2 \wedge b_1 \leq' b_2$

sorrendezés (HUN) – product order (ENG) – archimédeszi rendezés (HUN)

(AL.1) $\text{LEXICOGRAPHIC_ORDER}((a_1, b_1), (a_2, b_2)) \equiv (a_1, b_1) \leq (a_2, b_2) \stackrel{\text{def}}{=} a_1 <^* a_2 \vee (a_1 = a_2 \wedge b_1 \leq' b_2)$

lexikografikus rendezés (HUN) – alfabetikus rendezés (HUN) – ábécé szerinti rendezés (HUN) – lexicographic order (ENG) – dictionary order (ENG)

archimédeszi rendezés Lilly Görke, Halmazok, relációk, függvények, 275.o.

Preferenciaaggregálás mint másodrendű preferenciarendezés

Amikor a preferenciánkat fejezzük ki, akkor valamilyen szempont szerint valamely alternatívákon végrehajtunk egy preferenciarendezést. Eddig nem volt arról szó, hogy mit is jelent az, amire most valamilyen "szempont szerint"-ként hivatkoztam. De nem is az a lényeges, hogy meghatározzuk a szempontot, hanem inkább az a felismerés fontos, hogy könnyen előfordulhat és gyakran elő is fordul, hogy ugyanazt az alternatívahalmazt több szempont szerint is értékelhetjük (rendezhetjük). Ilyenkor azonnal felmerül a kérdés, hogy lehet egyszerre több szempontot is figyelembe venni. Mindez azért különösen fontos és nehéz kérdés, mert feltételezhető, hogy a különböző szempontok szerint más és más sorrendhez lehet jutni a rendezések során. Ilyenkor mi a teendő?

Az fontos "technikai" kérdés, hogy milyen módon lehet a preferenciarendezést végrehajtani több elem (és/vagy több szempont) esetén, erre később még visszatérek.

Korábbban már jeleztem, hogy többféle aggregálásról beszélhetünk, és a preferenciaaggregálást az aggregálás egyik típusaként határozhatjuk meg (↬191.o.).

preferenciaaggregálás (preference aggregation) és nem véleményaggregálás (judgement aggregation)

a közgazdaságtanban: **társadalmi jóléti függvény** (social welfare function)

az aggregálási operátoron belül az egyes preferenciarendezéseket komponenspreferenciaként is meghivatkozhatjuk

[199]

[30]

[561]

Where the preferences come from? [210]

[209]

alternatívák alternatívák tulajdonságai az ágens motivációs állapota motivációs szembeszökőség, szaliencia (motivationnally salient)

Arrow lehetetlenségi tétele a preferenciákon és azok aggregálásán alapul [31]

univerzális értelmezési tartomány gyenge Pareto-elv függetlenség a lényegtelen alternatíváktól diktátor-mentesség

✠ ⁎ ✠

4.4.1 Archimédeszi preferencia

archimédeszi preferencia (nincs végtelen nagy, illetve végtelen kicsi eleme, nincs ugrás a függvényben)

hasznosság

hasznosság modern fogalmát Jennings (1955) dolgozta ki

Jennings, R. (1955): Natural elements of political economy. Longmans, London. (hivatkozza: Georgescu - Roegen in. International encyclopedia of the social sciences, MacMillan Co. Free Press, New York, 1968)

Az arkhimédeszi preferencia meghatározásakor a preferenciarelációt fogyasztói kosarakra (matematikailag vektorokra) alkalmazzák, tehát az összehasonlítást, értékkifejeződést nem egyedi entitásokra, hanem entitások (mégpedig különböző típusú entitások) halmazaira vonatkoztatva fejezik ki. Erre a „módosított" preferenciarelációra írják elő az arkhimédeszi axióma érvényesülését, amely a következőképpen formalizálható (az arkhimédeszi tulajdonsággal rendelkező preferenciarelációt az 'APREF', a X és Y termék különböző mennyiségeit x_i és y_j szimbólumokkal jelölve):

a formula forrása [181]

(AL.2) $$\text{ARCHIMEDEAN_PREFERENCE} \stackrel{\text{def}}{=} \forall i \forall j \exists k (\mathbf{R}((x_i, y_i), (x_j, y_j)) \rightarrow \mathbf{R}((x_j, y_k), (x_i, y_i)))$$

archimédeszi preferencia (HUN) – archimedean preference (ENG)

A képlet azt fejezi ki, hogy bárhogy is állítunk össze két csomagot az X és Y termékekből, amelyek közül az egyik csomag (x_i, y_i) többet ér a másiknál (x_j, y_j), ha az Y jószág mennyiségét növeljük (y_j-t kicseréljük y_k-ra), akkor megfordulhat a preferenciareláció iránya, azaz (x_j, y_k) csomag értékesebbé válhat az (x_i, y_i) csomagnál.

A felhasználónak azonban lehetnek olyan preferenciái is, amelyekre nem teljesül az arkhimédeszi axióma. Ezek közül a legfontosabbak a *lexikografikus preferenciák*. Ha van két termékünk (szolgáltatásunk vagy bármink, amire a preferenciareláció alkalmazható), akkor azok különböző mennyiségeit egy-egy vektorral reprezentálhatjuk: $(x_1, x_2, \ldots, x_i)$, $(y_1, y_2, \ldots, y_j)$, ahol x_i az X jószág csökkenő nagyság szerinti sorrendbe állított i. egyede, míg y_j az Y jószág j. példánya. Ebben az esetben a lexikografikus preferencia fogalmának formális meghatározását a következőképpen adhatjuk meg (a lexikografikus tulajdonsággal rendelkező preferenciarelációt 'LPREF'-fel jelölve):

a formula forrása [181]

(AL.3) $$\text{LEXICOGRAPHIC_PREFERENCE} \stackrel{\text{def}}{=} \forall j \exists i (j < i \rightarrow (\mathbf{R}(x_j, y_j) \wedge \mathbf{R}(y_j, x_j) \wedge \text{LPREF}(x_i, y_i)))$$

lexikografikus preferencia (HUN) – lexicographic preference (ENG)

A formula értelmezésekor ki kell használnunk azt a feltételt, hogy a két jószágból vett különböző mennyiségeket – nagyság szerint csökkenő – sorrendbe állítva jelöljük x_i-vel és y_j-vel. Így a fenti képlet azt fejezi ki, hogy akármekkora mennyiséget veszünk is a két

jószágból, azaz vehetjük akárhányadik (j.) elempárt a sorrendbe rakott termékekből, a két terméksor j. elemeit megelőző párjaira még lehet döntetlen az állás köztük (ezt fejezi ki az $\text{LPREF}(x_j, y_j) \wedge \text{LPREF}(y_j, x_j$ részformula), de az i. elemtől kezdve már az X termék a preferáltabb az Y-nal szemben ($\text{LPREF}(x_i, y_i)$). Ezt azt is jelenti egyben, hogy az X termék „összességében" kívánatosabb a másikhoz viszonyítva.

Azért nevezik ezt a rendezést lexikografikus rendezésnek, mert a szavak ábécé szerinti sorrendjét is ilyen logika szerint alakítjuk ki. Az 'a' betűvel kezdődő szavakat előbbre tartjuk a többinél, de köztük még nem tudunk sorrendet kialakítani (tehát az 'a' betűs szavak között „döntetlen" van). A rendezés teljessé tételéhez az kell, hogy vegyük a szavak következő betűjét, és aszerint soroljuk tovább a szavakat. Lehet, hogy már ekkor sikeresen el tudunk minden szót rendezni. Ha mégsem, akkor venni kell a szavak harmadik betűjét és így tovább. Ezzel a módszerrel minden szót sorba tudunk állítani, mert minden egyes szóra igaz, hogy valahányadik betűje esetén már különbözni fog az összes többi szótól.

A lexikografikus preferencia terminust *N. Georgescu-Roegen* vezette be [285], [286], a fogalom elméleti alkalmazásáról, a vele kapcsolatos kísérleti eredményekről lásd: [434], [169], [231].

4.4.2 Lexikografikus preferencia

Natura non facit saltum.

„a természet nem szereti az ugrást"

[94]

példák: A Mastercard 'Felbecsülhetetlen (Priceless)' reklámjaiban magát az – archimédeszi preferenciát reprezentáló – bankkártyát azzal értékelik magasra, hogy a reklámban rámutatnak egy értékre (barátságra, családra, szerelemre, meghitt pillanatra), ami megfizethetetlen, majd a kártyát rögtön emögé pozícionálják. A reklámokat lezáró szlogen is ezt a kettősséget fejezi ki:

„Van, amit nem lehet pénzért megvenni. Minden másra ott a MasterCard."

"eladó az egész világ"

"van az a pénz, amiért korpás lesz a hajam"

Az archimédeszi rendezés dominánssá válásának és ezzel szoros összefüggésben a lexikografikus rendezés háttérbe szorulásának folyamatát Csontos László és társai "uralmi" tézisnek nevezték el [181].

[408, 235]

Lexikografikus döntések egy általános döntési modellben [213]

Sen 35. Lexikografilus preferencia példája: A) maximald x-t, B) ha x-nek ugyanaz az erteke, akkor maximald y-t

4.4.3 Formális modellek

A közgazdasági szakirodalomba Nicholas Georgescu-Roegen vezette be a fogalmat [285], [286]. A fogalom formális elemzésével Peter C. Fischburn foglalkozott a legtöbbet. [253] [254] [255] [252] [257] [433] [432] [116] The Mathematics of Preference, Choice and Order. Essays in Honor of Peter C. Fishburn

A lexikografikus rendezés matematikai modelljét Andréka Hajnal, M.D. Ryan és P.Y. Schobbens dolgozta ki [20]. Andrékáék abból indulnak ki, hogy a preferenciarendezés fogalma nem elég a lexikografikus rendezés fogalmának definiálásához, szükség van a *prioritás* operátor bevezetésére. Azért, mert többféle szempont szerint is rendezhetünk dolgokat (autóvásárláskor figyelembe vehetjük az autó árát, gyorsulását, fogyasztását, vezethetőségét, színét, méretét, alakját stb.). Minden egyes szempont szerinti rendezés egy-egy preferenciarendezésnek felel meg, tehát sok rendezési relációnk is lehetséges. Ha pedig több szempont szerint kell rendezni ugyanazt az elemkészletet, akkor más és más sorrendek alakulhatnak ki a különböző szempontok (rendezések) szerint, amikor pedig szükség van arra, hogy a szempontok (rendezések) is valamilyen prioritást alakítsunk ki (vagyis valamilyen sorrendbe állítsuk őket.

lehet az, hogy többszerelős döntési folyamatban kell közösségi döntést hozni, amikor az egyéni preferenciákat kell aggregálni, de az is lehet, hogy egy személy szempontjait kell figyelembe venni.

(JAL.4) $\mathscr{L}_{JAL}(\mathscr{L}_{PL}, \mathbb{S}, \mathbb{A}, \Phi, \mathbb{C}, \{x,y,z\}, \{\alpha,\beta,\gamma\}, \mathbf{F})$, ahol

$\mathscr{L}_{PL}$	propozícionális logika
$\mathbb{S}$	világállapotok halmaza
$\mathbb{A}$	ágensek halmaza
Φ	vélemények halmaza
$\mathbb{C}$	választások halmaza
$\{x,y,z\}$	ágensek
$\{\alpha,\beta,\gamma\}$	világállapotok
$\mathbf{F}$	választási függvény

Kiinduláshoz egy $\mathbf{R}$ tetszőleges reláció, valamint az $\underset{\iota}{\mathbf{F}}$ teljes reláció és az $\underset{\iota}{\mathbf{E}}$ üres reláció fogalmára van szükség (↬31.o.). Bár az $\mathbf{R}$-re preferenciarelációként hivatkozunk, a korábbi erős feltételeket (↬89.o.) itt nem követeljük meg. Nem feltételezzük sem azt, hogy $\mathbf{R}$ reflexív, sem azt, hogy tranzitív lenne, semmilyen tulajdonságot nem követelünk meg tőle. A "legalább olyan jó" (jobb vagy ugyanolyan jó) jelentés mentén értelmezett $\mathbf{R}$ relációra támaszkodva a következő további relációkat vezethetjük be:

(BA.5) $\mathbf{R}(m,n)$

$\mathbf{R}$ a **legalább olyan jó** (jobb vagy ugyanolyan jó; jobb vagy közömbös) reláció; $\mathbf{R}$ a $P = \{m,n,\ldots,o\}$ (világ)állapotok halmazán van értelmezve.

(BA.6) $\overline{\mathbf{R}}(m,n) \stackrel{\text{def}}{=} \neg\mathbf{R}(m,n)$

$\overline{\mathbf{R}}$ a **rosszabb** (se nem jobb, se nem közömbös) reláció; $\overline{\mathbf{R}}$ az $\mathbf{R}$ negáltja.

(BA.7) $\text{BETTER}(m,n)\mathbf{R}^{<}(m,n) \stackrel{\text{def}}{=} \mathbf{R}(m,n) \wedge \neg\mathbf{R}(n,m)$

$\mathbf{R}^{<}$ a **jobb** (szigorúan jobb; erősen preferált; szigorúan preferált) reláció; $\mathbf{R}^{<}$ aszimmetrikus.

(BA.8) $\text{INDIFFERENT}(m,n)\mathbf{R}^{\equiv}(m,n) \stackrel{\text{def}}{=} \mathbf{R}(m,n) \wedge \mathbf{R}(n,m)$

$\mathbf{R}^{\equiv}$ a **közömbös** (indifferens) reláció; $\mathbf{R}^{<}$ szimmetrikus.

(BA.9) $\text{INCOMPARABLE}(m,n) \equiv \mathbf{R}^{\#}(m,n) \stackrel{\text{def}}{=} \neg\mathbf{R}(m,n) \wedge \neg\mathbf{R}(n,m)$

$\mathbf{R}^{\#}$ az **összehasonlíthatatlan** (inkompatibilis) reláció.

Az univerzális és üres reláció, valamint a fenti preferencia-vonatkozású relációk között megállapíthatóak a következő összefüggések:

(BA.10) $\overline{\underset{\sim}{\mathbf{F}}} = \underset{\sim}{\mathbf{F}}^{<} = \underset{\sim}{\mathbf{F}}^{\#} = \underset{\sim}{\mathbf{E}}^{<} = \underset{\sim}{\mathbf{E}}^{\equiv} = \underset{\sim}{\mathbf{E}}$

(BA.11) $\underset{\sim}{\mathbf{F}}^{\equiv} = \underset{\sim}{\mathbf{E}}^{\#} = \overline{\underset{\sim}{\mathbf{E}}} = \underset{\sim}{\mathbf{F}}$

Ezekután az $\mathbf{R}(m,n)$ preferenciarendezés mintájára vegyük az $\{\mathbf{R}_v(m,n)\}_{v\in V}$ preferenciarelációk valamely halmazát (ahol $m,n \in M$). A preferenciarelációinkra legyen adott még egy $|V|$-edrendű $<$ (szigorúan részben rendezett) **prioritás** operátor is.

(BA.12) $< (\mathbf{R}_x(m,n), \mathbf{R}_y(m,n)) \equiv <(x,y) \equiv x < y \stackrel{\text{def}}{=} \text{STRICT_PARTIAL_ORDER}(x,y)$

Ekkor a $\{\mathbf{R}_v(m,n)\}_{v\in V}$ preferenciarelációk halmazán, azok kombinációjaként értelmezett $\langle \mathbf{R}_v(m,n)\rangle_{v\in V}$ rendezett $|V|$-es, amit innentől $\text{LPREF}(m,n)$ **lexikografikus preferenciarendezésnek** nevezünk, így definiálható:

(BA.13) $\langle \mathbf{R}_v(m,n)\rangle_{v\in V}(m,n) \equiv \text{LPREF}(m,n) \stackrel{\text{def}}{=} \forall m \in V(\mathbf{R}_x(m,n) \vee \exists y \in V(y < x \wedge \mathbf{R}_y^{<}(m,n)))$

A formula egy diszjunkció. Az első komponens akkor lesz igaz, ha az aggregátum összes eleme ($\mathbf{R}_x(m,n)$) ugyanúgy rendezi el a két elem (m és n) viszonyát. Ennek kiderítéséhez végig kell menni az aggregátumba bevont összes preferenciarendezésen, hiszen univerzális kvantor zárja le ezt a részformulát. A képlet második összetevője megadja a kiértékelési folyamat gyorsabb lezárásának lehetőségét (ebben az értelemben rendkívül praktikus szabálykomponenesnek minősíthetjük). A diszjunkció második eleme ugyanis azt rögzíti, hogy ha elindulunk a prioritási sorba állított preferenciarendezések kiértékelésében, és ezek között van olyan reláció (a prioritási sor y. pontján), amely az összehasonlított két elem (m és n) közül az egyiket szigorúan preferálja a másikkal szemben ($\mathbf{R}_y^{<}(m,n)$), akkor ez a tény eldönti az teljes lexikografikus rendezést is (az éppen adott prioritási ponton szigorúan preferált alternatíva vagyis az m javára).

Értelmezéséhez segíthet, ha a szótári rendezés példájára gondolunk. A szótárakban (lexikonokban) szavakat kell sorba állítanunk, ezeket jelöljük itt az m,n változókkal. A $<$ prioritás operátor itt a betűk szavakon belüli pozícióinak sorrendjét határozza meg (európai írásmód esetén balról jobbra halad a sorrend). A $\mathbf{R}_x(m,n)$ preferenciarelációink mind megegyeznek, és az ábécé szerinti rendezést jelentik. Annyi preferenciarelációnk van ($|V|$), ahány betűje van a szótár legtöbb betűből álló szavának. Minden szó tetszőleges x. pozíciójában a $\mathbf{R}_x(m,n)$ preferenciareláció alapján meg tudjuk mondani, hogy az ott található betű alapján az adott szó mely szavakat előz meg, mely szavak vannak előtte, illetve mely szavakkal van azonos helyzetben (amikor ugyanaz a betű szepel az x. pozíción). Ez azonban kevés a lexikon szavainak sorrendbe állításához. A végleges rendezést a $<$ prioritás operátor segítségével lehet csak kialakítani. Nézzük meg ezt a következő példa segítségével:

> E példában a pék a példa arra, hogy ki pótolható, de ő példázza azt is, hogy kit nem lehet pótolni.

A későbbi azonosíthatóság miatt érdemes megjelölni a szavak helyét (sorrendjét a mondaton belül). Erre azért lesz szükség, hogy kezelni tudjuk a szavak többszöri előfordulásait is. Mivel a szótárba nem kerülnek be a központozási jelek, csak a szavak, így a fenti mondatot húsz szóalakra bonthatjuk fel:

E(1) példában(2) a(3) pék(4) a(5) példa(6) arra(7), hogy(8) ki(9) pótolható(10), de(11) ő(12) példázza(13) azt(14) is(15), hogy(16) kit(17) nem(18) lehet(19) pótolni(20).

Ezt a húsz szót a lexikografikus elv alapján rendezve a szabály két összetevője közül két esetben alkalmazhatjuk az első komponenst. Az ‘a’ és a ‘hogy’ szavak esetében a formula első komponesét kell mindvégig alkalmazni, vagyis e két-két szóelőfordulásra a szabály azt hozza ki, hogy legalább olyan jók, mint a másik, hiszen a szavak minden pozíciójában megegyeznek a betűk. A többi esetben a lexikografikus szabály mindkét összetevőjét figyelembe kell venni, és mindig találhatunk egy olyan pozíciót az egyes szavakon belül, amelyekben az éppen összehasonlított szavak valamelyikének betűje megelőzi a másik szó betűjét. Az ‘azt’ és az ‘arra’ szavak esetében például a második pozícióban levő betűk esetében az ‘r’ előbb van a ‘z’ betűhöz képest az ábécé szerinti rendezés szerint, amiatt az ‘arra’ szavunkat előbbre soroljuk az ‘azt’ szavunkhoz képest. Az egybetűs szavakat már az első pozíció alapján rendezni lehet (ezért kerül előbbre az ‘a’ az ‘e’ vagy ‘ő’ szavakhoz képest), a több betűből álló szavaknál a későbbi pozíciók alapján lehet csak sorrendet kialakítani a szavak között. “Példaszótárunk” összeállításában a hatodik pozícióig el kellett menni ahhoz, minden szóról el tudjuk dönteni, hol a helye (a ‘példában’ és ‘példázza’, illetve a ‘pótolni’ és ‘pótolható’ szavak esetében).

#	szó	(#)	PR_1	PR_2	PR_3	PR_4	PR_5	PR_6	PR_7	PR_8	PR_9
1-2	a	(3)	a								
	a	(5)	a								
3	arra	(7)	a	r	r	a					
4	azt	(14)	a	z	t						
5	de	(11)	d	e							
6	e	(1)	e								
7-8	hogy	(8)	h	o	g	y					
	hogy	(16)	h	o	g	y					
9	is	(15)	i	s							
10	ki	(9)	k	i							
11	kit	(17)	k	i							
12	lehet	(19)	l	e	h	e	t				
13	nem	(18)	n	e	m						
14	ő	(12)	ő								
15	pék	(4)	p	é	k						
16	példa	(6)	p	é	l	d	a				
17	példában	(2)	p	é	l	d	á	b	a	n	
18	példázza	(13)	p	é	l	d	á	z	z	a	
19	pótolható	(10)	p	ó	t	o	l	h	a	t	ó
20	pótolni	(20)	p	ó	t	o	l	n	i		

majd annak általánosításaként a prioritás operátor formulája

prioritás reláció (priority relation), reflexív, tranzitív és antiszimmetrikus

(BA.14) $<$

Miután definiáltuk a pritoritás relációt, és magát az aggregáló függvényt, itt az idő rögzíteni, hogy milyen elvárásaink lehetnek az aggregálási folyamat eredményével szemben. Korábban annyit mondtunk, hogy az **R** relációval szemben nem kell megkövetelnünk a “szokásos” relációtulajdonságokat (reflexivitást, tranzitivitást stb.). Most érdemes

rögzíteni egy elvárást: ha az $\mathbf{R}_v$ reláció "mégis" tranzitív, akkor elvárjuk, hogy ez a tulajdonság mindvégig megőrződjék, vagyis ha tranzitív relációkból aggregálunk egy "közös" relációt, akkor ez utóbbi is maradjon tranzitív. Ez a kritérium új az eddig megfogalmazott feltételekhez képest. A többi tartalmát tekintve már ismerős lehet a korábbiakból. Andrékáék IBUT-kritériumoknak nevezik el az általuk felállított feltételrendszert. Nézzük meg, pontosan miből is áll.

(BA.15) *a*

(I) független az irreleváns alternatíváktól (independency of irrelevant alternatives) az összehasonlításba be nem vont alternatívák nem befolyásolják a végeredményt

(BA.16) *a*

(B) ?

(BA.17) $\text{UNANIMOUS_WITH_ABSTENTIONS}(prior) \stackrel{\text{def}}{=}$

(U) (unanimous with abstentions)

(BA.18) $\text{PRESERVATION_OF_TRANSITIVITY}(prior) \stackrel{\text{def}}{=}$

(T) a tranzitivtás megőrzése (preservation of transitivity): a komponens preferencia-relációk tranzitivitása megmarad az aggregálás után

(BA.19) $\text{NON-DICTATORIAL}(prior) \stackrel{\text{def}}{=}$

(N) a diktatúra kizárása (non-dictatorial)

A fenti szempontok alapján a továbbiakban olyan preferenciaaggregálási módszereket keresünk, amelyek kielégítik a fenti posztulátumokat. Ehhez a lexikografikus preferencirendezés általánosításával jutnak el a szerzők.

IBUT-feltételeknek megfelelő operátor (operator satisfying conditions IBUT)
végesrendű operátor (finitary operator)

(BA.20) $m\, o((R_x)_{x\in V})n \Leftrightarrow \forall i \in N(mR_{v(i)}n \vee \exists j \in N(j < i \wedge mR^{<}_{v(j)}n))$

(BA.21) $o\{R_x\}_{x\in V}(m,n) \Leftrightarrow \forall i \in N(R_{v(i)}(m,n) \vee \exists j \in N(\text{PRIOR}(j,i) \wedge R^{<}_{v(j)}(m,n)))$

ahol .

Ennek alapján megfogalmazzák az ASR-tételt [20, Theorem 14].

(BA.22) $\text{PRIORITY} \leftrightarrow IBUT$

ASR-tétel A finitely operator satisfies conditions IBUT iff it is a priority operator.

Következő lépésként Andrékáék bemutatják, hogy a prioritás operátor alapján felépíthető gráf mindig kifejezhető két bináris operátor segítségével, amelyekre aztán egy egész algebrát lehet építeni. A két operátort az 'x/y' és az '$x\|y$' szimbólumokkal, illetve a **de** (but) és a **másfelől** vagy másrészt (on the other hand) nevekkel jelölhetjük, és a következő módon határozhatjuk meg őket:

(BA.23) $x/y \equiv \text{BUT}(x,y) \stackrel{\text{def}}{=} (x \cap y) \cup y^{<}$

(BA.24) $x\|y \equiv \text{ON_THE_OTHER_HAND}(x,y) \stackrel{\text{def}}{=} x \cap y$

Az operátorok jelentésének értelmezéséhez segítséget jelenthet, ha arra gondolunk, hogy miként vesszük figyelembe a hierarchiába szervezett emberek véleményét. Amikor emberek egyetértenek, akkor bárhol is legyenek a szervezeten belül, a "közös véleményük" is az lesz, amit egyenként helyeselnek. A kérdés az, hogy miként összesítsük az eltérő véleményeket. Ezen a ponton lehet megérteni a két operátor szemantikáját. Amikor alá-fölérendeltségi viszonyban levő emberek véleménye eltér, akkor a magasabb ponton levő (a szervezeten belül nagyobb prioritással rendelkező) ember véleményét vesszük figyelembe. Ha i és j a két személy, és i van az alárendelt szerepben j-hez képest ($i < j$), akkor azt mondjuk, hogy "i véleménye az, hogy x, *de* j azt gondolja, hogy y", és ez utóbbi számít. Alá-fölérendeltségi viszony esetén tehát a **de** (BUT) operátort alkalmazzuk. Amikor viszont két ember azonos hierarchikus szinten van, akkor az eltérő véleményeik között nem tételezünk prioritási különbséget, azonos mértékben vesszük figyelembe őket. Ekkor azt mondjuk, hogy "(egyfelől) i azt gondolja, hogy x, *másfelől* j úgy véli, hogy y". Ilyen esetekben a **másfelől** (ON_THE_OTHER_HAND) operátort használjuk.

Érdemes megjegyezni, hogy bár matematikai értelemben elégséges a fenti két operátort definiálni, létezik egy harmadik is, amit leginkább akkor lehet értelmesen igénybe venni, amikor az elmélet alkalmazni akarjuk a társadalmi cselekvések leírásakor. A BUT operátorban ugyanis megcserélhetjük az argumentumok változóit. Ez a reláció (operátor) **konverzének** felel meg, amit logikailag kifejezhetünk új kategória felvétele nélkül is. Azért érdemes mégis említeni (és felvenni a modellbe) ezt a "másik irányú" relációt (operátort), mert közel sem mindegy, hogy melyik végén áll az ember egy alá-fölérendeltségi viszonyban. Az új operátor definíciója a következő:

(BA.25) $y/x \stackrel{\text{def}}{=} (y \cap x) \cup x^{<} \equiv x/^{-1}y$

Mivel az új operátor matematikai értelemben redundáns, az elmélet további kifejtésekor nem számolunk vele, noha később még hivatkozhatunk rá. A két alapoperátor viszont két ok miatt is nagyon fontos. Egyrészt bizonyítható, hogy minden IBUT-feltételeknek megfelelő, végesrendű operátor meghatározható a két alapoperátor és a változók segítségével ([20, Corollary 46]), másrészt a két operátorra támaszkodva definiálhatunk egy **preferenciális algebrát** (preferential algebra) is, ami az elméletet nagyon erőssé teszi. Ehhez az alábbi axiómákat kell rögzítenünk:

(BA.26) $x \| x = x$

(BA.27) $x \| (y \| z) = (x \| y) \| z$

(BA.28) $x \| y = y \| x$

(BA.29) $(x/x) = x$

(BA.30) $x/(y/z) = (x/y)/z$

(BA.31) $(x \| y)/z = (x/z) \| (y/z)$

(BA.32) $(x/y) \| x = x \| y$

SSSS

jólfundáltság (érvényes a regularitás axiómája)

a **BUT** operátor a benne levő alá-, fölérendeltségi mozzanat miatt diktatórikus minőséget tartalmaz

4.4.4 egyéb

ez a modell a hierarchia leírása is egyben. A hierarchia fontosságáról, szerepéről én is írtam eyg tanulmányt: [728] A hierarchia két fontos típusa: beágyazott és nem-beágyazott hierarchia az élő rendszerek beágyazott hierarchiával jellemezhetőek [245]

A hierarchia fogalmáról bővebben: [430]

[557]

Sven Ove Hansson preferencia változása [335]

prioritizálás = preferenciák aggregálási módszere (Grosof)

precedencia vs. prioritás vs. preferencia ???

???????

A lexikografikus preferencia másodrendű preferencia, hiszen a prioritás operátor maga is egy preferenciarendezésnek felel meg, amit preferenciarendezéseken hajtunk végre. Richard Jeffrey, másodlagos preferencia [377] The Logic of Decision

Self-Control through Second-Order Preferences [530]

a személyen belüli preferenciakonfliktusok azonos lényegűek a személyek közti preferenciaütközésekkel, így a társadalmi választások elméletének (social choice theory) relevanciája van az egyéni preferenciakezelésre vonatkozóan.

a weberi racionaltásfogalmak megfeleltetése Az ezredvégi nagy borzongás [181]

Lexikografikus döntések egy általános döntési modellben [213]

[212]

az archimédeszi preferencia kompenzatórikus, a lexikografikus nem-kompenzatórikus: Dombi

kompenzatórikus, ha bármelyik értékelési kritérium szerinti eredmény kompenzálható a többi szempont szerinti értékelésekkel

ha egy oktatási szituációban egy diák nem mehet tovább a következő félévre, ha nem teljesítette egy tantárgy minimálkövetelményét, akkor ott nem-kompenzatórikus rendszerről beszélhetünk, mégha az összes többi tantárgyból jelest szerzett is.

az outranking módszer a kompenzatórikus egyik típusa vétó küszöb

az aggregálási problémát oldják meg

grosof

Preference, priorities and belief [197]

[210] [209]

Dietrich és List:

desire or preference as volitional attitude szemben áll belief as cognitive attitude

fundamental preferences vs. derived preferences

motivating reason vs. normative reason

Chapter 5

Társadalmi cselekvés

gyakorlati következtetések, Georg Henrik von Wright [793] [790]

A cselekvéselmélet központi kategóriája a **társadalmi cselekvés**. Ezt a fogalmat Max Weber is saját elmélete középpontjába helyezte, és Weber meghatározása azóta is igazodási pontot jelent a társadalomtudósok számára [810]. Én is a weberi értelmezésből fogok kiindulni.

A társadalmi cselekvés weberi értelmezését akkor érthetjük meg pontosan, ha követjük azt az logikát, ahogyan Weber összekapcsolja egymással a cselekvéselmélet négy alapfogalmát, a **viselkedés**, a **cselekvés**, a **társadalmi cselekvés** és a **társadalmi kapcsolat** kategóriáját. A viselkedés fogalmát explicit módon ugyan nem definiálja Weber, de a cselekvés fogalmának meghatározásából kibontható, mit gondol róla. Nézzük először a cselekvés definícióját:

> „ . . . *cselekvésnek* nevezünk minden emberi viselkedést (mindegy, hogy külső vagy belső ténykedésről, valaminek az elmulasztásáról vagy eltűréséről van-e szó), ha és amennyiben a cselekvő, illetve a cselekvők valamilyen szubjektív értelmet kapcsolnak vele össze" [810, 38]

Ebben a definícióban benne van egyfelől a viselkedés fogalmának implicit meghatározása, másfelől a cselekvés fogalmának explicit minősítése. Viselkedés lehet valamilyen ténykedés (tevés), mulasztás vagy tűrés (tartózkodás), illetve mindezek lehetnek belső vagy külső viselkedések. Viselkedés lehet, ha valaki énekel, eszik, sétál, olvas, becsukja az ajtót, gondolkozik, segít valakin, adót fizet vagy nem fizet adót. A viselkedés (magatartás) a világ lehetséges állapotai között lehetséges transzformációkkal hozható kapcsolatba. Ha az asztalon van egy virsli és én megeszem, akkor a viselkedésem eredményeként a következő világállapotban már nem lesz a virsli az asztalon (a gyomromban lesz). A viselkedésnek ezt az értelmezését persze Weber nem bontja így ki, de megteszik ezt helyette mások. Hogy hogyan, azt a következő fejezetben fogom bemutatni.

A viselkedés és a világállapot-transzformációk összekapcsolásával csak annyi a "baj", hogy ezt a meghatározást kiterjeszthetjük minden élőlényre vonatkozóan. Nem véletlen ezért az, hogy Weber még hozzátesz valamit a cselekvés fogalmának definíciójához. Szerinte ugyanis a cselekvésről csak akkor beszélhetünk, ha a "cselekvő szubjektív értelmet kapcsol" a viselkedéséhez. A cselekvés ezáltal egyfelől a viselkedés fogalmának generikus alárendeltje lesz, másfelől a "szubjektív értelem" minőségével egy olyan értelmezési réteget hoz be Weber a cselekvéselméletbe, amely segítségével – talán – el lehet különíteni az emberi és állati viselkedéseket egymástól.

Ezután határozza meg Weber a társadalmi cselekvés kategóriáját: a társadalmi cselekvés az a cselekvés, amely „szándékolt értelme szerint mások viselkedésére van vonatkoztatva, és menetében is mások viselkedéséhez igazodik" [810, 38]. Sajnos a weberi definícióban a cselekvés "szubjektív értelme" helyett a "szándékolt értelem" kifejezés szerepel, ami kicsit zavaró, mert utóbbi minőség más és több, mint az előbbi. Ezzel a kérdéssel is foglalkozom majd később, az intencionalitásról szóló fejezetben. Ami nagyon fontos: a társadalmi cselekvés fogalmába a "másikhoz való igazodás" minőségével már "belekerül" a társadalmiság mozzanata – még ha gyengébb formában is. A társadalmiságot igazán jól a társadalmi kapcsolat fogalmával tudja Weber megragadni, amikor már arról van szó, hogy a társadalmi cselekvők kölcsönösen egymáshoz igazodnak a cselekvéseik során.

tűrés-tűrés-tűrés-tűrés-tűrés-tűrés-tűrés-tűrés-tűrés

William Hazlitt: An Essay on the Principles of Human Action

Ebben a fejezetben arra a kérdésre szeretnék választ találni, hogy mi a társadalmi cselekvés lényege. Ehhez először a cselekvés fogalmát kell értelmeznünk. Vizsgáljunk meg pár példát, hogy a jelenség mélyén rejlő kérdéseket és problémákat jobban érzékelhessük.

> Kiszáradt a fű.
>
> Az eső elmosta a nyomokat.
>
> Nagyot nőtt ez a Peti gyerek.
>
> A betörő ott hagyta az ujjlenyomatát az ajtókilincsen.
>
> Profi volt az elkövető. Nem hagyott nyomokat.
>
> A sofőr látta a balesetet, de nem állt megsegíteni.
>
> Peti beteg volt, egész nap köhögött.
>
> Péter köhögött egyet, ezzel jelezte, hogy ő is ott van a szobában.

5.1 Cselekvés

Bertalan László [77]

Georg Henrik von Wright Norm and Action c. könyvében [792]

K. Segerberg [655]

A Companion to the Philosophy of Action [544]

MichaelThompson, Life and Action. Elementary Structures of Practice and Practical Thought [747]

DanielVanderveken, Logic, Thought and Action [785]

Donald Davidson: [188]

Rescher szerkesztésében Davidson, von Wright, Anderson, Simon [601]

Alvin I. Goldman cselekvéselmélete [302]

Seumas Miller, Social Action. A Teleological Account [508]

5.1.1 Szent Anzelm cselekvéstipológiája

Anzelm cselekvésről szóló fejtegetése

S.L. Uckelman [761]

Douglas N. Walton [807], [805], [806]

Howard L. Dazeley és Wolfgang L. Gombocz [196]
Nuel Belnap [62]
Ricoeur [670]

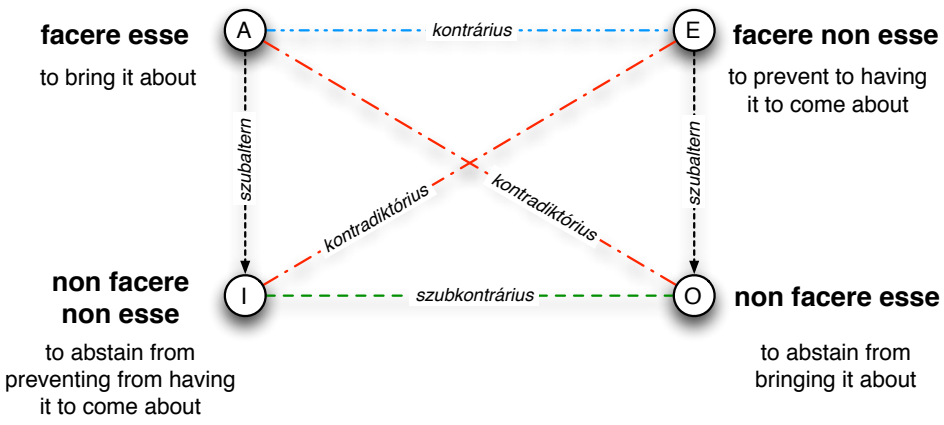


Figure 5.1: a cselekvés logikai négyszöge Szent Anzelm kategóriáival

facere = valamivé tesz verifikáció = verus (igaz) + facere (valamivé tesz) = igazzá tesz
inf. impf. act: facere inf. perf. act: facisse inf. inst. act: facturum esse inf. impf. pass: faci inf. perf. pass: factum esse inf. inst. pass: factum iri
agere - agent
facere - létrehoz, teremt, csinál értelemben agere - általános cselekvés értelemben
act -agere
fact - facere
bring about = előidézni, okozni, létrehozni

Facere esse	to bring it about	megcsinálni
Non facere esse	to abstain from bringing it about	
Facere non esse	to prevent from having it come about	
Non facere non esse	to abstain from preventing from having it come about	

»dsdsd «

„A latin nyelvben a ‘facere’ (= tenni, csinálni) és az ‘agere’ (= cselekedni) szavakban a gyakorlati tevékenység két sajátos formája nyer megfogalmazást. A ‘csinálni’ szó a művészet és a mesterségek gyakorlati tevékenységére utal. A »cselekedni«viszont erkölcsi tevékenységet fejez ki. Aquinói Szent Tamás azonban még egy fontos megállapítást tesz a két tevékenységi terület világos elkülönítésére. Arisztotelész nyomán az esztétika számára is jól felhasználhatóan fogalmazza meg, hogy amíg a művészet és a mesterség területén a “mű” a fontos, addig az erkölcs világában a cselekvő személyen van a hangsúly. Miközben az erkölcsi személy jót tesz másokkal, maga válik jóvá cselekedetei által, éppen mert tevékenysége jó szándékból fakad. Az esztétika világában viszont előfordul, hogy a mű ára esetleg a személy

> erkölcsi mivoltának lealacsonyodása. A sajátosan emberi cselekedetek jelölésére ajánlható a "személyes cselekedet", amely tudatos és szabad elhatározáson alapuló, tehát szándékos tettet, illetve állásfoglalást jelent, adott helyzetben. Ebben a meghatározásban a tudatosság a cselekvés emberi jellegének lényeges feltétele. Még nagyobb nyomatékot kap azonban a "szándékos" jelző, mert ha a "tudatos" teszi a cselekedetet emberivé, a szándékos teszi erkölcsivé. A tudatosság még nem jelent egyben szándékosságot is. Sok minden van magunkban is és a világban is, amiről tudunk, de amit nem akarunk." [96, 48]

nem tudom, kell-e ez?

5.1.2 Tevés és tartózkodás

> A népi bölcsesség szerint a gyereknevelésben két mondatot ismételgetünk folyamatosan: 'Kezdj már bele!' és 'Hagyd már abba!'

A hétköznapi, reflektálatlan gondolkodás könnyen egybe mossa, azonosítja a cselekvés és a tevékenység, tevés kategóriáit, holott nyilvánvaló, hogy a cselekvés fogalma alá két másik fogalmat kell rendelnünk: a **tevés**, megcsinálás, illetve a **tartózkodás**, tétlenség két kategóriáját. Szokás ezeket **pozitív** és **negatív cselekvéseknek** is nevezni. Ha valaki jelentéktelen, érdektelen kérdésnek tartaná a tartózkodás jelenségét, akkor annak érdemes belegondolni abba, hogy a sztrájk jelensége a nem-dolgozást (tehát a dolgozástól való tudatos tartózkodást) jelenti, amely igen sok, igen mély konfliktus forrása volt eddig a történelemben. Egy – Csuang-ce, taoista gondolkodótól származó – gyönyörű mondás szerint:

> „. . . az ember nem képes a sáros, zavaros vizet megtisztítani, de ha békén hagyja, az magától letisztul." [349, 27]

A taoizmus fontosnak tartja, hogy az emberi viselkedés a Wu-Wei elvhez igazodjon, aminek jelentése: cselekvés nem-cselekvés által. Ezzel voltaképp a tartózkodást írják le. De a passzív ellenállás jelensége is azt példázza, hogy a tartózkodás értelmes, jelentéssel bíró fogalom.

Gandhi **szatjágraha** mozgalma (**satyagraha**) erőszakmentes polgári engedetlenség (non-violent civil disobedience)

Gandhi, M.: Satyagraha. Non-Violent Resistance. Ahmedabad, 1951. King, Martin Luther: Nem hallgathattam. Bp., 1968. J. Ruskin: Unto this last, H. D. Thoreau: Civil Disobedience, L. Tolsztoj: Isten királysága bennünk

A mottóban idézet bölcsesség is azt fejezi ki, hogy a nevelés kétfajta cselekvés irányulhat: vagy azt szeretnénk elérni, hogy a gyerekeink megtegyenek valami, vagy azt, hogy tartózkodjanak valamitől.

Danto [186]

Bruce Vermazen négyféle negatív cselekvést különít el [787]

a negatív cselekvés pozitív cselekvéspár segítségével értelmezhető (100-102.o.)

"... resisting and displacement refrainings will count as acts in the same sense as positive acts, and simple and disobedient refrainings will require a separate account, one that does not require that they be events." [787] 104.o.

a negatív cselekvés vagy a "nem-cselekvés" az intenció fogalmával határozható meg a "nem cselekvéshez" képest.

a cselekvés, a nem-cselekvés fogalmához kell az intenció, tehát a könyvben előbbre kell venni

de ő még nem tudta jól megragadni a tartózkodásban rejlő intencionalitást. A tartózkodás nem-cselekvést jelent. Utóbbi fogalom azonban sok mindent jelenthet, és el kell tudnunk választani egymástól a nem-cselekvés kétféle értelmezését. Amikor állok egy ajtó előtt és nem lépek ki rajta, mert nem akarom ezt valamiért, akkor ez az "nem-kilépés" nem-cselekvésnek és tartózkodásnak számít. De ugyanebben a pillanatban még milliónyi más dolgot nem csinálok: nem énekelek, nem beszélek, nem írok könyvet, nem eszem stb. Ezekre a nem-cselekvésekre nem mondhatjuk, hogy tartózkodom az adott cselekvések végrehajtásától, mert ha így tennénk, akkor végtelen számú cselekvést lehetne (kéne) még felsorolnunk. Valahogy meg kell különböztetnünk a tartózkodást az ilyen "nem-cselekvésektől".

von Wright a tartózkodást a képességgel kapcsolta össze

A Stit-elmélet segít ebben:

(ZL.0) $\quad \text{REFRAIN}(\alpha) = \text{STIT}(\neg\text{STIT}(\alpha))$

A képlet valamely cselekvés elvégzésére irányuló törekvés hiányára való törekvést vagyis a cselekvés nem elvégzésére való törekvést fejezi ki.

A tartózkodást természetes nyelvi kontextusban sokféle alakban kifejezhetjük. Először is bármikor használhatjuk a tevőleges igék tagadó formáját: 'nem csinál meg valamit', 'nem segített neki'. De a negatív cselekvést kifejezhetjük "önálló" módon is. Mondhatjuk, hogy 'tétlenül szemlél', 'tétlenkedik', 'elmulaszt', 'mellőz', 'hagy valamit megtörténni', 'visszafogja magát'. Ha valaki 'enged másnak valamit megcsinálni' vagy 'tűri, hogy más megcsináljon valamit', akkor is tartózkodásról beszélhetünk, bár ekkor már másodrendű cselekvésről beszélhetünk. Olykor a 'kiváltság', a 'privilégium', a 'mentesség', az 'immunitás' is jelenthet valamilyen kötelezettség alóli mentességet, ami ebben az értelemben a tartózkodáshoz is kapcsolható.[1] Viszont ha valaki valamit elfelejt, akkor azt nem sorolhatjuk a tartózkodás kategóriája alá, mert ebből hiányzik az intencionáltság.

Egy formális cselekvéselméleti modellben nyilvánvalóan kulcsszerepe van a cselekvés fogalmának. Az alapkategóriák jelölésére többféle konvenció is létezik, ezek közül én a következő formulákat választottam a modellem számára. A **cselekvés** tehát lehet **tevés**, illetve **tartózkodás**. A cselekvés általános fogalmát a $\pm\text{DO}$ operátorral jelölöm, de megengedem, hogy olykor – von Wright megoldását követve – az általában vett cselekvésre az ACT terminussal hivatkozzak. Amennyiben nem zavaró, vagyis nem okoz többértelműséget, akkor a DO terminust is használhatónak gondolom. Ha a kontextus megkívánja, akkor a DO (esetleg a +DO) kifejezés a pozitív cselekvést vagyis a tevést, míg a $\neg$DO (esetleg a -DO) terminus a negatív cselekvést vagyis a tartózkodást jelöli.

Kálmán László posztja:

[1] Angolul is elég sok nyelvi olyan alak van, amely a tartózkodást fejezi ki, például 'not-doing', 'omission', 'forbearance', 'refraining', 'inaction', 'abstention', 'abstaining', 'avoidance', 'nonaction', 'letting happen'. És a sor biztosan sokáig folytatható lenne.

A Pride-dal kapcsolatban meg egy blokkolt FB-felhasználóval kapcsolatban jött elő megint ez, hogy toleránsnak kell-e lenni a homofóbokkal és más intoleráns emberekkel.

Egyik legkorábbi publicisztikámban (már nem emlékszem, hol jelent meg, de valamikor röviddel a rendszerváltás után volt, talán Csoóri hírhedt cikke után) Nyíri Tamásnak válaszoltam. Ő egy olyan szofizmát adott elő, aminek az volt az értelme, hogy aki a rasszistákkal és egyéb intoleráns emberekkel szemben nem toleráns, az nem nevezheti magát igazán toleránsnak. (Amúgy általában normális embernek tartották, de mégis leírt egy ilyet.) Ez a paradoxon elsőre arra a régesrégi vitára emlékeztet, hogy a semmittevés vajon cselekvés-e, de valójában sokkal egyszerűbb. Mert a klasszikus logikában érvényes a tertium non datur és így a kettős tagadás eliminálható, így a tolerancia igenis azonos az intolerancia elutasításával. Az már csak játék a szavakkal, nem pedig paradoxon, hogy ezt úgy is át lehet fogalmazni, hogy nem tolerálom az intoleranciát.

Ahogy Ligeti Gábor írta egy másik kommentben: ha valaki korlátozni akarja a magyar kormányt abban, hogy tovább csökkentse a szabadságot, az csak a szavak szintjén tűnik paradoxonnak, valójában nem korlátozás, ha valaki nem akarja tétlenül nézni a szabadság csökkentését.

Von Wright

A cselekvések és normák kapcsolatát elemző könyvében von Wright természetesen foglalkozik a cselekvés fogalmával. Von Wright a cselekvés fogalma alatt valamilyen ágens általi intencionális tevést vagy tartózkodás ért, amelynek eredményeként két világállapot közti átmenet (transition) megvalósul vagy elmarad.

(világ)állapotok ↣ állapotváltozások (átmenetek, tranzíciók) ↣ cselekvések

Az állapotokat p, q, r betűkkel jelöli von Wright, amit az esetek egy részében én is követek, de egy ponton eltérek ettől, és az állapotok jelzésére az F betűt fogom alkalmazni. Az állapotok fennállásának tagadását pedig a $\neg p, \neg q, \neg r$ terminusokkal jelölöm. Az átmenetek, tranzíciók jelölésére követem von Wrightet, és a T-operátorral hivatkozom rájuk, melynek két állapotváltozó jelzésére alkalmas paraméterét kell feltüntetni a formulákban: $T(p, \neg p)$ vagy $T(p, q)$. Ha azt akarom kifejezni, hogy az állapotváltozás operátor valamelyik paraméterében előfordulhat egy állapot vagy annak tagadása, akkor ezt a $\pm p$ szimbólummal jelölöm. Ekkor így néz ki egy állapotváltozást leíró nyitott formula: $T(\pm p, \pm p)$. Ezzel a nyitott formulával négy lehetséges állapotváltozást ragadhatunk meg egyszerre: $T(p, p)$, $T(\neg p, p)$, $T(\neg p, \neg p)$, $T(p, \neg p)$. Von Wright a cselekvés általános fogalmára az **act** (ACT) terminust használja, míg a tevés és tartózkodás fogalomkettősét a **do** és **forbear** igék segítségével jelöli, de képletekben a rövidebb d és f szimbólumokat is alkalmazza. Az általam épített formális modelbe nem ezt a jelölési konvenciót viszem tovább, hanem a $\pm$DO (illetve a DO és a $\neg$DO) szimbólumokkal jelölöm a cselekvésoperátorokat. Ezek alapján a cselekvésre vonatkozó legáltalánosabb nyitott formulát így fejezhetjük ki: $\pm\text{DO}(T(\pm p, \pm p))$. Ezt a következő négyféle módon jelölhetjük és értelmezhetjük (ahol a '|' szimbólum az **exklúzió** műveletét jelzi):

(ZL.1) DOING-FORBEARING

(i)	$\pm\text{DO}(T(\pm p, \pm p)) = +\text{DO}(T(\pm p, \pm p)) \mid \text{-DO}(T(\pm p, \pm p))$
(ii)	$\pm\text{DO}(T(\pm p, \pm p)) = \text{DO}(T(\pm p, \pm p)) \mid \neg\text{DO}(T(\pm p, \pm p))$
(iii)	$\pm\text{ACT}(T(\pm p, \pm p)) = do(T(\pm p, \pm p)) \mid forbear(T(\pm p, \pm p))$
(iv)	$\pm\text{ACT}(T(\pm p, \pm p)) = d(T(\pm p, \pm p)) \mid f(T(\pm p, \pm p))$

A cselekvések feltételei, a cselekvések, az átmenetek és a cselekvések eredményei közti kapcsolatokat von Wright alapján táblázatba rendezve mutatom be [792, 49], azonban von Wright megoldásához képest három apróbb módosítást is végrehajtok. Egyrészt más jelölési rendszert alkalmazok az operátorokra vonatkozóan, átírom a posztfix (vagy lengyel) jelölési rendszer szerint a kifejezéseket (például a '$pT\neg p$' helyett a '$T(p,\neg p)$' kifejezést használom). Másrészt a cselekvés paraméterei közé felveszek egy ágenst, és ezt 'Ego'-val jelölöm. Harmadrészt a cselekvés operátort a négy lehetséges jelölési mód közül a *(ii)*-vel jelölöm (szemben a von Wright által alkalmazott *(iv)* jelöléssel).

cselevés feltétele	cselekvés	cselekvés eredménye
$T(p,\neg p)$ p állapot fennáll, de megszűnik, hacsak nem tartják fent	$\text{DO}(Ego, T(p,p))$ Ego a p állapotot megőrzi	$T(p,p)$ p megmarad
$T(p,\neg p)$ p állapot fennáll, de megszűnik, hacsak nem tartják fent	$\neg\text{DO}(Ego, T(p,p))$ Ego hagyja, hogy a p állapot megszűnjön	$T(p,\neg p)$ p megszűnik
$T(p,p)$ p állapot fennáll, és meg is marad, hacsak nem változtatják meg	$\text{DO}(Ego, T(p,\neg p)$ Ego a p állapotot megváltoztatja	$T(p,\neg p)$ p megszűnik
$T(p,p)$ p állapot fennáll, és meg is marad, hacsak nem változtatják meg	$\neg\text{DO}(Ego, T(p,\neg p))$ Ego hagyja, hogy a p állapot megmaradjon	$T(p,p)$ p megmarad
$T(\neg p,\neg p)$ p állapot nem áll fenn, és nem is következik be, hacsak nem állítják elő	$\text{DO}(Ego, T(\neg p,p))$ Ego a p-t állapotot előállítja	$T(\neg p,p)$ p megtörténik
$T(\neg p,\neg p)$ p állapot nem áll fenn, és nem is következik be, hacsak nem állítják elő	$\neg\text{DO}(Ego, T(\neg p,p))$ Ego hagyja, hogy a p állapot hiánya megmaradjon	$T(\neg p,\neg p)$ p hiánya megmarad
$T(\neg p,p)$ p állapot nem áll fenn, de bekövetkezik, hacsak nem akadályozzák meg	$\text{DO}(Ego, T(\neg p,\neg p))$ Ego megakadályozza, hogy a p állapot bekövetkezzen	$T(\neg p,\neg p)$ p hiánya megmarad
$T(\neg p,p)$ p állapot nem áll fenn, de bekövetkezik, hacsak nem akadályozzák meg	$\neg\text{DO}(Ego, T(\neg p,\neg p))$ Ego hagyja, hogy a p állapot bekövetkezzen	$T(\neg p,p)$ p megtörténik

keveset írni a cselekvés feltételéről

C. Castelfranchi, Modeling social action for AI agents [137]

5.1.3 Magasabbrendű cselekvés

szankcionálás vs. büntetés

5.2 Cselekvő

A könyv elején már felvettem a modellbe az ágens (↬145.o.), illetve a self fogalmát (↬145.o.). behozni a pszichológia 'self' fogalmát

ágentivitás

Frankfurt: concept of a person [265] The Problem of Action [266]

Gary Watson Agency [809]

Alfred R. Mele, Motivation and Agency [492]

Michael E. Bratman, Structures of Agency [118]
E.J. Lowe, Personal Agency. The Metaphysics of Mind and Action [459]
ChristineM.Korsgaard:2008 [412]
TakingOurselvesSeriously.GettingitRight:2006 [268]

5.2.1 Ego és Alter

Ego a fokális ágens

A társadalmi cselekvés elemzése során ágensekről beszéltem, de eddig nem volt fontos sem az, hogy több cselekvőről beszéljünk, sem az, hogy megkülönböztessük őket egymástól. A társadalmi cselekvések során – akaratlanul is – modellezzük magát a társadalmat is. A társadalom minimális modelljéhez két szereplőre van szükség. Sokáig elég ennyit feltételeznünk, hogy a legelemibb vonásokat értelmezhessük. Ehhez viszont meg kell különböztetnünk a két szereplőt egymástól, leginkább azért, mert a társadalmi cselekvések megértését – módszertani individualizmus talaján állva – úgy remélhetük, ha valamelyik cselekvő szempontrendszerét, optikáját magunkévá tesszük. Ezt a szereplőt 'Ego'-nak fogom nevezni, és amíg csak lehet, kitartok e név mellett. A társadalmiságot "reprezentáló" másik szereplő neve 'Alter' lesz.[2] Az Alter fogalmára azért van szükség, mert vele tudjuk a társadalmiságot "kifejezni". Mindig Ego cselekvését akarjuk értelmezni, megmagyarázni, tehát egy személyre fókuszálunk, de a társadalmi cselekvéshez szükséges másik szereplőt nem hagyhatjuk ki a modellből – mégha tudjuk is, hogy ezzel a névvel egy passzív szerepet/szereplőt jelölünk csak.

Amikor a későbbiekben játékelméleti apparátust fogok alkalmazni a társadalmi kapcsolatok modellezésére, akkor látható lesz, hogy a játékelmélet különböző ágaiban a cselekvőkre olykor más neveket aggatnak (donor, recipiens, actor stb.). Ilyen esetekben igazodni fogok az éppen bemutatott elméletek névhasználati konvencióihoz, de minden esetben jelzem majd, hogy az egyes szerepnevek melyik "főszerepnek" felelnek meg (Egonak vagy Alternek).

5.3 Formális modellek

5.3.1 stit-elmélet

cselekvés, döntés

Belnap és Perloff cikke első megjelenés: [63], javított változat: [64]
axionatizálás: Xu cikkei
történeti elemzés: Michael Perloff [567] és Krister Segerberg [655]
két összegző mű:
Nuel Belnap, Michael Perloff és Ming Xu 'Facing the Future' című könyve [65]
John F. Horty 'Agency and Deontic Logic' című könyve [363]
Stit Cstit (Chellas Stit) Dtit (Deliberative Stit) John F. Horty és Nuel Belnap 1995 [364]

[2] Ezt az elnevezési konvenciót már a cselekvéselméleti könyvemben is alkalmaztam, és a vonatkozó referenciák közül akkor is, most is Talcott Parsons hasonló megoldását, illetve az ezt támogató érvelést emelném ki [562].

Horty

$\models^{\mathscr{M}}_{m/h}$
$|A|^{\mathscr{M}}_{m} \equiv |A|_{m} = \{h \in H_m : \mathscr{M}, m \vDash A\}$

horty-cikkek

Mark Brown. On the logic of ability. Journal of Philosophical Logic, 17:1-26, 1988.

Peter Geach. Whatever happened to deontic logic? Philosophia, 11:1-12, 1982.

Risto Hilpinen. On action and agency. In Eva Ejerhed and Sten Lindstrom, editors, Logic, Action, and Cognition, pages 3-27. Kluwer Academic Publishers, 1997.

Anthony Kenny. Human abilities and dynamic modalities. In Juha Manninen and Raimo Tuomela, editors, Essays on Explanation and Under- standing: Studies in the Foundations of Humanities and Social Sciences, pages 209-232. D. Reidel Publishing Company, 1976.

5.3.2 program-elmélet

dinamikus logika

5.3.3 Davidson

Action sentence inkább event sentence viszont nagyon jó javalslat de nem mondja meg, mitől cselekvés a cselekvés

a végén csak mond valamit Davidson: a cselekvésben intencionalitás van [188] 94.o.

'it was intentional of x that p' x a person, p action sentence (uo. 95.o.)

5.4 Interdependens cselekvések

A társadalmi cselekvésről szóló fejezetben eddig azt a kérdést jártam körül, hogy mit jelent a cselekvés fogalma, milyen típusai vannak a cselekvésnek, és kik a társadalmi cselekvések ágensei, a megelőző fejezetben pedig foglalkoztam az intencionalitás, az akarat problémájával. A cselekvés fogalmi lényegének megragadásakor az intencionáltságra fókuszáltunk. Eddig azonban az egyéni cselekvőre koncentráltunk, és nem figyeltünk arra a lehetőségre, amikor az egyéni cselekvések megvalósulása, azok sikeressége a kapcsolatba került felek mindegyikétől függ, vagyis a felek kölcsönös függőségben vannak. Ezt a társadalmi állapotot kölcsönös függésnek vagy interdependenciának, interdependens helyzeteknek nevezhetjük.

A továbbiakban az interdependens helyzetek, és az ilyen helyzetekbe került cselekvések alapvonásait szedem össze. Ha már tisztázva van, hogy

- a cselekvés elméletén belül mit jelent a tevés és a tartózkodás fogalma,
- az intencionalitás elméletén belül mit jelent az akarat fogalma,
- a cselekvés elméletén mit jelent az ágens fogalma,

akkor a cselekvők saját és mások cselekvésére irányuló akaratait, és ezek egymáshoz való viszonyait táblázatba foglalhatjuk. Valamelyik ágens cselekvésére Ego és Alter akarata vagy nem-akarata irányulhat (mindkettő intencionális!).

Table 5.1: Ego és Alter akaratainak és cselekvéseinek viszonyai, 1.

	Alter akarja	Alter nem akarja
Ego akarja	ágens cselekvését	ágens cselekvését
Ego nem akarja	ágens cselekvését	ágens cselekvését

A kölcsönös függőség itt abban érhető tetten, hogy a táblázatot elemeit úgy kell értelmezni, hogy egy cselekvésre vonatkozóan két akaratot akarnak "érvényesíteni" ('Ego akarja, Alter akarja, hogy ágens cselekedjen' stb.), és e két akarat metszetéből, egymáshoz való viszonyából kerekedhet ki valahogy a cselekvés tényleges kimenete, eredménye.

A fenti táblázat azonban még többértelmű, hiszen mind az ágens, mind a cselekvés fogalmába két-két értéket helyettesíthetünk (az Ego/Alter, illetve a tenni/tartózkodni párokat), ezért konkrét értékekkel kell kitöltenünk a táblázatot. Ha így teszünk (beírjuk, hogy 'Ego akarja, Alter akarja, hogy Ego megtegye' vagy 'Ego akarja, Alter nem akarja, hogy Ego megtegye' stb.), akkor az alábbi tizenhat állapothoz jutunk.

Table 5.2: Ego és Alter akaratainak és cselekvéseinek viszonyai, 2.

	Alter akarja	Alter nem akarja
Ego akarja	Ego megtegye (1)	Ego megtegye (2)
Ego akarja	Ego tartózkodjon (3)	Ego tartózkodjon (4)
Ego akarja	Alter megtegye (5)	Alter megtegye (6)
Ego akarja	Alter tartózkodjon (7)	Alter tartózkodjon (8)
Ego nem akarja	Ego megtegye (9)	Ego megtegye (10)
Ego nem akarja	Ego tartózkodjon (11)	Ego tartózkodjon (12)
Ego nem akarja	Alter megtegye (13)	Alter megtegye (14)
Ego nem akarja	Alter tartózkodjon (15)	Alter tartózkodjon (16)

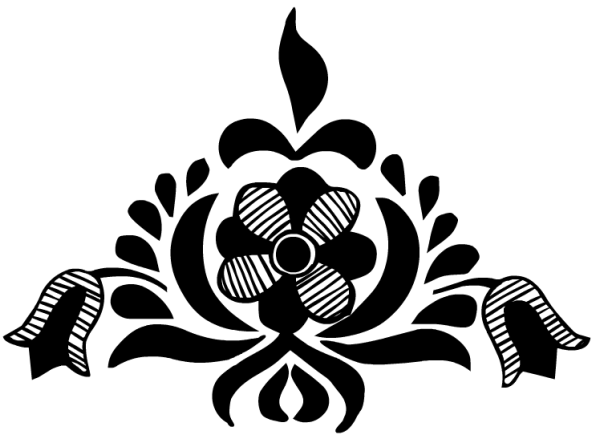
A tizenhat állapot közül a felében a két fél akarata megegyezik (vagy mindketten akarják vagy egyikük sem akarja valamelyikük cselekvését), tehát konszenzus (akaratkonszenzusról) van közöttük (az 1., 3., 5., 7., illetve a 10., 12., 14., 16. esetekről van itt szó). Az ilyen helyzetekben nyilván az várható, hogy a közös akarat szerint valósul meg (marad el) a cselekvés, hiszen mindkét fél egyaránt akarja (van nem akarja) azt. A másik nyolc esetben viszont, ahol a felek akaratai nem egyeznek meg, ott konfliktusról (akaratkonfliktusról) beszélhetünk. Az ilyen helyzetekben a cselekvés feltétele az, hogy a felek valamilyen módon rendezik a konfliktust (ez a helyzet a 2., 4., 6., 8., illetve a 9., 11., 13., 15. esetben). Hogy hogyan, azt egyelőre nem vizsgálom.

Érdemes kicsit átrendezni a táblázatot, hogy jobban látszódjék benne az, hogy kinek a cselekvéséről van szó.

Ebben az elrendezésben a táblázat első két oszlopában az Ego, a következő kettőben pedig az Alter cselekvésére irányuló akaratok egymáshoz való viszonyát láthatjuk. Mindkétszer létezik 4-4 olyan helyzet, amikor konszenzus, illetve konfliktus van a felek között. Ego cselekvése esetében konszenzus van az 1.,3.,10.,12., konfliktus van a 2.,4.,9.,11. esetben. Alter cselekvése esetében konszenzus van a 5.,7.,14.,16., illetve konfliktus van a 6.,8.,13.,15. esetben.

Table 5.3: Ego és Alter akaratainak és cselekvéseinek viszonyai, 3.

	Alter akarja	Alter nem akarja	Alter akarja	Alter nem akarja
Ego akarja	Ego megtegye (1)	Ego megtegye (2)	Alter megtegye (5)	Alter megtegye (6)
Ego akarja	Ego tartózkodjon (3)	Ego tartózkodjon (4)	Alter tartózkodjon (7)	Alter tartózkodjon (8)
Ego nem akarja	Ego megtegye (9)	Ego megtegye (10)	Alter megtegye (13)	Alter megtegye (14)
Ego nem akarja	Ego tartózkodjon (11)	Ego tartózkodjon (12)	Alter tartózkodjon (15)	Alter tartózkodjon (16)



Chapter 6

A társadalmi cselekvés típusai

Beállítódás, irányultság, motiváció a tipizálási szempont

6.1 Fiske relációs elmélete

négyféle reláció, négyféle mérési skála, négyféle matematikai struktúra: (reláció, modell, séma, szabály, nyelvtan, struktúra)

- ekvivalencia
- lineáris rendezés
- rendezett Ábel-csoport
- archimédeszi rendezett test

modell	leírás	skála	tendencia
communal sharing (CS)	World id divided into two groups (members or non-members). Non-members are excluded. Members give what they can and take what they need.	nominális skála	módusz
authority ranking (AR)	One's location in a social hierarchy determines one's status and claim on resources	ordinális skála	módusz + medián
equality matching (EM)	Relationship such as collegial or friendship networks in which in-kind or 'tit-for-tat' reciprocity dominates	intervallumskála	módusz + medián + számtani közép
market pricing (MP)	A single utiliy metric (e.g. money) makes valuation of large number of diverse entities possible	arányskála	módusz + medián + számtani közép + mértani közép

Jack Hirshleifer 1985-ös cikkében a rank-oriented motivációról azt írja, hogy az rosszindulatú, zérus-összegű játékra utal, hiszen vagy A van B felett vagy fordítva. (lásd: szervezet és/vagy státus fejezet)

6.2 Szabályszerűségek

alakzat divat erkölcs etikett habitus illem jog jogszabály jogszerűség konvenció minta mintaszerűség mintázat morál ösztön reflex rend rutin séma szabály szabályosság sz-

ekvivalencia	lineáris rendezés	rendezett Ábel-csoport	archimédeszi rendezett test
reflexive	reflexive	linear ordering	ordered Abel group
transitive	transitive	additive identity (0)	multiplicative identity (1)
symmetric	antisymmetric	additive inverse	multiplicative inverse
		associative addition	associative multiplication
		commutative addition	commutative multiplication
		addition is order preserving	multiplication with positive number is order preserving
			addition and multiplication combination is distributive
			archimedean property

Table 6.1: mérési skálák tulajdonságai

abályszerűség szabályzat szokás szokáserkölcs szokásnorma törvény törvényszerűség tradíció

Fizikai szint Biológiai szint Társadalmi szint

6.3 Ego és a világ

Motiváció, beállítódás, attitűd

Franz Dietrich és Christian List A reason-based theory of rational choice [209]

propozícionális attitűd olyan mentális állapot, amelyben egy személy valamilyen propozícióval (tehát olyan tartalommal, mely igazságértékkel rendelkezhet) áll relációban.

propozícionális attitűd: vélekedés és vágy a vélekedés doxasztikus attitűd a vágy konatív attitűd lásd: MEO-szöveg

belief, attitude, proattitude (Davidson)

motiváció Mérő, MacLelan

a négy motívum különbségét a szavazás paradoxonával (Downs-paradoxonal) bemutatni [214] nem racionális elmenni szavazni, hiszen az egyetlen szavazat befolyása a választás végső kimenetére elhanyagolható, ezért nem érdemes elmenni

ez egyfelől az egoista beállítódást feltételez, tehát altruista beállítódással már meg lehetne magyarázni, miért megy el valaki szavazni

másfelől sok esetben érzelmi elköteleződés miatt megy el valaki (bár ez miért érzelem, felindulás? - nem hirtelen felindulásból megy el, hiszen esetleg már hónapok óta készül rá)

lehet normatív kötelességből is szavazni

6.3.1 Célracionalitás

commitment a kulcs a döntésben (célracionális mérlegelésben) ott van a preferencia, a preferenciában pedig ott van az érzelem lásd: [691]

6.3.2 Értékracionalitás

Emotions: from psychological theories to logical formalization and implementation in a BDI agent [2]

Zajonc az érzelem, az érték és a preferencia fogalmait összemossa
[840] Zajonc, Feeling and thinking. Preferences need no inferences
[307] The Neural Bases of Cognitive Conflict and Control in Moral Judgment
Manipulations of Emotional Context Shape Moral Judgment [767]
The Roots of Morality [507]

6.3.3 Érzelem

az érzelem és értelem különbségéről: hiába tudja valaki az eszével, hogy a repülés a legbiztonságosabb közlekedési eszköz (a statisztikák alapján fajlagosan ez jár a legkevesebb halálos áldozattal), a félelem (az érzelem) ennek ellenére a hatalmába kerítheti őt.

propositional attitudes are not the only cognitive states. A more basic feature of cognition is that is has a ?mind-to-world direction of fit.? The expression is meant to sum up the contrast between cognition and the conative orientation, in which success is defined in terms of the opposite, world-to-mind, direction of fit (Searle 1983). We will or desire what does not yet exist, and deem ourselves successful if the world is brought into line with the mind's plan

cognitive and conative functions

James-Lange elmélet minden fiziológia

Stanley Schacter and Jerome Singer (1962) kísérletei: nem minden az. A kísérleti alanyoknak befecskendezett stimuláló anyag attól függően váltott ki eufórát vagy dühöt a páciensekből, hogy a milyen mintákat tapasztaltak a környezetükben más embereken.

Zajonc az affektív ítélet: - inescapable - elkerülhetetlen - irrevocable - visszavonhatatlan - implicate self - - difficult to verbalize - nehezen verbalizálható - nem kell a gondolkodástól függenie - külön válhat a tartalomtól

impervious

When we evaluate an object or an event, we are describing not so much what is in the object or in the event, but something that is in ourselves.

discriminanda preferenda

Paivio duál kódolás elmélete imagens és logogens

"Az affektív válaszkészség – eltérően a nyelvtől vagy a gondolkodástól – általánosan megtalálható valamennyi állatfajnál. Egy nyúl egy kígyóval szembekerülve nem vizsgálhat meg minden szóbajöhető lehetőséget a kígyóval kapcsolatban, hogy azokból sikeresen következtessen a kígyó támadására, annak időpontjára és irányára. A nyúl nem állhat meg a kígyó hosszáról, méregfogáról vagy bőrénekgeometriaia mintázatáról elmélkedni. Ha a nyúl menekül, ezt jóval az emlékezetijegyek mérlegelése előtt kell megtennie; valójában az előtt, mielőtt megállapította és tisztázta volna, az elkövetkező mozgást, amire egy összetekert kígyó képes lehet. A futási elhatározásnak minimális kognitív emlékek alapján kell születnie." ([840] 16.o.)

A New Hampshire-i Egyetem kutatói John Sparrow vezetésével harminc csecsemő bevonásával végezték el a legújabb kísérletet: a csecsemőket és szüleiket különböző helyzetekbe hozták, majd elemezték a videóra rögzített eseményeket. A kutatók megállapították, hogy fél- és egyéves koruk között a csecsemők a szüleik reakcióit figyelve tanulják meg, hogy mi a mulatságos.

http://minerva.elte.hu/gaboros/erzelem1718.htm

ok-os, ok-talan
??????????????????????????????

„Because unreflective action is typically rational, we normally do not need to consciously consider our reasons, stand back from our desires, and deliberately choose to do what is most rational, what we have most reason to do. We need not do this because acting without deliberation nearly always reflects the reasons we have or the values of the actions and their likely outcomes at the time for us." [301, 2]

„We can trust our unreflective decisions because they reflect our motivations and hence the reasons that derive from our motivations. We need not deliberate because our actions automatically flow from our desires or concerns, and our reasons too derive from these motivations, reflecting shifting priorities among immediate desires, which in turn derive from and specify our longer standing, deeper, and more general concerns." [301, 2-3]

„Serious deliberation occurs when one is pulled in opposite directions by conflicting reasons affecting important concerns and relating to the same prospect." [301, 3]

„They hold that we are rationally motivated because we recognize the independent force of reasons, while internalists believe that reasons are such because they motivate us or relate to previously existing motivations. The question, in short, is whether, ultimately, motivations generate reasons or reasons generate motivations." [301, 9]

„" [301,]

„" [301,]

„" [301,]

„" [301,]

„" [301,]

„The function of paradigmatic or basic emotions is to displace (temporarily) ordinary rational calculation based on such reasons for and against actions. Paradigm emotions are irruptive motivational states. Their adaptive value lies in their being reactions to events for which rational calculation, even of the more automatic sort, is too slow. On the basis of very rapid and coarse cognitive evaluations, they prompt a limited repertoire of behavior that has been successful in the evolutionary past." [301, 87]

„In regard to the paradigm emotions, fear, for example, bypasses ordinary rational calculation to prompt immediate reactions to perceived dangers. Its adaptive value lies in its immediacy. It prompts sudden movements away from more objects than there is good reason to flee, but from an adaptive point of view this is a good play-it-safe strategy. Better safe than sorry."[301, 88]

> „But while emotions may have adaptive value, they do not give rise directly to reasons as desires do, and actions based on them do not typically arise in response to or reflect reasons despite being often in accord with them, as when one flees a dangerous object out of fear. Nor do emotions alter in response to reasons as readily as beliefs do. They can sometimes be cognitively suppressed or controlled, but one may well continue to feel fearful, angry, or envious after deliberately concluding that one ought not to. Finally, emotions do not typically cohere with each other or with the stable motivational set that generates one?s practical reasons."[301, 88]

sentiment

affekció affektivitás affektus alkat (lelki alkat, héxisz) alkat (testalkat - Mauss) antipátia diszpozíció ellenszenv emóció empátia érzelem (emotion) érzület észjárás (habit of mind) felindulás gerjedelem habitus hajlam (disposition) hajlandóság hangulat (mood) indulat intellektuális szenvedély (Polányi Mihály) kedély kedélyállapot kedv kellem kény lelkület megindulás mentalitás mód passzió pátosz prediszpozíció személyiség szentiment szenvedelem szenvedély (passion) szimpátia temperamentum tendencia

Damasio-kritika

Damasio [183] szomatikus markerek segítenek a racionális gondolkodás számára azzal, hogy a korábbi tapasztalatokat érzelmeken, zsigereken keresztül közvetítik

Mérő alapérzelmek kognitív érzelmek (reflexiót ? kívánnak, test nélkül, az elméből is lehetnek) emberi érzelmek (self jelenlétét kívánják)

Mérő habituálódás

Hirschman szenvedélyek és érdek könyve [356]

50.o. idézet fontos

érdek az ésszerű önzés

Orthmayr, szenvedélymentesség [553] Steiger Kornél válasza, [710]

Susan James, Passion and Action [375]

szenvedély, érzelem, vágy, indulat vs. akarat, értelem, megértés, cselekvés

Susan James, szenvedélyek vágyak cselekvés, [374]

a szenvedély nincs az akarat közvetlen irányítása alatt

nem lehet akarni, hogy szeressünk, csodáljunk, tiszteljünk, féljünk valakit

Descartes: tevékenység és tétlenség cselekvés és szenvedély

Hobbes: a cselekvést megelőző gondolatokat 'rendszerint törekvésnek nevezzük'. "Ez a törekvés, amikor valamire irányul, étvágynak vagy vágynak nevezik, amikor valamitől távolodni igyekszik, általában ellenszenvnek hívják." Hobbes, Leviathan 38.o.

"akinek nincs vágya, az halott" Leviathan, 54.o.

Paul Ekman szerint az érzelem elsődleges funkciója "felkészíteni" az egyént az interperszonális helyzetek gyors kezelésére a múltban sikeresnek, adaptívnak bizonyult megoldások, tevékenységek "felidézésével, újrahasznosításával" [228, 46].

TenHouten tisztázza az egymáshoz kapcsolódó "érzelemfogalmak" jelentését [746]: "affect realm"

Damasio Spinoza könyve: érzelem és érzés megkülönböztetése, érzelem az egyén és a környezete érzés az agy és

people react emotionally to objects and events, usually in social interactions, and these emotional reactions are followed up by a pattern of feelings whose necessary components include some levels of pain and pleasure.

feelings refer to a person?s own state of mind, especially with reference to an evaluation of what is agreeable and disagreeable, pleasant or unpleasant.

Emotions involve actions and movements, often in public view, revealed in facial expression, posture, gesture, specific behaviors, and conversation. Feelings, which follow emotions, are in comparison private, playing out not in the body but in the mind and at a higher level. az érzelem rövidebb idejű, pár perces, pár órás az időtartama, a többi kategória hosszabb

sentiment Steve Gordon (1981) meghatározása szerint: sentiments are emotions with long duration érzület, érzékenység, hosszabb időszakon keresztül jellemző

mood - kedély, kedélyállapot, hangulat, temperamentum

affect, affective traits személyiségjegy, hosszabb ideig (hónapokig, évekig, akár egy egész életen át) áll fent

TenHouten egy táblázatba rendezte [746, 4] a legfontosabb szerzők alapérzelem-listáit[1]:

Plutchik átveszi Darwin nézetét az érzelmekre vonatkozóan, miszerint az érzelmeknek adaptív funkciójuk van

négyféle életprobléma van, amelyekkel szembe kerülve minden állat, így az ember is megfelelő érzelmekkel reagál

identitás, temporalitás (reprodukció), hierarchia, territorialitás

MacLean négydimenziós modellje kísértetiesen hasonlít Plutchik által felvázolt elméletre. TenHouten 118. o. előtt

három kémiai vegyület és az alapérzelmek kapcsolata:

Picard 1995-ben először írt az érzelmek mérnöki, számítógéptudományi vizsgálatának lehetőségéről, szükségességéről [571]

Economy of Esteem [122]

Esteem Engine [121]

makacs érzelmek [603]

Elster érzelmek és a racionalitás [230]

Bernard Williams, Deciding to believe [818] belief - 1) mental state, 2) a content

6.3.4 Habitus

diszpozíció észjárás ethosz (szokás) éthosz (lelki alkat) explicit tudás implicit hit, hiedelem, előfeltevés, tudás jártasság hajlam hallgatólagos tudás, belefoglalt tudás, implicit tudás (tacit knowledge) jellem készség képesség habitualizáció habitus hészix rejtett tudáselemek személyes tudás szokás társadalmi létfeltételek tudatalatti (tudás) tudatos tudás

opus operandi vs. modus operandi

mivel a hallgatólagos tudás a mindennapi fizikai tevékenységek, az eszközökkel való bánás mellett a társadalmi kapcsolatokban és interakciókban megjelenő készségekben és jártasságokban is testet ölt, benne szintén megvan az aktív és a passzív oldalnak ugyanaz a kettőssége, amelyet már Arisztotelésznél is láthatunk, így ez a tudás is egyszer [735]

Polányi Mihály Szívós Mihály [735]

Body techniques Mauss describes 'techniques of the body' as highly developed body actions that embody aspects of a given culture. Techniques may also be divided by such as

[1] TenHouten táblázatát kiegészítettem két további, tőle származó sorral, a kínai tradíció és Darwin nézeteivel.

szerző	Proposed primary emotions							
Descartes (1647)			joy	sadness				wonder, love, hatred, desire
Tomkins (1962, 1963)	fear-terror	anger-rage	enjoyment-joy	distress-anguish	interest-excitement	surprise-startle	disgust, dissmell	shame-humiliation
Plutchik (1962, 1980)	fear	anger	joy	distress-sadness	anticipation	surprise	acceptance-trust	disgust
Osgood et al. (1975)	fear	anger	joy, quiet pleasure	anxiety-sorrow	interest-expectancy	amazement	boredom, disgust	
Arieti (1970)	fear	rage	satisfaction					tension, appetite
Izard (1972)	fear	anger	joy	distress	interest	surprise	disgust	shame (shyness, guilt), contempt
Ekman (1973, 1980)	fear	anger	happiness	sadness		surprise	disgust	
Emde (1980)	fear	anger	joy	sadness + distress	interest	surprise	disgust	shame, shyness, guilt
Scott (1980)	fear	anger	pleasure	loneliness + anxiety				love
Panksepp (1982)	fear	rage			expectancy			panic
Epstein (1984)	fear	anger	joy	sadness				love
Trevarthen (1984)	fear	anger	happiness	sadness				
Johnson-Laird, Oatley (1992)	fear	anger	happiness	sadness			disgust	
Turner (2002)	aversion-fear	assertion-anger	satisfaction-happiness	disappointment-sadness				
kínai tradíció	fear	anger	happiness	sadness				
Darwin	fear	anger	happiness	sadness		surprise	disgust	

gender and class (for example in the manner of walking or eating). These include such as eating, washing, sitting, swimming, running, climbing, swimming, child-rearing, and so on. The techniques are adapted to situations, such as aboriginal squatting where no seats are available. Techniques are thus a 'craft' (Latin: habilis) that is learned. The teaching of these methods is what embeds the methods and the teaching is embedded within cultures and schools of teaching. A pupil who becomes a teacher will likely teach what they are taught. Norbert Elias and Pierre Bourdieu developed the ideas further in habitus, the non-discursive aspects of culture that bind people into groups, including unspoken habits and patterns of behavior as well as styles and skill in body techniques.

Bourdieu, Pierre. 1977. Outline of a Theory of Practice, Cambridge: Cambridge University Press Elias, N. (1978). The History of Manners. The Civilizing Process: Volume I. New York: Pantheon Books Elias, N. (1982). Power and Civility. The Civilizing Process: Volume II. New York: Pantheon Books Mauss, Marcel. 1934. Les Techniques du corps, Journal de Psychologie 32(3-4). Reprinted in Mauss, Sociologie et anthropologie, 1936, Paris: PUF, angolul: [475]

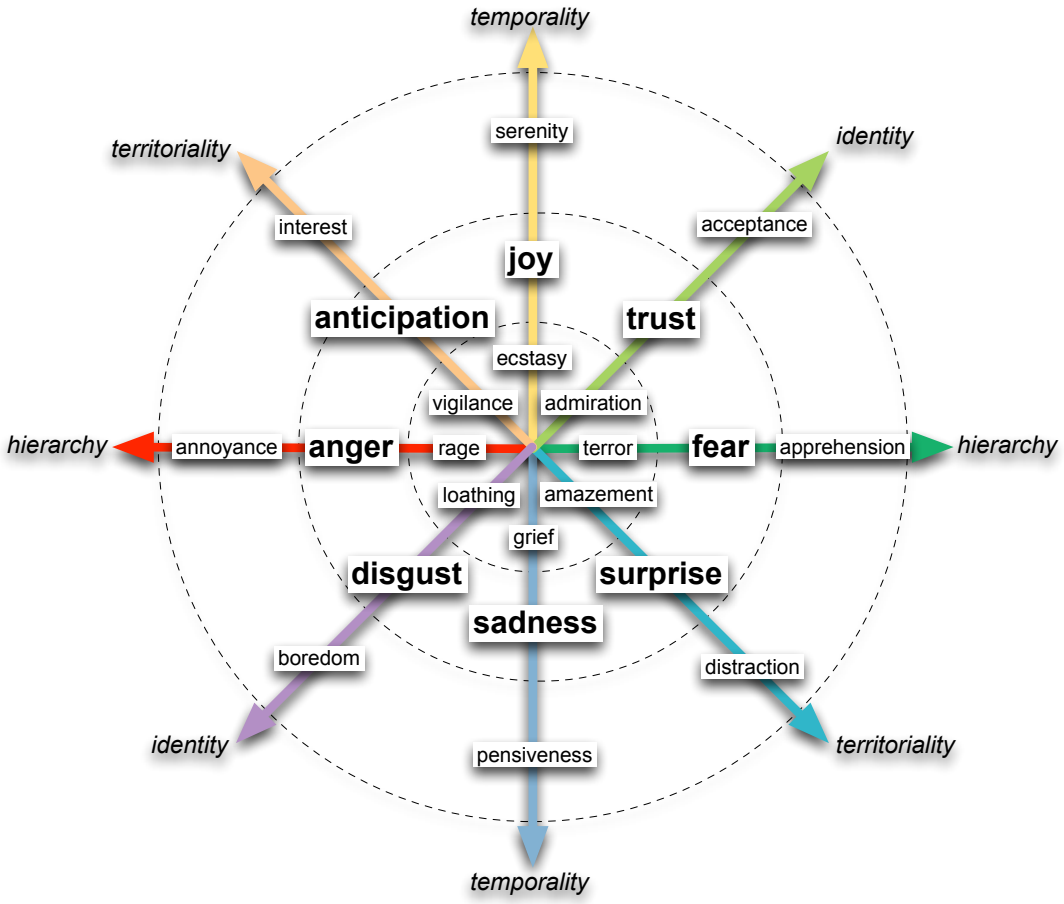


Figure 6.1: Plutchik alapérzelmei

Pokol Béla, Bourdieu elméletének alapkategóriái, a habitus [578]
"tudattalanított rutin"
egyénre zuhanó döntési teher
habituson belüli diszpozíciók
Habitus, Freedom, and Reflexivity [354]
Bourdieu and Wacquant [109]
Loïc J.D. Wacquant [802]

Pierre Bourdieu habitusértelmezése

"A habituson olyan tartós és átvihető hajlamok rendszerét értjük, amely, valamennyi múltbeli tapasztalatot beépítve, minden pillanatban mint *észlelési, értékelési és cselekvési minta* működik."[108, 213]

"... a habitus nem más, mint emberi természetté vált történekem, vagyis második természetté vált és így letagadott történelem. A „tudatalatti" valójában nem más, mint a történelem elfelejtése, ami abból fakad, hogy a történelem a habitusokba, e kvázi természetekbe ágyazza bele az általa létrehozott struktúrákat."[108, 215]

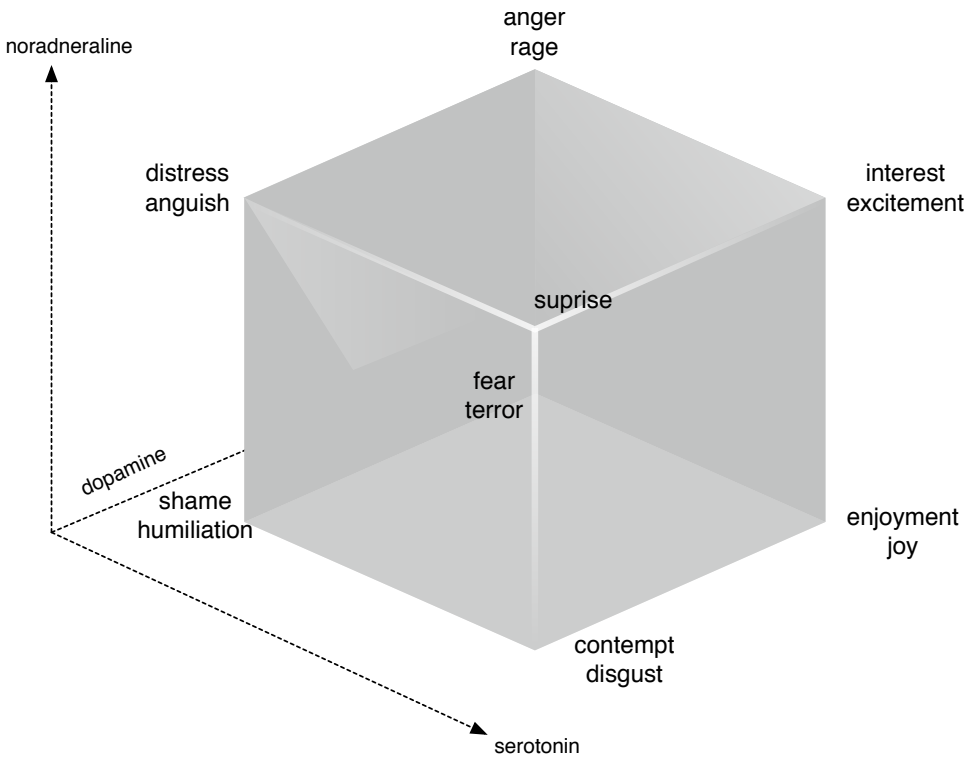


Figure 6.2: Löwheim érzelemtipológiája

Polányi Mihály személyes tudás fogalma

akarat és döntés és neuronok: http://www.informationphilosopher.com/solutions/scientists/eccles/

Duhigg

[222]

Eugene Pauli, szokáshurok (habit loop)

Larry Squire,

Claude Hopkins, szokásteremtés "Tudományos reklámozás"

cue-rooutine-reward + craving

Control and automaticity in social life: [811]

controlled vs. automatic process controlled process 4 jellemzője: 1 a conscious intention of what the control will accomplish 2 a sense or feeling of control 3 an expenditure of effort in the control output 4 a (closed-loop) monitoring of the control output

automatic process 1 autonomous (once started it runs by itself without the need for conscious guidance) 2 fast 3 efficient (requires little attentional capacity)

---

a könyv elején olvasható Whitehead idézet – amelynek tartalmával mélységesen egyetértek – azt sugallja, hogy a gondolkozásunk során is rengetegszer rögzített gondolatmeneteken, hivatkozásokon keresztül haladunk.

Az előszóban már idéztem Neumann János, aki azt mondta, „ ... a matematikában az ember nem megérti a dolgokat, hanem megszokja.”

6.3.5 Miért kell a racionalitás fogalma és miért hamis a mítosza?

Dan Kahan és munkatársai, Motivated Numeracy and Enlightened Self-Government című cikkükben [389]

When Corrections Fail: The persistence of political misperceptions [543]

Áttekintő ábrán a különböző motivációk, beállítódások egymáshoz való viszonyát a logika hatszög segítségével

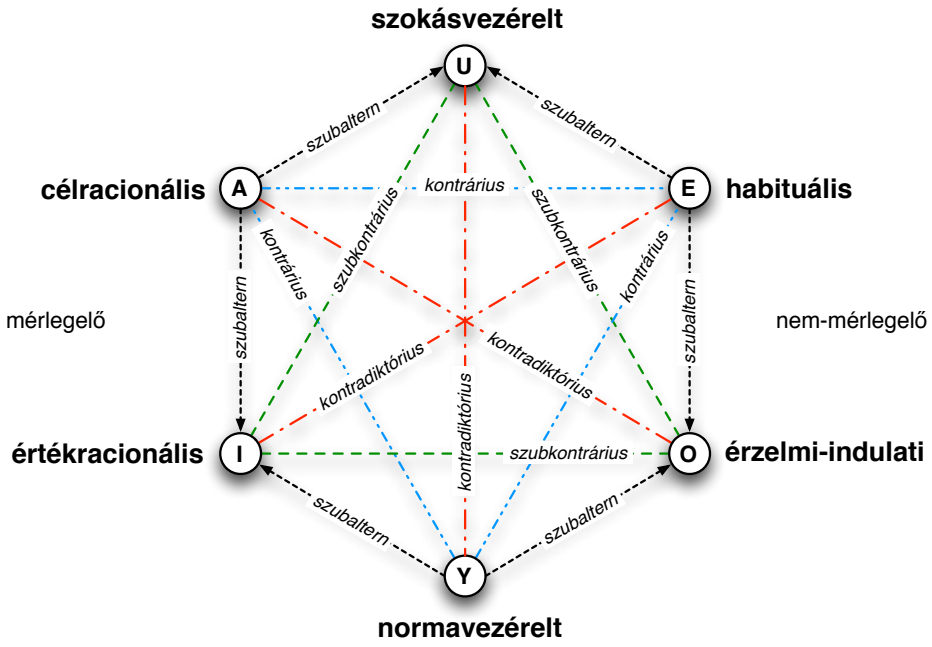


Figure 6.3: cselekvéstípusok beállítódás szerint, Max Weber

“Rationality may fail in two quite distinct ways. First, individuals often commit errors in logical inference even when doing their best to reason logically. Second, what is quite a different matter, actions are often "unthinking"; when governed by habit or passion, people do not even attempt rational self-control. (I will be suggesting below that such failures of rationality, like violations of the self-interest postulate, may have proved functionally adaptive in the genetic and cultural evolution of the human species.)” [358]

Mérő könyve, Hirshleifer cikke a racionalitás, konzisztencia, a logika bukfenceiről

döntési paradoxonok a döntéselmélet területéről Tversky és Kahneman elmélete a korlátozott racionalitásról, kilátáselméletről

Herbert Simon korlátozott racionalitás [683]

raconalizálás, kognitív disszonancia “When a person is made aware of a disharmony be- tween his actions and his preferences and beliefs, the economist would expect him to revise his choice of action - but the cognitive dissonance theorist predicts that he is likely instead to modify his preferences or beliefs.” [358]

mert habituális vagy érzelmi módon cselekszenek

racionális önbecsapás [743]

Chapter 7

Társadalmi kapcsolat

> „*Sir Dannis Robertson* ... mondja, hogy a közgazdász feladata megmutatni, hogyan lehet minimalizálni a legszűkösebb erőforrás, a szeretet iránti igényt. Azt követelte közgazdásztársaitól, hogy azonnal kezdjenek figyelmeztető csaholásba, ha olyan javaslatokkal találkoznak, melyek hatékony megvalósításához szeretetre van szükség." [130]

A társadalmi cselekvés fogalmának elemzésekor kiderült, hogy a cselekvés fogalma annyiban szűkebb terjedelmű az egyszerű cselekvés fogalmához képest, hogy itt "megköveteljük" azt is, hogy a cselekvőn túl legyen még valaki, akihez a cselekvő – valamilyen módon – igazodik a cselekvésével (↬267.o.). Ekkor azonban ez a másik személy még csak közvetítő szerepet játszik: a társadalom potenciális hatását fejezi ki. Ez a másik személy a társadalom jelzése. Ahogy a cselekvő igazodik hozzá, úgy már a társadalmiságot érhetjük tetten. De ez még csak hatás, és nem kölcsönhatás. A társadalmiságban benne rejlő kölcsönösséget, kölcsönhatást, kölcsönös egymáshoz igazodást a **társadalmi kapcsolat** fogalmával lehet igazán jól megragadni.

> „Társadalmi "kapcsolatról" beszélünk, ha többek magatartása értelmi tartalmának megfelelően kölcsönösen egymáshoz igazodik, és magatartásukat ez a kölcsönös igazodás irányítja." [810, 54]

A cselekvés és a társadalmi cselekvés fogalmába már bele kellett építeni az **ágencia**, valamint az **intencionalitás** minőségét, a most következő feladat a **kölcsönösség**, és egy kicsit később majd az **interdependencia** összetevőjének felvétele az elméleti keretrendszerbe. A következő fejezetben a kölcsönös igazodásnak, egymásra figyelésnek egy különösen fontos esetét vizsgálom meg. Azt a beállítódást elemzem, amely szabályozza bennünk, hogy miként viszonyulunk magunkhoz (Egohoz) és másokhoz, a másikhoz (Alterhez) a cselekvéseink során.

7.1 Ágensre irányultság: társas preferencia

Sen, Goals, Commitment, and Identity: different components of privateness of behavior, specifically self-centered welfare, self-welfare goal, and self-goal choice. [657]

Egyszer New Yorkban egyedül sétáltam a Central Parkban, amikor elém termett egy fiatal gyerek, akiről lerítt, hogy pénzt akart kérni tőlem. Már nem emlékszem pontosan, mi, hogyan történt. Olyasmit kérdezhetett, hogy honnan jöttem, mire azt feleltem neki, hogy Magyarországról ('from Hungary'). Lehet, hogy a kiejtésem nem volt tökéletes, lehet, hogy a fiú nem hallott még arról, hogy van ilyen nevű ország, de az biztos, hogy félreértett engem. Félreértett és megszánt. Nem mondott semmit, csak nézett rám egy ideig, miközben azt gondolta rólam, éhes lehetek ('hungry'). Ezt onnan tudom, hogy a zsebébe nyúlt, elővett egy csokit és odanyújtotta nekem.

Ezzel a történettel a közgazdászok nem igazán tudnának mit kezdeni. A közgazdaságtan követőinek túlnyomó részére ugyanis nagyon jellemző az a beállítódás, amely szerint az ember – gazdasági – cselekvését az 'önző ember' hipotézisére támaszkodva lehet (kell) megmagyarázni. Pedig ez nem igaz. Tagadhatatlan, hogy sokszor, nagyon sokszor valóban működőképes a hipotézis. Lehet, hogy az esetek nagyobb részében. Talán. De ugyancsak sokszor, sőt, nagyon sokszor találkozhatunk olyan helyzetekkel, amelyekben önzetlen beállítódások mentén cselekszenek az emberek. A társadalmi cselekvés modelljébe ezt a kétfajta beállítódást mindenképpen fel kell vennünk. De ha már beszélünk **önző** és **önzetlen** beállítódásokról, akkor érdemes azt is megvizsgálni, nem lehet-e általánosabban kezelni ezt a kérdést, nem lehet-e egy olyan elméleti keretet felállítani, amelyben az egoizmus és altruizmus is értelmezhető, de a modellbe fel lehet venni további fogalmakat is. A beállítódás, a motiváció fogalmát kell tehát pontosítani, és ezen belül lehet majd értelmezni a cselekvők egymáshoz való viszonyát kifejező fogalmakat. Utóbbiak terjedelme szűkebb a beállítódás általánosabb fogalmának terjedelméhez képest, meg kell tehát mondanunk azt, hogy milyen jellemzők mentén tudjuk meghatározni, tipizálni az ágenciával kapcsolatos beállítódásainkat.

A beállítódásaink, motivációink sok esetben abban az értelemben (és irányban) szűkülnek le a cselekvéseink során, hogy gyakran "csak" arra figyelünk, hogy a cselekvéseink következményei hogyan érintik a társadalmi kapcsolat résztvevőit, az ágenseket (Egot és Altert). A kérdés itt az, hogy a cselekvő mennyire veszi figyelembe azt, hogy a döntése (cselekvése) hat(hat) részben saját magára, részben a társára (a másikra, a többiekre). Kézenfekvőnek tűnik erre felállítani azt a fogalmi kettősséget, amely szerint a cselekvő (Ego) lehet önző vagy önzetlen. Ezek a fogalmak azt ragadják meg, hogy a döntés/cselekvés következményei mennyire jók, előnyösek vagy károsak, előnytelenek az Ego, illetve az Alter számára. Ha az Ego a hasznok maximalizálása, illetve a károk minimalizálására törekszik (és csak arra), akkor egoizmusról ('Ego'-izmusról), egoista beállítódásról beszélhetünk, ha döntései során figyelembe veszi Alter szempontjait is, akkor altruizusról ('Alter'-izmusról), altruista beállítódásról. Ha így gondolkozunk, akkor az önző, egoista (selfish, egoist) és az önzetlen, altruista (unselfish, selfless, altruist) beállítódás egymás ellentétének, komplementerének tekinthetjük.

Ez az ellentét azonban megszűnik, pontosabban átalakul, ha a másikhoz (Alterhez) való értékviszonyulást úgy kezeljük, hogy eközben figyelünk az értékelés (mint olyan)

szükségszerűen **bipoláris** jellegére. Ha feltételezzük, hogy a döntés/cselekvés nem mindig az Ego és/vagy Alter javára irányulhat, de előfordulhat az is, hogy bizonyos helyzetekben Egonak vagy Alternek (vagy mindkettőnek) kára keletkezik, és ezt a lehetőséget is értékelni akarjuk, akkor a modellünket bővíteni, általánosítani kell ezen a ponton. Ha a bipolaritást figyelembe akarjuk venni, akkor a **jóindulatú** (benevolent) és **rosszindulatú** (malevolent) beállítódást kell szembe állítanunk egymással. A jóindulat az ágens javát, a rosszindulat az ágens kárát preferálja.

A bipolaritás modellbe emelése magával vonja azt, hogy az egoizmus-altruizmus ellentétpár helyett másfajta kettősséget kell alkalmaznunk. Az **ágensre irányultság** általános fogalmából kell kiindulnunk. Ennek segítségével a **magára irányuló** (self-regarding, self-serving), illetve a **másikra irányuló** (other-regarding, other-directed) fogalmakat használva megragadhatjuk a kétfajta ágens (Ego és Alter) irányába történő viszonyulás mozzanatát.[1] Ha ki akarjuk fejezni az ágensre irányultság pozitív és negatív típusait, akkor további fogalmakat kell felvennünk. Később részletesebben kifejtem, mivel, miért és hogyan érdemes bővíteni a fogalmi modellünket, de már előzetesen érdemesnek tartom itt jelezni, hogy az Alter felé irányuló rosszindulatú beállítódás jelzésére az **agresszió**, az Egora vonatkozó rosszindulatra pedig az **autoagresszió** (önagresszió) fogalmakat vehetjük igénybe. Ha így teszünk, akkor lesz két dimenziónk, az ágensre irányultság, illetve e viszony polaritása, valenciája. A két dimenzió mentén négyfajta beállítódást fejezhetünk ki: ha a jóindulat az Egora irányul, akkor egoizmusról, ha Alterre, altruizmusról, ha a rosszindulat az Egora, agresszivitásról, ha Alterre, autoagresszivitásról beszélhetünk. Kiindulásként felrajzolhatunk egy ábrát, amelyet persze a későbbiekben még sok minden mással ki kell egészíteni.

Az ágensre irányuló beállítódások, a szociális preferenciák ábrájára tekintve egyből hiányérzetünk támadhat amiatt, hogy ez a keret olyan fogalmakat határoz meg, amelyek a szociális preferenciák tiszta végpontjait jelentik egy-egy dimenzió mentén, de nem mutat semmit az átmenti állapotokból. Márpedig jogos az a feltételezés, hogy vannak, sőt, sokkal gyakrabban vannak olyan esetek, amikor valakinek a másikhoz való viszonya egy kicsit altruizmusból, egy kicsit egoizmusból áll össze, és van úgy, hogy az ember társas irányultsága olykor a pozitív, máskor inkább a negatív pólushoz áll közelebb valamelyik dimenzióban. Nagyon steril az az elképzelés, ami szerint vagy a más javát (kárát) vagy a saját javamat (káromat) akarhatom. Sokkal életszerűbb azt feltételezni, hogy a döntéseink során valamilyen módon és mértékben egyaránt figyelembe vesszük a magunk és a másik javát (kárát).[2] Ehhez viszont arra van szükség, hogy a fenti ábrát (és az eddigi modellünket) további fogalmakkal kiegészítsük.

[1] Stilisztikai szempontok miatt, ha nem zavaró, használni fogom még a *másra irányuló beállítódás* kifejezést is. Az ágensre irányuló beállítódásokat szokták még **szociális preferenciáknak** (social preferences), **társas motivációknak** is nevezni, ezen belül pedig a magára irányulást a **proszelf** (pro-self), a másikra irányulást pedig **proszociális** (prosocial) beállítódásként értékelik. Bár utóbbi kettő fogalom értelmezésében nincs igazán konszenzus, előfordulhat, hogy a továbbiakban – megint csak stilisztikai szempontok miatt, és csak akkor, ha nem okoz félreértést – olykor használni fogom ezeket a terminusokat is.

[2] A későbbiekben tárgyalandó játékelmélet számára sokkal természetesebb, használhatóbb ez a megközelítés.

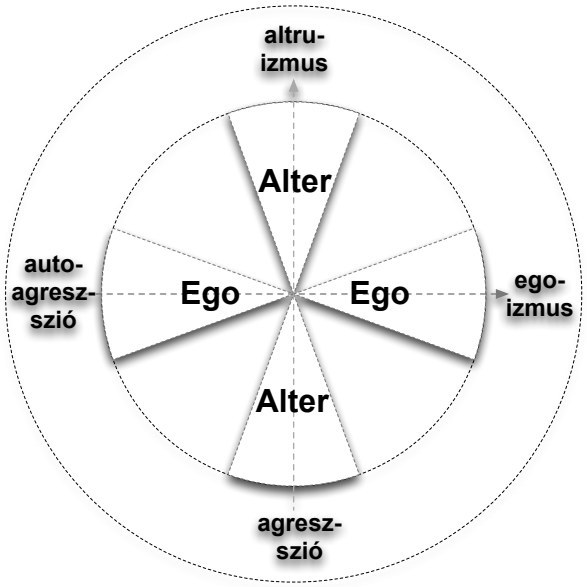


Figure 7.1: magára irányuló, illetve másikra irányuló beállítódások

7.2 Magasabbrendű beállítódás: büntetni akarás

Az előző fejezetben a szociális preferenciák osztályozásához használható két dimenziót különítettem el. Az értékirányultságok bipoláris jellegének megragadására a jóindulatú, illetve rosszindulatú beállítódások szembeállítását javasoltam, hogy ezáltal definiálhassuk az agresszió (másiknak kárt okozni) és az autoagresszió (magunknak kárt okozni) fogalmát. Később még tárgyalni fogom az agresszivitás jelenségét, és akkor kitérek arra a kérdésre, hogy hol érhetjük tetten és hogyan magyarázhatjuk az agressziót az állati és emberi társadalmakban. Szükségesnek érzem azonban, hogy egy fontos kiegészítést, vagy inkább valamiféle pontosítást tegyek ezen a ponton. Ha ugyanis az agresszivitáson azt értjük, hogy nem az ágens javát, hanem kárát akarja valaki, akkor e jelenség létezését sok esetben csak úgy tudjuk megmagyarázni, ha figyelembe vesszük azt a tényt, hogy az ember büntetni akarja magát és/vagy a másikat. Ha azonban a büntetés jelenségét is figyelembe vesszük itt, akkor az agresszivitás és a rosszindulat kapcsolatáról is kicsit másként kell gondolkoznunk.

A társadalmi cselekvésekről szólva korábban már említettem, hogy a cselekvés fogalmát magasabb rendben is értelmezhetjük (↬**??**.o.). Amikor az emberek szankcionálják, büntetik vagy jutalmazzák a másikat, akkor az ilyen cselekvéseket magasabbrendű fogalmakkal írhatjuk le, mert önmagukban nem értelmezhetők, csak akkor, ha valamilyen másik cselekvésre "vonatkoztatjuk". A büntetés viszont a legtöbb esetben valamiféle agresszió alkalmazását jelenti. Ebben nem különbözik a "sima" agressziótól. A büntető beállítódásában viszont lényeges különbség figyelhető meg: a büntető jóindulat mentén büntet, azaz saját büntető cselekvését jónak tartja. Kárt okoz a büntetett félnek, de ez a károkozás más.

Erről a fontos szemléletváltásról később még hozok példákat, a fogalmi elemzésben is mélyebben kitérek a problémára, itt csak annyit akarok jelezni, hogy az agresszivitás

jelenségét, az agresszív beállítódások létezését elég gyakran (bár nem mindig) a büntetési szándékkal tudjuk megmagyarázni.

———————-

A büntetni akarás magasabbrendű fogalom.

...

Hogy lehet a szociális preferencia fogalmát definiálni? Nyilván a PREFERENCIA fogalmából kell kiindulnunk.

Agent szelf Ego Alter

ASD

———————

7.3 Társas értékorientáció

> „We typically expect a rational agent to behave in a manner that best furthers its own interests. However, an artificial agent might be expected to act in the best interests of a user (or designer) who has somehow communicated its wishes to the agent." [110]

A korábbiakban felvettem a modellbe azt a két dimeziót, amelyek mentén leírhatjuk a szociális preferenciákat. Az ágensre irányultság, valamint az értékpolaritás tengelyei mentén négy "szélsőértéket" definálhatunk: egoizmus, altruizmus, agresszió, autoagresszió. Következő lépésként bővíteni szeretném a társas motivációk fogalmi modelljét. Ezen a ponton azonban egy – talán – szokatlan módszertani probléma nehezíti a továbblépést. A szociális preferenciák modelljének kiegészítése során a **társas értékorientáció** (social value orientation, SVO) elméletének eredményeit szeretném használni úgy, hogy a fogalmi építkezés logikáját nem törjem meg ezzel. A társas értékorientáció elmélete alkalmas erre, mert zökkenőmentes továbblépést biztosít. Egy olyan módszertant ad a kezünkbe, amely segítségével kísérleti úton lehet mérni az emberek társas beállítódásának irányát és mértékét (minőségét, mennyiségét). Történelmi tény viszont, hogy az egész elmélet azzal a céllal jött létre, hogy valahogy meg lehessen szabadítani a kísérleti mérési szituációkat az ún. interdependencia jelenségétől azért, hogy csak az egyéni cselekvők beállítódásait tudjuk tisztán mérni. Hogy az interdependencia pontosan mit jelent, és miért kell kilépni az ilyen helyzetekből, azt a játékelmélet alapfogalmainak ismeretében lehet elmagyarázni. Ezt viszont – a fogalmi modellezés építkezési logikája miatt – csak jóval később tudom megtenni. Ezt a problémát úgy tudom csak feloldani, hogy nagyon röviden itt is el fogom mondani azt, amit tudni kell a játékelméleti helyzetek leírásáról, hogy érteni lehessen a társas értékorientáció elméletének kapcsolódását a játékelmélethez.

A társas értékorientáció elméletének alapjait David M. Messick és Charles G. McClintock fektették le 1968-ban megjelent cikkükben [493]. A tanulmányból két fontos eredményt érdemes megemlíteni. Egyfelől elkülönítettek háromféle motivációt, társas értékorientációt egymástól, amelyeket **individuális**, **kooperatív** és **versengő** irányultságnak neveztek. Az individuális beállítódású ember csak a saját hasznát (own gain) nézi a döntései során, a kooperatív ember a maga és a társa együttes hasznát akarja maximalizálni (joint gain), míg a versengő típus arra tör, hogy minél nagyobb legyen kettejük közti különbség a maga javára (relative gain). Egy példát mutatva erre az alábbi táblázatban azt

ábrázolhatjuk, hogy a vizsgált alany (Ego) két döntési lehetőség közül választhat (X és Y), mindkét alternatívában a maga és a társa számára lehetséges **kifizetéseket** hasonlítja össze:

	X	Y
Ego kifizetése	5	8
Alter kifizetése	2	6

Ha a táblázatban szereplő számok megnyerhető pénzt jelentenek, akkor az X alternatíva választása azt jelenti, hogy Ego jobbnak tartja azt az állapotot, amikor 5 egységnyi pénzt kap, miközben Alter 2-t ahhoz az Y alternatívához képest, amely során Ego 8, Alter 6 egységhez jut. Egy ilyen helyzetben az individualista irányultság mentén Egonak az Y alternatíva a jobb, hiszen 8-at kap 5 helyett, és ugyancsak az Y-t fogja választani, ha kooperatív beállítódással rendelkezik, mert $(8+6) > (5+2)$. A versengő értékorientáció esetén azonban változik a helyzet, és ekkor már Ego az X alternatívát fogja preferálni, hiszen $(5+2) > (8+6)$. Egy ilyen kicsi táblázat arra is alkalmas, hogy különböző számokat írogatva bele, segítségével kísérleti helyzetekben egyszerűen mérni lehet az emberek a társas beállítódását.

A tanulmány másik eredménye az, hogy a három irányultság elkülönítése után a szerzők egyfelől megmutatják, hogy összesen hatféle módon lehet két-szereplős, két-alternatívás helyzetekben a háromféle beálíítódást reprezentálni, másfelől – és számunkra most ez a fontosabb – rámutatnak arra, hogy izomorf megfeleltetést lehet teremteni a szociális motivációk mérésére alkalmas táblázatok, illetve a játékelméletben használatos preferenciamátrixok között. Erre pedig azért van (volt) szükség, mert a játékelmélet azokat a társadalmi interakciókat (játékokat) próbálja megragadni és megmagyarázni, amelyekben a résztvevők (játékosok) együttes döntéseitől függ a helyzet megoldása. A játékok mindig olyan helyzeteket írnak el, amelyekben hiába tudjuk, hogy külön-külön mit szeretnének elérni a játékosok, a végeredmény mindkettejük döntésétől függ. Erre mondjuk azt, hogy interdependencia, kölcsönös függőség jellemzi ezeket a kapcsolatrendszereket. Mindez azért érdekes a társas értékorientáció elmélete szempontjából, mert a játékelméleti helyzetekben a kölcsönös függőség létezése miatt sosem tudhatjuk biztosan, hogy miért döntött valaki így vagy úgy a játékban (mert a játék végeredménye mindig mindkét fél döntése alapján alakul ki). Ezt a eltüntetni lehet feloldani a fent jelzett izomorfia segítségével. Egy adott játékelméleti játszma helyett ugyanis elegendő a neki megfeleltethető egy-személyes kísérleti vizsgálatot elvégezni (SVO-vizsgálat), hogy megtudjuk a kísérleti alany beállítódást. A két-személyes játékelméleti játszmákat így vissza lehet vezetni egy-személyes SVO-kísérletekre. Erre az egyszerűsítésre vezették be a szerzők a dekomponált játék (decomposed game) fogalmát. Hatféle dekomponált játékot különítettek el egymástól, amelyek mind izomorfak valamelyik hagyományos játékkal. A megfeleltetést az alábbi kapcsolatrendszer mentén végezték el:

	X	Y
Ego kifizetése	c_1	c_2
Alter kifizetése	d_1	d_2

$\mapsto$

	X_{Alter}	Y_{Alter}
X_{Ego}	$(c_1+d_1),(c_1+d_1)$	$(c_1+d_2),(c_2+d_1)$
Y_{Ego}	$(c_2+d_1),(c_1+d_2)$	$(c_2+d_2),(c_2+d_2)$

Az izomorfiára konkrét példaként a fogolydilemma játék esetét mutatom be. A következő bal oldali SVO-táblázatba beírt négy érték – fenti "kulcs" alapján – pontosan megfelel-

tethető a jobb oldali táblázatban látható preferenciamátrixnak, ami pedig a jól ismert fogolydilemmához tartozik.

	X	Y
Ego kifizetése	1	2
Alter kifizetése	2	0

$\rightarrowtail$

	X_{Alter}	Y_{Alter}
X_{Ego}	3,3	1,4
Y_{Ego}	4,1	2,2

Messick és McClintock konverziós megoldását használva a hetvenes évektől kezdve rengeteg kísérletet végeztek, amelyekben a társas motivációkat vizsgálták sokféle szempont alapján. Az induló cikk megjelenése után pár évvel McClintock bővítette a modelljüket és a 'motivációs sémák' közé felvette az 'altruista' irányultságot is [482], ami addig nem volt benne a modellben. Ekkortól tehát már négy motívumot (orientációt) különítettek el egymástól. A négyféle beállítódást úgy mutatom be alábbi táblázatban, hogy (a táblázat utolsó oszlopában) melléjük rendelem azt a döntési alternatívát, amelyeket az adott irányultságú kísérleti személy választana magának az ott látható négy lehetőség közül (a zárójelbe tett két érték közül az első az Ego, a második az Alter nyereményét mutatja).

individualizmus	saját nyeremény maximuma	max(Ego)	(100,70)
kooperáció	együttes nyeremény maximuma	max(Ego+Alter)	(90,90)
versengés	nyereménykülönbözet maximuma	max(Ego-Alter)	(80,40)
altruizmus	másik nyereményének maximuma	max(Alter)	(70,100)

Az (Ego+Alter) kifejezés a két ágens együttes hasznát jelenti, míg az (Ego-Alter) kifejezés az Ego relatív hasznát Alterhez képest. A 'max' függvényjel pedig azt fejezi ki, hogy az adott irányultság a jelzett haszon maximumát szeretné elérni.

A versengés fogalmát McClintock a bevett felfogáshoz – főleg a játékelméletben megszokott értelmezéshez – képest rendhagyó módon értelmezi, amikor a nyeremények különbségére teszi a hangsúlyt. Ez az orientáció kisebb nyereményt ér el, mint a tisztán önző, individualista beállítódás, viszont az lehet az "előnye", hogy a másikhoz képest jobban jön ki a helyzetből (hiszen így többet szerez a partnerénél, míg az individualista döntés után). Ez a hozzáállás értelmezhető úgy is, mint a kisebb veszteségre való törekvés. A játékelméletben versengésen az itteni felfogás helyett sokkal inkább az individualista beállítódást értik, de ezt többértelműség nem lesz annyira zavaró a későbbiekben, mert a játékelmélet modellje nem képes ezt a finom társas motivációs felosztást "kezelni". Az SVO jól elkülönített versengés és individualizmus fogalmai gyakorlatilag egybemosódnak a játékelmélet világában. Érdemesnek tartom viszont a társas értékorientáció individualizmus terminusát kicserélni az *egoizmus* terminusra a fogalom jelentésének meghagyása mellett. Az indokom csak annyi, hogy az individualizmus fogalmához sok más konnotációs réteg is tapad, ami az egoizmussal kapcsolatban nem jelentkezik, valamint az egész cselekvéselméleti modell legfontosabb ágensszerepéhez, az 'Ego'-hoz sokkal jobban illeszkedik hangulatában az egoizmus terminusa akkor, amikor a magára irányuló beállítódást akarom kifejezni.[3]

[3]Még egy indokom a terminuscsere mellett az, hogy az individualizmus fogalma az evolúcióbiológiában széles körben használt, ami szemben áll a társadalomtudományok által prefereált egoizmus terminusával. Ebben a tudományszakok közti terminológiai versengésben – ahol csak lehet – inkább saját szakterületem gyakorlatát követem.

A társas motivációk elméletében két változásról kell még említést tennem. Az első változás az elmélet modelljét érintette, ugyanis 1973-ban Donald W. Griesinger és James W. Livingston azt javasolták [309], hogy az addigi motivációs sémákat érdemes lenne kibővíteni. Mivel a bővítés szemléltetésére geometriai módszert alkalmaztak (a nyolc eleműre bővített motivációs listát egy körön elhelyezve mutatták be), ezért pár évvel később Wim B. G. Liebrand ezt a reprezentációs technikát beemelte az elmélet központjába [448]. Az új geometriai keretrendszer, a McClintock-féle motivációs gyűrűt (Ring measure) már lehetővé tette, hogy ne kategoriális, hanem már mértéken alapuló nominális változóként lehesen a vizsgált beállítódásokat kezelni.

———————-

DEdedededededededee

de Griesingerék csak hat motivációról beszéltek, és gyűrű sem volt még, csak geometriai ábrázolás, koordinátarendszer

———————

Amikor Griesinger és Livingston négy új motiváció beemelését tartotta szükségesnek, akkor arra hivatkoztak, hogy a szociális irányultságok lehetnek negatív valenciájúak, és az elmélet már elfogadott négy beállítódástípusa mellé (azokkal szemben) fel kell venni további négy típust, a **szadizmus** (sadism), a **mazochizmus** (masochism), az **önfeláldozás** (martyrdom) és a **szadomazochizmus** (sadomasochism) kategóriáit. Ezekben az esetekben a döntéshozó (Ego) negatív módon "áll a dolgokhoz", tehát a haszon helyett az ágens(ek) kárát akarja. Ezt úgy fejezhetjük ki, hogy a haszon maximalizálása helyett a haszon minimializálása (a kár maximalizálása) a "célja".[4] A modell bővítésének szükségességét elfogadva, de az alkalmazott terminusokat kicserélve a korábban már bemutatott, általam preferált terminusokra, az alábbi táblázatba rendezhetjük az új kategóriákat:

autoagresszív (mazochizmus)	saját nyeremény minimuma	min(Ego)	(10,50)
destrukció (szadomazochizmus)	együttes nyeremény minimuma	min(Ego+Alter)	(20,30)
önfeláldozás (mártíromság)	nyereménykülönbözet minimuma	min(Ego-Alter)	(30,80)
agresszió (szadizmus)	másik nyereményének minimuma	min(Alter)	(50,20)

A táblázatban zárójelben feltüntettem az eredeti névjavaslatokat is, és négy újabb alternatíva feltüntetésével jeleztem azt is, hogy az egyes irányultságok mentén melyik lehetőségeket választanák a kísérleti alanyok.

Ha egyesítjük az eredeti (négyes) és az új motivációs sémakészletet, akkor egyszerre láthatjuk őket a melléjük rendelt, rájuk jellemző kifizetésekkel együtt. A könnyebb összehasonlítás végett a kifizetéseket teszem a táblázat első oszlopába:
A nyolc szociális preferencia közül a Griesinger-Livingston-féle negatív motivációk modellbe emelését közel sem minden kutató helyeselte. Az agressziót mára már a legtöbb kutató elfogadta, de a további három negatív beállítódás mérését sokan fölöslegesnek, megalapozatlannak tartották (és tartják). Ezzel szemben a társas értékorientáció elmélet egyik jeles képviselője, Paul A. M. van Lange egy új irányultságot is felvett a McClintock-féle motivációs gyűrű elemei közé. A kooperáció mellé javasolta betenni az egyenlőségre

[4] A kármaximalizálás vs. haszonminimalizálás fogalmai között ekvivalenciát teremthetünk ilyenfajta megfeleltetéseken keresztül:
max(Ego-Alter) = max(Ego) & min(Alter)
max(Alter-Ego) = min(Ego-Alter) = min(Ego) & min(-Alter) = min(Ego) & max(Alter)

(100,70)	egoizmus	max(Ego)
(90,90)	kooperáció	max(Ego+Alter)
(80,40)	versengés	max(Ego-Alter)
(70,100)	altruizmus	max(Alter)
(10,50)	autoagresszió	min(Ego)
(20,30)	destrukció	min(Ego+Alter)
(30,80)	önfeláldozás	min(Ego-Alter)
(50,20)	agresszió	min(Alter)

törekvő *egalitarianizmust* (egalitarianism) mint a proszociális beállítódás másik megnyilvánulási formáját [782], azonban egy, a munkatársaival írott, 2007-es cikkében a kooperatív és egalitárius irányultságot már összevonja [783]. Viszont ők is csak 5 orientációról beszélnek (a négy pozitív motiváció mellett vizsgálják még az agressziót).

A SVO-kísérletek egyik nagy hozadéka az az ismeret, amit a különböző szociális preferenciák elterjedtségéről tudnak mondani a kutatók (ilyen eloszlásra vonatkozóan adatokat közöl például[784]). A kísérletek tanusága szerint a proszociális (kooperatív, az egalitárius és altruista) beállítódás a teljes népesség majd 60%-ra jellemző, akik között több a nő, és az életkorral együtt nő az proszociális irányultság aránya is. A sorozatos felmérések ismétlődő tapasztalata, hogy elenyésző számban vannak azok, akik tisztán altruista szempontok szerint döntenek. A proszelf beállítódású emberek durván 35-40 %-nyian vannak, akik közül az egosita (individualista) motivációjú személyek közel 30 %-ot tesznek ki. Az egoisták között többségben vannak a férfiak. A versengő típusú emberek alkotják a népesség 10-15 %-át, míg az agresszív orientáció aránya gyakorlatilag nulla, bár az ilyen esetek olykor előfordulnak kísérleti szituációkban is. Az agresszió fakadhat bosszúvágyból, felháborodásból is. A kísérleti személyeket például feldühítheti az, ha valaki más sosem kooperál, ha kihasználja a többieket, és a felháborodott alanyok ilyenkor a saját érdekeiket is félretéve igyekeznek neki minél nagyobb kárt okozni a másiknak. Meg kell még itt jegyezni, hogy az agresszív motivációt eleve kevesebbet is vizsgálták, többek között azért, mert vannak, akik nem sorolják a proszelf orientációk közé ezt a beállítódást mondván, hogy nem az ágens javára, hanem kárára tör.

Akárhogy is ítélik meg a különböző motivációkat a társas értékorientáció elméletén belül, a társadalmi cselekvések modelljébe a teljes motivációs gyűrűt érdemes felvenni – még akkor is, ha vannak olyan típusok, amelyek a gyakorlatban kevésbé elterjedtek, kevésbé jellemzik a cselekvőket. Az alábbi ábra azt mutatja, hogy a társadalmi cselekvőknek milyen társas motivációikat tulajdoníthatunk elméletileg.
Bár ezt a nyolcas tagolású, finom felbontást nem tudom majd kihasználni a későbbiekben, az elméleti teljesség miatt meghagyom így a modellt. A játékelméleti fejezetbe "csak annyit" viszek tovább belőle, hogy akkor majd megengedem, sőt, tudatosan alkalmazni fogom azt, hogy a játékosok nemcsak proszelf (egoista vagy versengő) beállítódással játszhassák a játékaikat, de proszociálisak (altruisták vagy kooperatívak) is lehessenek.

Ez a kettőség viszont felvet egy további komoly kérdést. Ha elfogadjuk, megengedjük, hogy a – játékelmélettel leírt – társadalmi helyzetekben különböző társas motivációkkal léphessenek be emberek, akkor ezt a különbözőséget, az irányultságok, a társas preferenciák különbözőségeit meg kell tudnunk ragadni a fogalmi modellünkben is. Ehhez — Franz Dietrich és Christian List javaslatát megfogadva – kettős ontológiát (esetleg többes ontológiát) kell kezelnünk, hogy egyszerre leírhassuk, értelmezhessük a játékosok (és a kívülálló modellezők, megfigyelők) különböző beállítódásait ugyanazon játékszituációra

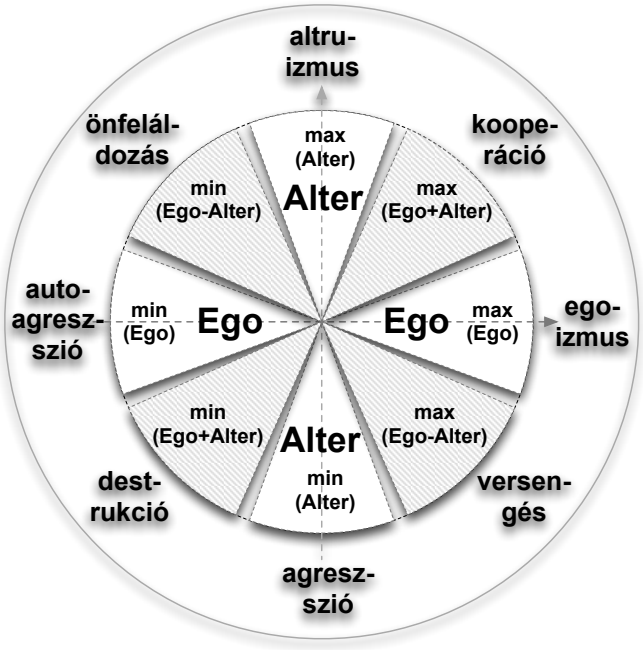


Figure 7.2: a társadalmi értékorientáció típusai

vonatkozóan:

> „ . . . our property-based account of preference formation can be re-expressed as a 'double-ontology' account, in which the agent's 'ontology' of alternatives is distinguished from that of the modeller." [210]

A nagy kérdés itt az, hogy miként akarjuk modellezni a társas preferenciákat. Ha azt állítjuk, hogy a cselekvő, az Ego proszociális értékekkel rendelkezik, akkor ezt hogyan írjuk le? Nem is az az igazi kérdés itt, hogy milyen súlyokat rendeljünk a magára irányuló, illetve másikra irányuló beállítódások erősségének kifejezésére, hanem az, hogy mit fejezünk ki pontosan a proszociális motívumokkal. Az Ego vagy a megfigyelő vagy az Alter véleménye, preferenciarendezése számít-e ilyenkor? A különböző válaszok esetén mást és mást kell majd a proszociális irányultságokat leíró formulákba, képletekbe beírni.

A válaszokat nem most próbálom meg kifejteni, hanem majd akkor, ha már áttekintettem a leggyakoribb társas orientációkról, az egoizmusról, altruizmusról, rosszindulatról szóló legfontosabb elméleteket.

A társas értékorientációval kapcsolatos fogalmakat két szinten vizsgálhatjuk: a tényleges cselekvések szintjén, illetve a cselekvésekkel kapcsolatos motivációk, irányultságok szintjén [698]. Az altruista minősítéhez hozzátartozik egyfelől Ego olyan cselekvése, amely Alter számára előnyös, Egonak költségei (elmaradt haszna?????) volt vele, viszont azt is "elvárjuk", hogy másfelől Ego altruista motiváció miatt tegye meg azt, amit megtesz.

A vélekedés és vágy mindkettő propozícionális attitűd. Van propozícionális tartalmuk, és viszonyt fejeznek ki az egyén és a propozíció között.

———————

valahova betenni: [5]

———————

a vágy mindig propozícióra irányul, nem objektumra, példa erre [698, 350].

amikor azt mondjuk: Kati akar egy pohár vizet. akkor rövidebben fejezzük ki azt, hogy Kati azt akarja, hogy legyen egy pohár vize. Kati azt akarja, hogy igaz legyen az az állítás, hogy Katinak van egy pohár vize.

Még jobban látszik ez a propozicionális viszony, ha a következő mondatot vesszük: Éva azt akarja, hogy Ádám egye meg az almát.

The following sentence is well-formed with it, but not with her: "John wants a mistress, but his wife won't allow it/her." This suggests that the grammatical object of wants is not a feminine singular, such as a mistress; he has a mistress is the obvious candidate.

majmok tudatossága kapcsán 216.o. öntudatosság (self-awareness): önfelismerés ??? (self-recognition):

ultimate vs. instrumental desire

„(1) S wants M solely as a means to satisfying S's desire for E if and only if (a) S wants M, (b) S wants E, and (c) S wants M only because S believes that obtaining M will promote obtaining E." 217.o.

„(2) If S desires M solely as a means to attaining E, then, if S comes to believe that M won't provide E, S will stop desiring M, continue to desire E, and try to find new means for attaining E; what S won't do is stop desiring E, continue to desire M, and try to find new means for attaining M (Batson 1991)." 218.p.

„(3) If S desires M solely as a means to attaining E, then, if S attains E but does not attain M (or more precisely, if S believes that this has happened), S will stop desiring M; however, if S attains M but does not attain E, S will not stop desiring E." 219.o.

az eszközvágyakat könnyebben feladjuk, a célvágyakat nehezebben

2-3 tétel prediktív, következtetni tudunk a jövőbeni viselkedésekre

olyasmi, mint a diszpozíció: itt is tudunk mondani valamit a várható dolgokról

de egyik sem retrodiktív, azaz nem tudunk semmit mondani arról, miért lett olyan, amilyen a múltban 221.o.

3 fő motiváció: hedonizmus, egoizmus, altruizmus

hedonizmus: the only ultimate desires that people have are the desires to obtain pleasure and avoid pain.

egoizmus: the only ultimate goals an individual has are self-directed; people desire their own well-being, and nothing else, as an end in itself.

hedonizmus altípusa az egoizmusnak

altruizmus: have irreducible, ultimate other-directed ends.

az egoizmus minden vágyat magára irányulónak tekint, az altruizmushoz "elegendő" egy másikra irányuló vágy, és már beszélhetünk róla. Ez módszertanilag fontos különbséget takar a két fogalom között, bár mindkét definíció nem egy cselekvési motivációt, hanem inkább egy tartós ás általános beállítódást mint személyiségjegyet határoz meg.

———————-

[5] Az egoista beállítódás magára irányul, azzal nincs elméleti kiterjesztési probléma. A másikra irányuló orientáció esetében azonban felmerül a kérdés, hogy lehet-e, érdemes-e tágítani a 'másik' fogalmának terjedelmét. Az általam kifejtett megközelítésben a 'másik' egy 'ágens'. A cselekvéselmélet számára ez az értelmezés elegendő. Elméletileg lehetne tágítani ezen az értelemzésen, ha megengednénk, hogy a 'másikra irányulás' lehessen egy eszme, az emberiség egésze, a Föld stb. Angolul ezekre lehetne az irányultságokra a 'selfless' terminust alkalmazni, magyarul nem tudok ennek megfelelőjéről.

A társas értékorientációk értelmezése

harmónia, mutualizmus (kölcsönyös előnyök) kontra kooperáció kérdése

ha a kooperációt úgy értelmezzük, hogy abban mind az Egonak, mind az Alternek előnyei származnak (hiszen a kooperáció az egoizmus-altruizmus tengelyek pozitív értéktartományában helyzkedik el), akkor a kooperációt harmóniának kell tartanunk. Ha viszont arra gondolunk, hogy bár a kooperációban mindkét ágens számára előnyös, de Egonak egy ilyen helyzet kevesebb nyereséggel, s ebben az értelemben veszteséggel jár a "tiszta" egoizmushoz képest, akkor látható, hogy motivációs szinten Egonak – részlegesen – altruista beállítódás mentén lehet csak kooperáció mellett dönteni.

Alter

Ego Ego

Alter

T(p/nem-p) p/p nem-p/p nem-p/nem-p

előny, haszon, java (benefit, advantage), elmaradt kár

hátrány, kár (felmerült kár, elmaradt haszon, indokolt költség), ártalom, veszteség (disadvantage, malefic, cost)

kooperáció = segítés a másiknak (Egonak költség/hátrány, Alternek előny) dezertálás= nem segítés a másiknak (Egonak nincs költség/hátrány, Alternek nincs előny)

„When the interest of others diverges from one's own, acting 100 altruistically involves acting to promote the interest of others at cost to one's own interest, whereas acting selfishly involves acting to promote one?s own interest at cost to that of others." [521]

önző= saját érdekét elégíti ki, altruista=másik érdekét elégíti ki önérdekkövető vs. másérdekkövető

someone motivated by self-hatred or a lack of self-respect is not altruistic (although she promote the interest of others at cost to his own) [521]

1. condition: interests diverge (divergence condition) 2. condition: promote own interest at cost to others's interest or (cost condition) 3. condition: someone acts in a morally bad way (evaluation condition)

a hagyományos interperszonális fogalmak mellett létezik *intraperszonális egoizmus* és *intraperszonális altruizmus* a diakróm időben

7.3.1 Egoizmus

A társas értékorientációk áttekintését az egoizmussal (individualizmussal) kell kezdeni. Egyszerűen azért, mert mind az evolúcióbiológia, mind a társadalommagyarázat számára az egoizmus jelenti a természetes kiindulópontot – még akkor is, ha ez egy könnyű kiindulópont.

Thomas Hobbes, Adam Smith és még sokan mások úgy gondolták, hogy az emberiség alapvetően és eredendően önérdekkövető. Hobbes szerint az ember alapállapotát a "mindenki harca mindenki ellen" kifejezéssel írhatjuk le, Adam Smith "láthatatlan kéz által irányított gazdasága" is azért működik, mert az egyéni cselekvők valamennyien saját önérdeküket követik. A gaz-

dasági ember (homo oeconomicus) haszonmaximalizáló és kárminimalizáló lény „Economic man is characterized by *self-interested goals* and *rational choice of means*." [358]

az egoizmus lehetségessége (The Possibility of Selfishness) [549] how is selfishness is possible?

„*Egoism* depends on what I shall call a concept of the self as a *particular*; it involves the capacity to discriminate among particulars that might be exactly similar, to identify one of those particulars as oneself, and to base wants, evaluations, and policies upon this discrimination." 26

„*Impersonal morality* depends on what I call a concept of the self as (an instance of) a *kind of thing*; here K_1 identifies himself for evaluative purposes in terms of sets of features which, in principle, K_2 or other things could also have." 26

„ . . . only an entity with a concept of the self as a particular can have selfish policies. Which raises the question, "What kind of entity can have a concept of the self as a particular?" I want to suggest that whereas any entity that can be said to follow policies can act on the basis of the kinds it belongs to, only to a conscious entity can it matter which particular of a given kind it is." 26

„The separation of conscious states into those that are one's own and those that are not provides a third way (and the egoist's way) of distinguishing among particulars, the first being to notice differences in qualities or features and the second to apprehend different spatial/temporal predicates." 27

„By a Kantian policy I mean any policy or directive that is universalizable in the sense that the entity that follows the policy takes its indexical "I" to be replaceable by general terms." 27.

a tudatosság nélküli lények perfekt kantiánusok a robotok is Kant szerint a tökéletesen racionális lények, mint például Isten, kizárólag univerzális törvények szellemében cselekednének és nem lenne szükségük morálra. De Oldenquist szerint univerzális törvények szerint lehet (kell) cselekedni abban az esetben is, amikor a cselekvőnek hiányzik a szelf mint partikularitás fogalma.

mind az altruizmus, mind az egoizmus fogalma megköveteli a szelf partikuláris fogalmát. Mindegyik esetben arra van szükség, hogy az individuum képes legyen eltérően értékelni magának és másoknak esetleg teljesen megegyező állapotát. Ilyen megkülönböztetésre nincs mód kizárólag az egyedek – külső szemlélők által is felismerhető – tulajdonságaira támaszkodva. Ehhez a szelf tudata, vagyis öntudat szükséges.

Az ember, amikor kanti értelemben univerzálisan cselekszik, akkor nem veszi figyelembe saját szelfjének létezését 28.o.

az önző gén elmélet nem a partikuláris, invididuális gén önzőségén, hanem a géntípus önzőségén alapul.

szuperindividuum részrehajló szándékközösség, amely a típus bármelyik tagjának és a közösség egészének, ezáltal a típusnak az előnyére törekszik

Orthmayr Imre,

Elster: az egoizmus fogalmilag előbb van, mint az altruizmus

> „Bizonyos értelemben az önérdek mégis alapvetőbb, mint az altruizmus. A természeti állapot – jóllehet gondolatkísérlet – logikailag koherens helyzet. Olyan világot viszont nem tudunk logikailag koherensen elképzelni, amelyben mindenkinek kizárólag altruista indítékai vannak. ... Ha senkinek sem lennének elsőfokú, önző örömei, senkinek sem lehetnének magasabb rendű, altruista indítékai sem. ...

> Ez tisztán logikai kérdés. Ahhoz, hogy egyesek altruisták lehessenek, másoknak legalább időnként önzőnek kell lenniük, de mindenki mindig lehetne önző. ...De ebből – sem általában, se bármely adott esetben – nem vonhatjuk le azt a következtetést, hogy az önzés a legelterjedtebb motiváció." [232, 61]

harc, szembenállás, konfliktus, érdekellentét lásd: [295] bevezetőjének 2. lábjegyzetében (33.o.)

Hamilton négyes felosztása

mutuality selfish spite altruism

az érdek ésszerű önzés

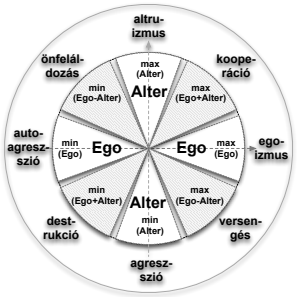
Albert O. Hirschmann: érdekek és szenvedélyek eszmetörténetének elemzése [356]

a szenvedélyeket az érdekek által lehet megzabolázni, az érdek kiszámítható:

Jhering egoizmus és jog viszonya kapcsán: [273]

Frank W. Marlowe és szerzőtársai [473]

7.3.2 Altruizmus



fals dichotomy? [424]

kár másik modell a (Gulyás-disszertációból

Gary Charness és Matthew Rabin [142]

Cox és társai [176]

mi befolyásolja a másikra irányuló döntést: a státus és a reciprocitás

Thomas Nagel The Possibility of Altruism [527]

motivációról és Nagel könyvéről [126]

az altruizmussal kapcsolatos vizsgálatok, kísérletek áttekintése: [423]

Az emberiség különböző (tudományos, vallási, művészeti, erkölcsi) diskurzusaiban mindig kiemelt helyet kapott az önzetlenség, az altruizmus kérdése.

Garret Hardin: diszkriminatív altruizmus típusai [338]

egyre növekvő számosság, egyre csökkenő mértékű altruizmus

- individuum (szerelem, barátság)
- család
- bajtársiság, kronizmus
- tribalizmus, törzsiség
- hazafiság, patriotizmus

áttekintés az altruizmusról: [506] [574]

> „1. Segítségnyújtás veszély esetén (pl. baleset, ragadozók, fajon belüli agresszió)
> 2. Élelem megosztása
> 3. A beteg, a sebesült, a nagyon fiatal és az öreg megsegítése
> 4. A szerszámok megosztása
> 5. A tudás átadása" [753]

többszemélyes altruizmus (multiparty altruism) az altruista rokonával szembeni viszonzás

jóvátételi altruizmus [753]

forms of other-directed behavior: megosztás (sharing), a segítségnyújtás (helping), az adás (giving), a kooperáció (cooperation)

> „I define altruism as behavior intended to benefit another, even when doing so may risk or entail some sacrifice to the welfare of the actor. There are several critical points in this definition. First, altruis must entail action. It can not merely be good intentions or well-meaning thoughts. Second, the goal of the act must be furthering the welfare of another. If another's welfare is treated as an unintended or a secondary consequence of behavior designed primarily to further one's own welfare, the act is not altruistic. Third, intentions count more than consequences. If I try to do something nice for you, and it ends up badly or with long-term negative consequences for you, this does not diminish the altruism of my initial action. Fourth, the act must carry some possibility of diminution to my welfare. An act that improves both my own welfare and that of another person would not be altruistic but would instead fall into the category of collective welfare." [514]

Hamilton [322], [323]
6

Gulyás disszertációjából átvenni irodalmat méltányosság
M. Dufwenberg és G. Kirchsteiger 2004 [220]
Az altruizmus fogalmának elemzése: Khalil [397]
1) egoista altruizmus az altruista cselekvést a várható jövőbeni haszon magyarázza (például a direkt reciprocitás esetében) párhuzamba állítható az evolócióbiológia triversi reciprock altruizmus fogalmával [753]
2) egocentrikus altruizmus ha az Ego abból eredő hasznosságérzete, hogy élvezi Alter hasznát/örömét valami jószág elfogyasztásakor, nagyobb, mint az ahasznosság, mint amit Ego az adott jószág saját fogyasztása révén érhetne el
vicarious pleasure – más öröméből vkinek jutó rész (vicarious - helyettesítő, mögöttes) de: vicarious liability – mögöttes felelősség
3) altercentrikus altruizmus
a méltányosság két egyenlőségcentrikus alapmodellje:
Ernst Fehr és Klaus M. Schmidt 1999 [244]
illetve
G.E. Bolton és A. Ockenfels modellje theory of equity, reciprocity and competition [102]
a méltányosság szándékcentrikus modellje:
M. Rabin [587]
illetve
Levine altruizmus-rosszindulat modellje [438]
az altruizmus megjelenése a különböző kultúrákban Gintis: [293]
Ulmann-Margalit, Considerateness [764] Family Fairness, [763]
Herbert SImon evolúciós magyarázata az altruizmus kialakulására [682]
kiváló áttekintést ad magyar nyelven Janky Béla [376]
az altruizmus család, vásott kölyök modell Gary Becker altruizmusmodellje
Altruism, Egoism, and Genetic Fitness: Economics and Sociobiology [59]

[6] Az önzetlenség kialakulásáról magyar nyelven ad kiváló áttekintést Bereczkei Tamás [67], a könyvről én is írtam egy recenziót [723].

későbbi Becker cikk Altruism in the Family and Selfishness in the Market Place [60]
a vásott könyök modell továbbfejlesztése [69]
a csoportszelekciós magyarázatok problémáiról: [633]
barátság kategóriája [272]
etikai egoizmus lehetősége: [275]
H.Simon [680]
Why Can't Altruism Be Selfish? [515]
Marosán cikk [474], a nők gyűjtögető életmódra berendezkedése miatti munkamegosztás [428]

Az altruizmus modellje

másodrendű fogalom, mert a hasznosságérzetbe be kell venni a másik hasznosságérzetét de ha kettős ontológiát használunk (lásd még méltányosságlemélet), akkor nem minden modellben lesz másodrendű a fogalom ha csak a kifizetések kerülnek be a paraméterek közé, akkor még nem lesz beágyazódás, de ha már Ego hasznosságérzetébe beépül Alter hasznosságérzete, akkor felmegyünk másodrendbe

A preferencialogika úgy határozta meg a preferenciarendezés fogalmát, hogy csak a rendezés tárgyaira fókuszált (a lehetséges világállapotokra, *m*-re és *n*-ra), így nem vette figyelembe a rendezést végrehajtó személyt (PREF(m,n)). Amikor a társadalmi preferenciákat akarjuk megragadni, akkor szükségünk van arra, hogy a cselekvőt, vagyis a döntést hozó, a preferenciarendezést végrehajtó személyt is beemeljük a preferencia fogalmának meghatározásába. Mondhatjuk azt is, hogy perszonifikálni kell a fogalmat. Első lépésben fel kell venni a fogalom összetevői közé a preferenciarendezést végrehajtó személyt, valamelyiket az ágensek közül, Egot vagy Altert.

$\text{PREF}(Ego, m, n)$ – *Ego preferálja m-t n-nel szemben*

Ez azonban nem elégséges a társadalmi preferenciák fogalmának leírásához, mert így még csak annyit tudunk kifejezni, hogy ki az értékelést végző személy. Azt is jelezni kell tudnunk, hogy

x= BENEFIT(Ego) & BENEFIT(Alter)

$\text{PREF}(Ego, \text{BENEFIT}(Agent), \text{BENEFIT}(AlterAgent))$ – *Ego preferálja m-t n-nal szemben*

David K. Levine empirikus kísérletekben vizsgálta az altruizmus jelenlétét Modeling Altruism and Spitefulness in Experiments [438]

R. Sethi és E. Somanathan a preferenciák változásáról (fejlődéséről) és a reciprocitás kapcsolatáról [674] a preferenciák evolúciója: [632]

Becker Rotten kid theorem [376]

Krebs - Miller: Altruism and aggression [422]

[450]

7.3.3 Kooperáció

Competition among cooperators: Altruism and reciprocity [185]

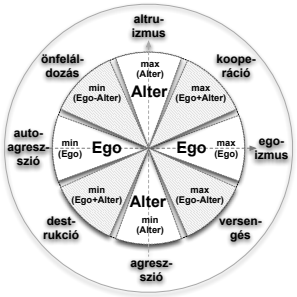
a kooperáció értékelésekor kétféle téves beállítódás tapasztalható mivel az altuizmus és az egoizmus "között" helyezkedik el félúton, ezért e két irányból a kooperációra tekintve könnyen az ellentétes pólushoz tartozónak lehet minősíteni azt. Az altruista szemszögéből nézve valamilyen mértékben önérdeket is

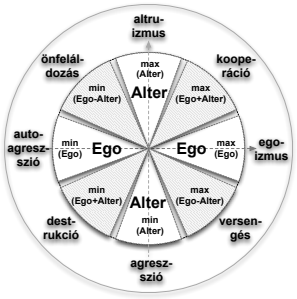
követ a kooperáló fél, ezért csak erre az összetevőre hivatkozva egoistának lehet tartani őt, és fordítva, amilyen mértékben valaki elmozdul a tiszta egoizmus felől az altruizmus irányába, olyan mértékben már altuistának is lehet minősíteni őt (elfeledkezve a "még megmaradt" egoista hányadról) ez más, mint amikor az altruizmust egoizmusnak értelmezi valaki mondván, hogy az altruista is azt csinálja, ami neki jó, tehát egoista. itt hivatkozni kell arra a részre, ahol ezt kritizálom

7.3.4 Versengés

7.3.5 Agresszió

> „A szellem kiszabadul a palackból, és azt mondja a parasztnak, hogy egyet kívánhat, de akármit is kíván, a szomszédja kétszer annyit fog kapni. 'Szúrd ki az egyik szemem!' – jön a válasz."





A tudományos vizsgálatok sokáig ugyanúgy elhanyagolták a rosszindulatú cselekvések jelenségét, mint tették ezt az altruista cselekvések esetében [689]. Pedig más (például vallási, erkölcsi, művészeti) diskurzusokban mindig is jelen volt a rosszindulat, az agresszió, a gonoszság kérdésköre (akárcsak az altruitmusé). A tudományos magyarázatok az egoizmus kizárólagosságára támaszkodtak, és sok idő kellett ahhoz, hogy a kutatók szép lassan beemeljék az altruista motívumokat is a társadalmi cselekvést leíró modelljeikbe. Még többet kellett várni a rosszindulat jelenségének tudományos vizsgálatára.

W.D. Hamilton tágította ki először az elemzési fókuszt, amikor vizsgálandónak tartotta a rosszindulat (spite) jelenségét az állatvilágban [324]. Hamilton rövid cikke elején még azt a kérdést tette fel, vajon lehetséges-e, hogy egy állat kárt okozzon magának úgy, hogy eközben egy másiknak még nagyobb kára keletkezzék, és ha ez előfordul, akkor hogyan lehetne modellezni. Hamilton a kérdésbe foglalt állítást nevezte el rosszindulatnak. Később azonban, írása egy másik pontján már kicsit másként határozta meg ugyanezt a fogalmat, amikor azt állította, hogy "rosszindulatú az a viselkedés, amely kárt okoz másokat, de nem jár előnnyel az egyén számára". A felállított modell szempontjából végülis az utóbbi értelmezés volt a fontos, a modell eredménye pedig az lett, hogy az átlagos

rokonsági szint alatti egyedekkel szemben evolúciós szempontból olykor előnyös lehet a rosszindulatú magatartás. A rosszindulattal kapcsolatos jelenségeket figyeltek meg az állatvilágban, bár nem túl sokat, és mindig csak fajok közti kapcsolatokban. Hamilton maga is azt állítja saját modelljéről, hogy az alapján a rosszindulatú magatartás megjelenését csak ritkán lehet elvárni, mert egyrészt minden rosszindulatú tevékenység költségekkel jár, másrészt nehezen lehet felismerni az átlagos rokonsági szinthez viszonyított "távolságot", harmadrészt a modell olyan, hogy relatíve kis méretű populációk esetében "teszi lehetővé" a rosszindulat megjelenését és elterjedését, ami viszont könnyen a populáció gyors kihalásához is vezethet.

agresszióról:

Punishment and spite, the dark side of cooperation [380]

[590]

[484]

> „We define human aggression as any behavior directed toward another individual that is carried out with the proximate (immediate) intent to cause harm. In addition, the perpetrator must believe that the behavior will harm the target and the target is motivated to avoid the behavior." [133]

a fogorvos fájdalmat okoz a gyökérkezelés során, de nem ez a szándéka, ezért nem agresszió a szadomazo kapcsolatban sincs agresszió, mert az "áldozat" nem akarja elkerülni a fájdalmat

agresszió, autoagresszió (önagresszió) S.A. West és A. Gardner [814]

The 'spiteful' origins of human cooperation [473]

animal conflict limited war? [477]

egér, héja, bika/szájhős, megtorló, szondázva megtorló stratégiák

Skyrms: an individual hurts himself by defecting, but hurts his partner more. This is a case of spiteful behavior. Hamilton and Price also showed how negative correlation is the key to the evolution of spite. Both spite and altruism on their face appear to violate the rational- choice paradigm. Both have an evolutionary explanation in terms of correlated interactions. Yet spite rarely receives the attention given to altruism. a search on Google scholar for "evolution of altruism" gets 1,570 hits, while one for "evolution of spite" gets 32. [689]

?minden olyan szándékos cselekvést, amelynek indítéka, hogy ? nyílt vagy szimbolikus formában ? valakinek vagy valaminek kárt, sérelmet vagy fájdalmat okozzon.? (Ranschburg, 1977, 90.o.)

Alfred Adler: hatalmi törekvések mások fölé kerekedés és a mások fölötti uralkodás vágyától egészen mások leigázásáig húzódhat.

Agresszivitás jelentkezhet kín, fájdalomokozás, gyakran szadizmus formájában is, mások megalázásában, megfenyegetésében, tényleges vagy verbális bántalmazásában.

rokon fogalmak: - Agresszió: viselkedés, szándékos (intenciója van)

- Harag: érzelmi állapot, intenzitásban változik Hiányzik a specifikus cél Nem szükséges, hogy az agresszió megjelenjen

- Hosztilitás: attitűd, kognitív konstruktum Az esemény interpretációját és az inger értékelését tartalmazza Verbálisan és fizikailag is kifejezett

E. Fromm felosztása: 1) Benignus (jóindulatú) - Önvédelem - Biológiailag megalapozott - önfenntartás

2) Malignus (rosszindulatú) kín és kár okozás bosszú, szadizmus személyiségjegy is (pszichopátia)

FESHBACH (1964), Bandura (1973), Green (2001) Instrumentális - célért elkövetett ?hidegvérrel? elkövetett - Tanult

- Indulati - emocionális (DÜH) - nem tervezett - automatikus folyamat - ártás

Bushman és Anderson, 2001 Nincs értelme ennek a megkülönböztetésnek, mert átfedőek, mindkettő tartalmaz kontrollált és automatikus folyamatokat is

gyáva nyúl játék az állatvilágban, héja-galamb: Dawkins önző gén könyvéből

Spite was defined as a form of behaviour that harms conspecifics without a direct adaptive benefit to the self. More specifically, spite is costly (has a negative effect) for the actor, but this cost is considerably smaller that the harm caused to the non-kin individual targeted. [620]

spiteful behaviour, which is costly for the both actor and recipient [383]

“üss vagy fuss” angol kifejezéssel “fight or flight”

verseny vetélkedés viaskodás viadal harc viszály vetekedés

A játékelméleti, közgazdasági gondolkodás világába

David K. Levine empirikus kísérletekben vizsgálta az altruizmus és rosszindulat jelenlétét Modeling Altruism and Spitefulness in Experiments [438]

gonoszság, a másik kárára irányuló motiváció

jóindulat vagy rosszindulat (benevolence vs. malevolence) [358]

nem túl gyakori, de azért létezik kétféle értelmezés van a fogalomra

Spite: Hamilton’s unproven theory [263]

Wilson könyvében ír a spite és az agresszió jelenségéről [826]

Hamiltonian spite (pozitív) Wilsonian spite (negatív)

[282]

Krebs - Miller: Altruism and aggression [422]

Conflict and aggression [586]

7.3.6 Autoagresszió

önmaga ellen irányuló rosszindulat

mazochizmus

vallási áldozat, áldozat, emberáldozat, Dugonocs Titusz,

Az autoagresszióra hozható legegyértelműbb példa az öngyilkosság jelensége. És nem igaz, hogy az öngyilkosság kizárólag emberi sajátosság volna. Bár nem túl gyakori, de előfordul az állatvilágon belül [384, 113]. A kutatók adaptív öngyilkosságnak nevezik, és a létezését úgy magyarázzák, hogy az egyed a saját génjének sikeres továbbörökítése érdekében pusztítja el magát. Előfordul, hogy a borsótetű (Acyrthosiphon pisum) testébe darazsak beszúrják a tojásaikat, hogy az utódaik parazitaként nevelkedjenek fel a borsótetű testéből táplálkozva. Ilyen esetekben előpordul, hogy az így megfertőzött borsótetű ledobja magát a levélről, ahol addig a társaival együtt élt. A földön biztosan áldozatává válik bogaraknak, de önfeláldozásával saját génjeit hordozó rokonai számára előnyöket biztosít, mert a parazita darázslárvák nem vándorolhatnak tovább a társaira. Az ausztráliai vöröshátú pók (Latrodectus hasseltii) nősténye is megeszi a hím párját, de ezt

a kutatók nem a nőstény agresszivitásával, hanem a hím önfeláldozásával magyarázzák [384, 114]. Az önfeláldozásnak ugyanis adaptív előnye van, mert a magát halálba kopuláló hím saját génjeit továbbörökíti, míg a halálos párzástól megmenekülő hímtársai erre nem képesek – mindaddig, amíg nem vállalkoznak saját halálos kalandjukra.

7.3.7 Önfeláldozás

Ez az igazi altruizmus, amikor a cselekvő magának kárt, költséget okoz, a másiknak előnyöket

Nem érzem feladatomnak, hogy kitérjek itt az emberi öngyilkosság kérdésére, de annyit azért érdemes megemlítenem, hogy Durkheim az öngyilkosság típusai között megemlíti az *altuista öngyilkosság* fogalmát, amit úgy jellemez, hogy erős társadalmi integráció létezésekor a közösség elvárása olyan erős az egyénnel szemben, hogy utóbbi maga megteszi az önfeláldozó lépést[226].

7.3.8 Destrukció

szadomazochizmus

Othello

a mottó is ezt ragadja meg

dögöljön meg a szomszéd tehene is

Ambrus-Lakatos Lóránd, Meszerics Tamás az ultimátumjátékban az egyik alapmotivációnak tekinthetjük a méltánytalannak tekintett ajánlattal szembeni megtorlást, amely a kölcsönösen káros visszautasítás formájában érvényesül. Az igazságosság normáját megsértővel szembeni büntetés igazságosnak tekinthető. [10]

7.4 Dinamikus kapcsolat: reciprocitás

a reciprocitás altruista/kooperatív és egoista/versengő/destruktív beállítódások időben váltakozó egymásutánisága

pozitív és negatív reciprocitás

7.5 Méltányosságelméletek

A méltányosságelméletekről magyar nyelvű áttekintést adnak Gulyás Attila cikkei [312], [313]

mi a különbség a méltányosság és az altruizmus között?

az altruizmus modelljében az Ego hasznosságérzetében változóként szerepel az Alter hasznosságérzete is, míg a méltányosság modelljében az Ego hasznosságérzetében Alter kifizetése szerepel csak [312]

kettős ontológiát kell használni [210]

erős reciprocitás

Journal of Personality and Social Psychology [155]

Manipulating Allocation Justice: How Framing Effects can Increase the Prevalence of the Talmudic Division Principle "Shnaim Ohazin" [78]

Janky Béla a méltányossági normák két nagy csoportját különíti el [376]. egyrészt cselekvés- vagy hozzájárulás-alapú, másrészt elosztás-alapú

Sugden szerint nem lehet az altruizmus vagy a normák tételezésével megmagyarázni a karitatív tevékenység bizonyos formáit a cselekvésre alapoz [714]

M. Rabin méltányosságmodellje a szándékra alapoz [587]

ide csatolni Fiske elméletét

igazságosság:

- formális
- szubsztantív
- retributív
- korrektív
- kommutatív
- disztributív

[35] 456.

7.6 Szolidaritás

Lindenberg három cikke

Janky könyve

7.7 Dominancia, státus

a rosszindulat, versengés egyik fontos motivációs példája, esete

”As the desire for status notoriously pervades all human activities, it is quite remarkable that economists up to quite recently have so signally failed to incorporate this phenomenon into their models.” [358]

Becker tipizálásában szerepel az irigység, gyűlölet, de a cikk húzós [58]

7.8 A társadalmi kapcsolat típusai

„Szerintem, amint a háború és a viszály két különböző szó, úgy két külön lényeg is, két különböző tényre vonatkozik. Ez az úgynevezett »saját«-ra, vagy »rokon«-ra, az meg az »idegen«-re és »elütő«-re. Ha saját vérünkkel tusakodunk, e harc neve: viszály, ha idegenekkel: háború.” (Platón: Az állam)

Az előző fejezetben a beállítódásokkal kapcsolatos fogalmakat elemeztem. Különösen sokat foglalkoztam az ágensre irányultság különböző formáival. A fejezet bevezetőjében két dolgot is jeleztem. Egyfelől azt, hogy ezek a fogalmak mind a társadalmi kapcsolat fogalma alá sorolhatók be, másfelől azt, hogy az ágensre irányultsággal kapcsolatos fogalmak szűkebb terjedelműek a társadalmi kapcsolat fogalmához képest. Természetesen minden generikusan alárendelt fogalom szűkebb terjedelmű a fölöttes kategóriához

képest, de itt nem erre a trivialitásra utalok. Az ágensre irányultság mint beállítódás figyelembe vétele fontos a társadalmi kapcsolat tipizálása szempontjából, de a szükség van valami másra is. Amikor Max Weber meghatározza a társadalmi kapcsolat típusait, akkor két dimenzió mentén különíti el az új kategóriákat. Egyrészt használja azt a dimenziót, amit én az ágensre irányultság fogalmával ragadtam meg, másrészt hivatkozik a társadalmi cselekvések típusaira (racionális, illetve aracionális cselekvések két nagy osztályára), amelyeket korábban én is bemutattam. Mivel nagyon fontos kategóriákat vezet be ezen a ponton, érdemes megnéznünk, hogyan teszi mindezt.

A társadalmi kapcsolat meghatározása után Max Weber a fogalom három altípusát különíti el egymástól. Először a **harc** (konfliktus) fogalmát definiálja:

> „*Harcnak* nevezünk egy társadalmi kapcsolatot, amennyiben a cselekvőt az a szándék vezeti, hogy saját akaratát a másik vagy a többi féllel szemben keresztülvigye." [810, 64],

majd kicsivel később pedig megadja a *közösség* és a *társulás* meghatározását:

> „ *Közösségnek* nevezzük a társadalmi kapcsolatot, ha a társadalmi cselekvést ... a résztvevők szubjektíve átérzett (érzelmi-indulati vagy tradicionális) összetartozásán alapuló beállítódás jellemzi."
> „ ... *társulásnak* nevezzük a társadalmi kapcsolatot, ha a társadalmi cselekvést racionálisan (érték- vagy célracionálisan) motivált érdekkiegyenlítődésen vagy ugyanígy motivált érdekkapcsolódáson alapuló beállítottság jellemzi." [810, 66].

Ez a weberi tipizálás bár nagyon fontos, de terminológiai szempontból közel sem teljes, érdemes tehát kiegészíteni.[7] Bár Weber erre tudatosan nem igazán reflektál, mégis mondhatjuk azt, hogy a weberi kategóriarendszeren belül a fenti három fogalom egy szembenállást fejez ki. A harc fogalma van az egyik oldalon, míg a közösség és társulás fogalmai a másikon. A szembenállás lényege az, hogy a harcot az ellenállással szembeni akaratérvényesítés, tehát az "kizárólagos" önérdekérvényesítés, a közösséget és a társulást pedig az összetartozás, az érdekkiegyenlítődés, érdekkapcsolódás jellemzi. Ezt a dichotómiát az ágensre irányultság fogalmának két pólusával jól megragadhatjuk, bár a terminológiai konzisztenciához szükségünk van egy-két új fogalom bevezetésére. Mondhatjuk ugyanis azt, hogy a harc a magára irányuló, míg a közösség és kapcsolat a másikra irányuló beállítódással kapcsolható össze, de ha ezt tesszük, akkor az is nyilvánvalóvá válik, hogy nem ugyanazon a fogalmi szinten vannak ezek a kategóriák. A weberi fogalomkészletből hiányzik a harc fogalmi párja a másra irányuló beállítódás területén, illetve alá kell bontani a harc kategóriáját is olyan logika mentén, ahogyan a közösség és kapcsolat kategóriái elkülönülnek egymástól.

A három weberi fogalom az alábbi módon rendezhető táblázatba (kérdőjelekkel jelzem a hiányzó kategóriákat vagy rendezési dimenziókat):
Szükség van tehát egy új kategóriára a közösség és társulás fölöttes fogalmaként, illetve szükség van két új kategóriára a harc fogalma alatt. Kicsit egyszerűbb az első feladat megoldása. Ha a magára irányuló beállítódás főfogalma a harc, akkor ennek párja (ellentéte) az **együttműködés** (kooperáció) lehet a másikra irányuló beállítódás oldalán.

[7] Ezzel a kiegészítéssel korábban másutt már megpróbálkoztam [724], de itt érdemesnek tartottam újra neki futni a problémának. Egyrészt azért, mert itt részletesebben és alaposabban meg tudom indokolni a mondanivalómat, másrészt most fogalmilag általánosabb módon oldom meg ugyanazt a feladatot.

	magára irányuló	másikra irányuló
?	harc	?
?	?	társulás
?	?	közösség

Kicsivel nehezebb a harc fogalmának alábontása – egyszerűen azért, mert a magyarban nem áll rendelkezésre megfelelő terminológia a feladat megoldásához. Tartalmilag, szemantikailag az **ellenfél** és az **ellenség** kategóriái megfelelőek lennének, de szintaktikailag, konnotációjukat tekintve nem illeszkednek jól a kiegészítendő fogalomkészlet elemeihez. Igaz, a harc és a közösség és társulás fogalmi közt is érezhető már ez a fajta diszkrepancia, ugyanis míg az utóbbi kettő valamiféle csoportra utal (nyilván azokra a csoportokra, amelyek tagjai a másikra irányuló beállítódás "kedvezményezettjei"), addig a harc inkább a magára irányuló beállítódás következményeként kialakuló állapotot írja le. Ugyanez a baj az ellenfél és az ellenség kategóriáival is. Ha viszont túllendülünk ezen a problémán, akkor már nyugodtan párba állíthatjuk az ellenfél kategóriáját a társulás fogalmával, míg az ellenségét a közösségével. Az összekapcsolási igényt pedig úgy védhetjük meg, ha megnézzük, hogyan tett különbséget a közösség és a társulás fogalmai között.

Orthmayr Imre, konfliktus [556]

ellenség és ellenfél, barát vagy ellenség fogalmáról George Schwab [647] közellenség, a köz, közössség ellensége (enemy), illetve személyes ellenség (foe) fogalmának szembeállítása e két fogalom történeti jelentésváltozásának bemutatása fontos mozzanat a semleges szereplő megjelenése a szembenálló felek között

fel kell venni a harmadik pólust, a semlegességet ez később fontossá válhat, a triád elemzésben: amikor a semleges fél békítő, közvetítő, esetleg ügydöntő szerepet tölt be vagy amikor felhatalmazottsági konfliktus van

---

??????? reifikációs integráció ?????????

Sok segítséget jelent, hogy Weber maga zárójelben jelezte, hogy a közösséghez hozzá lehet kapcsolni az érzelmi és a tradicionális (habituális) beállítódást, a társuláshoz pedig a kétfajta racionalitást. Ha a közösség és társulás fogalmai között a mérlegelő, illetve nemmérlegelő beállítódás szerint lehet tehát különbséget tenni, akkor ugyanezt a felosztást használhatjuk az ellenfél és ellenség fogalmai között is. A mérlegelő vagy nem-mérlegelő beállítódás helyett a racionális vagy aracionális beállítódás kategóriáit használva a fenti hiányos táblázatot az alábbi módon tehetjük teljessé:

	magára irányuló	másikra irányuló
	harc	együttműködés
racionális	ellenfél	társulás
aracionális	ellenség	közösség

A weberi felosztás explikációja (illetve kiegészítése) után tovább lehet lépni a társadalmi kapcsolatok tipizálásában.

nemzetközi kapcsolatok területéről származó felosztás [396] viszály (dischord) harmónia (harmony) kooperáció (cooperation)

Aumann, War and Peace [40]

	magára irányuló	másikra irányuló	egyikre sem irányuló
	harc	együttműködés	semlegesség
racionális	ellenfél	társulás	kívülálló
aracionális	ellenség	közösség	kívülálló

Rapport a konfliktus három típusa [592]

1) harc, háború, verekedés

2) játék, játszma, önérdekkövetés

3) vita (a meggyőződésekről) A rational debate is one which is truly aimed at modifying the outlook of the Other.

„... in a fight the urge is to eliminate the opponent; in a game the problem is to outwit the opponent; in the debate, the goal is to convince the opponent."

7.9 Neurológiai folyamatok

Alan Sanfey, Science, 2003.06.13. 1755. agyi folyamatok fMRI-vizsgálata ultimátumjáték közben

in: Neuron, 2003/4 Tania Singer, fMRI

Fehr et al. fMRI vizsgálat IPD játék után the opportunity to punish seems too activate similar kinds of reward-related pleasure circuits as does eating sweets.

Chapter 8

Kényszeren alapuló társadalmi kapcsolat

– Adj király katonát!
– Nem adok.
– Ha nem adsz, szakítok.
– Szakíts, ha bírsz!

harc, szembenállás, konfliktus, érdekellentét lásd: [295] bevezetőjének 2. lábjegyzetében (33.o.)

8.1 Érdek

A harc logikailag vagy biológiailag (?) megelőzi a kooperációt ezért a hatalmat előbb kell tárgyalni

a céltételezés, preferencia függvénye

nem egyenlő az önzésel, inkább a beállítódástól függ

8.2 Hatalom

olyan harci helyzetek, amelyben kényszeralkalmazás történik, történhet.

WALTER BENJAMIN: Az erőszak kritikájáról (ford. Bence György). in: UŐ: Angelus Novus. Budapest: Magyar Helikon, 1980, 25-56.

Hirshleifer '[358]:

konfliktuskezelés, a konfliktusok okai preferenciák, lehetőségek, percepciók "potenciális megegyezési terület"

a háború művészete, hatalommérnökség, a harci konfliktus feloldásának, megnyerésének technikái

a csatatéren a nyertes mindent visz (increasing-return), de ez nem terjeszthető ki a legyőzött ország teljes területére (decreasing-return). Előbbi miatt lehet kialakítani az

állami erőszak-monopóliumot, utóbbi miatt érthető meg, miért nem alakul ki a világállam, vagy miért esnek szét a nagy birodalmak

mérethozadék vagy skálahozadék

növekvő mérethozadék (increasing-return) csökkenő mérethozadék (decreasing-return)

8.2.1 Hatalom az állatok közt

biológiai minták, példák

alfahím, dominancia, csípkedési sorrend

erőszak mint hatalomforrás

lehet tudás?

megküzdenek a szűkös erőforrásokért, élelemért, területért, szállásért, szexuális partnerért.

8.2.2 Hatalom emberek közt

A hatalom weberi fogalma

ütköző akaratok, konfliktus, nyílt akaratkonfliktus

valamelyik akarat csak akkor teljesíthető, ha a szembenálló másik akarat nem teljesül

Kiindulásként vegyük Max Weber meghatározását:

> „Minden olyan esetben hatalomról beszélünk, ha egy társadalmi kapcsolaton belül van rá esély, hogy valaki saját akaratát az ellenszegülés ellenére is keresztülvigye, függetlenül attól, hogy min alapul ez az esély." [810, 77]

A hatalom fogalmának explikációja

Korábban bemuttam, hogy milyen interdependens társadalmi kapcsolatba kerülhet két cselekvő (↬277.o.). Ezek közül voltak konszenzusos, voltak konfliktusos helyzetek. Definiáltam a harc (konfliktus) fogalmát is (↬312.o.). Az akaratkonfliktus lényege, hogy az egymásnak feszülő (egymásnak ellentmondó) akaratok csak egymás kárára valósulhatnak meg. A kérdés tehát az, hogy miként "győzhet" az egyik akarat a másik "fölött", vagy fordítva: miért nem érvényesül valaki saját akarata a másik ágens akarata ellenében. Ezen a ponton nagyon körültekintőnek kell lennünk. A pontos ítéletkezéshez arra van szükség, hogy figyelembe vegyük azt, hogy az éppen vizsgált helyzetben melyik ágens cselekvésével kapcsolatos akaratkonfliktusról van szó. E szempont figyelembe vételével kétféle akarattípust különböztethetünk meg. Beszélhetünk saját cselekvésre irányuló 'cselekedni akarás'-ról, illetve a másik cselekvésére irányuló 'befolyásolni akarás'-ról.

Az akaratkonfliktus ténye még csak az ellentétes akaratok létezését fejezi ki, semmi többet. A nagy kérdés az, hogy milyen módon oldódik fel ez a konfliktus. A konfliktus kétféle lehetséges kimenete közül az egyik az, amikor a cselekedni akarás érvényesül (tehát a befolyásolási akarat kudarcot vall), másik az, amikor a befolyásolási akarat érvényesül. Az első esetben a befolyásolási akarat sikertelen marad, a másodikban sikeres. Ha valaki az utcán elém áll, és pénzt követel tőlem, és én nem akarok adni neki, kitérek az útjából, tovább megyek, akkor a másik nem volt képes érvényesíteni az akaratát velem szemben. Ha viszont a másik fél egy fegyvert fog rám, és azzal fenyeget, hogy lelő, akkor odaadom a pénzemet neki, saját – eredeti – akaratom ellenére.

Konfliktus esetén a mindkét fél elsődleges akarata valamelyikük cselekvésére irányul, de a két akarat eltér egymástól. Az akarat jelzése azonban még nem elég a megvalósuláshoz. A saját akarat érvényesítéséhez szükség van egy második akaratra is, mégpedig arra az akaratra, hogy a másik fél változtassa meg a saját elsődleges akaratát, illetve a partnerben az a másodrendű akarat keletkezik ilyenkor, ami tartalma nem akarni a saját elsődleges akaratának megváltoztatását. A konfliktushoz ez a kétfajta és kétszintű akaratkülönbség elengedhetetlen. Lehetséges, hogy valaki az elsődleges akaratkonfliktus konstatálása után inkább visszalép, nem akarja érvényesíteni saját akarát a másikkal szemben. Viszont ha mégis erre törekszik, akkor az akaratok kétszintű konfliktusa feloldásához szükség van valamire.

A befolyásoló akarattal rendelkező ágens oldaláról szükség van arra a pótlólagos cselekvésre, amelynek eredményeként a cselekedni akaró fél megváltoztatja a saját eredeti akaratát, és a befolyásoló akarat szerint fog cselekedni. A befolyásoló akarattal a kapcsolatba belépő ágensnek elfogadható **indokot** (**gyakorlati indokot**) kell szolgáltatnia arra vonatkozóan, hogy a cselekedni akaró (vagy nem akaró) ágensnek miért kellene megváltoztatnia a saját cselekvésére irányuló eredeti akaratát.

A befolyásoló akarat kifejtése, kifejezése elsőrendű cselekvés, az megfelelő indok szolgáltatása viszont már másodrendű cselekvés (ami egy másodrendű akarattal kapcsolható össze), hiszen a befolyásolni akaró ágens indokot szolgáltató cselekvése (illetve az erre irányuló akarata) a másik fél elsőrendű cselekvésére irányuló akaratának megváltoztatására irányul. Ezt az akaratkonfliktust, illetve annak feloldásának folyamatát az alábbi ábrával próbálom meg szemléltetni:
Háromféle módszer, háromféle indok szolgáltatása lehetséges a cselekvő eredeti akaratának megváltoztatására. Az egyik módszer a — weberi értelemben vett – **hatalom** (POWER), a másik a – weberi értelemben vett - **uralom** (AUTHORITY) gyakorlása, a harmadik a "vitán keresztüli" **meggyőzés** (BELIEF-REVISION). A hatalom esetében az akaratot megtörik ("ellenszegülés ellenére viszik keresztül az akaratot"), az uralom esetében az ellenkezés úgy iktatják ki, hogy egy norma által előírt kötelességre vagy tilalomra hivatkoznak. A vita során érveket hoznak fel egymással szemben a másik akaratának megváltoztatása érdekében.

$will(Agent_1, Do(Agent_2, T(will_1, will_2)))$

1) Ego és Alter között akaratkonfliktus létezik.

$\text{EXIST}(e_2) \wedge \text{will'}(e_2, Ego, e_1) \wedge \text{fly'}(e_1, Ego) \wedge \neg \text{exist}(e_1) \wedge \neg \text{will'}(e_3, Alter, e_1) \wedge \text{exist}(e_3)$ A képletben redundáns az a feltétel, amely annak az eseménynek a létezését tagadja, amely szerint Ego száll ($\text{EXIST}(e_1)$). Azért kell ezt tagadni, mert az Egonak az az akarata, hogy Ego szálljon, ugyan létezik, de maga az az esemény, hogy ti. Ego száll, még nem létezik. És azért redundáns, mert az $\text{EXIST}(e_i)$ propozíció alapértelmezés szerinti igazságértéke 'hamis', amit nem kell a képletekben megjelenítenünk.

2) Ego akarata megvalósul.

$\text{PAST}(e_2) \wedge \text{WILL'}(e_2, Ego, e_1) \wedge \text{FLY'}(e_1, Ego) \wedge \text{EXIST}(e_1) \wedge \neg \text{WILL'}(e_3, Alter, e_1) \wedge \text{PAST}(e_3)$ —

$\text{PAST}(e_2) \wedge \text{WILL'}(e_2, Ego, e_1) \wedge \text{FLY'}(e_1, Ego) \wedge \text{EXIST}(e_1) \wedge \neg \text{WILL'}(e_3, Alter, e_1) \wedge \text{PAST}(e_3)$

HATALOM

POWER(Ego,Alter,Act(Alter,F)) =

hatalom az állatok között: galambrajok belső erőviszonyai
csípkedési sorrendről is valamit összeszedni

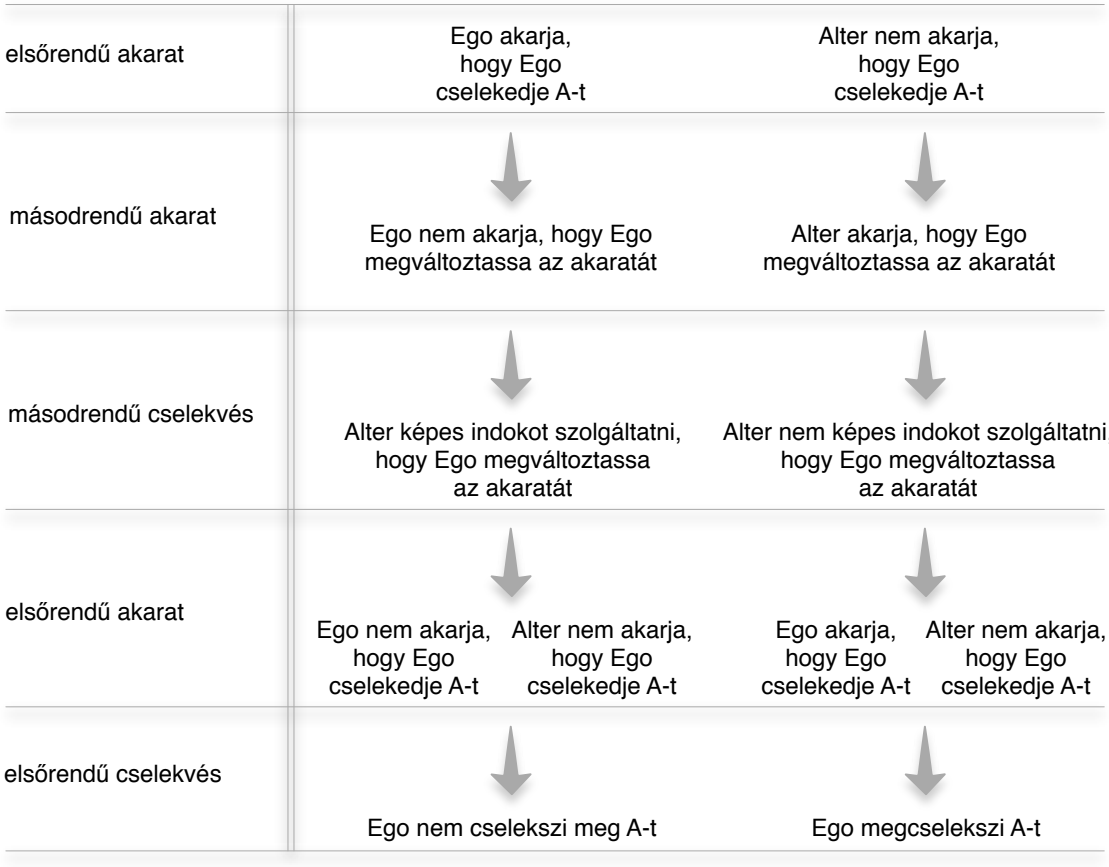


Figure 8.1: a konfliktusfeloldás folyamata

8.3 A kooperáció biológiai mechanizmusai

Cooperation within and among species [629]

Eusociality: origin and consequences [824]

[813]

[814]

ahol b a rokonsági fok, b a recipiens előnye, c a donor költsége r rokonsági fok a genetika, a b/c arány ecology, ökológiai paraméterek változásától függ

8.3.1 Rokoni altruizmus

Hamilton rokoni altruizmus

$rb > c$

reciprocitásban nincsen altruizmus, csak a jövőbe eltolt viszonzás, tehát egoizmuson alapul

haploidipoidy hipotézis

monogámia elősegíti a

poliandria esetén is lehet, de akkor bizonyos feltételeknek teljesülniük kell

kritika: Wilson and Wilson:Rethinking the theoretical foundation of sociobiology

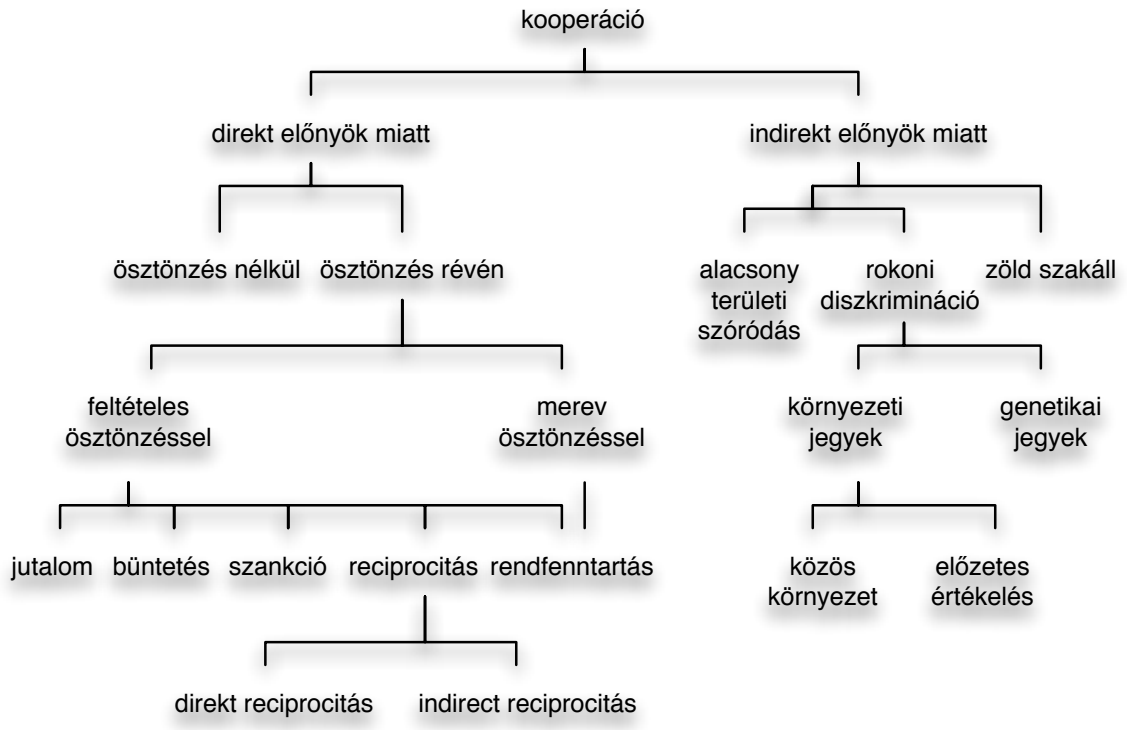


Figure 8.2: a kooperáció típusai – West és tsai ábrája

„Selfishness beats altruism within single groups. Altruistic groups beat selfish groups." [825]

a kooperáció, a reciprocitás sokszor nem kölcsönösséget, viszonzást, hanem inkább azonnali mutualizmust, érdekeltséget jelent: "Cooperation between non-kin in animal societies" tételes listát ad az állatok közti direkt reciprocitás jelenségre, és megmutatja, hogy azonnali kölcsönös előnyök vannak jelen [162]

a rokoni altruizmus (kin selection) az inkluzív fittnesz hamiltoni magyarázósémája helyett lehetséges eredeti evolúciós magyarázatot adni, fölösleges az inkluzív fitnesz elmélet The Evolution of Eusociality [540]

8.3.2 Zöld szakáll hatás

zöld szakáll (green beard) [195, 115]

8.3.3 Spite

két magyarázat

kétszereplős, illetve háromszereplős

a c itt is negatív (vagyis költséges a donor számára, de a b is negatív a recipiens számára, ha a donor kárt okoz a recipiensnek. Ez akkor lehetséges, ha a donor és a recipiens között az átlagoshoz képest nagyon kicsi a rokonsági kapcsolat.

a háromszereplős értelmezése úgy szól, hogy a harmadik szereplő a másodlagos recipiens, akinek előnyös lesz, ha a donor árt a recipiensnek, mivel a másodlagos recipiens helyzete javul (kisebb lesz a versenyképessége a recipiensnek). Ezt a jövőbeni versenyképesség-növekedést egyébként ki lehet terjeszteni akár a donorra is.

8.3.4 Csoportszelekció

greenbeard [195] [814]
skyrms cikke
Wilsonék elmélete

8.3.5 A costly signaling és a handicap principle

Bereczkei Tamás magyar áttekintése [67]

The Handicap Principle. A Missing Piece of Darwin's Puzzle [839]

Zahavi, A. (1975). Mate selection: Selection for a handicap [836]

Zahavi, A. (2003). Indirect selection and individual selection in sociobiology: My personal views on theories of social behaviour. [837]

Costly signaling and cooperation H Gintis, EA Smith, S Bowles - Journal of theoretical biology, 2001 [296]

Zahavi, A. (1995). "Altruism as a handicap – The limitations of kin selection and reciprocity". Avian Biol 26 (1): 1-3. [838]

Altruism as a Signal: Zahavi's Alternative to Kin Selection and Reciprocity Jonathan Wright [829]

High-quality male field crickets invest heavily in sexual display but die young [368]

Egyéb

Mersch DP, Crespi A, Keller L. Tracking Individuals Shows Spatial Fidelity Is a Key Regulator of Ant Social Organization. Science. 2013 Apr 18. [Epub ahead of print]

Chapter 9

Kényszermentes, interdependens társadalmi kapcsolat

Olyan társadalmi kapcsolat, amelyben nincs kényszeralkalmazás, csak szabad akaratkövetés. Ennyiben tehát békés harcról is lehet szó, de itt megjelenhet már a kooperáció jelensége is.

nem kizárólag érdekkövetésről van szó, van viszont interdependencia

A társadalmiság a társadalmi kapcsolatba került cselekvők közti

9.1 Strukturális kényszerek: játékelmélet

játékelmélet társadalmi helyzet, interakció, szituáció

kulcsmozzanat: interdependencia kölcsönös függőség

Amit viszont nem szokták hangsúlyozni, pedig fontos: Kényszermentes társadalmi kapcsolatok

Ebből következően nincs se külső, se belső kényszer tehát hatalommentes és uralommentes (utóbbi persze csak módjával igaz) nincs norma, nincs kényszer, nincs uralom, nincs hatalom

Ebből következően a játékelmélet az érdek (belső társadalmi motiváción) alapú társadalmi kapcsolatok elmélete

stratégia versengés

A játékelméleti építkezés során a társadalom lehető legegyszerűbb modelljéből indulok ki. A legegyszerűbb itt egyszerre sok feltétel teljesülését jelenti. Egyrészről a lehető legkisebb társadalmat tételezem fel, ami két főből áll: Ego és Alter alkotja. Ego az a személy, akinek a cselekvését meg akarjuk érteni, magyarázni, Alter pedig az a személy, aki szükséges a társadalmi kapcsolat létezéséhez (és ezzel a társadalom legkisebb egységének létrejöttéhez).[1]

A játékok bemutatása során három korábbi játéktaxonómiára hivatkozom. A 2x2-es játékok első teljességre törekvő leírását Rapoport és Guyer adták meg, akik a preferenciamátrix sajátosságai alapján végezték az elemzésüket, de nem foglalkoztak komolyabban játékosztályok kialakításával, a játékok tipizálásával, illetve nem adtak neveket sem a

[1] A játékosok megragadására sokféle jelölési konvenció létezik, amelyek közül olykor használom a 'Sor' és 'Oszlop' játékos megnevezést.

játékoknak (sőt, nem is mutatták be őket – például történetek segítségével). A játékok számosságára vonatkozóan Rapoport és Guyer 1966-os cikke alapján közel negyven éven keresztül elfogadott volt az a nézet, hogy 78-féle játékot lehet elkülöníteni egymástól. Ez a szám úgy jött ki, hogy a 2x2-es preferenciamátrix 2x4 értékét 576 (4!x4!) féle módon lehet elrendezni, amelyek közül viszont sokatat ekvivalensnek lehet tekinteni egymással. Rapoporték úgy gondolták, hogy a mátrix sorait és oszlopat felcserélhetjük anélkül, hogy a játék logikáját, belső viszonyrendszerét megváltoztatnánk. Ezért 4-4 mátrix összekapcsolhatóvá vált azon az alapon, hogy ezek csak formai reprezentációjukat tekintve különböznek egymástól. Ezentúl azonban úgy gondolták Rapoporték, hogy attól is eltekinthetünk, hogy melyik játékos szemszögéből "nézzük" a játékot, vagyis ugyanazt a játékot kapjuk, ha a sorjátékos (Ego) vagy az oszlopjátékos (Alter) felől tekintünk a preferenciamátrixra. Ez szimmetrikus játékok alapján mindenképpen igaz, de lehet amellett érvelni, hogy az aszimmetrikus játékok esetében igenis van különbség az eltérő nézőpont miatt a játékok értékelésében.

Ugyan kevesebb játékra figyelt, viszont a legtöbb játékleírást Steven J. Brams adta meg 1994-ben. Brams kizárta az elemzésből azokat a játékokat, amelyek nem konfliktusosak (amikor a kölcsönösen legjobbnak rangsorolt kimenet a játék végeredménye), így csak 57 játékot vizsgált [115].

A közel negyven éves konszenzust rúgta föl 2005-ben David Robinson és David Goforth könyve a 2x2-es játékok topológiájáról [608]. A két szerző 78 helyett 144 játékra kiterjesztve mutatta be azt, hogy milyen topológiai leírás adható a 2x2-es játékokra, milyen játékosztályokat lehet elkülöníteni közöttük.

A játékokat több szempont szerint jellemezhetjük. Figyelhetünk arra, hogy:

- hányan játszák a játékokat (lehet két- vagy többszereplős a játék)
- hány döntési alternatíva van a játékosok előtt (lehet két- vagy többalternatívás játék)
- hányszor játszanak egymással a játékosok (lehet egyszer vagy többször játszott, tehát egyszeri vagy ismételt a játék)
- azonosan értékelik-e a játékosok a játékot vagy sem (lehet szimmetrikus vagy aszimmetrikus a játék)
- a preferenciarendezések szigorú rendezésen alapulnak-e vagy sem, vagyis megengedhető-e az, hogy két döntési alternatívát azonosnak minősítsen egy játékos, esetleg egy alternatívát ne értékeljen (lehet szigorúan rendezett vagy nem szigorúan rendezett a játék)
- ismerik-e egymást a játékosok vagy sem, aminek az ismételt játékok esetében van kiemelt jelentősége (lehet direkt vagy indirekt reciprocitás játék)
- kommunikálhatnak a játékosok a döntések meghozatala során vagy sem (lehet nemkooperatív, "néma" vagy kooparatív, "hangos" a játék[2])
- egyszerre döntenek-e a játékosok vagy egymás után, egyik a másik döntésének ismeretében (lehet szinkron, szimultán, statikus vagy aszinkron, szekvenciális, alternáló, dinamikus a játék)

a kooperatív játék: [684, 33]??????? a) a játékosok egymástól függetlenül választják stratégiáikat b) a játékosok nem tehetnek a játék megkezdése előtt olyan ígéreteket, ame-

[2]Meg kell jegyezzem, hogy ezen a ponton nagyon szerencsétlen a kooperatív minősítés, mert a terminusnak egész más jelentése van itt ahhoz képest, ahogy a játékelmélet egészében szóba kerül a kooperativitás, a kooperáció kérdése.

lyeket be lehetne tartatni a játék folyamán. Ez nem jelenti azt, hogy esetleg ne kooperálnának, de ezt önkéntesen kell tenniük. ————

és a többi szemponttal mi van??? tiszta stratégia vs. kevert stratégia információs aszimmetria teljes információs játék, mindent tudnak a játékosok, köztudott tudás vs. hiányos információs játék árverési helyzetek (Simonovits 27.o., Binmore)

méltányosságelméletek, erős reciprocitás

(IL.0) $\mathscr{L}_{IL}(\mathscr{L}_{??}, \mathbb{N}, \mathbb{S},)$, ahol

---

$\mathscr{L}_{??}$
$\mathbb{N}$
$\mathbb{S}$

Formális keretrendszer

(IL.1) $\text{PLAYER} \stackrel{\text{def}}{=} \mathbb{N} = \{1,2,\ldots,n\}$

---

játékosok (HUN) – player (ENG) – Ego, Alter ((???))

(IL.2) $\text{STRATEGIES} \stackrel{\text{def}}{=} \forall i \in \mathbb{N} \exists S_i (S_i = \{s_i^1, \ldots, s_i^n\})$

---

játékos stratégiahalmaza (HUN) – player's strategy set(ENG) – lépéshalmaz (HUN)

(IL.3) $\text{PAYOFF} \stackrel{\text{def}}{=} \forall i \in \mathbb{N} \exists \mathbf{u}_i (\prod_i S_i \mapsto R)$

---

kifizetés(i függvény) (HUN) – payoff (function) (ENG)

(IL.4) $\text{DOMINANT_STRATEGY} \stackrel{\text{def}}{=}$

---

domináns stratégia(HUN) – dominant strategy (ENG)

A játékosok szociális preferenciái

A társas értékorientáció elmélete olyan eszköztárat fejlesztett ki, amely alkalmas a játékelméleti szituációkba került játékosok valós szociális preferenciáinak feltárására (↬295.o.). Erre történetileg azért volt szükség, mert a játékok interdependens mivolta miatt nem lehetett kísérleti eszközökkel vizsgálni a résztvevők tényleges motivációit, mivel az interdependencia ténye "elfedte" a motivációkat. Hogy az egyes játékosok irányultságai feltárhatók legyenek, a dekomponált játékok módszerét dolgozták ki.

A társas értékorientáció elmélet keretében végzett rengeteg kísérlet számszerűen is kimutatta, milyen arányban vannak a társadalomban különböző társas beállítottságú emberek. Ezt a tudást a játékelmélet bemutatása során is érvényesíteni kell.

9.2 Magjátékok

A társadalmi kapcsolatok játékelméleti modellezését úgy építhetjük fel, hogy kiindulunk a lehető legegyszerűbb, legkisebb méretű társadalom modelljéből, majd lépésenként egyre több kezdeti feltételt "elengedünk", hogy a társadalom egyre összetettebb, egyre bonyolultabb modelljéhez juthassunk. Ebben az építkezésben a preferenciamátrix van a központban. Kiindulásként rögzítjük, hogy egyszerre érvényesül a kétszereplős, kétalternatívás, egyszeri, szimmetrikus, szigorúan rendezett, szinkron, nem kooperatív játékok

feltételegyüttese. Ezeket a játékokat korábban magjátékoknak neveztem el [722], és – ahol lehetett – alapos leírást is adtam rájuk (sok példával), ezért itt csak a legfontosabb vonásait mutatom be a játékoknak.

magjátékok	
aszimmetrikus játékok	szimmetrikus
többalternatívás játékok	kétalternatívás
többszereplős játékok	kétszereplős
nem szigorúan rendezett játékok	szigorúan rendezett
aszinkron játékok	szinkron
ismételt játékok	egyszeri
direkt reciprocitás játékok	ismerős játékosok
indirekt reciprocitás játékok	ismeretlen játékosok
kooperatív (hangos) játékok	nem kooperatív (néma)

A játékokat a preferenciamátrix alapján jellemezhetjük. A preferenciamátrix a játékosok alternatívákra vonatkozó rendezéseit fejezi ki, amit számokkal jelölünk. A preferenciamátrix segítségével kifejezhető strukturális kapcsolatokat (kényszereket) ábrázolhatjuk grafikusan is. Kétféle geometriai (topológiai) reprezentációt is figyelembe veszek.

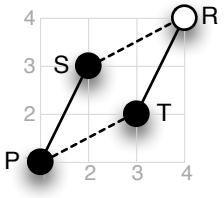
Egyfelől átveszem Robinson-Goforth könyvéből a játékokat azonosító (és egyben jellemző) topológiai ábrákat, amelyek úgy reprezentálják az egyes játékokat, hogy a két játékos preferenciaértékeit egy közös 4x4-es koordinátarendszerben tüntetik fel. A sorjátékos (Ego) értékeit a vízszintes, az oszlopjátékos (Alter) értékeit a függőleges tengely mentén fejezik ki, és az értékek metszetébe egy-egy pontot rajzolnak. Ezután összekötik azokat a pontokat egymással (folytonos, illetve szaggatott vonallal), amelyek egy-egy játékos döntései által összekapcsolódnak, ami másfelől egyfajta döntési kényszert is jelent a másik játékos számára [608].

4,4	2,3
3,2	1,1

$\rightsquigarrow$

R	S
T	P

$\rightsquigarrow$



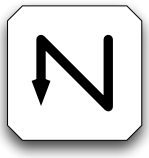
Az ábrában az egyensúlyi kimenet(ek) létezését a fehér közepű köröcske jelzi.

A játékok képi azonosítására dolgozott ki Pierre Dragicevic egy olyan vizuális jelölési rendszert, amely azzal a megoldással különíti el egymástól az egyes játékokat, hogy a preferenciamátrix négy cellájában szereplő preferenciaértékeket összeköti egymással növekvő sorrendben, és így mindig a játékra jellemző, egyedi, vonalszerű alakzatot kap [216]. Dragicevicnek ezt az egyszerű, de látványos megoldását (még tovább egyszerűsítve, a "végletekig stilizálva") szintén alkalmazom az egyes játékok jellemzésekor.
A játékok leírásakor feltüntetem még a preferenciamátrixukat, valamint a széles körben elterjedt cellamegnevezési konvenció alapján kialakítható SRPT-rendezéseiket. A preferenciamátrix celláit megnevezhetjük és azonosíthatjuk az S, R, P, T betűk segítségével (sorrendben a Sucker, Reward, Punishment, Temptation kifejezések kezdőbetűivel a játékdilemmára jellemző állapotok jellemzésére alkalmazva), és a játékokat egyértelműen jel-

3,3	1,4
4,1	2,2

↦

3	1
4	2

↦ 

lemzi az, hogy a négy cellához milyen preferenciaértékek vannak rendelve, tehát a négy betű sorrendje egyedi módon jellemzi a játékot, így megfelel azonosítónak is.

A 12 szimmetrikus játék között mindegyiknek van egy párja, amellyel szorosabb kapcsolatban van a többiekhez képest. Amennyiben a kisátlóban szereplő értékeket felcseréljük, miközben a főátló értékeit változatlanul hagyjuk, akkor mindig az adott játék komplementer párjához jutunk. Robinson és Goforth antijátéknak nevezi az így kapott játékokat, amit én nem tartok szerencsés megnevezésnek.[3] A komplementer játékokat úgy is lehet értelmezni, hogy a kisátló preferenciaértékeinek felcserélése voltaképpen azt fejezi ki, hogy a játékot két olyan játékos játssza, akik az alapjáték játékosaihoz képest ellentétes ágens-beállítódással rendelkeznek (azaz a játékban két egoista beállítódású helyett két altruista játékos van).

Mivel az alapjátékokat máshol korábban alaposan elemeztem [722], itt nem fogok sok példát adni a játékokra, csak akkor, ha valamilyen fontos vonás érzékeltetése végett van szükség egy új példa bemutatására, elemzésére, vagy olyan játékot elemzek, amelyre a korábbiakban nem tudtam példával szolgálni.

————- kooperáció és dezertálás kettőssége

terminológiai tisztázás:

kooperáció = két (vagy több) játékos között fennálló, kooperációs mintázatot mutató, hosszabb folyamat, cselekvés/döntéssorozat, tehát 1) interakció + 2) folyamat, állapot

kooperál = egy játékos kooperál (egy cselekvés, döntés) ehhez közeli jelentéssel: kooperatív cselekvés, viselkedés

„ We define cooperation as an interaction between individuals that results in net benefits for all of the individuals involved. Parasitism is defined as an interaction that yields a net benefit for one individual but a net cost for other(s), and competition as an interaction that entails net costs for all players. Following traditional terminology (Bronstein, 2001), cooperation refers only to intraspecific interactions; we use the term mutualism to refer to equivalent interspecific interactions.” [68]

cooperative behaviour is: an individual behaves cooperatively if the resulting payoff for its partner is higher than the average payoff for the relevant type of interaction. According to this definition, even parasites can behave cooperatively if they exploit their host less than do their conspecifics, even though the outcome would still be called parasitism.

cooperative behaviour may be due to by-product mutualism, pseudo-reciprocity, negative pseudo-reciprocity (i.e. sanctions), positive reciprocity, negative reciprocity (i.e. punishment) or indirect reciprocity.

Hamilton négyes klasszifikációja azon alapszik, hogy Ego és Alter eredményét hogyan befolyásolja Ego cselekvése De a társadalmi kapcsolatban interdependencia van, értelmetlen egyetlen cselekvés hatásáról beszélni

[3] Például azért nem, mert a koordinációs játék antijátékát anti-koordinációs játéknak kéne nevezni, ami azért nem lenne jó, mert ezen kifejezéssel már régóta a koordinációs helyzetek egyik típusát írják le.

———

Van-e játékoknak valamilyen konzisztens elnevezési logikája? Sajnos, nincs.

játékokhoz kapcsolódó szerepnevek gyakran előfordulnak ezek olykor negatív minőségűek, máskor meg pozitívak Rapoport egy cikkében négy játékra vonatkozóan konzisztens gyakorlattal rukkol elő arra figyel, hogy "mi lenne ha?" ki tudja megoldani a játékot? Rapoport négyféle játékostípust ír le, és háromféle (????ez biztos?????) játékra hivatkozik, miközben kétféle játék agoista és altruista interpretációjáról van szó

a fogolydilemma komplementerét sokféleképpen nevezik (holtpont, PD delight, mártír)

többféle elnvezést érezhetünk intuitívnak: kizsámányoló vagy mártír a PD-ben, de a balek vagy potyautas ugyancsak a PD-ben

gyáva nyúl vagy önkéntes a chicken játékban, de

a magasabb konzisztencia reménye miatt nem érdemes eltérni a már kialakult szóhasználati gyakorlattól, mégha igény lenne rá

———

konfliktusos vagy konfliktusmentes, harmónikus játék

társadalmi dilemma, társadalmi csapda, közlegelők tragédiája [575], volunteer's dilemma [205]

——— [684, 33] játékok végeredménye, egyensúlyi helyzete,

Pareto-optimalitás: Egy elosztást Pareto-optimálisnak nevezünk, ha nem le?tezik egy olyan másik elosztás, amely szintén megvalósítható, amelyben senki sem jár rosszabbul és legalább egy szereplő jobban jár, mint a szóban forgó elosztásban.

Nash-egyensúly fogalma, domináns stratégia/lépés, dominált vs. domináns (Ha van domináns stratégia, akkor minden más stratégia dominált. Általában feltehetjük, hogy a játékosok nem játszanak szigorúan dominált stratégiákat. Sőt, érdemes a játékban egymás után, iteratívan kiküszöbölni a szigorúan dominált stratégiákat. Belátható, hogy a végeredmény független a kiküszöbölés sorrendjétől. A maradék játék már jóval kisebb és áttekinthetőbb, szélső esetben egyetlen egy stratégiaegyüttesből áll.), szigorúan dominált/domináns stratégia, gyengén dominált/domináns stratégia

stratégia mint döntés, lépés, cselekvés stratégia mint egyszeri döntés vagy döntéssorozat

Személyiség, neveltetés, kapcsolati tőke, társadalmi beágyazódás, és az ezekből fakadó döntési különbségek: kockázatkerülő, kockázatkövető, kockázatdomináns. kifizetésdomináns személyiségek

———

csomagolás hatása (mérő könyv, Simonovits példa:[684, 33]

	magának ad	másiknak ad
C	1	2
D	2	0

Ha a játékos a C gombot nyomja meg, akkor magának 1-et ad, a másiknak 2-t; ha a verseng gombot nyomja meg, akkor magának 2-t ad, a másiknak 0-át.

A játék fogolydilemma. Ennek ellenére ezt a játékot a kísérletekben sokkal kooperatívabban játsszák, mint az eredetit, mert nyilvánvaló benne, hogy csak a másiktól jöhet az igazi nyereség.

———-

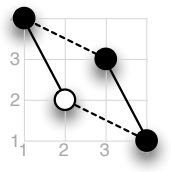
korábban esetszétválasztás módszereként hivatkoztam rá most: **függetlenségi axióma**, lényegtelen alternatívák függetlenségének elve, feltételes valószínűségek függetlenségének elve, biztos dolgok elve (sure thing principle) [363], Savage az alap ez csak akkor alkalmazható, amikor a döntési teret fel lehet bontani egymástól fügetlen világállapotokra

Az elv téves alkalmazásra nézzük meg a következő példát [363]. A vizsgája előtt egy diáknak két lehetősége van (tanulni vagy nem tanulni), és a vizsga kimenete is kétesélyes (vagy átmegy a diák vagy nem). Ezt a döntési teret felrajzolhatjuk a következő mátrixba:

	tanul	nem tanul
átmegy a vizsgán	2	3
megbukik a vizsgán	0	1

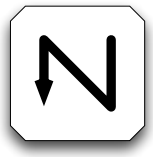
Mind Azért nem alkalmazható a sure thing elv, mert itt nem függetlenek egymástól

Fogolydilemma játék



3,3	1,4
4,1	2,2

TRPS



a játék száma	
Rapoport-Guyer:	12
Brams:	32
Robinson-Goforth:	111

A játékelmélet története is azt mutatja, és sok kutató elméletileg megalapozott véleménye szerint is a fogolydilemma áll a szimmetrikus játékok (sőt, többek szerint, az egész játékelmélet) középpontjában. Sok szempontból osztom én is ezt a véleményt, de azt azért meg kell jegyeznem, hogy van egy olyan elfogultság ebben a minősítésben, amely a konfliktusos jelleg fakad. Amikor John Platt a társadalmi csapdákról szóló, híres, 1973-as cikkében "metaforát keres" a fogolydilemma számára, akkor Adam Smith 'láthatatlan kéz' (invisible hand) fogalmát állítja szembe az által kreált 'láthatatlan ököl' (invisible fist) kifejezéssel [575]. Az "áldást hozó láthatatlan kéz" fogalmával Adam Smith azt az állapotot ragadja meg, amikor az egyéni érdekek megegyeznek és a közösségi érdek nem sérül. Ezzel áll szemben John Platt szerint a "láthatatlan ököl, amely lesújt a közjóra", amikor az egyéni érdekek teljesen ellentétesek, és a közösségi érdek is sérül. Mindez azért érdekes itt, mert a láthatatlan kéz harmóniát kifejező fogalmához is rendelhető egy játék (a harmóniajáték), amit persze a játékelmélet szintúgy mostohán kezel, mivel a konfliktus hiányában, a teljes harmónia miatt "nincs benne semmi érdekes".

A fogolydilemma legfontosabb tanulsága, hogy vannak olyan helyzetek, amelyekben a kétszeres dezertálás lesz a játék várható kimenete, mert a bizalmatlanság és/vagy erős egoista beállítódás miatt senki meri/akarja a kooperációt felválllani. Ekkor viszont mindenki kisebb nyereséggel jön ki a játékból a kettős kooperáció esetén elérhető kifizetéshez képest. A játék várható szuboptimális végeredménye adja a játék társadalmi dillema jellegét.

A fogolydilemma játék eredeti meséje bár életszerű, de nem mondható túl gyakorinak, viszont számtalan társadalmi helyzetről lehet állítani, hogy ebbe a játékosztályba tartozik. A közös társadalmi javak, a közlegelők egyéni túlfogyasztása (szemetelés a

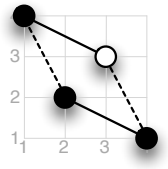
köztereken, az ingyen parkolóhelyek “feltöltése”), a potyautasság jelenségének számtalan megnyilvánulása (bliccelés a tömegközlekedésben, lakótelepi lakásokban a fűtés “lopása” a szomszédoktól), a piaci csökkenő árspirál jelenségének kialakulása, vagy az árucsere helyzetek mélyén mindig is ott rejtőzködő késztetés a partner “átverésére” – olyan példák, amelyek jelzik, hogy a társadalmi cselekvések minden szegmensében felbukkannak fogolydilemma-szerű konfliktusok [722].

társadalmi csapda, szuboptimális megoldás, társadalmi dilemma

potyautas, balek, mártír, kizsákmányoló Rapoport javasolta az utóbbi nevet. nem azért, mert a fogolydilemmában valóban az egyik játékos kizsákmányolója lenne a másiknak, bár ez is előfordulhat ebbena játékban. De tudjuk, hogy nem ez az alapértelmezés szerint elvárt végeredmény, hanem a kétszeres dezertálásból eredő szuboptimális kimenet, amikor egyik sem zsákmányolja ki a másikat. Mégis

Ha a játékosok cselekvési lehetőségeit a végletekig absztraháljuk, akkor az együttműködés, kooperálás (C), illetve a nem-együttműködés vagy másként a dezertálás (D) dichotómiájához jutunk. Mindkét játékos kooperálhat vagy dezertálhat, így a játéknak négyféle kimenete lehetséges: kölcsönös kooperálás (CC), kölcsönös dezertálás (DD), az egyiknek balekság, a másiknak potyázás (CD), illetve fordítva (DC). A fogolydilemma-helyzetek tanulsága az, hogy hiába lenne a kölcsönös kooperáció mindenkinek előnyösebb a kölcsönös dezertáláshoz képest, a bizalmatlanság miatt a dezertálás (a potyautasság, a másik kizsákmányolásának szándéka) válik uralkodó motívummá. Ezt az az összefüggést fejezi ki a következő egyenlőtlenség (preferenciarendezés): DC > CC > DD > CD.

Mártírjáték



3,3	4,1
1,4	2,2

SRPT

a játék száma	
Rapoport-Guyer:	9
Brams:	9
Robinson-Goforth:	166

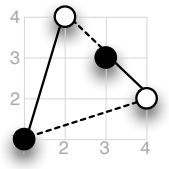
A mártírjáték a fogolydilemma komplementár párja. Szokták holtpontjátéknak (Deadlock) is nevezni [581], Robinsonék az ‘Anti-PD’ nevet használják, Ken Binmore pedig a ‘Prisoner’s Delight’ nevet húzta rá [86, 131]. A mártírjáték legfontosabb vonása az, hogy egy fogolydillema-szerű szituációból képes kihozni a “közösségi optimumot” a játékosok számára. Ezzel egyben arra is például szolgál, hogy nemcsak az önzés nem mindig racionális, de előfordulhat az is, hogy az altruizmus nagyobb társadalmi nyereséget biztosít az önzéshez képest [722].

balek, mártír azért mártírjáték, mert két altruista beállítódású személy játékaként lehet felfognunk. Pontosabb lenne a ‘mártírok játéka’

Robert Frank, Thomas Gilovich és Dennis Regan 1993-as tanulmányukban leírták, hogy amikor fogolydilemma játékot játszattak egyetemistákkal, akkor a végzős közgazdászhallgatók között sokkal nagyobb arányban fordult elő a kölcsönös dezertálás, mint más hallgatók körében, amiből azt a következtetést vonták le, hogy a dezertálás – legalább részben – tanult lehet.

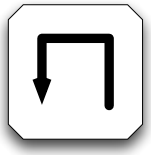
[264]

Gyáva nyúl játék



3,3	2,3
4,2	1,1

TRSP



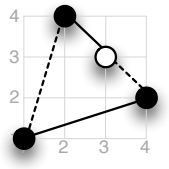
a játék száma	
Rapoport-Guyer:	66
Brams:	57
Robinson-Goforth:	122

Ha a fogolydillemmában a két legkisebb preferenciaértéket felcseréljük egymással, akkor jutunk a gyáva nyúl játékhoz. Ezáltal a kétszeres dezertálás lesz mindkét játékos számára a legrosszabb alternatíva, de ez csak azért fontos, mert emiatt tolódik el a játékosok választása a kisátlóban levő két cella felé. Ez viszont azt jelenti, hogy a játék döntetlen jellegű lesz, mert két ekvivalens egyensúlyi kimenettel fog rendelkezni, amelyek közül nem lehet előzetesen megjósolni, melyik lesz a játék tényleges végeredménye. Ez az első koordinációs játék, amikor az a játék alapproblémája, hogy milyen módon lehet úgy koordinálni a játékosok döntését, hogy ugyanazt a kimenetet válasszák mindketten.

A gyáva nyúl játékra is rengeteg valós példát lehet találni. A “névadó” filmbeli történet az egymással szembe (vagy a szakadék felé) haladó bandavezérek versengéséről szól, amikor az egyikből fenyegyerek, a másikból gyáva nyúl válhat, de ebbe a játékosztályba tartozik a hidegháború korszakát jellemző kölcsönös elrettentés politikája, vagy mindazok a hétköznapi történetek, amikor az a kérdés, hogy ki vállalja fel az önkéntes szerepét (aki leviszi a szemetet, elmosogat, kitakarít, lemegy a boltba stb.).

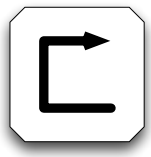
gyáva nyúl (negatív szerep), önkéntes (pozitív)

Önkéntes játék



3,3	4,2
2,4	1,1

SRTP

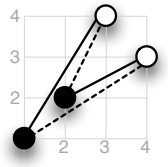


a játék száma	
Rapoport-Guyer:	7
Brams:	7
Robinson-Goforth:	155

Az önkéntes játék a gyáva nyúl játék komplementer párja, és legfőbb vonása az, hogy a gyáva nyúl játék döntetlenjét képes feltörni és az optimális közösségi megoldást adja a játékosok számára. A játék kevés figyelmet kap a szakirodalomban. Az ok ugyanaz, mint a mártírjáték esetében: sokak számára a konfliktusmentessége miatt érdektelennek tűnik a játék.

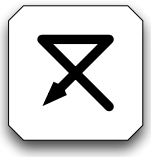
Robinson és Goforth anti-gyávanyúl játéknak (Anti-Chicken) nevezi a játékot. Az általam preferált önkéntesjáték megnevezés arra utal, hogy ezt a játékot a gyáva nyúl játék keresett/remélt önkéntesei játsszák. Amíg azonban a gyáva nyúl játékban az önkéntes magára vállalja a második legrosszabb kimenetet biztosítva, hogy ezzel a partnere a legnagyobb nyereménnyel zárhassa a játékot, addig itt két önkéntes játékos kerül egymással szembe, akik kölcsönösen a második legjobb eredménnyel szállhatnak ki a játékból.

Vezérjáték



2,2	3,4
4,3	1,1

TSRP



a játék száma	
Rapoport-Guyer:	68
Brams:	54
Robinson-Goforth:	133

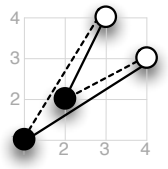
Talán ennél a játéknál (illetve akomplementer párjánál) tapasztalható a legnagyobb névkavarodás a játékelméleti szakirodalomban. A nemek harca megnevezést alkalmazzák erre a játékra és a párjára egyaránt. Bár ez a keveredés önmagában mutatja, hogy valóban értelmes és érdemes a komplementer játékokat úgy felfogni, hogy azokban a játékosok ágens-beállítódásai vannak felcserélve, hiszen amíg a nemek harca játékban az önző beállítódású játékosok a vezér szerepre törnek, addig az önzetlenek az áldozatvállalásra hajlandó hős szerepét vállalják magukra. Az egyértelműség kedvéért ezt a játékosztályra a vezérjáték megnevezést veszem át Rapoport egyik cikkéből [594], míg a játék komplementer párjára a hősjáték nevet fogom használni. A vezérjátékra egyébként magyarul használják még a vezérürü megnevezést is, de ezt én hanyagolnám.

A vezérjáték is koordinációs problémák közé tartozik, akárcsak a gyáva nyúl játék, de annyiban azért más a játék mintázata, hogy a két ekvivalens egyensúlyi kimenet itt a kisátlóban levő két cellában van. Ez annyit tesz, hogy az egyensúlyi állapotok valamelyikét úgy lehet csak elérni (és ezáltal a sikeres koordinációt megvalósítani), ha a két résztvevő eltérő cselekvést hajt végre (az egyik kooperál, a másik dezertál). Ezért van az, hogy ezeket a játékokat valamilyen szerepmegosztás (nemi, családi, foglalkozási munkamegosztás) modellezésére tartják alkalmasnak. A játéknak azért javasolta Rapoport a vezér megnevezést, mert a döntetlent egy dominanciára törekvő játékos tudja feltörni azzal, hogy képes elérni a partnerénél azt, hogy fogadja el őt vezérnek, aki így a játék legjobb eredményét szerezheti meg magának, de ezzel párhuzamosan "képes" biztosítani a második legjobb kifizetést a partnerének. A komplementer párjához képest érdekes és fontos vonása még a játéknak, hogy a kétszeres versengés adja a legrosszabb kifizetést mindkét játékos számára (míg a hősjátéknál ez pont fordítva van, ott a kétszeres kooperáció a legrosszabb).

A játékra sok hétköznapi példa hozható. Ide tartoznak azoka helyzetek, amelyekbe az útkereszteződésbe érő, a kanyarodásban egymást zavaró gépkocsivezetők vagy az egy nyomvonalat, két irányból használni akaró közelekedők (sofőrök, gyalogosok) kerülnek, de vezérjátékbe kerülnek azok is, akik között megszakad a telefonbeszélgetés, és el kell dönteniük, ki hívja vissza a másikat.

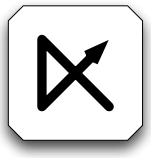
Hősjáték

A hősjáték a nemek harca, a kultúrák harca, a szerepmegosztást modellező játék altruista beállítódású játékosok által játszott változata (és egyben a vezérjáték komplementere). A mintázata ugyanolyan, mint a vezérjátéké: két egyensúlyi kimenete van. A döntetlent itt akkor lehet feltörni, ha a két altruista hős közül az egyik "felvállalja" az egoista szerepét.



2,2	4,3
3,4	1,1

STRP

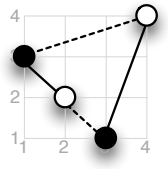


a játék száma	
Rapoport-Guyer:	69
Brams:	55
Robinson-Goforth:	144

Ez a játék idézi elő azt a furcsa helyzetet, hogy a kétszeres kooperálás a legrosszabb mindenki számára.

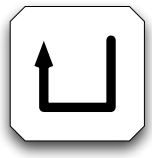
A játék illusztrálására a klasszikus nemek harca történetét érdemes legelőször meghivatkozni (amikor a férjnek és feleségnek külön-külön kell eldöntenie, hogy hova menjenek estére, boxmeccsre vagy színházi előadásra). A gazdaság területéről származó példaként lehet említeni azt a helyzetet, amikor két szabványt használó vállalatnak kellene megegyezniük egy közös szabvány követésében. Hősjáték példája még az is, amikor két tudósnak kell megegyeznie arról, melyikük anyanyelvén írják meg a közös cikkük egy-egy fejezetét, vagy amikor két cég egy hosszú szerződéselőkészítési folyamat végén meg kell egyezzen a kötbér nagyságáról.

Szarvasvadászat játék



4,4	1,3
3,1	2,2

RTPS



a játék száma	
Rapoport-Guyer:	61
Brams:	-
Robinson-Goforth:	322

A szarvasvadászattal kezdődő hatos sorozat tagjai abban mutatnak közös vonásokat, hogy mindegyik játékosztály preferenciamátrixában a kétszeres kooperációról elnevezett R-cellában érheti el mindkét játékos a legnagyobb nyereményt. Emiatt ez a cella mindegyik játékban az egyik lehetséges egyensúlyi kimenetet adja (amely mellé három esetben még egy másik egyensúlyi kimenet is "bekerül").

A szarvasvadászat játékot több szerző szerint a fogolydilemmából lehet "származtatni" úgy, hogy a két legmagasabb preferenciaértéket (3-t és 4-t) kicseréljük egymással. Ebben az esetben mindkét játékos számára a kölcsönös kooperálással lehet megszerezni a játék legnagyobb nyereményét. Bár ebben a játékban a dezertáló fél még mindig alaposan "ki tudja használni" a kooperáló játékost, hiszen a második legnagyobb kifizetést szerezheti meg (3-t), miközben a kooperáló partnere a legrosszabb nyereményhez juthat csak hozzá (1-t), a játékra mégsem ez az asszimmetria az igazán jellemző. Első látásra talán meglepő, de az a lehetőség, hogy a kölcsönös kooperálással mindkét játékos megszerezhetné a legmagasabb kifizetést, ez mégsem történik meg automatikusan, mert a játéknak két egyensúlyi kimenete van: a kétszeres dezertálás is Nash-egyensúlynak felel meg. A kölcsönösen legjobb kifizetéshez szükséges a kölcsönös kooperálás, és ezzel szemben "kényelmes" alternatívát jelenthet mindkét játékos számára az, ha a partner kooperálása mellett ő dezertál (hiszen ekkor a második legjobb kifizetéshez juthat hozzá). Ez a másik kárára szerezhető "könnyű nyeremény" elfogadható teszi a legnagyobb nyereményről való lemondást, ezért a játékban a kétszeres dezertálás is egyensúlyi helyzetté

válik. Ebben az értelemben és ezen a ponton még mindig a fogolydilemma "hatása" érezhető.

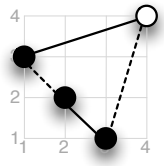
Lehet azonban másfajta értelmezést is adni a szarvasvadászat játék származtatására vonatkozóan. Kiindulhatnánk ugyanis a harmóniajátékból, amelyben ha megcserélnénk a két legkisebb értéket (az 1-t és 2-t), akkor a szarvasvadászat játékhoz jutnánk. Ezt úgy lehetne interpretálni, hogy ha a kölcsönösen legjobb kifizetést biztosító harmóniajátékban a kétszeres dezertálás kifizetését növeljük "eggyel" (már nem a legrosszabb, hanem a második legrosszabb lesz a jutalom), akkor ez elég lehet ahhoz, hogy a kétszeres dezertálás"összemérhető" és vállalható alternatíva a kétszeres kooperáláshoz képest legyen mindkét játékos számára. Ez az elmozdulás pedig kibillentheti a játékot az egyszeres egyensúlyi helyzetből és két Nash-egyensúlyi pontot eredményezhet.

A példák között a "névadó" szarvasvadászat játék mellett említeni érdemes a matrózlázadás történetét.

Brian Skyrms egy egész könyvet szentelt annak a tézisének, miszerint igazából a fogolydilemma áll a játékok központjában, belőle vezethető le a fogolydilemma, és nem fordítva, ahogy sok más szerző ezt gondolja [688]. Robinson és Goforth említi, hogy bizalmi dilemmaként (trust dilemma), illetve koordinációs játékként (coordination game) is felfogható a játék [608], Amartya Sen 1967-es cikke óta nevezik olykor biztosítási játéknak (Assurance Game [660]), Sam Perlo-Freeman pedig Deterrer játéknak (Deterrer) nevezi [566].

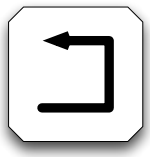
jobb ma egy veréb, mint holnap egy túzok.

Háztulajdonos játék



4,4	3,1
1,3	2,2

RSPT



a játék száma	
Rapoport-Guyer:	5
Brams:	-
Robinson-Goforth:	355

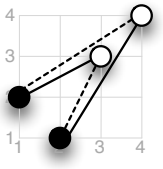
A játék a szarvasvadászat komplementere, és ezt is kétféle módon származtathatjuk.

A játéknak azért adtam a 'háztulajdonos' nevet, mert azt a logikát írja le, ami a vigalmi negyedek kialakulásának hátterében áll. Kétszereplős játékként kicsit életszerűtlen ugyan a történet, de még így is jól visszaadja azokat a belső kényszereket, amelyek a játék preferenciamátrixában meghúzódnak. Van két egymás mellett fekvő ház, a tulajdonosaik dönthetnek úgy, hogy bordélyházat csinálnak a házából vagy kiadják bérlőknek a házaik lakásait. A legvonzóbb kimenet mindkettejük számára az, ha bordélyházat nyitnak, mert sokat kereshetnek vele, nem zavarják egymást, sőt a két háztömbből álló "mini vigalmi negyed" kialakulásával még nagyobb forgalomra is számíthatnak, hiszen nagyobb híre megy a szolgáltatásuknak. Ha csak az egyikük vág bele ebbe az üzletbe, akkor neki anyagilag még mindig nagyon megéri (ez lesz a második legjobb opció számára), de a másik háztulajdonosnak ez válik a legkevésbé kívánatos lehetőséggé, mivel a szomszéd bordélyház zavarná a háza lakásait bérbe vevő lakosokat, tehát alacsonyabb lakbért tudna tőlük beszedni. Ehhez képest a második legrosszabb alternatíva az lenne, ha mindketten

bérlőknek adnák ki a lakásokat, mert akkor a jól szituált környéken magasabb lakbért kérhetnének a bérlőiktől.

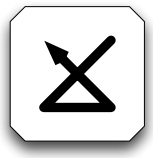
Robinson-Goforth könyv kevert harmóniajátéknak (Mixed Harmony) nevezi

Koordinációs játék: ügyfél



4,4	1,2
2,1	3,3

RPTS



a játék száma	
Rapoport-Guyer:	63
Brams:	-
Robinson-Goforth:	333

Ennél a játéknál (illetve a komplementer párjánál) valójában nem értelmes kooperálásról vagy dezertálásról beszélni, itt ugyanis pusztán arról van szó, hogy két cselekvési alternatíva közül érdemes ugyanazt választania mindkét játékosnak. A cselekvési lehetőségeknek abban az értelemben nincs értékvonzata, hogy az egyiket kooperálásnak, tehát kívánatosnak, a másikat dezertálásnak, tehát nem kívánatosnak tartanánk, mi, a játék megfigyelői. Mégis használom a kétféle minősítést, d ecsak azért, mert így a már megszokott módon tudok a cellákra, kimenetekre hivatkozni.

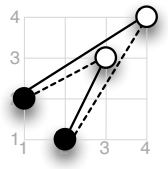
A játék két legmagasabb kifizetése a kétszeres kooperálással vagy kétszeres dezertálással érhető el. Ezért nevezik ezt koordinációs játéknak. Ennek a játéknak is két egyensúlyi helyzete van. A kétszeres kooperáláshoz szükséges kooperálás választása azzal is együtt jár, hogy ha mégsem jön össze a koordináció, akkor a legkisebb kifizetéshez lehet csak hozzájutni, míg ha a másik egyensúlyi állapot eléréséhez szükséges dezertálást választva nem jön össze a koordináció, akkor a második legrosszabb nyereményt lehet elérni. Az első egyensúlyi állapot elérése hatékonyabb, a másiké kockázatmentesebb. Az ilyen játékok nagy kérdése, hogy kockázatdominán/kockázatérzékeny vagy kifizetésdomináns/kifizetésérzékeny (risk dominant vagy payoff dominant) egyensúly alakul-e ki [737].

Bár a játék elterjedt megnevezése a koordinációs játék, én mégis az ügyféljáték nevet választottam. Az olyan fogyasztói vásárlások esetén, amikor a hálózati hatás érvényesül (minél többen választják az adott terméket, szolgáltatást, annál jobban nő a szolgáltatás önértéke) vagyis az ügyfélkör nagysága hozzájárul a jószág értékéhez. Ilyen esetekben koordinációs kérdéssé válik, hogy a versengő szolgáltatások közül melyiket választják a vásárlók, kiknek lesznek az ügyfeleik.

Példa lehet erre a szituációra, amikor két mobilszolgáltató (ATel és BTel) van a piacon [737]. Az egyik percdíja 10, a másiké 20, ATel felől BTel felé érkező hívások ára 70, fordítva 40 egység. Ezt a piaci feltételrendszert a tiszta koordinációs játék logikája szerint rendezik el a vásárlók.

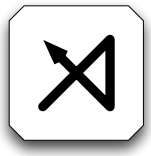
Koordinációs játék: balszélső

Ha egy focimeccsen a bal oldali középpályás szöktetni akarja a balszélsőt, akkor vagy a bal vagy a jobb szélre rúghatja a labdát, és a balszélső is ugyanúgy vagy bal vagy a



4,4	2,1
1,2	3,3

RPST

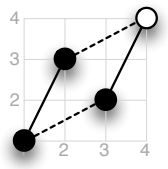


a játék száma	
Rapoport-Guyer:	60
Brams:	-
Robinson-Goforth:	344

jobb szél irányába kezdhet el rohanni. Ha mindketten ugyanazt az irányt “választják”, akkor mindkét esetben gólhelyzetbe kerülhet a csatár. Mivel balszélső ballábas, azért egy kicsivel jobb számára a bal oldali támadás a jobb oldalihoz képest. Ha nem sikerül az akció, akkor a balszélső számára kicsivel jobb az a helyzet, amikor a bal oldalra fut, mert bár nem kapja meg a labdát, de legalább ott van, ahol otthonosan érzi magát, míg a job oldalon idegenebbül mozog. A bal oldali középpályásnak az a legrosszabb opció, amikor jobbra küldi a labdát, de a szélső balra fut, mert mert ekkor neki kell a jobb oldalra a labda után futni és a neki “nem fekvő” területen kijavítania a hibáját.

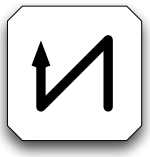
Ez a játék az ügyféljáték (koordinációs játék) kompleneter párja.

Harmóniajáték: csónakos



4,4	2,3
3,2	1,1

RTSP



a játék száma	
Rapoport-Guyer:	6
Brams:	-
Robinson-Goforth:	311

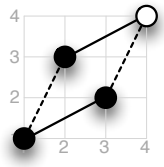
Említettem már, hogy a benne rejlő harmonikusság miatt kevés figyelmet szenteltek eddig ennek a játéknak. Adam Smith láthatatlan kéz elméletéhez ez illeszkedik a legjobban. Nem véletlenül nevezik Robinsonék harmóniajátéknak, és korábban – Olson nyomán – én is privilégiumjátékként azonosítottam, míg Skryms Prisoner’s Delight-ként hivatkozik rá [689]. Kicsit más nézőpontból tekintve a játékra Sam Perlo-Freeman az ‘Appeaser’ (békéltető, lekenyerező) nevet használja [566].

A csónakos játék meséje Brian Skyrmstől származik [689]. Két halász egy szigeten horgászott egész nap, és nap végén egy csónakkal indulnak haza úgy, hogy mindkettejüknek van egy-egy pár evezője, és mivel éhesek, minél gyorsabban szeretnének hazajutni. Ekkor mindketten a kölcsönös kooperációt tartják a legjobbnak, mert ekkor érnek haza leggyorsabban. A második alternatíva mindkettejük számára az, ha maguk nem eveznek, de a parnerük igen. mert akkor is hazaérnek, de nem kell erőt kifejteniük. A legrosszab kimenet számukra a kétszeres dezertálás lenne, hiszen ekkor sosem érnének haza.

A békéltető (appeaser) megnevezés azért találó, mert azt a beállítódást fejezi ki, amely arra törekszik, hogy a partner kooperációját (békés hozzáállását) minden áron fenn tudja tartani. Ez abban mutatkozik meg, hogy a legjobban preferált kölcsönös kooperálás (béke) után a második legfontosabb alternatívának a partner kooperálását tartják a felek. Így jellemezték a második világháború előtti angol diplomácia stratégiáját a németekkel szemben. Winston Churchill (aki ezt a politikát tagadta meg) azt mondta erről, hogy

“An appeaser is one who feeds a crocodile – hoping it will eat him last” [566]. Más megközelítésből tekintve a játékra, észrevehetjük, hogy egy jól ismert másik mondás, “Ha békét akarsz, készülj a háborúra!”, szintén ehhez a játékhoz illeszthető, csak ekkor a hangsúly eltolódik úgy, hogy a legjobban preferált kölcsönös kooperálás (béke) után a másodikra helyre az az alternatíva kerül, amikor a játékos már dezerál a partner kooperálásával szemben.

Harmóniajáték: pacifista



4,4	3,2
2,3	1,1

RSTP

a játék száma	
Rapoport-Guyer:	
Brams:	
Robinson-Goforth:	366

Ez a játék a csónakos játék komplementere, és a sajátosságát a második legjobb alternatívával kapcsolatos elmozduládban van. Azon túl, hogy mindkét fél a kölcsönös kooperálást tartja a legmagasabbra, a második helyre mindketten a dezertálással szembeni kooperálást teszik, vagyis a saját kooperálást preferálják a leginkább. Ez az igazi pacifista hozzáállás: a pacifista békét akar, még akkor is, amikor a partnere már dezertálna. A pacifista úgy akarja a békét, hogy maga mindent meg is tesz érte.

Sam Perlo-Freeman adta a játéknak a pacifista elnevezést [566], míg – más szerzőkkel ellentétben – Amir N. Licht ezt nevezi harmóniajátéknak [445].

A magjátékok áttekintése

A tizenkét magjáték között érdekes összefüggések vannak, amelyek még inkább észrevehetőek, ha valamilyen összesítő ábrában mutatjuk meg őket. Tanulságos áttekinteni egyszerre is az egyes játékoknál egyenként bemutatott vizuális reprezentációkat, jellemzéseket. Elsőként a játékok útjelző tábláit mutatom meg.
Az ábrán jól látszik, hogy háromféle mintázat van, a z-, az n- és az x-betűs. A függőleges tengelyen van négy játék, melyek középpontjában a fogolydilemma (potyautas) és a harmónia (csónakos) játék van (mellettük ott vannak még a tengelyen a komplementer párjaik). A fogolydilemmaból (és komplementer társából, a mártírjátékból) kaphatjuk meg a gyáva nyúl és a szarvasvadászat játékokat (illetve ezek komplementereit), majd még egy “lépés téve” jobbra , illetve balra eljuthatunk a kétféle nemek harca, illetve kétféle koordinációs játékhoz.

Ha a játékok topológiai sajátosságait gyűjtjük össze, akkor a tizenkét játékot az alábbi módon mutathatjuk be egy közös, áttekintő ábrában.
Az összes magjátékok egyszerre nézve még jobban szembeötlő lehet a tény, hogy a fogolydilemma (potyautas) játék adja, ígéri a legkisebb nyereményt a játékosok számára. Még a szarvasvadászat játékban fordulhat elő az, hogy a szuboptimális egyensúlyi pontot (2,2) választják a résztvevők, egyébként minden más esetben csak magasabb nyereményekkel lehet a játákokból kijönni.

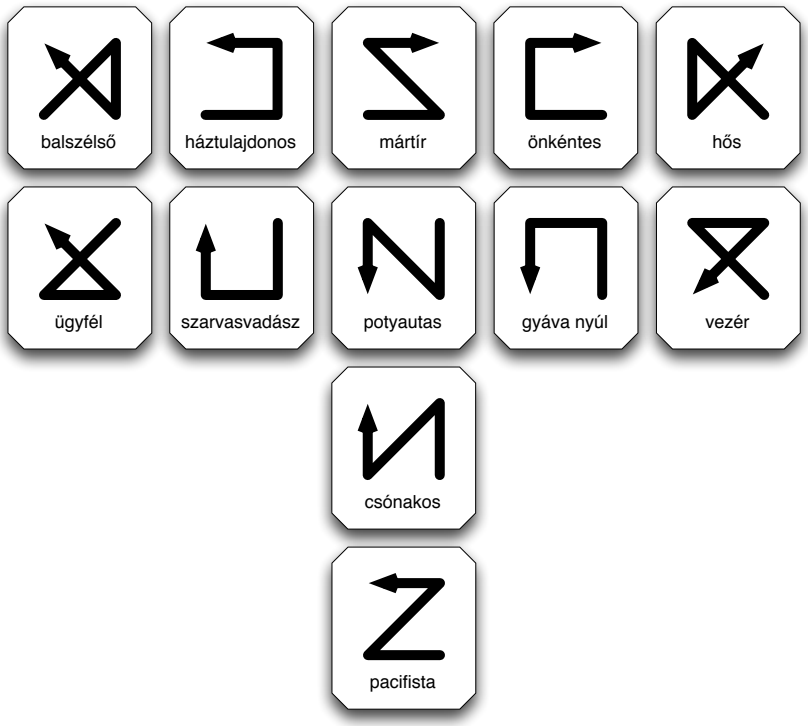


Figure 9.1: magjátékok útjelzői táblái összesítve

A fenti két ábra érdekes összefüggéseket mutat a magjátékok között, de mindkettő csak egyetlen vonását mutatja be a játékoknak. Érdemes lenne olyan ábrát is szerkeszteni, amely a lehető legtöbb információt képes egyszerre megmutatni. Ehhez segítségül lehet hívni azt a vizuális technikát, amelyet már évezredek óta használ az emberiség személyiségtípusok megjelenítésére. A zodiákus tizenkét állatövi jegyét, tizenkét személyiségtípusát sokszor jelenítik meg kördiagram segítségével.[4] Ezt a megoldást veszem át én is itt a tizenkét magjáték ábrázolására. Természetesen egy percig sem állítom, hogy bármilyen tartalmi kapcsolat lenne a kétfajta reprezentáció között, kizárólag a vizuális megjelenítés áttekinthetőséget és megérthetőséget könnyítő szerepe miatt veszem igénybe ezt a technikát. Az összes fontos vizuális információt, a preferenciamátrixokat, a játékok neveit, valamint a Robinson-Gofort könyvben használt azonosítószámot feltüntetve íme a magjátékok kördiagramja:

A magjátékok tipizálása

Ulmann-Margalit [762] fogolydilemma normák koordinációs normák partial/impartial normák (részrehajló, elfogultsági/elfogulatlansági normák)

Robinson-Goforth [608] technikai tipizálás

Wim B.G. Liebrand társadalmi dilemmák osztályozása [447] hármat különít el: fogolydilemma gyáva nyúl bizalmi (szarvasvadászat)

[4]Érdekességként jegyzem csak meg, hogy vannak még további 12-osztatú ábráink: a nap 12 órából, az év 12 hónapból áll.

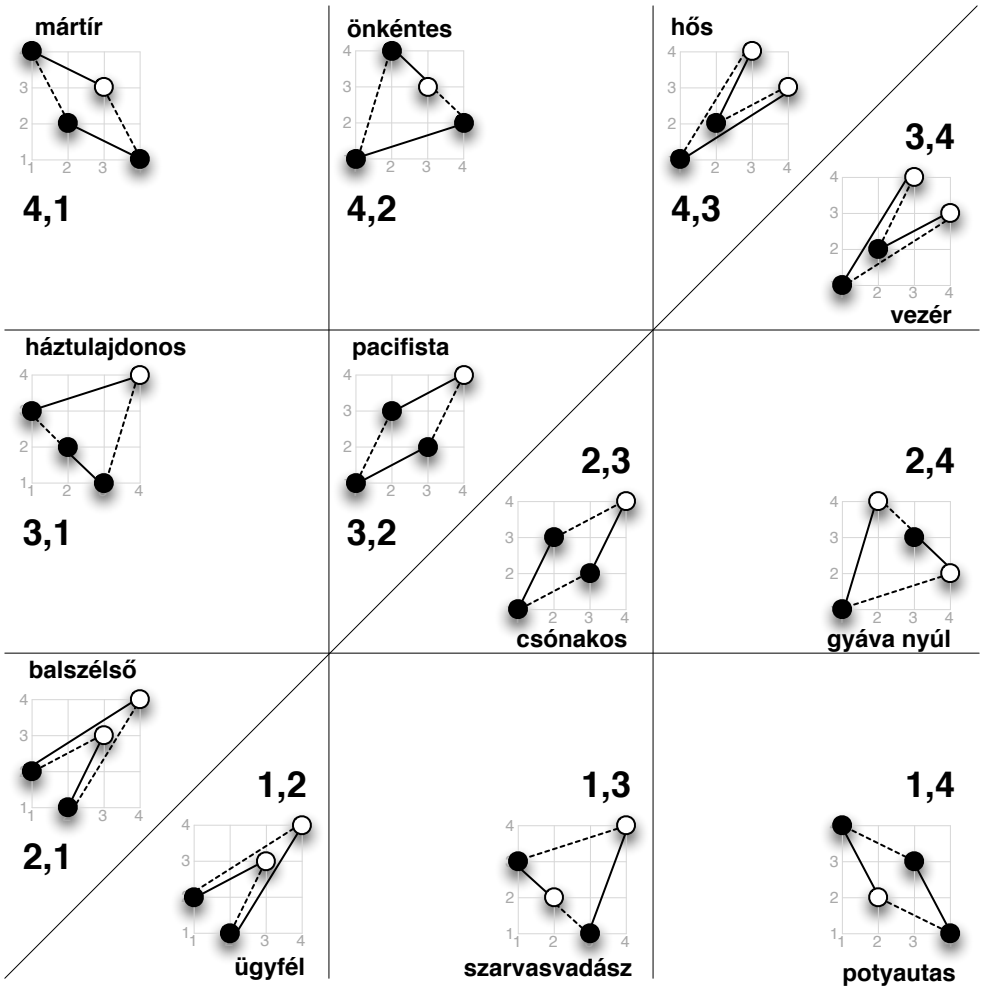


Figure 9.2: magjátékok geometriája – áttekintő ábra

9.3 A játékok statikus kiterjesztése

A magjátékok bemutatása után úgy léphetünk tovább a társadalmi cselekvések játékelméleti modellezésében, hogy fokozatosan "elengedjük" a kezdetben megfogalmazott feltételeinket. Ezáltal egyre összetettebb társadalommodellhez jutunk. A modell ilyenfajta kiterjesztését két nagyobb részben érdemes elvégezni. Először

9.3.1 Aszimmetrikus játékok

Az aszimmetrikus játékoknak aszimmetrikus preferenciamátrixuk van, minden másban megegyeznek a szimmetrikus játékokkal. A kétszemélyes, kétalternatívás játékok közül csak 12 szimmetrikus, az összes többi aszimmetrikus. Hosszú időn keresztül széles körben elfogadott volt Rapoport és Guyver nézetét, amely szerint 66 aszimmetrikus (és a tizenkét szimmetrikus játékkal együtt összesen 78) alapjátékot különíthetünk el egymástól. Ez a szám úgy jött ki, hogy Rapoporték ekvivalensnek minősítették azokat a játékokat, amelyek a két játékos felől nézve ugyanúgy "néznek ki". Ha arra gondolunk, hogy mindegy, hol is "áll" a két játékos, a Sor vagy az Oszlop játékos szerepében játszik-e, akkor

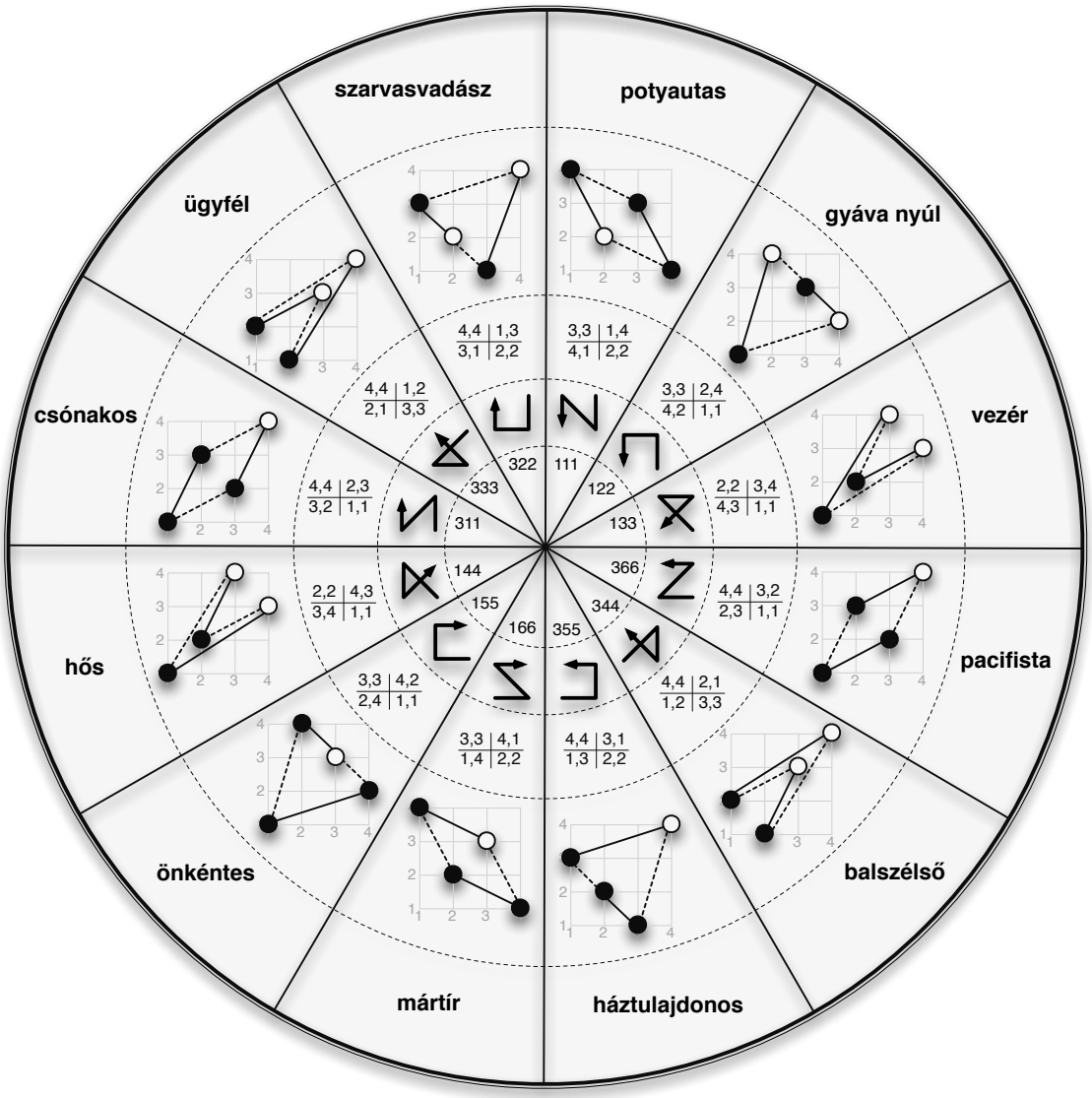


Figure 9.3: a tizenkét magjáték kördiagramja

ez az ekvivalenciatételezés elfogadható. Ha azonban Egoról és Alterről beszélünk, vagyis elkülönítjük egymástól azt a két szerepet, amelyek közül az egyiket (Egot) kitüntetett helyzetbe tesszük, hiszen mindig az ő cselekvését, döntéseit akarjuk megérteni az adott játék értelmezésekor, akkor már igenis “láthatunk” különbséget két olyan játék között, amelyek közül az egyikben Ego a Sor, a másikban Ego az Oszlop játékos szerepébe kerül. Egonak nem mindegy, hogy az aszimmetrikus helyzet melyik pólusán van. Ezt a szemléletváltást hozta el Robinson és Gofort 2005-ös könyve [608], és ezt fogadom el én is. Ebből következően az alapjátékok száma 144, az aszimmetrikus játékoké pedig 132.

A szimmetrikus játékok ismeretében már könnyebb dolgunk van az aszimmetrikus játékok modellezésekor. Ha ugyanis elválasztjuk egymástól, és külön tekintünk a két játékos, Ego és Alter preferenciarendezéseire, akkor ezek alapján mindkettőjükhöz hozzárendelhetünk egy-egy magjátékot. Mivel az aszimmetrikus játékoknak aszimmetrikus a preferenciamátrixuk, ezért a szétválasztás eltérő magjátékokat eredményez. Azt mondhatjuk tehát,

magjátékok	
aszimmetrikus játékok	aszimmetrikus
többalternatívás játékok	kétalternatívás
többszereplős játékok	kétszereplős
nem szigorúan rendezett játékok	szigorúan rendezett
aszinkron játékok	szinkron
ismételt játékok	egyszeri
direkt reciprocitás játékok	ismerős játékosok
indirekt reciprocitás játékok	ismeretlen játékosok
kooperatív (hangos) játékok	nem kooperatív (néma)

hogy az aszimmetrikus játékok lényege az, hogy a játékosok – ugyanazt a játékhelyzetet másként értékelve – eltérő magjátékokat játszanak egymással.

Példák aszimmetrikus játékokra

felvett kesztyű (called bluff) Falkland-szigeteki háború Argentína és Nagy-Britannia között
a hidegháború mint percepcionális dilemma szarvasvadászat és fogolydilemma (amiből a fegyverkezés, dezertálás következik)
pártfogó
projektfeladatból KalmarDaniel:2009

Az aszimmetrikus játékok osztályozása

Robinson és Grofort

9.3.2 Többalternatívás játékok

A kétalternatívás játékok abban az értelemben nagyon egyszerű társadalomképet feltételeznek, hogy ebben a világban csak "igen-nem" döntések hozhatnak a játékosok. Sok esetben elég ez a megoldás. De természetesen nem kevésszer kerülhetünk olyan helyzetbe, amelyekben a döntéseink nem ilyen egyszerű szerkezetűek, amikor több alternatíva közül kell választanunk. Gondoljunk csak a kő-papír-olló néven ismert, egyébként nagyon egyszerű játékra, amelyben három döntési lehetőség közül kell választaniuk a játékosoknak. Az ilyen helyzeteket a többalternatívás játékokkal modellezhetjük.

magjátékok	
aszimmetrikus játékok	szimmetrikus
többalternatívás játékok	többalternatívás
többszereplős játékok	kétszereplős
nem szigorúan rendezett játékok	szigorúan rendezett
aszinkron játékok	szinkron
ismételt játékok	egyszeri
direkt reciprocitás játékok	ismerős játékosok
indirekt reciprocitás játékok	ismeretlen játékosok
kooperatív (hangos) játékok	nem kooperatív (néma)

A többalternatívás játékok modellezését könnyen és röviden el lehet intézni. Más tudásterületeken gyakran alkalmazott technikát hasznosíthatunk itt újra. A több alternatíva

kezelését ugyanis könnyeb visszavezethetjük kétalternatívás döntések egymásba fűződő sorozatára. Csak annyit kell tennünk, hogy az összes döntésre váró alternatívát besoroljuk két csoport valamelyikébe, majd a két csoportot "mérkőztetjük meg" egymással. A győztes csoportba tartozó alternatívákat újra két csoportra osztjuk szét, a csoportokat újra összemérjük egymással és így tovább. Az alternatívák csoportba sorolást és a csoportok összehasonlítását addig folytatjuk, ameddig két alternatíva marad a végére, amelyek közül egy utolsó összehasonítással már végső győztes lehet kihozni.

9.3.3 Többszereplős játékok

A többalternatívás játékok modellezése nagyon könyen megoldható a kétalternatívás döntéssorozat technikájának alkalmazásával. Nem ez a helyzet a többszereplős játékok esetében. Sok minden megváltozik ugyanis akkor, amikor a társadalmat már úgy modellezzük, hogy két játékos helyett sok szereplő alkotja azt. Ezt a lépést meg kell tennünk, mert bár sokszor találhatjuk magunkat kétszemélyes játékszituációkban, de még többször kerülünk olyan helyzetekbe, amikor a helyzet minősítő sajátossága az, hogy sokan, nagyon sokan vagyunk benne.

magjátékok	
aszimmetrikus játékok	szimmetrikus
többalternatívás játékok	kétalternatívás
többszereplős játékok	többszereplős
nem szigorúan rendezett játékok	szigorúan rendezett
aszinkron játékok	szinkron
ismételt játékok	egyszeri
direkt reciprocitás játékok	ismerős játékosok
indirekt reciprocitás játékok	ismeretlen játékosok
kooperatív (hangos) játékok	nem kooperatív (néma)

A többszereplős játékok több szempontból is átalakítják azt a döntési-cselekvési teret, amelyben a játékosok egymást és a többiek lehetőségeit értékelik.

hivatkozni a könyvemre, Davis tanulmányára és itt csak a gondolatmenet vázlatát bemutatni.

közlegelők tragédiája név és problémafelvetés Garret Hardin: [339]

DE: az n-szereplős PD-t belőle Russel Hardin csinálta meg [340]

többszereplős önkéntes dilemma: Andreas Diekmann [205]

+szempont a megváltozott helyzetre: a többszereplős helyzetek gyakorlati szempontból időben szétcsúsznak, aszinkronná válnak, így az interdependencia egyidejűségi feltétele nem szorító a modellezésben is könnyít ez

Háromszereplős játékok

> „ . . . nekem nem az oroszlánnál kell gyorsabban futnom, hanem a nálad" (vicc)

többszemélyes altruizmus (multiparty altruism) az altruista rokonával szembeni viszonzás negatív formája a vendetta, amikor a negatív reciprocitás a rokonra irányul

koalícióelmélet

harmadrendű büntetés indirekt altruizmus spite

9.3.4 Nem szigorúan rendezett játékok

Elméleti szempontból érthetőnek és könnyen elfogadhatónak tűnik az az elvárás, amely szerint feltételezzük, hogy a játékosok egyértelműen sorba tudják rendezni a játék lehetséges kimeneteit, alternatíváit. A preferenciarendezésre vonatkozó szigorú rendezési feltétel azonban a tényleges gyakorlatban sok esetben túlzottan szigorú elvárásnak minősül. Nem tudunk mindig minden alternatívát egyértelműen sorba rendezni. Van, hogy nem tudunk valamely alternatívát megítélni és egyszerűen kihagyunk a rendezésből, de az is előfordul, hogy nem vagyunk képesek dönteni két vagy több alternatíva esetében, és ilyenkor "döntetlenre hozzuk ki" a sorrendet köztük. Az ilyen esetek modellezésére az olyan játékok alkalmasak, amelyekben nem követeljük meg a szigorú rendezés feltételének fennállását.

magjátékok	
aszimmetrikus játékok	szimmetrikus
többalternatívás játékok	kétalternatívás
többszereplős játékok	kétszereplős
nem szigorúan rendezett játékok	nem szigorúan rendezett
aszinkron játékok	szinkron
ismételt játékok	egyszeri
direkt reciprocitás játékok	ismerős játékosok
indirekt reciprocitás játékok	ismeretlen játékosok
kooperatív (hangos) játékok	nem kooperatív (néma)

Az ilyen játékok modellezésekor rögtön egy számíthatósági problémával kerülünk szembe, mert ugrásszerűen megnő a lehetséges rendezési mintázatok számossága. Hivatkozni a cselekvéselméleti könyvemre + Goforth-ék 2007-es cikke

9.4 A játékok magasabbrendű kiterjesztése: büntetés

A játékelmélet első szintjén a játékosok stratégiáit úgy értékelik, hogy a cselekvések (döntések) tartalmát absztrahálják, vagyis lecsupaszítják a kooperálás-dezertálás dichotómiájára. A kooperálás a jó, a dezertálás a rossz, és az a nagy kérdés, hogy mikor, hogyan, miért lehet elérni, remélni, hogy kooperáljanak a játékosok. Ez a séma persze nem pontos. Hiszen a koordinációs játékoknál láthattuk, hogy a kooperálás vagy dezertálás fogalma nem használhatók tartalmas módon. Az összes alapjátékra igaz viszont, hogy ha van valamilyen feloldandó társadalmi dilemma vagy koordinációs probléma, akkor mindig van olyan kimenet, amely jobb, kívánatosabb a többinél, és a kérdés úgy szól, hogy milyen eszközzökkel lehet biztosítani a társadalmilag kívánatosnak tartott kimenetek választását. Az egyik válasz a döntések szankcionálása. Ismételt vagy aszinkron játékok esetében, amikor az egyik játékos ismerheti a másik döntését, szankcionálhatja már ismert döntést, vagyis saját cselekedetét a másik játékos előző döntéséhez igazíthatja.

Olson vezette be a **szelektív ösztönző** (selective incentive) fogalmát, de tovább már nem pontosította.

a fogalom elemzése Oliver 1980 [551] citeoliver:jcr:1980

pozitív és negatív ösztönzők: **jutalom** (reward) és **büntetés** (punishment)

a hétköznapi kicsit szóhasználat számára kicsit furcsa negatív ösztönzésről beszélni.

Fogalmilag ugyanezt fejezi ki a szankció fogalma is, és ezt a fogalmat is értelmezhetjük negatív és pozitív értelemben. Ahogy idegenül hangzik a negatív ösztönzés kifejezés, ugyanúgy szokatlan lehet sokak számára a pozitív szankció terminusa.

Peters, Richard Stanley (1966). Ethics and Education pp. 267?268 Inflicting something negative, or unpleasant, on a person or animal, without authority is considered either spite or revenge rather than punishment

9.4.1 Másodrendű játékok – költséges büntetés

a szankció etimológiája

A szankcionálásnak két típusa lehetséges: pozitív és negatív. A negatív szankció vagyis a büntetés a rossz, nem kívánatos cselekvésre adott negatív tartalmú válasz. A pozitív szankció ennek ellentéte, vagyis a kívánatos cselekvésre adott pozitív tartalmú válasz.

másodrendű fogolydilemma játék [551] [559]

büntetések másfajta tipizálása:

2PP (second-party punishment) másod büntetés spite büntetés rosszindulatú ?? kárt okozó ??

3PP (third-party punishment) [241] altruista büntetés harmadik

megkülönböztetik egymástól a csoporton belüli (within-group) és a csoportok közötti (inter-group) altruista büntetést

az evolúcionalista magyarázat szerint a csoportok közti versengésben alakulhat ki és szilárdulhat meg az altruizmus jelensége,

parokiális altruizmus (parochial altruism) is megjelenik a közösségen belül [71]

Human punishment is motivated by inequity aversion, not a desire for reciprocity [589]

az altruista viselkedés, beállítódás neurológiai alapjai [198] [511]

Ernst Fehr és Urs Fischbacher (2002) [242]megkülönbözet

- stratégiai büntetés: Fehrék feltevése szerint a stratégiai büntetés hátterében nem állnak olyan erős érzelmi motiváló tényezők, mint a nem-stratégiai büntetésnél: itt a cél pusztán a haszonszerzés a többi játékos nyereségeinek csökkentése útján. A szerzők elismerik ugyan a stratégiai büntetés létezését, de véleményük szerint nincs rá bizonyíték, hogy az szerepet játszana akár a laboratóriumi vizsgálatokban, akár a valós élethelyzetekben
- nem-stratégiai büntetés (A nem-stratégiai büntetés hátterében motiváló tényezőként a normatív agresszió jelenik meg, és a cél az együttmködési normák betartatása, a fair viselkedés kikényszerítése: ez a büntetési típus jelenik meg az erős negatív reciprocitás elvében.)

(Fehr, Fischbacher, 2002, p. 21).

Más tanulmányok ezt a szankcionálási formát rosszindulatú (spiteful) vagy irigységen alapuló (envious) büntetésnek nevezik: ezt a haszonleső játékosok alkalmazzák azokkal a résztvevőkkel szemben, akik hozzájuk képest sikeresebbek, akik fenyegetik az érdekeiket (Falk et al., 2005; Fehr et al., 2008).

Falk, A., Fehr, E., Fischbacher, U. (2005). Driving forces behind informal sanctions. Econometrica, 73 (6), 2017-2030.

Fehr, E., Hoff, K., Kshetramade, M. (2008). Spite and development. American Economic Review ? Papers and Proceedings, 98 (2), 494-499.

Szemben az altruista büntetéssel, a hagyományos Közjavak-játékban ez a büntetési forma csak akkor jelenik meg, ha a büntet költsége alacsony a kiosztott büntetés mértékéhez képest: tehát ahol kevés veszteség árán lehet másoknak igen nagy kárt okozni (Falk et al., 2005). Az irodalomban megjelenik ezen kívül az ún. antiszociális (vagy perverz) büntetés is: ez kimondottan azokra a résztvevőkre irányul, akik sokat adnak a közösbe, ugyanakkor nem nevezhető előnyösnek a büntetőre nézve sem (Gaechter, Hermann, 2009; Hermann et al., 2008; Nicklisch, Wolff, 2010).

Gaechter, S., Herrmann, B. (2009). Reciprocity, culture, and human cooperation: Previous insights and a new cross-cultural experiment. Philosophical Transactions of the Royal Society B ? Biological Sciences, 364, 791-806.

Herrmann, B., Thöni, C., Gaechter, S. (2008). Antisocial punishment across societies. Science, 319, 1362-1367.

Nicklisch, A., Wolff, I. (2009). Cooperation norms in multiple-stage punishment. Working Paper Series of the Max Planck Institute for Research on Collective Goods, 2009, 40, 1-35.

nagyon fontos látni, hogy a szankció másodrendű cselekvés. Szankcionálást, büntetést, jutalmazást önmagában nem tudunk végrehajtani, fogalmilag mindig szükségünk van rámutatni arra az – elsőrendű – cselekvésre, amely kiváltotta a szankcionálás igényét, és amelyre hivatkozva igazoltnak érezhetik a szankcionálók saját cselekvésüket.

a negatív szankció olyan cselekvés

[300] [236]

jutalmazás

utánanézni ennek:

„Létezik egy hihetetlenül bölcs afrikai törzsi szokás. Amikor a törzsből valaki rosszat tesz a törzs egy másik tagjával, vagy fájdalmat okoz neki, akkor az illetőt kiállítják a falu közepére, és az egész törzs köré gyűlik. De ahelyett, hogy valami nyilvános büntetésben részesítenék, két napig folyamatosan sorolják neki mindazokat a jó dolgokat, amiket valaha is tett. Hogy miért?

Mert a törzs hisz abban, hogy minden egyes ember jónak születik, és szeretetre, békére és boldogságra vágyik. Csak néha, a mindezekre való törekvés közben hibákat vétünk. És mivel a rossz cselekedeteket segélykiáltásként értelmezik, a törzs összegyűlik a társukért, és segítenek neki újra megtalálni önmagát, a jó természetét. Emlékeztetik őt arra, hogy ki is valójában. És erre néha mindenkinek szüksége van!”

de a gyereknevelésből elég egyértelmű, hogy a jutalom erősebb szankcionálási eszköz, mint a büntetés.

9.4.2 Szankcióparadoxon

Ha már látjuk és elfogadjuk, hogy a kívánatos cselekvési mintákat azzal (is) remélhetjük általánossá tenni és fenntartani a társadalomban, hogy vannak, akik magukra vállalják a dezertálók büntetését. A büntetésnek azonban sokfajta módja lehet, amelyeket érdemes alaposabban megvizsgálni [723]. A büntetés során olyan erkölcsi kérdések merülnek fel,

melyekre különböző válaszok adhatók – a cselekvők, értékelők eltérő értékválasztásaiból, eltérő kultúrájából adódóan. Ezek pedig nyilván befolyásolják az egyes stratégiák tartalmát. A büntetés "célja" általában a dezertáló, a nem-segítőkész magatartás szankcionálása. Igen ám, de itt a büntetés maga a segítésnyújtás megtagadása vagy testi, lelki erőszak alkalmazása, ami önmagában véve rossznak, dezertálásnak minősülne. Ezt az "ellentmondást" többféleképpen lehet értékelni. Ha mindenekfölött álló értéknek tartjuk a segítést, a kooperálást, akkor nem szabad büntetni senkit, mert az nem kooperatív hozzáállás. Ezzel az elvvel azt lehet szembeállítani, hogy ha nem büntetünk, akkor ezzel "jóváhagyjuk" mások dezertálását, helytelen magatartását, ami a kooperáció fenntartása, a közösség egésze szempontjából nem előnyös. Ha erre az érvre hivatkozva mégiscsak büntetünk, akkor árnyalni kell a büntetést végrehajtó személy megítélését, hiszen ebből a szempontból mondhatjuk, hogy a büntető a közösség érdekét követve büntet. Ha így tesz valaki, akkor a büntetés "jogos, igazolt", amit szembe lehet állítani azzal a helyzettel, amikor valaki "jogtalan, igazolatlan" módon büntet.

Ez pedig megkívánja, hogy beemeljük a modellünkbe a diszkrimináció jelenségét (és fogalmát), hiszen különböző irányban, különböző megítélésű szereplőkkel szemben másként kell alkalmazni a büntetés eszközét (a diszkrimináció fogalmának itt nincs negatív felhangja, egyszerűen csak annyit lehet vele kifejezni, hogy a segítségnyújtás, illetve annak megtagadása vagyis a büntetés melletti döntés szempontrendszere nem homogén).

Ha táblázatba foglaljuk, hogy a cselekvések kétféle rendjében, kétféle szintjén milyen döntések elé kerülhetnek a cselekvők, akkor az alábbi cselekvési lehetőségeket különíthetjük el egymástól:

	másodrendben kooperál, büntet	másodrendben dezertál, nem büntet
első rendben kooperál	kooperál és büntet	kooperál, de nem büntet
első rendben dezertál	nem kooperál, de büntet	nem kooperál és nem is büntet

Az elsőrendű cselekvések szintjén megvalósuló altruista/kooperatív magatartást segítségnyújtásnak, a nem segítést pedig egoista/dezertáló magatartásnak nevezhetjük. A büntetést viszont a segítségnyújtás megtagadásaként lehet értelmezni. Akkor érthetjük meg, hogy miért fontosak ezek a cserék, ha az új terminusokat beleírjuk a fenti táblázatba:

	másodrendben kooperál (büntet = nem segít)	másodrendben dezertál (nem büntet = segít)
első rendben kooperál (segít)	elsőrendben segít, másodrendben nem segít	elsőrendben segít, másodrendben segít
első rendben dezertál (nem segít = büntetés	elsőrendben nem segít, másodrendben nem segít	elsőrendben nem segít, másodrendben segít

Mivel két esetben egymásnak ellentmondó állítások állnak egymással szemben, ami nyilván nem követhető döntési/cselekvési alternatíva, csak két választás lehetséges: első rendben dezertálni és másodrendben kooperálni (azaz nem segíteni), vagy fordítva (azaz segíteni). Ha tehát az értelmezési keretrendszerbe beemeljük a büntetést mint másodrendű cselekvést, akkor ami első rendben önzetlenség, az másodrendben önzés, illetve fordítva. Mindez azért különösen fontos, mert ezt a két opciót ellentétes módon lehet értékelni még akkor is, ha egyértelműen az altruizmust, a kooperációt preferáljuk az egoizmussal, a dezertálással szemben.

Látható tehát, hogy a szankcionálás jelensége paradoxonhoz vezet. Ugyanazt a cselekvést elsőrendben kooperálásnak (jónak), másodrendben dezertálásnak (rossznak), vagy fordítva elsőrendben dezertálásnak, másodrendben kooperálásnak lehet minősíteni. Vagy-vagy. Értékválasztás kérdése. Az indirekt reciprocitás jelenségének tárgyalásakor pedig bemutatom majd, hogy milyen – egymásnak ellentmondó – válaszokat termeltek ki erre a különböző kultúrák az emberiség történelmében

9.4.3 Aszinkron játékok

A magjátékok kevésbé reflektált, kevesebbet elemzett feltétele a játékok szinkronitásának elvárása. Pedig vannak olyan esetek, amelyekben nem tartható fent ez a feltétel. A szimultán játszott, szinkron jellegű játékokkal szembe állíthatóak azok a helyzetek, amelyekben az aszinkron (szekvenciális, alternáló) döntéshozatal a jellemző. Ekkor a játékosok nem egyszerre, szimultán módon hoznak döntést, hanem egymás után, így a másik döntésének ismeretében választanak az alternatíváik közül.

magjátékok	
aszimmetrikus játékok	szimmetrikus
többalternatívás játékok	kétalternatívás
többszereplős játékok	kétszereplős
nem szigorúan rendezett játékok	szigorúan rendezett
aszinkron játékok	aszinkron
ismételt játékok	ismételt
direkt reciprocitás játékok	ismerős játékosok
indirekt reciprocitás játékok	ismeretlen játékosok
kooperatív (hangos) játékok	nem kooperatív (néma)

Azért is érdekesek ezek a játékok, mert aszinkron döntési helyzetekben a büntetés jelensége is értelmezhetővé, lehetővé válik. A játékosok időben egymástól elváló döntési helyzetekbe kerülnek, és a döntéseikben figyelembe vehetik a partnerük már ismert döntéseit.

Az állatvilágból származó reciprok altruizmus egyik leghíresebb példájaként említik a vámpírdenevéreket, akik között a sikeres táplálékszerző körútról visszaérkező egyedek megosztják beteg vagy hoppon maradt társaikkal az élelmet, míg egy másik alkalommal ezek a szerepek cserélődnek [817]. Ebben a példában a reciprocitás időben eltolt módon jelentkezik, amit az aszinkron játékokkal lehet modellezni. Kétfajta modellt is lehet állítani az aszinkron játékokra [536]. Az alternálás, a szerepváltás lehet szigorú vagy laza. A szigorúan alternáló játékok során pontos rendben váltják egymást a játékosok, míg a laza változatban véletlenszerűen dől el, ki milyen szerepet kap egy adott fordulóban. Utóbbi sok eset életszerűbb, hiszen a reciprocitáson alapuló kapcsolatokban sokszor nem lehet biztosítani, hogy a segítségnyújtó és támogatott (donor és recipiens) szerepek pontos rendben kövessék egymást.

Ahogy a fenti bekezdésben írottakból látszik, az aszinkronitás vagyis az időben eltolt döntési helyzetek könnyen összekeverhetők, gyakran össze is csúsznak az ismétlődő játékhelyzetekkel, a reciprocitás jelenségével. Mivel a reciprocitás jelenségével később bővebben foglalkozom, ezért inkább ott mutatom be azt a stratégiát, ami aszinkron játékokban bizonyult eredményesnek (↬358.o.).

9.4.4 Alkujátékok

Az aszinkronitást azonban nem csak az ismételt játékok esetén lehet feltételeznünk. A kísérleti közgazdaságtan, viselkedés-gazdaságtan vizsgál olyan helyzeteket is, amelyek csak egyszer játszott játékoknak felelnek meg, ám a játékosok nem egyszerre, hanem egymás után döntenek, így a későbbi döntéshozók már ismerik a partnereik választásait. Ezeket is aszinkron helyzeteknek minősíthetjük.

Az ilyen játékokat a kísérleti közgazdaságtan régóta és elég nagy intenzitással vizsgálja, igaz, nem elsősorban a helyzetek aszinkronitása miatt, hanem inkább azért, mert ezekben a játékokban jól lehet vizsgálni a kísérleti alanyok beállítódásait, értékelköteleződéseit, normatív struktúráit. Ezért tárgyalhatnám máshol is ezeket a játékokat,[5] de mivel vannak érvek amellett, hogy itt szerepeljenek, ezért itt mutatom be őket. Problémásnak érzem a játékok általános megnevezését is. A szakirodalomban gyakran **alkujátékoknak** (bargaining games-nek) nevezik az ilyen játékokat, ami több szempontból is vállalható, megmagyarázható, de az **alkuelmélet** kifejezéssel párba állítva már kicsit félrevezetőnek, könnyen félreérthetőnek tartom az alkujátékok terminusát. Az 'alku' fogalmának kétszeres használatával "sugallt" szemantikai közelség itt félrevezető, mert amíg az alkuelmélettel egy folyamatot, egy procedurát írunk le, addig az alkujátékok az egyének beállítódásainak egyszeri kinyilvánításait "mérik". Mivel azonban az alkufolyamat mélyén is ott lapulnak az alkudozó egyének beállítódásai, értékelkötelződései, ezért – fenntartásokkal ugyan, de csak – elfogadhatónak tartom az alkujátékok megnevezés használatát (ezért szerepel már a fejezet címében is). Az is amellett szól, hogy az alkujátékokat ebben a fejezetben tárgyaljam, hogy ezek aszinkron játékok. A játékosok itt egyszemélyes játékokat játszanak szekvenciális módon:

> „ . . . in ultimatum bargaining games strategic interaction occurs only in the form of anticipating future decisions.There is no mutual interdependence resulting from simultaneous moves and infinite plays." [315]

Ezekben a játékokban a méltányosság (fairness, equity) jelenlétét, mértékét, százalékokban is kifejezhető erősségét keresik. Azt vizsgálják, hogy milyen mértékű a másik személy iránti reciprok beállítódás mértéke, azaz mennyire hajlandó a kísérleti személy kooperálni a másikkal (lemondani saját előnyei egy részéről a másik javára), illetve hajlandó-e büntetni a másikat, ha az nem kooperál, még akkor is, ha a büntetés saját magának költségbe kerül. Mivel a személyiségben rejtőzködő reciprocitási hajlamot vizsgálják, gyakran nevezik 'erős reciprocitás' (strong reciprocity) elméletnek ezt a vizgálati irányt.

Egyszeri és ismételt játékok egyaránt előfordulnak, az a közös bennük, hogy empirikus kísérleteken, vizsgálatokon alapul az egész. Éppen ezért ezek az empirikus vizsgálatok nagyon szépen kiegészítik a formális játékelméleti modellek, gondolatkísérletek eredményeit. Nagyon sok kultúrában, szubkultúrában elvégezték, megismételték már a leghíresebb kísérleteket, és nagyjából hasonló eredményeket kaptak – akárhol is, akármilyen kultúrában is gyűjtötték az adatokat. Természetesen a különböző kultúrákban, szubkultúrákban felvett adatok között volt különbség, ami fontos vizsgáltai kérdéseket vet, vetett fel, de az erős reciprocitás elmélet körébe tartozó kísérletek talán legfontosabb tanulsága az az általánosíthatónak tűnő állítás, miszerint minden kultúra valamilyen mértékben beleteszi, beleneveli a tagjaiba a reciprocitásérzet, a méltányosságérzet valamilyen

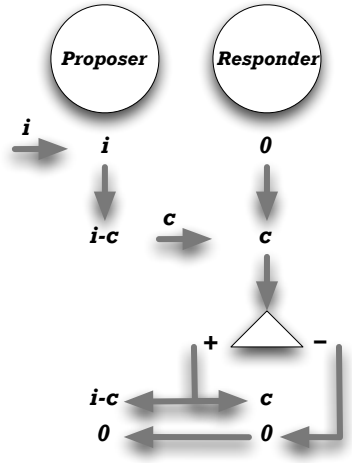
[5] Például a társas preferenciák elmélete környékén (↬295.o.)

fokát. Ez a tétel pedid megintcsak azt a közgazdaságtan széles körben elterjedt vélekedést kezdi ki, amely szerint az ember – alapértelmezés szerint – önző beállítódással lép másokkal kapcsolatba.

Első körben bemutatom az ultimátumjátékot, a diktátorjátékot, a nyilvános javak játékát és a bizalmi játékot, majd utána röviden kitérek még néhány más játékra is.

Ultimátumjáték

Az alkujátékok talán legismertebb és egyben paradigmatikus példája az ultimátumjáték (Ultimatum Game – UG).



A játék lényege a következő.[6] Adott összegből (i) az egyik játékos, az ajánlattevő (Proposer, Allocator) javaslatot tesz a másik játékos, a kedvezményezett (Responder) számára, hogy milyen arányban osztaná fel kettejük között a pénzt (i-c, illetve c). A második lépésben a másik játékos csak annyit tehet, hogy az ajánlatot elfogadja vagy visszautasítja. Amennyiben elfogadja, akkor mindketten megkapják az ajálattevő által javasolt összeget, visszautasítás esetén egyikük sem kap semmit.

A játék azért keltett (és kelt a mai napig) meglepetést sokakban, mert tisztán önérdekkövető játékosokat feltételezve már a legkisebb ajánlatot is el kellene fogadniuk a kezdeményezetteknek, ám a kísérletek tanulságai szerint ennél jóval magasabb ajánlatot fogadnak csak el átlagosan a játékosok. Az előzetes "elméleti" megfontolások alapján ugyanis azt lehet feltételezni, hogy a racionális (és önző) beállítódású kedvezményezettek úgy gondolkozhatnak, hogy a legkisebb egységnyi ajánlat is több, mint a semmi (amit a visszautasítás esetén kapnának), ezért érdemes elfogadni bármilyen ajánlatot, ami a nullánál több. Ezt tudva az ajánlattevők is gondolkozhatnak hasonlóképpen és nyugodtan adhatják a legkisebb ajánlatot. A kísérletek azonban egészen más eredményeket hoztak, hoznak. Az eddigi felmérésekben – kultúráktól függően – 20-50 százalék között mozgott mindenhol az átlagos ajánlattétel nagysága.

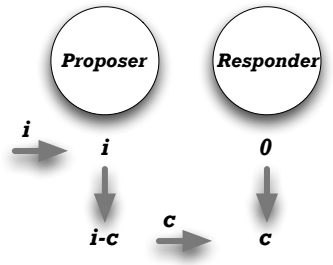
A kísérletek tanulsága szerint ezeket a helyzeteket nem lehet leírni a racionalitás és önző ember hipotézisére támaszkodva. Valami más magyarázó tényezőt kellett tehát találni. Ezen a ponton lehetett felvenni az ember, a társadalom modelljébe a méltányosságérzet fogalmát. Az emberben levő méltányosságérzettel tudjuk csak megmagyarázni az a büntetési szándékot, ami a visszautasítások tényében érhető tetten. A kedvezményezettek ugyanis saját anyagi érdekeikkel szemben vállalják a visszautasítást, ami így büntetésként fogható fel, hiszen így a méltánytalan ajánlata miatt az ajánlattevő sem kap semmit a játékban. A kedvezményezettben létezik valamiféle méltányosságérzet, ami meghatároz egy elfogadható ajánlati százalékot, és ha a konkrét ajánlat ennél az összegnél kisebb, akkor a kedvezményezett ezt a partner önzésének tudja be, és a méltánytalan ajánlatot visszautasítással bünteti. Az ajánlattevők pedig ezt a helyzetet racionálisan mérlegelve részben a büntetéstől (és ezáltal a teljes veszteségtől) való félelmükben, részben a bennük

[6] Az ultimátumjátékról (és a többi alkujátékról) kiváló áttekintést ad [351], magyar nyelven a problémakört bemutatja [10], [312].

is létező méltányosságérzettől vezéreltetve elmozdulnak a “fifty-fifty” elv irányába.

Diktátorjáték

A diktátorjáték (Dictator Game – DG) neve nagyon félrevezető, szerencsétlen, mert a játéknak nincs semmi köze a diktátorsághoz, a diktátori szerephez. Nyilván csak azt akarták kifejezni ezzel a névvel, hogy ebben a játékban egyszemélyi döntésről van szó, a másiknak nincs beleszólási lehetősége a döntésbe.

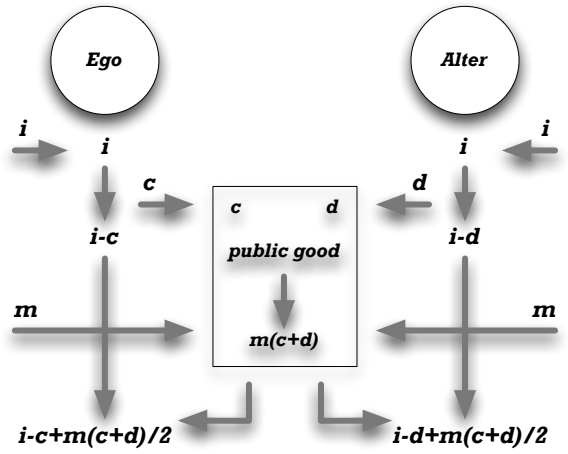


A diktátorjáték hasonlít az ultimátumjátékhoz. A különbség csak annyi közöttük, hogy a kedvezményezettnek nincs mérlegelési joga a második lépésben. Pontosabban nincs második lépés, vagyis az ajánlattevő ajánlata a játék végeredménye, a javak végleges megosztása is egyben. A kedvezményezetnek semmilyen formában nincs módja befolyásolni a végeredményt. Mivel ezt tudja az ajánlattevő is, ezért arra lehetne számítani az ilyen kísérletek előtt, hogy itt már igenis érvényesülhet az önző racionális ember hipotézise, és az ajánlattevők a minimális ajánlat környékén teszik majd meg ajánlataikat.

Az eredmények azonban mégsem ezt mutatják. Bár már kisebb mértékben, mint az ultimátumjátékban, de még mindig elég magas arányban ajánlanak fel összeget a kedvezményezett számára. Még akkor is (vagy inkább: annak ellenére is) így van ez, hogy ezekben a helyzetekben már azt sem lehet feltételezni, hogy a "büntetéstől" való félelem miatt lennének nagylelkűek az ajánlattevők. Ezek a kísérletek egyértelműen az ajánlattevő altruista beállítódását mérik (és mutatják egyben).

Nyilvános javak játék

A nyilvános javak játék (Public Goods Game – PGG) felfogható az előzőek sokszereplős kiterjesztéseként is (nevezik Simultaneous-Decision Prisoners’ Dilemma játéknak is).



Van x (az egyszerűség és összehasonlíthatóság kedvéért legyen megint csak 2) játékos (Ego és Alter), mindegyiknek van valamennyi pénze (i), és mindenki beadhat a közösbe valamennyi pénzt (legyen ez c és d). Az egyéni döntések után a közös kasszába tett teljes összeget megtöbbszörözik (m), és a közös javakból egyenlő arányban visszakapják a játékosok a saját részüket.

A kísérleti eredmények szerint kezdetben nagy az altruista hajlam a játékosokban, de idővel (vagyis a játékok számának növekedésével párhuzamosan) ez a kooperatív hajlandóság csökken. A kezdetben mért 40-50%-os felajánlások leesnek nulla közeli értékekre. Az is kiderült azonban ezekből a felmérésekből, hogy

ha bevezetik a büntetés lehetőségét (amikor a játékosok saját pénzükből büntethetik a potyautasokat vagyis azokat a társaikat, akik keveset ajánlottak fel a közösbe, akkor a csoport egészén belül jelentősen megugrik a kooperálási hajlandóság. Erre a játékra is igaz, hogy a világ minden pontján, sokféle kultúrán belül elvégezték a kísérleteket, és mindenhol nagyjából hasonló eredményeket kaptak.

Ez a játék sokban hasonlít a fogolydillema játékosztályba tartozó 100 dolláros játékhoz (bár nem azonos vele), amelynek a döntési folyamatábrája a következő:

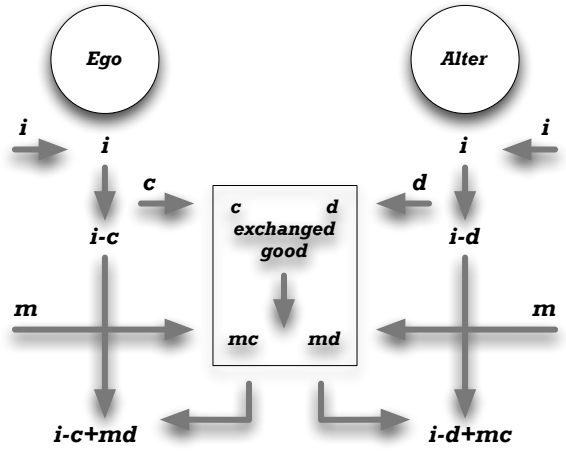
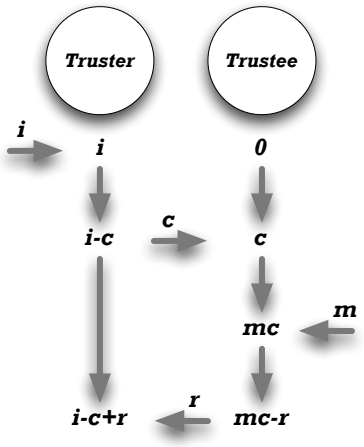


Figure 9.4: 100 dolláros játék

Bizalomjáték

A bizalomjáték (Trust Game – TG, Investment Game – IG) név alatt más játékot értenek az alkujátékok, az erős reciprocitás elméletében, mint a kétszemélyes magjátékok körében, ahol az egyik magjátékot is ezzel a névvel illetik.



Ebben a játékban is két játékos tesz ajánlatokat egymásnak. Az ajánlattevő lépése előtt a kedvezményezett (Truster) kap egy összeget (i), amiből lehetősége van egy részösszeget (c) "visszaajánlani" a partnerének (Trustee). Ezt az előajánlott összeget megtöbbszörözik (m), és a megnövelt összegű pénzt az ajánlattevőnek meg kell osztani kettejük között, amiből a kedvezményezett visszakap valamit (r).

A játék konstrukciója olyan, hogy leginkább a viszonzási hajlandóságot, a negatív reciprocitást lehet mérni benne, hiszen az ajánlattevő (akinek a viselkedését vizsgálják) az őt ért impulzus (az előzetes ajánlat) hatására dönt (legalább részben). Ebben a kísérleti játékban is az előző játékokhoz hasonlóakat tapasztaltak, vagyis a mérési eredmények szerint az ajánlattevő személyekben – kisebb, nagyobb mértékben – jelen van az altruizmus.

További játékok

Selten áruházlánc paradoxona (chain store paradox) [656]

called bluff játék megoldásához is kapcsolni lehet elrettentést néha beváltja valaki, mert további érvek vannak a dezertálás, háború mellett, Falkland háború

simonovits könyve [684, 33]

9.5 A játékok dinamikus kiterjesztése: társadalmi emlékezet

Ha elengedjük az egyszer játszott játékok feltételét, akkor ismételt játékokhoz jutunk. Ekkor ugyanazt a játékot többször lejátszák egymással a játékosok, vagyis játék helyett játéksorozatot kell vizsgálni. Az egyes játékokat fordulóknak is nevezhetjük, és szükségünk lesz a stratégia fogalmának átértelmezésére is.

Az ismételt játékok bevezetésével a társadalmiság modellezése szempontjából nagyon fontos lépést tehetünk meg. Talán a legjelentősebb változást az jelenti, hogy a modellben be kell emelni a társadalmi emlékezet fogalmát, és ezáltal meg lehet ragadni a múlt és a jövő döntésbefolyásoló szerepét. Amíg csak egyszer játszott játékokról beszélünk, addig nincs idő a modellben, a szereplők csak a "pillanatban élnek", döntéseikben nem befolyásolja őket sem a múlt ismerete, sem a "a jövő árnyéka". Ezt a nyilvánvaló modellezési hiányosságot korrigálja az a kiegészítés, hogy az ismételt játékokban a döntéshozatal során már figyelembe lehet venni mind a múlt megtörtént, mind a jövő várható eseményeit, és ehhez lehet igazítani a döntési stratégiákat.

magjátékok	
aszimmetrikus játékok	szimmetrikus
többalternatívás játékok	kétalternatívás
többszereplős játékok	kétszereplős
nem szigorúan rendezett játékok	szigorúan rendezett
aszinkron játékok	szinkron
ismételt játékok	ismételt
direkt reciprocitás játékok	ismerős játékosok
indirekt reciprocitás játékok	ismeretlen játékosok
kooperatív (hangos) játékok	nem kooperatív (néma)

Azáltal, hogy a játékból játéksorozat lesz, a stratégia fogalma módosul. Az egyszer játszott játékokban a statégia fogalma a döntés (vagy a lépés, a cselekvés) fogalmával volt azonos (de legalábbis összemosható). Az ismételt játékokban viszont stratégia fogalma alatt azt a döntési szabályt értjük, amely szerint a játékos dönt az egyes játékokban. A stratégia tehát egy döntéssorozatot meghatározó szabály. A döntés viszont cselekvésként is értelmezhető, így a stratégiát felfoghatjuk cselekvési szabályként is egyben.

Egy stratégia lehet **feltételes szabály** és lehet **feltétel nélküli szabály**. Feltételességen itt azt kell értenünk, hogy valamely játékban (tehát a döntéssorozat valamely pontján) a döntés során figyelembe veszik-e más – nyilván előző – játékok eredményeit, pontosabban: a játékoknak a játékosok cselekvéseiben megnyilvánuló eredményeit. Ha igen, akkor beszélhetünk feltételes stratégiáról, ha nem, az jelenti a feltétel nélküli stratégiát. Feltétel

nélküli stratégia például az, amikor valaki mindig kooperál vagy mindig dezertál mindentől függetlenül, de ebben az értelemben feltétel nélküli az a stratégia is, amely minden fordulóban véletlenszerűen dönt, mert nem a játékosok korábbi döntésétől, azaz más játékok (más döntések) eredményétől függ az aktuális döntés maga, hanem csak a véletlentől. A feltételességet itt az ágensek cselekvéseihez, illetve az azokat meghatározó döntésekhez, szándékokhoz kötjük.

Ha megvizsgáljuk a stratégiák, szabályok logikai szerkezetét, észrevehetjük, hogy a feltételes stratégia másodrendű fogalom, hiszen az aktuális döntés változójában figyelembe kell venni egy korábbi döntés eredményét is. A startégia itt valamilyen cselekvési szabályt jelent (DO(*Agent*)). A stratégiák jellemzésekor Ego cselekvését írjuk le, de ennek során olykor figyelembe kell venni az Alter és az Ego korábbi cselekvéseit is. A cselekvők előtt két döntési, választási lehetőség van: vagy kooperálnak (C) vagy dezertálnak (D). Ebben a dichotóm helyzetben a két alternatíva egymás tagadásának, komplementumának is tekinthető, tehát igaz a következő:

$$D = \neg C$$

Az ismételt játék n lépésig (fordulóig) tart, és a stratégiát leíró szabálynak meg kell adnia minden i. lépésre (vagy másként: mindent t_i időpontra), hogy Ego mit lép (DO(Ego, t_i)). A feltételes cselekvési stratégiák olyan formulákkal írhatók le, amelyekbe valamelyik ágens korábbi cselekvése is be van ágyazva paraméterként. A következőkben minden elemzett stratégia esetében bemutatom a cselekvési szabályt reprezentáló formulát, és ezekben sokszor lesz majd látható ez a beágyazottság vagyis a másodrendű vonás.

9.5.1 Reciprocitás

A feltételes stratégia (cselekvési szabály) összekapcsol két döntést, a jelen aktuális döntését a múlt vagy a jövő döntésével (esetleg döntéseivel). A feltételes stratégia így egymásra "feleselő", egymásra hivatkozó döntéseket, cselekvéseket köt össze, és ezzel lehetővé teszi azt, hogy modellezni tudjuk a reciprocitás jelenségét. A reciprocitás fogalmát cselekvések időben eltolt viszonzási rendszereként értelmezve a feltételes szabály fogalmával tudjuk megragadni a reciprocitás fogalmának legfontosabb összetevőjét. A reciprocitás jelenségében ugyanis összekapcsolódik az aktuális – jelenbeni – döntés/cselekvés a múltbeli, illetve a jövőbeni döntésekkel/cselekvésekkel, a jelen, a jövő és a múlt összerendelése adja az időben eltolt viszonzási rendszer lényegét (lásd Gouldner cikkét a reciprocitás értelmezéséről [304]).

Fontosnak tartom azonban megjegyezni, hogy egy szempontból nagyon félrevezető itt reciprocitásról beszélni. A reciprocitás fogalmának gouldneri értelmezése ugyanis normatív szinten zajlik (Gouldner összekapcsolódó jogosultságokról és kötelezettségekről beszél), míg a játékelméletnek ezen a szintjén még sehol nincsenek normák. A feltételes szabályokban megtestesülő szankciók, a jutalomként felfogható kooperációk, illetve a büntetésként felfogható dezertálások bár nyilván hasonlítanak a normativitás lényegi mozzanatát alkotó szankcionálási mechanizmushoz, de van egy nagy különbség. A játékelméletben itt még nincs norma, nincs előre rögzített előírás, nincs senkiben igény mások cselekvésére irányuló kötelezettségek vonatkozóan. És mindezen "hiányok" miatt természetesen a legitimitás kérdése sem vethető fel egyelőre.

A reciprocitás fogalmát máshol már bemutattam [722], itt csak néhány fontos vonást akarok kiemelni. A jeleség sokrétű és ezért nem meglepő, hogy sokféleképpen próbálták meg tipizálni. Számomra kétféle szempont szerint tipizálás fontos, amelyek azonban nem ortogonálisak egymásra, van átfedés közöttük. Szokás **pozitív reciprocitás** és **negatív reciprocitás**, illetve **direkt** és **indirekt reciprocitásról** beszélni. Utóbbiaknak külön fejezetet szentelek, előbbiek pedig a "szokásos" módon értelmezhetőek.

Fehr és Gächter pozitív és negatív reciprocitást különböztet meg aszerint, hogy valaki kooperál, jutalmaz vagy dezertál, büntet [243]. Herbert Gintis ugyanerre a kettősségre a gyenge, illetve erős reciprocitás fogalmakat alkalmazza [292]. Az altruizmus és reciprocitás kapcsolatát elemzi még Gintis társaival az 'Explaining altruistic behavior in humans' című cikkében [294], amelyből kiderül, hogy a kutatók között abban a tekintetben nincs vita, hogy a pozitív és negatív reciprocitás fogalmát a jutalmazó, kooperatív, illetve a büntető viselkedéshez kötik. Ez a szembeállítás persze korántsem újkeletű. Még jóval a játékelmélet, a kísérleti közgazdaságtan "megjelenése" előtt az antropológus Marshall Sahlins is hasonló felosztást javasolt (bár részben más terminusokat használt).

ellenőrizni az idézetet:

> „Az általános reciprocitás olyan tranzakciókat jelent, amelyek altruisztikusak, segítségnyújtás jellegűek ...
> A kiegyensúlyozott reciprocitás egyensúlyú cserét jelent. A tökéletes egyensúly azt jelenti, hogy a viszonzás egyenértékű a kapott jószággal, és azonnal történik ...
> A negatív reciprocitás próbálkozás büntetlenül semmiért valamit szerezni. Ilyen az erőszakos eltulajdonítás sok formája, vagy a személytelen és tisztán haszonelvű piaci tranzakció." [630, 264].

A pozitív és negatív reciprocitás mint jutalmazás és büntetés fontos szerepet kap a direkt és indirekt reciprocitás fogalmainak feltárásában is. De ezt már önálló fejezetekben mutatom be.

9.5.2 Direkt reciprocitás játékok

A direkt reciprocitásról akkor beszélhetünk, ha a játékosok ismerik egymást. Kérdés persze, hogy miért is fontos mindez, mit is jelent ez az ismeretség. Amikor ismételt játékokat engedünk meg, akkor azon túl, hogy a modellezni tudjuk a társadalmi emlékezet jelenségét, valami más is lehetővé válik. A játékok ismételhetősége, az a tény, hogy a játékok sorban követik egymást, az teszi lehetővé, hogy a játékosok reagálni tudjanak a másik cselekvésére. Figyelembe vehetik a partnerük (és esetleg a saját) múltbeli cselekvéseit, illetve mérlegelhetik azt is, vajon a jövőben mit lépne a partnerük, ha ők a jelenben így vagy úgy cselekednének. Innentől kezdve a játékosok már alkalmazhatnak feltételes stratégiákat is.
Az a feltételezés, hogy az emberek ismerik egymást, nyilván csak kis méretű társadalmat feltételezve valószerű. A direkt reciprocitás érvényességének fontos feltétele ez, amit Anatol Rapoport már a hatvanas években rögzített [593].

A direkt reciprocitás jelenségét **ismételt fogolydilemma** játékként lehet modellezni. Az egyszer játszott fogolydilemma játékban a kölcsönös dezertálás a domináns stratégia. A sok-sok megerősítő példa mellett azonban gyakran tapasztalhatjuk azt, hogy tipikus fogolydilemma szituációkban valamiért mégiscsak megjelenik a kooperáció. Ezt azzal

magjátékok	
aszimmetrikus játékok	szimmetrikus
többalternatívás játékok	kétalternatívás
többszereplős játékok	kétszereplős
nem szigorúan rendezett játékok	szigorúan rendezett
aszinkron játékok	szinkron
ismételt játékok	ismételt
direkt reciprocitás játékok	ismerős játékosok
indirekt reciprocitás játékok	ismeretlen játékosok
kooperatív (hangos) játékok	nem kooperatív (néma)

magyarázhatjuk, hogy ezek a szituációk ismétlődő jellegűek, vagyis a benne résztvevők ismerik egymást és gyakran kerülnek hasonló helyzetekbe. A játékelméletben ezt ragadhatjuk meg az ismételt játékok fogalmával. Az ismételt játékok azzal írják le pontosabban a társadalmi cselekvéseinket, hogy ebben a modellben már megjelenik a társadalmi emlékezet jelensége. Ezzel persze rögtön bonyolultabbá is válik a modell, hiszen innentől kezdve már számon kell tartani azt is, hogy a játékosok miket léptek (hogyan döntöttek) az előző fordulókban. Szerencsére az ismételt játékok (ismételt fogolydilemmák) helyzeteit a hetvenes évektől kezdve már lehetett számítógépes szimulációk segítségével is vizsgálni. Az ismételt fogolydilemma helyzetek elemzésének klasszikusa műve Robert Axelrod híres szimulációs tornájáról szóló beszámoló [44].[7]

A direkt reciprocitás játékelméleti modellezéséről a legtöbben a híressé vált tit-for-tat stratégia miatt és révén hallottak, írtak. Nem tagadható, hogy ez a stratégia valóban központi szerepet játszik a játékelméletben, de szeretném bemutatni azt is, hogy emellett vannak még más stratégiák is, amelyek egyrészt bizonyos feltételek között a tit-for-tat stratégiával "összemérhető" módon képesek teljesíteni, másrészt egy teljesnek mondható fogalmi tér hiányzó elemeit, részeit alkotják. Ezért a továbbiakban a direkt reciprocitás körülményei között elméletileg létező, értelmezhető összes másodrendű stratégiát elemezni fogom. Előtte azonban röviden ki kell térnem a stratégiák, pontosabban az evolúciósan sikeres stratégiák kérdésére.

Evolúciósan stabil stratégia

A stratégia fogalmát már bemutattam, és a későbbiekben is lesz még szó róla. Az **evolúciósan stabil stratégia** fogalmát azonban még nem vettem tárgyalásba, most viszont megteszem, mert szükségem lesz rá.[8] A fogalom az evolúciós elméletek területéről szivárgott át a játékelmélet vidékére.[9] Az evolúciósan stabil stratégia fogalmával speciális minőséget lehet kifejezni a stratégiák egymáshoz való viszonyában.

Az evolúcióbiológiától átvehető egyik fontos módszertani eszköz a populációdinamikai

[7] A direkt reciprocitás játékelméleti modellezéséről, benne a tit-for-tat stratégiáról, bővebben írtam a cselekvéselméleti tankönyvemben [722].

[8] Az evolóciósan stabil stratégia (ESS) fogalmát John Maynard Smith emelte be a tudományos diskurzusba [696].

[9] A játékelmélet és az evolúcióbiológia kölcsönhatásáról, a tanulás és a kiválasztódás különbségéről máshol már elmondtam a véleményemet [722]. Megengedhetőnek tartom az evolócióbiológia fogalmait, módszereit igénybe venni a játékelmélet társadalomtudományi alkalmazása során, mert tudományos eszközként hasznosnak tartok őket, de mindez egy pillanatig sem jelenti, jelentheti azt, hogy elfogadnám és alkalmaznám a társadalomtudomány világában az evolócióbiológia determinisztikus szemléletmódját.

modellezés technikája. Ha vizsgálni akarjuk, hogy valamely viselkedésmód, magatartási stratégia hogyan képes elterjedni egy nagyobb populációban, akkor azt úgy tehetjük meg, hogy a "versengő" stratégiákat "szétosztjuk" a különböző csoportok között, és szimulációs számításokat végzünk arra vonatkozóan, hogy a csoportok, egyének interakciói során a különböző stratégiák követése mennyi nyereséget, veszteséget hoz a stratégiákat követő egyének, csoportok számára. A stratégiák sikerességét azzal kifejezhetjük ki, hogy a népesség egészében egyre több egyed követi a sikeres viselkedési szabályt. Erre a "kiválasztódási folyamatra" nyugodtan mondhatjuk azt, hogy társadalmi tanulással, nem pedig "biológiai" kiválasztódással valósul meg. A népesség egészében elterjedni látszó – ebben az értelemben – domináns stratégiák azonban még nem biztos, hogy hosszú ideig képesek lesznek stabilan fennmaradni. Fontos kérdés az, hogy mi történik, ha a populáción belül egy domináns stratégia mellett megjelenik egy másik stratégiát követő csoport. Meg tudja-e az új – mutáns – stratégia ingatni a korábban domináns stratégia helyzetét? Saját magával szemben hogy viselkedik a stratégia, megerősíti-e saját magát, vagy esetleg utat nyit más stratégiák számára, amelyek később 'lenyomhatják' őt? Ilyen kérdések merülhetnek fel a domináns stratégiákkal kapcsolatban. Az evolúciósan stabil stratégiák képesek ellenállni az ilyenfajta "támadásoknak", vagyis az adott viselkedési mintázatot tartósan fenn tudják tartani a közösség egészén belül.

A négy másodrendű stratégia

Az ismételt játékok modelljében az az egyik döntő mozzanat, hogy megjelenik benne az emlékezet. A feltétel nélküli stratégiák esetében nincs szükség még erre, de a tit-for-tat és a többi feltételes stratégia már az emlékezet segítségével tudja a stratégia feltételét szabályba foglalni. És ha van már emlékezet, akkor – érdekes módon – megjelenik a jövő is, hiszen ha megismerem a partnerem stratégiáját, akkor ez alapján saját döntésemet formálhatja a jövőre vonatkozó elvárás is. Axelrod szép kifejezésével élve: a jövő árnyéka visszahat a jelenre.

Nem mindegy azonban, hogy mit veszek figyelembe a múltból. A tit-for-tat arra figyel, hogy a partner mit lépett az előző fordulóban. Ez gazdaságos megoldás, mivel nem kell hozzá sok információt tárolni. Minél több patnerünk lehet, annál fontosabbá válik ez a szempont. Két irányban is tágíthatjuk a megjegyzendő múltbeli információk körét. Egyfelől hátranézhetünk messzebb az időben, vagyis több lépésre visszamenőleg figyelhetjük partnereink cselekedeteit. A tit-for-two-tat stratégia például két forduló adatait jegyzi meg: a partner első dezertálása után még kooperál és "csak" a sorban második dezertálást bünteti megtorlással. A távoli múltba nézés azonban nem igazán elterjedt a szimulációs elemzések világában – vélhetőleg azért, mert túlzottan nagy memóriaigénnyel járna, és – talán – nem is lenne annyira életszerű, hiszen a való életben sem tartjuk minden partnerünkről számon azt, hogy pontosan mit tett velünk szemben korábban. Sokkal inkább valamiféle általános benyomást alakítunk ki róluk, de ez már a reputációs elméletekhez vezetne minket el, amivel csak a következő fejezetben foglalkozom majd (↬364.o.).

A memória bővítésének másik útja az lehet, hogy továbbra is megmaradunk az előző forduló adatainál, de nemcsak a partnerünk lépését, hanem a sajátunkat is figyelembe vesszük. A tit-for-tat szabály ezt nem teszi meg, de vannak olyan stratégiák, amelyek figyelnek erre. Ha tudni szeretnénk, hogy a direkt reciprocitás feltételei mellett hány

stratégia lehetséges elméletileg és mik ezek, akkor érdemes egy egyszerű formalizmust alkalmazni a stratégiák leírására [725].

Egy cselekvési szabály abból áll, hogy bizonyos feltételekhez kooperálást (C), másokhoz dezertálást (D) ír elő. Ezt kifejezhetjük úgy is, hogy a kooperáláshoz az '1', dezertáláshoz '0' értéket rendeljük. Mivel itt a stratégiákat mindig a játékos és partnere előző fordulóban tett lépései alapján határozhatjuk meg, egy cselekvési szabály úgy adhatunk meg, hogy az előző forduló négy lehetséges állapotához, a kölcsönös kooperáláshoz (CC) és dezertáláshoz (DD, illetve a két aszimmetrikus cselekvéshez (CD és DC) egy-egy cselekvésértéket rendelünk, vagyis a (CC,CD,DC,DD) állapotsorhoz hozzátesszük a szabály által előírtakat. Ezt 16-féleképpen tehetjük meg, tehát ennyi stratégia létezik. és az egyes stratégiákat a (0,0,0,0), (0,0,0,1), ..., (1,1,1,1) számsorozatok valamelyikével reprezentálhatjuk.

Nem minden elméletileg lehetséges stratégia "értelmes". Vannak olyanok, amelyek gyakran felbukkannak a szimulációs elemzésekben, és ami fontosabb: a gyakorlati életben. Ilyen a feltétel nélkül kooperáló, az **allC** (1,1,1,1), a feltétel nélküli dezertáló, az **allD** (0,0,0,0), a kockadobó, a **random** (0.5, 0.5, 0.5, 0.5) vagy az alternáló, az Alt (0,0,1,1). A stratégiák között előfordulnak olyanok is, amelyek sosem szerepelnek jól a szimulációk során és a gyakorlati előfordulásuk esélye is kicsi. Ezeknek nevet sem szokás adni, ezért csak a számnégyesükkel hivatkozhatunk rájuk. Ilyen például ez a stratégia: (0,0,0,1). A következő fejezetekben röviden megvizsgálom a négy legfontosabb stratégiát, amelyek valamilyen értelemben sikeresnek mondhatóak a szimulációs számítások és a gyakorlati élet tapasztalatai szerint.

Tit-for-tat, a tiszta viszonzó stratégia

A direkt reciprocitás elmélethez kapcsolható stratégiák legismertebb példája, a tit-for-tat a Robert Axelrod által meghirdetett számítógépes torna kapcsán került be a köztudatba. Axelrod nyilt versenyt hirdetett ismételt fogolydilemma játéksorozatban való részvételre. A versenybe benevezett stratégiák (pontosabban a stratégiákat megvalósító számítógépes programok) sok meneten keresztül egymás ellen játszottak, mindenki mindenki ellen megmérkőzött, a játékosok (stratégiák) minden játék után pontokat kaptak, ezeket folyamatosan összesítették, aztán a torna végén győztest hirdettek. A végeredmény mindenkit meglepett. Az általánosnak mondható elvárásokkal szemben, amelyek valamilyen dezertáló jellegű stratégia győzelmét jósolták, a **tit-for-tat** stratégia lett a torna nyertese, amely viszont tartalmazott kooperatív elemeket is. Mellesleg a tit-for-tat stratégia volt a legegyszerűbb az összes között, hiszen pusztán két szabályból áll:

1) kezdj kooperálással,

2) mindig ismételd meg a partnered előző fordulóban megtett lépését.

A tit-for-tat terminusra nem született szélesebb körben elfogadott magyar fordítás. Nem azért, mert nincs a tartalmára megfelelő szöveg. Inkább az a gond, hogy túl sok van. Vannak olyan közmondásaink, amelyek egyértelműen megfelelnek a tit-for-tat stratégia lényegének és széles körben elterjedtek:

> Jótett helyébe jót várj.
> Ki mint vet, úgy arat.
> Amilyen az adjonisten, olyan lesz a fogadjisten.
> Aki bottal köszön, annak doronggal felelnek.

> Amilyen a kérdés, olyan a felelet.
> Amit nem kívánsz magadnak, te se tedd embertársadnak.
> Kölcsön kenyér visszajár.
> Valamit valamiért.
> Aki nem dolgozik, ne is egyék!
> Aki kardot ránt, kard által vész el.
> Ki gyomot vet, gazt arat.

A Bibliában is több helyen találhatunk olyan passzusokat, amelyek "ugyanerről beszélnek":

> „A ki ember-vért ont, annak vére ember által ontassék ki; . . . " [Mózes I. 9.6]

> „Szemet szemért, fogat fogért, kezet kézért, lábat lábért,
> Égetést égetésért, sebet sebért, kéket kékért." [Mózes II. 21.24-25]

> „Törést törésért, szemet szemért, fogat fogért; a milyen sérelmet ő ejtett máson, olyan ejtessék rajta is.
> A ki barmot üt agyon, fizesse meg azt, de a ki embert üt agyon, halállal lakoljon." [Mózes III. 24.20-21]

> „És a bírák vizsgálják meg jól a dolgot, és ha hazug tanú lesz a tanú, a ki hazugságot szólott az ő atyjafia ellen: Úgy cselekedjetek azzal, a mint ő szándékozott cselekedni az ő atyjafiával. Így tisztítsd ki közüled a gonoszt; Hogy a kik megmaradnak, hallják meg, és féljenek, és többször ne cselekedjenek te közötted ilyen gonosz dolgot. Ne nézz reá szánalommal; lelket lélekért, szemet szemért, fogat fogért, kezet kézért, lábat lábért. [Mózes V. 19.18-21]

Más kultúrkörből, az Edda énekekből származik a következő idézet [545]:

> 42.
> Barátnak barátja légy,
> adományt adománnyal
> hálálj híven.
> Gúnyt fizess meg gúnnyal,
> s ne maradj fösvény,
> fizetvén gonoszért.
>
> 43.
> Barátnak barátja légy,
> s barátod barátjának
> szintén szívese;
> ám aki ellened,
> egyetlen hívét
> se szíveld soha.

A magyar nyelvű játékelméleti szakirodalomban a tit-for-tat leginkább elterjedt fordítása a 'szemet szemért' kifejezés lett, amit azonban a leszűkítő – csak a negatív reciprocitást kifejező – értelme miatt nem tartok szerencsésnek. A tit-for-tat szabály lényege a viszonzás, ami miatt lehetne 'utánzó' vagy 'tükröző stratégiának' is nevezni, de még pontosabbnak érzem az **viszonzó stratégia** terminusát, mert ez fejezi ki leginkább a partner korábbi cselekvésének ismétlésében rejlő mechanikus mozzanatot. Pontossága ellenére sem akarom azonban az általam legjobbnak érzett terminust használni, és maradok inkább az eredeti kifejezés (tit-for-tat) mellett.

A fentebb idézett közmondások, maximák értelmezésében egyébként használhatjuk a már hivatkozott megkülönböztetést a pozitív és negatív reciprocitás között. Ha a kooperációra kooperáló válasz a pozitív, míg a dezertálásra dezertáló válasz a negatív reciprocitás, akkor a 'jó tett helyébe jót várj' vagy a 'kölcsönkenyér visszajár' közmondások a pozitív, a 'szemet szemért', a 'Ki kardot ránt, kard által vész el' vagy a 'Ki gyomot vet, gazt arat' elvek a negatív reciprocitást írják le, míg a magyar fordítások közül az 'amilyen az adjonisten, olyan lesz a fogadjisten' (illetve kicsit gyengébb formában a 'Ki mint vet, úgy arat', illetve az 'Amilyen a kérdés, olyan a felelet') közmondások írják le a reciprocitás mindkét oldalát.

A szemet szemért elvről annyit még fontosnak tartok megjegyezni, hogy az – legalább Hammurapi törvénykönyve óta – a lex talionis elvének köznyelvi kifejeződéseként is fefogható. És ami még fontosabb: a lex talionis elv az arányos büntetés kívánalmát és gyakorlatát hozta el, amikor megjelent. Éppen ezért ez az elv, megjelenésekor humánus volt a korábbi állapotokhoz, a korábbi büntetési elvekhez és gyakorlatokhoz képest. A "szemet szemért" elvet akkor úgy kellett érteni, hogy "csak szemet szemért", nem pedig úgy, ahogy korábban tették: "életet szemért". Ezt a történelmi tényt csak azért nehéz észrevenni és értékelni, mert a keresztény kultúrában a szemet szemért elv volt az, amely időben később szembe került egy másik híresssé vált maximával, a jézusi aktív szeretet parancsával.

A tit-for-tat stratégia jelentősége abban rejlik, hogy ez egy kooperatív elemeket (is) tartalmazó stratégia, szemben a dezertáló stratégiákkal. A kooperálás mozzanata abban érhető tetten, hogy a stratégia szerint az ellenfél előző menetbeli kooperálását meg kell ismételni vagyis kooperálni kell. A tit-for-tat játékos egy barátságos, kooperatív partnerrel szemben maga is kooperatív módon fog mindvégig viselkedni. Igaz, a tit-for-tat tud barátságtalan is lenni, hiszen az ellenfele dezertálását azonnal megtorolja a saját dezertálásával. Mindezekért jellemezte úgy Axelrod a tit-for-tat stratégiát, hogy az:

- barátságos (mert az első menetben kooperál),
- megtorló (mert a dezertálást azonnal megbosszulja),
- megbocsátó (mert a kooperálásra azonnal kooperálással válaszol, bármi is volt azt megelőzően) és
- nem irigy (mert egy játékban sosem győzi le a másikat, maximum annyi pontot szerez, mint a partnere).

A tit-for-tat stratégia ezeken felül még feltételes stratégia is egyben, hiszen mindig a partner lépéséhez igazítja a saját lépését. Ebben a minőségében szemben áll a feltétel nélküli stratégiákkal (mint a mindig kooperáló allD, illetve mindig dezertáló allD szabály). Ez a feltételesség a stratégia reciprok jellegére utal, s mivel a feltételes kooperálásban legtöbbször tetten érhetünk valamiféle altruizmust is, a tit-for-tat stratégiát a reciprok altruizmus és a direkt reciprocitás) fogalmával is összekapcsolhatjuk.

A korábban javasolt jelölési konvenció alapján a tit-for-tat stratégiát az alábbi táblázatos formában reprezentálhatjuk:

Ego és *Alter* cselekvései az előző fordulóban	(CC,	CD,	DC,	DD)
Ego cselekvése az aktuális fordulóban	(C,	D,	C,	D)
	(1,	0,	1,	0)

Ezt a (CC,CD,DC,DD) séma alapján lehet értelmezni, ami az előző forduló lépéseit jelzi. A CC azt jelenti, hogy az Ego (tehát a tit-for-tat stratégiát követő játékos) és Alter (tehát

a partner) egyaránt kooperált, a CD azt, hogy Ego kooperált, Alter dezertált, DC esetén Ego dezertált, Alter kooperált és a DD a kétszeres dezertálásra utal. A tit-for-tat stratégia mindig a partner (Alter) lépéseit veszi figyelembe, és azt ismétli meg, így adódik a fenti séma szerint hogy a (1,0,1,0) reprezentáció. A cselekvési szabályra felírhatunk egy formulát is:

(IL.5) $$\text{DO}(Ego, t_i) = (t_i = t_1 \rightarrow C) \wedge (t_i > t_1 \rightarrow \text{DO}(Alter, t_{i-1}))$$

tit-for-tat (HUN) – szemet szemért (HUN) – viszonzó stratégia (HUN) – tit-for-tat (ENG) – TFT (ENG)

A formula azt fejezi ki, hogy mit tesz az Ego az i. lépésben. A viszonyzó stratégiát két állítással definálták, ezt fejezi ki a jobb oldali kifejezésben található részformula. Az első lépésben (t_1 időpontban) Ego kooperál (C), a többi fordulóban ($t_i > t_1$) a játékostársa ($Alter$) előző (t_{i-1} időpontbeli) lépését ismétli.

A tit-for-tat stratégia tehát sikeresnek bizonyult mind a szimulációban, mind a gyakorlatban. A kezdeti ragyogását azonban idővel kissé beárnyékolta az a tény, hogy kiderült, bár a stratégia kollektíve stabil, de evolúciós értelemben már közel sem az [112]. Ha például a hibázás lehetőségét is beépítjük a modellbe, figyelembe vesszük a szimulációk során, akkor a stratégia jóval gyengébben teljesít. Ezt persze tudtuk már jóval a szimulációs számítások előtt is, hiszen a szemet szemért (tit-for-tat) elv kritikája kétezer év óta jól ismert a keresztény kultúrkörön belül. Hiszen a jézusi 'ki kardot ránt, kard által vész el' mondat ugyan egyfelől a negatív reciprocitás tényszerű megállapítás, de másfelől egy figyelmeztetés, normatív felszólítás gesztusa is egyben.

> "An-eye-for-an-eye-for-an-eye-for-an-eye ... ends in making everybody blind"
> (Mahatma Gandhi)
> "The old law of an eye for an eye leaves everyone blind"
> (Martin Luther King Jr.)

Nagy kérdés, hogy az aranyszabály tényleg ide tartozik-e. Mert ha igen, akkor be kell mutatni azt, hogy az összes világvallásban felbukkan, és idézni a megfelelő könyvet.

Bűnbánó stratégia

Ha aszinkron játékokból képezünk játéksorozatot, akkor meglehetősen sikeres stratégiának bizonyul a **szigorú, de igazságos** ('firm-but-fair' – **FBF**) szabály) [269]. Bár a 'szigorú, de igazságos' kifejezést a magyar közgondolkodás is ismeri, alkalmazza, mégis úgy gondolom, hogy jobban járunk, ha a **bűnbánó stratégia** (CONTRITE TIT-FOR-TAT STRATEGY) kifejezést használjuk helyette.

Az FBF stratégia erősen kooperatív, mert csak egy esetben dezertál: ha az előző menetben a játékos kooperált, miközben a partnere dezertált. Ezt úgy értelmezhetjük, hogy nem is annyira a nagylelkűség, mint inkább a "bűnbánás" a stratégia igazi sajátossága, hiszen ha a játékos korábban dezertált, utána mindkét esetben kooperál (mintegy megbánva korábbi "bűneit"). A tit-for-tat stratégiától a kétszeres dezertálásra való reakcióban tér el azzal, hogy a bűnbánó ekkor kooperál, míg a tit-for-tat dezertál. A tit-for-tat szabály a partner minden dezertálását bünteti, ám a bűnbánó stratégia nem. A kölcsönös

dezertálás helyzetét már "szégyenli" (hiszen ő is dezertált), ezért próbálja azt "kijavítani" a kooperálással. A bűnbánó stratégia táblázata a következő:

Ego és *Alter* cselekvései az előző fordulóban	(CC,	CD,	DC,	DD)
Ego cselekvése az aktuális fordulóban	(C, (1,	D, 0,	C, 1,	C) 1)

A stratégiát leíró formula következőképpen néz ki:

(IL.6) $\text{DO}(Ego,t_i) = (t_i = t_1 \rightarrow C) \wedge (t_i > t_1 \wedge \text{DO}(Ego,t_{i-1}) = C \rightarrow \text{DO}(Alter,t_{i-1})) \wedge (t_i > t_1 \wedge \text{DO}(Ego,t_{i-1}) = D \rightarrow C)$

bűnbánó stratégia (HUN) – szigorú, de igazságos (HUN) – contrite TFT (ENG) – Firm-But-Fair (ENG) – FBF (ENG)

A formula három mondatra tagolható: az első állítás az első fordulóban, a következő kettő részformula az összes többi fordulóban érvényes előírást rögzíti. Az első állítás azt fejezi ki, hogy az első fordulóban ($t_i = t_1$) *Ego* kooperál (C). A második részformula írja le azt a részszabályt, hogy amennyiben az előző fordulóban ($t_i = t_{i-1}$) az *Ego* kooperált (C), akkor az i. menetben a partnere (*Alter*) előző lépését ismétli. A harmadik összetevőben pedig az szerepel, hogy amennyiben az *Ego* az előző fordulóban dezertált (D), akkor most kooperál (C).

Meg kell még itt jegyezzem, hogy amikor Frean bevezette az FBF-stratégiát, akkor a zajos játékokhoz igazodó sztochasztikus stratégiát tervezett, így a kétszeres dezertálás esetében nem a biztos, csak a nagyon valószínű kooperálást írta elő, amit ezzel a mintázattal fejezhetünk ki: (1,0,1,p). A fent elmondottak logikája és érvényessége azonban azonban nem változik meg amiatt, hogy a bűnbánó stratégiát a FBF-szabály determinisztikus változataként értelmezem. Érdemes még azt is megjegyezni, hogy ugyanez a cselekvési szabály felbukkan a 'bűnbánó tit-for-tat' ('contrite TFT') [111], valamint a hírnévápoló ('standing') stratégia név alatt is [716]. Ezekkel azért nem foglalkozom itt, mert mindkettő stratégia az indirekt reciprocitás körébe tartozik.

Számító stratégia

Ahogy Axelrod tornáján Anatol Rapoport pszichológus volt az, aki beküldte a nyertes tit-for-tat programot, úgy az ő nevéhez köthető annak a stratégiának az első megfogalmazása is, amit **együgyű** (simpleton) stratégiának nevezett el [605]. Ez a stratégia később más neveket is kapott, és inkább ezek terjedtek el szélesebb körben. A 'nyer-marad-veszít-vált' (angolul win-stay-lose-shift – WSLS) stratégia [535], valamint a Pavlov megnevezések [415][416] egyaránt a Rapoport-féle együgyű stratégiának felelnek meg, aminek lényegét a 'Győztes csapaton ne változtass!' mondás fejezi ki a legjobban. Ez a szabály a kölcsönösség után vagyis a szimetrikus cselekvések esetén (CC és DD) kooperál, aszimmetrikus helyzetekben (CD és DC) dezertál, amit így írhatunk fel:

Ego és *Alter* cselekvései az előző fordulóban	(CC,	CD,	DC,	DD)
Ego cselekvése az aktuális fordulóban	(C, (1,	D, 0,	D, 0,	C) 1)

A stratégia logikája, a ‘győztes csapaton ne változtass’ elv helyénvalósága azáltal érthető meg, ha belátjuk a stratégia egy fontos vonását. Ez a szabály ugyanis úgy működik, hogy mindaddig megismétli az előző fordulóban tett lépését, amíg nyer (a két legtöbbet fizető helyzetben), de amint veszít (a két legkevesebbet fizető helyzetben), ellenkezőjére váltja a cselekvését. Ha a játékok lehetséges kimeneteinek kívánatosságát egyre növekvő számokkal jelöljük (1-2-3-4), akkor ez a stratégia akkor ismétli meg a korábbi cselekvését, ha 3-at vagy 4-et szerzett az előző fordulóban, és akkor vált cselekvést, ha előzőleg 1-t vagy 2-t kapott. A fogolydilemma esetében a négyféle (CC,CD,DC,DD) helyzetben a játékos (C,C,D,D) lépéseihez a (3,1,4,2) nyereményeket rendelhetjük hozzá. A cselekvési szabály az első és harmadik esetben megismétli az előző lépését, vagyis (C,?D,?), a második és negyedik esetben magatartást vált, vagyis C helyett D-t, illetve D helyett C-t lép. Így alakul ki a teljes szabály mintázata: (C,D,D,C). Ezt a számokkal kifejezve a fentebb már jelzett formában írhatjuk fel: (1,0,0,1).
A stratégia formulája:

(IL.7) $\text{DO}(Ego,t_i) = (t_i = t_1 \rightarrow C) \wedge (t_i > t_1 \wedge \text{DO}(Ego,t_{i-1}) = C \rightarrow \text{DO}(Alter,t_{i-1})) \wedge (t_i > t_1 \wedge \text{DO}(Ego,t_{i-1}) = D \rightarrow \neg\text{DO}(Alter,t_{i-1}))$

számító (HUN) – győztes csapaton ne változtass (HUN) – Win-Stay-Lose-Shift (ENG) – Simpleton (ENG) – WSLS (ENG)

A háromrészes formulából az első szabály írja le az első forduló teendőjét (Ego kooperál). A második részformula szerint, amennyiben az előző fordulóban ($t_i = t_{i-1}$) az Ego kooperált (C), akkor az i. menetben a partnere ($Alter$) előző lépését ismétli. A harmadik részformulában pedig az szerepel, hogy amennyiben az Ego az előző fordulóban dezertált (D), akkor a jelenlegiben pont az ellenkezőjét csinálja annak, mint amit az $Alter$ csinált az előzó menetben ($\neg\text{DO}(Alter,t_{t-1})$).[10]

Ez a szabály tanuló stratégia, mivel az előző forduló tapasztalatait értékelve alakítja ki az aktuális lépését. Mivel a szabály arra figyel, hogy a mindig a legtöbb pontszámot érje el, ezért **számító stratégia** megnevezés tűnik a legmegfelelőbbnek. Ez a stratégia is akkor igazán sikeres, ha a tit-for-tat játékosok előzőleg megtisztítják a terepet az önző allD stratégiáktól [605].

Haragtartó stratégia

Az elemzésekben időről időre felbukkan a Grim trigger (GT) stratégia [270], amit **haragtartó**, engesztelhetetlen, zord vagy zordra váltó stratégiának fordíthatnánk (Axelrod 1984-es könyvében ezt Friedman stratégiának nevezi [44]). Ez a cselekvési mintázat barátságosan kezd, és mindaddig kooperál, amíg a másik is ezt teszi, de a partner első dezertálása után ő is azonnal dezertálásba vált, és attól kezdve már sosem tér vissza a kooperáláshoz. Ez a stratégia engesztelhetetlenül haragtartó, sosem bocsát meg. Azok, akik ezt a stratégiát javasolják, az tartják az előnyének, hogy nagy a visszatartó ereje, hiszen a kooperálás feladását az örökös dezertálással fenyegeti, vagyis a legerősebb büntetést helyezi kilátásba. A haragtartó stratégiára a következő mintázat jellemző:
A stratégiát leíró formula következőképpen néz ki:

[10] Itt kihasználhatjuk azt, hogy a játékosok cselekvései csak két értéket vehetnek fel, vagy kooperálnak vagy dezertálnak, s emiatt a két cselekvésérték egymás tagadásának tekinthető.

Ego és *Alter* cselekvései az előző fordulóban	(CC,	CD,	DC,	DD)
Ego cselekvése az aktuális fordulóban	(C,	D,	D,	D)
	(1,	0,	0,	0)

(IL.8) $\text{DO}(Ego,t_i) = (t_i = t_1 \rightarrow C) \wedge (t_i > t_1 \wedge \text{DO}(Ego,t_{i-1}) = C \rightarrow \text{DO}(Alter,t_{i-1})) \wedge (t_i > t_1 \wedge \text{DO}(Ego,t_{i-1}) = D \rightarrow D)$

haragtartó (HUN) – Grim Trigger (ENG) – Friedman (ENG)

A formula első állítása rögzíti, hogy *Ego* kooperál az első menetben. A második részformula szerint *Ego* előző fordulóbeli kooperálása esetén az aktuális fordulóban megismétli a partnere (*Alter*) előző lépését. A harmadik összetevő pedig azt írja le, hogy amennyiben az *Ego* az előző fordulóban dezertált (*D*), akkor most is dezertálni fog (*D*) függetlenül attól, mit tett korábban a partnere.

A szimulációs eredmények azt mutatják, hogy ez a stratégia sosem tud igazán jól szerepelni, mert nagyon könnyen keveredik megtorlási spirálba [43]. Különösen így van ez zajos játékokban (erről a következő fejezetben kicsit bővebben írok).

A négy alapstratégia

Ez előző fejezetekben bemutatott stratégiák mind olyanok voltak, hogy a szimulációs számítások és a gyakorlati tapasztalat szerint is a sikeres, bizonyos feltételek mentén gyakran alkalmazott cselekvési szabályoknak lehet minősíteni őket. Talán csak a haragtartó (Grim trigger) stratégia volt az, amire valamivel kevesebbszer hivatkoztak a többi háromhoz képest.

A bemutatott stratégiáknak van egy érdekes közös vonása, ami talán nem véletlen. Mind a négy szabály megegyezik abban, hogy mit tesz azután, hogy a játékos kooperált az előző fordulóban (modhatjuk azt is, hogy a játékos jó volt):

- kooperál, ha a partnere kooperált,
- dezertál, ha a partnere dezertált.

Ezt fejezi ki az, hogy a számnégyesük első két helyén mindannyiszor az ‘1,0’ érték szerepel. Ha a négy vizsgált stratégia megegyezik abban, hogy mit ír elő arra az esetre, amikor az előző fordulóban a játékos kooperált (jó volt), akkor a különbség csak abban lehet köztük, hogy mit ír elő a stratégia arra az esetre, amikor az előző fordulban a játékos dezertált (vagyis rossz volt). Ilyen esetből viszont pont négy van (nem lehet se több, se kevesebb), és az az érdekes, hogy pont ez alapján jellemezhetjük (különíthetjük el) a vizsgált stratégiáinkat is:

		(CC,	CD,	DC,	DD)
viszonzó	TFT	(1,	0,	1,	0)
bűnbánó	FBF	(1,	0,	1,	1)
számító	WSLS	(1,	0,	0,	1)
haragtartó	GT	(1,	0,	0,	0)

A direkt reciprocitáson alapuló kooperációs lehetőségeket keresve, a kooperatív stratégiákat értelmezve azt mondhatjuk, nem lehet véletlen, hogy a fenti négy stratégia szerepel a legtöbbet az elemzésekben (és tegyük hozzá: a mindennapi életünkben). A kooperációs

minták elterjedésének feltétele lehet az a "félszabály", hogy ha előzőleg kooperáltunk (jók voltunk), akkor az viszonzás elve szerint, tehát a tit-for-tat stratégia szerint kell eljárnunk. Az előző menetbeli dezertálásunk (tehát a rosszaságunk, a rossz múltunk) esetén adott négyféle cselekvési lehetőség pedig azt a négy stratégiát írja le, amelyek kooperációs potenciálját a játékelmélet is alátámasztotta, és amiket a mindennapi életben mi is megtapasztalhattunk, alkalmaztunk évezredek óta, alkalmazunk a jelenben is.

Érdemes a fejezet végén egy táblázatba összegyűjtve bemutatni az eddig elemzett stratégiák formuláit is. A kontraszt kedvéért a listába felveszem a korábban említett feltétel nélküli stratégiákat is, a stratégiákat reprezentáló formulákat pedig úgy tagolom, hogy jobban láthatóak (oszlopokba rendezve) legyenek a bennük levő összetevők.

a stratégia neve	a stratégia: $\text{DO}(Ego, t_i)$		
allC			C
allD			D
TFT	$t_i = t_1$	$\rightarrow$	C
	$\wedge \quad t_i > t_1 \wedge \text{DO}(Ego, t_{i-1}) = C$	$\rightarrow$	$\text{DO}(Alter, t_{i-1})$
	$\wedge \quad t_i > t_1 \wedge \text{DO}(Ego, t_{i-1}) = D$	$\rightarrow$	$\text{DO}(Alter, t_{i-1})$
FBF	$t_i = t_1$	$\rightarrow$	C
	$\wedge \quad t_i > t_1 \wedge \text{DO}(Ego, t_{i-1}) = C$	$\rightarrow$	$\text{DO}(Alter, t_{i-1})$
	$\wedge \quad t_i > t_1 \wedge \text{DO}(Ego, t_{i-1}) = D$	$\rightarrow$	C
WSLS	$t_i = t_1$	$\rightarrow$	C
	$\wedge \quad t_i > t_1 \wedge \text{DO}(Ego, t_{i-1}) = C$	$\rightarrow$	$\text{DO}(Alter, t_{i-1})$
	$\wedge \quad t_i > t_1 \wedge \text{DO}(Ego, t_{i-1}) = D$	$\rightarrow$	$\neg\text{DO}(Alter, t_{i-1})$
GT	$t_i = t_1$	$\rightarrow$	C
	$\wedge \quad t_i > t_1 \wedge \text{DO}(Ego, t_{i-1}) = C$	$\rightarrow$	$\text{DO}(Alter, t_{i-1})$
	$\wedge \quad t_i > t_1 \wedge \text{DO}(Ego, t_{i-1}) = D$	$\rightarrow$	D

A tit-for-tat szabályt redundáns módon írtam fel. Olyan alakra akartam hozni, hogy összehasonlítható legyen a többi feltételes stratégiával, ezért a második részformuláját kettébontottam, és – a többi stratégiát leíró formulához hasonlóan – külön fejeztem ki azt, amikor Ego kooperált, illetve dezertált a megelőző fordulóban. A redundancia abban érhető tetten, hogy mindkét részfeltétel esetén ugyanaz a teendő adódik: meg kell ismételni a partner korábbi lépését.

A táblázatra tekintve jobban szembetűnő, hogy olyan két stratégia van, amely teljesen a partner előző lépéséhez köti az aktuális cselekvését (a teljesen ugyanazt viszonzó tit-for-tat és a részben hasonló, részben ellenétes lépést viszonzó pavlovi szabály), és a két másik stratégia az Ego korábbi dezertálása esetén a partner előző lépésétől függetlenül cselekszik (a bűnbánó kooperál, a haragtartó dezertál). És természetesen szépen kitűnik a táblázatból az a korábban már említett fontos közös mintázat is, miszerint az $Alter$ megelőző fordulóbeli kooperálása ("jósága") esetén mindig a partner lépését "kell" megismételni.

A haragtartó és a bűnbánó stratégiák abban az értelemben merevebbek a másik kettő szabályhoz képest, hogy csak "egyszeresen" igazodnak a másik játékos korábbi lépéseihez (amikor az Ego kooperált az előző fordulóban). Ha úgy kerülnek a partnerükkel szembe, hogy az előző menetben dezertáltak, akkor mindketten "feltétel nélküli" stratégiát követnek (a bűnbánó stratégia kooperál, a haragtartó dezertál). Ehhez képest erősebben igazodik a tit-for-tat szabály a másikhoz, hiszen függetlenül attól, hogy Ego mit tett ko-

rábban, mindig az *Alter* lépését ismétli. Ez a szabály egy fokkal azonban még mindig merevebb a WSLW-stratégiához képest, hiszen az utóbbbi szabály hol a partner korábbi lépését ismétli, hol annak tagadását. Talán ez a "másság" vagy "képesség" magyarázhatja azt, hogy a szimulációk során – legalábbis bizonyos feltételek között, kiváltképp zajos környezetben – miért ez a stratégia bizonyult a legjobbnak. Amikor erre a pavlovi stratégiára azt mondják [415], [416], hogy tanul az előző fordulók eseményeiből, akkor ez a tanulási képesség talán épp ennek a "rugalmasabb" igazodásnak köszönhető.

Társadalmi hiba – zajos játékok

Az előző fejezetekben megírtam, hogy az ismételt fogolydilemma játék (Iterated Prisoner's Dilemma Game – IPDG) bevezetésével modellezni tudjuk a társadalmi emlékezetet és ezzel a cselekvéseink viszonzási rendszeren keresztüli összekapcsolását. Az egymásra hivatkozó, egymásba kapcsolódó cselekvési láncok megfelelő stratégiák követésével alakíthatók ki és tarthatók fenn. Ezen stratégiák között a legismertebb a tit-for-tat szabály, aminek megtalálhatjuk a nyomát minden kultúrában. És – legalább a keresztény-zsidó kultúrából – ismert a tit-for-tat stratégia talán legfontosabb hátránya is: a szemet szemért elv könnyen "rálöki" az embert a megtorlási spirál pályájára. A történelem, a mindennapi élet tapasztalatai azt mutatják, hogy ha mindig a tit-for-tat szabályhoz ragaszkodnánk mereven, akkor egyetlen tévedés, kicsiny kis hiba is elegendő lenne ahhoz, hogy az egymásra feleselő cselekvések a dezertálási-büntetési ágra váltsanak át (vagyis a megtorlási spirálba keveredjenek bele).

Axelrod tornájával szemben elég hamar megfogalmazták azt a kritikát, hogy ebből a szempontból nézve súlyos deficittel küzd, mivel nem képes kezelni a társadalmi hibázás lehetőségét. A számítógépes progamok sosem hibáznak. Az éppen adott szabályt minden lépésben pontosan végrehajtanak. Ebben az értelemben nagyon sterilek, és nem képesek modellezni azt a tényt, hogy a valós emberek – bár sokszor követnek valamely stratégiát – néha eltérnek a cselekvési szabály előírásaitól. A számítógépes szimuláció során a tit-for-tat stratégia így sosem kerülhet a megtorlási spirálba, míg a valós életben egy ember, aki általában követi a tit-for-tat szabályt, olykor-olykor mást léphet, mint amit a szabály szerint éppen lépnie kellene.

Ezt a hibázási lehetőséget a játékelméleti modellekbe úgy "vehetjük fel", hogy bevezetjük a zajos játékok fogalmát [274]. Ez pusztán csak annyit jelent, hogy amikor az egyes stratégiákat követő játékosok közöti játéksorozatot szimuláljuk, akkor a bizonyos valószínűség mellett mintegy "kívülről beavatkozva" megváltoztatjuk egy-egy lépés tartalmát (tehát ha valaki az adott szabály szerint kooperálna, annak a lépést dezertálásra váltjuk, és fordítva). Ezzel beviszünk egy hibázási faktort a rendszerbe, és a nagy kérdés az lesz, vajon milyen stratégiák lesznek sikeresek ilyen feltételek között.

A válasz tulajdonképpen nem meglepő. Olyan stratégiák képesek kezelni a megváltozott – tegyük hozzá: a valóságot jobban modellező – helyzeteket, amelyek egyrészt jóindulatúbbak a tit-for-tat stratégiához képest, másrészt tanulni képesek a múlt – jó és rossz – döntéseiből. Az előzőre a nagylelkű tit-for-tat (Generous Tit-For-Tat – GTFT) szabály a példa [539], az utóbbira a pavlovi ("győztes csapaton ne változtass" elvet megvalósító) stratégia [535].

A GTFT-t azért mondják nagylelkű tit-for-tat szabálynak, mert – a stratégia követésébe egy belső valószínűségi változót csempészve – bizonyos pontokon, amikor dezertálnia

kellene az előírás szerint, mégis kooperál. Néha nagylelkűen "elnézi" a partner dezertálását. Ez a szabálymódosítás elég lehet ahhoz, hogy a társadalmi hiba miatt beinduló - és egyébként végtelenné váló – megtorlási spirált meg lehessen szakítani, és vissza lehessen téríteni a kooperálási ágra a játék menetét. A pavlovi (vagy más néven Win-Stay-Lose-Shift) stratégiát már bemutattam korábban, a szabály logikáját egy korábbi fejezetben leírtak alapján lehet megérteni (↬359.o.).

9.6 A játékok közösségi kiterjesztése: reputáció

Említettem, hogy a direkt reciprocitás jelensége csak kis létszámú társadalmakban működik, hiszen az a feltétel, hogy a résztvevőknek ismerniük kell egymást, nem enged meg nagyobb létszámot. Viszont a helyi közösségek növekedés sok helyen nagyobb létszámú társdalmakat eredményezett, ám ezekben a társadalmakban is megtalálhatóak a kooperáció mintázata, itt is van reciprocitás. A kérdés tehát az, hogy miként, milyen formában, miért alakulnak ki mindezek?

A társadalmi modellünket ezen a ponton újra csak bővítenünk kell, mégpedig egy nagyon fontos társadalmi jelenséget kell felvennünk: a **reputáció** mint **társadalmi értékelés** jelenségét. Ahhoz pedig ezt modellezni tudjuk, szükségünk lesz még arra is, hogy a 'közösséget' mint az értékelést végző új "szereplőt" is felvegyük a modellbe a játékosok mellé.

---

a reputáció fogalma már korán felbukkan: D. Kreps és R. Wilson Reputation and imperfect information [425]

9.6.1 Indirekt reciprocitás játékok

A direkt reciprocitás jelensége kialakítható és fenntartható azzal, hogy kooperálás és dezertálás (büntetés) valamilyen egyvelegéből kialakított stratégiát érvényesítünk. Az ilyen feltételes stratégiák azonban csak akkor és azon az alapon működnek, ha feltesszük, hogy a játékosok ismerik egymást. A kooperációt azonban megtalálhatjuk olyan ismétlődő társadalmi helyzetekben is, amelyekben idegenek kerülnek kapcsolatba egymással. Meg kell tudnunk magyarázni az olyan eseteket is, amikor a játékok ismétlődnek, de a játékosok ismeretlenek egymás számára. Az ilyen helyzeteket nevezzük **indirekt reciprocitásnak**.

magjátékok	
aszimmetrikus játékok	szimmetrikus
többalternatívás játékok	kétalternatívás
többszereplős játékok	kétszereplős
nem szigorúan rendezett játékok	szigorúan rendezett
aszinkron játékok	szinkron
ismételt játékok	ismételt
direkt reciprocitás játékok	ismerős játékosok
indirekt reciprocitás játékok	ismeretlen játékosok
kooperatív (hangos) játékok	nem kooperatív (néma)

Az indirekt reciprocitás megragadásához nem elég az eddig fogalomkészlet, szükségünk van egy rendkívül fontos társadalmi jelenség, az értékelés, a reputációképzés modellbe

emelésére. Ez az értékelési jelenség a reputáció, ezért az indirekt reciprocitás elmélet megnevezés mellett előfordul, hogy a kutatók a reputációs elméletek kifejezést használják.

Innentől kezdve némiképp megváltozik a terminológia is. A cselekvők jelzésére eddig használt *Ego* és *Alter* nevek helyett a **Donor** és a **Recipiens** terminusokat alkalmazzuk. Ez a terminológiai váltás arra utal, hogy a reputációs elméleteken belül a kooperáció voltaképp az ismeretlenek megsegítését jelenti. Eddig az volt a kérdés, hogy 'miért kooperál *Ego* az Alterrel szembe kerülve', mosttantól viszont azt vizsgáljuk, hogy 'miért nyújt segítséget a *Donor* a *Recipiens* számára'. A *Donor* terminus aktívabban fejezi ki a kooperálás, segítségnyújtás tényét, elvárását, lehetőségét, míg a *Recipiens* terminus jóval inkább utal e szerep passzivitására az *Alter* terminushoz képest. *Donornak* és *Recipiensnek* speciálisabb a jelentése, viszont "messzebb vannak" egymástól, a két szerep sokkal kevésbé kicserélhető *Ego-Alter* szereppárhoz képest.

Az indirekt reciprocitás, a reputáció modelljében fel kell vennünk egy teljesen új szereptípust is, az értékelést, a reputációképzést végző közösség fogalmát. Erre a reputációs elméletek kevéssé, de legalábbis nem eléggé reflektáltak, pedig nagyon fontos ez a szerep. Hogy miért is? A direkt reciprocitás feltételei között a kooperáció kialakulásához még "csak" annyi kellett, hogy a játékosok emlékezzenek mind a saját, mind a társaik korábbi cselekedeteire, majd ezek alapján végrehajtsák azt, amit a stratégiájuk "előír". A gond az, hogy az indirekt reciprocitás helyzetében semmilyen információ nem áll rendelkezésre a partner korábbi cselekedetéről, hiszen ismeretlenekről van szó. Itt is kell valami fogódzó a döntés meghozatalához, a kérdés csak az, hogy mi lehet az. Megoldás lehet, ha valamilyen értékjelzést ragasszunk minden játékoshoz, ami orientálhatja a *Donort* abban, segítsen-e vagy sem partnerének. Ez az értékjelzés a **reputáció**, amit minden játékoshoz egyedi módon hozzá kell rendelni, hiszen ismeretlenek között sosem tudhatjuk, kivel hoz össze a sors, *Donorként* kikkel kerülhetünk olyan helyzetbe, amikor döntenünk kell arról, segítsünk-e neki. Ha mindenkin van értékjelzés, ha mindenkinek van reputációs címkéje, akkor ennek segítségével már képesek lehetünk dönteni. A *Donor* cselekvési stratégiájában nem a korábbi cselekvéseket, hanem a reputációs értékeket kell figyelembe venni.

A reputáció kezelésére egy "fekete-fehér világképen" alapuló mechanizmust működtethetünk. Elég azt mondani, hogy az emberek vagy jók (Good – G) vagy rosszak (Bad – B), és minden értékelés során e két értékből lehet választani és valamelyiket a *Donorhoz* rendelni. Talán meglepő, de majd látni fogjuk, hogy ez a primitív megoldás meglepően jól modellezi a társadalmi értékelés valós folyamatait.

Egy fogalom jelentését azonban át kell értelmeznünk: a stratégia fogalmát ki kell terjeszteni. Stratégia fogalma alatt eddig azt cselekvési szabályt értettük, amely alapján a játékos meghatározhatta, hogyan döntsön, miket lépjen a játéksorozat bármely pontján. Erre a viselkedési szabályra továbbra is szükség lesz, viszont emellett a stratégiának ki kell terjednie még arra is, hogy miként lehet meghatározni a játékosok reputációját. Ehhez szükség van valamilyen értékelési szabály (mechanizmus, módszer) alkalmazására. Ezt az értékelési szabályt azonban már nem a *Donor*, hanem a közösség veszi birtokba. Az a közösség, amely 1) "figyeli" azt, hogy a *Donor* saját cselekvési szabálya alapján miként dönt arról, segíti-e a *Recipienst* vagy nem, és ez alapján a közösség 2) az értékelési szabály segítségével értékeli *Donor* cselekedetét, vagyis felülvizsgálja a *Donor* reputációs értékét.

Az indirekt reciprocitás stratégiafogalmát kiterjesztve mindig tudnunk kell, hogy milyen viselkedési szabály szerint dönt a *Donor*, illetve milyen értékelési rendszer alapján

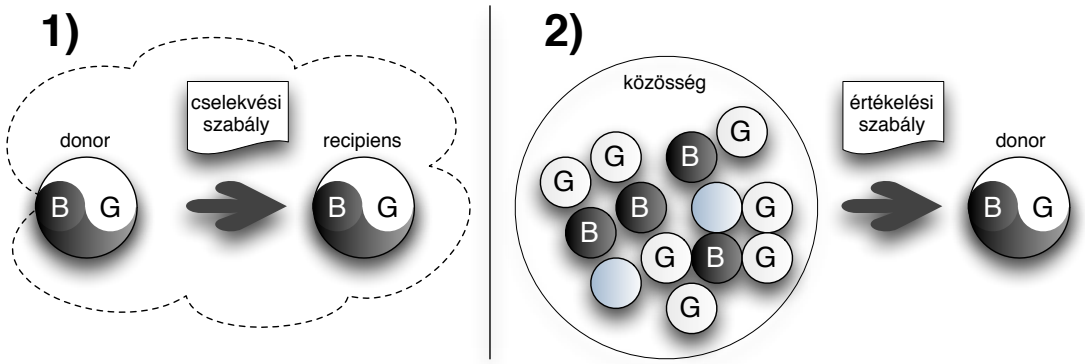


Figure 9.5: a stratégia két összetevője: cselekvési és értékelési szabály

értékeli (újra) a *Donor* reputációját a közösség a *Donor* cselekvése alapján. A cselekvési szabályokról bővebben írtam másutt, itt elegendőnek tartom hivatkozni az ott leírtakra [722]. Az értékelési mechanizmusok áttekintését viszont annyira fontosnak tartom, hogy itt röviden bemutatom a legfontosabb reputációs rendszereket.

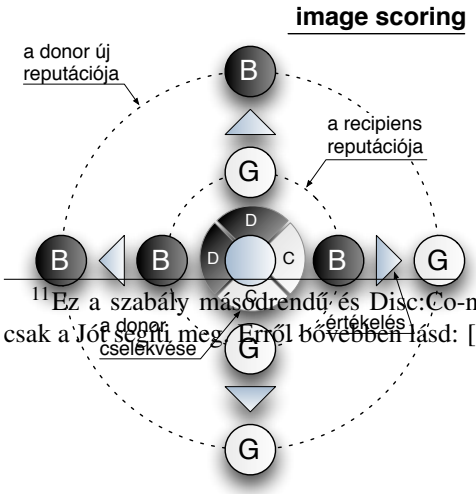
Első körben a négy másodrendű értékelési rendszert veszem sorba, majd utána bemutatom azt, hogy a harmadrendű értékelési rendszereket hogyan lehet modellezni a már ismert másodrendű rendszerek segítségével. A másodrendű értékelési szabály fogalma azt jelenti, hogy az értékelés során két dolgot veszünk figyelembe: a *Donor* cselekvését, illetve a *Recipiens* reputációját. A harmadrendű értékelés ezeken felül figyel a *Donor* reputációjára is [722].

Az indirekt reciprocitás, a reputáció típusainak megragadásához további formulákat kell felvennünk a formális modellünkbe. A

$$reputation = \text{VALUE}(agent,t) = Good(agent,t) \mid Bad(agent,t) = G \mid B \tag{IL.9}$$

Kép-mutató (Image Scoring) értékelési rendszer

Az indirekt reciprocitás első elmélete az image scoring módszertant mutata be, mint a reputációs rendszerek prototípusát [537]. Szigorú módszertani szempontok szerint ez az értékelési rendszer nem is másod-, hanem elsőrendű, de ez nem változtat azon a tényen, hogy ebben a minőségében is az elméletileg lehetséges négy másodrendű értékelési rendszer egyik típusát jelenti.



A többihez hasonlóan ez a módszertan is azzal indít, hogy mindenkinek 'jó' minősítése van a játékban. A cselekvési szabály egyszerű: Jóval szemben kooperálni, Rosszal szemben dezertáni kell.[11] Az Image scoring módszer lényege az, hogy minden egyes döntés után a *Donor* képe (imidzse) aszerint változik, hogy segítette-e a *Recipienst* vagy nem. Ha a *Donor*

[11] Ez a szabály másodrendű és Disc:Co-nak hívják azért, mert diszkriminatív módon kooperál, hiszen csak a Jót segíti meg. Erről bővebben lásd: [722].

kooperált, vagyis segített, akkor onnatól Jó lesz (akármi is volt addig), dezertálása esetén pedig, vagyis amikor nem segített, a *Donor* 'rossz' minősítést kap (akármi is volt addig a reputációja).

Az első reputációs elméletről szóló cikkben a szerzők (Nowak és Sigmund) azt hozták ki, hogy ez a módszertan képes stabil kooperációt kialakítani egy közösségen belül, de a későbbi szimulációs számítások árnyaltak ezen a képen. Egyesek olyan – katalizáló – szerepet tulajdonítanak ennek a stratégiának, mint amilyet szerepet a TFT-stratégia tölt be a direkt reciprocitást szimuláló modellekben [538]. Mindenesetre születtek később olyan publikációk is, amelyek az Image scoring módszer tényleges érvényesülését mutatták ki az állatvilág bizonyos szegmenseiben.[12]

A módszer formális reprezentációja:

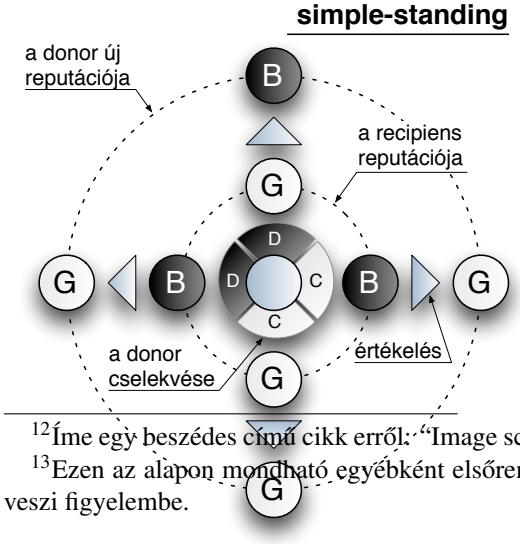
$$\text{VALUE'}(Ego, t_i) = \text{VALUE'}(\text{DO'}(Ego, t_i)) \tag{IL.10}$$

kép-mutató (HUN) – Image Scoring (ENG)

Az Image scoring módszert nemcsak azzal támadták később, hogy sok esetben nem képes a kooperáció tartos kialakítására, de "érték" alapon is kritizálták sokan. Ez a módszer ugyanis "érzéketlen" a *Recipiens* reputációjára.[13] Márpedig – szóltak az ellenérvek – ha nem segít valaki a *Recipiensnek*, az nem feltétlen jelent rosszat, amiért a *Donort* reputációját "le kéne húzni". Mert mi van akkor, ha a *Donor* büntetésből nem segít a rossz *Recipiensnek*? E kérdésből kiindulva kezdték el keresni a további értékelési rendszereket.

Egyszerű hírnévápoló (Simple Standing) értékelési rendszer

A Simple Standing értékelési rendszer abban tér el az Image Scoring módszertől, hogy a rossz *Recipienssel* szemben megengedi a büntetést. Ez abban nyilvánul meg, hogy ha a *Donor* egy rossz *Recipiensnek* nem segít (büntetésből), akkor ezt a Simple Standing módszer helyesli és a *Donort* ezek után Jó minősítéssel látja el.



Ez a módszertan az egyik legtöbbet sulykolt értékelési rendszerünket modellzi a valós társadalmi életünkből. A gyerekeinket úgy neveljük fel (és felnőttkorunkban is ezerszer elmondjuk magunknak, másoknak, egymásnak) azt a mondatot, amely a Simple Standing módszertant megalapozó elvnek tekinthető:

„a rossz elnyeri méltó büntetését"

A büntetés igazolását adja meg ez az értékelési rendszer, ezért olyan fontos. A mindennapi

[12] Íme egy beszédes című cikk erről: "Image scoring and cooperation in a cleaner fish mutualism": [127].

[13] Ezen az alapon mondható egyébként elsőrendűnek ez a módszertan, hiszen csak a *Donor* cselekvését veszi figyelembe.

narratíváink egyik legfontosabb összetevője ez az elv. Ezáltal tudjuk nagy létszámú társadalmakban is fenntartani a kooperációs mintákat, mert lehetővé teszi a büntetés intézményét az ismeretlenekkel, idegenekkel szemben úgy, hogy a büntetésben megnyilvánuló – elsőrendű – dezertálást társadalmilag igazoltnak, megengedhetőnek minősítsük.[14]

Ez a stratégia a Standing metódus másodrendű változata [546], [547], [548]. A különbség köztük abban van, hogy ez az egyszerűsített változat jónak minősíti a rossz személy visszautasítását akkor is, ha a *Donor* maga előtte rossz volt (az eredeti hírnévgyűjtő stratégia ezt a diszkriminációt csak a jó *Donorokkal* szemben alkalmazta). Ez annyit jelent, hogy ez a szabály jónak tartja azt is, ha a hipokrita játékos bünteti a Rosszat. A másodrendű változat alkalmazásának "célja" az értékelési feladat egyszerűsítése lehet, hiszen ebben az esetben kevesebb információt kell beszerezni és feldolgozni.

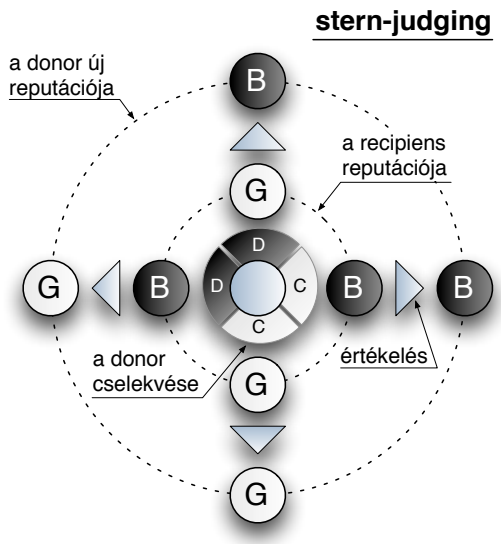
A szabály formulája:

(IL.11) $\text{VALUE'}(Ego,t_i)=(\text{VALUE'}(Alter,t_i)=G\rightarrow\text{VALUE'}(\text{DO'}(Ego,t_i)))\wedge(\text{VALUE'}(Alter,t_i)=B\rightarrow G)$

egyszerű hírnévápoló (HUN) – Simple Standing (ENG)

Kérlelhetetlen (Stern Judging) értékelési rendszer

A kérlelhetetlen (Stern Judging) értékelési rendszer még egy lépést tesz a büntetés irányába a Simple Standing módszertanhoz képest. Az utóbbi ugyanis egy ponton még "elnéző" a Rosszakkal szemben, és ezen a ponton lehet még szigorítani a rossz *Recipiensekkel* szembeni fellépésen.



A Stern Judging [558] megegyezik egy jóval korábban megfogalmazott elképzeléssel, amit kandori-módszernek neveztek el [391]. A stratégia két szabályból áll:

(i) nő a *Donor* reputációja, ha segíti a Jót és nem segíti a Rosszat, illetve

(ii) csökken a *Donor* reputációja, ha segíti a Rosszat és visszautasítja a Jó támogatását.

Ez a módszertan is egy tényleges társadalmi gyakorlatban létező elv reprezentációjaként fogható fel. Vannak társadalmi állapotok, amikor sokan a „rossz elnyeri a métó büntetését" elvnél is erősebb elvárásaink vannak a *Donorokkal* szemben, és megköveteljük tőlük,

[14] Idegeneken itt azokat kell érteni, akiknek a cselekvéseire nem látunk rá, ezért csak az általános reputációs címke alapján döntünk arról, hogy segítsünk-e nekik vagy sem.

hogy ne csak büntessék a Rosszakat, de ne is segítsék őket. Ezt fejezi ki az alábbi mondás:

„vétkesek közt cinkos, aki néma”

Ez az elv azért szigorúbb az előzőnél, mert a kooperációt (ami önmagában eddig jónak minősült) nem tartja elfogadhatónak, helyeselhetőnek akkor, ha az rossz *Recipiensre* irányul. A mondás azt fejezi ki, hogy vannak olyan helyzetek, amelyekben a hallgatás a Rosszak között beleegyezést, elfogadást, tehát kooperációt jelent, és ez – e módszertan szerint – elfogadhatatlan, tehát büntetendő.

A stratégia másodrendű értékelésen alapul, és csak annyiban tér el a hasonló – harmadrendű – ‘ítélkező’ (judging) stratégiától, hogy azt is jutalmazza, ha Rossz büntet Rosszat, ami annyit jelent, hogy a hipokrita magatartást is jónak ítéli. Mindegy, ki büntet (Jó vagy Rossz), csak büntesse meg a Rosszat.

Formula:

(IL.12) $\text{VALUE'}(Ego,t_i)=(\text{VALUE'}(Alter,t_i)=G\rightarrow\text{VALUE'}(\text{DO'}(Ego,t_i)))\wedge(\text{VALUE'}(Alter,t_i)=B\rightarrow\neg\text{VALUE'}(\text{DO'}(Ego,t_i)))$

kérlelhetetlen (HUN) – Stern Judging (ENG)

Elkerülő (shuning) értékelési rendszer

A Shunning értékelési rendszer próbálja továbbcsavarni a “büntetési faktort”, ami első pillanatra talán nehezebben vehető észre, ha a módszertan ábrájára tekintünk, mert a Stern Judging módszertanhoz képest annyi a változás benne, hogy a Rosszal szemben alkalmazott büntetést (nem segítést) most rossznak minősíti (ezzel “törli” a Simple Standing módszer által bevezetett szabályt a rossz megbüntetésének helyességére vonatkozóan). A módszertan kevesebb figyelmet kapott a többi három másodrendű értékelési rendszerhez képest, pedig fontos társadalmi jelenség modellezésére alkalmas [560], [738], [548].

A Shunning stratégia logikája a következő:

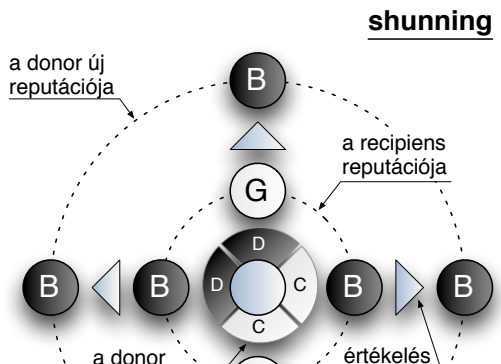
(i) mindig bünteti a dezertálást és

(ii) minden esetben bünteti a *Donort*, ha az Rossz személlyel kerül kapcsolatba – még akkor is, ha egyébként a *Donor* segíti a *Recipienst*

A stratégia másodrendű értékelést használ és a szimulációs kísérletek eredményei szerint hatékonyan hozzájárul a kooperációs minta elterjedéséhez. Ez a stratégia kicsit arra a közkeletű mondásra emlékeztet, amely szerint

„aki korpa közé keveredik, azt megeszik a disznók”

vagyis tehet az ember bármit, akár kooperál, akár dezertál, ha Rosszak közé kerül (értsd: rosssz *Recipiensekkel* kerül kapcsolatba), menthettlenül Rosszá válik ő maga is.



A módszer névválasztásával azt akarták kifejezni, hogy a stratégia a Rosszak kizárására, elkerülésére ösztönözhet, mivel a velük kapcsolatos bármilyen érintkezés “káros” lehet a *Donor* számára. Első hallásra talán meglepő,

nehezebben értelmezhető, elfogadható ez az értékelési módszer, pedig a Rosszakkal való érintkezés büntetése meglehetősen gyakran előfordult a történelemben [722], sőt, várhatóan a jövőben is találkozni fogunk ezzel a jelenséggel. Ez az értékelési rendszer a kizárásra törekvő, a szegregációs szándékok mélyén meghúzódó logika. Vannak társadalmi állapotok, amikor társadalmi csoportok között túl naggyá válnak a társadalmi, kulturális távolságok, és elérkezhet egy pillanat, amikortól az egyik közösség hirtelen átvált a Shunning módszertanra, mert a közösségközi kofliktusokat úgy vélik feloldhatónak, ha a két közösséget teljesen elszeparálják egymástól. Ilyenkor veszik elő ezt a módszertant. A szimulációs számítások egyébként ezt az értékelési rendszert kooperációs minták kihordására alkalmatlannak szokták minősíteni (és a történelmi tapasztalatok sem igazán arról szólnak, hogy a szegregáció képes lett volna megoldani a közösségközi konfliktusokat).

Formula:

(IL.13) $\text{VALUE'}(Ego,t_i)=(\text{VALUE'}(Alter,t_i)=G\rightarrow\text{VALUE'}(\text{DO'}(Ego,t_i)))\wedge(\text{VALUE'}(Alter,t_i)=B\rightarrow B)$

elkerülő (HUN) – Shunning (ENG)

A négy értékelési szabály áttekintése

Az értékelési szabályok bemutatása után érdemes egyetlen táblázatba rendezve összegezni a fenti négyféle szabályt, hogy jobban szembetűnővé válhassanak a köztük levő hasonlóságok és különbségek.

értékelési szabály ↓	$\text{VALUE'}(Alter,t_i)\rightarrow$	G		B	
	$\text{DO'}(Ego,t_i)\rightarrow$	C	D	C	D
Image Scoring	$\text{VALUE'}(Ego,t_i)\rightrightarrows$	G	B	G	B
Simple Standing		G	B	G	G
Stern Judging		G	B	B	G
Shunning		G	B	B	B

A táblázatból könnyen kiolvasható a négy szabály közös vonása. A *G* jó partnerrel szemben mindegyik értékelési módszer ugyanazt írja elő: ha a jó *Alterrel* szemben *Ego* kooperált, akkor *G* jó minősítést kap, ha dezertál, akkor *B* rosszat. A különbség a *B* rossz partnerrel szembeni cselekvés utáni értékelésben rejlik.

Harmadrendű stratégiák dekomponálása

4096 elméletileg lehetséges harmadrendű stratégia létezik ebből összesen 8 olyan van, amely kooperáció kialakítására és fenntartására képes

Ez visszavezethetőek a már ismert másodrendű szabályokra egy egyszerű dekomponálási technika alkalmazásával

átemelni a két tánlázatot a dióból

Ez a tény tovább növeli a másodrendű értékelési rendszerek jelentőségét.

Társadalmi hiba és indirekt reciprocitás

Generous Stern Judging Rosas cikkben van hivatkozás [610]

spite és indirekt reciprocitás

altruizmus és indirekt reciprocitás mellett

agresszió (spite) és indirect reciprocitás kapcsolata: Evolution of spite through indirect reciprocity [383]

9.6.2 Közös minták a direkt és indirekt reciprocitás sikeres stratégiáiban

Korábban bemutattam, hogy mind a direkt, mind az indirekt reciprocitás feltételei közt négy-négy másodrendű stratégia (cselekvési, illetve értékelési szabály) képes valamilyen formában és mértékben hozájárulni a kooperáció kialakulásához. A stratégiákhoz rendelt formulák hasonlósága alapján a cselekvési és értékelési szabályok párba kapcsolhatók egymással. Az egyszerű *TFT* és az *Image Scoring*, a *bűnbánó* és a *Simple Standing*, a *Pavlov* és a *Stern Judging*, valamint a *haragtartó* és a *Shunning* szabályok ugyanazt a mintázatot mutatják, logikájukat tekintve hasonlóképpen működnek [610]. Ha a direkt reciprocitáshoz tartozó stratégiák esetén a feltételes cselekevési szabályokban kicseréljük a kooperálásra, illetve dezertálásra utaló *C*, illetve *D* jelzéseket *1*, illetve *0* értékekre, valamint az indirekt reciprocitás értékelési szabályait összegző táblában szereplő *G*, illetve *B* jelek helyett szintén *1*, illetve *0* értékeket írunk, akkor a kétféle szabálynégyest egy táblában reprezentálhatjuk.

ACRs ↓	C		D		$\leftarrow \text{DO'}(Ego, t_{i-1})$
	C	D	C	D	$\leftarrow \text{DO'}(Alter, t_{i-1})$
Simple TFT	1	0	1	0	Image Scoring
Contrite TFT	1	0	1	1	Simple Standing
Pavlov	1	0	0	1	Stern Judging
Grim Tigger	1	0	0	0	Shunning
$\text{DO'}(Ego, t_i) \rightarrow$	C	D	C	D	↑ ASRs
$\text{VALUE'}(Alter, t_i) \rightarrow$	G		B		

A párba kapcsolható szabályok közötti azonos mintázat alapja a büntetéshez való viszony hasonlósága. Mind az *Image Scoring*, mind a *TFT* elsőrendű szabály. Saját "terepükön" ugyanaz a szerepük, mindkettő ugyanazt a funkciót tölti be:

> „We expect, however, that Scoring has a similar role in indirect reciprocity to that of tit-for-tat in direct reciprocity. Neither strategy is evolutionarily stable, but their ability to catalyze cooperation in adverse situations and their simplicity constitute their strength." [538].

Mindkét szabály "ellenzi" a büntetést, bár ezt hipokrita módon teszik. Közös elvük az, hogy nem fogadják el büntetést, tehát ha valaki büntet, akkor azt büntetik: 'büntesd, aki büntet'.

Amikor Sugden 1986-ben javasolta saját T_1 módszerét (a *Standing* szabályt) [716], akkor ő nem beszélt reputációról, bár ahhoz közel álló fogalmat használt: 'kiállás' ('standing'). Sugden nem igazán reflektált a kérdésre, de tény, hogy cselekvési és nem értékelési szabályokról beszélt. Jóval később Sugden szabálya *Standing* (*Simple Standing*) néven mint értékelési szabály az indirect reciprocitás elméletek részévé vált. Amikor 1989-ben Boyd újracímkézte Sugden szabályát *bűnbánó tit-for-tat* stratégiának [111], akkor ő is cselekvési szabályként hivatkozott rá. Ez a kettős kezelésmód persze valamiféle terminológiai zavar, fogalmi pongyolaság jele, de mindebben az a fontos most, hogy ez az "összemosás" nem véletlen. Az összekeverhetőség ugyanis a *bűnbánó* és a *Simple Standing* szabályok közötti szoros kapcsolatot jelzi. Amikor Frean bevezette a szigorú, de igazságos *Firm-But-Fair* szabályt, akkor a következőpen értelmezte a szabályt:

> „ . . . it is 'firm' in that it retaliates by defecting if it was a sucker in the previous round. It is also 'fair' in that it does not retaliate against a defector if it defected itself, and it cooperates with suckers rather than continuing to exploit them." [269].

Ez az indoklás ahhoz hasonlít, amit a *Contrite TFT* stratégia bemutatásakor már elmondtunk egyszer – bár az érvelésben tetten érhető egyfajta elmozdulás a büntetés irányába. *Ego* úgy értékeli *Alter* előző fordulóbeli dezertálását, hogy az igazolt büntetés volt *Ego*'s saját korábbi dezertálásával szemben, és mivel *Ego* egyfelől egyetért az igazolt büntetés intézményével ('firm'), másfelől igazságos ('but fair'), így *Ego* kooperál *Alter* lépésére válaszul. Közös szlogenjük ez lehetne: 'ne büntesd, aki büntet'.

A *Pavlov* vagy *WSLS* stratégia jellemzésekor többen kiemelték annak tanulási képességét, mi több, ezzel magyarázták a stratégia sikerességét is. Most, amikor ezt a cselekvési szabályt állítjuk párba a *Stern Hudging* értékelési szabállyal, akkor – a büntetéshez való viszonyra fókuszálva – kicsit átértelmezhetjük a *Pavlov* szabály logikáját, hatásmechanizmusát. Rosas úgy értékeli, hogy a *Pavlov/Stern Judging* szabálykettős hízelgő (flattery), alázatos (obsequious) stratégia [610]. Szerintem inkább arról van szó, hogy a *Pavlov* esetében is a *Stern Judging* módszer logikája érvényesül. Mindkét szabály hasonlóképpen értelmezhető. *Ego* úgy értékelheti *Alter*'s kooperálását *Egoval* szemben, hogy ezzel *Alter* tartózkodott megbüntetni *Egot*, aki pedig dezertált az előző fordulóban. A dezertálás büntetése viszont *Alter* kötelessége lett volna, s mivel ezt elmulaszotta, ezért *Ego* ezt a kooperálás helytelennek tartva böntetéssel (dezertálással) sújtotta. Ez volt a *Stern Judging* értékelési szabály logikája, és ez az indoklás érvényes, alkalmazható a *Pavlov* cselekvési szabály esetében is. Éppen ezért lehet mindkettejük receptje a következő szabály: 'büntesd, aki nem büntet'.

Mind a *haragtartó* (*Grim Trigger*), mind a *Shunning* szabály erősen, vagy inkább túlzottan büntetéspártinak mondható. Nem véletlen, hogy a szimulációk során kiderült, hogy ezek a stratégiák nehezen, ritkán vezettek a kooperációs minták elterjedéséhez. A haragtartó stratégia az első dezertálásra végnélküli büntetéssel reagál, azaz bármit is tesz a partnere, a stratégia képviselője válogatás nélkül dezertál. A Shunning módszer is ugyanezt a logikát követi, csak ebben az esetben a partner "rosszasága" váltja ki ezt az egyértelmű elkülönülést. A közös vonásuk tehát az, hogy büntetik mind a kooperációt, mind a dezertálást, ami úgy is értelmezhető, hogy nem az – ilyen vagy olyan – cselekvést, hanem magát a cselekvőt utasítják el. Elvük lehetne: 'büntess mindent'.

A négy stratégiapáros bemutatása alapján megállapíthatjuk, hogy négy stratégiapárhoz négy alapelvet társíthatunk, amelyek mindegyike a büntetéshez való eltérő viszonyulást fejez ki. Ezt összegzem a következő táblában.

szabálypár	közös elv	
egyszerű TFT & Image Scoring	büntesd, aki büntet	$P(P)$
bűnbánó TFT & Simple Standing	ne büntesd, aki büntet	$\neg P(P)$
Pavlov & Stern Judging	büntesd, aki nem büntet	$P(\neg P)$
haragtartó & Shunning	büntess mindent	$P(P \vee \neg P)$

A négyféle fenti logika állandóan megnyilvánul a mindennapjainkban, amikor cselekszünk, döntünk, amikor értékeljük mások cselekvését. Ez a négy logika beépül a személyiségünkbe, a társadalmi helyzeteinkben, a kulturális, szubkulturális emlékezetbe, a közösségek közös tudatába, és helyzetről helyzetre változóan, tehát sohasem determnisztikus módon, de általában valamelyik elv dominanciája mellett befolyásolja a cselekvők magatartását. A tudósi szereptől egy pillanatra elszakadva, a bulvár irányába lépve egyet mondhatjuk, hogy kicsit azok vagyunk, amelyik elvet választjuk magunknak a fenti négyből.[15]

9.6.3 Értékelési rendszerek

shame culture vs. guilt culture

Punishment, Cooperation, and Cheater Detection in 'Noisy' Social Exchange [103]

9.7 Koordinációs problémák, közös tudás, aszimmetrikus információ

lewisi értelemben vett konvenció

9.7.1 Szaliencia, fókuszpont, precedens

szaliens alternatíva vagy fókuszpont (focal point) keresése

szaliencia szembeötlőség, szembeszökőség

Ismert jelenség, hogy a kankutyák gyakran vizelnek sétáltatás közben, és előszeretettel teszik ezt fák tövében. Az etológiai magyarázat erre a területvédés: a szagminták elhelyezésével az általuk "őrzött" területet jelzik (és védik) így a többi "betolakodó" kutya számára. Ebben az egyszerű jelenségben egy komoly koordinációs problémát, feladatot vehetünk észre. Mind az területvédő kutya, mind a "behatoló" kutya számára az lenne kívánatos, hogy utóbbi megtalálja az adott területen elhelyezett szagmintákat. A területvédő kutya "érdekeltsége" egyértelmű, de a behatoló kutyának is fontos lehet az az információ, hogy "itt vannak más kutyák". A kérdés már csak az, hogy a sok lehetséges pont közül hova tegye le a mintákat a védő, és hol keresse azokat a behatoló kutya, hiszen a terület határán erre végtelen számú lehetőség adódik. Ez egyértelműen koordinációs probléma: "közös érdek", hogy ott szagolgasson a behatoló, ahol korábban vizelt a

[15] Nem számítva persze azt a beállítódást, stratégiát, amely a jézusi aktív szeretet parancsán alapul. Ezt a szabályt tudományos eszközökkel még nem elemezték.

területvédő kutya. Ebben a helyzetben jelentenek szaliens megoldást a környék fái, mivel azok egyértelmű, mindkét kutya számára szembeötlő pontot kínálnak.

az amerikai vasút nyomtávja

9.7.2 Közös tudás

9.8 minimal social situation

The minimal social situation (MSS) was first described by the US psychologist Joseph Sidowski in his doctoral dissertation and in articles based on it (Sidowski, 1957; Sidowski, Wyckoff and Tabory, 1956). The two-person or dyadic MSS is a game of incomplete information in which both players know their own strategy sets but neither knows the co-player?s strategy set nor either player?s payoff function. In its most extreme form, both players suffer from a profound and debilitating kind of ignorance: they are oblivious not only of the nature of their strategic interdependence but even of the fact that their decisions are choices in a game of strategy. The payoff matrix of the game generally used to study the minimal social situation, named Mutual Fate Control by Thibaut and Kelley (1959)

The Mutual Fate Control game can be interpreted as an impoverished Prisoner?s Dilemma game, in which the level of measurement in the payoff functions is reduced to a binary scale (Sozanski, 1992, p. 110). This is clarified in Figure 10.2. But the Prisoner's Dilemma game differs from the Mutual Fate Control game in important ways. In the Prisoner?s Dilemma game, there is a unique (pure-strategy, strict) Nash equilibrium at (D, D), representing joint defection. Furthermore, for each player, D yields a strictly higher payoff than C irrespective of the strategy chosen by the co- player, so that D is a strongly dominant strategy. In the Mutual Fate Control game, on the other hand, all four outcomes (C, C), (C, D), (D, C), (D, D) are (pure-strategy, weak) Nash equilibria, and there is no strongly dominant strategy. The characteristic feature of the Mutual Fate Control game, as its name suggests, is that the players? payoffs are unaffected by their own actions and are entirely in the hands of their co-players ? another feature not shared by the Prisoner?s Dilemma game.

[168]

J. W. Thibaut and H. H. Kelley (1959) The Social Psychology of Groups (New York: Wiley).

H. H. Kelley and J. W. Thibaut (1978) Interpersonal Relations: A Theory of Interdependence (New York: Wiley).

9.9 Kollektív cselekvés, kollektív döntés

Társadalmi választások elmélete [494]

közösségi döntési helyzetek: testületi döntések

szavazó testületek, döntési testületek, tanácsadó testületek választás elve kompromisszum elve

9.10 Kooperatív játékok: koalícióelmélet

A kooperatív, illetve nem-kooperatív játékok fogalmát John Nash vezette be. Kétféle módon szokták magyarázni, mit is jelent ez a megkülönböztetés. Egyfelől nem kommunikálhatnak egymással, ami azért fontos, mert nem így nem köthetenek egyezséget a partnereikkel a játékosok annak érdekében, hogy kikerüljenek a társadalmi csapda helyzetekből. Harsányi János hívja fel a figyelmet arra [343], hogy az igazán fontos mozzanat ebben nem az, hogy a kommunikáció révén biztosítható az információcsere, és így egy közösen jobb megoldásra számíthatnak a felek, mert a másik szándékának ismerete, tehát a többletinformáció még nem feltétlen elegendő a jobb megoldás eléréhez. Sokkal fontosabb az, hogy a kommunikáció révén kikényszeríthető elköteleződéseket tehetnek egymás felé a játékosok, amire a nem-kooperatív játékok során nem volt lehetőségük. Fontos, hogy ezek az elköteleződések kikényszeríthetőek, azaz az elköteleződések betarttatására vannak eszközeik a feleknek. Enélkül az egész semmit sem érne.

A kikényszeríthető elköteleződések azonban már átvezetnek minket a normativitás területére, amit a játékelmélet eddigi modelljéből még kizártunk.

átvezetés lehet a koalícióelmélet világa.

Simmel triádelmélete [679]

Somlai Péter cikke erről [701]

Komorita-Meek [409]

David Wilkinson Cohesion and Conflict könyve [816]

neutrális, semleges szereplő megjelenése csak triádkapcsolatokba lehetséges semlegesség megjelenése, kialakulása

Philip C. Jessup and Francis Deák, Neutrality: Its History, Economics, and Laws, Columbia Univ. Press, 1935-36 idézi [647]

magjátékok	
aszimmetrikus játékok	szimmetrikus
többalternatívás játékok	kétalternatívás
többszereplős játékok	kétszereplős
nem szigorúan rendezett játékok	szigorúan rendezett
aszinkron játékok	szinkron
ismételt játékok	egyszeri
direkt reciprocitás játékok	ismerős játékosok
indirekt reciprocitás játékok	ismeretlen játékosok
kooperatív (hangos) játékok	kooperatív (hangos)

(IL.14) ESSENTIAL_GAME $\stackrel{\text{def}}{=}$

esszenciális (HUN) – essential (ENG)

(IL.15) SUPER ADDITIVE_GAME $\stackrel{\text{def}}{=}$

szuperadditív (HUN) – super additive (ENG)

(IL.16) CORE $\stackrel{\text{def}}{=}$

mag (HUN) – core (ENG)

(IL.17) CORE $\stackrel{\text{def}}{=}$

mag (HUN) – core (ENG)

Az Arrow-féle lehetetlenségi tétel újraértelmezése kooperatív játékok világában. Ordeshook [552], Ordeshook [1992], Moulin [1995], Shapley-Shubik [678] [180]

Minden esszenciális, szuperadditív, karakterisztikus függvény formájában felírt állandó összegű játék magja üres.

9.11 Társadalmi dilemmák feloldási lehetőségei

boudon, nem szándékolt társadalmi következmények [106] perverz effektusok

Bernholz-tétel +

Externalities as a Necessary Conditions for Cyclical Social Preferences [75]

pozitív externália potyautas magatartás

negatív externália tiltakozás, kilépés

Chapter 10

Társadalmi pozíció

„A jogtudományt tanuló angol diákok döntő többsége a jog alapvető kategóriáival való ismerkedést Hohfeld ‘Alapvető jogi fogalmak’ című tanulmányával kezdi.” [345, 171]

Ebben a fejezetben a normativitás jelenségét szeretném megalapozni úgy, hogy még nem lesz szó a norma fogalmáról. Nem tudom persze elkerülni, hogy ne használjam olykor a norma, a jog, normativitás, legalitás fogalmait, de ezeket majd csak a következő fejezetben fogom pontosan definiálni. A norma jelenségének leírásához, a fogalom megalapozásához arra van szükségünk, hogy a deontikus logikát bizonyos értelemben kiterjesszük, és az eddig megismert imperszonális deontika helyébe felépítsük a perszonális deontikus logika elméletét. Ehhez fel kell használni a Wesley Newcomb Hohfeld jogértelmezését, és fel kell építeni a hohfeldi elmélet formális logikai reprezentációját. Ehhez a skandináv logikai iskola által kijelölt úton kell végigmennünk, miközben olykor a fogalmi modell és a formalizálási javaslatok korrekciójára, olykor pedig kiterjesztésére van szükség.

10.1 Perszonális deontika

A logikai alapozó fejezetben jeleztem (↬**??**.o.), hogy a deontikus logika évtizedeken keresztül személytelen, imperszonális módon építkezett. Megkockáztatom, hogy az imperszonális megközelítés dominanciája is hozzájárult ahhoz, hogy a deontikus logika hosszú időn át nem tudta beváltani a hozzá fűzött kezdeti reményeket. Sokan várták ugyanis, hogy a deontikus logika fejlődése során hamar eléri az a szintet, amikor már megfelelő alapokat képes nyújtani jogi, normatív szövegek formális leírásához, a jogi szövegek konzisztenciájának vizsgálatához és ezáltal jogi szövegek számítógépes feltárásához. Ez azonban nem következett be. Nem is nagyon következhetett be, mert amíg a deontikában nem lehetett személyközi viszonyokat kifejezni, kezelni, addig nem lehetett megragadni a jogban, normativitásban alapvető jogi, normatív viszonyokat. A társadalmi kapcsolatok feltárásához szükség volt arra, hogy az ágensek közti kapcsolatokat is kezelni lehessen a deontikán belül.

A deontikus logika a deontikus modalitások, a kötelezések, tiltások, megengedések, eltekintések logikájával foglalkozik. Amíg imperszonális deontikát "űzünk", azaz amíg egyetlen ágenst feltételezünk, amíg egyetlen deontikus operátoron van a fókuszunk, amivel ki tudunk fejezni egy normatív, deontikus állítást, addig nem tudjuk megragadni a jognak (normativitásnak) azt a minőségét, alapvonását, hogy a létező jogi (normatív) rendszerekben jogi (normatív) viszonyokat kell kezelnünk. Ez a deontika nyelvén azt jelenti, hogy a deontikus operátorokkal kifejezhető normatív aktusokat mindig "párosával", egymással összekapcsolódva értelmezhetjük.

A jogtudománynak már nagyon régóta az egyik alapvetése ez a tétel [394]. Ahol valakinek jogosultsága van valamire, ott van valaki más, akinek valamilyen kötelezettséggel rendelkezik az adott jogosultsághoz kapcsolódóan. Vagy fordítva: a kötelezettségek mindig mások jogosultságával járnak együtt. W.N. Hohfeld egyetértőleg idézi az alábbiakat:

> „Kötelesség vagy jogi kötelezettség az, amit valakinek meg kell tennie vagy nem szabad megtennie. A 'kötelesség' és a 'jog' korrelatív kifejezések. Ha megsértenek egy jogot, megszegnek egy kötelességet is." [360, 73]

A jogbölcseletben vita van arról, vajon csak általában vagy mindig érvényes ez a fajta komplementaritás. Vannak nagy hatású jogelméleti gondolkodók, akik vitatják a korrelativitás teljeskörűségének tézisét. Joseph Raz például így ír:

> „Egy népszerű nézet szerint minden joghoz tartozik egy korrelatív kötelezettség. Minthogy a jog a kötelezettség egyik alapja, jó adag igazság van az efféle korrelativitás-tézisben. A tézis legtöbb szokásos megfogalmazása mégis félrevezető." [597, 166]

Ebben a vitában nem tudok, nem is akarok állást foglalni, és mivel a fomális modell építésének jelen szakaszában nem is igazán lényeges ez a kérdés, el lehet tekinteni e szempont figyelembe vételétől. Lehet, hogy igaza van Joseph Raznak, és nem áll mindig a korrelativitás tézise, de azt Raz sem tagadja, hogy sok esetben igaz. Ha viszont azt állíthatjuk, hogy a jognak, normativitásnak vannak olyan területei, ahol diadikus viszonyokat, azaz jogi relációkat kell kezelnünk, akkor a jogi fogalmak fent hivatkozott kettősségét, korrelativitását valahogy be kell építeni a fogalmi modellünkbe. Ezt a relációs jelleget úgy tudjuk megragadni, ha a *imperszonális* szemléletmódot felváltjuk a *perszonális* megközelítéssel.

A továbblépés előtt azonban szükséges figyelmeztetnem arra, hogy az elemzendő fogalmak, terminusok mindennapi használatában sajnos sokfajta és elég széles körben elterjedt pongyolaság figyelhető meg. A terminusokat sokszorosan túlhasználják, vagyis egy-egy terminust – sokszor akaratlanul – többféle jelentéssel vesznek használatba, és ez sok esetben nagyon könnyen fogalmi zavarokhoz, félreértésekhez vezethet. Érdemes ezért itt egy-két előzetes megjegyzést tennem.

A jog a deontikus logika operátorait kifejező terminusok közül a kötelezést és a tiltást kiterjedt módon alkalmazza, ám a megengedés terminusát ritkábban veszi igénybe (ha egyáltalán használja). Sokkal elterjedtebb a jog, jogosultság terminusa. Ennek ellenére nyugodtan rögzíthetjük, hogy a jog, jogosultság, megengedés terminusai mind ugyanahhoz a jelentéstartományhoz sorolhatók, mint ahová a megengedés is tartozik. Ebből persze az is következik, hogy akárcsak a megengedés kategóriája esetében, úgy a jogosultság mindennapi használatában is keveredik az a két lehetőség, hogy a nemkötelezés vagy a nem-tiltás állapotát fejezzük-e ki a terminussal. Mivel a hétköznapi fogalomhasználati gyakorlatba mindkét értelmezés mélyen beleivódott, értelmetlen lenne

bármelyik interpretációt elsődlegessé tenni a másik ellenében, így csak arra törekedhetünk, hogy tudatosan figyelve erre a problémára, a terminusok használata során mindig egyértelművé tesszük, az adott pillanatban éppen melyik értelmezés tartjuk érvényesnek. A szinonimitásról beszélve megállapíthatjuk még, hogy a kötelezettség és kötelesség terminusait is egymással felcserélve lehet használni.

Sajnos létezik egy másik terminológiai pongyolaság jogi alapfogalmak területén, ami a laikusok számára még több problémát, félreértést okozhat – legalábbis a magyar nyelvben. A 'jog' terminusa ugyanis többféle értelemben használható. A jog egyfelől jelenthet jogosultságot, illetve másfelől normatív megnyilatkozást, és az utóbbi jelenthet egyetlen normatív kijelentést (mint jogszabály, jogtétel), de jelentheti akár az egész jogi rendszert is. Az igazi megkülönböztetést a jog mint jogosultság és a jog mint normatív rendszer (tételes jog) között kell megtennünk, a jogtétel és a tételes jog közti különbség inkább mennyiségi, nem annyira minőségi. Természetesen a jogtudomány réges rég elvégezte ezt a megkülönböztetést, hiszen az 'alanyi jog' (facultas agendi) és a 'tárgyi jog' (norma agendi) fogalmai lényegileg erre a kettősségre reflektálnak. Ami nagyon fontos (és ez sokak számára szokatlan, meghökkentő lehet), fogalmilag szükségünk van a 'jogi jog' és 'erkölcsi jog' elhatárolására egymástól, amit csak a "jog" fogalmának pontos explikációja, sokszálú értelmezése révén tehetünk meg. Amikor a későbbiekben a norma fogalmát fogom tipizálni a különböző fogalmi összetevők alapján (453.o.), akkor a különböző tipizálási lehetőségek egymás utáni kibontásával jobban lehet majd érzékeltetni, hogy valóban értelmes kategória akár a 'jogi jog', akár az 'erkölcsi jog' fogalma.

Másfajta fogalmi zavar jelentkezik a jogi viszonyok minősítésében. Itt nem terminológiai poliszémikus jellegű túlhasználásról, hanem fel nem ismert szinonimitásról van szó. Amikor a jogtudósok az összetartozó jogokról és kötelességekről beszélnek, akkor a köztük levő viszonyra a 'korrelativitás' terminusát alkalmazzák. Wesley Newcomb Hohfeld is a korrelativitás kategóriáját használta. A mai napig bevett ez a szóhasználat a jog világában. Később elemezni fogom a szociológusok által bevezetett társadalmi szerep fogalmát, szólok arról, hogy a társadalmi szerepek normatív összetevői közt ugyanúgy jogokat és kötelességeket találhatunk, mint a jog esetében. A szociológusok azonban az általuk vizsgált normatív komponensek közt megfigyelhető viszonyt más kategóriával ragadják meg, ők erre a komplementaritás terminusát használják (lásd Sir Ross [612], Gouldner, Parsons [562]). A korrelativitás terminusa helyett Stig Kanger már nem csak egy, de két másik kategóriát használ [394]. Később bemutatom, hogy Kanger milyen alapon minősíti hol inverz, hol konverz relációnak két jogviszony közt létező korrelativitást. Jogi szövegekben előfordul olykor a jogi viszonyok 'reciprocitása' kifejezés is, ami szintén a korrelativitás, komplementaritás jelenségéhez kapcsolódik.

Nagyon hasonló tartalom mentén van tehát négy-öt terminus, a korrelativitás, a komplementaritás, az inverzitás és konverzitás, illetve a reciprocitás, amelyek közül nekem is választanom kell majd. Mivel az empirikus társadalomtudományokban a korreláció fogalmának a korrelativitás itteni tartalmához képest nagyon más jelentése van, ezért a korrelativitás helyett inkább a komplementaritás fogom használni, illetve ahol szükség lesz a nagyobb pontosságra, ott a konverzitás és inverzitás terminusait is alkalmazni fogom. És mindezek mellett – ha történeti vagy olykor stilisztikai okokból szükségesnek látszik – megengedhetőnek tartom a korrelativitás terminus használatát is. Mivel a reciprocitás fogalma nagyon erősen használt a társadalomtudományokban, nem tartanám szerencsésnek ezt a fogalmat az itt tárgyalt jelentésben is használatba venni.

10.2 Normatív pozíciók, jogosultságok és kötelezettségek

A társadalmi cselekvések indoklásai során rengetegszer hivatkozunk a jogainkra, a jogosultság fogalmára:

> Jogom van a bolt előtti járdán sétálni.
>
> Jogában áll hallgatni.
>
> Jogában áll beszélni.
>
> Jogom van visszakapni a kölcsönadott pénzemet.
>
> Jogom van eladni a saját gitáromat.
>
> Jogom van nem eladni a saját gitáromat.
>
> Jogom van kidobni a nem használt ruháimat.
>
> Jogom van megváltoztatni a saját boltom nyitvatartási idejét.
>
> Nincs joga másnak arra, hogy eladja az én autómat valakinek.
>
> Jogom van szavazni.
>
> Jogom van nem szavazni.
>
> Nincs joga senkinek engem megverni.
>
> Ha megtámadnak, akkor jogom van megvédni magam.

A fenti példák azt mutatják, hogy egyrészt a jogaink irányulhatnak tevőleges cselekvésre (csinálásra), illetve nem-cselekvésre (tartózkodásra), másrészt a jogaink vonatkozhatnak a saját, illetve mások cselekvésére, harmadrészt a jogaink tetszőleges cselekvésre irányulhatnak, de olykor "csak" azt szabályozzák, hogy a jogi-normatív viszonyainkat megváltoztathatjuk-e vagy sem.

A fenti példák azt is mutatják egyben, hogy a jogosultságainkhoz valóban hozzárendelhetők mások kötelezettségei. Amikor például a kölcsönadott pénzt várjuk vissza, akkor mindig valaki más kötelezettségére gondolunk (akinek vissza kell adnia a pénzt), vagy ha nekem jogom van erőszakmentes életre, akkor azt is jelenti egyben, hogy mindenki másnak tilos megvernie engem (kötelező tartózkodnia az erőszak alkalmazásától).

A példák rávilágítanak arra, milyen sokféle jelentést sűrítünk a jog-jogosultság fogalmába. Itt az idő, hogy megnézzük, hogyan lehet szétszálazni ezt az összefonódó sokféleséget.

10.2.1 Hohfeld elsőrendű jogviszonyfogalmai

Wesley Newcomb Hohfeld volt az, aki miután észlelte, hogy a jog(osultság) és kötelesség fogalmát sokszorosan túlhasználják a jogászok, megpróbálta elválasztani a különböző jelentésrétegeket egymástól. Hohfeld két tanulmánya a jog és kötelesség fogalmainak többértelműségéről csak a halála után, 1913-ban és 1917-ben jelent meg.[1]

A jogosultságok és kötelezettségek közti korrelatív kapcsolat tézise már Hohfeld előtt is nyilvánvaló volt a jogtudományon belül [394]. Hohfeld azzal hozott újdonságot, hogy rámutatott, a jogosultság, illetve kötelezettség fogalmaiban meghúzódó jelentésrétegek hogyan különíthetőek el egymástól.

[1] Hohfeld tanulmánya magyarul is olvasható [360], a Hohfeld-recepcióból kiemelném Arthur Corbin [173], [174], Albert Kocourek [404], Thomas D. Perry [568], Frederic B. Fitch [259], Walter J. Kamba [390] cikkeit, illetve Matthew H. Kramer, N.E. Simmonds és Hillel Steiner tanulmánykötetét [420].

Hohfeld jogfogalmainak struktúrája

Hohfeld szerint négy-négy jogosultság-, illetve kötelezettségtípust definiálhatunk, és ezek mindegyikét kétféle módon is kapcsolatba hozhatjuk egymással. Minden jogosultságnak van egy ellentéte, illetve van egy korrelatív párja a kötelezettségek között (és természetesen mindez igaz fordítva is, a kötelezettségekhez is rendelhető jogosultság). Hohfeld a következő kategóriákat vezette be:

right	privilege	power	immunity
duty	no-right	liability	disability

Fontos megjegyezni, hogy a Hohfeld-recepció az idők során pár ponton változtatott az eredeti terminusokon. Hohfeld követői az eredeti 'right' terminus helyett a 'claim' (illetve 'claim-right'), míg a 'privilege' helyett a 'liberty' és a 'freedom' kategóriáit is használták, illetve volt szerző, aki az 'immunity' helyett az 'exemption'-t használta, más a 'no-right' terminust javasolta lecserélni az 'exposure'-re [392], [393].[2]

A hohfeldi kategóriák magyar fordítására több esetben olyan javaslat született, amelyek a laikusok számára ugyan nehezebben befogadhatóak, viszont sokkal egyértelműbbnek tűnnek. A következő magyar terminusokat vezették be a Hohfeld tanulmányának fordításában [360].

jog	szabadság	felhatalmazottság	mentesség
kötelezettség	joghiány	beavatkozásnak kitettség	beavatkozásképtelenség

A Hohfeld után megjelenő "versengő terminusok" közül a 'claim' fordítására az 'igény' vagy 'igényjog', míg a 'liberty/freedom' terminuspárosra a 'szabadság' kifejezése került használatba. A továbbiakban igyekszem követni a hohfeldi terminológia magyar fordítását, de egy ponton tudatosan eltérek tőle, néhány más esetben pedig – stilisztikai okok miatt – olykor alkalmazok majd szinonímákat. A hohfeldi 'right' terminus magyar fordítása ('jog') helyett inkább a 'igényjog' terminust veszem igénybe, hogy ezzel is elkerüljem a 'jog' terminusának sokrétűségéből fakadó veszélyeket. Sokkal pontosabb lenne a 'joghiány' helyett az 'igényhiány' terminusát "igénybe venni", ezért alkalmazom ezt a kifejezést is, de – mivel nem akarok túlzottan eltérni a magyar fordítástól – olykor használom az eredeti magyar fordítást is. Ha nem lesz zavaró, akkor – stilisztikai okokból – olykor-olykor lecserélem a 'szabadság'-ot a 'privilégium', a 'mentesség'-et az 'immunitás' terminusára (és csak zárójelben jegyzem meg, hogy a 'mentesség' helyett a 'kivételezettség' terminusát is nyugodtan használhatnánk).

Ugyan nem merült fel sehol fordítási javaslatként, mégis érdemesnek tartom megjegyezni, hogy a 'power' kategóriáját lehetne 'beavatkozás-képesség'-nek, az 'immunity' kategóriáját pedig 'beavatkozásmentességnek' is fordítani. Ha így tennénk, akkor egyértelműbben látszódna, hogy a második négyes hohfeldi fogalmi csoport a beavatkozással (vagy mondhatnánk azt is, hogy a rendelkezéssel) kapcsolatos.

Sajnálatos módon a hohfeldi elmélet formalizálási kísérletei során a terminológiai zavart inkább növelték, mint csökkentették azzal, hogy a kötelezettségfogalmakra is jogtípusként hivatkoztak, vagyis négy-négy jog-, illetve kötelezettségfogalom helyett nyolc

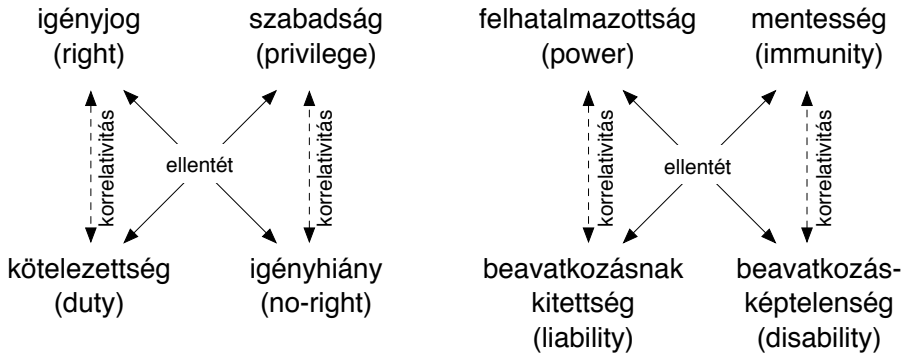
[2] További terminusjavaslatokat mutat be Walter J. Kamba cikke [390].

jogtípus fogalomról beszéltek. Igaz, az alkalmazott formális nyelvi szinten ezt a zavart fel lehetett oldani, de a természetes nyelvi szinten ez a megoldás nehezebbé tette és teszi az értelmezést, befogadást.

Mielőtt sorba venném, hogyan értelmezi Hohfeld a jogosultság- és kötelezettségfogalmait, be kell mutatnom azt, hogy milyen relációkkal hogyan köti össze Hohfeld a kategóriáit. Hohfeld szerint korrelatív viszonyban vannak egymással a táblázatba foglalt elemek közül az egymás alatt levő kategóriák (azaz az igényjog és kötelezettség, a privilégium és igényhiány, a felhatalmazottság és beavatkozásnak kitettség, valamint az immunitás és beavatkozásképtelenség), és ellentét van a "keresztkapcsolatokban" (azaz az igényjog és igényhiány, a privilégium és kötelezettség, a felhatalmazottság és beavatkozásképtelenség, illetve az immunitás és beavatkozásnak kitettség között).

legal relation	jogi viszony	korrelatív párja	ellentéte
right	igényjog	kötelezettség	igényhiány
duty	kötelezettség	igényjog	szabadság
privilege	szabadság	igényhiány	kötelezettség
no-right	igényhiány	szabadság	igényjog
power	felhatalmazottság	beavatkozásnak kitettség	beavatkozásképtelenség
liability	beavatkozásnak kitettség	felhatalmazottság	mentesség
immunity	mentesség	beavatkozásképtelenség	beavatkozásnak kitettség
disability	beavatkozásképtelenség	mentesség	felhatalmazottság

A táblázatos megjelenítés mellett érdemes grafikus reprezentációt is adni a hohfeldi kategóriáknak és a köztük levő viszonyoknak, mert így jobban lehet érzékeltetni a fogalmi viszonyrendszer belső szerkezetét.



Az ábra nemcsak azt mutatja, hogy mely fogalmak között van ellentétes, illetve korrelatív viszony, de az is "leolvasható" róla, hogy a fogalmak két csoportra oszthatók. Az egyik csoport az igényjog, a másik felhatalmazottság köré szerveződik, és mindkettő csoportnak hasonló belső mikrostruktúrája van. A későbbiekben majd kiderül, hogy milyen kapcsolat van a két csoport, illetve a két csoport elemei között.

Az igényjog köré szerveződő fogalmak értelmezése

A hohfeldi jogfogalmak struktúráját már bemutattam, de se nem definiáltam, se nem adtam példákat az egyes fogalmakra, pedig az értelmezésükhöz, modellbe építésükhöz

mindkettőre szükség lenne. Első körben bemutatom Hohfeld saját meghatározásait, miközben igyekszem tőle vagy a korai Hohfeld-recepcióból származó példákat hozni rájuk.

Az első fogalmi csoport középpontjában a jogigény kategóriája áll, aminek komplementere a kötelezettség és ellentéte a joghiány fogalma. Hohfeld példája így szól:

> „. . . ha *X*-nek joga van ahhoz, hogy *Y* az ő földjén kívül tartózkodjék, ezzel korrelatív (és ennek megfelel) *Y* azon kötelessége, hogy maradjon távol attól a helytől." [360, 73]

A fenti idézetben Hohfeld két jogi terminust nevez meg (az igényjogot és a kötelezettséget), de ezzel igazából nem két "különálló" jogviszonyt állít szembe egymással. Az igény és a kötelezettség voltaképpen ugyanannak a jogviszonynak két "nézetét" jelenti.[3] Ezért mondja rájuk Hohfeld, hogy korrelatív kapcsolatban vannak egymással. A két jogviszonyt így bonthatjuk fel összetevőkre:

jogosult	jogviszony	cselekvő	cselekvés	
Ego	igénye	Alter	nem	átjárni Ego telkén
Alter	kötezettsége	Alter	nem	átjárni Ego telkén

Az összetevőkre bontásból látszik, hogy mindkét fél (jogbirtokos) jogviszonya ugyanarra a cselekvésre (és ugyanarra a cselekvőre) vonatkozik. Ez fontos, később látni fogjuk, hogy nem minden párosítás esetében van ez így.

Hohfeld következő jogfogalma a szabadság, ami a kötelezettséggel van ellentétben és a joghiánnyal korrelatív kapcsolatban. Hohfeld újra az ingatlanos példát hozza erre.

> „ . . . ha *X*-nek joga vagy igénye van arra, hogy *Y* – valaki más – maradjon távol földjétől, akkor neki magának megvan az a szabadsága, hogy belépjen e területre; másként fogalmazva, *X*-t nem terheli a távolmaradás kötelezettsége. A belépés szabadsága a kívülmaradás kötelezettségének ellentéte." [360, 73]

Szabadság és kötelezettség tehát ellentétes fogalmak, de pontosan látni kell, mit is ért Hohfeld az 'ellentétes' fogalom jelentése alatt:

> „. . . e ponton nem árt némi óvatosság; hiszen amikor azt mondjuk: egy adott szabadság egy kötelezettség puszta negációja, akkor ezt természetesen mindig úgy értjük, hogy a kötelezettség tartalma az, ami pontosan ellentétes a kérdéses szabadságéval." [360, 74]

A telektulajdonosnak szabadsága van saját telkén átjárnia, másoknak kötelezettségük van tartózkodni az átjárástól. Utóbbi úgy is értelmezhető, hogy mások kötelesek nem átjárni a tulajdonos telkén. Ha Egora és Alterre vonatkoztatjuk mindezt, akkor a következő formában adhatjuk meg a két jogviszony összetevőit:
Látható, hogy az ellentétes jogviszonyok más cselekvőre és más cselekvésre irányulnak. Fontos kérdés, hogy mi a kapcsolat ugyanazon személy, ugyanazon jogbirtokos (itt az Ego) igényjoga és szabadsága között. Joel Feinberg ezt írja:

[3] Talán hasznos lehet itt egy másik példát hozni szemléltetésképpen. Amikor azt mondjuk, hogy *X* szülője *Y*-nak, akkor mondhatjuk azt is, hogy *Y* gyereke *X*-nek. Bármelyiket is mondjuk, ugyanazt a tartalmat fejezzük ki velük. A 'szülője' és a 'gyereke' reláció korrelatív párjai egymásnak (vagy másként: egymás inverzei). Ugyanez igaz az igényjog-kötelezettség párosra is.

jogosult	jogviszony	cselekvő	cselekvés	
Ego	szabadsága	Ego		átjárni Ego telkén
Alter	kötelezettsége	Alter	nem	átjárni Ego telkén

> „Amit a jog [az igényjog] hozzátesz a szabadsághoz, az másoknak az a kötelessége, hogy tartózkodjanak a beavatkozástól.” [246, 88]

Feinberg arra utal, hogy Hohfeld két kategóriáját szigorúan el kell választani egymástól, a szabadság nem feltételezi – sem logikailag, sem gyakorlatilag – a jogigény létezését. A szabadság önmagában még nem teremti meg, nem hozza magával azt az igényjogot, hogy az Alter nem avatkozik be, nem zavarja meg Egot abban a cselekvésben, amire Egonak szabadsága van. Hohfeld példája arra, hogy a szabadság nem keletkeztet korrelatív kötelezettséget a másik fél be nem avatkozásra, így szól:

> „*A, B, C* és *D* mint salátatulajdonos esetleg azt mondja *X*-nek: ‘Edd meg a salátát, ha tudod; engedélyezzük ezt neked, de nem értünk egyet azzal, hogy ne zavarjunk meg ebben.’ Ebben az esetben a szabadság létezik, s amennyiben *X*-nek sikerül megennie a salátát, akkor nem sérti meg egyik fél jogát sem. Világos azonban az is, hogy ha *A*-nak sikerülne olyan gyorsan elkapnia a tálat, hogy *X* egy falatot sem tudna abból enni, úgy *Y*-nek semmilyen joga sem szenvedne csorbát.” [360, 76]

A példa szemléletesen rávilágít, hogy mi a különbség a kétfajta jogviszony, az igény és a szabadság között. De vehetek egy másik példát a hétköznapok világából. Ha egy vonatra helyjegyet veszek, akkor igényjogom keletkezik arra, hogy leüljek a székemre. Ha valaki ül azon a széken, amire a helyjegyem szól, akkor igényelhetem tőle, hogy adja át a helyét, és ezzel “párhuzamosan” neki kötelessége lesz felállni. Ha viszont nem helyjegyes vonaton utazom, akkor szabadságom van a még szabad székek bármelyikére leülnöm, de nem igényelhetem senkitől sem, hogy adja át nekem azt a széket, amin ő ül. Viszont ha én leülök egy szabad székre, utána nem leszek köteles átadni az a helyt valaki másnak, ezt senki sem igényelheti tőlem. Amíg viszont vannak szabad székek, addig bárki elfoglalhatja azok bármelyikét, és megakadályozhat engem abban, hogy a kiszemelt szabad helyre leülhessek, ha ő előbb teszi ezt meg.

Nem könnyű észrevenni a szabadság “függetlenségét” az igényjogtól, már ha valóban magában áll. A tényleges gyakorlatban a jogviszonyok az esetek nagy részében összetettek és a szabadság és igényjog együttjár. Így van ez például a tulajdonjog esetében. De ez a gyakorlati együttjárás semmiképpen sem jelent logikailag szükségszerű összetartozást köztük. Ha viszont a szabadság hol magában, hol az igényjoggal együtt fordul elő a különböző gyakorlati jogfogalmak között, akkor ezt a kétféle helyzetet, pontosabban e kétféle helyzetre jellemző szabadságfogalmat érdemes lehet jelzős szerkezetbe téve elkülöníteni egymástól. Erre tett javaslatot Jeremy Bentham, aki a **nyers szabadság** (**naked liberty**), illetve a **szerzett szabadság** (**vested liberty**) fogalompárt javasolta alkalmazni erre a kétféle helyzetre (Bentham nézeteit ismerteti: [449, 17-19]). Amikor a szabadság csak magában áll, akkor nyers, védtelen ahhoz képest, amikor az igényjoggal kiegészülve védve van a mások általi beavatkozástól. A nyers szabadságban megtehetünk mindent tetszésünk szerint, de mindenki másnak szabadsága van megakadályozni minket saját szabadságunk (akaratunk) érvényesítésében. A védett szabadságban a beavatkozás tiltva van a többiek

részére.[4]

Nem ejtettem még szót az utolsó hohfeldi kategóriáról, a joghiányról. Ez a fogalom az igényjog ellentéte és a szabadság korrelatív párja. A korrelativitás azt jelenti itt, hogy az egyik fél (Ego) szabadsága megtenni valamit megegyezik azzal, hogy a másik félnek (Alternek) joghiánya van (nincs igénye) arra, hogy az egyik (Ego) ne tegye meg azt:

> „ ... *X* azon jogának korrelatív párja, hogy *Y* ne lépjen a földjére, azonos *Y* azon kötelezettségével, hogy ne tegye azt; *X* azon szabadságának korrelatív párja azonban, hogy saját földjére lépjen, azonos *Y* azon 'jogának hiányával', hogy igényelhesse *X* távolmaradását." [360, 74]

Az ingatlanos példa jogviszonyait kifejező mondatok összetevőkre bontásával a következő táblázatot kaphatjuk:

jogosult	jogviszony	cselekvő	cselekvés	
Ego	szabadsága	Ego		átjárni Ego telkén
Alter	igényhiánya	Ego	nem	átjárni Ego telkén

Hohfeld ezt is korrelatív kapcsolatnak nevezi, pedig ez a viszony különbözik az igényjog és kötelezettség közti korrelativitástól. Ebben az esetben ugyanis két jogviszony nem ugyanarra a cselekvésre irányul, hanem egyszer a pozitív, másszor a negatív cselekvésre (vagyis a tevésre, illetve a tartózkodásra). A különbséget felismerve Stig Kanger, aki később formalizálta Hohfeld elméletét, két fogalmat is bevezetett a korrelativitás e két típusára, és hol az inverz, hol a konverz reláció fogalmát alkalmazta (↬395.o.)

Az igényhiány az igényjog ellentéte. Hohfeld nem mond semmit erről a kapcsolatról, pedig ez eltér a szabadság és kötelezettség közt megállapított ellentét relációtól. Amíg utóbbira Hohfeld azt mondta, hogy „a kötelezettség tartalma az, ami pontosan ellentétes a kérdéses szabadságával", addig az egymással ellentétes joghiány és igényjog két jogviszonya ugyanarra a tartalomra vonatkozik [449, 27]. Az ellentéte fogalmat inkább az utóbbi értelmezésben szokás használni, de akárhogy is akarjuk ezt a fogalmat értelmezni, arra mindenképpen ügyelni kell, hogy a hohfeldi elmélet kétféle 'ellentéte' relációját megkülönböztessük egymástól.

jogosult	jogviszony	cselekvő	cselekvés	
Ego	igénye	Alter	nem	átjárni Ego telkén
Alter	igényhiánya	Ego	nem	átjárni Ego telkén

A hohfeldi jogfogalmak első csoportja az igényjog köré szervezhető, és a fent bemutatott kapcsolatok alakítják ki a köztük levő struktúrát. Összegzésképpen egy áttekintő táblázatban – az eddigi ingatlanos példákat felhasználva – egyesítem a négy fogalom összetevőkre bontását, hogy jobban látni lehessen a fogalmak egymáshoz való viszonyát.
A táblázat első kettő, illetve második kettő sorában egymás korrelatív párjai szerepelnek, míg az első és negyedik, illetve a második és harmadik sorban egymás ellentétei.

Bonyolódik azonban a kép, ha figyelembe vesszük azt is, hogy – a fenti példában – Egonak nemcsak az lehet szabadságában, hogy átjárjon saját telkén, de szabadsága lehet

[4] A formális elmélet bemutatása után kitérek majd H.L.A. Hart ide kapcsolható **védőövezet** (**protective perimeter**) fogalmára is (↬**??**.o.).

jogosult	jogviszony	cselekvő	cselekvés	
Ego	igénye	Alter	nem	átjárni Ego telkén
Alter	kötelezettsége	Alter	nem	átjárni Ego telkén
Ego	szabadsága	Ego		átjárni Ego telkén
Alter	igényhiánya	Ego	nem	átjárni Ego telkén

tartózkodni is ettől. Ennek pedig az a korrelatív párja, hogy Alternek igényhiánya van arra, hogy Ego átjárjon a telkén. Ezt a következő módon írhatjuk fel:

jogosult	jogviszony	cselekvő	cselekvés	
Ego	szabadsága	Ego	nem	átjárni Ego telkén
Alter	igényhiánya	Ego		átjárni Ego telkén

Ha csak az Ego telkére irányuló jogviszonyok eddig példáit nézzük, akkor is kiviláglik, hogy Hohfeld jogviszonyfogalmai nem elég pontosak. Nem mindegy adott eetben ugyanis, hogy valamely jogviszony relátumaként (paramétereként) milyen minőségű cselekvést írunk be. Hohfeld maga is figyelmeztet a szabadság és kötelezettség ellentéte kapcsán (↬383.o.), hogy ott egymás ellentéteit, a pozitív cselekvést (tevést), illetve a negatív cselekvést (tartózkodást) kell beleérteni a két ellentétes jogviszonyba. De Hohfeld nem megy túl ezen a figyelmeztetésen, nem igazán reflektál a problémára. A továbblépéshez viszont tisztáznom kell, hogy hogyan lehet pontosabban kezelnünk a hohfeldi sémán belül a cselekvés két típusát, ezért ezen a ponton egy rövid kitérőt kell tennem.

Az igényjog köré csoportosítható jogviszonyok felbontása

A hohfeldi jogviszonyfogalmak pontosítása, egyértelmű formalizálása érdekében első lépésként érdemes grafikai eszközökkel reprezentálni a kategóriák közti viszonyokat. Egy hohfeldi jogviszonyt négy változóval írhatunk le. Mindig meg kell adnunk a jog típusát, a jog birtokosát, a jog tartalmát adó cselekvés típusát és e cselekvés ágensét. Ezt az alábi formulában összegezhetjük:

(i) $-\,\textsc{right}(Agent, \pm\textsc{do}(Agent, T(\pm p, \pm p)))$

Ezt – az egyszerűbb kezelésmód miatt, hangsúlyozottan csak ideiglenes jelleggel – írhatjuk a következő rövidebb alakban is. Azt viszont kétféleképpen is tehetjük. Először kivehetjük a formulából a teljes állapotváltozást ($T(\pm p, \pm p)$), hogy csak azt hangsúlyozzuk, a cselekvésnek van egy Ágense.

(ii) $-\,\textsc{right}(Agent, \pm\textsc{do}(Agent))$

Később bemutatok olyan elméleteket, illetve azokban alkalmazott formalizmusokat, amelyekben fontos lesz valahogy jelezni azt, hogy mi a cselekvés révén "megvalósuló" állapotváltozás során előállt végállapot (ezt a későbbi elméletekhez való igazodás kedvéért az eddig alkalmazott p állapotváltozó helyett F-fel jelölöm).

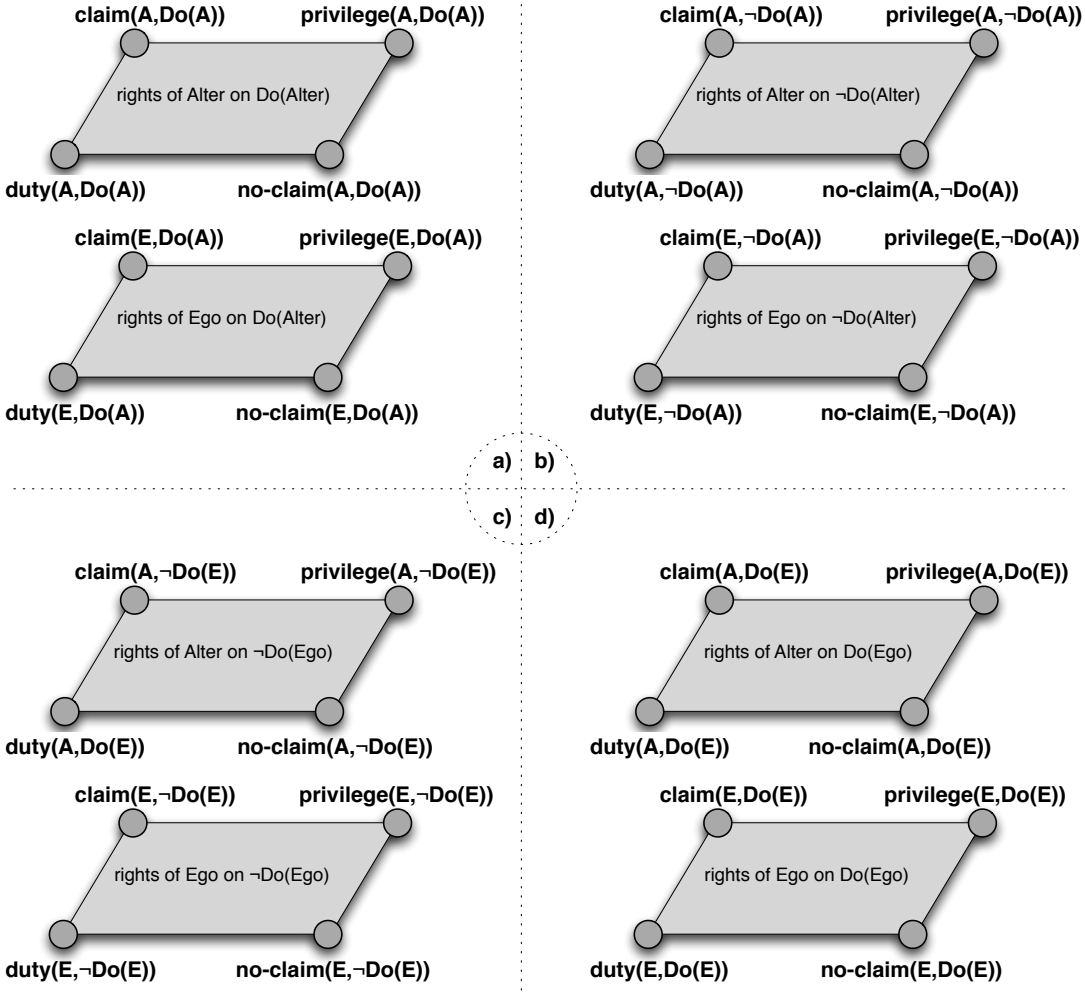
(iii) $-\,\textsc{right}(Agent, \pm\textsc{do}(Agent, \pm F))$

Ez kevesebb információt képes kifejezni, mint az *(i)* formula, de többet, mint a második *(ii)* képlet. A formula a következő értékeket veheti fel az egyes változókban:

A négy változó közül a Right-nak négy, a többinek két-két értéke lehet, így a formulának összesen 32 különböző kimenete lehet. Egy adott jogbirtokos adott ágens adott cse-

a jog (RIGHT) típusa:	CLAIM | LIBERTY | DUTY | NO-CLAIM
a jog birtokosa (*Agent*):	*Ego (E)* | *Alter (A)*
a jog tartalmát adó cselekvés (±DO):	DO | ¬DO
a cselekvés ágense (*Agent*):	*Ego (E)* | *Alter (A)*

lekvésére vonatkozó jogviszonyait egyetlen "síkban" összefogva, az alábbi nyolc jogviszonycsoportot kapjuk:

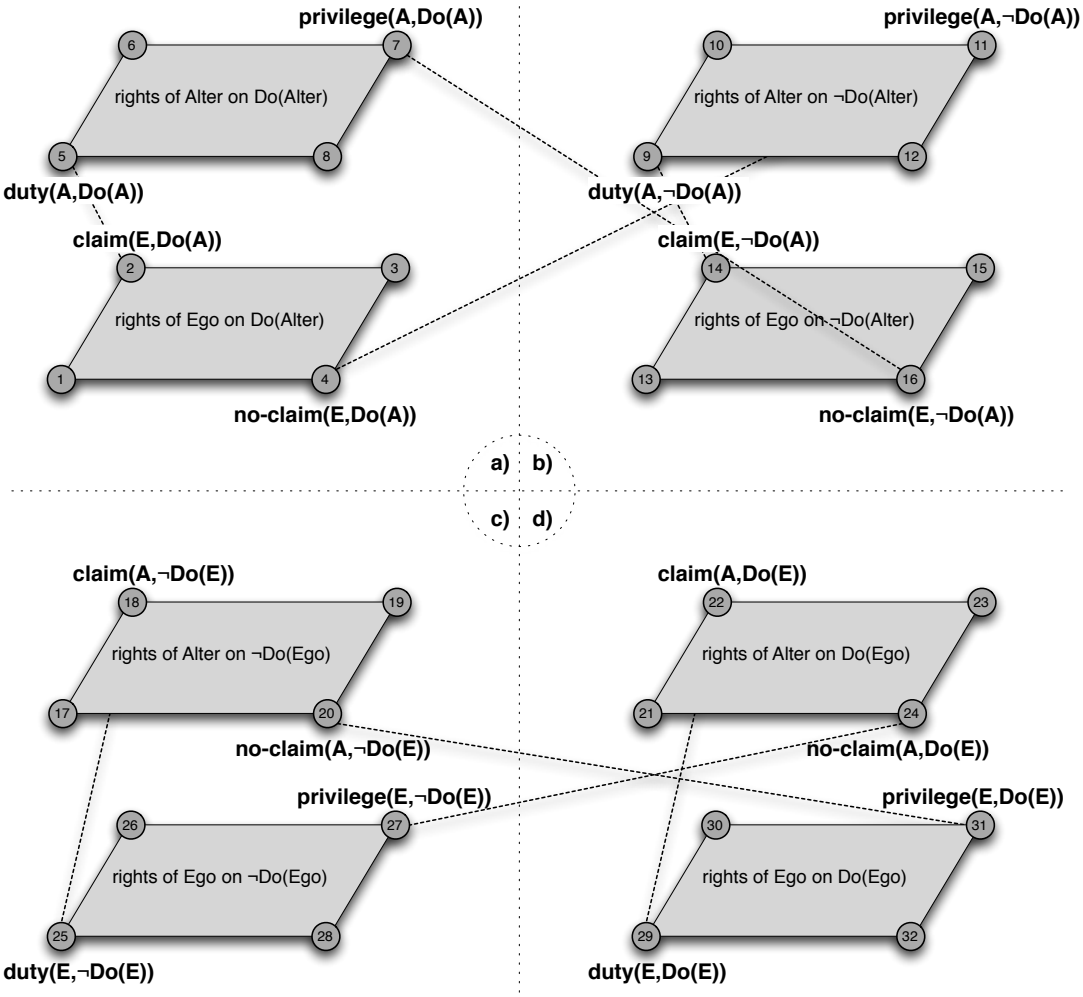


A nyolc síkot négy-négy térnegyedbe rendezhetjük azon az alapon, hogy milyen cselekvésről van szó (vagyis kinek a milyen cselekvéséről beszélünk). Az a) jelű térnegyedben Alter tevésével, a b) jelűben Alter tartózkodásával, a c) jelűben Ego tevésével, a d) jelűben Ego tartózkodásával kapcsolatos jogviszonyok láthatóak úgy, hogy egy-egy negyedbeli síkon Ego, illetve Alter jogai vannak "összekötve".

Észre kell vennünk azonban, hogy nem minden – logikailag lehetséges – jogviszony értelmes. Az igény (jogigény) például csak akkor értelmezhető, ha a jogbirtokos és a cselekvő különbözik. Nincs értelme saját cselekvésemre irányuló igényről beszélni. Ezen a ponton ugyanis feltételezhetjük, hogy egy ágens saját cselekvésére irányuló akarata

egyértelmű, ha valaki akar valamit, akkor azt végre is hajtja.[5] Ugyancsak (és ugyanazért) értelmetlen az igényhiány fogalma is azonos ágensek esetében. Tartalmatlannak tarthatjuk a szabadság fogalmát is – eltérő ágensekre vonatkozóan. A szabadság csak saját cselekvésemet illetően értelmes, mással cselekvésére vonatkozóan nekem nincs, nem lehet szabadságom. Ugyanez igaz a kötelezettségre is: csak magamat illetően lehetnek kötelezettségeim, másokra nézve csak a kötelezettségre vonatkozó igényem képzelhető el.

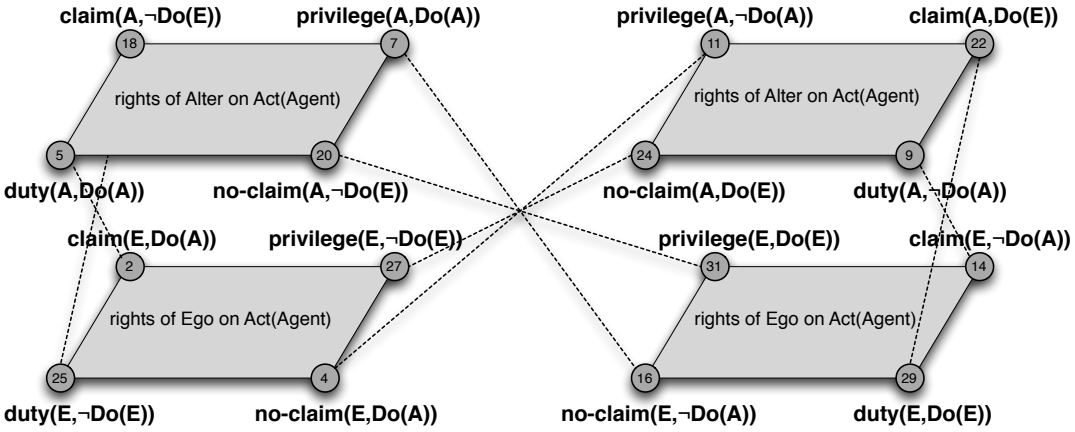
Ha mindezek alapján a "nem értelmes" fogalmakat töröljük a "palettáról", akkor már csak 16 jogviszonyhoz jutunk. Ha ezek között feltüntetem a hohfeldi korrelativitási kapcsolatokat, akkor a következő ábrát kapjuk:



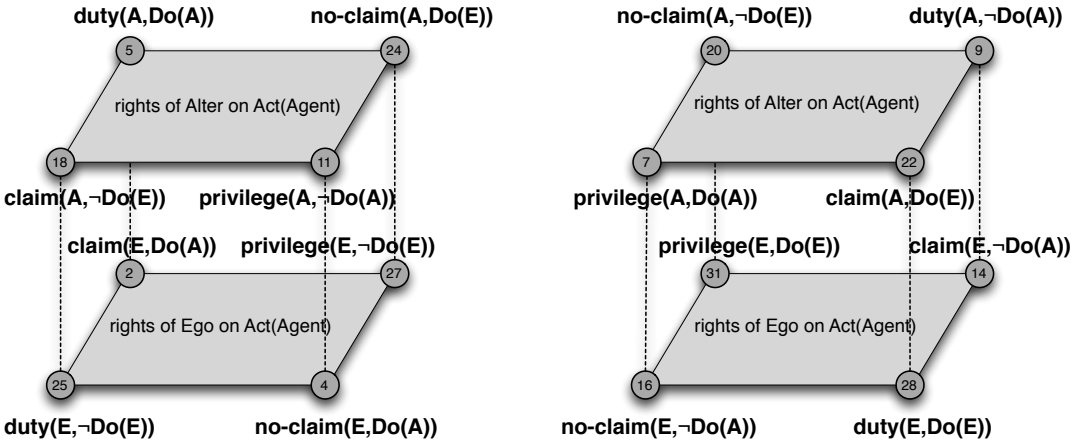
Az ábrán már nem láthatóak az "értelmetlen" jogviszonyok, csak azok szerepelnek, amelyek a hohfeldi értelmezésekben felbukkantak, felbukkanhattak. A korrelativitást jelző vonalak "átkötései" más negyedekbe szépen mutatják azt, hogy a jogfogalmak közti belső összefüggések másfajta cselekvéseket kapcsolnak össze. Az ábra tehát azt is szemlélteti egyben, hogy a fogalmak pontos értelmezéséhez mennyire fontos figyelembe venni azt, hogy milyen cselekvésre vonatkoznak a jogviszonyok. Hohfeld sajnos nem igazán reflektált erre a kérdésre. A jogviszonyok elméletileg teljesnek mondható ábráját tovább

[5] Az akaratgyengeség problémáját én ebben a kontextusban irrelevánsnak gondolom.

egyszerűsíthetjük úgy, hogy két-két negyedet "összecsúsztatunk. Ha az egymás alatt levő negyedekkel (tehát az a) és c), illetve a b) és d) jelűekkel) tesszük meg ezt, akkor olyan síkokhoz jutunk, amelyek ugyanannak a jogbirtokosnak a jogviszonyait mutatják. A két alsó síkon Ego, a két felső síkon Alter jogviszonyai láthatók. A jobb áttekinthetőség kedvéért a jobb oldali két síkot tükröztem a függőleges tengely mentén, mert így jobban láthatókká váltak a korrelativitási kapcsolatok:



Egy utolsó átrendezést végrehajtva a Hohfeldi sémához nagyon hasonló szerkezethez juthatunk. Csak annyit kell tennünk, hogy a két felső síkot a vízszintes tengelyük körül elforgatjuk 180 fokkal (vagyis a síkok felső és alsó csúcspontjait megcseréljük egymással), majd ezután a két felső sík privilege és no-claim jogviszonyait kicseréljük egymással, végül a második kocka csúcspontjait a lapra merőleges, függőleges síkra vetítve tükrözzük. Ekkor ugyanúgy az alsó két síkon helyezkednek az Ego, a felső két síkon az Alter jogviszonyai, viszont a korrelatív kapcsolatok egymás fölé kerülnek, és az egész struktúra két részre "esik szét". A két kocka elemei csak abban térnek el egymástól, hogy mindig az ellentétes cselekvésre vonatkozó jogviszonyt írják le.



Mindkét ágens jogviszonyai között (tehát a felső és alsó két síkon) találhatunk két-két négyes csoportot. Mindkét ágens esetében négy-négy jogfogalom összetartozik azon az alapon, hogy a jogviszony az ágens saját vagy a partnere cselekvésére irányul-e (mindig a DUTY és PRIVILEGE, valamint a CLAIM és NO-CLAIM fogalmak kapcsolódnak így össze).

Ezeket a fogalmi négyeseket lehet majd deontikus logikai formulák segítségével megragadni. Az ábrából az is kiderül, hogy Hohfeld négy-négy jogosultság, illetve kötelezettség fogalma kevés, valójában kétszer annyira van szükség.

Az utolsó ábra "eredményeit" táblázatos formában, szövegesen is reprezentálhatom. A pozitív és negatív cselekvést tevésnek és tartózkodásnak írva a már megszokott táblázatos formában az alábbi módon írhatjuk fel a nyolcféle jogviszonyt:

jogosult	jogviszony	ágens	cselekvés	jogosult	jogviszony	ágens	cselekvés
Ego	igénye	Alter	tartózkodni	Alter	kötelezettsége	Alter	tartózkodni
Ego	szabadsága	Ego	megtenni	Alter	igényhiánya	Ego	tartózkodni
Ego	igénye	Alter	megtenni	Alter	kötelezettsége	Alter	megtenni
Ego	szabadsága	Ego	tartózkodni	Alter	igényhiánya	Ego	megtenni
Ego	kötelezettsége	Ego	tartózkodni	Alter	igénye	Ego	tartózkodni
Ego	igényhiánya	Alter	megtenni	Alter	szabadsága	Alter	tartózkodni
Ego	kötelezettsége	Ego	megtenni	Alter	igénye	Ego	megtenni
Ego	igényhiánya	Alter	tartózkodni	Alter	szabadsága	Alter	megtenni

A táblázatban az elemi jogviszonyok mellett mindig feltüntettem a korrelatív párjaikat is. A táblázat baloldali részében van Ego, a jobboldaliban Alter nyolc-nyolc jogviszonya. Ezeket a 10.2.1 ábrán a két kocka két-két alsó, illetve felső síkján lehet megtalálni.

✠ ✠ ✠

Miután kiderült, hogy meg kell duplázni Hohfeld jogviszonyfogalmait, és nyolc alapfogalomról (illetve ezeknek nyolc korrelatív párjáról) kell beszélnünk, keresnünk kell olyan a terminusokat a jogi diskurzuson belül, amelyeket hozzárendelhetünk a kibővült fogalmi tér új elemeihez. A jogviszonyra vonatkozó tipizálási javaslatok közül a két legismertebb az **aktív jog** (ACTIVE RIGHT) és **passzív jog** (PASSIVE RIGHT), valamint a **pozitív jog** (POSITIVE RIGHT) és **negatív jog** (NEGATIVE RIGHT) megkülönböztetésén alapul.

Az aktív jog, passzív jog kettősségét David Lyons mutatja be a 'The Correlativity of Rights and Duties' című cikkkében [461]. Lyons szerint az aktív jog az, amikor az ágensnek joga van arra, hogy maga cselekedjen valamit, a passzív jogról pedig akkor beszélhetünk, ha az ágensnek joga (igénye) van arra, hogy egy másik ágens cselekedjen valamit. Hohfeld eddig elemzett jogosultságfogalmai közül az igényjog passzív jognak, a szabadság aktív jognak minősíthető. A passzív jog esetében a jogosultság korrelatív párja mindig valamilyen aktív kötelezettséget jelent, az aktív jog esetében pedig a korrelativitás valamely passzív jogigény hiányáról van szó. A korrelatív kapcsolatban levő jogfogalmak tehát mindig eltérő minőségűek az aktív-passzív dimenzióban.

A jogot szokás felosztani pozitív, illetve negatív jogra is. Innentől kezdve azonban nagyon ügyelni kell arra, hogy milyen jogfogalmakat hogyan használunk. Ezen a ponton ugyanis több forrásból, más megközelítésekből származó fogalmakat kell majd egységes keretben kezelni. A jog fogalmával foglalkozó szakirodalomban az eddig elemzett jogosultságfogalmainkra (tehát mind az igényjogra, mind a szabadságra) vonatkozóan ugyanúgy alkalmazzák a pozitív/negatív jog terminusait. Nagyon vigyázni kell tehát akkor, amikor ezeket egységes keretbe foglalva akarjuk értelmezni.

Jan Narveson volt az egyik olyan szerző, aki javasolta pozitív és negatív jog fogalmainak megkülönböztetését (amit átvett tőle például a Stanford filozófiai lexikon joggal foglalkozó szócikkében a cikk szerzője, Leif Wenar is [812]) is. Narveson a hohfeldi

igényjog kategóriájára gondolt, noha nem reflektált rá. Jól mutatja ezt az a mód, ahogy Narveson a jogot a kötelezettség ellenében határozta meg, de ezt erősíti az a tény is, hogy milyen példákat hoz a **pozitív jog** (POSITIVE CLAIM), illetve a **negatív jog** (NEGATIVE CLAIM) fogalmára [528, 57-58]. A negatív jogon Narveson azt érti, hogy az egyén (a jogbirtokos) "jogosult arra, hogy ne avatkozzanak bele mások a cselekvésébe". Ilyen például a "megtámadás elleni védelemhez", a "méltósághoz", a "jó hírnévhez" vagy a "magánlakás sérthetetlenségéhez" való jog. Ezzel szemben a pozitív jogon azt érthetjük, hogy az egyén (a jogbirtokos) jogosult valamire, valamilyen jószágra, szolgáltatásra. Példa rá a "szociális segélyre" való jog.

A definíciókból, a példákból egyértelműen látszik, hogy itt a hohfeldi igényjog fogalmának alábontásáról van szó. Rögzíteni is érdemes, hogy a passzív jog fogalma alatt – a fenti értelemben – létrehozhatjuk a pozitív és negatív jog fogalmait. Amíg a pozitív passzív jog mások aktív cselekvésére (valaminek a megtevésére) való igényt, addig a negatív passzív jog mások valamitől való tartózkodásra vonatkozó igényt jelenti. Nyilvánvalóan adódik azonban itt a kérdés, vajon a többi jogosultságfogalom esetében alkalmazható-e a pozitív/negatív dimenzió a további tipizálásra. Vannak, akik nem így gondolják. A Stanford online filozófiai lexikon joggal foglalkozó cikkében például ez szerepel:

> „Since both negative and positive rights are passive rights, some rights are neither negative nor positive. Privileges and powers cannot be negative rights; and privileges, powers, and immunities cannot be positive rights. The (privilege-) right to enter a building, and the (power-) right to enter into a binding agreement, are neither negative nor positive." [812].

Nem értek teljesen egyet ezzel a gondolatmenettel. Azon a ponton mindenképpen tévesnek tartom, amikor a pozitív/negatív jog kettősségét a passzív jog fogalma alá rendeli. A pozitív és negatív jog megkülönböztetése ugyanúgy értelmes az aktív jog esetében is. Ezt mutatom meg a továbbiakban.

Adja magát, hogy azonnal hivatkozzak itt a kérdés leghíresebb szerzőjére, Isaiah Berlinre, aki a 'A szabadság két fogalma' című tanulmányában definiálta a pozitív és negatív szabadság fogalmát [70], és ezzel megalapozta, hogy e két fogalom hosszú időre bekerüljön a jogok típusairól szóló diskurzusba. Arra azonban nagyon vigyázni kell Berlin kétféle szabadságfogalma kapcsán, hogy ne tévesszen meg minket a terminus szintű azonosság. Berlin ugyanis – részben vagy teljesen – más értelemben használja a fogalmait.

Talán úgy tudom legpontosabban megmutatni, hogy hol térnek el a hohfeldi és berlini értelmezések, logikák, ha előbb megvizsgálom Jan Narveson szabadságfelfogását. Narveson is kétféle szabadságfogalmat definiál, hasonló szellemben, mint ahogy azt Berlin is teszi, csak még általánosabb formában [528, 22-23]. **Pozitív szabadság** (POSITIVE FREEDOM) fogalma alatt Narveson azt az állapotot érti, amikor minden szükséges feltétel adott ahhoz, hogy valamely cselekvés végrehajtható legyen. Ezek a feltételek lehetnek külső és belső feltételek egyaránt. Utóbbira a foci szabályainak ismeretét, megfelelő testi kondíciót, előbbire a szükséges számú társ jelenlétét, focilabda, szabad focipálya meglétét hozza fel Narveson példaként. Ezzel szemben a **negatív szabadság** (NEGATIVE FREEDOM) fogalmát Narveson olyan állapotként definiálja, amelyből hiányzik minden olyan mozzanat, amely megakadályozná valamely cselekvés végrehajtását. Példa lehet erre egy

másik, ugyanazon a pályán focizni vágyó társaság hiánya. Narveson tehát a kétféle szabadságfogalom értelmezésekor bármilyen feltételt "megengedhetőnek", figyelembe vehetőnek tart.

Isaiah Berlin 'A szabadság két fogalma' című tanulmányában definiálta a pozitív és negatív szabadság fogalmát [70]. Berlin a negatív szabadság fogalmával kezdi az elemzését, és abban az értelemben Narveson definíciójához hasonlítható Berlin meghatározása is, hogy a cselekvést akadályozó feltételek hiányaként értelmezi a negatív szabadság fogalmát.

> „Kényszeríteni valakit annyi, mint megfosztani őt a szabadságától." [70, 340]

> „Általában azt mondják, hogy annyira vagyok szabad, amennyire más emberek vagy más emberekből álló csoportok nem avatkoznak be cselekedeteimbe. A politikai szabadság ebben az értelemben egyszerűen az a terület, amelyen belül az ember anélkül cselekedhet, hogy mások megakadályoznák ebben." [70, 342]

> „. . . a szabadság ebben az értelemben a *valamitől* való szabadságot jelenti, vagyis a változó, de mindig felismerhető határok áthágásának hiányát." [70, 352]

Narveson felfogásához képest Berlin szűkebbre fogja a negatív szabadságot biztosító, hiányzó feltételek körét – bár erre a kérdésre nem igazán reflektál tudatosan. Az mindenesetre kiderül a fenti idézetből, hogy Berlin a negatív szabadságot más emberek beavatkozásától való mentességként gondolja el. Ennyiben szűkebben definiálja a fogalmat Narvesonhoz képest, viszont ez a megközelítés gyakorlatilag megegyezik Hohfeld szabadságfelfogásával. Ahol más emberek nem avatkozhatnak be a cselekvésembe, ahol másoknak igényhiánya van velem szemben, ott vagyok szabad – Berlin negatív szabadsággal kapcsolatos értelmezésében.

Mivel sem Berlin, sem Narvenson nem különbözteti meg a pozitív és negatív cselekvéseket egymástól, ezért a negatív szabadság fogalmába sorolt példákban sem figyelnek erre a szempontra. Berlin negatív szabadságra vonatkozó példái között találhatunk a tevésre és a tartózkodásra irányuló utalásokat egyaránt. A narvesoni negatív szabadság értelmezés túl tág, a berlini felfogás már megfelelő terjedelmét tekintve, és jelentését tekintve az ő definíciója megegyezik Hohfeldével, de a két negatív szabadság fogalom egyike sem segít minket abban, hogy meg tudjuk nevezni a hohfeldi szabadság fogalom két altípusát, azt a két esetet, amikor pozitív, illetve negatív cselekvésre irányuló szabadságokról beszélünk.

Más a helyzet Berlin pozitív szabadság fogalma esetében. Itt nemcsak arról van szó, hogy az új kategóriát nem tudjuk elhelyezni a hohfeldi struktúrába, nem találunk neki megfelelő párt a hohfeldi fogalmak közül, hanem arról is, hogy az értelmezéséhez más dimenzióra van szükség a hohfeldi sémához képest. Vegyük a következő néhány rövid passzust:

> „Mit ér a szabadság annak, aki nem tud mit kezdeni vele? Mi a szabadság értéke, ha nincsenek meg a felhasználásához szükséges feltételek?" [70, 347]

> „Egy cél elérésére való képtelenség önmagában nem jelenti a politikai szabadság hiányát. . . . Azt mondják, nagyon meggyőzően, hogy ha egy ember túl szegény ahhoz, hogy elérjen valamit, amit jogilag nem tilos – egy vekni kenyeret, egy világ

> körüli utat, vagy azt, hogy a bírósághoz fordulhasson jogorvoslatért –, éppoly kevéssé szabad, mint hogyha ezen dolgok elérésétől a törvény tiltaná el." [70, 344]

> „A 'szabadság' szó 'pozitív' értelme az egyénnek abból a vágyából fakad, hogy saját magának az ura legyen. Azt szeretném, ha életem és döntéseim rajtam múlnának, nem pedig holmi külső erők határoznának rólam. Saját akarati aktusomnak kívánok eszköze lenni, nem pedig másokénak. Szubjektum, nem pedig objektum akarok lenni, szubjektum, amelyet saját gondolatai, saját tudatos céljai irányítanak, s nem külső okok mozgatnak." [70, 361]

A negatív szabadságra vonatkoztatható "szabadság valamitől" felfogással szemben itt a "szabadság valamire" értelmezés lehet az érvényes Berlin szerint. A pozitív szabadság fogalmát jóval hosszabban elemzi Berlin, mint teszi azt a negatív szabadsággal kapcsolatban. Sok olyan példát hoz (vagy lehetne hozni Berlin értelmezése alapján), amely kivezet a hohfeldi jogviszonyok fogalmi köréből, mert kulturális, gazdasági teendőket "nevez meg", de sok olyan pozitív szabadság körébe tartozó példa kerül elő, amely a hohfeldi séma igényjog kategóriája alá lenne besorolható. A pozitív szabadságot tehát Berlin úgy értelmezi, hogy vagy már ismert kategóriával állítható párhuzamba (az igényjog hohfeldi fogalmával), vagy nem a perszonalizált jogviszonyok fogalmi körén belül ragadható meg. Akármelyik esetről is van szó, egy a lényeg: ez a kategória nem hasznosítható a modellépítés során.

Bár más terminusokat használ, mások a kérdésfeltevései, mégis érdemes felidézni itt, amit Cass R. Sunstein írt a pozitív és negatív jogok különbségeiről, mert segít értelmezni a kétféle berlini szabadságfogalom viszonyát (nyugodtan be lehetne helyettesíteni a 'jog' helyébe a 'szabadság' terminusát, ugyanúgy érvényes maradna Sunstein mondandója a berlini fogalmakra vonatkoztatva.

> A pozitív jogok „. . . nem érvényesíthetők bírósági úton: különleges szerepet töltenek be, és elsősorban arra szolgálnak, hogy az állami tisztségviselők és az állampolgárok összességének általános céljait megfogalmazzák.
> A negatív jogok . . . azonban bírósági védelmet élveznek, és nem szükséges a bírói út igénybevétele előtt semmilyen más állami hatósághoz fordulni. Senki sem függhet az állami hatóságoktól, ha a jogait bíróság előtt kívánja étvényesíteni.
> A gazdasági *szabadságok* megőrzésére tett erőfeszítéseket kísérheti az a tudatos törekvés, hogy csökkentsük a gazdasági *jogok* szerepét." [717, 14]

Az idézetben nem a 'politikai üzenet' a fontos ("nem annyira kívánatos a pozitív jogokat az alkotmányba emelni"), hanem inkább az, ahogy a kétféle fogalmat szembeállítja egymással a szerző. A kétféle jogfogalom inkább a politika vagy a jog intézményrendszeréhez köthető, és e két kötődés eltérő "működési logikát" eredményez.

✠ ✠ ✠

A jogelméleti diskurzus kétféle jogtipológiája, a pozitív-negatív jog, illetve az aktív-passzív jog fogalomkettőse csak kisebb részben, de azért használható a hohfeldi elmélet pontosítására. A passzív-aktív dimenzió a későbbiekben segíthet még minket a hohfeldi jogviszonyok értelmezésében, de a modellépítésnek jelen fázisában nem alkalmazható. A pozitív-negatív jog megkülönböztetése egy az egyben ráhúzható az igényjog két altípusára, a pozitív és negatív cselekvésekre irányuló igényjogok eseteire. Amíg a pozitív

igényjog esetén a partner pozitív cselekvésére (tevésére) irányuló igényről beszélhetünk, addig a negatív igényjog esetén az igény a partner negatív cselekvésére (tartózkodására) vonatkozik. A szabadság esetében is ugyanezt megtehetjük, tehát értelmezhetjük a negatív és pozitív szabadság fogalmát (tevésre és tartózkodásra irányuló szabadságokat értve alattuk), de ez a felfogás teljesen eltér a Berlin-féle értelmezéstől.

Meg kell itt jegyeznem azt is, hogy könnyen zavart okozhat az, hogy mind a pozitív és negatív, mind az aktív és passzív minőségeket vonatkoztatjuk a cselekvésekre, és a gyakorlatban nehéz a pontos különbséget tenni köztük, ami sajnos sok félreértésre, félrehasználatra, olykor túlhasználatra ad alkalmat. Ennek oka egyszerű (és érthető): mindkét dichotómia könnyen "ráhúzható" a tevés és tartózkodás kettősségére, hiszen a tartózkodás a tevéshez képest értelemezhető egyfelől negatívnak a pozitívhoz, másfelől passzívnak az aktívhoz képest. Ezen a pongyolaságon nem tudok változtatni. Csak annyit tehetek a potenciális félreértések kizárása érdekében, hogy ebben a könyvben az aktív-passzív jelzőpárost csak a jogra vonatkoztatom, míg a pozitív-negatív minőségeket mind a jog, mind a cselekvés fogalma esetében arra használom, hogy a jog, illetve a cselekvés típusait képezzem a segítségükkel. Ez a kétfajta használat azonban erőteljesen összekapcsolódik, hiszen a pozitív és negatív jog (igény vagy szabadság) altípusait (pozitív, illetve negatív jog) a pozitív vagy negatív cselekvésre (tevésre vagy tartózkodásra) irányultság alapján definiálom.

modern generációs jogtipológiák

első-, másod- és harmadrendű jogok pozitív és negatív jogok

első generációs, klasszikus szabadságjogok, negatív jog

szólásszabadság - gondolat és kifejezés szabadsága vallásszabadság gyülekezési szabadság választójog sztrájkjog választhatóság joga a büntető igazságszolgáltatás visszaéléseivel szembeni védelem rendőri visszaélésekkel szembeni védelem joga a gyűlöletkeltő hátrányos megkülönböztetéssel szembeni védelem és ennek megelőzése tulajdonhoz való jog szerződési szabadság véleményszabadsághoz való jog megfelelő jogszolgáltatáshoz (due process) való jog nemi, faji, etnikai vagy vallási alapon történő megkülönböztetés tilalma tulajdonhoz való jog vállalkozáshoz való jog magánlakás sérthetetlenségéhez való jog jó hírnévhez való jog méltósághoz való jog

második generásciós, pozitív jogok (a szociális és gazdasági jogok):

a pihenéshez, a fizetett szabadsághoz, a táplálékhoz, a hajlékhoz, a minimálbérhez való jog egészsgéges környezethez való jó oktatáshoz való jog maximált munkaidő joga fizetett szülési szabadság ingyenes orvosi ellátáshoz való jog a legmagasabb szintű testi és lelki egészsgéhez való jog fogyatékosok képzéséhez való jog szociális biztonsághoz való jog

harmadik generációs, környezeti jogok tiszta, fenntartható környezethez való jog

negatív szabadság (privilégium): közeli rokon esetén a vallomástétel megtagadása, orvosi, ügyvédi vagy papi titoktartás, gyónási titok

Dénes Iván Zoltán [200] Benjamin Constant, A régiek és a modernek szabadsága [171] Hannah Arendt: A forradalom [27]

Egonak joga (szabadsága:FREEDOM) van sétálni Alter boltja előtt freedom(Ego, walk)

Egonak joga (szabadsága:FREEDOM) van bemenni Alter boltjába a nyitvatartási idő alatt

Alternek joga (felhatalmazottsága:AUTHOTITY) van megváltoztatnia a boltja nyitvatartási

idejét

Alternek joga (mentessége:IMMUNITY) van a boltja nyitvatartási idejének műsok általi megváltoztatásával szemben

Egonak joga (felhatalmazottsága:AUTHORITY) van publikálni saját kéziratát

Egonak joga (kiváltsága:IMMUNITY) van nem publikálni saját kéziratát

⁎⁎ ✻ ⁎⁎

Az igény, a kötelezettség, a szabadság és a joghiány fogalmainak, illetve ezek grafikus reprezentációjának bemutatása után következzék a hohfeldi jogfogalmak második csoportjának elemzése.

beavatkozás, kényszerítés

beavatkozásmentesség, kényszerítésmentesség interference, kényszerítés,

Makinson

Lindahl

Hart

10.2.2 Az elsőrendű jogviszonyfogalmak formalizálása

A hohfeldi jogfogalmak logikai leírását a skandináv logikai iskola tagjai végezték el a hatvanas-hetvenes években. A legfontosabb szereplője ennek a történetnek Stig Kanger, svéd logikus volt. Ebben a fejezetben Kanger formalizálási kísérletét mutatom be, majd rámutatok arra, hogy bár a Kanger által teremtett formalizmus konzisztens, ezért beemelhetőnek gondolom a társadalmi cselekvés általános modelljébe is, de a kangeri javaslat nem pontosan a hohfeldi fogalmakat ragadja meg. Ezek után megpróbálom korrigálni a kangeri elmélet problémáit, pótolni a hiányosságait. Eközben természetesen többször támaszkodni fogok Kanger követőinek észrevételeire, kritikáira, javaslataira.[6]

Hohfeld-Kanger-féle elemi jogtípusok

Stig Kanger az ötvenes évek végén és a hatvanas években írta meg azokat a tanulmányait, amelyek a hohfeldi jogviszonyok formállogikai feldolgozásával foglalkoztak [393], [394], [392]. Még a kangeri formalizmus bemutatása előtt hangsúlyoznom kell, hogy egy szempontból hibásnak tartom Kanger megoldását, amit mindenképpen javítani kell. Kanger ugyanis azzal, hogy egy logikai szintre helyezi el a két hohfeldi jogviszonycsoportot, elfogadhatatlan módon összecsúsztatja azokat, és így nem tudja a felhatalmazottság köré szerveződő jogviszonyfogalmakat jól megragadni. Ez nem azt jelenti, hogy a javasolt formalizmus nem alkalmazható. A logikai struktúra egy az egyben elfogadaható és átvehető, csak annyi változtatásra van szükség, hogy a kangeri sémát más fogalmak értelmezésére használjuk. Ebből persze az is következik, hogy mivel a felhatalmazottság körüli fogalomnégyest nem lehet figyelembe venni a kangeri séma interpretációja során, ezekre a fogalmakra új sémát kell kidolgozni.

[6]Kanger elméletét kimerítően tárgyalja Lars Lindahl könyve [449]. Kanger elméletén túl Lindahl bemutatja a Hohfed előtti legfontosabb szerzők jogfogalmakkal kapcsolatos nézeteit is, és saját elméletét is kifejti a jogi pozíciókról.

A hohfeldi jogviszonyfogalmak formalizálása és interpretációja előtt Kanger felállít egy logikai elméletet, egy öt axiómából álló sémát, amely két operátort definiál. Ezek segítségével fogja majd Kanger a hohfeldi jogfogalmakat meghatározni. Kanger a logikai sémája eredetileg ebből az öt formulából állt:

1)	$If\ F \longrightarrow G,\ then\ Shall\ F \longrightarrow Shall\ G$
2)	$(Shall\ F\ and\ Shall\ G) \longrightarrow Shall\ (F\ and\ G)$
3)	$Shall\ F \longrightarrow not\ Shall\ (not\ F)$
4)	$If\ F \longrightarrow G\ and\ G \longrightarrow F,\ then\ X\ causes\ F \longrightarrow X\ causes\ G$
5)	$X\ causes\ F \longrightarrow F$

Ez az öt axiómás elmélet erősnek mondható, amely deontikus logikai értelemben is megalapozottnak és jól-formáltnak tekinthető, valamint lehetővé teszi a hohfeldi jogviszonyfogalmak (illetve azok kiegészítéseinek) pontos formalizálását.[7] Később maga Kanger és a követői a két operátort az alábbi módon interpretálták:

$$\begin{array}{lcl} \textsc{Shall}\ F & = & \textit{it shall be that } F \\ X\ \textsc{causes}\ F & = & X\ \textsc{sees to it that}\ F\ \textit{holds} \end{array}$$

Az elméleti beágyazás megkönnyítése végett érdemes itt jelezni, hogy az első SHALL operátor a deontikus kötelezés (**O** – OBLIGATORY) operátorának feleltethető meg, a második 'sees to it that' pedig a cselekvés operátoraként értelmezhető (DO).

Mivel a jogfogalmak formalizálásához szükség van további összetevők jelölésére, Kanger a jogviszonyok típusaira temészetesen átveszi Hohfeld kategóriáit (bár a Right helyett a Claim, a Privilege helyett a Freedom terminusait használja). A jogviszonyhordozót (bearer, holder) jelöli X-szel, a jogi partnert, jogi társat (counter-party) Y-nal, és F(X,Y)-nal jelöli a jogviszony tartalmát, amit állapotként értelmez (object-matter, state of affairs, condition).

Fontos tény, hogy Kanger az állapotokat, nem pedig a cselekvéseket emeli be a formalizmusába. Mind a jogelmélet, mind a dentikus logika a cselekvésekre vonatkoztatja a jogviszonyokat (normatív viszonyokat). Hohfeld is is így értelmezi saját fogalmait (X-nek igénye van Y-nal szemben arra, hogy Y végezzen el egy cselekvést vagy tartózkodjon tőle. Kanger pedig egy jogviszonyt két személlyel és egy állapottal tart megragadhatónak. Jelöljük ezt a továbbiakban így:

$$R(X,Y,F(X,Y))$$

Ez a megoldás nagyobb kifejezőerővel bír. Hogy miért és hogyan, azt Lindahl alaposabban bemutatja könyvében [449]. Röviden: a nagyobb kifejezőerő onnan származik, hogy ha az állapotokra vonatkoztatjuk a jogviszonyokat, akkor el tudjuk különíteni egymástól azokat az eseteket, amelyekben egy állapot vagy annak tagadása áll fent. Egy pozitív vagy negatív cselekvés (tevés vagy tartózkodás) más és más jelentéssel bírhat attól függően, hogy egy állapotnak vagy annak komplementerének fennállására irányul. Egy aktív cselekvéssel létre lehet hozni egy állapotot vagy – épp' ellenkezőleg – meg lehet akadályozni, hogy fennálljon. Ugyanez fordítva is igaz lehet: egy tartózkodás eredményeként előállhat egy állapot, de meg is szűnhet (tehát az állapot ellentéte jöhet létre). Ezeket

[7] Erről bővebben lásd Lindahl könyvét [449, 38-54].

az eseteket a kangeri formalizmussal mind meg tudjuk ragadni. Ehhez arra van szükség, hogy a cselekvés operátor mellett az állapotokat vegyük be a jogviszonyok paraméterei közé.

Természetesen nem arról van itt szó, hogy Kanger a cselekvésektől eltekintene a jogviszonyok leírásakor, hanem inkább arról, hogy többet akar megragadni azáltal, hogy a cselekvések összetevőit is figyelembe veszi. Még egy lépést lehetne tenni ebben az irányban, ha a cselekvések összetevőkre bontását von Wright szellemében végeznénk el (↬272.o.). Von Wright a cselekvéseket az állapotok közti tranzíciókkal modellezte, vagyis egy cselekvés összetevői közé beemelte a kezdő és végállapotokat és a köztük fennálló átalakulás típusát. Kanger a végállapotot veszi figyelembe, ezáltal bővíti a modellje leíróerejét. Ha a von wright-i megoldást követnénk, akkor még finomabb modellhez juthatnánk, de félő, hogy túl finomszemcséssé válna már a modell és nehezebben áttekinthető, kevéssé használható lenne. Ezért itt csak a további finomítás elvi lehetőséget vetem fel, de nem törekszem a javaslat kibontására. A kangeri formulát a modellembe eddig bevezetett terminusokra átírva a következő alakot kapjuk (a második képlet megegyezik a 386. oldalon található *(iii)* jelű sor második formulájával):

$$R(X,Y,F(X,Y)) \quad \rightarrowtail \quad \text{RIGHT}(Agent, \pm\text{DO}(Agent, \pm F))$$

Fontos megjegyezni, hogy a kangeri formulához képest még van egy változás a második képletben. Kanger a cselekvés eredményeként "létrejövő" állapot leírásakor jelezni akarja azt is, hogy az adott állapot milyen ágensek között értelmezhető, ezért írja a képletébe az $F(X,Y)$ terminust. Ezt a két paramétert elhagyhatónak érzem, mivel Kanger érdemben nem használja ki a két paraméter alkalmazásában rejlő információt. Amikor a kangeri elméletet mutatom majd be, akkor – a "történeti hűség kedvéért – én is mutatni fogom ezeket a ágenciára utaló paramétereket, de bármikor, amikor továbbviszem valamilyen módon a kangeri formalizmust, akkor már elhagyom ezeket.

Szemléltetésképpen egy ábrában bemutatom, milyen módon lehetne formálisan leírni az *árucsere* aktusában rejlő jogviszonyokat úgy, hogy a von wright-i utat követjük (↬272.o.), amikor az állapotváltozásokat is beemeljük a formulába (10.1. ábra).
A hohfeldi jogviszonyokat négy egymásra épülő szinten írhatjuk le. Az első szinten az állapotok vannak, a másodikon az események (változások, átmenetek), a harmadikon a cselekvések, a negyediken a jogviszonyok. Az árucserét az állapot szinten úgy írhatjuk le, hogy a kezdőállapothoz képest $(p \wedge \neg q \wedge r \wedge \neg s)$ a végállapotban minden ellentétébe fordul $(\neg p \wedge q \wedge \neg r \wedge s)$. A két állapot közti átmenetet az esemény szint az alábbi formula írja le:

$$T(p \wedge \neg q, \neg p \wedge q, e_1) \wedge T(r \wedge \neg s, \neg r \wedge s, e_1) =$$
$$T(p, \neg p, e_1) \wedge T(\neg q, q, e_1) \wedge T(r, \neg r, e_1) \wedge T(\neg s, s, e_1)$$

A cselekvési szinten már külön kell leírni az – árucserét megvalósító – két szereplő cselekvését, hiszen két ágens cselekvése külön cselekvésnek számít. A cseréhez mindig két cselekvőre és két cselekvésre van szükség:

$$Do(Ego, T(p, \neg p, e_1) \wedge T(\neg q, q, e_1) \wedge T(r, \neg r, e_1) \wedge T(\neg s, s, e_1)) \wedge$$
$$Do(Alter, T(\neg p, p, e_1) \wedge T(q, \neg q, e_1) \wedge T(\neg r, r, e_1) \wedge T(s, \neg s, e_1))$$

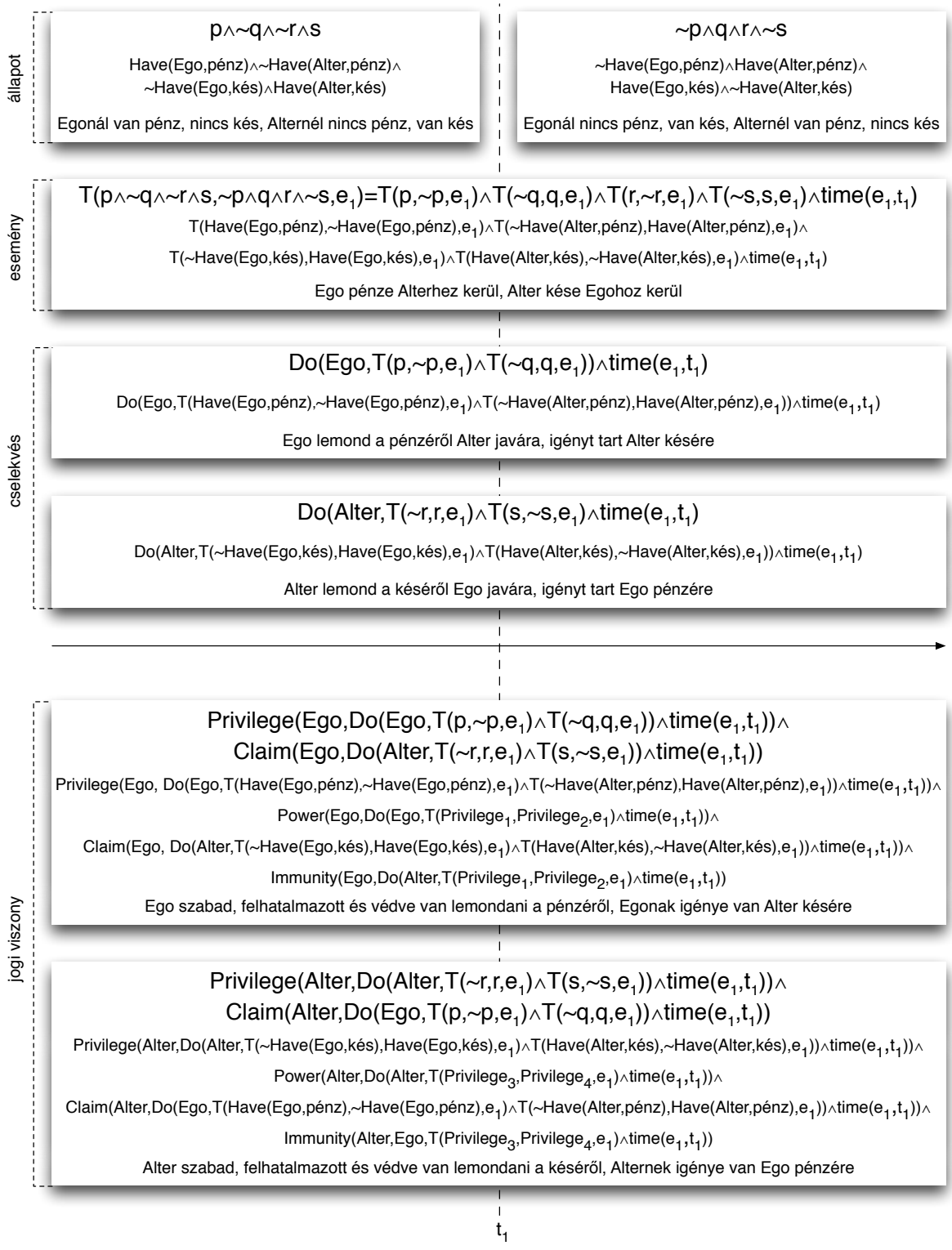


Figure 10.1: az árucsere logikai szerkezete

Ezután következhet a leírás negyedik szintje, ahol a jogviszonyokat reprezentálhatjuk. A négyféle jogviszony közül egyet kiemelve, Ego szabadságára vonatkozó formula a

következő:

$$Privilege(Ego, Do(Ego, T(p, \neg p, e_1) \wedge T(\neg q, q, e_1)) \wedge time(e_1, t_1))$$

Kanger jognak mond minden hohfeldi jogviszonyt (tehát a jogosultságokat és a kötelezettségeket egyaránt). A kötelezettségeket nem kezeli, legtöbb tanulmányában nem is nevezi meg őket, egyszerűen csak a jogosultságok tagadásaként hivatkozik rájuk. Viszont megduplázza a hohfeldi típusokat amiatt, hogy a jogosultság fogalmának egyik fogalmi összetevőjét, a két ágens közti – jogi – állapotot pontosabban megragadhassa, hogy ki tudja fejezni, vajon az állapot vagy annak tagadásának esete igaz-e. Az eredeti hohfeldi kategóriák közül tehát a tagadás segítségével "kiszorítja" a kötelezettségfogalmakat:

jogosultság	kötelezettség	(= jogosultság tagadása)
right	duty	(= not privilege)
privilege	no-right	(= not right)
power	liability	(= not immunity)
immunity	disability	(= not power)

A következő lépésben aztán (a right-claim, illetve privilege-freedom terminuscserék végrehajtása után) Kanger megduplázza a jogosultságok számát mondván, hogy a jogviszony más lesz akkor, ha a tartalmát az F(X,Y) állapottal vagy annak tagadásával írjuk le. Az ellentétes állapottal jellemezhető jogviszonyokat Kanger a 'counter-' előtaggal látja el, így a következő nyolc jogosultságtípust sorolja fel:

Hohfeld kategóriái	Kanger kiegészítései
claim	counter-claim
freedom	counter-freedom
power	counter-power
immunity	counter-immunity

Kanger a hohfeldi jogviszonyokat (illetve azok kiegészítéseit) **egyszerű jogtípusoknak** (**simple types of right**) nevezi. Zavaró lehet azonban az a tény, hogy amikor Kanger később a hohfeldi jogviszonyfogalmakból "összetett" jogtípusokat képez, akkor ezeket 'atomi jogtípusoknak' nevezi el. Ebből következően a kangeri elméleten belül a hohfeldi fogalmakat az atominál kisebbnek kell tételeznünk, ezért így is lenne érdemes megnevezni őket. Ha igazodni akarnánk Kanger névadási konvenciójához, akkor ezeket az egyszerű jogviszonyokat **elemi jogtípusoknak** (**elementary rights**) kellene inkább neveznünk.[8] A következőkben én így teszek.

Kanger logikai formalizmusa a hohfeldi jogviszonyfogalmak pontosabb megragadására alkalmas. Ennek egyik eredménye az lesz, hogy Kanger kétszer annyi fogalmat képes elkülöníteni, mint Hohfeld. Ehhez természetesen kétszer annyi terminusra van szükség. Ezt a bővítést a 'counter'-fogalmak felvételével valósítja meg Kanger. Kanger a hohfeldi fogalmak logikai formulákkal írja le, miközben arra az öt axiómából álló logikai sémára támaszkodik, amit korábban már bemutattam (↬396.o.). A SHALL és a DO fogalmak segítségével képezhető formulákkal akarja megfogni a hohfeldi jogviszonyfogalmakat. A

[8] Nevezhetnénk még szubatomi jogtípusoknak is ezeket.

passzív jog alá sorolható fogalmakat úgy lehet formulával leírni, hogy a jogviszony másik szereplőjének korrelatív kötelezettségfogalmával fejezzük ki (ezért minden ilyen viszonyt **partner kötelezettségnek** (COUNTER-PARTY OBLIGATIVE RIGHTS) vagy röviden **O-right**-nak is nevezhetünk – a deontikus kötelezés modalitása után). Kanger az alábbiakat javasolja:

claim(Ego,Alter,F(Ego,Alter))	SHALL DO$(Alter, F(Ego, Alter))$
counter-claim(Ego,Alter,F(Ego,Alter))	SHALL DO$(Alter, \neg F(Ego, Alter))$
immunity(Ego,Alter,F(Ego,Alter))	SHALL$\neg$DO$(Alter, \neg F(Ego, Alter))$
counter-immunity(Ego,Alter,F(Ego,Alter))	SHALL$\neg$DO$(Alter, F(Ego, Alter))$

Az aktív jog alá tartozó jogviszonyokat Kanger a deontikus megengedés modalitásával gondolja megragadni, ezért ezekre ráhúzhatjuk a **megengedő jogok** (PERMISSIVE RIGHTS) vagy röviden a **P-right** kifejezést is. Kanger ide a következő fogalmakat sorolja be:

power(Ego,Alter,F(Ego,Alter))	$\neg$SHALL$\neg$DO$(Ego, F(Ego, Alter))$
counter-power(Ego,Alter,F(Ego,Alter))	$\neg$SHALL$\neg$DO$(Ego, \neg F(Ego, Alter))$
freedom(Ego,Alter,F(Ego,Alter))	$\neg$SHALL DO$(Ego, \neg F(Ego, Alter))$
counter-freedom(Ego,Alter,F(Ego,Alter))	$\neg$SHALL DO$(Ego, F(Ego, Alter))$

Ez tehát a kangeri javaslat, ami a logikai formalizmust tekintve megfelelő, elfogadható, ám a formulákkal összekapcsolt fogalmak összerendelése önkényes. Kanger a második hohfeldi jogfogalom-csoportot (vagyis a felhatalmazottság köré szervezhető fogalmakat) is felhasználja, ám ezt úgy teszi, hogy ezek definiálásakor eltér Hohfeld eredeti intenciójától. A logikai formulák szintjén alkotott elmélet korrekt, tartalmas és teljesnek mondható, viszont a felhatalmazottság köré rendezhető fogalmakat beemelése annyiban elfogadhatatlan, mert Kanger és Hohfeld fogalomértelmezései e téren túl nagy mértékben térnek el egymástól. Ezzel Kanger komoly zavart okoz. Makinson szerint ezt már nem is annyira kölcsönzésnek, mint inkább lopásnak kell tartanunk („Kanger borrowed, or rather stole") [464, 408].

Már abból a mozzanatból is látszik, hogy Kanger más interpretációt ad a Hohfeldtől átvett terminusoknak, hogy Kanger a **claim** és az **immunity** jogfogalmakat állítja párhuzamba, és ezzel Hohfelddel ellentétesen cselekszik, utóbbi ugyanis az igényjog (**claim**) "analóg párjaként" a felhatalmazottságot (**power**) említi:

> „…Hohfeld fogalmainak pontos meghatározása érdekében a két csoport fogalmai közül igazából csak egyet-egyet kell explikálnunk – az igényjog és a mentesség fogalmait." [394, 101]

Ezzel szemben Hohfeld a két csoport "vezérfogalmaként" az igényjogot és a felhatalmazottságot emeli ki. Később fogom elemezni, bemutatni, hogy Kanger miért "tér el" ezen a ponton Hohfeldtől, itt csak annyit vetítek előre, hogy Kanger nem veszi észre, hogy a második hohfeldi jogviszonycsoport tagjait másodrendű fogalmakként lehet megragadni, ezért nem ismeri fel azt sem, hogy az új fogalmaknak kettős öröklődése és így kettős tulajdonsága lesz. A magasabbrendű fogalmiság abban a mozzanatban érhető tetten, ahogy Hohfeld a második jogcsoportot meghatározza. Az első és második jogviszonycsoport abban tér el egymástól – Hohfeld intenciója és meghatározása szerint –, hogy a második esetben a jogviszonyok szűkebb cselekvési tartományra vonatkoznak. Hohfeld

a **felhatalmazottság** fogalmával a jogviszonyok változtatásának lehetőségét, képességét akarta kifejezni, de a többi három új jogfogalom is más jogokra irányuló jogként értelmezendő. A **felhatalmazottság** és a **szabadság** egymáshoz való viszonyát tehát más szinten kell kezelni, és nem úgy, ahogy Kanger teszi, amikor a két fogalom közti különbséget a pozitív vagy negatív cselekvések (tevés vagy tartózkodás) hozzárendelésével írja le. Hétköznapi fogalmakra átfordítva Kanger formuláit azt kellene mondanunk, hogy a felhatalmazottság az, ami nem tilos és a szabadság az, ami nem kötelező, de ez egyáltalán nem tűnik intuitívnek. Újra hangsúlyozom: a logikai séma jó, csak az elemeihez fűzött interpretáció nem elfogadható.

A megoldás az lehet, hogy Kanger logikai sémáját megtartjuk, de a séma egyes elemeinek interpretálásakor nem alkalmazzuk a hohfeldi másodrendű fogalmakat. Ez azt jelenti, hogy a szabadság, igényjog, felhatalmazottág és mentesség helyett a szabadság és igényjog pozitív, illetve negatív altípusait használjuk a kangeri formulák értelmezéséhez. Ezzel a megoldással megmaradunk az elsőrendű logikai szintjén úgy, hogy a kangeri formalizmus erejét teljes mértékben ki tudjuk használni, így viszont a másodrendű hohfeldi fogalmak kikerülnek a formális modellből, ezért a későbbiekben még ki kell dolgozni az őket leíró formulákat.

POSITIVE-CLAIM(Ego,Alter,F(Ego,Alter))	SHALL DO$(Alter, F(Ego, Alter))$
COUNTER-POSITIVE-CLAIM(Ego,Alter,F(Ego,Alter))	SHALL DO$(Alter, \neg F(Ego, Alter))$
NEGATIVE-CLAIM(Ego,Alter,F(Ego,Alter))	SHALL$\neg$DO$(Alter, \neg F(Ego, Alter))$
COUNTER-NEGATIVE-CLAIM(Ego,Alter,F(Ego,Alter))	SHALL$\neg$DO$(Alter, F(Ego, Alter))$
NEGATIVE-FREEDOM(Ego,Alter,F(Ego,Alter))	$\neg$SHALL$\neg$DO$(Ego, F(Ego, Alter))$
COUNTER-NEGATIVE-FREEDOM(Ego,Alter,F(Ego,Alter))	$\neg$SHALL$\neg$DO$(Ego, \neg F(Ego, Alter))$
POSITIVE-FREEDOM(Ego,Alter,F(Ego,Alter))	$\neg$SHALL DO$(Ego, \neg F(Ego, Alter))$
COUNTER-POSITIVE-FREEDOM(Ego,Alter,F(Ego,Alter))	$\neg$SHALL DO$(Ego, F(Ego, Alter))$

Ez a kangeri séma annyiban féloldalas, hogy *Ego* nyolcféle jogviszonyát *Ego* és *Alter* négy-négyféle perszonális deontikus modalitással fejezi ki. Ez a megoldás a jogviszonyokat tekintve "teljesnek" mondható, de ha az ágensekre fókuszálunk (tehát Egora és Alterre), akkor mindkét esetben hiányos a fenti séma, hiszen mindkét ágens esetében csak feleannyi perszonális deontikus modalitást tartalmaz, mint amennyi valójában létezik. Ezt Kanger is tudta, kezelte is, hiszen – ahogy azt a következő fejezetben bemutatom majd – összesen tizenhat elemi jogviszonnyal számolt. A **kötelezés** ($\underline{\mathbf{O}}$) deontikus operátor és a DO cselekvésoperátor segítségével mindegyiküket megadhatjuk:

$\underline{\mathbf{O}}(x\text{DO}F)$	$\underline{\mathbf{O}}(x\text{DO}\neg F)$	$\underline{\mathbf{O}}(y\text{DO}F)$	$\underline{\mathbf{O}}(y\text{DO}\neg F)$
$\neg\underline{\mathbf{O}}(x\text{DO}F)$	$\neg\underline{\mathbf{O}}(x\text{DO}\neg F)$	$\neg\underline{\mathbf{O}}(y\text{DO}F)$	$\neg\underline{\mathbf{O}}(y\text{DO}\neg F)$
$\underline{\mathbf{O}}\neg(x\text{DO}F)$	$\underline{\mathbf{O}}\neg(x\text{DO}\neg F)$	$\underline{\mathbf{O}}\neg(y\text{DO}F)$	$\underline{\mathbf{O}}\neg(y\text{DO}\neg F)$
$\neg\underline{\mathbf{O}}\neg(x\text{DO}F)$	$\neg\underline{\mathbf{O}}\neg(x\text{DO}\neg F)$	$\neg\underline{\mathbf{O}}\neg(y\text{DO}F)$	$\neg\underline{\mathbf{O}}\neg(y\text{DO}\neg F)$

Makinson a fenti formulák általánosításaként egy olyan nyitott formulajavaslattal élt, amely segítségévek mind a tizenhat változatot kifejezhetjük [464]. A formulát bemutatom eredeti formájában is, illetve abban az alakban is, amelyben az ágensszerepek a modellemben használatos nevekkel vannak jelölve. A formulában látható $\pm$ jel azt fejezi ki, hogy vagy pozitív vagy negatív (tehát állító vagy tagadó) értelemben használjuk-e azt az összetevőt, amely előtt a jel áll.

$$(\pm)\underline{\mathbf{Q}}(\pm)\begin{pmatrix} X \\ Y \end{pmatrix}\text{DO}(\pm)F \quad \twoheadrightarrow \quad (\pm)\underline{\mathbf{Q}}(\pm)\begin{pmatrix} Ego \\ Alter \end{pmatrix}\text{DO}(\pm)F$$

Az a fenti javaslatom egyébként, amelyet a szabadság és igényjog pozitív és negatív altípusainak modellbe emelésére vonatkozóan teszek, megegyezik Jeremy Bentham elképzelésével is [449, 47].[9] A továbbiakban úgy mutatom be Kanger elméletének hátralevő részét, hogy az általa használt eredeti fogalmakat, terminusokat alkalmazom, de saját modellembe az itt jelölt fogalmakat fogom átemelni.

Kanger atomi jogtípusai

A jogok tipizálása már Jeremy Bentham számára is fontos volt [449, 63]. Bentham három jogtípust különített el, a **kötelezés**, a **tiltás** és a **nem tiltás és nem kötelezés** típusait. Ez a hármas felosztás megfelel a JEPD-elvnek, vagyis a három kategória terjedelme nem fedi át egymást, együttesen viszont kiadják a felettes fogalom (a jog) teljes terjedelmét. A hohfeldi jogviszonyfogalmak nem felelnek meg ennek az elvárásnak. Ezzel tisztában volt Kanger is, ezért a jogfogalmakra JEPD-kompatibilis felosztást dolgozott ki. Kanger kiindulópontja az volt, hogy az általa bevezetett nyolc elemi jogtípus, illetve ezek komplementer párjai összetételével összetett jogi viszonyokat lehet képezni. A hohfeldi-kangeri elemi jogviszonyok összekapcsolásával atomi jogviszonyokat definálhatunk, mégpedig úgy, hogy a 26 egymástól különböző, tehát egymást át nem fedő atomi jogtípus összességében teljesen kitölti a jog fogalmának teljes terjedelmét. Nézzük meg, hogy jutott el Kanger a 26 atomi jogtípusáig. A nyolc elemi összetevőt (a nyolc jogosultságot és kötelezettséget), valamint ezek tagadásait logikailag 2^8 módon tehetjük össze, így 256-féle, elméletileg lehetséges elemi típusunk (possible atomic types) lehet. Ha ezeket mind feírnánk egy táblázatba, akkor 256 sort kapnánk, és valahogy így nézne ki a táblázat:

	1.	2.	3.	4.	5.	6.	7.	8.
1.	claim	freedom	power	immunity	counter-claim	counter-freedom	counter-power	counter-immunity
2.	claim	not-freedom	power	immunity	counter-claim	counter-freedom	counter-power	counter-immunity
3.	claim	freedom	not-power	immunity	counter-claim	counter-freedom	counter-power	counter-immunity
4.	claim	freedom	power	not-immunity	counter-claim	counter-freedom	counter-power	counter-immunity
5.	claim	freedom	power	immunity	not-counter-claim	counter-freedom	counter-power	counter-immunity
...								
255.	not-claim	not-freedom	not-power	not-immunity	not-counter-claim	not-counter-freedom	not-counter-power	counter-immunity
256.	not-claim	not-freedom	not-power	not-immunity	not-counter-claim	not-counter-freedom	not-counter-power	not-counter-immunity

Kanger szerint azonban meg lehet szűrni ezeket a sorokat, mert vannak köztük olyanok,

[9] Annyi különbséggel csak, hogy Bentham nem foglalkozott a Kanger által javasolt "counter"-fogalmakkal.

amelyek vagy egymásnak ellentmondó elemeket tartalmaznak, így nem fordulhatnak elő, vagy egymással ekvivalensnek tekinthetők. Az ellentmondásos sorok kiesése már eleve csökkenti a 256-os számosságot, a további szűrés lehetőségét az teremti meg, hogy logikai összefüggések vannak egyes elemi jogtípusok között, és így némelyik sorban előfordulhat, hogy a nyolc összetevő közül néhányat el lehet hagyni a sorban szereplő elemi jogtípusok közötti logikai összefüggés miatt. Kanger egy ábrában mutatja meg, ezeket a logikai összefüggéseket.

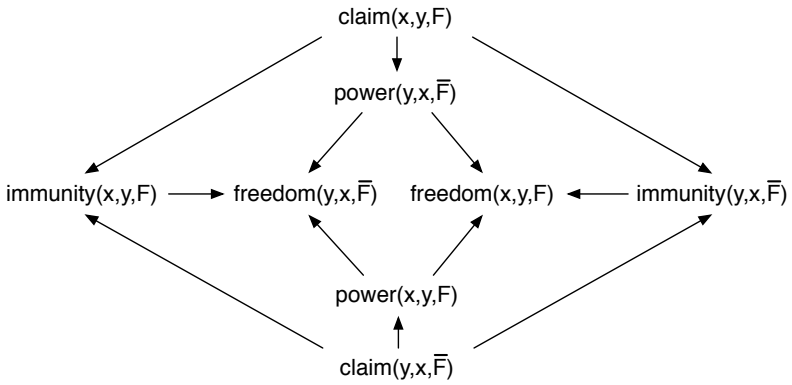


Figure 10.2: logikai összefüggések típusok között

Ha a fenti logikai kapcsolatok miatt mind a 256 sorból elhagyjuk a logikailag fölösleges összetevőket, akkor előfordulhat, hogy kettő vagy több sor meg fog egyezni. A 256 sor eredetileg mind különbözött egymástól, de a "ritkítás" után kevesebb egyedi sort kell kapjunk. Kanger szerint összesen 26 különböző sor marad a végén. Van olyan sor, amelyben kettő, de van olyan is, amelyben hat elemi jogtípus marad meg. Ezeket nevezi Kanger **atomi jogtípusoknak**. Az atomi jogtípusok és elemi összetevőik listája az alábbi táblázatban olvasható:
A lista meglepően hosszú. A konkrét atomi jogtípusok pedig nem igazán mondhatók intuitívnek. Nehezen értelmezhetőek a hétköznapi fogalmainkkal, tapasztalatainkkal, amit az is nehezít, hogy Kanger sajnos nem ad mindegyikre példát. Az atomi jogtípusok között azért léteznek összefüggések, és ezek felfedésével Kanger kicsit segít a 26 jogtípusból álló rendszer befogadásában.

Kanger ugyanúgy értelmezi saját atomi jogtípusai között azokat a relációkat, amelyeket Hohfeld tételezett a jogosultságok és kötelezettségek között. Saját formális elméletére támaszkodva azonban Kanger megállapítja (helyesen), hogy a hohfeldi korrelativitás relációja két altípusra osztható. Attól függően, hogy a korrelatív kapcsolatban levő jogviszonyok tartalma megegyezik-e vagy egymás tagadásai, Kanger **inverziónak**, illetve **konverziónak** nevezi a korrelatív relációt.[10] Az inverzió definícióját így adja meg Kanger:

$$\text{X has versus Y a right of type P that S(X,Y)} \leftrightarrow$$
$$\text{Y has versus X a right of type Q that S(X,Y)}$$

míg a konverzió fogalmát az alábbi módon határozza meg:

[10] Kanger eltérő módon definiálja az inverzió és konverzió műveletét ahhoz képest, ahogy a könyv logikai alapozó részében én bemutattam a két fogalmat (↬43.o.).

1)	Power, not immunity, counter-power, not counter-immunity
2)	Not power, immunity, not counter-power, counter-immunity
3)	Claim, not counter-freedom
4)	Not claim, power, immunity, counter-freedom, not counter-power, not counter-immunity
5)	Power, immunity, counter-power, counter-immunity
6)	Claim, power, counter-freedom
7)	Claim, not power
8)	Power, immunity, counter-freedom, not counter-power, counter-immunity
9)	Power, immunity, counter-power, not counter-immunity
10)	Power, not immunity, not counter-power, counter-immunity
11)	Not freedom, counter-claim
12)	Freedom, not power, not immunity, not counter-claim, counter-power, counter-immunity
13)	Freedom, counter-claim, counter-power
14)	Counter-claim, not counter-power
15)	Freedom, not power, immunity, counter-power, counter-immunity
16)	Power, not immunity, counter-power, counter-immunity
17)	Not power, immunity, counter-power, not counter-immunity
18)	Not power, not immunity, not counter-power, not counter-immunity
19)	Not claim, not counter-freedom, not counter-immunity
20)	Not counter-freedom, counter-immunity
21)	Not claim, not power, immunity, not counter-power, not counter-immunity
22)	Power, not immunity, not counter-power, not counter-immunity
23)	Not freedom, not immunity, not counter-claim
24)	Not freedom, immunity
25)	Not power, not immunity, not counter-claim, not counter-power, counter-immunity
26)	Not power, not immunity, counter-power, not counter-immunity

X has versus Y a right of type P that S(X,Y) ↔
Y has versus X a right of type Q that not-S(X,Y)

Kanger a fenti két "korrelativitási" reláció mellé felvesz még egy harmadik relációt, a **koordinációt**, de ezt itt nem tartom fontosnak bemutatni.

Kanger a 26 atomi jogtípust az alábbi ábrába rendezte:
A szokatlan alakú táblázat arra jó, hogy jelezze az egyes kangeri atomi jogtípusok közti kapcsolatokat. Az .... számú típusok rendre konverz kapcsolatban vannak a .... típusokkal, míg a ... típusok inverz viszonyban állnak a ... típusokkal.

10.2.3 Hohfeld másodrendű jogviszonyfogalmai

A hohfeldi elmélet bemutatásának talán nem "tett jót", hogy nem egyhuzamban mentem végig a hohfeldi kategóriákon, de a logikai építkezés szempontjából rákényszerültem erre. A Kanger által javasolt formalizmust én csak az elsőrendű hohfeldi fogalmakra tartom alkalmazhatónak. Ezért "mozogtam" eddig csak az elsőrendű logikai szinten. A további hohfeldi jogviszonyfogalmak tárgyalásához viszont fel kell majd lépnünk másodrendű fogalmak szintjére, és ez a tény megkívánja (és indokolja is egyben) a két szakaszra osztott tárgyalásmódot.

A felhatalmazottság köré szerveződő fogalmak értelmezése

Hohfeld a második – magasabbrendű – jogviszonycsoportba a **felhatalmazottság** (**power**), a **mentesség** (**immunity**), a **beavatkozásnak kitettség** (**liability**) és a **beavatkozásképtelen-**

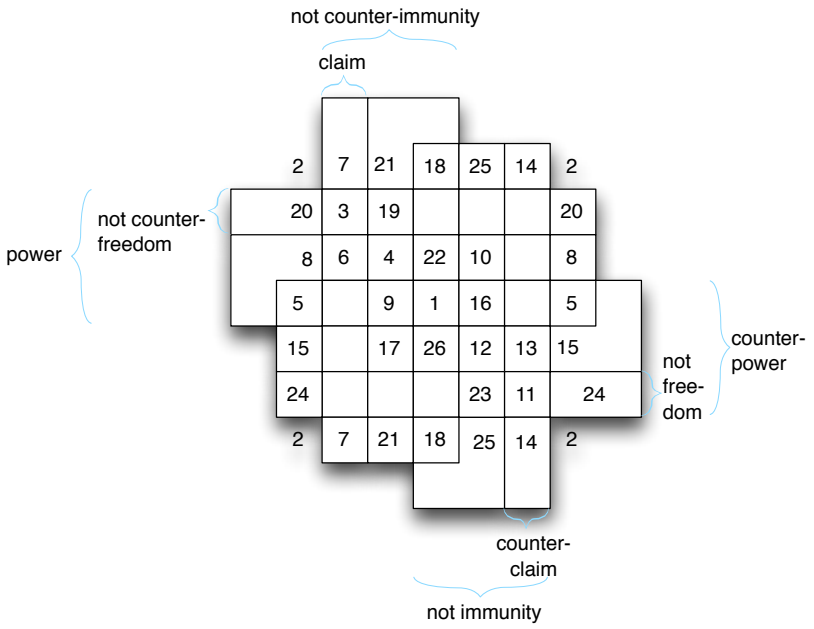


Figure 10.3: Kanger 26 atomi jogtípusa

ség (**disability**) kategóriáit sorolta be. Ebben a fogalmi csoportban a felhatalmazottság az a "kulcsfogalom", ami köré a többi három fogalmat szervezve az első csoport esetében megismert struktúrához jutunk újra. A jogi felhatalmazottság a jogi beavatkozásra képtelenség ellentéte és a jogi beavatkozásnak kitettség korrelatív párja. Több részletet is idézek itt, hogy a példák segítségével jobban érthetővé váljék, mit gondol Hohfeld a felhatalmazottságról:

> „Egy meghatározott jogi viszonyban bekövetkező változás fakadhat (1) olyan új tényből, illetőleg ilyen tények együtteséből, amely kívül esik az emberi lény(ek) akaratlagos befolyásán; vagy (2) olyan új tényből, illetőleg tények együtteséből, amely egy vagy több emberi lény akaratlagos befolyása alá esik. Az esetek utóbbi csoportjában az a személy (vagy azok a személyek), aki(k)nek akaratlagos befolyása meghatározó, (jogi) felhatalmazottsággal bír(nak) arra, hogy meghatározott változást idézzen(ek) elő a kérdéses jogviszonyban." [360, 84]

> „*X* fel van hatalmazva arra, hogy hogy *Y*-ra ruházza érdekeltségét – azaz megszüntesse saját érdekeltségét, és ezzel együtt új és azonos érdekeltséget keletkeztessen *Y* javára. *X* tehát felhatalmazottsággal rendelkezik különböző fajta szerződéses kötelezettségek keletkeztetésére." [360, 85]

> „A megbízási viszony keletkezése *inter alia* magában foglalja a jogi felhatalmazottság átruházását a megbízottra, s az ennek megfelelő beavatkozásnak kitettséget a megbízó oldalán. Másként fogalmazva: az egyik fél – *P* – felhatalmazottsággal rendelkezik arra, hogy megbízotti felhatalmazást hozzon létre a másik fél – *A* – oldalán, például arra, hogy átruházza *P* tulajdonát, úgynevezett szerződéses kötelezettséget keletkeztessen *P* terhére, feloldjon egy *P*-vel szemben fennálló tartozást, 'megszerezzen' egy tulajdonosi jogcímet *P* javára, és így tovább." [360, 85-86]

Bár Hohfeld a definícióban egyértelművé teszi, hogy a felhatalmazottság a jogi viszonyok változtatására irányul, igazából nem – vagy legalábbis nem eléggé – reflektál erre a mozzanatra. Pedig ez az új fogalmi négyes pontos értelmezését tekintve kulcsfontosságú. A felhatalmazottságot és a kapcsolódó másik három jogfogalmat ugyanis nem a logikai minőségét tekintve különíti el Hohfeld az igényjog köré épített fogalmaktól, hanem "csak" a fogalmak értelmezési tartományaiban tesz különbséget közöttük.

> „Míg a joghelyzetek első két párja elsőrendű viszonyt, addig a második kettő másodrendűt képez. Az elsőrendű viszonyok közvetlenül az emberek viselkedésére és kapcsolataira vonatkozhatnak, a másodrendű viszonyok közvetítése nélkül; ezzel szemben a másodrendű viszonyok közvetlenül csak az emberek jogcímeire vonatkoznak, s csak áttételesen (bár meghatározó módon) az emberek viselkedésére és kapcsolataira." [419, 20]

Magasabbrendű normákról von Wright is ír a normák formális elemzésekor [792, 189-207], bár ő nem a hohfeldi fogalmak közti viszonyok esetében látja alkalmazhatónak a magasabbrendű logikát. Azt, hogy miként lehet a másodrendű fogalmakat ráhúzni a hohfeldi elméletre, rövidesen bemutatom, de előtte még meg kell ismernünk Hohfeld meghatározásait a felhatalmazottság köré építhető jogviszonyfogalmakra.

A jogi felhatalmazottságra ott lehet példákat keresni, ahol jogi viszonyok változtatásáról van szó. Ilyen egy miniszter kinevezése vagy felmentése, ilyen a házasság, az örökbefogadás, a végrendelkezés, az örökség visszautasítása, az elővételi jog gyakorlása, a megbízási érdekek képviselete, az ajánlati kötöttség, a telki szolgalom biztosítása, de minden tulajdonviszony jogszerű változtatását eredményező aktus is jogi felhatalmazottságon alapul. Lehet felhatalmazottságból eredő jogviszonyváltozás az, amikor valaki elcseréli a lakását két másikra vagy eladja a boltját, de a felhatalmazottság megnyilvánulhat abban is, ha a boltos megváltoztatja a nyitvatartási idejét, a tulajdonos bezárja a vendéglőjét vagy valaki lemond valamely tulajdonáról.

A felhatalmazottság valamiféle jogosultságot jelent, amivel ha él valaki, akkor a felhatalmazottság tartalma, valamely konkrét cselekvés megvalósul, és ezt a többiek tudomásul veszik, "jogosnak" tartják. Ez természetesen nem jelenti azt, hogy ne lehetne valamely cselekvést megtenni úgy, hogy az embernek nincs rá felhatalmazottsága. Ékes példa erre az alábbi jól ismert történet:

> „Hamrák egy újságban olvasta, hogy Székesfehérvár vezetői megelégelték, hogy a piactéren vagy száll a por, vagy öles pocsolyákban tobzódnak a kofák, és elhatározták, hogy faburkolattal látják el a teret. Emberünk szépen levonatozott Fehérvárra, ahol egy kávéházban elhitette néhány fiatal városatyával, hogy a fa ára történelmi emelkedés előtt áll, ám neki van éppen tizenöt vagonnyi használt fatéglája, amit hajlandó 9000 aranykoronáért eladni. A hálás hivatalnokok rögtön a kezébe is nyomtak ezer aranykoronát előlegként, hősünk pedig már másnap munkásokat toborzott Pesten a Rottenbiller utca faburkolatának felszedésére.
>
> A módszer egyszerű, de nagyszerű volt; innen-onnan összeszedett néhány úttorlaszt, amikkel a Baross-térnél és a Lövölde-térnél lezárta az utcát, aztán a munkásokat rászabadította a faburkolatra. (Ekkor még sok helyen a macskaköves megoldásnál sokkal olcsóbb fakockás útburkolatot használták.) Mindenkinek azt mondta, hogy a városvezetésnek dolgozik, mely a faburkolat bazaltkőre cserélése mellett döntött. A munkások hat nap alatt szedték föl a majdnem egy kilométeres útszakasz

> burkolatát, amit a Kelenföldi pályaudvarnál vonatra raktak és elszállítottak Székesfehérvárra. A legszebb az egészben, hogy az egész művelet alig került 600 aranykoronába, így még az előlegből is maradt, a sikeres üzlet után pedig megkapta Székesfehérvártól a fennmaradó 8000 aranykoronát is." [5]

A történet értelmezésekor azt kell mondanunk, hogy Hamrák úrnak "nem volt felhatalmazottsága" az útburkolat felszedésére, értékesítésére, amit azonban úgy is mondhatnánk, hogy "nem volt megengedve neki" mindaz, amit megtett. Ez a helyettesíthetőség önmagában jelzi, hogy a felhatalmazottságot milyen deontikus logikai fogalomhoz lehet majd kötnünk, de ezt csak a későbbiekben fogom kifejteni.

A felhatalmazottság korrelatív párja a **beavatkozásnak kitettség** (LIABILITY) és ellentéte a **beavatkozásképtelenség** (DISABILITY). A beavatkozásnak kitettségre Hohfeld az alábbi példát adja:

> „A Booth v. Commonwealth esetben a bíróságnak Virginia egyik törvényi rendelkezését kellett értelmeznie, mely szerint «minden szabad fehér férfi, aki betöltötte huszonegyedik életévét és nem töltötte be a hatvanadikat, esküdtként köteles közreműködni, az alábbi kivételektől eltekintve». Világos, hogy e törvény csak beavatkozásnak kitettséget keletkeztetett, nem pedig kötelezettséget. Az beavatkozásnak kitettség, ha kötelezettséget róhatnak ránk. Ez utóbbi csak akkor következik be, amikor felhatalmazottságukkal élve a peres felek és a bíróság tagjai megtették mindazt, ami az esküdt funkciójának ellátására vonatkozó specifikus kötelezettség létrejöttéhez szükséges." [360, 91]

Az esküdtszéki munkában való részvétel mint a beavatkozásnak kitettség hohfeldi mintapéldájával párhuzamba állíthatjuk a katonai szolgálat feltételes kötelezettségét, ami számunkra talán ismerősebb lehet. Ha a beavatkozásnak kitettség a felhatalmazottság korrelatív párja, akkor minden olyan példa, amit a felhatalmazottság esetében adtam meg korábban, adekvát lehet a beavatkozásnak kitettségre is (természetesen az ágensszerepek megcserélésével).

A beavatkozásnak kitettség alanyának nem feltétlen jelent hátrányt az a jogviszonyváltozás, amit a partnere okoz felhatalmazottságával élve. Ha valaki lemond a tulajdonáról (mondjuk, az órájáról), amire természetesen fel van hatalmazva, akkor ez a jogviszonyváltozás előnyös lehet mindenki más számára, hiszen ezzel az aktussal bárki szabadságot és felhatalmazottságot nyer arra, hogy eltulajdonítsa az órát [360, 92-93]. Ezt a jelenséget figyelhetjük meg a nagyvárosi lomtalanítások során.[11] Azzal, hogy a tulajdonosok kiteszik a szükségtelenné vált ingóságaikat (bútoraikat, használt ruháikat, háztartási eszközeiket, limlomjaikat) a közterületre, egyben jelzik azt is, hogy lemondanak e tulajdonaikról, és bárki viheti azokat. Ekkor potenciálisan mindenki beavatkozásnak kitetté válik, ami azt jelenti, hogy mindenkinek tudomásul kell vennie, hogy a kirakott ingóságok fölötti jogosultság megváltozott, és onnantól kezdve bárki eltulajdoníthatja azokat (akár az eredeti tulajdonos is, bár ez nyilván nem valószínű opció). Az már egy másik kérdés, hogy ki lehet az új tulajdonos. Robert Sugden érdekes tanulmánya szerint az ilyen helyzetekben az elsőségnek lehet perdöntő szerepe [715].

[11] A magyar hulladékgazdálkodási törvény 2013. január elsejével megváltozott, és ekkortól a lomtalanítás során közterületre kirakott hulladék a közszolgáltató tulajdonába kerül, de korábban nem így volt, és sok más országban is szabadon vihető az az ingóság, amiről lomtalanításkor lemondott a gazdája.

Ha az eddig bemutatott két fogalmat, a felhatalmazottságot és a beavatkozásnak kitettséget összetevőikre bontjuk (ahogy tettük ezt az igényjoghoz kapcsolódó jogviszonyfogalmak esetében is), akkor a következő táblázathoz juthatunk:

jogosult	jogviszony	cselekvő	cselekvés
Ego	felhatalmazottsága	Ego	eladni Ego telkét
Alter	beavatkozásnak kitettsége	Ego	eladni Ego telkét

Ha felidézzük az elsőrendű hohfeldi jogfogalmak hasonló jellegű táblázatát (↬386.o.), akkor magyarázatra váró eltéréseket tapasztalhatunk. Ha ugyanis a felhatalmazottságot akár a szabadság, akár az igényjog fogalma alá akarjuk besorolni, mindenképpen eltérő mintázatot kapunk az első és másodrendű fogalmak összetevőikre bontott táblázataiban (a magyarázatot később adom meg).

A **mentesség** (IMMUNITY) mint jogosultság a beavatkozásnak kitettség ellentéte és a beavatkozásképtelenség korrelatív párja. Mindenféle adómentesség, végrehajtás alóli mentesség, büntethetőséget kizáró ok fennállása, szabályok alóli kivétel e fogalom alá tartozik. Hohfeld példája a következő:

> „*X*, egy ingatlan tulajdonosa . . . fel van hatalmazva, hogy elidegenítse az *Y* vagy bárki más számára. Ugyanakkor *X*-et számos mentesség is megilleti *Y*-nal vagy bárki mással szemben. *Y* ugyanis képtelen beavatkozni (azaz nincs felhatalmazottsága), amíg érdekeltséget nem szerez önmaga vagy bárki más számára. S ami igaz *Y*-ra, az hasonlóképpen igaz bárki másra is, aki sajátos keletkeztető tények folytán felhatalmazottságot nem szerez *Y* tulajdonának elidegenítésére.” [360, 93]

A Hohfeld-recepció jóval kevesebbet foglalkozik a mentesség fogalmával a felhatalmazottsághoz képest. A fogalom tartalmát akkor lehet jól megérteni, ha arra fókuszálunk, hogy a mentesség a beavatkozásnak kitettség ellentéte. A diplomáciai mentesség lehet talán a legnyilvánvalóbb példája ennek. A diplomáciai rendszámmal rendelkező kocsikat nem lehet büntetni nyilvánvaló szabálytalanság esetén sem, vagyis nem lehet a használójára, tulajdonosára nézve kötelező jellegű jogi döntést hozni, nem lehet ő(ke)t jogi beavatkozásnak kitenni.

A **beavatkozás-képtelenség** (DISABILITY) a felhatalmazottság ellentéte (annak hiánya) és a mentesség korrelatív párja. Példaként lehet hozni rá a hatóságok hatáskörének vagy illetékességének hiányát valamely ügyben, a bérbeadó rendelkezési jogának szűkítését a bérbe adott dolog fölött vagy a rendőr beavatkozásképtelenségét a szabálytalankodó, ám diplomáciai rendszámmal rendelkező autóssal szemben, de idetartozik az is, hogy a tulajdonom fölött más nem rendelkezhet, nem adhatja el, nem használhatja, hacsak én nem adtam erre engedélyt.

A mentesség és beavatkozásképtelenség fogalmainak összetevőkre bontása eredményeként az alábbi táblázatot kapjuk:

jogosult	jogviszony	cselekvő	cselekvés	
Ego	mentessége	Alter	nem	eladni Ego telkét
Alter	beavatkozásképtelensége	Alter		eladni Ego telkét

Ebben az esetben is ugyanazt tpasztalhatjuk a táblázat mintázatáról, mint a felhatalmazottság/beavatkozásképtelenség esetében: az elsőrendű fogalmak összetevőkre bontásához képest részben eltérő, részben egyező mintázat jellemzi a most elemzett fogalmakat. Jobban láthatóvá válik mindez, ha egyesítem a másodrendű fogalmak táblázatait:

jogosult	jogviszony	cselekvő	cselekvés	
Ego	felhatalmazottsága	Ego		eladni Ego telkét
Alter	beavatkozásnak kitettsége	Ego		eladni Ego telkét
Ego	mentessége	Alter	nem	eladni Ego telkét
Alter	beavatkozásképtelensége	Alter		eladni Ego telkét

illetve megmutatom újra, hogy nézett ki ugyanez az elsőrendű fogalmakra vonatkozóan (a sorokat némiképp átrendezve, hogy az összetartozó fogalompárok a táblázatok ugyanazon soraiban legyenek láthatóak):

jogosult	jogviszony	cselekvő	cselekvés	
Ego	szabadsága	Ego		átjárni Ego telkén
Alter	igényhiánya	Ego	nem	átjárni Ego telkén
Ego	igénye	Alter	nem	átjárni Ego telkén
Alter	kötelezettsége	Alter	nem	átjárni Ego telkén

A két táblázat összehasonlításából látszik, hogy az első-, illetve másodrendű fogalmak esetében más a korrelatív fogalmak viszonya egymáshoz. Elsőrendben az igény és kötelezettség között ugyanarra a cselekvésre irányult a jogviszony, míg a szabadság és igényhiány esetében a cselekvés és annak tagadására. A korrelativitás fogalmában megbúvó kettősséget kezelte le Stig Kanger az inverz és konverz relációk elkülönítésével (↬403.o.). Kanger szerint így elsőrendben az igény és kötelezettség közti korrelativitás inverziót jelent, míg a szabadság és igényhiány közti kapcsolat konverzió. A másodrendű fogalmak esetében a felhatalmazottság, illetve a beavatkozásnak kitettség között láthatjuk az inverz viszonyt, és a menteség, illetve beavatkozásképtelenség között a konverziót érvényesülni.

Mindez azért is problematikus, mert ha azt keressük, hogy a másodrendű fogalmakat melyik elsőrendű fogalmak alá sorolhatjuk be, akkor ellentmondásos helyzethez jutunk. Ha a felhatalmazottságot a szabadság alá rendeljük azon az alapon, hogy mindkettő **aktív jognak**[12] számít, akkor a korrelatív párjaikhoz való viszonyban lesz eltérés (a felhatalmazottság inverz, a szabadság konverz viszonyban van a párjával). A mentesség esetében is hasonlót tapasztalhatunk. A **passzív jogi** minősítésük okán a mentesség és az igényjog között kell a generikus kapcsolatot feltételeznünk, ám ha így teszünk, ez esetben is eltérő lesz a korrelativitás tartalma köztük.

Meg kell említenem egy további problémát Hohfeld fogalommeghatározásaival kapcsolatban. Hohfeld ugyanis hangsúlyosan elkülöníti egymástól a kötelezettséget és a beavatkozásnak kitettséget, de nem tisztázza kellő alapossággal azt, mi a köztük lévő különbség lényege. Egy helyen azt írja, hogy:

> „A beavatkozásnak kitettség nem más, mint felelősség.” [360, 91]

[12] Az aktív és passzív jog meghatározását lásd 390. oldaltól.

de ez a mondat nem igazán segíti a tisztánlátást.[13] A felhatalmazottság és a beavatkozásnak kitettség magyarázata során kétféle értelemben is félrevezetőnek, sőt hibásnak tartom a hohfeldi megközelítést, amit sajnos a későbbi Hohfeld-recepció is folyamatosan életben tartott. Egyfelől ezen a ponton kerül be az nézet a hohfeldi elméletbe, miszerint a felhatalmazottság (illetve korrelatív párja a beavatkozásnak kitettség) nem jogok, hanem kapacitások, képességek létezéséről vagy azok hiányáról szól. Az angol terminusok közül a **disability** már önmagában is ezt a "sugallja". Ennek kritikájára a következő fejezetben fogok kitérni. Másfelől a felhatalmazottság és a beavatkozásnak kitettség kapacitív fogalmait Hohfeld (és utána még sok más szerző is) úgy értelmezi, mintha abban definitív szerepe lenne a feltételességnek. Megismétlem az előző Hohfeld-idézet egy részletét:

> „. . . beavatkozásnak kitettség, ha kötelezettséget róhatnak ránk. Ez utóbbi csak akkor következik be, amikor felhatalmazottságukkal élve a peres felek és a bíróság tagjai megtették mindazt, ami az esküdt funkciójának ellátására vonatkozó specifikus kötelezettség létrejöttéhez szükséges." [360, 91]

Az idézetből úgy tűnik, mintha Hohfeld szerint a beavatkozásnak kitettség feltételes kötelezettség lenne. Ez a feltételesség azonban nem köthető kizárólagosan a beavatkozásnak kitettséghez, előfordulhat az "egyszerű" kötelezettségek esetén is. Ha én – felhatalmazottságommal élve – kölcsönadom valakinek egy könyvemet, akkor folyamatos igényem (a partneremnek pedig folyamatos kötelezettsége) lesz arra, hogy, amikor kérem tőle, adja vissza. Ezt értelmezhetem úgy is, hogy a partnerem folyamatos beavatkozásnak kitettségben van és csak akkor keletkezik kötelezettsége arra, hogy visszaadja a könyvemet, amikor én – felhatalmazottságommal újra élve – ezt valahogy jelzem felé. Amíg azonban nem kérem vissza, addig csak feltételes, potenciális kötelezettséggel rendelkezik.

Az itt jelzett problémákat csak akkor tudjuk megoldani, ha megadjuk a másodrendű fogalmak formális definícióit, és így láthatóvá tesszük egyfelől azt, hogy milyen módon kapcsolódnak, hasonlítanak az elsőrendű fogalmakhoz, illetve másfelől bemutatjuk azt, hogy milyen új vonások felvételével lehet őket saját minőségükben megragadni. A következő fejezetben ezt kísérlem meg bemutatni.

10.2.4 A másodrendű jogviszonyfogalmak formalizálása

A felhatalmazottság köré építhető hohfeldi fogalmak eddigi elemzéséből kiderült, hogy *(i)* ezek fogalmak másodrendűek, *(ii)* ellentmondás látszik a passzív vagy aktív jog szerinti osztályozásuk, az igényjog vagy szabadság alá való besorolásuk és a korrelatív kapcsolataik eltérő minősége között, valamint *(iii)* a feltételesség, az állapot- vagy pontszerűség kérdésében eltérő minőséget lehet érezni a komplementer fogalompárok között.

A Hohfeld-recepcióban sokan úgy gondolják, hogy azon az alapon különíthetjük el az első- és másodrendű fogalmakat egymástól, hogy az első csoportot jogosultságnak, a másodikat képességnek (potenciálnak, kapacitásnak) minősítjük. Sok példát lehetne hozni arra, hogy miben különbözik a két hohfeldi jogviszonycsoport, vagy másként a

[13] Az elemzett téma szempontjából nem tartom szükégesnek bővebben bemutatni, de jelezni mindenképpen érdemes, hogy a felelősség fogalmának elemzésekor H.L.A. Hart négy jelentésréteget is elkülönít egymástól úgy, hogy közben a hohfeldi értelmezést, a beavatkozásnak kitettséghez való viszonyát nem is igazán tisztázza ([347]).

jogosultság és a felhatalmazottság egymástól. Makinson azt a példát említi, hogy lehet valakinek megengedett vagy tiltott meztelenül fürdeni a tengerparton, de ettől még nem rendelkezik azzal a felhatalmazottsággal, amely alapján ezen a jogosultságon változtatni tudna (megengedhetné vagy megtilthatná azt) [464, 409]. A különbségek megragadása végett egyes szerzők (például Makinson [464]) az elsőrendű fogalmakat deontikus, a másodrendűeket kapacitív képességeknek nevezik.

Amikor Frederic B. Fitch egy 1967-es cikkében Hohfeld fogalmaira ajánl egy formalizmust [259], akkor, bár nem mondja ki, de a formalizmusból kiderül, hogy eltérő minőségűeknek tartja Hohfeld első és második jogcsoportjába tartozó fogalmakat. Az igényt, szabadságot, kötelességet és joghiányt deontikus operátorokkal, a felhatalmazottságot és a köré rendezett jogfogalmakat a 'can' operátor segítségével definiálja.[14] Lars Lindahl is ugyanezt a különbséget tartotta fontosnak:

> „ . . . power should be analyzed as ability rather than a permission 'can' rather than 'may'." [449]

Makinson figyelmeztet rá [464], hogy maga Hohfeld is erős analógiát tételezett a deontikus és kapacitív képességek egymáshoz való viszonyának tárgyalásakor, amikor így írt:

> „Ha például a közönséges tulajdonos elidegenítési felhatalmazottságára gondolunk, gondosan meg kell különböztetnünk a jogi felhatalmazottságot, a jogi felhatalmazottság 'gyakorlásához' szükséges dolgok megtételének fizikai lehetőségét, s végül e dolgok megtételére irányuló szabadságot, már amennyiben ez a szabadság tényleg fennáll. Lehet, hogy fennáll, lehet, hogy nem. Így, ha X mint egy ingatlan tulajdonosa szerződést köt Y-nal, miszerint ingatlanát nem adja el Z-nek, akkor X-nek a Z javára történő értékesítési felhatalmazottság gyakorlásához szükséges magatartások gyakorlására vonatkozó szabadsága X és Y-on kívül mindenki más között fennáll. Világos azonban, hogy X és Y viszonyában az előbbi nem rendelkezik a kérdéses magatartás tanúsítására vonatkozó szabadsággal, vagy megfordítva: Y-nal szemben az a kötelezettség terheli, hogy ne tegye meg azt, ami felhatalmazottsága gyakorlásához szükséges." [360, 90-91]

Makinson szerint ez a Hohfeld idézet is azt mutatja, hogy a felhatalmazottság és a szabadság nem mindig "jár együtt". A tézis megerősítésére Makinson hoz egy másik példát. Amennyiben egy egyházi közösségen belül a papok csak akkor adhatnak össze két embert, akik közül az egyik nem az egyházközösség tagja, ha mindketten megígérik, hogy a gyereküket a vallási előírások szellemében nevelik fel, akkor "elválhat" egymástól a szabadság és felhatalmazottság kérdése. Ha ugyanis egy pap úgy végzi el a házassági szertartást, hogy nem kapja meg a két házasulandó fél ígéretét, attól még a szabályosan elvégzett szertartás eredményeként a házasság érvényessé válik. Az pedig már egy másik kérdés, ha a pap utána büntetést kap a feletteseitől, mert nem megengedett dolgot művelt.[15]

Makinson szerint két strukturális különbség is van a deontikus és kapacitív modalitások között [464]. Az egyik különbség az, hogy *(i)* más következményekkel jár az, ha valaki olyat tesz, amire nem volt jogosultsága, illetve nem volt felhatalmazottsága.

[14] Az már egy másik kérdés, hogy Fitch ezzel a megoldással túláltalánosítja a felhatalmazottság fogalmát.

[15] Makinson példája szerintem nem a két fogalom ellentmondásáról szól, hanem inkább arról, hogy két norma, kétféle jogosultság ütközésbe került egymással, tehát ez az eset sokkal inkább a normaütközés, a normakollízió problemakörébe tartozik.

A másik különbség pedig (amit Makinson *Triponodo elvnek* nevezett el) azzal kapcsolatos, hogy *(ii)* mi történik akkor, ha a megengedés, illetve a felhatalmazottság negatív cselekvésre vagyis tartózkodásra vonatkozik.

(i) Ha valaki olyan dolgot hajt végre, ami nem megengedett, akkor ez azt jelenti, hogy a cselekvés – a tiltás ellenére – megvalósult (tehát a rá vonatkozó állítás igaz). A nem megengedett cselekvés végrehajtásának következménye valamilyen szankció lesz. Ellenben ha valaki olyan dolgot tesz, amire nem volt felhatalmazottsága, akkor a cselekvés – bizonyos értelemben – nem valósul meg, érvénytelennek minősül. Felhatalmazottság hiányában is elvégezheti valaki pontosan ugyanazt a cselekvést, mint a társa, akinek felhatalmazottsága van rá, de a felhatalmazottság hiányában cselekvése érvénytelen lesz. Az útlevélhamisító pont ugyanazt a cselekvést végzi el, mint az okirat hivatalos kiadója, mégsem lesz érvényes a cselekvése, hiszen nem volt rá felhatalmazottsága.

> „Amikor végrendelkezünk, akkor a tanúk számát illetően vagy 'eleget teszünk' a Végrendeletekről szóló 1837. évi törvény 9. szakaszának, vagy nem. Ha nem teszünk eleget, akkor okiratunk nem lesz jogokat és kötelességeket teremtő 'érvényes' végrendelet; azaz 'joghatás' vagy 'jogkövetkezmény' nélküli 'érvénytelen' végrendelet lesz. De annak ellenére, hogy érvénytelenné válik, a törvényi rendelkezés be nem tartása nem valamilyen kötelezettség 'megszegése', nem is 'vétség', és zavaró lenne erről ilyen fogalmakban gondolkoznunk." [346, 41]

Ha a cselekvő nem rendelkezik felhatalmazottsággal egy adott cselekvésre, akkor nem is végezheti el azt. Formálisan kifejezve ezt (feltételezve, hogy a cselekvések tartománya a jogi viszonyok változtatására van leszűkítve):

$$\neg\textbf{power}(Agent, \textsc{do}(Agent, F)) \rightarrow \neg\textsc{do}(Agent, F)$$

Ennek a tézisnek igaz kell legyen a **kontraponáltja** is:

$$\textsc{do}(Agent, F) \rightarrow \textbf{power}(Agent, \textsc{do}(Agent, F))$$

A fenti állítás úgy interpretálható szabad szöveges formában, hogy ha egy ágens végrehajtott egy cselekvést, akkor hatalma is volt végrehajtania azt. A kérdés itt az, hogy mi az az értelem, ami szerint érvénytelennek kell minősíteni egy cselekvést még akkor is, ha sok szempontból teljesen megegyezik azzal, ami érvényes.

(ii) A deontikus és kapacitív modalitások közötti másik fontos strukturális eltérés alapja a tartózkodással szemben "mutatott különbözőség". Ha X-nek kötelessége törekedni F elérésére, akkor azt úgy is mondhatjuk, hogy számára nem megengedett nem törekedni F elérésére. Értelmes tehát arról beszélni, hogy X-nek megengedett megtenni valamit, és arról is, hogy X-nek megengedett nem megtenni valamit. Ezzel szemben azt mondhatjuk ugyan, hogy X-nek felhatalmazottsága van végrehajtani egy cselekedetet, de az értelmetlennek, jelentés-nélkülinek tűnik, ha azt állítjuk, hogy X-nek felhatalmazottsága van nem megtenni ugyanazt. Makinson szerint ez az állítás nem is jelentés-nélküli, mint inkább triviális, és ezt nevezi Triponodo elvnek.[16]

[16] Makinson – kis pontosítással – úgy tartja ellentmondásmentesnek ezt a tézist, hogy „... amikor az F jogi állapot nem egy logikai igazságot ír le, akkor mindig triviálisan igaz, hogy bárkinek felhatalmazottsága van nem megtenni F-t." [464, 413].

*

Megítélésem szerint a felhatalmazottságot nem képességként, kapacitásként kell értelmezni, és nem is a feltételesség mentén lehet igazán különbséget tenni a beavatkozásban kitettség és a kötelezettség között. Azok, akik a képeséget, a kapacitív modalitást állították szembe a deontikus modalitással, nem tudtak igazán számot adni arról, hogy milyen ontológiai szerkezetbe, hová lehetne beállítani és milyen formulákkal lehetne megragadni a képességet, kapacitív modalitást leíró operátorokat: a felhatalmazottság leírására Fitch ugyan felvett egy 'can' operátort [259], de nem definiálta, von Wright [792], Makinson [464], Bulygin [132] pedig még csak kísérletet sem tett a formalizálásukra.

Lindahl Hansson

A megoldás annak fel- és elismerésében van, hogy a **felhatalmazotság** (és a többi "társfogalom") egyfelől **másodrendű**, másfelől összetett, ún. **polihierarchikus fogalom**. Érdemes ezt az állítást alaposabban kibontani, és levonni a konzekvenciáit. A felhatalmazottság és a "társfogalmai" a jogviszonyok megváltoztatásával kapcsolatosak, míg az "igényjogi csoport" jogfogalmai bármilyen cselekvésre vonatkozhatnak. Ez nagyon fontos mozzanat, amit pontosan kell értelmeznünk. Ha az új fogalmak (a felhatalmazottság és "társai") csak annyiban térnek el a már bevezetett fogalmaktól (az igényjogcsoportba tartozóktól), hogy az egyik fogalmi összetevőjüket, a cselekvés fogalmát tekintve szűkebb terjedelműek, akkor ez az új fogalmak modellbeli helyét is meghatározza. Az új jogviszonyfogalmak nem az eddig megismert fogalmak szintjén, azokkal mellérendelt viszonyban, hanem alárendelt módon jelennek meg, tehát mind a négy új kategória (a felhatalmazottság, a beavatkozásnak kitettség, a mentesség és a beavatkozásképtelenség) valamelyik igényjogi csoportba tartozó kategória alá rendelt fogalomként (altípusként) értelmezhető. Ezt azt is jelenti egyben, hogy amikor a felhatalmazottságot és a mentességet mint a jogosultságok új típusait akarjuk majd formálisan "definiálni", akkor ezeket a már "meglevő" jogosultságfogalmak, tehát az igényjog és a szabadság fogalmai alá sorolhatjuk csak be.[17]

Ha a másodrendű jogosultságfogalmakat az elsőrendű jogosultságfogalmak alá kell rendelnünk, akkor adódik a kérdés, mi hová tartozik. A válasz közel sem egyértelmű. Azt már említettem, hogy ha az aktív/passzív jogi szempontra figyelnünk, akkor a felhatalmazottságot aktív jognak minősítve a szabadság alárendeltjeként kell értelmeznünk, a mentességet mint passzív jogot az igényjog alá kell besorolnunk. Kanger tipizálási javaslata mögött is ez a megfontolás állhatott (↬400.o.). Maga Hohfeld azonban ezt a kérdést másként válaszolta meg:

> „ . . . a felhatalmazottság általában ugyanúgy viszonyul a mentességhez, mint a jog [igény] a szabadsághoz. A jog [igény] mással szembeni pozitív igény, a szabadság pedig valaki más jogával vagy igényével szembeni szabadság. Hasonlóképpen: a felhatalmazottság más jogi viszonyának 'alakítására' irányuló lehetőség; míg a mentesség másvalakinek egy jogviszony 'alakítására' szóló jogi felhatalmazottságától való szabadság." [360, 93]

Az előzőekhez képest pont fordított kapcsolatot tételez Hohfeld, amikor az igényjogot és a felhatalmazottságot, illetve a szabadságot és a mentességet köti össze egymással.

[17] Természetesen ugyanígy kell eljárnunk a kötelezettségfogalmak alá besorolandó beavatkozásnak kitettség és beavatkozásképtelenség esetében is.

Igaz, Hohfeld nem beszél alá-fölérendeltségi relációról, és nem is igazán indokolja meg javaslatát. Mindkét másodrendű kategóriát valamiféle megengedésnek (egyszer lehetőségnek, másszor szabadságnak) tartja.

Hart is tartja magát a Hohfeld által javasolt struktúrához (vagyis az igény a felhatalmazottsággal, a szabadságot a mentességel köti össze), de ő megpróbál valamiféle indoklást is adni az összerendelések magyarázatául:

> „A következő lehet az egyesítő mozzanat: a jogrendszer mind a négy esetben kifejezetten elismeri az individuum választását vagy negatív értelemben azáltal, hogy nem akadályozza, nem gátolja meg ebben (szabadság és mentesség), vagy megerősítő értelemben azáltal, hogy joghatást rendel hozzá (igényjog és felhatalmazottság). A negatív esetekben nincs olyan törvény, amely akadályozná az egyént abban a választásában, hogy megtegyen valamit vagy tartózkodjon valamely cselekvéstől (szabadság), illetve mindent változatlanul tartson fent saját jogi helyzetét illetően (mentesség); a megerősítő esetekben a jogrendszer joghatást rendel az egyén döntéséhez, amely alapján valaki másnak meg kell tennie valamit vagy tartózkodnia kell valamilyen cselekvéstől, illetve valaki más jogi helyzete megváltozik." [344, 35]

Hart egy lépéssel többet tesz meg, mint Hohfeld, amikor rögzíti, hogy az igényjog éa a felhatalmazottság esetében „ a jogrendszer joghatást rendel az egyén döntséhez", és ez az a mozzanat, ami összeköti e két jogviszonytípust. Az igényjog esetében egyértelmű, hogy ez a hozzárendelt joghatás a másik fél kötelezettsége valamilyen cselekvésre (valaminek a megtételére vagy valamitől való tartózkodásra),) viszont a felhatalmazottság esetében Hart nem mondja meg pontosan, hogy mi is a hozzárendelt joghatás tartalma. Legalábbis abból, hogy Hart csak annyit mond, hogy ilyen esetekben „valaki jogi helyzete megváltozik", nem lehet egyértelműen kikövetkeztetni, milyen joghatásról is van szó.

Ahhoz, hogy az egymással "vitatkozó" osztályozási, alá-fölérendelési javaslatok közti ellentmondást feloldjuk, illetve a harti joghatás valós természetét feltárhassuk, szükségünk van a fogalmak formális leírására.

*

Eddig arról beszéltem, hogy a másodrendű fogalmakat az elsőrendű fogalmak alá kell besorolnunk. Ezt mondhatjuk másként is: Hohfeld első és második jogviszonycsoportjába tartozó fogalmak között **generikus alárendeltségi reláció** (GENERIC SUBORDINATE RELATION) létezését kell feltételeznünk.[18] A generikus alárendeltségi viszony lehet **polihierarchikus** (POLYHIERARCHICAL), illetve **monohierarchikus** (MONOHIERARCHICAL) jellegű [728]. Előbbi azt jelenti, hogy a leszármaztatott fogalomnak lehet több fölérendeltje is, utóbbi esetben csak egy fölöttes fogalom lehetséges. Amellett fogok érvelni a következőkben, hogy Hohfeld felhatalmazottság fogalma polihierarchikus módon rendelhető mindkét elsőrendű jogviszonyfogalom, a szabadság és az igényjog kategóriája alá.

Abból kell kiindulnunk, hogy a felhatalmazottság (és a többi másodrendű fogalom) jogviszonyfogalom. Egy jogviszonyt a következő nyitott formulával írhatunk le:

$$\text{RIGHT}(Agent, \pm\text{DO}(Agent, T_1))$$

[18] Ezt Hohfeld nem állítja, de a meghatározásaiból mégiscsak ez következik.

Amikor – Hohfeld nyomán – a felhatalmazottságot úgy értelmezzük, hogy az jogviszonyok megváltoztatására irányuló jogviszony, akkor ezt úgy ragadhatjuk meg, hogy valamely elsőrendű jogviszony (tehát kétféle szabadság vagy kétféle igényjog) megváltoztatásának jogát jelenti.

$$\text{RIGHT}(\text{Agent},\pm\text{DO}(\text{Agent},\text{RIGHT}(\mathit{Agent},\pm\text{DO}(\mathit{Agent},T_1))))$$

A fenti képletbe ekkor a RIGHT helyett POWER-t írhatunk, az *Agent*-változó helyébe *Ego*-t, a T_1 helyébe pedig valamely korábbi jogviszony megváltoztatását mint cselekvést helyettesíthetjük:

$$\mathit{Change}(\text{RIGHT}(\mathit{Agent},\text{DO}(\mathit{Agent},T_0))).$$

A felhatalmazottság definíciójában már nem kell, hogy szerepeljen a ± jel a DO cselekvésoperátor előtt. Azért nincs rá szükség, mert a felhatalmazottság esetén nincs értelme negatív cselekvésről vagyis a felhatalmazottságtól való tartózkodásról beszélni (erre később még visszatérek). Az új paraméterekkel a fenti nyitott és általánosított formula erre a speciálisabb alakra hozható:

$$\text{POWER}(\mathit{Ego},\text{DO}(\mathit{Ego},\mathit{Change}(\text{RIGHT}(\mathit{Agent},\text{DO}(\mathit{Agent},T_0)))))$$

Ez a képlet szemléletesen mutatja a fogalom másodrendű jellegét, hiszen egy jogviszonyba beágyazva egy másik jogviszony szerepel paraméterként. Ezen a ponton viszont fel kell tennünk a kérdést, vajon milyen módon lehetne kifejezni a **power** fogalmát elsőrendű jogviszonyfogalmak segítségével, vajon a szabadság (FREEDOM) vagy az igényjog (CLAIM) alá tartozik. Az első – kézenfekvőnek tűnő – válasz az kell legyen, hogy a **power** a FREEDOM alárendeltje, hiszen mindkettő **aktív jog**, és a felhatalmazottság gyakorlásához valamilyen formában szükséges feltételnek tűnik, hogy a jogbirtokos rendelkezzen a jogviszony változtatásához szükséges szabadsággal. Vélhetőleg ez a szerkezeti hasonlóság vezethette Kangert is arra, hogy a szabadság mellé a felhatalmazottságot emelje be a deontikus logikai sémájába. Bár von Wright elsősorban kompetencianormának nevezte, de ő is permisszív normának minősítette a hatalom/felhatalmazottság fogalmát:

> „A magasabbrendű *kompetencianormákat* permisszív normáknak nevezem. Azt mondhatjuk, hogy egy kompetencianorma vagyis egy magasabbrendű permisszív norma kibocsátásának aktusában a magasabbrendű felettes hatóság *delegálja a felhatalmazottságát* egy alacsonyabbrendű *alárendelt hatóság* számára. A 'felhatalmazottság' itt nem más, mint 'normával teremtett kompetencia normák kibocsátójaként cselekedni'. Úgy is beszélhetünk erről, mint *normatív* kompetencia vagy felhatalmazottság." [792, 192]

Észre kell azonban azt is vennünk, hogy van egy másik szerkezeti hasonlóság is a hohfeldi fogalmak között. A szabadság korrelatív párja az igényhiány, aminek a hohfeldi másodrendű fogalmak között a beavatkozásképtelenség felelne meg – miközben Hohfeld a felhatalmazottság korrelatívumára, a beavatkozásnak kitettségre állandóan úgy hivatkozik, mint ami szerkezetileg, deontikus modalitását tekintve a kötelezettségnek felelne meg. A feltételes kötelezettség hohfeldi elképzelése is ezt sugallja. Ha a felhatalmazottság mint speciális szabadság aktív jog, akkor annak korrelatív párjának passzív jognak kéne lenni,

viszont a beavatkozásnak kitettség szintén aktív jognak tűnik. Ellentmondásos helyzetnek tűnik, de nem az.

A látszólagos ellentmondás feloldódik, ha észrevesszük, hogy a felhatalmazottság fogalmában valójában két elsőrendű jogviszonyt érhetünk tetten. A felhatalmazottság "alapja" az a szabadság, amely "lehetővé teszi", hogy a jogbirtokos megváltoztathasson egy jogviszonyt. De ezzel még nincs vége. A jogbirtokos ezzel párhuzamosan egy igényjoggal is fellép a másik féllel szemben arra vonatkozóan, hogy a partnere ismerje el a jogviszony-változtatás tényét (érvényességét). A beavatkozásnak kitettség az ebben az értelemben vett elismerési igénnyel szembeállítható elismerésére irányuló kötelezettséget jelenti. A felhatalmazottság fogalma tehát két összetevőre bontható fel, amit így fejezhetünk ki formális módon:

$$\begin{gathered}\textbf{power}(Ego,\text{DO}(Ego,\text{CHANGE}(\text{RIGHT}(Agent,\text{DO}(Agent,T_0)))))\\ =\\ \text{FREEDOM}(Ego,\text{DO}(Ego,\text{CHANGE}(\text{RIGHT}(Agent,\text{DO}(Agent,T_0)))))\\ \wedge\\ \text{CLAIM}(Ego,\text{DO}(Alter,\text{RECOGNITION}(\text{DO}(Ego,\text{CHANGE}(\text{RIGHT}(Agent,\text{DO}(Agent,T_0)))))))\end{gathered}$$

Ez az összetettség azt jelenti, hogy a felhatalmazottság polihierarchikus fogalom, vagyis nem egy, hanem több **generikus fölérendelt** (GENERIC SUPERORDINATE) fogalom alá lehet besorolni.[19] Ha így értelmezzük a felhatalmazottság fogalmát, akkor már magyarázhatóak a korábban jelzett "anomáliák", érthetővé válik a beavatkozásnak kitettség és a kötelezettség oly közeli viszonya is. Ha a felhatalmazottságban van egy igényjogi összetevő, akkor annak korrelatívuma egy kötelezettség lesz, ami a beavatkozásnak kitettség fogalmába "kerül be", és ez a mozzanat alapozza meg a kötelezettség és a beavatkozásnak kitettség fogalmai között érzett hasonlóságot.

---

kicsit előbb felvenni az ontológiába: RECOGNITION CHANGE

a beavatkozásnak kitettség nem feltétlen szül kötelezettséget: lomtalanítás példában a tulajdonos ingóságáról való lemondó gesztusa szabadságot teremtett mindenki számára az eltulajdonításra, itt nincs kötelezettség

*

A felhatalmazottság meghatározása alapján már könnyen definiálni lehet annak inverz fogalmát, a beavatkozásnak kitettséget. A két fogalom közt fennálló inverzitás tényét kihasználva, a beavatkozásnak kitettséget is polihierarchikus fogalomként kell értelmeznünk, csak épp a felhatalmazottsághoz rendelt szabadság és elismerési igényjogi összetevők inverzeit kell a beavatkozásnak kitettség két összetevőjeként definiálni.

$$\begin{gathered}\text{LIABILITY}(Alter,\text{DO}(Ego,\text{CHANGE}(\text{RIGHT}(Agent,\text{DO}(Agent,T_0)))))\\ =\\ \text{NO-CLAIM}(Alter,\neg\text{DO}(Ego,\text{CHANGE}(\text{RIGHT}(Agent,\text{DO}(Agent,T_0)))))\\ \wedge\\ \text{DUTY}(Alter,\text{DO}(Alter,\text{RECOGNITION}(\text{DO}(Ego,\text{CHANGE}(\text{RIGHT}(Agent,\text{DO}(Agent,T_0)))))))\end{gathered}$$

[19] A generikus fölérendeltje relációt szokták még GENERIC SUPERIOR relációnak is nevezni.

*

Azt lehet mondani, hogy a szakirodalomban a felhatalmazottsághoz képest alig foglalkoztak a mentesség fogalmával.

a mentesség formális definíciója:

$$\begin{gathered}\text{IMMUNITY}(Ego, \neg\text{DO}(Alter, \text{CHANGE}(\text{RIGHT}(Agent, \text{DO}(Agent, T_0)))))\\ =\\ \text{CLAIM}(Ego, \neg\text{DO}(Alter, \text{CHANGE}(\text{RIGHT}(Agent, \text{DO}(Agent, T_0)))))\\ \wedge\\ \text{CLAIM}(Ego, \text{DO}(Alter, \text{RECOGNITION}(\text{DO}(Ego, \text{CHANGE}(\text{RIGHT}(Agent, \text{DO}(Agent, T_0)))))))\end{gathered}$$

10.2.5 A jogviszonyfogalmak és szabályfogalmak viszonya

Az eddigiekben jogviszonytípusokról, azok értelmezési lehetőségeiről volt szó, és a deontikus operátorokat használtam a fogalmak formális leírásához. A másodrendű jogviszonyfogalmakkal egy olyan pontra értünk, amikor megpróbálhatjuk hasznosítani azokat az elméleteket, fogalmakat, amelyeket az általában vett szabályok tipizálása során alkottak meg, de ehhez az is szükséges, hogy innentől kezdve az általánosabb **szabály** fogalomról beszéljünk egy darabig az eddigi jogviszonyfogalmak helyett. A jelen kontextusban a szabály fogalmát a norma fogalmával szinonim módon fogom használni, noha a végső modellben a szabály főfogalma alá fogom majd berendelni a norma alfogalmát. Azt azonban fontos rögzíteni, hogy a szabály fogalmát az egész könyvben az emberi cselekvésekre vonatkozó szabályként használom, ebben az értelemben ez a fogalomhasználat szűkebb jelentésű a teljesen általános szabályfogalomhoz képest (a **szabály** terminus jelentése tehát ebben a könyvben megegyezik a **cselekvési** vagy **viselkedési szabály** fogalmának jelentésével).

A cselekvési (viselkedési) szabály fogalmát már sokat használtam, több helyen is elemeztem a könyvemben (♀→**??**.o.)., illetve **??**, így itt csak egy idézettel szemléltetem, miért is fontos a szabály mind a társadalmi cselekvések végrehajtásában, mind azok magyarázatában.

> „Két jel utal arra, hogy a viselkedés szabályvezérelt, nem pedig csupán rendszeresen ismétlődő: a mintától való eltérést valamiképpen hibának vagy tökéletlenségnek tartjuk, valamint a szabályt automatikusan alkalamazzuk új esetekre is. Ha a cselekvő szembekerül egy új esettel, amellyel korábban nem találkozott, tudni fogja, mit tegyen." [653, 55]

Bár ez a szabályfogalom megengedi, hogy mind a normához, mind a szokáshoz igazodó cselekvésre ráhúzzuk, a továbbiakban csak a normativitás körébe tartozó szabályfogalmakkal foglalkozom. Az egyik fontos megválaszolandó kérdés az az, hogy milyen típusai vannak a viselkedési szabály fogalmának. Eugenio Bulygin miután tisztázza, hogy a szabály (rule) és a norma fogalma között nem tesz különbséget, valamit egymás szinonímájaként használja a kompetencia, kapacitás, potencia, hatalom, uralom, képesség, befolyás teminusokat, két szabályfogalmat különböztet meg egymástól: a **viselkedési normák** (norms of conducts) és a **kompetencianormák** (competence norms) kategóriáit [132].

H.L.A. Hart a kompetencianorma (norm of competence) terminusa helyett inkább a **felhatalmazottság-kiszabó szabály** (power-conferring rule) fogalmát javasolja használni [**?**]

Hart szerint a két kategória egymást kizárja, Hart 28, 77,

Von Wright háromféle normát különít el: *(i)* a preskripciókat vagy regulációkat (prescriptions, regulations), *(ii)* a determinatív szabályokat (determinative rules) és *(iii)* a technikai normákat vagy direktívákat (technical norms, directives) [792]. A von wright-i hármas felosztás első két eleme, a preskripciók és a determinatív szabályok megfeleltethetők a harti kettős szabályfelfogásnak, a harmadik von wright-i szabályfogalom, a technikai normák fogalma viszont eltér minden eddigi tipizálási kísérlettől. Sokan mások, én magam sem tartom igazán a norma fogalma alá sorolhatónak ezt a fogalmat, mert pont a normativitáshoz szükséges legfontosabb elem hiányzik belőle.

Bulygin azt javasolja, hogy von Wright determinatív szabály kifejezése helyett inkább a fogalmi szabályok kifejezést használjuk, és tekintsünk el a harmadik von wright-i normatípustól. ezek után pedig kijelenti, hogy szerinte a kompetencianormákat nem lehet visszavezetni sem a kötelesség sem a megengedés fogalmára. A viselkedési norma megsértése esetén szankció várható (következik), ezzel szemben a kompetencianorma megsértésekor érvénytelen lesz a felhatalmazottságot közvetítő, átruházó, deklaráló cselekvés [132, 208]. Ha nem az előírt feltételekhez igazodva születik meg valami, akkor **érvénytelenség**, semmisség (NULLITY) lesz a következménye ugyanúgy, mint bármely játékban (sakk, sport) ez történik, ha megsértik a játékszabályokat. Így ír Bulygin:

> „. . . Hart lényeges pontra tapint rá. Szankció és érvénytelenség [nullity] két radikálisan különböző fogalom és bármely kísérlet, amely beolvasztaná az érvénytelenséget a szankció fogalmába, kudarcra van ítélve. Ráadásul a szankcióra és érvénytelenségre jellemző eltérő működésmódok azt is mutatják, hogy lényegi különbség van a két szabálytípus között is. Első pillantásra a konstitutív szabály az, ami definál egy fogalmat vagy egy játékot mint tevékenységet, és ezáltal értelmessé teszi az érvénytelenségről való beszédet. Ha azonban létezik egy definíció, amely szerint ahhoz, hogy egy norma, egy cselekvés vagy egy dokumentum érvényessé váljék, bizonyos feltételeket ki kell elégíteni, akkor bármely feltétel hiánya az adott norma, cselekvés, dokumentum érvénytelenségét idézi elő. Az érvénytelenség nélkül nem létezik maga a definíció sem, így a konstitutív szabály és az érvénytelenség nem választhatóak el egymástól. E kétfajta intézmény, a büntetés és az érvénytelenség együttes létezése egyben jelzi a jogrendszer (legalább) kétféle szabályának létezését: vannak viselkedési normák és konceptuális szabályok, vagy, mindezt searle-i terminusokban kifejezve, vannak regulatív és konstitutív szabályok.” [132, 214]

A **regulatív** és **konstitutív szabályok** fogalmainak bevezetését John Searle nevéhez kötik, ami a terminológiai szinten természetesen helyénvaló, de Searle-t jócskán megelőzte John Rawls, aki már 1955-ben szembeállította a **summázati** és **gyakorlati szabályokat** egymással úgy, hogy Rawls fogalomkettőse fogalmilag teljesen megfeleltethető Searle párosának [595]. Searle regulatív szabályai (és Rawls summázati szabályai) olyan cselekvésekre vonatkoznak, amelyek akkor is “léteznének”, ha nem irányulna rájuk szabályozás. Ezzel szemben a konstitutív szabályok (Rawlsnál a gyakorlati szabályok) maguk “teremtik”, maguk teszik lehetővé a cselekvéseket, az ilyen szabályok híján az ilyen cselekvések nem léteznének, értelmetlenekké, értelmezhetetlenekké válnának. A regulatív szabályok alá sorolható cselekvésekre deontikus operátorok irányulhatnak, és ha a szabályok által

előírt cselekvések nem valósulnak meg, tehát "szabálysértés" történik, akkor ezt szankció követi. Ez utóbbi összefüggést lehet kifejezni az alábbi két formulával:

$$\text{OBLIGATORY}(A) \wedge \neg A \rightarrow \text{S}$$
$$\text{FORBIDDEN}(A) \wedge A \rightarrow \text{S}$$

A konstitutív szabályok által teremtett cselekvéseket máshogy lehet jellemezni. A konstitutivitás mozzanatát John Searle a **counts as** (valaminek számít) operátor segítségével véli megragadhatónak, és ezért azt javasolja, hogy az alábbi módon értelmezzük a konstitutív szabályt [653, 47]:

$$X \text{ counts as } Y \text{ in } C.$$

Természetes nyelvi mondatban kifejezve:

'valamely X dolog valamely C kontextuson belül Y dolognak számít'.

Y az az új dolog, entitás, amelyet a konstitutív szabály alkalmazása révén konstituálunk, X pedig az a dolog, entitás, amely a szabály alkalmazása nélkül (előtt) is létezik a világban. Utóbbira – Searle terminológiáját használva – a **nyers tények** (**brute facts**), előbbire az **intézményes tények** (**institutional facts**) vonatkoznak [653, 63]. Sokan elfogadták, helyeslően átvették Searle formuláját, és persze voltak, akik kritizálták őt ezért.[20] Joseph Ransdell egy 1971-es tanulmányában azzal "támadta" Searle-t, hogy az intézményes tényeket "hordozó" entitást kifejező Y terminusban összemos két fontos mozzanatot, amit pedig inkább szét kellene választani és érdemes lenne külön reprezentálni őket [591]. Ransdell kritikája annyiban jogos, hogy amikor Searle a fenti formuláról beszél, akkor ahhoz képest szövegszerűen maga is további összetevőkre utal, ám ezeket nem emeli be a formalizmusába. Searle szerint azzal, hogy X-re "mutatva" konstituálunk egy intézményes tényt, miközben "létrehozzuk" Y-t, ez utóbbi terminussal egyben specifikálunk egy új státust is, amihez hozzárendelünk egy vagy több funkciót [651, 46-47]. Ezek a funkciók adják a konstitúció értelmét. Ezek a státusfunkciók (status function) olyan szabályok, amelyek a konstituált új entitásra vonatkozóan rögzítik, hogy mit lehet vagy mit nem lehet tenni velük. Hiába állapítja meg ezeket Searle, a konstitutív szabály formalizmusába már nem emeli át.

Ransdell azt javasolja, hogy a konstituált új entitás két fontos összetevőjét nevezzük meg, válasszuk el egymástól [591, 388-389]. Szerinte ugyanis a konstituált entitásnak (terminusnak) mindig van mellékjelentése, *konnotációja* (*connotation*) egyfelől, illetve *tartalma* (*import*) másfelől. A baseball szabályain szemléltetve Ransdell a következő példákkal világítja meg a két új fogalom jelentését. Vannak Y-terminusok (Ransdell szavaival: játékterminusok/game-terms), mint Ball (labda), Bat (baseballütő), Pitcher (dobó), Batter (ütő), Catcher (elkapó), Inning (menet) stb. Ezeknek legtöbb esetben van konnotációja és tartalma (import-ja). A Bat konnotációjába tartozik a mérete, formája, felépítése, míg a tartalmával azt lehet rögzíteni, hogy mire lehet, mire kell, mire nem lehet használni: például lehet vele ütni a labdát, de nem lehet vele elgáncsolni a futó játékost. Ransdell fontos felismerése, hogy ez utóbbiak elsőrendű játékszabályok (cselekvési szabályok), amelyek "be vannak ágyazva" a másodrendű konstitutív szabályokba.

[20] F.A. Hindriks egy tanulmánya elején rövid, de tartalmas áttekintést ad a searle-i felfogás kritikájáról [355].

Jobban meg tudom világítani Ransdell gondolatát, ha a szemléltető példákat a baseball helyett inkább a labdarúgás területéről veszem.

*

A labdarúgás szabályait egy több, mint százharminc oldalas könyvben rögzítik [251]. Ez a szabálygyűjtemény tele van konstitutív szabályokkal. A játék játszhatósága végett mindent pontosan definiálni kell, így meghatározás olvasható arról, hogy mit jelent a 'játékos', 'cserejátékos', 'lecserélt játékos', a 'csapat', a 'labda', a 'pálya', a pályára felfestett 'vonalak' és 'pontok', a 'kapu', a 'leshelyzet', a 'szabadrúgás' stb. fogalma. A játék értelme, célja és ezért központi fogalma a 'gól'. Nézzük meg a meghatározását:

> „Gól akkor van, amikor a labda teljes terjedelmével áthalad a kapuvonalon a kapufák között és a keresztléc alatt, feltéve, hogy a gólt elérő csapat előzően nem követett el szabálysértést." [251, 30]

A definíció támaszkodik olyan, már előzőleg konstituált fogalmakra, mint a 'labda', 'csapat', 'kapuvonal', 'kapufa', 'keresztléc', valamit 'szabálysértés'. Searle formulája úgy konkretizálható, hogy a C kontextus itt a 'labdarúgást játszani', az Y a 'gól' és az X a fenti definíció szinte teljes egészében. A definíciót természetesen csak úgy lehet pontosan és egyértelműen értelmezni, ha ismerjük a meghatározásba ágyazott további konstutitív fogalmak jelentését is. De azok is mind a searle-i logika mentén határozhatóak meg. Eddig tehát tűnhetne úgy is, hogy Searle leírása megfelelő és elegendő, de ez nincs így. Vannak ugyanis olyan szabályok a labdarúgás szabályrendszerében, amelyek kilógnak a searlei formalizmusból.

Nézzük meg a büntetőrúgás (tizenegyes) végrehajtásának szabályait! A szabálykönyv korábbi pontjaiban rögzítették, hogy mi az a tíz 'szabálytalanság', amely esetén 'közvetlen szabadrúgást' kell ítélni a szabálysértő játékos csapata ellen [251, 32]. Ezek alapján lehet bevezetni a 'büntetőrúgás' terminusát. A büntetőrúgás konnotációjaként értelmezhetjük azokat a sorokat, amelyekben segítenek abban, hogyan kell a tizenegyesrúgás fogalmát értelmezni.

> „Büntetőrúgás jár az ellen a csapat ellen, amely elköveti a közvetlen szabadrúgást maga után vonó tíz szabálytalanság valamelyikét a saját büntetőterületén, amikor a labda játékban van." [251, 41]

Miután megadtuk, hogy mit jelent a büntetőrúgás fogalma, milyen esetben kell büntetőrúgást ítélni egy focimeccsen (tehát hogyan, mikor konstituálódik a fogalom), utána kell megadnunk azt, hogy milyen körülmények között és hogyan kell végrehajtani a büntetőrúgást. Ezektől az előírásoktól lesz, lehet érvényes vagy érvénytelen a tizenegyes végrehajtása. Először megadják a büntetőrúgás során előírt elhelyezkedést:

> „A labdát:
> - a büntetőpontra kell helyezni.

> A büntetőrúgást végrehajtó játékost:
> - egyértelműen azonosítani kell.

> A védőcsapat kapusának:
> - a kapuvonalon a kapufák közt, a rúgó játékossal szemben kell maradnia a rúgás elvégzéséig

A többi játékosnak:
- a játéktéren
- a büntetőterületen kívül
- a büntetőpont mögött
- a labdától legalább 9,15 méter távolságra kell elhelyezkedni." [251, 40]

Az elhelyezkedésre vonatkozó szabályok mind a játékosok cselekvésére irányuló, elsőrendű szabályok, melyek a szokásos módon megragadhatóak deontikus operátorokkal. A következő feltételsorban azt rögzítik, hogy milyen módon lehet magát a büntetőrúgást szabályosan végrehajtani. Ezek a szabályok nyilván a büntetőrúgást végző játékosra vonatkoznak csak.

- „Miután minden játékos a szabály által előírt helyen tartózkodik, a játékvezető jelt ad a büntetőrúgás végrehajtására.
- A büntetőrúgást végző játékosnak a labdát előre kell rúgnia.
- Nem játszhatja meg ismét a labdát, amíg az egy másik játékost nem érintett.
- A labda akkor van játékban, amikor elrúgták és előre elmozdult." [251, 41]

Ebben a definíciós részben is kötelezések és tiltások vannak, vagyis itt is alkalmazhatjuk az elsőrendű cselekvésekre irányuló deontikus operátorokat. A szabályok harmadik csoportjában a büntetőrúgással kapcsolatos szabálysértéseket és a rákövetkező büntetéseket kapcsolják össze – egymás után felsorolva azokat az eseteket, amikor a büntetőrúgást végző játékos, a kapus, a rúgást végző játékos csapattársa, a kapus csapattársa, illetve mindkét csapat játékosa vét a szabályok ellen. Ha a támadó csapat játékosa sérti meg a játékszabályokat, akkor a játékvezetőnek az alábbiak szerint kell döntenie:

„Ha a büntetőrúgást végző játékos . . . [vagy] a rúgást végző játékos csapattársa sérti meg a játékszabályokat:
- Ha a labda kapuba jut, a rúgást meg kell ismételni.
- Ha a labda nem jut a kapuba, a játékvezető állítsa meg a játékot, és a mérkőzés a védő csapat javára megítélt közvetett szabadrúgással indul újra arról a helyről, ahol a szabálysértés történt." [251, 42]

Ha a védekező csapat valamelyik játékosa szabálytalankodik, akkor a bírónak a következőképpen kell eljárnia:

„Ha a kapus . . . [vagy] a kapus csapattársa sérti meg a játékszabályokat:
- Ha a labda kapuba jut, gólt kell ítélni.
- Ha nem jut a kapuba, a rúgást meg kell ismételni." [251, 42]

Ezek a szabályok rögzítik azokat a szempontokat, amelyek alapján a bíró dönthet az érvényesség-érvénytelenség kérdéséről. Fontos különbség, hogy ezek az előírások a bíró tevékenységére vonatkoznak, míg a korábbiak a játékosok cselekvéseire irányultak.

Az, hogy mikor lesz gól a büntetőrúgásból, két tényezőtől függ. Egyfelől attól, hogy szabályos volt-e a büntetőrúgás végrehajtása, másfelől attól, hogy a büntetőrúgás eredményeként a labda bekerült-e a kapuba. Feltételezve, hogy a büntetőrúgásra vonatkozó szabályokat betartotta a rúgó játékos, az érvényesség kérdését "leszűkíthetjük" arra a szabályhalmazra, amely a büntetőterületre való belépés tilalmával kapcsolatos (illetve a gól definíciójára). Az alábbi táblázat mutatja, mi a bíró teendője a különböző esetekben.

belép-e valaki:	a labda a kapuba kerül	a labda nem kerül a kapuba
támadó játékos	a büntetőrúgás megismétlése	közvetett szabadrúgás
védő játékos	gól	a büntetőrúgás megismétlése
mindkettő játékos	a büntetőrúgás megismétlése	a büntetőrúgás megismétlése
senki	gól	nem gól, megy tovább a játék

A táblázat első oszlopa mutatja, hogy követ-e el valaki szabálysértést a belépésre vonatkozóan, és ha igen, melyik csapatba tartozó játékosról van szó. A következő két oszlopban a különböző esetekben meghozandó bírói döntést láthatjuk aszerint, hogy a büntetést végrehajtó játékos berúgta-e a labdát vagy sem a kapuba. A gól egyik feltétele az, hogy a kapuba kerüljön a labda. Ezeket a helyzeteket mutatja a táblázat második oszlopa. A gól érvényességéhez azonban arra is szükség van, hogy ne előzze meg azt a támadócsapat szabálytalansága. A szabálytalanságot itt a büntetőterületre való korai belépés jelenti. Ezért van az, hogy "csak" két esetben van gól (a lehetséges négyből): akkor, amikor a támadó csapat nem szabálytalankodott (a második és negyedik sorban). Ha a gól elsődleges feltétele nem teljesül (ha a labda nem kerül be a kapuba), akkor a a szabálytalanságok figyelése csak arra jó, hogy el lehessen dönteni, meg kell-e ismételni a büntetőrúgást vagy sem (és akkor csak kell ismételni, ha a védekezőcsapat szabálytalankodott). A játékvezető kompetenciájába tartozik, hogy megismételteti-e a büntetőrúgást vagy sem. A tapasztalatok szerint az ismételtetések elég ritkán fordulnak elő.[21]

Előfordulhat, hogy az érvényesség fogalma mellett más fogalmat is használatba kell venni. Sok esetben meg kell különböztetnünk egymástól az **érvényességi feltételt** és az **eredményességi feltételt**. A különbséget a büntetőrúgás példáján szemléltetve azt mondhatjuk, hogy ha a támadó berúgja a labdát, de szabálytalanság történt, akkor érvénytelen lesz a büntetőrúgás, míg ha a támadó nem rúgja be a labdát és nem történt szabálytalanság közben, akkor érvényes lesz a büntető, de eredménytelen. Végül, ha szabályos körülmények között rúgja be, akkor érvényes és eredményes lesz a tizenegyes. A

[21] A YouTube-on látni olyan videót, amelyen a bíró ötször ismételteti meg a büntetőrúgást, és csak hatodik kísérletet fogadja el érvényesnek, lásd: http://youtube.com/watch?v=VJwAE5Mtrxc.

1982. december 5-én az Ajax – Helmond Sport labdarúgó mérkőzésen az Ajax – 5:0-ás vezetésnél – büntetőhöz jutott. Johan Cruyff egy társa, Jesper Olsen segítségével különös módon hajtotta végre a tizenegyest. Cruyff nem egyetlen rúgással próbált meg gólt rúgni, ahogy előtte mindenki a labdarúgás történetében, hanem lassan oldalra gurította a labdát, amit az elrúgás után már szabályosan a tizenhatoson belülre befutó Olsen kapott meg, aki az öt és feles vonalhoz érve visszaadta a középen tartózkodó Cruyffnak, ő pedig közvetlen közelről bevágta a labdát a kapuba. Az eset mindenben megfelelt a szabályoknak, a bíró meg is adta a gólt. Az esetről több videó is látható a YouTube-on (például: http://youtube.com/watch?v=GDoqv74egBM). Később kiderült, hogy az elsőség nem Cruyffot illeti, hanem Rik Coppenst, aki már 1957-ben megcsinálta a két-lépéses büntetőt egy társával (lásd: http://youtube.com/watch?v=hFk0xqO4VkU).

Mivel „ . . . a játékvezető dönti el, mikor tekintendő a büntetőrúgás végrehajtottnak", adódik a kérdés, hogy meddig tart egy tizenegyes végrehajtása. Ez az esetek döntő részében egyértelműen megválaszolható, de rendkívül ritkán előfordul az a jelenség, amikor a sikeresnek tűnő védéssel nem "zárul le" a büntetőrúgás eseménye, mert a védés után valahogy mégis bekerül a labda a hálóba. Előfordult olyan eset, hogy miután a kapus kivédte a büntetőt, felpattant és a közönség felé fordulva ünnepelni kezdte magát, miközben a kapus testén gellert kapó labda szép lassan bepörgött a kapuba (http://youtube.com/watch?v=r4bnoMka0SE). Egy másik alkalommal, egy büntetőpárbaj során a felső kapufáról magasba pattanó labda visszahullott az öt és feles vonalára, mire a kapus örömében kirohant a pálya szélén levő társaihoz, ám eközben a visszafele pörgő labda a kapu irányába pattogva szép lassan áthaladt a gólvonalon (http://youtube.com/watch?v=s5hE6PK67bc).

parlamenti választások alkalmával is van érvényességi, illetve eredményességi feltétel. Az érvényességi feltételként azt lehet kikötni, hogy a szavazók x (például 25) százaléka érvényes szavazatot adjon le. Az eredményességhez az kell, hogy valamely jelölt megkapja az érvényes szavazatok valamilyen többségét (például az első fordulóban a szavazatok abszolút, az esetleges második fordulóban a szavazatok relatív többségét).

*

A labdarúgás példáján láthattuk, hogy a konstitutív szabályokat leírásához nem elégséges a searle-i 'counts as' operátor alkalmazása. A szabályok konstitutív jellegét meg lehet ragadni vele, de logikai formalizmus alapjául nem képes szolgálni. A searlei funkció (vagy a ransdelli import) kiegészítő fogalmának szükségessége pedig arra mutat rá, szükség van arra, hogy elsőrendű regulatív szabályokat kapcsoljunk a konstituált fogalmakhoz (társadalmi intézményekhez). Két kérdést is fel kell itt tenni, amelyekre a válaszokat egymással összekapcsolva lehet megadni: *(i)* milyen logikai formulákkal ragadhatjuk meg a konstitutív szabályokat; *(ii)* hogyan kapcsolhatjuk össze a konstitutív szabályokat a felhatalmazottság köré szerveződő jogviszony-kategóriákkal vagyis a másodrendű deontikus fogalmakkal.

Először is észre kell vennünk, hogy a játékvezető teendőit másodrendű szabályok írják elő, és így a játékvezető cselekvéseit (döntéseit) is másodrendű cselekvésként írhatjuk le. „Ha a kapus megsérti a játékszabályokat büntetőrúgásnál, de közben kivédi a tizenegyest, akkor a rúgást meg kell ismételni." – szól a bíróra vonakozó szabály. Egy előírásba be van ágyazva egy másik előírás (ami a kapusra vonatkozik) – ez egyértelmű jele a másodrendű szabálynak. De milyen másodrendű fogalomról van szó? A választ akkor kapjuk meg, ha a hohfeldi fogalmak segítségével próbáljuk meg jellemezni a büntetőrúgás helyzetét. Mit is jelent a büntetőrúgás? Azt a jogot a támadócsapat számára, hogy egy játékosa egyedül rúghassa kapura a labdát a kapussal szemben (és minden más játékos kötelezettségét arra vonatkozóan, hogy ezt az igényjogban kifejezett cselekvést ne akadályozza). Ha a konstitutív szabályok valamelyikét megsértik a büntetőrúgás során, akkor a bíró hozhat döntés az érvényességre vonatkozóan. Amikor a játékvezető úgy ítél, hogy megismételteti a be nem rúgott büntetőt a védőcsapat szabálytalansága miatt, akkor a támadócsapat büntetőrúgásra vonatkozó jogát teremti újra. Tehát a létező jogviszonyok valamelyikét változtatja meg. Ez azt jelenti, hogy a bíró – a hohfeldi értelemben vett – felhatalmazottsággal rendelkezik.

Mivel ez a tézis nagyon fontos mozzanat a gondolatmenetemben, ezért szeretnék még egy példát bemutatni, amikor a labdarúgáshoz képest jóval egyszerűbb versenyről szól, ezért jobban áttekinthetőek a viszonyok a játékban. 200 éves hagyomány az angliai Gloucester városában a Cooper-dombi sajtkergetés.[22] Hosszú ideig hivatalos sportversenynek számított, ám 2009-ben visszavonták a hivatalos minősítést a rendezvénytől a sok balaset miatt. Azóta "nem hivatalosan" gyűlnek össze az évente megrendezett versenyen a versenyzők, nézők. A versenyen a gloucesteri domboldal tetejéről legurítanak egy 3 és fél kilós kerek sajtot, és a résztvevők közül az lesz a nyertes, aki utoléri és el tudja kapja a sajtot. Ha ilyen nincs, akkor az nyer, aki legelőbb leér a domb aljára. A verseny célja a sajt megszerzése, de ezt gyakorlatilag lehetetlen teljesíteni, tehát a győzelem igazi feltétele a leérés elsősége.

[22] http://www.cheese-rolling.co.uk/index1.htm

A sajtkergetési versenyt mint társadalmi intézményt az alábbi – searle-i értelemben vett ‘counts as’ – konstitutív szabályokat révén definiálhatjuk:

Y	(is counted as) *X*
versenynap	meghirdetett dátum
startvonal	Cooper-domb teteje
célvonal	Cooper-domb alja
pálya	startvonal és célvonal közti terület
sajt	hivatalosan kijelölt sajt
kezdés	elindítják a sajtot a startvonaltól a dombon lefelé
verseny	kezdés után (elkapni a sajtot vagy végigmenni a pályán)
győzelem	versenyen elsőre (elkapni a sajtot vagy végigmenni a pályán)

Az első oszlopban láthatóak a kontituált, intézményes tényekkel leírt entitások(Y) , a második oszlop tartalmazza a fizikai világban létező, nyers tényekkel leírt entitásokat (X). Utóbbi minősítés persze csak úgy igaz, ha figyelembe vesszük, hogy az (X) entitások tartalmazhatnak “beágyazott” intézményes tényeket, amelyek azonban – esetleg több lépésben – mindig visszavezethetők fizikai tényekre, entitásokra. A pálya meghatározása például tartalmaz két másik intézményes tényt (a startvonal és a célvonal kategóriáit), azonban ezek egy-egy további lépéssel már visszavezethetők nyers tényekre.

Ezek az intézményes tények, ezek a konstituált entitások még nem elegendőek ahhoz, hogy le lehessen bonyolítani a versenyt. Szükség van még olyan regulatív szabályokra, amelyek a meghatározzák, hogy mit, hogyan kell, lehet használni a konstituált fogalmak által definiált, konstituált környezetben, vagyis hogyan kell a versenyt lebonyolítani. És ha ezen szabályok közül valaki megsért egyet, akkor – nagy valószínűséggel – nem lesz érvényes “valami”. Ilyesfajta szabályokról lehet szó:

előírások	szabálytalan/érvénytelen
rajtvonal mögül kell indulni	ha valaki lejjebb indul
kezdés előtt tilos indulni	ha valaki előbb indul
a hivatalos sajtot kell elkapni	ha valaki a saját sajtját kapja el
versenynapon kell versenyezni	ha valaki más napon versenyez

Hogy az ilyen szabályok nélkül nem lehetne a versenyt sem definálni, sem lebonyolítani, azt felfedhetjük azzal, ha belegondolunk, hogy milyen reakciókra számíthatna az, aki nem igazodna a regulatív szabályokhoz. Ha valaki az általa vitt sajtot maga elé ejti, majd azt felkapva felkiált, hogy ‘nyertem’, senki nem fogja győztesnek tartani, és ha megkérdezi, miért nem, akkor azt a választ kapja, hogy azért nem, mert nem a hivatalos sajtot kapta el vagy másként: a győzelemhez a hivatalos sajtot kell elkapni. Ha valaki előbb elindul a többiekhez képest, és ezért el tudja kapni a sajtot vagy elsőként ér le, akkor azt fogják neki mondani, ‘azért nem te vagy a győztes, mert tilos volt a kezdés előtt elindulni, te pedig ezt tetted’.

Vannak, lehetnek hallgatólagos, nem reflektált szabályok is. Például van egy tényleges szabály, miszerint bárki indulhat a versenyen. Nyilván mindenki úgy gondolja ezt, hogy bárki indulhat, aki ember. Ez azonban így nincs rögzítve. A szabályzat e “hiányossága” csak akkor derülhetne ki, ha egy sajtimádó hegyikecske is pont akkor futna arra, amikor elhangzik a kezdésjelzés, és ügyesen lefele szökdellve elkapná a guruló sajtot. Nem tűnik

valószínűnek ennek bekövetkezte, de elméletileg nem lehet kizárni, és elég biztosnak mondható, hogy ha mégis megtörténne, akkor nem hirdetnék ki a kecskét győztesként, viszont beleírnák a szabályba, hogy kecskék nem versenyezhetnek.

A konstitutív szabályok működési mechanizmusa a következő. Konstituálunk bizonyos társadalmi intézményeket (amelyeket intézményes és nyers tényekkel írunk le).[23] A konstituált fogalmakat sok esetben egymásba ágyazzuk. A konstituált fogalmakhoz, társadalmi intézményekhez sok esetben elsőrendű, regulatív cselekvési szabályokat rendelünk. Ezeket a regulatív szabályokat lehet betartani és lehet megsérteni őket. Amennyiben az utóbbi valósul meg, vagyis megsértik a konstituált intézményekhez rendelt elsőrendű szabályok valamelyikét, akkor a szóban forgó társadalmi intézmény konstitív feltételerendszere nem lesz teljes, tehát a konstitutív szabály alapján létrehozni kívánt intézmény érvénytelen lesz. A kulcsmozzanat tehát az érvényesség kérdése. De mit is jelent ez?

a fehatalmazottságban benne rejlő elismerési igény érvénytelenségéről van itt szó.

igényt

felhatalmazottsága a bírónak van, aki a szabálytalanságokat megállapíthatja dönthet az elsőség kérdésében

A labdarúgás az előzetesen rögzített szabályok által konstituált játék. A felhatalmazottság először ott érhető tetten, amikor változtatnak valamit a játék szabályain. A 2010-es évek egyik nagy kérdése a videobírók bevezetése. Ha valamikor elfogadják a videobíró intézményét, akkor ehhez módosítani kell majd a szabálykönyvet, és további fogalmakat, illetve hozzájuk kapcsolódó szabályokat kell felvenni. A felhatalmazottság szinte definitív erővel jellemzi a játékvezetők tevékenységét. Ők vannak azokkal a jogosítványokkal ellátva, amelyek alapján folyamatosan változtathatják a játékot játszók jogviszonyait (jogokat adnak ide, kötelezettségeket rónak ki oda). Felmerül a kérdés, vajon van-e felhatalmazottsága a játékosnak. Nos, ritkán, kivételes alkalmak esetén még ez is lehetséges. Előfordul olykor, hogy a bíró sérülés miatt megszakítja a játékot. A megszakítás ténye azt jelenti, hogy valamelyik csapat volt a labda birtokában, és a bíró "elvette tőlük a jogot" a további birtoklás lehetőségét, hogy a sérült játékost ápolni lehessen. Az ápolás után a bíró feldobással hozza játékba a labdát, ezzel indítja újra a játékot. Ilyen esetekben, ha nem annak a csapatnak a játékosa szerzi meg a labdát, amely csapat a megszakítás előtt birtokolta a labdát, akkor az a konvencionális elvárás (ami nem a játékszabály része), hogy az ellenfél "önként" (tehát nem valamely játékszabályi előírása miatt) adja vissza a labdát. Ha ezt megteszik, akkor mondhatjuk, hogy éltek a felhatalmazottságukkal, és a labdabirtoklás jogát átadták a másik félnek.

Érdekes helyzetet produkált egy dél-koreai labdarúgó-mérkőzés.[24] Miután az egyik csapat játékosa társa sérülése miatt kirúgta a labdát a pályáról, az ápolás után a másik csapat folytathatta a játékot. A bedobás után az ellenfél támadója a konvenciónak megfelelően "visszaadta" a másik csapat kapusának a labdát, de a lövés túl erősre és "pontosra" sikerült, a kapus nem tudta megfogni a labdát, gól lett belőle. Bár a rúgó játékos mozdulatából látszódott, hogy nem így akarta, elnézést is kért érte, a bíró nem tehetet mást,

[23] Vita van arról, hogy a nyers tények (a fizikai világ X entitásai) feltétlen szükségesek-e az Y társadalmi intézmények definiálásához vagy sem. Barry Smith – John Searle-lel szemben – azt állítja, hogy nem szükséges, vannak olyan társadalmi intézmények, amelyeket nyers tényekre való hivatkozás nélkül definiálunk. A modellem felépítése szempontjából nem perdöntő e kérdésben állást foglalni, így ezt a kérdést itt nem elemzem.

[24] Az esetről beszámoló cikk: [372], a mérkőzés videója: http://www.youtube.com/watch?v=vLfNFZQdvE4

mint hogy gólt ítélt. A középkezdés után az “érdemtelenül”, “erkölcsileg helytelenül” gólt szerzett csapat kapusa a hozzákerült labdát direkt a saját kapujába lőtte, hogy ezzel az öngóllal kárpótolja az ellenfelet.

*

Barry Smith és John Searle [694]

Tummolini és Castelfranchi cikke [758]

David-Hillel Ruben, John Searle’s The Construction of Social Reality [621]

Steven E. Boerl Speech Acts and Constitutive Rules [101]

Barry Smith 2008 [693]

a konstituált entitások, terminusok egy része szerepet definiál a szerep nem rigid fogalom Guarino OntoClean módszertanára hivatkozni, a rigid és nem-rigid fogalmak különbségére (↬58.o.)

[34] On the Notion of an Empowered Agent

juristic act judicial body legislative body in rem in personam

A magyar szabályozás szerint egy házasság akkor jön létre, ha c_1) az anyakönyvvezető hivatalos helységében jelenlevő házasulók c_2) az anyakönyvvezető előtt, c_3) két tanú jelenlétében, c_4) személyesen, c_5) szabad akaratukból, c_6) egybehangzóan kijelentik, hogy házasságot kötnek, és c_7) ezt a tényt az anyakönyvvezető a házassági anyakönyvbe bejegyzi.

$$c_1 \wedge c_2 \wedge c_3 \wedge c_4 \wedge c_5 \wedge c_6 \wedge c_7 \rightarrow \text{RIGHT}$$

A házasság fogalmát definiáló konstitutív szabályrendszer tartalmaz beágyazott társadalmi intézményeket, mint például a tanu vagy az anyakönyvvezető. Ezeket korábbi konstitutív aktusok definiálták. Ebben a kontextusban az anyakönyvvezető, illetve a két házasuló fél a felhatalmazottsággal rendelkező szereplő, és a konstitutív szabályokat át lehet forgatni az anyakönyvvezetőre vonatkozó másodrendű szabályok

a házasulókra nem kötelezettség van előírva, hanem “kijelentési” szabadság

interdependens konstitutív szabály, mert három felhatalmazottsággal rendelkező szereplő kell hozzá

mi van az élettársakkal? ott nem kell az anyakönyvvezető, elég a két fél ráutaló magatartása?

———-

valahova betenni Hansson deklarációs operátorát

———

Az egyéni írásbeli magánvégrendeletetre vonatkozó magyar jogszabály az alábbi alaki érvényességi feltételeket szabja:

> „Az egyoldalas „ . . . írásbeli magánvégrendelet alaki szempontból érvényes, ha készítésének ideje az okiratból kitűnik, továbbá ha a végrendelkező azt
> **a)** sajátkezűleg írt végrendelet esetén elejétől végéig maga írja és aláírja;
> **b)** más által írt végrendelet esetén két tanú együttes jelenlétében aláírja, vagy ha azt már aláírta, az aláírást két tanú előtt, azok együttes jelenlétében a magáénak ismeri el, és a végrendeletet mindkét esetben a tanúk is – e minőségük feltüntetésével – aláírják; vagy
> **c)** a sajátkezűleg írt vagy más által írt végrendeletet aláírja, és akár nyílt, akár zárt

> iratként a közjegyzőnél – végrendeletként feltüntetve – személyesen letétbe helyezi." (PTK 7:17. § Az írásbeli magánvégrendelet alaki érvényességi feltételei [1])

A fentiek alapján ha saját kézzel megírom a végrendeletemet, és azt személyesen zárt iratként közjegyzőnél letétbe helyezem, akkor alaki szempontok alapján érvényessé válik a végintézkedés. Ha halálom után kiderül, hogy a házamat az egyik gyermekemre hagytam, és úgy rendelkeztem, hogy a másik gyermekemé legyen a szomszéd háza, akkor tartalmilag az első rendelkezésem érvényes lehet, de a második nyilván nem lehet az. Ha a két rendelkezési aktust kettébontjuk, hogy külön kezelhessük őket, akkor lesz egy érvényes (a felhatalmazottságot példálózó) és egy érvénytelen (a felhatalmazottság hiányát példálózó) cselekményünk. A két cselekmény megegyezik alakilag. Ezt mondhatjuk úgy is, hogy mindkét esetben szabadságom volt megtenni a szóban forgó cselekvést. Amiben különböznek, az az aktushoz rendelhető igény létében van. Az első esetben érvényes elismerési igényt kapcsolhatunk a szabadsághoz, a második esetben ez hiányzik, vagyis nem rendelhető érvényes elismerési igény az – egyébként szabadon elvégezhető – tevékenységhez.

Hasonló példaként szolgálhatna az a két eset, amikor az egyikben a hivatalban levő miniszterelnök menti fel valamelyik miniszterét, míg a másikban én próbálnám meg ugyanezt. Mindkettőnknek szabadságában állna ugyanazt a mondatot kimondani, mindketten meg is tehetnénk ezt, de az én elismerési igényemet nem (vagy csak kevesen) fogadnák el, míg a miniszterelnökét igen. Az egy másik kérdés, amit még eddig nem vizsgáltam, hogy mi számít az elismerési igény társadalmi elfogadásának. Ha megnézzük az alábbi hírt a kétezres évek Magyarországáról, akkor láthatjuk, hogy különleges és konkurrens elismerési igények olykor felbukkanhatnak a történelemben.

> „Borbély József végre elkezdi a beszédét. Elmeséli, hogy ő volt a Tempó, vagyis Tiszta Energiával Magyarországért nevű párt elnöke, mígnem 2009. június 13-án hajnalban 'arra nem ébredtem, hogy Magyarország királyság.' Bármelyikünket megzavarna egy ilyen hír, de Borbély dicséretesen hamar ráeszmélt arra, hogy nemcsak királyság van, de pont ő az 'önkéntes kormányzó'. ... 2011. november 11-én véget ért a kettős hatalom időszaka, a Kormányzó visszahelyezte jogaiba az Apostoli Magyar Királyságot, Orbán Viktor pedig azóta nem más, mint egy amúgy is illegitim párt elnöke.
>
> December 28-án Kecskeméten a Kormányzó felavatta az Apostoli Magyar Királyság Kormányzói Hivatalát, egyben betiltotta minden magyarországi párt működését. Ha ez nem lett volna elég, idén február 25-én partra szállt Visegrádon és bátor gerillaakcióval levonta a köztársasági zászlót." [734]

Lehet azt mondani a történet főszereplőjére, hogy "bolond", de ez nem változtat azon a tényen, hogy a hír egy olyan eéismerési igény kinyilatkoztatásáról (is) szól, amelyet kevesek azért csak elfogadtak.[25]

A táblázat harmadik sorában az opcionális deontikus operátor szerepel, nem véletlenül. Nem lehet sem a megengedett, sem a mellőzhető deontikus operátorokat használni, mert előbbi magába foglalja a kötelező, míg az utóbbi a tiltás deontikus tartományát. Emiatt viszont

Felmerülhet a kérdés, vajon a 'nem felhatalmazottságnak' lehet-e kétféle értelmezése? Az egyértelmű, hogy értelmes azt állítani, hogy valakinek nem szabad megtenni valamit

[25] A hír szereplőit ezen a videón lehet látni: http://youtube.com/watch?v=avQKJJRDvtA

kötelező	OBLIGATORY$(Ego, \text{DO}(Ego, F))$	$\wedge$	$\neg\text{DO}(Ego, F)$	$\rightarrow$	S
tiltott	FORBIDDEN$(Ego, \text{DO}(Ego, F))$	$\wedge$	$\text{DO}(Ego, F)$	$\rightarrow$	S
opcionális	OPTIONAL$(Ego, \text{DO}(Ego, F))$	$\wedge$	$\text{DO}(Ego, F) \mid \neg\text{DO}(Ego, F)$	$\rightarrow$	$\neg$S
megengedett	PERMITTED$(Ego, \text{DO}(Ego, F))$	$\wedge$	$\text{DO}(Ego, F)$	$\rightarrow$	$\neg$S
mellőzhető	OMISSIBLE$(Ego, \neg\text{DO}(Ego, F))$	$\wedge$	$\neg\text{DO}(Ego, F)$	$\rightarrow$	$\neg$S
felhat.-tt	AUTHORITATIVE$(\text{DO}(Ego, F))$			$\rightarrow$	$\neg F$
szabadság	OPTIONAL$(Ego, \text{DO}(Ego, F))$	$\wedge$	$\text{DO}(Ego, F)$	$\rightarrow$	$\neg$S
igény	OBLIGATORY$(Alter, R(F))$	$\wedge$	$\neg\text{DO}(Ego, R(F))$	$\rightarrow$	S
nem felhat.	$\neg$AUTHORITATIVE(F)			$\rightarrow$	$\neg F$
szabadság	OPTIONAL(Ego, F)	$\wedge$	$\text{DO}(Ego, F)$	$\rightarrow$	$\neg$S
igény	OBLIGATORY$(Alter, R(F))$	$\wedge$	$\neg\text{DO}(Ego, R(F))$	$\rightarrow$	S

és nincs is rá felhatalmazottsága. Az okirathamisító esete lehet a példa erre. De lehetséges-e, hogy valamit szabad megtenni, de nincs rá felhatalmazottság?

logikailag ez nem lehetséges, hiszen a felhatalmazottság fogalma monohierarchikusan generikus alárendeltje a szabadság fogalmának.

A festményhamisítónak lehet szabadsága hamisítani, nekem szabadságom van leváltani az aktuális magyar pénzügyminisztert, de nem lehet igényem a változtatás elismertetésére.

A pénzhamisítónak nincs szabadsága hamisítani.

vannak olyan konstituálások, amelykre lehet szabadságom, vannak, amelyekre tiltás van.

De ha egyik esetben sincs felhatalmazottság, akkor nem lesz érvényesség, viszont ha volt tiltás, akkor akkor nemcsak érvénytelenség, de szankció is lesz (mivel tiltás volt).

---

jogviszony	tartalom	jogkövetkezmény
tilos a dohányzás	dohányzom kocsmában	kötelező a szankció
megengedett a dohányzás	dohányzom otthon	tilos a szankció
kötelező büntetni	dohányzom kocsmában, rendőr büntet	érvényes a szankció
tilos büntetni	dohányzom otthon, rendőr büntet	érvénytelen a szankció
felhatalmazott később nyitni	később nyitok	érvényes
felhatalmazott pénzt nyomni	pénzt nyomtat	érvényes
nem felhat. pénzt nyomni	pénzt nyomtat	érvénytelen+szankció
felhat. lemondani	kiteszi a fotelét lomtalanításkor	érvényes
nem felhat. eladni	eladja a Rottenbiller utca burkolatát	érvénytelen+szankció
felhat. összeadni	anyakönvvez. összeadja a feleket	érvényes
felhat. végrendeletet írni	tanuk nélkül írja meg	érvénytelen
	apaság elismerése	
nem felhat. kinevezni	kinevezem a minisztert	érvénytelen

az útlevélhamisító is pont ugyanolyan útlevelet állít elő, mint az állam, de azért nem lesz útlevél, mert nem érvényes az érvényesség feltételei nem teljesültek

Az öröklés szabályait polgári törvénykönyv egyes fejezetei és más törvények (például a családjogi törvény) ide vonatkozó részei határozzák meg.

Az olyan fogalmakat, mint hogy ki számít rokonnak, házastársnak, tanúnak, ki lesz öröklésre érdemtelen, mi az öröklési sorrend stb. mindig konstitutív szabályokon keresztül definiálja a jogalkotó.

a bírónak, a köztársasági elnöknek kötelező felhatalmazottsága van. A bírónak – ha megállapította a vádlott bűnösségét, akkor – nincs mérlegelési joga, vagyis köteles elítélni a vádlottat, és csak abüntetés mértékében mérlegelhet (ott – relatív – szabadsága van). A köztársasági elnöknek ki kell neveznie a beterjesztett személyeket az adott pozícióba, nem mérlegelhet, vagy alá kell írnia a felterjeszett tövényeket.[26]

F. B. Fitch [258, 259]

Brian F. Chellas [143, 144]

a tárgyi jog szabályai vagy feljogosítanak (más oldalról: köteleznek) vagy felhatalmaznak jogosító vagy felhatalmazó szabályok De miért is fontos a felhatalmazó és jogosító szabályok kettőssége?

10.2.6 A jogviszonyfogalmak teljes logikai rendszere

---

tisztázni, hogy a permitted helyett nem az optional-t kell-e írni a formulákban
ha a pozitív szabadság nem tilosként van értelmezve, a negatív szabadság nem kötelezőként, akkor az előzőben benne van a kötelezőség, az utóbbiban a tiltás, ami nem intuitív

---

mi a két jogcsoport közti viszony?

A jog és kötelezettség fogalmainak hohfeldi „finomszemcsés” felbontása “megkívánja” a fogalmak osztályozását, tipizálását. A logikai rend alapján már utaltam egy fontos tipizálási lehetőségre: a hohfeldi jogfogalmakat két típusba rendezhetjük aszerint, hogy elsővagy másodrendű fogalmak-e (elsőrendű jogosultságfogalom a jogigény és szabadság, másodrendű a felhatalmazottság és mentesség). A logikai rend alapján történő megkülönböztetés mellett természetesen vannak másfajta tipizálási javaslatok a jogfogalmakra vonatkozóan. A két legismertebb ezek közül az aktív és passzív jog, valamint a pozitív és negatív jog megkülönböztetésén alapul.

A logikai rend, illetve az aktív-passzív dimenzió alapján a négy jogosultságtípusunkat az alábbi módon rendezhetjük el:

	aktív jog	passzív jog
elsőrendű jog	szabadság	igényjog
másodrendű jog	felhatalmazottság	mentesség

Bár a két hohfeldi jogcsoport között nagy a strukturális hasonlóság, látnunk kell a különbségeket is köztük. Többször is volt róla szó, hogy mindkét csoportnak van egy-egy központi fogalma, amelyre támaszkodva a csoport további három fogalmát is értelmezni lehet. Az első csoport vezérfogalma az igényjog, a másodiké a felhatalmazottság. Analóg szerkezeti helyzetük mellett azonban figyelnünk kell a két fogalom közti különbségre is. Egy fontos szempont szerint a két fogalom nagyon különbözik egymástól. Az igényjogot ugyanis a passzív jog, a felhatalmazottságot pedig az aktív jog típusába kell sorolnunk.

[26] Ezen a helyzeten nem változtat lényegileg az, hogy első körben részleges vétójoga van.

Ez a különbségtétel természetesen csak akkor lesz érthető, miután már bemutattam a jogfogalmak másfajta tipizálási kísérleteit.

	Alter akarata nem teljesül	Ego akarata teljesül
elsőrendű jog	szabadság	igényjog
másodrendű jog	mentesség	felhatalmazottság

	aktív jog	passzív jog
pozitív jog	pozitív szabadság	pozitív (igény)jog
negatív jog	negatív szabadság	negatív (igény)jog
pozitív jog	pozitív felhatalmazottság	pozitív mentesség
negatív jog	negatív felhatalmazottság	negatív mentesség

szöveg, szöveg

*

tulajdonjog
A hohfeldi jogtipológia ugyan nem a JEPD-elvnek megfelelően különíti el egymástól a négy jogtípust, hiszen nem egymást kölcsönösen kizáró jelentést rendelhetünk hozzájuk, ám ettől még jól megkülönböztethető egymástól a négyféle jogfogalom. A jog terminusában sűrűsödő négyféle jelentésréteg elkülönítése egymástól semmiképpen sem jelenti azt, hogy ezeket csak "magukban" lehetne a tényleges joggyakorlatban jogfogalmainak (tulajdonjog, birtokjog, szerzői jog, vallásszabadság, szólásszabadság stb.) meghatározásához használni. A hohfeldi jogfogalmak segítségével "összetett jogosultságokat" is definiálhatunk. A **tulajdonjog** példáján kiválóan lehet ezt szemléltetni.

Ego privilégiuma	Ego birtokolja, használja, hasznosítja saját tulajdonát, hasznot szed belőle
Ego igénye	Alter önhatalmúlag nem birtokolja, nem használja, nem hasznosítja Ego tulajdonát, nem szed hasznot belőle
Ego felhatalmazottsága	Ego rendelkezik a saját tulajdonával (eladja, elcseréli, elajándékozza, haszonbérletbe adja, végrendelkezik vele)
Ego immunitása	Alter nem rendelkezik Ego tulajdonával (nem adja el, nem cseréli el, nem ajándékozza el, nem veszi haszonbérletbe, nem végrendelkezik vele)

10.2.7 A jogviszonyfogalmak kiterjesztése

A jogviszonyok eddigi tárgyalása során jogelméleti szerzők gondolataira támaszkodtam. Ez gyakorlati szempontból talán érthető, hiszen ezen a tudományterületen született a legtöbb elemzés e kérdés kapcsán, elméletileg azonban megalapozatlan ez a leszűkítő tárgyalásmód. Az lett volna pontos (és, tán, a helyénvaló), ha a 'jogszonyok' fogalma helyett mindvégig a 'normatív viszonyok' kifejezést használtam volna, de ettől azért tekintettem el, mert nagyon nehézkesnek, kicsit elidegenítőnek, "kontextusvesztőnek" éreztem volna egy jogelméleti gondolatmenet során.

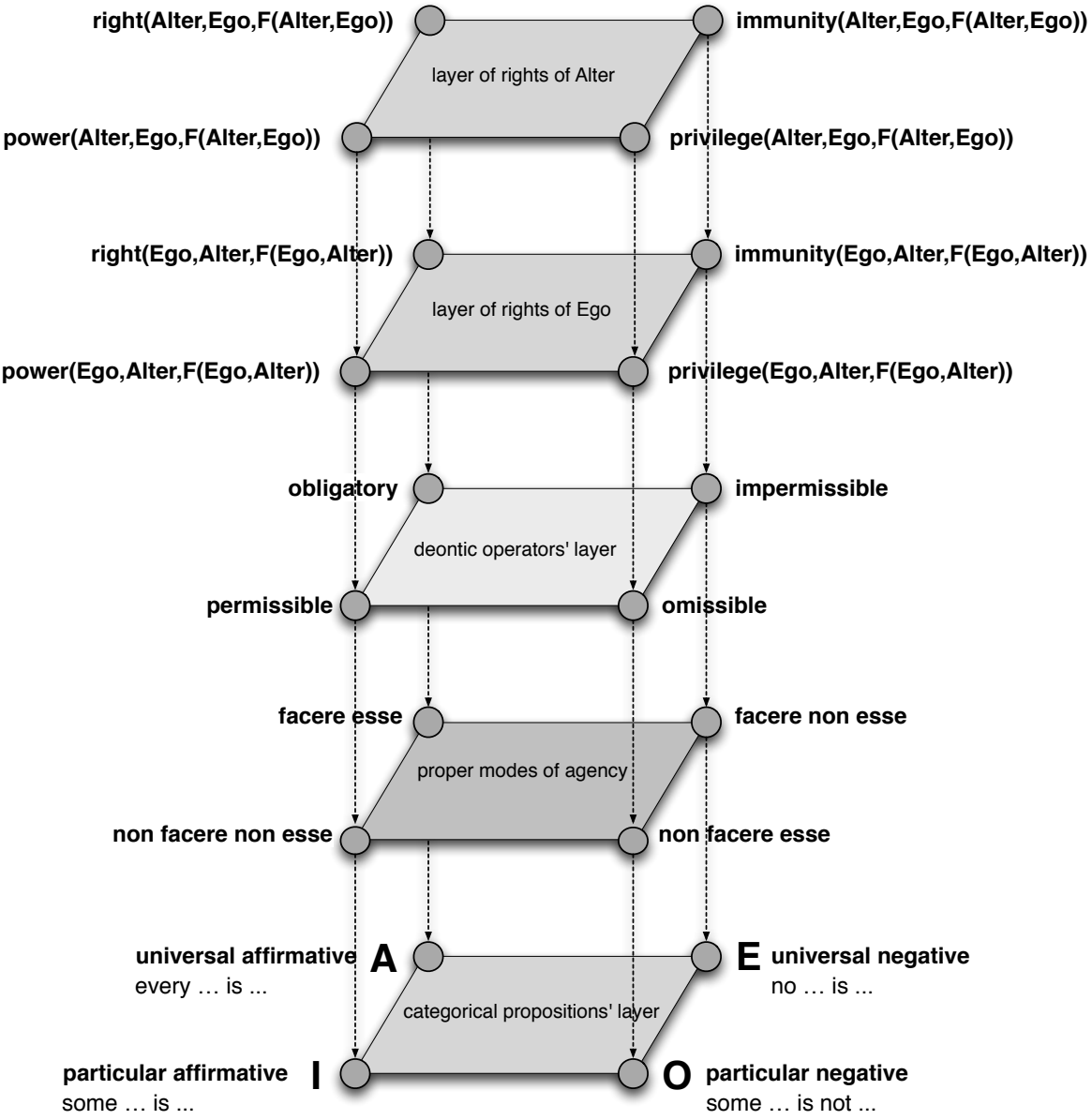


Figure 10.4: jogtípusok logikai négyzete

A továbblépéshez azonban érdemes ezt az általánosítást megtenni, mert ezzel mód nyílik arra, hogy egyrészt más tudományterületek tudását hasznosíthassuk, másrészt megpróbálkozhítunk azzal, hogy a jogviszonyok elemzése során eddig elért eredményeket máshol is "értékesíthessük". Az első lépést még egy jogász, Frederick Schauer "segítségével" tehetjük meg. Schauer ugyanis – szigorúan jogelméleti kérdéseket tárgyalva – az intézményeket az általuk hozható döntések típusai alapján különíti el egymástól.

> „Maga a hatáskör fogalma adja azt a hasznos eszközt, mely segítségével különbséget tehetünka döntéshozásra adott felhatalmazás és a meghozott döntés tartalma között. Következésképpen kétféle döntéshozó intézményt különböztethetünk meg: egyfelől azokat, melyek rendelkeznek mind a hatáskörrel felruházó, mind a döntések eredményét meghatározó tartalmi szabályokkal; másfelől pedig azokat, melyek csak az

előbbi csoportba tartozó szabályokkal rendelkeznek, s így az eredmény meghatározását a döntéshozó szabályok által nem vezérelt mérlegelésére bízzák." [640, 414]

Azért fontos Schauer megkülönböztetése, mert azáltal, hogy az intézményekhez rendelt jogviszonyok – jogelméleti – tárgyalását "áthelyezzük" egy döntéshozatali – szociológiai – kontextusba, láthatóvá, elemezhetővé válhatnak az eltérő jogosultságokkal való rendelkezés szociológiai konzekvenciái. Minden intézményt jogosultságok és kötelezettségek adott rendszerével jellemezhetünk, és minden intézményben a rá jellemző jogviszonyrendszer határozza meg azt, hogy miként cselekszenek, cselekdhetnek a szereplők. Ezek a cselekvések a szereplők döntésein keresztül is értelemezhetők, tehát a döntési kompetenciák különbségei határozzák meg az intézmények közti különbségeket.

„Azokban a döntéshozó intézményekben azonban, melyekben csak hatáskört megállapító felhatalmazó szabályokat alkalmaznak, hiányoznak a döntések kimenetelét megszabó tartalmi korlátok. A döntéshozó felhatalmazása szélesebb annál, semhogy egyszerűen szabályokat kelljen alkalmaznia a konkrét esetekre. A társadalmi és morális elvek és célkitűzések [*policy*] teljes készletére támaszkodva kell eldöntenie, hogy mi a teendő az előtte álló konkrét esetekben." [640, 415]

Schauer szemléletes példája a rendőr "intézménye", akinek a cselekvését felhatalmazó és tartalmi szabályok egyaránt meghatározzák (Schauer tartalmi szabályon a regulatív szabályokat "érti", bár erre nem igazán reflektál). A kétféle szabály közti viszonyra fény derül az alábbi idézetből:

„Az a szabály, mely felhatalmaz egy rendőrt arra, hogy mindazokat, de csakis azokat beidézze, akik átlépték az óránkénti kilencven kilométeres sebességhatárt, példaként szolgálhat arra az esetre, amikor a hatáskört megállapító felhatalmazó szabály és az eredményt meghatározó tartalmi szabály együttesen van jelen. Bár a rendőr fel van hatalmazva arra is, hogy idézést adjon ki, s arra is, hogy megállapítsa az autók sebességét, nincs felhatalmazva sem annak meghatározására, hogy mennyi legyen a megengedett legnagyobb sebesség, sem annak meghatározására, hogy mikor mentesülhet valaki a következmények alól, amennyiben átlépte a sebességhatárt." [640, 414-415]

Attól a pillanattól kezdve azonban, amikor már nem jogokról, hanem döntésekről, társadalmi intézményekről beszélünk, hirtelen "kinyílik" a fogalmi, elemzési tér, és a jogrendszeren kívüli normativitásban is ugyanúgy alkalmazhatjuk ezeket a distinkciókat. Nemcsak a döntéshozati jogosítványok alapján megkülönböztetett intézményfogalmakat, hanem minden más, a jogviszonyokról eddig bemutatott fogalmakat is alkalmazhatunk azokon a területeken, ahol jogosultságokról, kötelezettségekről általában beszélhetünk. A jog mint jogrendszer fogalma felől az általánosabb norma fogalma felé elmozdulva más dimenziókba léphetünk át. Bár Somlai Péter a családon belüli társadalmi kapcsolatok világát elemezte "csak", az alábbi meglátása általánosítható mindenféle normatív cselekvést vizsgáló tudományos tevékenység számára.

„A célracionális cselekvések modelljeivel dolgozó kutatónak tehát nemcsak vágyaktól és szokásoktól, értékvezérelt vagy irracionális döntésektől kell elvonatkoztatnia. Mindezekkel elvonatkoztat a szabályalkotás és intézményesedés folyamataitól is. Attól tehát, hogy *(a)* lehetnek olyan szabályok, amelyek valamiképp hozzáférhetők a

> résztvevők számára (akik formálhatják ezeket, illetve felismerhetik érvényességükben a saját kompetenciájukat, jogaikat és felelősségüket) és *(b)* a szabályok és intézmények is változhatnak a szereplők kapcsolatának és konfliktusainak eredményeként." [701]

Bármely társadalmi intézményben, ahol normák szabályozzák a résztvevők cselekvéseit, a kétféle szabály- vagy jogfogalom elemzése adhatja az eszközt az intézményen belüli dinamikák megértéséhez. A jogrendszert elemezve tudtuk az alapvető jogviszonyfogalmakat értelmezni, pontosan definiálni, de ezek a fogalmak – kellő általánosítás után – sok más területen is használhatók. Mindenhol, ahol társadalmi intézményekben emberek saját normatív viszonyaikat próbálják meg szabályozni, fenntartani, változtatni. Vagyis mindenhol, ahol emberek cselekszenek.

A.V. Cicourel Basic and Normative Rules in the Negotiation of Status and Role [159]
Pamela Oliver

> „It is this possibility of altering the parameters that determine others' (and evenone's own) behavior which makes selective incentive such an important part of the dynamic processes of collective action. People need not merely react to their situation. They can act to alter the very situation which constrains their choices." [550]

antropológia konstitutív intézményei, fogalmai [184]
Hart a 54-55. oldal: jó összefoglaló a 3. fejezetről
70-73
másodrendű szabályok
elismerési szabály változtatási szabály ítélkezési szabály
még hohfeldről.

Allan, From the Fundamental Legal Conceptions of Hohfeld to Legal Relations: Refining the Enrichment of Solely Deontic Legal Relations [8]

Munindar P. Singh, An ontology for commitments in multiagent systems: Toward a unification of normative concepts [686]

Castelfranchi Cristiano, Commitments: from individual intentions to groups and organizations [138]

megengedő normákról

Lewis ?master and slave? game [30] shows that the notion of permission alone is not enough to build a normative system, because only obligations divide the possible actions into two categories or spheres: the sphere of prohibited actions and the sphere of permitted (i.e., not forbidden) actions or ?the sphere of permissibility?. More importantly, Bulygin [13] explains why permissive norms are needed in normative systems using his ?Rex, Minister and Subject? game. ?Suppose that Rex, tired of governing alone, decides one day to appoint a Minister and to endow him with legislative power. [...] an action commanded by Minister becomes as obligatory as if it would have been commanded by Rex. But Minister has no competence to alter the commands and permissions given by Rex.? If Rex permits hunting on Saturday and then Minister prohibits it for the whole week, its prohibition on Saturday remains with no effect. [97]

ezt összekapcsolni a megengedés erős és gyenge értelmezésével

Substantive and procedural norms in normative multiagent systems [100] „Procedural norms are instrumental norms, addressed to agents playing roles in the normative system, which aim at achieving the social order specified in terms of substantive norms."

Magasabbrendű normák

10.2.8 pozíciók

Anderson [11, 13, 16, 17]
Pörn [579, 580]
Krogh és Herrestad [353, 426, 427, 353]
Makinson [464], [466]
Marek Sergot és Jones [672, 385, 671, 673]
On the Characterization of Law and Computer Systems: The Normative Systems Perspective [386]
A Formal Characterisation of Institutionalised Power [385]
Swan Ove Hansson [337]
Vranas [801, 799, 800, 797, 798]

10.2.9 Jogi szemantika

Giovanni Sartor [638], [637]

10.2.10 Deontikus logika és az informatikusok

konferenciasorozat deontic logic in computer science (ΔEON) könyvek: Deontic Logic and Artificial Normative Systems [299]
Deontic Logic in Computer Science, 2004 [457]
Deontic Logic in Computer Science, 2008 [776]
Games, Norms and Reasons. Logic at the Crossroads [774]
G. Boella, L. van der Torre és H. Verhagen 2006-os cikke: Introduction to Normative Multiagent Systems
A deontikus logika és számítástudomány több évtizedes viszonyáról áttekintést nyújt Guido Boella, Gabriella Pigozzi és Leendert van der Torre cikke [99]

> „With 'norm' we mean 'a principle of right action binding upon the members of a group and serving to guide, control, or regulate proper and acceptable behavior'." [99]

> „A normative multiagent system is a multiagent system together with normative systems in which agents on the one hand can decide whether to follow the explicitly represented norms, and on the other the normative systems specify how and in which extent the agents can modify the norms." [97]

The mechanism design definition:

> „A normative multiagent system is a multiagent system organized by means of mechanisms to represent, communicate, distribute, detect, create, modify, and enforce norms, and mechanisms to deliberate about norms and detect norm violation and fulfillment." [98]

10.3 Társadalmi szerep

szerep, komplementaritás, reciprocitás

Talcott Parsons, Toward a General Theory of Action. Theoretical Foundations for the Social Sciences, [562]

sir David Ross [612]

Alvin W. Gouldner [304]

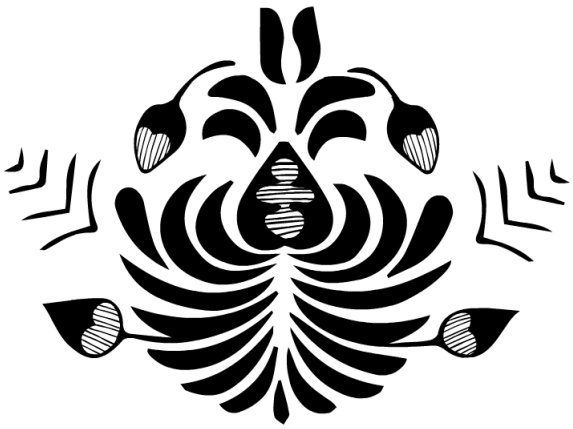
státus és privilégium

A.V. Cicourel Basic and Normative Rules in the Negotiation of Status and Role [159]

10.4 Privacy

Hirshleifer, Privacy: Its Origin, Function, and Future, három fémszabály, arany-, ezüst-, vasszabály [357]

social evolution has instilled in man elements of at least three distinct social ethics: one associated with the Golden Rule of sharing,a second with the Silver Rule of private rights and reciprocity, and a third with what might be called the Iron Rule of dominance and subordination



Chapter 11

Uralom

Talán sokan gondolhatnák azt, hogy a normatív viszonyok vizsgálata után a norma általános fogalmának elemzésével kellene folytatnom. Én mégsem ezt teszem, hanem az **uralom** (AUTHORITY) fogalmát veszem elő. Ennek az az oka, hogy mind a fogalmi építkezés szempontjából, mind a keletkezéstörténeti szempontból előbbrevalónak tartom az uralom fogalmát a normához képest.

Joseph Raz a jog preemtív jellege, mindent kizáró, az érvelési teret teljesen kiüresítő, és oda egyedül sajátmagát belehelyező, minden más gyakorlati indok helyébe lépő "képessége"

Az első utalmi mozzanat akkor született meg, amikor valaki (Ego) valamilyen engedelmeskedési igénnyel lépett fel egy másik emberrel (Alterrel) szemben, és ezt az igényt utóbbi elfogadta, amire azt mondhatjuk, hogy érvényes volt számára a

nem gondolom, hogy az uralom, a legitim érvényesség kérdését a jogra, a politikára, az államra kellene szűkíteni

Az uralom ott van minden normatív viszonyban, és ott van minden társadalmi intézményben, a családban, a párkapcsolatban, az iskolában, a munkahelyen stb.

uralom - autoritás, tekintély

uralom elismerése = a kiadott paramcsot az alávetettek elismerésre méltónak, példaszerűnek, mintaszerűnek, illetve belsőleg érvényesnek tartják

Milyen gyakorlati indokokat vehetünk figyelembe a cselekvésbefolyásolás során?

Fontos rögzítenünk, hogy itt elsődlegesen cselekvésbefolyásolásról beszélünk, aminek a központi mozzanata az egyén cselekvésére irányuló **akarata**. A cselekvésbefolyásolás lényege abban van, hogy valakinek a saját cselekvésére irányuló akarata megváltozik egy másik ember hatására. Lehet valakit vitán keresztül befolyásolni, aminek eredményeként megváltoztatja a véleményét. Bármennyire is fontos ez a helyzet a társadalmi cselekvések világában, nem foglalkozom vele. A véleménybefolyásolás egész kérdéskörét kihagytam, kihagyom ebből a könyvből. Ha a vélemény megváltozatásán kereszüli hatást kizárom a vizsgálódásom fókuszából, akkor marad a cselekvésre irányuló akarat más eszközökkel történő megváltoztatása.

Az uralom fogalmának legfőbb összetevője a mindenféle érdektől független **engedelmeskedni akarás** mozzanata.

ez jelenti/jelzi/konstituálja az uralom érvényességét

Balázs Zoltán A politikai közösség, Osiris Kiadó, hatalom-uralom fogalma Hatalom könyve

11.1 Érvényesség, legitimitás

hopp ezt hogy kell értékelni?

sztrájkjog: legális föllépés érvényes szerződésekkel szemben
Garai Zsolt, értékek, normák [281]
Eliade, szent és a profán [229]
von Wright az érvényesség két értelmezéséről [792, 195]
validity as existence – law in force (hatályos)
validity as legal – legal law
itt von Wright legitimitás kérdését feszegeti, de leszűkíti annak értelmezését.

> „Amikor egy társadalmi csoportban valamely szokás általános, ez az általánosság csupán a csoport legtöbb tagjának megfigyelhető magatartására vonatkozó tény. Ahhoz, hogy a csoportban éljen egy bizonyos szokás, a csoport egyetlen tagjának sem kell tudatosan szem előtt tartania az általános magatartást, vagy akár tudnia, hogy a szóban forgó magatartás általános, s még kevésbé kell arra törekednie, hogy tanítsa vagy fenntartsa ezt a mintát. Elég, ha mindenki a maga részéről úgy viselkedik, ahogy a többiek. Ezzel szemben ahhoz, hogy egy társadalmi szabály létezzen, legalább néhány embernek úgy kell tekintenie a kérdéses magatartást, mint a csoport egésze részéről követendő általános mintát. Egy társadalmi szabálynak létezik tehát egy 'belső szemlélete' is, azon a külső szemléleten túl, melyben nem különbözik egy társadalmi szokástól, és ami nem más, mint az ismétlődő viselkedés, amit egy megfigyelő észlelhet." [346, 73]

Hart 'belső szemlélet' (internal aspect) fogalma a közös tudás
a szokást csak a 'külső szemlélet' (external aspect) jellemzi
The Sources of Normativity [413]
Performatives in a rationally based speech act theory, [165]

> „Az iráni kormány nem ismer el más tekintélyt vagy hatalmat, csak a Mindenható Istenét, és semmiféle jogi hagyományt, csak az iszlám törvényét. Ilyen feltételek mellett az iráni delegáció újra megerősíti, hogy a nemzetközi szervezetek iszlámmal ellentétes nyilatkozatainak, határozatainak vagy döntéseinek semmiféle érvényessége nincs az Iszlám Köztársaságban. Az 'Emberi Jogok Egyetemes Nyilatkozata', amely a zsidó-keresztény hagyomány világi koncepcióját illusztrálja, nem alkalmazható a muzulmánok által, és egyáltalán nem kapcsolódik az Iráni Iszlám Köztársaság által elismert értékrendhez; az Iráni Iszlám Köztársaság ezért nem habozhat e rendelkezések megsértésekor, ha az ország isteni törvényének megsértése vagy a világi egyezségokmányok megsértése között kell választania." [619]

> „A természetjog arra törekszik, hogy a célok igazságos voltával 'igazolja' az eszközöket, a pozitív jog pedig arra, hogy az eszközök jogos voltával 'szavatolja' a célok igazságosságát." Walter Benjamin 29.

kötelezőség és kívánatosság elemzése Durkheimnél:
http://www.tankonyvtar.hu/hu/tartalom/tkt/durkheim-durkheim/ch12.html

„ . . . a szentet a profánhoz viszonyítva már csakis különneműségük alapján határozhatjuk meg. E különneműség azért elegendő a dolgok ezen osztályozásának jellemzéséhez és az összes többitől való megkülönböztetéséhez, mert igen sajátos: abszolút. Az emberi gondolkodás történetében nincs még egy példa a dolgok ennyire mélyen különböző, ennyire ellentétes két kategóriájára." (Durkheim, 1912m, 46)

„Amikor azt állítjuk róluk, hogy szentek, ezen azt értjük, hogy értékük nem mérhető össze a többi értékkel. Szent az, amit elkülönítettek, aminek nem lehet közös mértéke a profánnal." [224, 190]

Hart: érvényesség, elismerési szabály
belső szemlélet közös tudás és norma
jogászi terminológia:
érvényesség, belső szempont vs. érvényesülés (létezés, igazodás), külső szempont

„Mindenki vesse alá magát a fölöttes hatalomnak. Mert nincs hatalom, csak az Istentől, ami van, azt Isten rendelte. Ezért aki a hatósággal szembeszáll, Isten rendelésének szegül ellene. Az ellenszegülő pedig magára vonja az ítéletet.
A felsőbbség nem arra való, hogy a jótettől ijesszen el, hanem a rossztól. Azt szeretnéd tehát, hogy ne kelljen a hatalomtól tartanod? Tedd a jót és megdicsér:
Mert az Isten eszköze a javadra. De ha gonoszságot művelsz, rettegj, mert nem hiába viseli a kardot. Isten eszköze, hogy a gonosztevőt megbüntesse.
Alá kell tehát magadat vetned neki, nemcsak a megtorlás miatt, hanem a lelkiismeret szerint is.
Ezért fizettek adót is, hiszen akik behajtják, Isten szolgái. Adjátok meg mindenkinek, ami jár neki: akinek adó, annak az adót, akinek vám, annak a vámot, akinek hódolat, annak a hódolatot, akinek tisztelet, annak a tiszteletet." (Róm 13.1-13.7)

11.1.1 Legitimitástipológia

szabályhoz vagy emberhez igazodás
nemcsak szabály alapú legitimitás van
nemcsak a legalitás szabályalapú
emberhez igazodás: karizma
Jézus paradigmatikus példáját idézve:

„Hallottátok, hogy a régiek ezt a parancsot kapták: . . .
Én pedig azt mondom nektek . . . " [348]

Weber szerint az uralmat a legitimitásigény szerint lehet tipizálni.

„ . . . célszerű az uralom fajtáit a rájuk jellemző legitimitásigény szerint megkülönböztetni. . . . Itt az a mérvadó, hogy maga az uralom milyen fajta legitimitásra tart igényt, és saját legitimitásigénye releváns mértékben érvényesül-e . . . " [810, 222]

A parancs és az érték a szabályhoz hasonlóan normatív. Azonban míg a parancs nélkülözi az általánosságot (?Légy olyan jó, vidd le a szemetet!?), a szabály általános (?A jó gyerek mindig elpakol maga után!?), az érték pedig feltétel (előtag) nélküli kellést fogalmaz meg (?Légy jó mindhalálig!?)

11.2 Uralomtípusok

Amikor a *Gazdaság és társadalom*-ban Max Weber kifejti az uralomtipológiáját, akkor három típust mutat be, a tradicionális, a karizmatikus és a racionális uralmat [810, 224]. Korábbi munkáimban én négyelemű utalomtipológiát használtam [731], [722], de nem igazán támasztottam alá, miért gondolom kiegészítendőnek a weberi tipológiát. Most megpróbálkozom vele.

A legitimitás fogalmának elemzésekor már jeleztem, hogy Weber az uralomtípusokat a hozzájuk kapcsolható legitimitásigény alapján tartja elkülöníthetőnek egymástól. Ezt a weberi elképzelést úgy lehet interpretálni, hogy

11.2.1 Tradicionális uralom

„ . . . elsődlegesen az emberemlékezet óta érvényes tradíciók szentségébe és a tradíciók által autoritással felruházott személyek uralmának legitim voltába vetett mindennapi hiten alapul.” [810, 224]

11.2.2 Racionális uralom

„ . . . elsődlegesen a legalitásba – a tételes rendnek és a tételes rend által az uralom gyakorlására kijelölt személyek utasítási jogának legális voltába – vetett hiten alapul.” [810, 224]

11.2.3 Karizmatikus uralom

„Charisma is the ability to influence without logic.” (Quentin Crisp)

„ . . . elsődlegesen nem mindennapi odaadáson alapul egy szent, hősies vagy példamutató személy, illetve az általa kinyilatkoztatott vagy megteremtett rend iránt.” [810, 224]

karizmatikus uralom + közös tudás

„A karizmatikus vezető, ahelyett, hogy – ahogy az a szimbolikus küzdelemben alulmaradottak esetében fennáll – saját maga számára is az lenne, aminek a többiek tartják, eléri azt, hogy az őt vezetővé választó csoport számára is azzá váljék, aminek saját maga véli magát. Ő teremti meg az őt magát megteremtő közvéleményt.” [107, 156]

„A tradíciókhoz köthető korszakokban a karizma jelenti *a* nagy forradalmi erőt. Eltérően a másik, ugyancsak forradalmasító erőtől, a »ratio«-tól amely vagy egyenest kívülről hat – úgy, hogy megváltoztatja az életkörülményeket és az életproblémákat, és ezzel közvetve a hozzájuk való viszonyulást is – vagy azon keresztül hat, hogy intellektualizálja a problémákat, a karizma kínból vagy lelkesedésből fakadó belső átalakulást *tud* előidézni, amely különféle viszonyulásokban – minden egyes életformához és általában a »világhoz« való viszonyulásban – kialakuló tökéletesen új beállítottság hatása alatt teljesen megváltoztatja a leglényegesebb érzületi és cselekvési irányokat.” [810, 251-252]

„A karizmatikus tekintély tényleges érvényessége ugyanis valójában teljes egészében azon múlik, hogy karizmája »beigazolódik-e«, mert ez a feltétele annak, hogy az alávetettek *elismerjék*, ami persze a karizmatikus képességekkel rendelkező *és ebből következően* legitim úrral szemben *kötelességük*. A szervezeti kapcsolatok növekvő racionalizálása esetén azonban kézenfekvő, hogy ezt az elismerést már nem a legitimitás következményének, hanem e helyett a legitimitás *alapjának* fogják tekinteni . . . ” [810, 271-272]

11.2.4 Átértelmezett karizmatikus uralom

11.2.5 Áttekintés

táblázatot írhatjuk fel [732, 219]:

	az uralmi szabályok változtatásnak módja:	
a legitimititásigény:	szabályértelmező	személycserélő
földöntúli akaratra hivatkozik	tradicionális	karizmatikus
emberi akaratra hivatkozik	racionális-legális	átértelmezett karizmatikus

ezért – bármennyire is ellenétetesnek tűnik az intuícióinkkal – a kommunista-bolsevik legitimitásra azt kell mondjuk, hogy az tradicionális uralomnak feleltethető meg [727].

A kandidátusi értekezésemben megírtam, hogy a kommunista-bolsevik uralom nagyon sok hasonlóságot mutat a tradicionális uralommal [727]. Éppen ezért azt javasoltam, hogy azt a történelmi formációt is – legalább részben – besorolhatjuk a tradicionális uralom fogalma alá. Ez sokaknak elfogadhatatlan volt, nekem sem tetszett, de nem tudtam jobbat javasolni helyette.

	a szabályok változtatásnak módja:	
a legitimititásigény:	szabályértelmező	személycserélő
nem emberi akaratra hivatkozik	?????	karizmatikus
emberi akaratra hivatkozik	racionális-legális	átértelmezett karizmatikus

Chapter 12

Norma

„Freedom from a bad norm is a public good that is often very difficult to bring about." Cristina Bicchieri [83]

Nagyon sok jelentésű fogalom

Social norms become the moral rules of the group and specify what should be done to remain a member of a group as well as maintain a self-image of being good, kind and helpful (Asch 1955; Berkowitz 1972; Cialdini and Trost 1998; Pepitone 1976; Schaffer 1983; Schwartz 1977). In addition to being motivated to be helpful and to maintain self-image, individuals conform to norms to avoid feelings of shame, guilt and fear (Aronson 1969; Baumeister and Leary 1995; Cialdini, Kallgren and Reno 1991; Eagly and Chaiken 1993; Pepitone 1976; Reno et al. 1993; Staats et al. 2009a). Conformity to such norms is rewarded repeatedly in society (Opp 1982) since helpful individuals are held up to others as examples and heroes. Hence, people conform to such norms even when the consequence for providing this assistance is unpleasant (Latane and Darley 1968; Schwartz 1970). Schwartz (1970) showed that a fairly high percentage (59%) of people are willing to donate their bone marrow to save another, even though they would suffer some pain during the procedure and physical discomfort for several days.

Pepitone A, Toward a normative and comparative biocultural social psychology. [565]

Toward Pepitone's vision of a normative social psychology: What is a social norm? [676]

A **norma** fogalma alatt a **társadalmi normát** értem, és amikor nem okoz zavart, akkor a két kifejezést szinonímnak tekintem.

Arisztotelész: phronészisz

A sztoikus filozófiában a phronészisz, mint erény azoknak a dolgoknak a tudása, amelyeket meg kell tennünk ill., amelyeket nem szabad megtennünk.

Georges Canguilhem a normalitásról [136]

The Normativity of the Natural. Human Goods, Human Virtues, and Human Flourishing [147]

Sartor et al. owl ontology of norms [622] [284]

A 'szemet szemért' elvben létezik egyfajta deontikus többértelműség, mert nem derül ki egyértelműen, hogy milyen deontikus modalitás mentén kell értelmezni. Akkor érhető

tetten az a többértelműség, ha megpróbáljuk átfogalmazni az elvet. Ezt ugyanis megtehetjük többféleképpen is:

> Ha kiszúrják a szemed, szúrd ki a szemét! ↣ kötelező kiszúrni a szemét
>
> Ha kiszúrják a szemed, kiszúrhatod a szemét! ↣ megengedett kiszúrni a szemét

Lehetséges kötelezésként vagy megengedésként is értelmezni a jólismert erkölcsi parancsot.[1]

ez egy feltételes norma

lehet úgy is értelmezni, hogy az az ősi parancs felhatalmazottságot teremt: abban az esetben, ha a feltételben megfogalmazott cselekvés normasértőnek minősül, és ennek szankcionálására hatalmazza fel a parancs a

talió elv lex talionis

Hammurapi törvénykönyve [325]

Tokics Imre, phd-dolgozat: [749]

12.1 A norma leírása

Ebben a fejezetben bemutatom azt, miként lehet a norma jelenségét leírni. Először áttekintem, hogy a norma fogalmát milyen összetevőkre bonthatjuk fel. Ennek során a von Wright javaslatából indulok ki [792], de néhány új elemmel kiegészítem az eredeti von wright-i elképzelést. A kiegészítéseket viszont nem sorolom a norma összetevői közé, ezért a kibővített készletre a **norma jellemzőiként** hivatkozom. Ezek a normajellemzők arra is alkalmasak lesznek, hogy segítségükkel a fejezet második részében egy sok szempontú, kiterjedt nomatipizálást tudjak elvégezni.

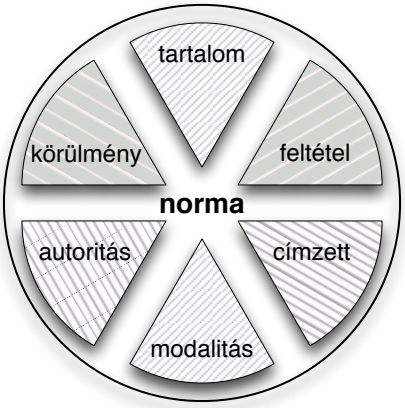
12.1.1 A norma jellemzői

A norma leírásához összetevőire kell bontanunk a fogalmat, hogy ezáltal láthatóvá váljék, milyen elemi adatok szükségesek ahhoz, hogy egyértelműen megadhassunk egy adott normát. Von Wright hat összetevőt tartotta elkülönítendőnek [792, 70]:

- norma tartalma (content of norm)
- norma modalitása (character of norm)
- norma feltétele (condition of norm)
- norma alkalma (occasion of norm)
- norma autoritása (authority of norm)
- norma címzettje (subject of norm)

Nem vitás, hogy ezek az összetevők szükségesek a normák értelmezéséhez. A kérdés csak az, hogy elégségesek-e, nincs-e szükség újabb összetevők felvételére. Szerintem van még két olyan szempont, ami mentén további adatokat rendelhetünk a norma jelenségéhez. Mielőtt azonban ezekre kitérnék, egy ábra segítségével is rögzítsük az eredeti – von wright-i – összetevők körét.

[1] Azt is fel lehetne vetni, vajon a megengedett vagy az opcionális modalitást kellene itt használnunk, de ennek a kétértelműségnek a kezelése most nem annyira fontos.



Ezek után nézzük meg, mivel is lehetne kiegészíteni a fentieket. Von Wright és a logikusok mindig is érzéketlenek voltak bizonyos társadalmi mozzanatok iránt a normatív jelenségek vizsgálata során. Az egyik ilyen vakfolt annak fel nem ismerése, hogy a normativitás csak egy részét jelenti a cselekvésbefolyásolás jelenségének, és a fenti von wright-i keretrendszer nemcsak a normákra húzható rá, de alkalmas lehet másfajta cselekvésbefolyásolási helyzetek leírására is. Ha viszont ez így van, akkor valahogyan jelezni kell tudnunk azt, hogy a normativitás és más cselekvésbefolyásolási jelenségek, mint például a hatalmi helyzetek miben térnek el egymástól. Ennek "kezelésére" valamilyen módon kezelnünk kell a legitimitás, a **legitim érvényesség** kérdését (az érvényesség létét vagy nem létét, illetve annak típusait).

A másik hiányossága a fenti hat-elemű összetevőre bontásnak az, hogy nem veszi figyelembe a normativitáshoz elengedhetetlenül hozzátartozó szankcionálás jelenségét.[2] A normativitásban tetten érhető, jövőre irányuló cselekvésbefolyásolási szándék önmagában nem elégséges ahhoz, hogy a szándékolt cselekvés megvalósuljon. Ehhez szükség van arra, hogy amennyiben az elvárt cselekvés nem valósul meg, akkor valamilyen szankció kövesse azt. Korábban bemutattam, hogy egy teljes deontikus kalkulust lehet definiálni a szankció fogalmának felvételével, ami már önmagában is elég okot szolgáltat arra, hogy ezt a kategóriát is figyelembe vegyük a norma jelenségének leírásakor. De azért is szükség lehet erre a fogalomra, mert a normasértéshez kötött szankció jelensége biztosítja a talán legfontosabb normatipizálási lehetőséget is egyben. És ha már a norma fogalmi modelljébe beemeljük a szankcionálás mint másodrendű cselekvés fogalmát, akkor meg kell tennünk ezt a normákhoz szükségszerűen kapcsolódó másik cselekvéstípussal, a **kibocsátás** fogalmával is. A norma kibocsátása, illetve a normasértés szankcionálása másodrendű cselekvésnek számít, míg a normához való igazodás, a norma követése, illetve a normasértés cselekvései elsőrendű fogalmakkal ragadhatók meg. Utóbbiakra érdemes egy átfogó fogalmat alkalmazni, hogy könnyebben, egységesen hivatkozni tudjunk rájuk. Erre a célra fogom használni a **normavezérelt cselekvés** kifejezést.

A fenti hat összetevő közül von Wright a modalitást, a tartalmat és a feltételt együttesen minősíti a **normamagnak** (norm-kernel) azzal a tézissel együtt, hogy ez minden előírás-

[2] Von Wright ugyan mint fontos jelenséget megemlíti a **szankció** jelenségét, ám azzal a kikötéssel, hogy az nem tartozik a norma összetevői közé [792, 70].

nak (preskripciónak) része. A további három összetevőről tehát azt állítja, hogy nem feltétlen találjuk meg őket bizonyos normatípusokban.

öööööööööööööö

A norma tartalma

A **normatartalom** fogalma alatt azt a cselekvést kell értenünk, amire a normatív befolyásolási szándék irányul. Ebben a felfogásban a tartalom valamilyen cselekvés, nem más. Nem világállapot, nem folyamat, nem változás, nem puszta esemény. Cselekvés, mint például a dohányzás, az autóvezetés, a szobába belépés, a köszönés stb. Azt, hogy a cselekvést hogyan lehet leírni a világállapotok, az átmetek, valamint a cselekvésoperátorok segítségével, a könyv egy korábbi fejezetében bemutattam (↬267.o.). A normák leírásakor nincs szükség többre, mint, hogy az ott bemutatottak szerint írjuk le a cselekvéseket, és ha így teszünk, egyúttal megadjuk a norma tartalmát is.

Fontos figyelni arra, hogy a normatartalom megragadásakor nem vehetjük bele a norma deontikus értékét. A 'Ebben a helységben tilos a dohányzás!' norma esetében a norma tartalma "csak" a dohányzás lesz, a tiltás modalitását a norma másik összetevőjével írjuk le.

A norma modalitása

Amit von Wright a norma karaktereként (character of norm) ad meg, azt én a **norma modalitásaként** nevezem. Modalitáson pedig azt a deontikus operátorral kifejezhető normatív szándékot ragadhatjuk meg, amellyel a normát kibocsátó (ellenőrző, számonkérő) személy a norma címzettjét cselekvésre akarja bírni. A deontikus modalitás kérdését alaposan elemeztem a könyvben, bővebb áttekintésekért ezekre a fejezetekre hivatkozhatok **??**, illetve 377. oldalaktól).

A norma feltétele

Amilyen fontos fogalom a norma felétele, annyira könnyű félreértelmezni a jelentését. A 'feltételesség' sokjelentésű fogalma miatt könnyen zavar alakulhat ki. A **norma feltételének** – legalábbis a von wright-i interpretáció szerint – nincs köze a **norma körülményeihez**, ennek ellenére könnyű "összekeverni" őket a két fogalomban "megbúvó" feltételesség mozzanata miatt. Von Wright szerint minden normának van valamilyen **alkalmazási feltétele** (condition of application), ami lényege szerint a norma tartalmától függ. A sokszor idézett példa szerint ha egy normál ajtóval szemben állunk, akkor kiadhatják nekünk azt a parancsot, hogy 'Csukd be az ajtót!', de ennek az a feltétele, hogy *(i)* az ajtó nyitva legyen, és *(ii)* ne csukódjon be magától. Ellenben ha egy magától bezáródó ajtóról van szó, akkor értelmetlen kiadni az előző parancsot, hiszen az világban végbemenő állapotváltozás eredményeként becsukódik az magától. Egy ilyen világállapotban más parancsnak lehet csak értelme, például annak, hogy 'Ne engedd becsukódni az ajtót!'.

Fontos hangsúlyozni, hogy minden normához tartozik a fenti értelemben vett alkalmazási feltétel, sőt, pontosabb lenne úgy fogalmazni, hogy tartoznia kell. De természetesen az alkalmazási feltételen kívül lehet még további feltételeket is a normákhoz rendelni, amelyek már nem szükségszerűen kapcsolódnak a normákhoz. Von Wright példája erre

a ‘Ha esik az eső, csukd be az ablakot!’ felszólítás. Egy ilyen feltétel esetlegesen van a normához rendelve – szemben a norma alkalmazási feltételével.

A kategorikus normák esetében az alkalmazási feltétel a norma cselekvéstartalma ismeretében is megállapíthatjuk, ezért mondja azt von Wright, hogy a kategorikus normák esetében elég az alkalmazási feltétel a tartalommal együtt adott [792, 74], ezért nincs szükség annak külön jelzésére a normát leíró logikai formalizmusban. Ezzel szemben a norma további feltételeit – ha vannak – meg kell tudnunk adni, ezért van szükség külön paraméter felvételére a normák összetevőkre bontása során.

A hipotetikus normák feltételei tetszőlegesek lehetnek, elméletileg bármely világállapot lehet egy norma feltétele. A cselekvéselmélet számára azonban kiemelten fontosak azok a feltételek, amelyek egy másik cselekvésre, még pontosabban: egy másik ágens cselekvésére hivatkoznak (ilyen például a ‘szemet szemért’ parancs). Ezek a normák a reciprocitás és komplementaritás társadalmi helyzeteiben kiváltképp gyakoriak és fontosak.

A norma alkalma

Vigyázni kell, hogy a norma alkalmazási körülményét, **alkalmát** (occasion of norm) ne keverjük össze a norma feltételével. Persze mondhatjuk azt, hogy a norma körülménye az a feltételrendszer, amely mellett a normát érvényesnek és követendőnek tartják az érintettek (vagy elgalább azok egy része), de ilyenkor is speciális feltételekre kell gondolnunk. A dolgokat kicsit leegyszerűsítve mondhatjuk, hogy a norma tér-idő meghatározottságáról van itt szó.

A körülmény azért is fontos összetevő, mert a fogalmi terjedelmébe tartozó adatok alapján tudjuk tipizálni a normákat a partikularitás/univerzalitás dimenziója mentén (legalábbis részben).

A norma autoritása

A norma mint a jövőbeni cselekvésekre vonatkozó szándékkinyilvánítás logikailag megköveteli, hogy legyen egy olyan ágens, aki a normát kibocsátja. Ezt a társadalmi szereplőt nevezzük a **norma autoritásának**. Vannak normák, amelyek esetében könnyen és egyértelműen meghatározhatjuk a norma kibocsátóját, de előfordulhatnak olyan esetek is, amikor igencsak nehéz megmondani, ki volt az adott norma autoritása. Az erkölcsi parancsok, a szokáserkölcs előírásai szinte mind ilyenek: a kibocsátójuk személye a történelem homályába vész.

Részben a norma autoritása alapján válik lehetővé, hogy az autonómia kérdését meg tudjuk fogalmilag ragadni, és így egy újabb normatipogizálási lehetőséghez jussunk.

A norma címzettje

A **norma címzettje** (subject of norm) az az ágens, akire vonatkozik a normában megtestesülő cselekvésbefolyásolási szándék, akinek a normában előírt cselekvástartalomhoz a norma modalitásértéke szerint kell (vagy lehet) viszonyulnia. Von Wright másfajta megfogalmazásában: a címzett hajtja végre azt a cselekvést, amelyre vonatkozóan az autoritás **toleráns** vagy **intoleráns** viszonyt nyilvánít ki.

A címzett lehet partikuláris, egyedi (particular), illetve univerzális, általános (general). Ez a megkülönböztetés később – részben – alapul szolgálhat a normák egyfajta

tipizálásához.

A normasértés szankciója

Von Wright, bár megemlítette, mint fontos, a normához kapcsolható jelenséget, de nem vette fel külön a szankcionálás jelenségét, a **társadalmi szankció** fogalmát a norma összetevői közé. Mégha el is fogadom, hogy a szankcióra nem a norma összetevőjeként kell tekinteni, ez nem jelenti, nem jelentheti azt, hogy ne kellene mindenképpen beemelnünk a fogalmat a norma modelljébe. Szankció hiányában nem beszélhetünk normáról, ezért szükség van rá a norma jellemzőinek tárgyalásakor.

Amikor von Wright a normák főbb típusait mutatta be, akkor normaként értelmezte a technikai normákat is, de én ezt az interpretációt nem fogadom el. Azon az alapon, hogy a technikai normák, a főzési receptek, a kezelési, szerelési útmutatók (például az IKEA-bútorok összeszerelését támogató vizuális segédletek), a használati leírások, szoftverkezelési kézikönyvek jelenségében nincs ott a társadalmi szankció mozzanata, ezért nem nevezhetjük ezeket társadalmi normának. Ezt Émile Durkheimhez is így gondolta, ahogy az kiderül a következő idézetből:

> „Ha ügyeimet nem a mesterség szabályainak megfelelően viszem, azt kockáztatom, hogy nem érek el sikert; de a társadalom nem emel szót az ellen, hogy így cselekedjek. Hagyja, hogy cselekvésem teljesen szabadon végbemenjen. Ez a cselekvés nem juthat el azokhoz a célokhoz, amelyek felé törekszik, de ezért még nem tiltatik meg." (Durkheimet idézi: [532])

Hiába van egyfajta "előírás", valamiféle mesterségbeli útmutató (technikai norma), és hiába nem követem az abban leírtakat, nem lesz rá társadalmi reakció. Ez nem jelenti azt, hogy ne lesz semmi következménye az útmutatótól való eltérésnek. Lehet, hiszen a technikai norma követése által remélt célt esetleg így nem érem el. De az eltérést ilyen esetekben nem követi szankció. Ha nem a recept szerint főzök, akkor ehetetlen lesz az étel, és ez következményként lehet igen kellemetlen, de társadalmi szankcióra nem kell számítanom. Nagyon fontos tehát, hogy a technikai előírások megsértéséből származó következményeket ne azonosítsuk, ne keverjük össze azzal, amikor a normasértést társadalmi szankció követi.

> „Ha megsértem a higiénia szabályait, amelyek azt parancsolják nekem, hogy tartózkodjam a gyanús érintkezésektől, e cselekedetet automatikusan követi a káros következmény, vagyis a betegség. Az elkövetett cselekmény magától váltja ki a következményt . . . " [224, 172]

Ezzel szemben a normasértés mindig valamilyen társadalmi helytelenítéssel találkozik, és a társadalom (közösség) vagy igyekszik megakadályozni a normasértőnek ígérkező cselekvést, vagy - ha ez már nem sikerült, akkor – valamilyen szankcióval sújtja a normától való eltérést.

> „Ha egy cselekvés, amelynek természeténél fogva egy morális szabályhoz kell alkalmazkodnia, attól eltér, a társadalom, ha erről tájékoztatva van, közbelép, hogy megakadályozza ezt az eltérést. . . . Ez a társadalmi reakció igazi szükségszerűséggel követi a normasértést; gyakran még módjában is előre meghatározott." (Durkheimet idézi: [532])

A normasértés helytelenítése és szankcionálása elválaszthatatlan a norma jelenségétől, de persze mindkét jellemzőt (a helytelenítést, illetve a szankcionálást) megfelelő módon kell értelmeznünk. Nem követelhetjük meg azt, hogy akár a helytelenítés, akár a szankcionálás konszenzuális legyen, vagyis mindenki egyetértsen, illetve részt vegyen ezekben. Azt sem állíthatjuk, hogy minden normasértést szankció követne, sőt, ki kell mondanunk egy fontos tételt: amilyen mértékben előzetesen várható a szankcionálás tényleges bekövetkezése, olyan mértékben fognak az emberek igazodni a normákba foglalt előírásokhoz. Az is fontos kérdés, hogy a szankcionálásnak milyen eszközei, lehetőségei vannak, de ezt a normák tipizálásakor fogom megválaszolni (és a típusokat bemutatni).

A norma érvényessége

A norma érvényessége alatt legitim érvényességet értek. Ezt a kérdéskörét az uralomról szóló fejezetben alaposabban tárgyalom (↬438.o.).

A normához való igazodás

A normavezérelt magatartásnak két fő típusa van: a normasértő vagy a normához igazodó cselekvés. A normasértésről már szóltam, most a normakonform magatartás jellemzésén van a sor. Előtte azonban tisztáznom kell valamit. A normához való igazodás nem feltétlen vonzza maga után a normakövető magatartást. Mind elméletileg, mind gyakorlatilag lényegi és szoros kapcsolatot tételezhetünk a normakövetés és a normához való igazodás között, de ez nem jelent kizárólagosságot. Csontos László a következő tényezőket sorolja fel a normakövetés magyarázata során:

> „A normakövető egyéni cselekvés magyarázata során - durván - négy lehetséges mechanizmusra hivatkozhatunk: (1) a normától eltérő vagy normasértő (deviáns) viselkedés szankcionálása; (2) az egyéni stratégiaválasztástól függő nyeremények pozitiv interdependenciája (a normakövető magatartás - pozitív - externális hatása); (3) az egyéni preferenciákba épített konformizmus; és végül (4) ex ante kommunikáció az ex post normakövető csoport tagjai között." [179]

A pontosabb értelmezés érdekében bemutatok egy hétköznapi példát arra vonatkozóan, hogy milyen cselekvési stratégiákat lehet követni a nagyvárosi tömegközlekedési eszközök igénybevétele során. Az alapnorma úgy szól, hogy minden használatért fizetni kell (bérlet- vagy jegyvásárlással), a normasértés esetén pedig büntetést kell fizetni. Ebben a helyzetben az alábbi stratégiákra könnyedén találhatnánk valós példákat az utazóközönség tagjai közt:

1.) valaki havi bérletet vált
2.) valaki minden utazási alkalommal előre jegyet vesz és érvényesíti
3.) valaki minden utazási alkalommal előre jegyet vesz, de csak akkor érvényesíti, ha ellenőrökkel találkozik
4.) valaki nem vesz sem jegyet, sem bérletet, de ha a járművön találkozik a kalauzzal, tőle vesz jegyet (ilyen rendszer működik a Szentendrei HÉV-en)
5.) valaki nem vesz sem jegyet, sem bérletet, ha ellenőrök elkapják, büntetést fizet

Az 1.) és 2.) esetben az a közös, hogy az ágens érvényesnek tartja a normát, és igazodik hozzá, ezért fizet a szolgáltatás igénybevételéért. A két eset közti kölönbség pedig abban van, hogy az utazási szokások ismerete alapján vagy a bérlet- vagy a jegyvásárlás kerül kevesebbe. Egyszerű kalkulációval kiszámolható, hogy hány utazás után éri meg jobban a bérletet megvásárolni. Ha valaki általában autóval jár, és nagy ritkán száll csak át tömegközlekedési eszközre, akkor megérheti az egyedi jegyvásárlás, különben nem nagyon.

Más a helyzet a 3.) eset ágensével, aki igazodik az alapnormához, de belsőleg nem tartja azt érvényesnek, csak a büntetéssel való külső fenyegetés tényleges létezése esetén cselekszik normakonform módon. Azzal, hogy az ellenőrök felbukkanása esetén érvényesíti a jegyét, öntudatlanul is "mutatja", hogy igazodik az alapnormához, de a normakonformitása nyilván csak olyan mértékű, amilyen mértékű az ellenőrzés gyakorisága. Amennyiben az ellenőrzések gyakorisága annyira alacsony, hogy az ilyenkor érvényesített jegyekért összesen kevesebbet kell fizetni havonta, mint amennyi a havi bérlet ára, akkor ez a "módszer" megérheti – csak racionális kalkuláció alaján.

A 4.) és 5.) esetben a normakonformitás nagyon gyenge, ha van egyáltalán. A 4.) eset annyiban különbözik a 3.) esettől, hogy az előbbiben az utazási szabályzat jogszerűvé teszi azt a magatartásmintát, ami az utóbbiban jogszerűtlen (hogy ti. valaki úgy szálljon fel a járműre, hogy előzetesen jegyet vegyen). A körülményekben bekövetkezett változás lehetővé teszi a teljes racionalizálást, vagy más oldalról tekintve a teljes "normátlanítást", mert ebben az esetben a normatív vonatkozásoktól teljesen eltekintve csak anyagi mérlegelés kérdése, hogy valakinek a bérlet vásárlása vagy a jegyek időnkénti kifizetése kerül-e többe. Az 5.) eset annyiban megegyezik a 4.) esettel, hogy mindkettőt a "csak racionális" mérlegelés dominálja. Az 5.) eset ágense kalkulálhat úgy magában, hogy ha statisztikai értelemben kifizetődőbb a ritkán előforduló büntetéseket kifizetni, mint folyamatosan bérleteket, jegyeket venni, akkor megteszi ezt a lépést akkor is, ha ennek a normasértő magatartást övező társadalmi helytelenítés az ára (esetleg ezeket a pszichológiai költségeket is beárazhatja).

Láthatjuk tehát, hogy a racionális mérlegelésen alapuló érdekkövetés és a – norma érvényességének belső elfogadása esetén – nem mérlegelésen alapuló normakövetés, mint két szélső pólus között számtalan átmeneti, kevert eset is elképzelhető.

még további értékelés kell ide .....

Amíg az ember nem gondol bele, talán paradoxonnak tűnhet azt állítani, hogy a norma társadalmi elfogadottsága, követése pozitív kapcsolatban állhat a normasértés jelenségével. Pedig Durkheim szerint a normasértés indirekt módon igenis hozzájárulhat a normakövetés erősödéséhez és ezáltal a társadalmi rend – legalább részleges – megszilárdulásához:

> „A kollektív érzelmek ugyanis, amelyeket a történelem adott időszakában egy nép büntetőjoga védelmez, csak akkor hatolhatnak be azokba az emlékekbe, amelyek eladdig zárva voltak előttük, s csak akkor tehetnek szert akkora hatalomra, amellyel addig nem rendelkeztek, ha a korábbinál nagyobb lesz az intenzitásuk. Az egész közösségnek erősebben kell éreznie ezeket az érzelmeket. Csakis innen meríthetnek ugyanis ezek az érzelmek olyan, a korábbbinál nagyobb erőt, amelynek révén

> rákényszeríthetik magukat azokra az egyénekre is, akik ezt megelőzően leginkább ellenük szegültek. A gyilkosok csak akkor tűnhetnek el, ha a vérontástól való irtózás a korábbinál jóval erősebbé válik azokban a társadalmi rétegekben is, ahonnét a gyilkosok rendszerint kikerülnek." [225, 87]

Rögtön ezek után persze Durkheim hozzáteszi azt is, hogy a kollektív érzelmek megerősödése, amely – jó esetben – valamely bűntípus visszaszorulását, esetleg eltűnését eredményezi, "hozzájárul" ahhoz, hogy más területeken újfajta bűncselekmények "jelenjenek meg". Durheim persze ennek is megadja a magyarázatát:

> „ . . . a közös tudat ilyen intenzív állapotai nem erősödhetnek meg anélkül, hogy ne erősítenék fel azokat a gyengébb állapotokat, amelyek áthágása addig csupán erkölcsi vétségnek számított. . . . [a kollektív] tudat sokkal érzékenyebben reagál majd azokra a sérelmekre, amelyek korábban alig érintették: erőteljesebben lép fel ellenük, jobban elítéli őket, s némelyiküket egyszerű erkölcsi fogyatékosságból bűncselekménnyé változtatja." [225, 88]

Ezen a ponton kiváltképp "tetten érhető" a normavezérelt magatartás mélyén kitapintható, elemi erővel hatni képes értékalapja, érzelmi töltete.

mmm

Rorbert B. Cialdini és szerzőtársai megkülönböztetnek kétféle kategóriát, a **leíró norma** (descriptive norm) és **rendelkező norma** (injunctive norm) fogalmát[158]. Az általam javasolt modellben Cialdiniék leíró normája nem igazán norma, hiszen nincs mögötte szankcionáló erő, de a szociálpszichológiában meglehetősen népszerű vált nézetet nem hagyhatom teljesen figyelmen kívül.

ők elsősorban arra gondolnak, hogy amit sokan követnek, azt mindenkinek érdmes lesz követni. Ezt nevezzük hálózati hatásnak. Ezen nincs mit vitatni. A nagy kérdés az, hogy ezt a jelenséget a norma fogalma alá kell-e, szabad-e sorolnunk.

Lényegét tekintve nem, de azért figyelembe kell venni.

———— valahova betenni: ————-

Egy dolog, hogy a normát kifejezhetjük előíró, preskriptív és leíró, deskriptív formában, és egy másik dolog az, hogy ugyanazt a formát olykor felfoghatjuk kétféle értelemben is. A "Szemet szemért!" elv sűrített formájú előírás, nem lehetne azt mondani rrá, hogy kijelentés lenne, de a preskriptív jelleget ki lehet érezni belőle. Biztos leíró formája van annak a másik híres idézetnek, amely így szól: "Ki kardot ránt, kard által vész el." Ez a mondás elsődleges értelmében egyszerűen csak megállapítja, hogy a "Szemet szemért!" elv alkalmazása megtorlási spirálba taszítja, taszíthatja a résztvevőket. Lehet azonban egy olyan értelemzést is adni neki, hogy az előző az utóbbi egyfajta tiltása.

A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places: [158]

The Transsituational Influence od Social Norms: [599]

Descriptive social norms as underappreciated sources of social control [156]

Managing social norms for persuasive impact: [157]

The Constructive, Destructive, and Reconstructive Power of Social Norms: [646]

An Explication of Social Norms [431]

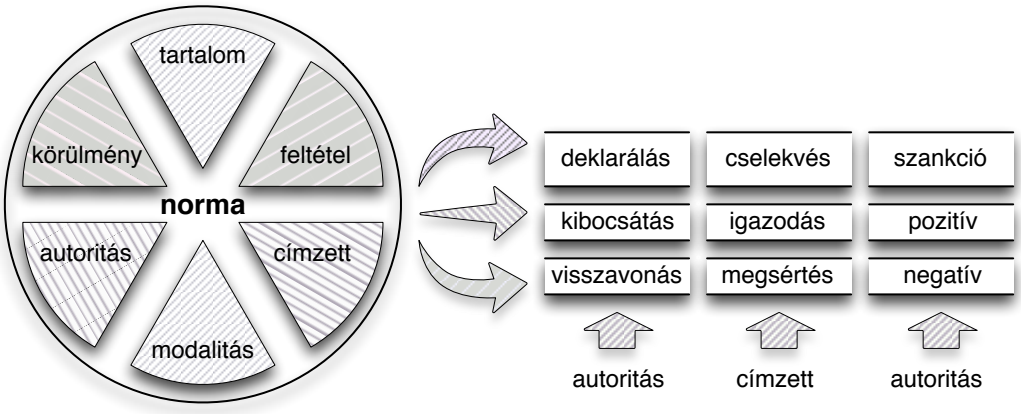
A norma másodrendű kontextusa

A von wright-i normaösszetevők kiegészítését akkor érthetjük meg igazán, ha látjuk (és a modellünkben feltüntetjük), hogy a **normavezérelt cselekvések** sosem magukban állnak, így nem is értelmezhetők magukban. Az elemzés központjába természetesen a normavezérelt cselekvést kell helyeznünk, de emellett figyelembe kell vennünk még két másik cselekvéstípust is. A normavezérelt cselekvést a norma címzettje hajtja végre. Ennek során vagy igazodik a normába foglalt szándékhoz vagy megsérti a normatív elvárást.

A normához való igazodáshoz azonban meg kell előzze az a cselekvés, amelynek során a normát magát kibocsátja valaki.[3] Ha viszont azt a szempontot is figyelembe vesszük, hogy a normákat nemcsak **kibocsátani** lehet, de megszüntetni vagy **visszavonni** is, akkor szükség van egy olyan általánosabb fogalomra, amely mind a kibocsátás, mind a visszavonás aktusát magába foglalja. Erre alkalmas lehet a **deklarálás** fogalma. A deklarálás tevékenységét az autoritáshoz köthetjük, és fontos kényszerként rögzíthetjük, hogy ennek a cselekvésnek időben meg kell előznie a normavezérelt cselekvést magát. Mivel pedig a norma deklarálása olyan cselekvés, amely más cselekvést magukba foglaló normákra vonatkozik, így logikai minőségét tekintve a deklarálást másodrendű cselekvésfogalomnak kell tekintenünk.

A normavezérelt cselekvéshez tartozik még a szankcionálás aktusa is, ami a cselekvést követően, tehát ahhoz képest időben később valósul meg, de fogalmilag ugyanúgy elválaszthatatlanul kapcsolódik hozzá, mint a a kibocsátás cselekvése. A szankcionálás is másodrendű, hiszen mindig csak valamilyen elsőrendű (alacsonyabbrendű) cselekvésre hivatkozva értelmezhető, önmagában vett szankcióról sosem beszélhetünk. A szankció lehet pozitív, valamilyen jutalom, illetve negatív, valamilyen büntetés. A szankcionálást az autoritás végzi el, de fontos látni, hogy az autoritás szerepfogalom, és nem feltétlen egyezik meg az a két ágens, amely egy adott normával kapcsolatos kibocsátó, illetve szankcionáló szerepeket viszi.

A normavezérelt cselekvés másodrendű kontextusán a fent bemutatott másodrendű cselekvéseket értem. A normatív jelenségek értelmezéséhez a norma összetevői mellé fel kell venni ezeket a jelemzőket is.



hiányzik az érvényesség problémaköre, az ábrát is ki kellene egészíteni

[3] Itt most újra eltekintek attól a problémás kérdéstől, hogy ki bocsátja ki a morális normákat.

nincs megoldva a felhatalmazottsággal kapcsolatos érvénytelenség problémája, azt is ide kell majd vezetni

*

12.1.2 A norma típusai

A norma összetevői és egyéb jellemzői alapján már módunk nyílik arra, hogy a normákat különböző szempontok szerint tipologizáljuk.

Normatípusok a szankció típusa szerint

A normavezérelt cselekvés következménye valamilyen joghatás lesz. Joghatás alatt érthetjük a szankció bekövetkeztét is, nem részletezve azt, mit is érünk szankció alatt. Még azt is mondhatnánk, hogy a szankció lehet pozitív vagy negatív, tehát beszélhetünk jutalomról és büntetésről. A tényleges társadalmi gyakorlatok ismeretében azonban nem tévedünk nagyot, ha a szankció fogalma alatt itt csak a büntetést értjük, és a talán legfontosabb három normatípust a rájuk jellemző szankciók különbsége alapján határozzuk meg. Ezt a hármas tipológiát Max Webertől vehetjük át, akitől most egy kicsit hosszabban idézek:

> „6.§. Valamely rend legitimitása biztosítható:
> I. tisztán belülről, mégpedig
> 1. tisztán indulati-érzelmi alapon: érzelmi átéléssel, belső odaadással;
> 2. értékracionálisan: a végső és kötelező értékek (erkölcsi, esztétikai vagy bármely más értékek) kifejeződéseként fölfogott rend abszolút érvényességébe vetett hittel;
> 3. vallásos hittel: azzal a meggyőződéssel, hogy a rend betartásán keresztül el lehet jutni az üdvözüléshez;
> II. továbbá úgy is (vagy csak úgy), hogy sajátos külső következményekhez kapcsolódnak bizonyos várakozások.
> A rendet
> a) konvenciónak nevezzük, ha érvényessége külsődlegesen, azáltal van biztosítva, hogy az emberek megadható körén belül a rendtől való eltérés várhatóan (viszonylag) általános és gyakorlatilag érezhető helytelenítésbe ütközik;
> b) jognak nevezzük, ha ugyancsak külsődlegesen, de azáltal van biztosítva, hogy az emberek egy külön erre fölkészült csoportja várhatóan (fizikai vagy pszichikai) kényszert alkalmaz a rend betartásának kikényszerítésére, illetve megszegésének megtorlására.” [810, 60-61]

Ahogy az idézetből kiderül, Weber először elkülöníti egymástól azokat a normatív cselekvéseket, amelyek során a cselekvők belső, illetve külső tényezők miatt igazodnak a normához. A belső normakövetéshez rendeli hozzá az erkölcs kategóriáját. A külsődleges hatásokra történő normakövetés során a cselekvő társadalmi környezete, a közösség valamilyen befolyása érvényesül. Az erkölcsi normát megsértő cselekvés belső folyamatokat indít el az emberben (önmarcangolás, lelkiismeret-furdalás). A konvenció és a jog megsértése külsődleges büntetéssel jár. A konvenció esetében normasértéskor a helytelenítésen keresztül közösségi nyomás nehezedik az egyénre (aminek legszélsőségesebb esetében a közösség kizárja soraiból a normát megsértő egyént), a jogsértés pedig az

adott közösség lelki vagy fizikai erőszakon alapuló szankcióit vonja maga után. Utóbbi általában úgy valósul meg a mindennapokban, hogy a közösség vezetése a börtönbe záráshoz szükséges fizikai erőszak-alkalmazás lehetőségére hivatkozik egy kifejezetten erre szakosodott apparátus segítségével – amely egyébként a folyamat végén el is végzi az olykor valóban szükséges erőszak-alkalmazást.

A normasértésekhez rendelhető szankciók két szempont figyelembevételével írhatók le. Egyfelől tudni kell, hogy ki szabja ki a szankciót a normasértőre, másfelől azt kell számításba venni, hogy milyen szankciót lehet kiszabni. Az első szempont szerint külső vagy belső szankcionálásról beszélhetünk. A második szempont szerint normasértés esetén háromféle lehet a büntetés: az egyénben ébredő lelkiismeret-furdalás, vagy a normafenntartó közösség által alkalmazott lelki, illetve fizikai kényszer (vagy az azzal való fenyegetés). Mindezek alapján megkapjuk Weber három alapvető normatípusát:

- **erkölcs** (belső szankció: az egyén lelkiismeret-furdalása)
- **konvenció** (külső szankció: a közösség helytelenítése, kritikája)
- **jog** (külső szankció: az erőszak-alkalmazással való fenyegetés)

A három szankciótípus és persze a három normatípus mindegyike vonatkozhat ugyanarra a cselekvésre, ezért minden egyes konkrét társadalmi kapcsolat elemzésekor meg kell mondani, hogy éppen akkor melyik normatípus érvényesül a leginkább. Előfordulhat ugyanis, hogy egy társadalmi helyzeten belül egyidejűleg többféle normatípus áll fenn vagy érvényesül. Betarthatom én a sebességkorlátot azért, mert *(i)* látom az út szélén a rendőrt (tehát a joghoz igazodom), *(ii)* a feleségem szúrós szemekkel tekinget felém (konvencionális nyomást fejt ki rám), illetve *(iii)* sajátmagam belső késztetésére (rosszul érezném magam, ha normát sértenék).

Szükségesnek tartok pár további megjegyzést. Először is fontos rögzíteni, hogy Weber konvenciófogalma jelentős mértékben eltér attól az értelmezéstől, ami David Lewis 1969-ben megjelent könyve óta elterjedt a játékelmélettel foglakozó kutatók körében [444]. Lewis a koordinációs helyzetekben jelentkező teljesen vagy közel egyenrangú döntési alternatívák közti kollektív választás egyik eszközének tekinti a konvenciót. A koordinációs helyzetek döntetlen jellegét oldja fel az, hogy valamilyen módon (szaliencia-, fókuszpont- vagy precedenstechnikákra támaszkodva) konvenciók alakulnak ki egy közösségen belül.[4]

A másik megjegyzésem Weber erkölcsfogalmára vonatkozik. Először is rögzítem, hogy a könyvemben egymás szinonimájaként kezelem az erkölcs, az etika és a morál terminusokat. Akármelyik terminust is használjuk, azt is látni kell, hogy a weberi definíció se nem igazán reflektált, se nem igazán hasonlít a megszokott, széles körben elterjedt erkölcsfelfogáshoz. Reflektálatlanságon azt értem, hogy ritkán azonosítják erkölcsként a fenti weberi idézet I. pontjában leírtakat, de szokatlan az a definíció is, amely az erkölcsöt a belső lelkiismerettel támogatott normával azonosítja.[5] Pedig maga Weber erről a hármasságról ír könyvének más oldalán.

> „A »kívülről« biztosított rendnek azért még »belső« garanciái lehetnek. A szociológia számára nem probléma a jog, a konvenció és az »etika« kapcsolata. Az »etikai« mérce itt azt jelenti, hogy az emberek bizonyos speciális értékracionális hite normaként szolgál az olyan emberi cselekvés számára, amely az 'erkölcsi jó' predikátumára tart igényt . . . " [810, 62]

[4] A david lewis-i értelemben vett konveciófelfogásról ír magyarul Pete Krisztián[569].

[5] Az erkölcsről magyarul olvasható Orthmayr Imre tanulmánya [554].

Ebben a gondolatmenetben már egy szinten, egyenrangúan (és kizárólagos módon) szerepel a rend három típusa, a jog, a konvenció és az etika, csak az nincs megfogalmazva, hogy az etikához (erkölcshöz) kapcsolható speciális szankció a lelkiismeretfurdalás.

Machine ethics: Creating an ethical intelligent agent [19]

Normatípusok a logikai rend szerint

A norma tartalma és szankciója alapján **elsőrendű** vagy **magasabbrendű normákat** különíthetünk el egymástól. Elméletileg tetszőlegesen magas rendű normát el lehet képzelni, de a gyakorlatban vélhetőleg harmadrendű normáknál nem nagyon találunk magasabbrendűeket. A szakirodalomban az első- és másodrendű normákkal sokat foglalkoznak, harmadrendűekkel már jóval kevesebbet, és arról nem tudok, hogy a harmadrend fölött bármilyen elemzés született volna. Pedig azt gyanítom, hogy egy alaposabb elemzés kideríthetné, hogy – főleg, mióta az európai uniós jogkörnyezetbe kerültünk – előfordulnak negyed-, netán ötödrendű normák tényleges joggyakorlatban, csak ez – már, ha tényleg létezik – azért nem okoz gondot, mert az adott kontextusban, a tényleges jogi környezetben "közelről" nem lehet érzékelni a magasabb rendbe való kerülést. De nem is az az igazán fontos kérdés, hogy milyen másodrendnél magasabb normák vannak (egyáltalán vannak-e), hanem sokkal inkább azt kell látnunk, hogy hol, milyen másodrendű normákat találhatunk.

Von Wright a normáról szóló könyve egyik fejezetét a magasabbrendű normáknak szenteli [792, 189-207]. Itt vezeti be a kompetencianorma fogalmát, amire már utaltam a hohfeldi elmélet elemzésekor (↬415.o.). A fogalom definíciójából következik, hogy a kompetencianormát (a hohfeldi felhatalmazottságot) delegáló normáknak is tekinthetjük, ami egyfajta normahierarchia kialakítására és kezelésére lehet alkalmas. Jog esetében az állam az alárendelt hatóságok számára átengedhet felhatalmazottságot, hogy így a közjót jobban ki tudja szolgálni.

> „Előfordul, hogy a legfelsőbb hatóság bizonyos dolgokat szabályozni akar, mondjuk úgy, a közjó érdekében. Ilyenkor a jogszabályi részletek kidolgozását olyan alacsonyabb szintű hatóságra hagyhatja, amely megfelelő ismeretekkel rendelkezik a cél, az adott közjó megvalósításához szükséges speciális követelményekről." [792, 193]

Persze, von Wright nem volt se jogász, se szociológus, se antropológus, így a társadalmi cselekvések elemzésével kapcsolatban nem tudott, nem is akart továbblépni. A magasabbrendű normatív fogalmakat a jogviszonyok, jogrendszer elemzése kapcsán először jogászok bontották ki alaposabban. Már Hans Kelsen is beszélt első- és másodrendű jogszabályokról, de ő nem logikai értelmezés szerint sorolta ilyen típusokba a jogszabályokat, hanem a jogrendszer működési logikája szerinti fontosságot tekintette meghatározónak [395]. Ezért ő a logikai interpretációhoz képest pont fordítva minősítette a jogszabályokat, és a bírókat az ítélkezési folyamat során kötelező szabályokat nevezte elsőrendűnek, míg az egyszerű állampolgároknak, a jog alanyainak tetszőleges, mindennapi tevékenységét szabályozó normákat tekintette másodrendűeknek. Ez a megközelítés pont az ellentéte a logikai-fogalmi gondolkodásnak. Természetesen itt csak eltérő terminológiai választásról van szó, fogalmi hibáról nem beszélhetünk.

Amikor H.L.A. Hart elsődleges és másodlagos szabályokról írt, akkor ő már a szokásos logikai interpretációnak megfelelő szemléletben osztotta fel a jogszabályokat [346]. A

három híres másodrendű szabálya közül az **elismerési szabály** abban az értelemben különbözik a másik kettőtől, az **ítélkezési** és a **változtatási szabályoktól**, hogy az előbbi a jog egészének érvényességével kapcsolatos, míg az utóbbi kettő a jogrendszer fontos szereplőinek tevékenységére vonatkozik. A bírónak, rendőrnek másodrendű kötelessége van a szankciókat kiosztani a normasértések esetén, az ítélkezési szabályok a rájuk vonatkozó jogviszonyokat szabályozzák. A törvényhozónak, a jogalkotónak magasabbrendű joga (felhatalmazottsága) van az elsőrendű (olykor a másodrendű) jogviszonyokat megváltoztatni.

A felhatalmazottságról szóló fejezetben már jeleztem, hogy a jog elemzése során kialakított jogviszonyfogalmak, az azokkal kapcsolatos tanulságok általánosíthatók és kiterjeszthetők az általánosabb normafogalmakra is. A legkülönfélébb társadalmi intézményekben, a családban, a munkahelyeken, az iskolákban, a társadalmi, kulturális, politikai szervezetekben mindenhol találhatunk elsőrendű szabályokat, normákat, amelyek a résztvevők mindennapi cselekvéseit szabályozzák, és vannak másodrendű normák, szabályok, amelyek az éppen aktuális normarend megváltoztatásának módját határozzák meg. Ezek lehetnek konvenciók, lehetnek szervezeti szabályzatok mint gyenge jogszabályok, de lehetnek akár morális szabályok is. A másodrendű szabályoknak pedig minden esetben, minden területen ugyanazok a funkciói, mint amit már a jog esetében korábban megismertünk: elismerni (eredeztetni), változatni, ítélkezni (fenntartani).

Azt jeleztem ugyan, hogy léteznek a másodrendű szabályoknál magasabbrendűek is, de eddig igazából csak a másodrendű normákról beszéltem. Bár a másodrendű szint alapvető fontosságát (elsődlegességét) nem akarom kétségbe vonni, szükségesnek tartom jelezni, hogy minden normatív rendszerben vannak, kellenek, hogy legyenek harmadrendű szabályok. Ezért bemutatok egy egyszerű, hétköznapi példát a harmadrendű norma létezésére, működésére. Egy családnál gyerekzsúrt tartanak, és a kisgyerekeket egyedül hagyják egész délután a nagyszobában játszani. Este, miután minden vendég elment, az anyuka észreveszi, hogy a csokitorta egy részét valaki belepaskolta a drága perzsaszőnyegbe. Amikor kérdőre vonja ezért a kisfiát, a gyerek azzal vág vissza, hogy nem ő volt, aki ezt csinálta, hanem Zsoltika. Az első megszólalásával az anyuka másodrendű cselekvést hajt végre, hiszen elsőre szóbeli feddésben részesíti vagyis szankcionálja a fiát. Amikor azonban kiderül, hogy az elsőrendű cselekvés szintjén nem a fia volt a rendetlen (a normasértő), akkor az anyuka visszakérdez: 'Miért nem szóltál rá Zsoltikára, hogy nem szabad ilyet csinálni?' Ez a feddő kérdés már harmadrendű cselekvésnek számít, hiszen az anya feddése annak a mulasztásnak szól, hogy a fia nem szankcionálta (másodrendű cselekvés) a vendégét, aki rosszalkodott (elsőrendű cselekvés).

Normatípusok a deontikus modalitás szerint

A norma modalitása alapján is adott egy természetes tipizálási lehetősége, hiszen a deontikus logikai kategóriák alapján a normákat élesen elkülöníthetjük egymástól. Attól függően, hogy milyen deontikus logikai operátorokat veszünk fel a deontikus logikai modellünkbe, beszélhetünk három, négy vagy hat modalitás szerinti normatípusról (úgymint: **kötelező, megengedő, tiltó, mellőzhető, opcionális vagy preskriptív normákról**).

Ahhoz képest, hogy eldöntjük, hányféle deontikus modalitással akarjuk kezelni a normatív teret, illetve melyek ezek, az már egy másik kérdés, hogy ezeket a deontikus operátorokat fogalmilag, logikailag hogyan írjuk le, és egy formális ontológiába miket

hogyan veszünk fel a természetes nyelvi szinten használt kifejezésekből. Von Wright a kötelezést és a tiltást egymásból levezethetőnek tartja, ezért elegendőnek tartja csak az egyiket felvenni a modelljébe [792, 71]. Ezt a kötelező normával (O, O-NORM teszi meg. A megengedés aktusát viszont – jelezve annak vitatott státusát – önállóan kezelendőnek ítéli, és ezért külön felveszi saját ontológiájába a permisszív normát (P, P-NORM). Mások szerint a logikai szinten elegendő egyetlen deontikus operátor, és minden más egyéb levezethető belőle (bővebben erről (↬**??**.o.).

A normák deontikus modalitás szerinti tipizálásában érdemes Hohfeld elméletét, fogalmait is figyelembe venni. Bár a hohfeldi másodrendű jogfogalmak deontikus logikai minőségüket nem jelentenek új típusokat, gyakorlati jelentőségük miatt érdemes felvenni őket a normák deontikus modalitás szerinti osztályozása során. Eközben persze azt is figyelembe kell venünk, hogy a két másodrendű hohfeldi fogalmat, a felhatalmazottságot és a mentességet nemcsak itt használhatjuk önálló altípusként, de a normák logikai rend szerint csoportosításakor is.

A klasszifikációhoz ugyan nem járul hozzá a Hohfeld által megállapított jogok és kötelezettségek közti komplementaritás figyelembe vétele, hiszen a hohfeldi korrelativitás csak a normatípus közti viszonyokat ragadja meg, nem ad új alosztályt. De a korrelatív kapcsolatban levő normaviszonyok mégiscsak a deontikus modalitás szerinti normatípusokat kapcsolják össze egymással, így a korrelativitás "követése" révén további tudást lehetséges a normatív viszonyok

Normatípusok a tartalom szerint

A logikai elemzésekben nemigen találhatunk olyan kísérleteket, amelyek a norma tartalma alapján tipizálná a normákat. Vélhetőleg azért nem, mert az ilyen vizsgálat kivezet a logika területéről. Ebből a szempontból tekintve nyilván sokféle tipizálás elképzelhető. A legtermékenyebbnek a játékelmélet játékosztályai szerint elvégezhető tipizálási lehetőségeket érzem, amit nagy vonalakban be is mutattam a játékelméleti fejezetben (↬336.o.), ezért itt csak vázlatosan idézem fel az ott leírtakat.

Robinson és Goforth a 144 kétszemélyes alapjátékot azok "geometriai jellemzői" alapján osztályozva ??????? játékosztályt definiált [608]. ?????? Ulmann-Margalit az alapjátszmákat három nagy osztályba sorolva be elkülönítette a **koordinációs normákat** (coordination norms), a **fogolydilemma normákat** (Prisoners' Dilemma norms) és a **részrehajló normákat** (inequality or partiality norms) egymástól [762].

Normatípusok az érvényesség alapján

A normákat – a jogi normákat biztosan, de a konvencionális és morális normákat is, mégha kicsit nehezebben – tipizálhatjuk az érvényesség, pontosabban a hatályba lépés szempontja szerint. Egy norma lehet **hatályos**, illetve **hatályon kívüli**. Hatályos norma esetén az autoritás (a kibocsátói vagy a szankcionáló szerepben) a legitim érvényességre hivatkozva igényli a normához való igazodást a címzettektől, míg a hatályon kívüli norma esetében nincs ilyen igény.

Normatípusok az autoritás alapján

Fogalmilag már kivezet a módszertani individualizmus világából, ha az autoritástípusok alapján próbáljuk meg tipizálni a normákat, ám a komoly szociológiai relevanciája miatt mindenképpen be kell mutatnom ezt a klasszifikációs sémát. Ha a széles értelemben vett jogra gondolunk, arra a tételes normarendre, amit emberek hoznak saját életük szabályozására ideértve mind az egységes és centralizált állami jogrendszer, mind az osztottabb, korlátozottabb hatókörű, ám nagyobb számosságú szervezeti világ működési szabályait, akkor a normák kibocsátói szerint elkülöníthetjük az alábbi normatípusokat: **törvény, rendelet, ítélet, szabályzat**. Az elkülönítés alapja az, hogy milyen típusú szerv a norma kibocsátója. Ha a törvényhozó szerv, akkor a norma **törvény**, ha államigazgatási hatóság, valamilyen végrehajtó szerv az autoritás:, akkor a norma **rendelet**, ha bírósági szerv az autoritás, akkor a norma bírósági **ítélet**, és ha az autoritás nem egy államigazgatási szerv, hanem valamilyen gazdasági, társadalmi, politikai, de államtól független szervezet, akkor a norma szervezeti és működési **szabályzat**.

Normatípusok a címzett és autoritás viszonya alapján

A norma címzettje és autoritása az a két társadalmi szerep, amely mentén az ágenseket a normákhoz kapcsolhatjuk. E két szerep egymáshoz való viszonya alapján egy olyan felosztáshoz juthatunk, amely egy jól ismert és nagyon széles körben használt kettősséget eredményez, jelesül az **autonóm** (AUTONOMOUS NORM), illetve a **heteronóm norma** (HETERONOMOUS NORM) dichotómiáját [792, 76]. Ha a norma kibocsátója, tehát az autoritás és a norma címzettje megegyezik egymással, akkor beszélhetünk autonómiáról, autonóm normáról, ellenkező esetben pedig a norma heteronóm.

Ez a felosztás jól alkalmazható a jog, a racionális uralom világában, ahol sok helyen megtalálhatjuk mind az autonómia, mind a heteronómia jelenségét (ehhez presze a jogot általánosabb értelemben kell interpretálnunk, mint az állam által teremtett normatív rendszert). Nem ilyen egyértelmű a helyzet a konvenció és a morál területén, hiszen ilyen esetekben nem mindig lehet (vagy szinte sosem lehet) megmondani, hogy ki a norma kibocsátója. Sokszor nem is beszélhetünk kibocsátóról, kibocsátásról. Von Wright ezt pont ellentétesen látja. Szerinte a morál lehet autonóm, a preskripció (durván: a jog) nem igazán. Ezt azon az alapon állítja, hogy a morálnak (és a konvenciónak, bár ezt ő nem említi) nem ismert a kibocsátója, és így az autonómia azt jelenti, hogy nem valaki mástól (nem egy konkrét ágenstől) származik a norma. Ezzel szemben von Wright nem tartja lehetségesnek, hogy valaki önmagának bocsásson ki egy normát, vagyis a preskripció szerinte nem lehet autonóm – legfeljebb csak valami másodlagos, analógiás, átvitt értelemben [792, 77]. Ez a von wright-i ellenvetés hamar súlyát veszti, ha az egyéni cselekvők szintjéről egy kicsit elmozdulunk a kollektív cselekvések irányába. Mégha egyet is értenénk von Wright nézetével, a szervezetek világában már egészen biztosan értelmesnek, használhatónak kellene gondolnunk az autonóm és heteronóm minősítéseket.

Normatípusok a címzett és az alkalom szerint

A norma összetevői között a címzett, az alkalom és a feltétel az a három elem, amelyben egyedi, partikuláris dolgokat is kifejezhetünk. A másik három öszetevőben általános, generális információt adhatunk meg. A normák egyedi, partikuláris minőségét értelem-

szerűen csak az ennek kifejezésére képes összetevők segítségével ragadhatjuk meg, míg a normák általános, univerzális minőségéhez reprezentálására bármely normaösszetevőt használhatjuk. Mivel nem minden normához tartozik feltétel, ezért egy norma partikularitását annak címzettjén, illetve alkalmán keresztül jelezhetjük. Ezt teszi von Wright is, amikor a normákat a partikuláris-univerzális dimenzió mentén tipizálja [792, 77-81]. Igazodva az általa javasolt terminológiához elkülöníthetjük a partikuláris, **egyedi normákat** (PARTICULAR NORM) és az univerzális, **általános normákat** (GENERAL NORM) egymástól.

Egy norma esetében mind a címzettre, mind az alkalomra teljesen egyedi, teljesen konkrét adatokat megadhatunk. Mondhatjuk azt, hogy egy vagy több konkrét emberre vonatkozik a norma, de egyedi tér-idő adatok megadásával megszabhatjuk a különös vagyis egyedi alkalmazási körülményeket is. Eszerint a norma címzettje és alkalma lehet partikuláris, egyedi, illetve generális, általános. Ezek alapján definiálhatunk új normatípusokat. Az **egyedi normákhoz** (PARTICULAR NORM) akkor jutunk, ha mind a norma címzettje, mind az alkalma partikuláris, egyedi. Ezzel szemben egy **általános normát** (GENERAL NORM) úgy adhatunk meg, hogy vagy a címzett vagy az alkalom (vagy mindkettő) általános [792, 81]. Ehhez hozzáteszi még von Wright azt is, hogyha mindkét összetevő általános, akkor a normát **teljesen általánosnak** (EMINENTLY GENERAL NORM) lehet nevezni.

Normatípusok a feltételesség alapján

A norma feltételének fogalma a normák természetes tipizálásra is módot ad [792, 74-75]. Von Wright elkülöníti egymástól a **kategorikus** (CATEGORICAL NORM) és a **hipotetikus normákat** (HYPOTHETICAL NORM). Ezt azon az alapon teszi meg, hogy van-e a természetes alkalmazási feltételeken kívül további feltétel a normához rendelve. Ha nincs, akkor beszélhetünk kategorikus normákról, ha van, akkor hipotetikus normákról van szó.

A kategorikus és hipotetikus norma bemutatásakor már jeleztem, nagyon vigyázni kell arra, hogy ne keverjük össze egymással *(i)* a norma kategorikus feltételeit, *(ii)* a hipotetikus norma "feltételességébe tartozó" feltételeket, valamint *(iii)* a norma alkalma, alkalmazási körülményei fogalma alá sorolható feltételeket. További nehézséget okozhat az a további körülmény, hogy *(iv)* a normához tartozó szankciót is kondicionális logikai kapcsolóval, azaz egy feltétellel kötjük össze a normasértés tényét rögzítő propozícióval. A 'feltétel' terminust tehát legalább négyféle értelemben alkalmazhatjuk, és ebben a helyzetben könnyen lehet tévedni. Jakab András a norma szerkezetének vizsgálatáról szóló tanulmányában azt nyilván helyesen állapítja meg [373], hogy a tényállás és jogkövetkezmény között van egy kondicionális kapcsoló, amit lehet többféleképpen is ki lehet fejezni (és ezáltal kategorikus vagy hipotetikus formában reprezentálni), de a von wright-i értelemben vett hipotetikusság-kategorikusság csak a tényállásra, még pontosabban arra a diszpozícióra vonatkozik, amely a cselekvésre vonatkozó előírást tartalmazza. Jakab állításával ellentétben tehát továbbra van értelme hipotetikus normáról beszélni. A hipotetikusság értelmességére (és szükségességére) vonatkozóan bemutatok egy ismert példát, hogy rajta keresztül érzékeltethessem a hipotetikus és kategorikus normák közti különbséget. Vegyük az alábbi bibliai példamondatokat (normákat):

1.) Ha megdobnak kővel, dobd vissza kenyérrel!

2.) Szemet szemért, fogat fogért!

Akárhogy is akarnánk átformálni ezeket a példamondatokat, csak hipotetikus normákként értelmezhetnénk csak őket. A bennük megtalálható hipotetikusság megszüntethetetlen-

sége pedig akkor válik nyilvánvaló, ha átalakítjuk őket kategorikus normákká. Ekkor, tehát kategórikus normaként, így néznének ki:

1'.) Dobd meg kenyérrel!

2'.) Szúrd ki a szemét, üsd ki a fogát!

Mindenki látnia kell a különbséget 1.) és 1'.). illetve 2.) és 2'.) mondatok között, nem szorul különösebb magyarázatra a kétféle megnyilatkozás szembeállítása. A hipotetikus normáknak igen nagy szerepe van a társadalmi érintkezésben. Említettem már, hogy a reciprocitás jelenségéhez kapcsolható cselekvési stratégiák, viselkedési szabályok a legtöbb esetben hipotetikus normaként (vagy logikai struktúráját tekintve annak megfelelő érdekvezérelt szabályként) vannak megfogalmazva.

A normatípusok áttekintése

A különböző tulajdonságok, összetevők mentén megalkotott normatípusokat egyetlen áttekintő ábrában bemutatva láthatóvá válik, hogy az egyes szempontok, dimenziók szerint milyen típusokat különíthetünk el egymástól. Ha pedig arra is gondolunk, hogy ezeket a különböző típusképző szempontokat egyszerre is érvényesíthetjük, tehát további altípusokat kaphatunk, akkor látszik, hogy egy igencsak kiterjedt, sokszempontú normatipológia áll rendelkezésünkre.

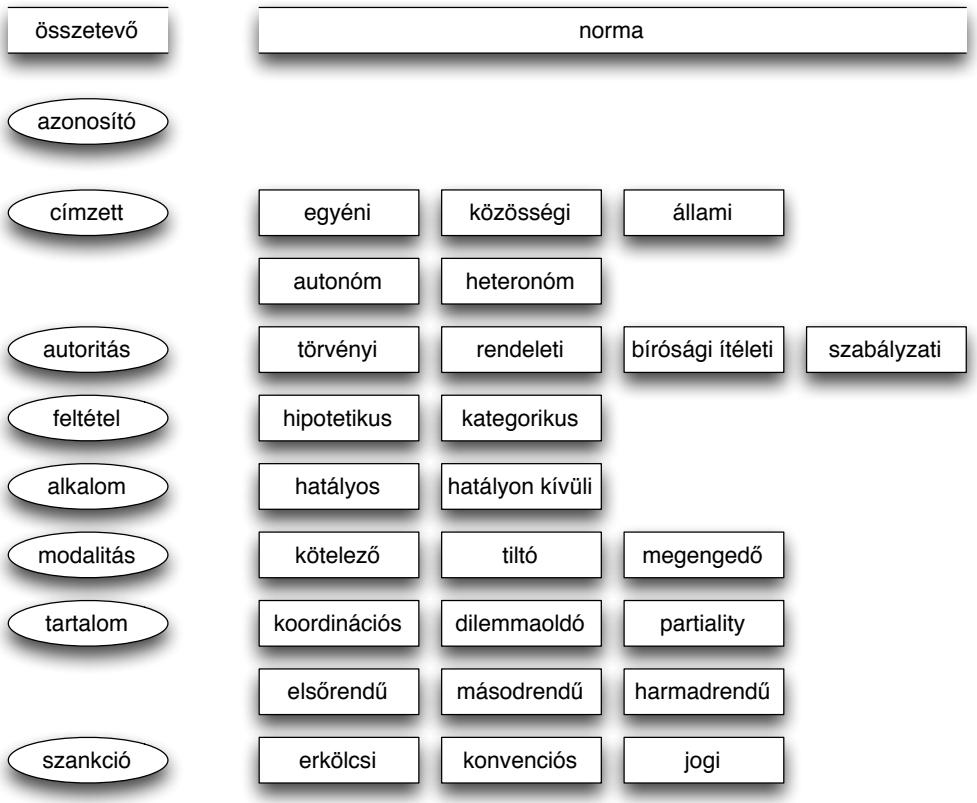


Figure 12.1: normatípusok a norma összetevői alapján

egyéb

HÉV-en utazas Csak egy negyes ulesben ul egy par Felteteles norma: egy uj utas nem ul mellejuk, hanem uj negyest valaszt. Ha nagy a valasztasi ter, akkor nagyobb a privat ter merete is. Ha a valasztasi/dontesi ter csokken, akkor mar be lehet ulni a foglalt negyesbe is. 2) Ha ketten ulnek egy boxban, az ulesre felteszik a labukat, es egy uj utas megkerdezi, szabad-e a hely, akkor ezzel inkabb szankcionalni akar

12.2 Játékelméleti magyarázat a normák keletkezésére

norma emergencia

Luca Tummolini: konvenció [755] [756] [757]

Dawid Lewis konvenció [444]

Sugden három konveciótpíust határol el [716]: koordinációs helyzetekkel kapcsoaltos konvenciók tulajdonnal kapcsolatos kovenciók reciprocitással kapcsaolatos konvenciók

Payton Young három cikke: Evolution of Conventions [834] Economics of Conventions [835] The dynamics of social innovation [832]

Bicchieri, Rationality and Coordination [82]

Cristina Bicchieri, The Grammar of Society. The Nature and Dynamics of Social Norms [83]

Skyrms, Evolution of the Social Contracts, [687]

The Stag Hunt and the Evolution of Social Structure [688]

Pete Krisztián a konvenciókról, közös tudásról, intencióról [569]

Ullmann-Margalit, The Emergence of Norms [762] fogolydilemma játékok koordinációs játékok impartiality (pártosság, részrehajlás vs. pártatlanság,)

Binmore a méltányossági normák kialakulásáról: [85]

Binmore a preferenciák és normák különbségeiről, a normák funkcióiról: [87]

Orthmayr Imre, társadalmi normák döntéselméleti és evolúciós magyarázata [555]

12.3 Rítus, narratíva

narratív megközelítés

[745]

Chapter 13

Normavezérelt társadalmi kapcsolat

13.1 Hatalom és uralom viszonya

KIss Balázs, Foucault hatalomfelfogása [398]

Hart:

> „ . . . de még az ilyen törvény is különbözik a szóban forgó utasításoktól abból a fontos szempontból, hogy rendszerint nem csupán másokra vonatkozik, hanem megalkotóira is." [346, 98]

magasabbrendű kötelezettségek Michael McKinsey a kötelezettség szintjei [489]

[140] Cristiano Castelfranchi, The "Logic" of Power. Hints on How my Power Becomes his Power

The Micro-Macro Constitution of Power [139]

13.2 Szervezet

a társadalmi szerep részt összekötni ezzel

Norms are not usually addressed to individual agents, but rather they are addressed to roles played by agents Goffman

G. Boella and L. van der Torre. The ontological properties of social roles in multi-agent systems: Definitional dependence, powers and roles playing roles. Artificial Intelligence and Law Journal (AILaw), 2007.

„In organizing a system all types of norms are necessary, in particular, constitutive norms, which are used to assign powers to agents playing roles inside the organization. Such powers allow to give commands to other agents, make formal communications and to restructure the organization itself, for example, by managing the assignment of agents to roles." [99]

legal institutions are defined by Ruiter as „systems of [regulative and constitutive] rules that provide frameworks for social action within larger rule-governed settings" Ruiter: A basic classification of legal institutions [624]

'This shift of interest marks the passage of focus from the more static legalistic view of norms (where power structures are fixed) to the more dynamic interactionist view of norms (where agent interaction is the base for norm related regulation).' This ties in to

what Strauss [42] called “negotiated order”, Goffman?s view on institutions, and Giddens? structuration theory.

A. Strauss. Negotiations: Varieties, Contexts, Processes and Social Order. San Francisco, Jossey-Bass, 1978.

Fiske ordinális rendezése

“Rank-oriented motivations are intrinsically malevolent, since one person’s rise is another’s fall - the process is a zero-sum game.” [358]

13.3 norma és érdek

a fine is a price: [298]

In a widely discussed example of the so- called centipede game, there is a pile of thousand pennies, and two agents can in turn either take one or two pennies. If an agent takes one then the other agent takes turn, if it takes two then the game ends. A backward induction argument implies that it is rational only to take two at the first turn. Norms and trust: M. Hollis. Trust within reason. Cambridge University Press, Cambridge, 1998.

Chapter 14

Egyéb

14.1 erkölcsi elvek

kettős hatás doktrinája (doctrine of double effect)

Aquinoi Szent Tamás Anscombe Philippa Foot Sanders [634] Taurek [739] plato.stanford.edu

14.2 Antinómiák, önreferencia

[690]

14.3 Hibák, javítandók

bibliográfia: harvard stílusú hivatkozás kell a cikkcímekben most kiveszi a nagybetűket, ezt kikapcsolni von előnevet figyelembe veszi a névrendezésnél (von wright) nem-magyar neveknél családnév+,+keresztnév sorrend kéne az oldalszám rövidítése 'p' helyett 'o' legyen amikor magyar kiadásban magyar nevű szerkesztő van, de nem-magyar nevű szerzők, akkor hogy lehet a magyar névre érvényesíttetni a huname mezőt?

ábrák: feliratok kisebb betűvel magyarítani (figure helyett ábra stb.) a caption szöveg menjen közelebb az ábrához

magyar elválasztás

layout: a definíciós paragrafusok előtt és után kevesebb soremelés legyen ne legyen olyan nagy behúzás

Bibliography

[1] *2013. évi V. törvény a Polgári Törvénykönyvről*. 2013.

[2] Carole Adam: *Emotions: from psychological theories to logical formalization and implementation in a BDI agent*. PhD értekezés (Docteur de l'Institut National Polytechnique de Toulouse). 2007.

[3] Icek Ajzen: From intentions to actions: A theory of planned behavior. In Julius Kuhl – Jürgen Beckmann (szerk.): *Action control: From cognition to behavior*. 1985, Springer-Verlag, 11–39. p.

[4] Icek Ajzen: The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50. évf. (1991), 179–211. p.

[5] Aki eladta a Rottenbiller utcát – Legendás szélhámosok 5., 2012. július. URL `http://player.hu/biznisz/aki+eladta+a+rottenbiller+utcat+legendas+szelhamosok+5/#.UZEcjJWgSlA`.

[6] Carlos E. Alchourrón: Detachment and defeasibility in deontic logic. *Studia Logica: An International Journal for Symbolic Logic*, 57. évf. (1996) 1. sz., 5–18. p.

[7] Carlos E. Alchourrón – Peter Gärdenfors – David Makinson: On the logic of theory change: Partial meet contraction and revision functions. *The Journal of Symbolic Logic*, 50. évf. (1985) 2. sz., 510–530. p.

[8] Layman E. Allen: From the fundamental legal conceptions of hohfeld to legal relations: Refining the enrichment of solely deontic legal relations. In Mark A. Brown – José Carmo (szerk.): *Proceedings of the 3rd International Workshop on Deontic Logic in Computer Science (DEON 1996)*. 1996, Springer, 1–26. p.

[9] Altrichter Ferenc: A tudat két aspektusa: intencionalitás és qualia. In Vizi E. Szilveszter – Altrichter Ferenc – Nyíri Kristóf – Pléh Csaba (szerk.): *Agy és tudat*. 2002, BIP Kiadó, 135–143. p.

[10] Ambrus-Lakatos Loránd – Meszerics Tamás: Az ultimátumjáték elemzéséhez. *Közgazdasági Szemle*, L. évf. (2003. június), 505–518. p.

[11] Alan Ross Anderson: The logic of norms. *Logique et Analyse*, 2. évf. (1958), 84–91. p.

[12] Alan Ross Anderson: A reduction of deontic logic to alethic modal logic. *Mind, New Series*, 67. évf. (1958) 265. sz., 100–103. p.

[13] Alan Ross Anderson: Logic, norms and roles. *Ratio*, 4. évf. (1962), 36–49. p.

[14] Alan Ross Anderson: The formal analysis of normative systems. In N. Rescher (szerk.): *The Logic of Decision and Action*. 1967, University of Pittsburgh Press, 147–213. p.

[15] Alan Ross Anderson: Some nasty problems in the formal logic of ethics. *Noûs*, 1. évf. (1967) 4. sz., 345–360. p.

[16] Alan Ross Anderson: The logic of hohfeldian propositions. *Logique et Analyse*, 1970., 231–242. p.

[17] Alan Ross Anderson – Omar Khayyam Moore: The formal analysis of normative concepts. *American Sociological Review*, 22. évf. (1957) 1. sz., 9–17. p.

[18] Alan Ross Anderson – Omar Khayyam Moore: A normatív fogalmak formális elemzése. In Irving M. Copi – James A. Gould (szerk.): *Kortárs-tanulmányok a logikaelmélet kérdéseiről*. 1985, Gondolat Kiadó, 561–583. p.

[19] M. Anderson – S. Leigh Anderson: Machine ethics: Creating an ethical intelligent agent. *AI Magazine*, 28. évf. (2007) 4. sz., 15–26. p.

[20] Hajnal Andréka – M. D. Ryan – P.-Y. Schobbens: Operators and laws for combining preference relations. *Journal of Logic and Computation*, 12. évf. (2002) 1. sz., 13–53. p.

[21] G. E. M. Anscombe: On frustration of the majority by fulfilment of the majority's will. *Analysis*, 36. évf. (1976) 4. sz., 161–168. p.

[22] G.E.M. Anscombe: *Intention*. 1963, Harvard University Press.

[23] Stephen Ansolabehere – James M. Snyder Jr.: Valence politics and equilibrium in spatial election models. *Public Choice*, 103. évf. (2000) 3/4. sz., 327–336. p.

[24] Lennart Åqvist: Good samaritans, contrary-to-duty imperatives, and epistemic obligations. *Noûs*, 1. évf. (1967) 4. sz., 361–379. p.

[25] Lennart Åqvist: Chisholm-sosa logics of intrinsic betterness and value. *Noûs*, 2. évf. (1968) 3. sz., 253–270. p.

[26] Lennart Åqvist: Some results on dyadic deontic logic and the logic of preference. *Synthese*, 66. évf. (1986) 1. sz., 95–110. p.

[27] Hannah Arendt: *A forradalom*. 1991, Európa Könyvkiadó.

[28] Dan Ariely: *Kiszámíthatóan irracionális. A racionálisnak vélt döntéseinket alakító rejtett erőkről*. 2011, Gabo Kiadó.

[29] Horacio Arló-Costa – Arthur Paul Pedersen: Social norms, rational choice and belief change. In E.J. Olsson – S. Enqvist (szerk.): *Belief Revision meets Philosophy of Science*. Logic, Epistemology, and the Unity of Science sorozat, 21. köt. 2010, Springer, 163–212. p.

[30] Gustaf Arrhenius: Superiority in values. *Philosophical Studies*, 123. évf. (2005), 97–114. p.

[31] Kenneth J. Arrow: A difficulty in the concept of social welfare. *Journal of Political Economy*, 58. évf. (1950) 4. sz., 328–346. p.

[32] Kenneth J. Arrow: Rational choice functions and orderings. *Economica*, 26. évf. (1959), 121–127. p.

[33] Kenneth J. Arrow: *Social Choice and Individual Values*. 1963, John Wiley & Sons.

[34] Alberto Artosi: On the notion of an empowered agent.

[35] Robert Audi (szerk.): *The Cambridge Dictionary of Philosophy*. 1999, Cambridge University Press.

[36] Aurelius Augustinus: *Vallomások*. Budapest, 1982, Gondolat Kiadó.

[37] Robert Aumann: Backward induction and common knowledge of rationality. *Games and Economic Behavior*, 8. évf. (1995), 6–19. p.

[38] Robert J. Aumann: Agreeing to disagree. *The Annals of Statistics*, 4. évf. (1976) 6. sz., 1236–1239. p.

[39] Robert J. Aumann: Rationality and bounded rationality. *Games and Economic Behavior*, 21. évf. (1997), 2–14. p.

[40] Robert J. Aumann: War and peace. nobel prize lecture, 2005. December.

[41] Robert J. Aumann: Rule-rationality versus act-rationality. Jelentés, 2008, Center for the Study of Rationality at The Hebrew University of Jerusalem.

[42] David Austen-Smith – Jeffrey Banks: Cycling of simple rules in the spatial model. *Social Choice and Welfare*, 16. évf. (1999) 4. sz., 663–672. p.

[43] Robert Axelrod: On six advances in cooperation theory. *Analyse & Kritik*, 22. évf. (2000) 1. sz., 130–151. p.

[44] Robert Axelrod: *The Evolution of Cooperation*. 2006 (1984), Basic Book.

[45] Balkányi Péter: A cselekvések racionalitásának kérdéséről. *Világosság*, 44. évf. (2003) 3-4. sz., 171–181. p.

[46] Miguel A. Ballester – Guillaume Haeringer: A characterization of the single-peaked domain. *Social Choice and Welfare*, 36. évf. (2011) 2. sz., 305–322. p.

[47] Salvador Barberà: Pivotal voters. a new proof of arrow's theorem. *Economic Letters*, 6. évf. (1980), 13–16. p.

[48] Salvador Barberà: Strategy-proofness and Pivotal Voters: A direct proof of the Gibbard-Satterthwaite theorem. *International Economic Review*, 24. évf. (1983), 413–417. p.

[49] Salvador Barberà: Indifferences and domain restrictions. *Analyse und Kritik*, 29. évf. (2007) 2. sz., 146–162. p.

[50] Salvador Barberà – Dolors Berga – Bernardo Moreno: Individual versus group strategy-proofness: when do they coincide? *Journal of Economic Theory*, 145. évf. (2010), 1648–1674. p.

[51] Salvador Barberà – Dolors Berga – Bernardo Moreno: Domains, ranges and strategy-proofness: the case of single-dipped preferences. *Social Choice and Welfare*, 39. évf. (2012), 335–352. p.

[52] Salvador Barberà – Dolors Berga – Bernardo Moreno: Two necessary conditions for strategy-proofness: On what domains are they also sufficient? *Games and Economic Behavior*, 75. évf. (2012), 490–509. p.

[53] Salvador Barberà – Bernardo Moreno: Top monotonicity: a common root for single peakedness, single crossing and the median voter result. *Games and Economic Behavior*, 73. évf. (2011) 2. sz., 345–359. p.

[54] Salvador Barberà – B. Peleg: Strategy-proof voting schemes with continuous preferences. *Social Choice and Welfare*, 7. évf. (1990) 1. sz., 31–38. p.

[55] Brian Barry: *Political Argument*. 1965, Routledge and Kegan Paul.

[56] J. Bartholdi – J. Orlin: Single Transferable Vote resists strategic voting. *Social Choice and Welfare*, 8. évf. (1991), 341– 354. p.

[57] J. J. Bartholdi III – C. A. Tovey – M. A. Trick: The computational difficulty of manipulating an election. *Social Choice and Welfare*, 6. évf. (1989) 3. sz., 227–241. p.

[58] Gary S. Becker: A theory of social interactions. *Journal of Political Economy*, 82. évf. (1974) 6. sz., 1063–1093. p.

[59] Gary S. Becker: Altruism, egoism, and genetic fitness: Economics and sociobiology. *Journal of Economic Literature*, 14. évf. (1976. Sep) 3. sz., 817–326. p. URL `http://www.jstor.org/stable/2722629`.

[60] Gary S. Becker: Altruism in the family and selfishness in the market place. *Economica, New Series*, 48. évf. (1981. Feb) 189. sz., 1–15. p.

[61] Bednay Dezső – Ollár Mariann: Anonim és semleges társadalmi választási függvények. *Szigma*, 40. évf. (2009) 1-2. sz., 67–77. p.

[62] Nuel Belnap: Backwards and forwards in the modal logic of agency. *Philosophy and Phenomenological Research*, 51. évf. (1991) 4. sz., 777–807. p.

[63] Nuel Belnap – Michael Perloff: Seeing to it that: A canonical form for agentives. *Theoria*, 54. évf. (1988) 3. sz., 175–199. p.

[64] Nuel Belnap – Michael Perloff: Seeing to it that: A canonical form for agentives. In Jr. Henry E. Kyburg – Ronald P. Loui – Greg N. Carlson (szerk.): *Knowledge Representation and Defeasible Reasoning*. 1990, Kluwer Academic Publishers, 167–190. p.

[65] Nuel Belnap – Michael Perloff – Ming Xu: *Facing the Future. Agents and Choices in Our Indeterminist World*. 2001, Oxford University Press.

[66] Martin Mose Bentzen: *Stit, Iit, and Deontic Logic for Action Types*. PhD értekezés (Section for Philosophy and Science Studies Roskilde University). 2009.

[67] Bereczkei Tamás: *Az erény természete*. Budapest, 2009, Typotex.

[68] Ralph Bergmüller – Rufus A. Johnstone – Andrew F. Russell – Redouan Bshary: Integrating cooperative breeding into theoretical concepts of cooperation. *Behavioural Processes*, 76. évf. (2007), 61–72. p.

[69] Theodore Bergstrom: A fresh look at the rotten kid theorem – and other household mysteries. *Journal of Political Economy*, 97. évf. (1989), 1138–1159. p.

[70] Isaiah Berlin: *Négy esszé a szabadságról*. 1990, Európa Könyvkiadó.

[71] Helen Bernhard – Urs Fischbacher – Ernst Fehr: Parochial altruism in humans. *Nature*, 442. évf. (2006) 7105. sz., 912–915. p.

[72] Peter Bernholz: Logrolling, arrow paradox and cyclical majorities. *Public Choice*, 15. évf. (1973), 87–95. p.

[73] Peter Bernholz: Logrolling and the paradox of voting: Are they really logically equivalent? a comment. *The American Political Science Review*, 69. évf. (1975) 3. sz., 961–962. p.

[74] Peter Bernholz: Prisoner's dilemma, logrolling and cyclical group preferences. *Public Choice*, 29. évf. (1977) 1. sz., 73–84. p.

[75] Peter Bernholz: Externalities as a necessary conditions for cyclical social preferences. *The Quarterly Journal of Economics*, 97. évf. (1982) 4. sz., 699–705. p.

[76] Peter Bernholz: A general constitutional possibility theorem. *Public Choice*, 51. évf. (1986) 3. sz., 249–265. p.

[77] Bertalan László: Előadások Max Weberről: Max Weber cselekvéselmélete és cselekvéstipológiája. *Szociológiai Szemle*, 11. évf. (2001) 3. sz., 113–136. p.
URL `http://www.szociologia.hu/dynamic/0103bertalan.htm`.

[78] Yevgeni Berzak – Michael Fink: Manipulating allocation justice: How framing effects can increase the prevalence of the talmudic division principle "shnaim ohazin". Jelentés, 2008, Center for the Study of Rationality at The Hebrew University of Jerusalem.

[79] Th.G.G. Bezembinder – P. Van Acker: The ostrogorski paradox and its relation to nontransitive choice. *Journal of Mathematical Sociology*, 11. évf. (1985) 2. sz., 131–158. p.

[80] Jean-Yves Béziau: New light on the square of oppositions and its nameless corner. *Logical Investigations*, 10. évf. (2003), 218–233. p.

[81] Jean-Yves Béziau: The power of the hexagon. *Logica Universalis*, 6. évf. (2012), 1–43. p.

[82] Cristina Bicchieri: *Rationality and Coordination*. 1993, Cambridge University Press.

[83] Cristina Bicchieri: *The Grammar of Society. The Nature and Dynamics of Social Norms*. 2006, Cambridge University Press.

[84] Meghyn Bienvenu – Jérôme Lang – Nic Wilson: From preference logics to preference languages, and back. In *Proceedings of the Twelth International Conference on the Principles of Knowledge Representation and Reasoning (KR 2010)* (konferenciaanyag). 2010, 214–224. p.

[85] Ken Binmore: The evolution of fairness norms. *Rationality and Society*, 10. évf. (1998), 275–301. p.

[86] Ken Binmore: *Game Theory. A Very Short Introduction*. 2007, Oxford University Press.

[87] Ken Binmore: Social norms or social preferences? *Mind & Society*, 9. évf. (2010), 139–157. p. ISSN 1593-7879. URL `http://dx.doi.org/10.1007/s11299+010+0073+2`. 10.1007/s11299-010-0073-2.

[88] Biró Péter: Stabil páros´ási modellek és ezeken alapuló központi páros´ó programok. *Szigma*, XXXVII. évf. (2006) 3-4. sz., 153–175. p.

[89] Biró Péter – Csóka Péter – Kóczy Á. László – Radványi Anna Ráhel – Sziklai Balázs: Közgazdasági nobel-emlékdíj2012: Alvin e. roth és lloyd s. shapley. *Magyar Tudomány*, 174. évf. (2013) 2. sz., 190–199. p.

[90] Thomas Bittner – Maureen Donnelly – Barry Smith: Individuals, universals, collections: On the foundational relations of ontology. In Achille C. Varzi – Laure Vieu (szerk.): *Formal Ontology in Information Systems. Proceedings of the Third International Conference (FOIS 2004)*. Amsterdam, Berlin, Oxford, Tokyo, Washington DC, 2004, IOS Press, 37–48. p.

[91] Duncan Black: On the rationale of group decision-making. *Journal of Political Economy*, 56. évf. (1948), 23–34. p.

[92] Patrick Blackburn – Maarten de Rijke – Yde Venema.: *Modal Logic*. 2001, Cambridge University Press.

[93] Julian H. Blau: The existence of social welfare functions. *Econometrica*, 25. évf. (1957) 2. sz., 302–313. p.

[94] Lawrence Blume – Adam Brandenburger – Eddie Dekel: An overview of lexicographic choice under uncertainty. *Annals of Operations Research*, 19. évf. (1989), 231–246. p.

[95] Jean Beer Blumenfeld: Goldman's account of intentional action. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 29. évf. (1976) 6. sz., 391–396. p.

[96] Boda László: *A keresztény erkölcs alapkérdései*. 1997, Szent József Kiadó.

[97] G. Boella – L. van der Torre – H. Verhagen: Introduction to normative multiagent systems. *Computation and Mathematical Organizational Theory*, 12. évf. (2006) 2-3. sz., 71–79. p.

[98] G. Boella – H. Verhagen – L. van der Torre: Introduction to the special issue on normative multiagent systems. *Journal of Autonomous Agents and Multi Agent Systems*, 17. évf. (2008) 1. sz., 1–10. p.

[99] Guido Boella – Gabriella Pigozzi – Leendert van der Torre: Normative systems in computer science. ten guidelines for normative multiagent systems. In Guido Boella – P. Noriega – Gabriella Pigozzi – H. Verhagen (szerk.): *Normative Multi-Agent Systems* (konferenciaanyag). 2009.

[100] Guido Boella – Leendert van der Torre: Substantive and procedural norms in normative multiagent systems. *Journal of Applied Logic*, 6. évf. (2008) 2. sz., 152–171. p.

[101] Steven E. Boer: Speech acts and constitutive rules. *The Journal of Philosophy*, 71. évf. (1974) 6. sz., 169–174. p.

[102] G. E. Bolton – A. Ockenfels: Erc: a theory of equity, reciprocity and competition. *American Economic Review*, 90. évf. (2000) 1. sz., 166–193. p.

[103] Gary Bornstein – Ori Weisel: Punishment, cooperation, and cheater detection in "noisy" social exchange. *Games*, 1. évf. (2010) 1. sz., 18–33. p.

[104] Walter Bossert – Kotaro Suzumura: Social norms and rationality of choice. ISSN 0709-9231. Jelentés, 2007. 7, Département de sciences économiques Université de Montréal.

[105] Walter Bossert – Kotaro Suzumura: *Consistency, Choice, and Rationality*. 2010, Harvard University Press.

[106] Raymond Boudon: *The Unintended Consequences of Social Action*. New York, 1982, St. Martin's Press.

[107] Pierre Bourdieu: *A társadalmi egyenlőtlenségek újratermelődése*. 1978, Gondolat.

[108] Pierre Bourdieu: *A gyakorlat elméletének vázlata. Három kabil etnológiai tanulmány*. 2009, Napvilág Kiadó.

[109] Pierre Bourdieu – Loïc J.D. Wacquant: *An Invitation to Reflexive Sociology*. 2009, Polity Press.

[110] Craig Boutilier: Toward a logic for qualitative decision theory. In Jon Doyle – Erik Sandewall – Pietro Torasso (szerk.): *Principles of knowledge representation and reasoning*. 1994, Morgan Kaufmann Publishers, 75–86. p.

[111] Robert Boyd: Mistakes Allow Evolutionary Stability in the Repeated Prisoner's Dilemma Game. *Journal of Theoretical Biology*, 136. évf. (1989), 47–56. p.

[112] Robert Boyd – Jeffrey P. Lorberbaum: No pure strategy is evolutionarily stable in the repeated prisoner's dilemma game. *Nature*, 327. évf. (1987. May), 58–59. p.

[113] Ronald J. Brachman – Hector J. Levesque: *Knowledge Representation and Reasoning*. 2004, Morgan Kaufmann Publishers.

[114] S.J. Brams – D.M. Kilgour – W.S. Zwicker: The paradox of multiple elections. *Social Choice and Welfare*, 15. évf. (1998) 2. sz., 211–236. p.

[115] Steven J. Brams: *Theory of moves*. 1994, Cambridge University Press.

[116] Steven J. Brams – William V. Gehrlein – Fred S. Roberts (szerk.): *The Mathematics of Preference, Choice and Order. Essays in Honor of Peter C. Fishburn*. 2009, Springer.

[117] Michael Bratman: *Intention, Plans, and Practical Reason*. Cambridge, MA, 1987, Harvard University Press.

[118] Michael E. Bratman: *Structures of Agency*. Oxford, 2006, Oxford University Press.

[119] Robert Bredereck – Jiehua Chen – Gerhard J. Woeginger: Are there any nicely structured preference profiles nearby? In *Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI '13)* (konferenciaanyag). 2013, AAAI Press.

[120] Robert Bredereck – Jiehua Chen – Gerhard J. Woeginger: A characterization of the single-crossing domain. *Social Choice and Welfare*, 41. évf. (2013), 989–998. p.

[121] Geoffrey Brennan: The esteem engine: A resource institutional design. 1. Occasional Paper Series, 2005, Academy of the Social Sciences in Australia - Cunningham Lecture 2004.

[122] Geoffrey Brennan – Philip Pettit: *The Economy of Esteem*. 2004, Oxford University Press.

[123] Franz Brentano: A pszichikai és fizikai jelenségek közötti különbségrõl. In Csaba Pléh (szerk.): *Pszichológiatörténeti szöveggyûjtemény*. I. köt. 1983, Tankönyvkiadó, 157–163. p.

[124] Franz Brentano: *Az erkölcsi ismeret eredete*. 1994, Kossuth Könyvkiadó.

[125] John Broome: Reasoning with preferences? In Serena Olsaretti (szerk.): *Preferences and Well-Being*. 2006, Cambridge University Press, 183–208. p.

[126] John Broome: Motivation. *Theoria*, 75. évf. (2009), 79–99. p.

[127] R. Bshary – A.S. Grutter: Image scoring and cooperation in a cleaner fish mutualism. *Nature*, 441. évf. (2006) 22. sz., 975–978. p.

[128] James M. Buchanan: An ambiguity in sen's alleged ptoof of the impossibility of a pareto libertarian. *Analyse & Kritik*, 18. évf. (1996), 118–125. p.

[129] James M. Buchanan – Gordon Tullock: *The Calculus of Consent: Logical Foundations of Constitutional Democracy*. Ann Arbor, Michigan, 1962:, University of Michigan Press.

[130] James N. Buchanan: A törvény és a láthatatlan kéz. In *Piac, állam, alkotmányosság*. 1992, KJK, 39–51. p.

[131] Robert A. Bull – Krister Segerberg: Basic modal logic. In Dov M. Gabbay – Franz Guenthner (szerk.): *Handbook of Philosophical Logic*. 3. köt. 2nd edition. kiad. 2001, Kluwer Academic Publishers, 1–82. p.

[132] Eugenio Bulygin: On norms of competence. *Law and Philosophy*, 11. évf. (1992) 3. sz., 201–216. p.

[133] B.J. Bushman – C.A. Anderson: Is it time to pull the plug on the hostile versus instrumental aggression dichotomy? *Psychological Review*, 108. évf. (2001) 1. sz., 273–279. p.

[134] Sarah Buss – Lee Overton (szerk.): *Contours of Agency. Essays on Themes from Harry Frankfurt*. 2002, The MIT Press.

[135] Donald E. Campbell – Jerry S. Kelly: A simple characterization of majority rule. *Source: Economic Theory*, 15. évf. (2000) 3. sz., 689–700. p.

[136] Georges Canguilhem: *A normális és a kóros*. Budapest, 2004, Gondolat.

[137] C. Castelfranchi: Modeling social action for ai agents. *Artificial Intelligence*, 103. évf. (1998) 1-2. sz., 157–182. p.

[138] Cristiano Castelfranchi: Commitments: from individual intentions to groups and organizations. In Victor R. Lesser – Les Gasser (szerk.): *Proceedings of the First International Conference on Multiagent Systems, ICMAS 1995* (konferenciaanyag). 1995, The MIT Press, 41–48. p.

[139] Cristiano Castelfranchi: The micro-macro constitution of power. *ProtoSociology, An International Journal of Interdisciplinary Research*, 18-19. évf. (2003), 208–265. p.

[140] Cristiano Castelfranchi: The “logic” of power. hints on how my power becomes his power. In *AISB 2011* (konferenciaanyag). 2011.

[141] Cristiano Castelfranchi – Fabio Paglieri: The role of beliefs in goal dynamics: Prolegomena to a constructive theory of intentions. *Synthese,*, 155. évf. (2007) 2. sz., 237–263. p.

[142] Gary Charness – Matthew Rabin: Understanding social preferences with simple tests. *Quarterly Journal of Economics*, 117. évf. (2002), 817–869. p.

[143] Brian F. Chellas: Imperatives. *Theoria*, 37. évf. (1971) 2. sz., 114–129. p.

[144] Brian Farrell Chellas: *The logical form of imperatives*. 1968, Dept. of Philosophy, Stanford University.

[145] Brian Farrell Chellas: *Modal Logic: An Introduction*. 1980, Cambridge University Press. ISBN 0-521-22476-4.

[146] Herman Chernoff: Rational selection of decision functions. *Econometrica*, 22. évf. (1954) 4. sz., 422–443. p.

[147] Mark J. Cherry (szerk.): *The Normativity of the Natural. Human Goods, Human Virtues, and Human Flourishing*. 2009, Springer.

[148] Roderick M. Chisholm: Contrary-to-duty imperatives and deontic logic. *Analysis*, 24. évf. (1963) 2. sz., 33–36. p.

[149] Roderick M. Chisholm: Supererogation and offence: A conceptual scheme for ethics. *Ratio*, 1963. 5. sz., 1–14. p.

[150] Roderick M. Chisholm: The ethics of requirement. *American Philosophical Quarterly*, 1. évf. (1964) 2. sz., 147–153. p.

[151] Roderick M. Chisholm: *The Encyclopedia of Philosophy*. 3 and 4. köt. Intentionality fejezet. 1967, MacMillan, 201–204. p.

[152] Roderick M. Chisholm: The structure of intention. *The Journal of Philosophy*, 67. évf. (1970) 19. sz., 633–647. p. Sixty-Seventh Annual Meeting of the American Philosophical Association Eastern Division.

[153] Roderick M. Chisholm – Ernest Sosa: Intrinsic preferability and the problem of supererogation. *Synthese*, 16. évf. (1966) 3/4. sz., 321–331. p.

[154] Roderick M. Chisholm – Ernest Sosa: On the logic of "intrinsically better". *American Philosophical Quarterly*, 1966., 244–249. p.

[155] Shoham Choshen-Hillel – Ilan Yaniv: Agency and the construction of social preference: Between inequality aversion and prosocial behavior. *2011*, Journal of Personality and Social Psychology.

[156] Robert B. Cialdini: Descriptive social norms as underappreciated sources of social control. *Psychometrika*, 72. évf. (2007) 2. sz., 263–268. p.

[157] Robert B. Cialdini – Linda J. Demaine – Brad J. Sagarin – Daniel W. Barrett – Kelton Rhoads – Patricia L. Winter: Managing social norms for persuasive impact. *Social Influence*, 1. évf. (2006) 1. sz., 3–15. p.

[158] Robert B. Cialdini – R.R. Reno – C.A. Kallgren: A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology*, 58. évf. (1990) 6. sz., 1015–1026. p.

[159] A.V. Cicourel: Basic and normative rules in the negotiation of status and role. In D. Sudnow (szerk.): *Studies in Social Interaction*. 1972, Glencoe, 229–258. p.

[160] Edward H. Clarke: Multipart pricing of public goods. *Public Choice*, 11. évf. (1971), 17–33. p.

[161] Edward H. Clarke: Some aspects of the demand-revealing process. *Public Choice*, 29. évf. (1977), 37–49. p.

[162] Tim Clutton-Brock: Cooperation between non-kin in animal societies. *Nature*, 462. évf. (2009) 5. sz., 51–57. p.

[163] Philip R. Cohen – Hector J. Levesque: Speech acts and rationality. In *Proceedings Twenty-third Annual Meeting Association for Computational Linguistics* (konferenciaanyag). Chicago, IL, 1985, 49–59. p.

[164] Philip R. Cohen – Hector J. Levesque: Intention is choice with commitment. *Artificial Intelligence*, 42. évf. (1990) 213-261. sz.

[165] Philip R. Cohen – Hector J. Levesque: Performatives in a rationally based speech act theory. In *ACL '90 Proceedings of the 28th annual meeting on Association for Computational Linguistics* (konferenciaanyag). 1990, 79–88. p.

[166] Philip R. Cohen – Hector J. Levesque: Rational interaction as the basis for communication. In P.R. Cohen – J. Morgan – M.E. Pollack (szerk.): *Intentions in communication*. Cambridge, ma. kiad. 1990, The MIT Press, 221–255. p.

[167] Tom Coleman – Vanessa Teague: On the complexity of manipulating elections. In *Proceeding CATS '07 Proceedings of the thirteenth Australasian symposium on Theory of computing* (konferenciaanyag), 65. köt. 2007, Australian Computer Society, 25–33. p.

[168] Andrew M. Colman: Evolution of cooperation without awareness in minimal social situations. In N. Gold (szerk.): *Teamwork: Multi-disciplinary perspectives*. 2005, Palgrave Macmillan, 216–235. p.

[169] Andrew M. Colman – Jonathan A. Stirk: Singleton bias and lexicographic preferences among equally valued alternatives. *Journal of Economic Behavior & Organization*, 40. évf. (1999), 337–351. p.

[170] Vincent Conitzer – Tuomas Sandholm – Jérôme Lang: When are elections with few candidates hard to manipulate. *Journal of the Association for Computing Machinery*, 54. évf. (2007) 3. sz.

[171] Benjamin Constant: *A régiek és a modernek szabadsága*. 1997, Atlantisz.

[172] D. Copp: Reasons and needs. In R.G. Frey – Christopher W. Morris (szerk.): *Value, Welfare, and Morality*. Cambridge, 1993, Cambridge University Press, 112–137. p.

[173] Arthur Corbin: Legal analysis and terminology. *Yale Law Journal*, 1919., 163–173. p.

[174] Arthur Corbin: Jural relations and their classification. *Yale Law Journal*, 1921., –238. p.

[175] Gary W. Cox – Kenneth A. Shepsle: Majority cycling and agenda manipulation: Richard mckelvey's contributions and legacy. In John H. Aldrich – James E. Alt – Arthur Lupia (szerk.): *Positive Changes in Political Science*. 2007, The University of Michigan Press, 19–40. p.

[176] J.C. Cox – D. Friedman – S. Gjerstad: A tractable model of reciprocity and fairness. *Games and Economic Behavior*, 59. évf. (2007), 17–45. p.

[177] Charles B. Cross: The modal logic of discrepancy. *Journal of Philosophical Logic*, 26. évf. (1997) 2. sz., 143–168. p.

[178] Csáji Balázs Csanád – Rédei Miklós: A racionális demokratikus véleményösszegzés korlátairól. *Magyar Filozófiai Szemle*, 55. évf. (2011) 2. sz., 97–121. p.

[179] Csontos László: Szempontok a privatizáció és a hierarchikus gazdasági szervezetek belső szerkezetének változásai közti összefüggések tanulmányozásához. *Szociológiai Szemle*, 4. évf. (1994) 4. sz., 83–95. p.
URL `http://www.mtapti.hu/mszt/19944/csontos.htm`.

[180] László Csontos: A politika tanulmányozása és a közgazdaságtan. *Közgazdasági Szemle*, XLIV. évf. (1997), 557–568. p.

[181] Csontos László – Király Júlia – László Géza: Az ezredvégi nagy borzongás. *Közgazdasági Szemle*, XLIV. évf. (1997) július-augusztus. sz., 569–596. p.

[182] A. Damasio: Understanding the mind's will. *The Behavioral and Brain Sciences*, 8. évf. (1985) 4. sz., 589. p.

[183] Antonio Damasio: *Descartes tévedése. Érzelem, értelem és az emberi agy*. 1996, Aduprint.

[184] Roy D'Andrade: Cultural meaning systems. In R.A. Shweder – R. LeVine (szerk.): *Culture Theory. Essays on Mind, Self, and Emotion*. 1986, Cambridge University Press, 88–119. p.

[185] Peter Danielson: Competition among cooperators: Altruism and reciprocity. *PNAS*, 99. évf. (2002), 7237–7242. p.

[186] Arthur Danto: Freedom and forbearance. In K. Lehrer (szerk.): *Freedom and Determinism*. 1966, Random House, 45–63. p.

[187] Mehdi Dastani – Joris Hulstijn – Leendert van der Torre: How to decide what to do? *European Journal of Operational Research*, 160. évf. (2005), 762–78. p.

[188] Donald Davidson: The logical form of action sentences. In Nicholas Rescher (szerk.): *The Logic of Decision and Action*. 1966, University of Pittsburgh, 81–95. p.

[189] Donald Davidson: The logical form of action sentences. In *Essays on Actions and Event*. 1967, Clarendon Press, 105–121. p.

[190] Donald Davidson: Cselekvések, indokok és okok. *Thél@geme*, 1999. 4. sz., 7–21. p.

[191] Donald Davidson: Actions, reasons, and causes. In *Essays on Actions and Events*. 2002, Clarendon Press, 3–19. p.

[192] Otto A. Davis – Morris H. DeGroot – Melvin J. Hinich: Social preference orderings and majority rule. *Econometrica*, 40. évf. (1972) 1. sz., 147–157. p.

[193] Wayne A. Davis: A causal theory of intending. *American Philosophical Quarterly*, 21. évf. (1984) 1. sz., 43–54. p.

[194] Wayne A. Davis: The two senses of desire. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 45. évf. (1984) 2. sz., 181–195. p.

[195] Richard Dawkins: *Az önző gén*. 2005, Kossuth Könyvkiadó.

[196] Howard L. Dazeley – Wolfgang L. Gombocz: Interpreting anselm as logician a. *Synthese*, 40. évf. (1979) 1. sz., 71–96. p.

[197] Dick de Jongh – Fenrong Liu: Preference, priorities and belief. In T. Grüne-Yanoff – S. O. Hansson (szerk.): *Preference Change: Approaches from Philosophy, Economics and Psychology*. 2009, Springer, 85–108. p.

[198] Dominique J. F. de Quervain – Urs Fischbacher – Valerie Treyer – Melanie Schellhammer – Ulrich Schnyder – Alfred Buck – Ernst Fehr: The neural basis of altruistic punishment. *Science*, 305. évf. (2004) 5688. sz., 1254–1258. p.

[199] Gerard Debreu: Representation of a preference ordering by a numerical function. In Thrall – Davis – Coombs (szerk.): *Decision Process*. XI. köt. 1954, John Wiley, 159–165. p.

[200] Dénes Iván Zoltán: Beavatkozástól mentes cselekvés vagy uralomtól mentes emberi állapot? (a liberális és a republikánus szabadságfelfogás hasonlóságai és különbségei). *Jogelméleti Szemle*, 2004. 2. sz.
URL `http://jesz.ajk.elte.hu/denes18.html`.

[201] Daniel C. Dennett: *Content and Consciousness*. 1969, Routledge.

[202] Daniel C. Dennett: *Az intencionalitás filozófiája*. 1998, Osiris.

[203] Jacques Derrida: Mit jelent a “tanítani” szó? *Iskolakultúra*, 2007. 3. sz., 3–16. p.

[204] John Dewey: *Experience and Nature*. 1929, George Allan and Unwin.

[205] Andreas Diekmann: Volunteer's dilemma. *Journal of Conflict Resolution*, 29. évf. (1985) 4. sz., 605–610. p.

[206] Franz Dietrich: Judgment aggregation: (im)possibility theorems. *Journal of Economic Theory*, 126. évf. (2006), 286–298. p.

[207] Franz Dietrich – Christian List: Arrow's theorem in judgment aggregation. *Social Choice and Welfare*, 29. évf. (2007) 1. sz., 19–33. p.

[208] Franz Dietrich – Christian List: Majority voting on restricted domains. *Journal of Economic Theory*, 145. évf. (2010) 2. sz., 512–543. p.

[209] Franz Dietrich – Christian List: A reason-based theory of rational choice. *Noûs*, forthcoming. évf. (2012).

[210] Franz Dietrich – Christian List: Where do preferences come from? *International Journal of Game Theory*, 2012.

[211] Elad Dokow – Ron Holzman: Aggregation of binary evaluations. *Journal of Economic Theory*, 145. évf. (2010) 2. sz., 495–511. p.

[212] Dombi J. – Vincze N.: A lexikografikus döntési eljárás kapcsolata a hagyományos döntési eljárásokkal. In *XXV. Magyar Operációkutatási Konferencia Előadáskivonatok, Debrecen, Hungary* (konferenciaanyag), 32. köt. 2001.

[213] Dombi József – Vincze Nándor: Lexikografikus döntések egy általános döntési modellben. *Szigma*, 36. évf. (2005) 3-4. sz., 149–162. p.

[214] A. Downs: *An Economic Theory of Democracy*. 1957, Harper and Row.

[215] Jon Doyle – Michael Wellman: Representing preferences as ceteris paribus comparatives. In Steve Hanks – Stuart Russell – Michael Wellman (szerk.): *Working Notes of the AAAI Spring Symposium on Decision-Theoretic Planning* (konferenciaanyag). 1994, AAAI Press.

[216] Pierre Dragicevic: Játékelméleti jelölések.
URL `http://www.lri.fr/~dragice/gameicons/`.

[217] Fred Dretske: Epistemic operators. *The Journal of Philosophy*, 67. évf. (1970) 24. sz., 1007–1023. p.

[218] Lester E. Dubins – David A. Freedman: Machiavelli and the Gale-Shapley algorithm. *American Mathematical Monthly*, 88. évf. (1981) 7. sz., 485–494. p.

[219] Ho Ngoc Duc: *Resource-Bounded Reasoning about Knowledge*. PhD értekezés (Fakultät für Mathematik und Informatik der Universität Leipzig). 2001.

[220] M. Dufwenberg – G. Kirchsteiger: A theory of sequential reciprocity. *Games and Economic Behavior*, 47. évf. (2004) 2. sz., 268–298. p.

[221] John Duggan – Thomas Schwartz: Strategic manipulability without resoluteness or shared beliefs: Gibbard-satterthwaite generalized. *Social Choice and Welfare*, 17. évf. (2000) 1. sz., 85–93. p.

[222] Charles Duhigg: *A szokás hatalma*. 2012, Casparus Kiadó.

[223] Michael Dummett – Robin Farquharson: Stability in voting. *Econometrica*, 29. évf. (1961) 1. sz., 33–43. p.

[224] Émile Durkheim: Az erkölcsi tény meghatározása. In *A társadalmi tények magyarázatához*. Budapest, 1978, KJK, 161–195. p.

[225] Émile Durkheim: A szociológia módszertani szabályai. In *A társadalmi tények magyarázatához*. 1978, KJK, 21–160. p.

[226] Émile Durkheim: *Az öngyilkosság. Szociológiai tanulmány*. 1982, KJK.

[227] Douglas Ehring: "normal" intentional action. *Philosophy and Phenomenological Research*, 46. évf. (1985) 1. sz., 155–157. p.

[228] Paul Ekman: Basic emotions. In T. Dalgleish – M. Power (szerk.): *Handbook of Cognition and Emotion*. 3 fejezet. 1999, John Wiley and Sons, 45–60. p.

[229] Mircea Eliade: *A szent és a profán*. Budapest, 1996, Európa Könyvkiadó.

[230] J. Elster: Rationality and the emotions. *Economic Journal*, 106. évf. (1996) 438. sz., 1386–1397. p.

[231] Jon Elster: *Ulysses and the Sirens*. 1979, Cambridge University Press.

[232] Jon Elster: *A társadalom fogaskerei*. 1997, Osiris.

[233] Herbert Enderton: Second-order and higher-order logic. URL `http://plato.stanford.edu/entries/logic+higher+order/`.

[234] Ulle Endriss: Logic and social choice theory. In A. Gupta – J. van Benthem (szerk.): *Logic and Philosophy Today*. 2. köt. 2011, College Publications, 333–377. p.

[235] E. E. Evans-Pritchard: Science and sentiment an exposition and criticism of the writings of pareto. *Revue européenne des sciences sociales, Herméneutique paretienne*, 16. évf. (1978) 43. sz., 51–73. p.

[236] Jeffrey Fagan – Tracey L. Meares: Punishment, deterrence and social control: The paradox of punishment in minority communities. *Ohio State Journal of Criminal Law*, 6. évf. (2008), 173–229. p.

[237] Ronald Fagin – Joseph Y. Halpern – Yoram Moses – Moshe Y. Vardi (szerk.): *Reasoning about Knowledge*. 1995, MIT Press.

[238] Piotr Faliszewski – Edith Hemaspaandra – Lane A. Hemaspaandra: How hard is bribery in elections? *Journal of Artificial Intelligence Research*, 35. évf. (2009), 485–532. p.

[239] Piotr Faliszewski – Ariel D. Procaccia: AI's war on manipulation: Are we winning? *AI Magazine*, 31. évf. (2010) 4. sz., 53–64. p.

[240] Robin Farquharson: *Theory of Voting*. 1969, Yale University Press.

[241] Ernst Fehr – Urs Fischbacher: Third-party punishment and social norms. *Evolution and Human Behavior*, 25. évf. (2004), 63–87. p.

[242] Ernst Fehr – Urs Fischbacher – Simon Gächter: Strong reciprocity, human cooperation and the enforcement of social norms. *Human Nature*, 13. évf. (2002), 1–25. p.

[243] Ernst Fehr – Simon Gächter: Fairness and Retaliation: The Economics of Reciprocity. *Journal of Economic Perspectives*, 14. évf. (Summer 2000) 3. sz., 159–181. p.

[244] Ernst Fehr – Klaus M. Schmidt: A theory of fairness, competition and cooperation. *The Quarterly Journal of Economics*, 114. évf. (1999. August) 3. sz., 817–865. p.

[245] Joel Feinberg: Supererogation and rules. *Ethics*, 71. évf. (1961) 4. sz., 276–288. p.

[246] Joel Feinberg: *Társadalomfilozófia*. 1999, Osiris Kiadó.

[247] Todd E. Feinberg: *Altered Egos. How the Brain Creates the Self*. 2001, Oxford University Press.

[248] Allan M. Feldman: Strongly nonmanipulable multi-valued collective choice rules. *Public Choice*, 35. évf. (1980) 4. sz., 503–509. p.

[249] Adam Feltz: The knobe effect: A brief overview. *Journal of Mind and Behavior*, 28. évf. (2008), 265–278. p.

[250] Ferenczi Miklós: *Matematikai logika*. 2002, Műszaki Könyvkiadó.

[251] FIFA. *A labdarúgás játékszabályai 2009/2010*. 2010. URL `http://jatekvezetok.blsz.hu/szabalyzatok/szk2009magyar.pdf`.

[252] Peter C. Fishburn: On the foundations of game theory: The case of non-archimedean utilities. *International Journal of Game Theory*, 1. évf. (1971) 1. sz., 65–71. p.

[253] Peter C. Fishburn: A study of lexicographic expected utility. *Management Science*, 17. évf. (1971. July) 11. sz., 672–678. p.

[254] Peter C. Fishburn: Lexicographic orders, utilities and decision rules: A survey. *Management Science*, 20. évf. (1974. July) 11. sz., 1442–1471. p.

[255] Peter C. Fishburn: Axioms for lexicographic preferences. *The Review of Economic Studies*, 42. évf. (1975. July) 3. sz., 415–419. p.

[256] Peter C. Fishburn – Steven J. Brams: Paradoxes of preferential voting. *Mathematics Magazine*, 56. évf. (1983) 4. sz., 207–214. p.

[257] Peter C. Fishburn – Irving H. LaValle: Subjective expected lexicographic utility: Axioms and assessment. *Annals of Operations Research*, 80. évf. (1998), 183–206. p.

[258] Frederic B. Fitch: A logical analysis of some value concepts. *The Journal of Symbolic Logic*, 28. évf. (1963) 2. sz., 135–142. p.

[259] Frederic B. Fitch: A revision of hohfeld's theory of legal concepts. *Logique et Analyse*, 1967., 269–276. p.

[260] Anton Ford – Jennifer Hornsby – Frederick Stoutland (szerk.): *Essays on Anscombe's Intention*. 2011, Harvard University Press, 170–197. p.

[261] Forrai Gábor: Dennett intencionális hozzáállása. *Replika*, 10. évf. (1999) 36. sz., 59–70. p. URL `http://www.scribd.com/doc/74387228/Dennett+intencionalis+hozzaallasa`.

[262] James William Forrester: Gentle murder, or the adverbial samaritan. *The Journal of Philosophy*, 81. évf. (1984) 4. sz., 193–197. p.

[263] K. R. Foster – T. Wenseleers – F. L. W. Ratnieks: Spite: Hamilton's unproven theory. *Ann. Zoo. Fenn.*, 38. évf. (2001), 229–238. p.

[264] Robert Frank – Thomas Gilovich – Dennis Regan: Does studying economics inhibit cooperation? *Journal of Economic Perspectives*, 7. évf. (1993) 2. sz., 159–171. p.

[265] Harry G. Frankfurt: Freedom of the will and the concept of a person. *The Journal of Philosophy*, 68. évf. (1971) 1. sz., 5–20. p.

[266] Harry G. Frankfurt: The problem of action. *American Philosophical Quarterly*, 15. évf. (1978) 2. sz., 157–162. p.

[267] Harry G. Frankfurt: *The Importance of What We Care About. Philosophical Essays*. 1998, Cambridge University Press.

[268] Harry G. Frankfurt – Christine M. Korsgaard – Michael E. Bratman – Meir Dan-Cohen: *Taking Ourselves Seriously. Getting it Right*. 2006, Stanford University Press.

[269] Marcus R. Frean: The prisoner's dilemma without synchrony. *Proc. R. Soc. B*, 257. évf. (1994), 75–79. p.

[270] James W. Friedman: A non-cooperative equilibrium for supergames. *The Review of Economic Studies*, 38. évf. (1971. Jan) 1. sz., 1–12. p.

[271] Frivaldszky János: *Klasszikus természetjog és jogfilozófia*. 2007, Szent István Társulat.

[272] Frivaldszky János: Politikai barátság és jogi viszony. *Iustum Aequum Salutare*, V. évf. (2009) 4. sz., 5–52. p.

[273] Frivaldszky János – Karácsony András: Jog, jogosultság, erőszak – jogfilozófiai kérdésfeltevések jhering tanai nyomán. *Jogelméleti Szemle*, 2010. június. 2. sz. URL http://jesz.ajk.elte.hu/frivaldszky42.html.

[274] Drew Fudenberg – Eric Maskin: Evolution and cooperation in noisy repeated games. *The American Economic Review*, 80. évf. (1990) 2. sz., 274–279. p.

[275] Fülöp Endre: Az etikai egoizmus lehetőségéről. *Phronesis*, 1. évf. (2007) 1. sz., 47–54. p.

[276] Wulf Gaertner: *A Primer in Social Choice Theory*. Oxford, 2006, Oxford University Press.

[277] David Gale – Lloyd S. Shapley: College admissions and the stability of marriage. *American Mathematical Monthly*, 69. évf. (1962) 1. sz., 9–15. p.

[278] David Gale – Marilda Sotomayor: Ms. machiavelli and the stable matching problem. *American Mathematical Monthly*, 92. évf. (1985) 4. sz., 261–268. p.

[279] Francis Galton: Vox populi. *Nature*, 75. évf. (1907), 450–451. p.

[280] Joshua S. Gans – Michael Smart: Majority voting with single-crossing preferences. *Journal of Public Economics*, 59. évf. (1996), 219–237. p.

[281] Garai Zsolt: Partikularista érvek az értékek realizmusa mellett. *Világosság*, 44. évf. (2003) 5-6. sz., 123–134. p.

[282] A. Gardner – S.A. West: Spite. *Current Biology*, 16. évf. (2006) 17. sz., R662–R664. p.

[283] John Geanakoplos: Three brief proofs of arrow's impossibility theorem. *Economic Theory*, 26. évf. (2005), 211–215. p.

[284] Jonathan Gelati – Guido Governatori – Antonino Rotolo – Giovanni Sartor: Declarative Power, Representation, and Mandate. a Formal Analysis. In T.J.M. Bench-Capon – A. Daskalopulu – R.G.F. Winkels (szerk.): *Legal Knowledge and Information Systems. Jurix 2002: The Fifteenth Annual Conference* (konferenciaanyag). 2002, IOS Press, 41–52. p.

[285] Nicholas Georgescu-Roegen: The pure theory of consumer's behavior. *Quarterly Journal of Economics*, 50. évf. (1936) 3. sz., 545–593. p.

[286] Nicholas Georgescu-Roegen: Choice, expectations, and measurability. *Quarterly Journal of Economics*, 68. évf. (1954) 2. sz., 503–534. p.

[287] Edmund Gettier: Is justified true belief knowledge? *Analysis*, 23. évf. (1963) 6. sz., 121–123. p.

[288] Allan Gibbard: Social choice and the arrow conditions (1968-69). mimeograph, 1969. URL http://www+personal.umich.edu/~gibbard/Gibbard1968+Social+Choice+Arrow+Conditions.pdf.

[289] Allan Gibbard: Manipulation of voting schemes: A general result. *Econometrica*, 41. évf. (1973), 587–601. p.

[290] Allan Gibbard – Aanund Hylland – John A. Weymark: Arrows theorem with a fixed feasible alternative. *Social Choice and Welfare*, 4. évf. (1987), 105–115. p.

[291] Margaret Gilbert: Modelling collective belief. *Synthese*, 73. évf. (1987) 1. sz., 185–204. p.

[292] Herbert Gintis: Strong reciprocity and human sociality. *Journal of Theoretical Biology*, 206. évf. (2000) 2. sz., 169–179. p.

[293] Herbert Gintis: The hitchhiker's guide to altruism: Gene-culture coevolution, and the internalization of norms. *Journal of Theoretical Biology*, 220. évf. (2003), 407–418. p.

[294] Herbert Gintis – Samuel Bowles – Robert Boyd – Ernst Fehr: Explaining altruistic behavior in humans. *Evolution and Human Behavior*, 24. évf. (2003), 153–172. p.

[295] Herbert Gintis – Samuel Bowles – Robert Boyd – Ernst Fehr (szerk.): *Moral Sentiments and Material Interests: On the Foundations of Cooperation in Economic Life*. 2005, MIT Press.

[296] Herbert Gintis – Eric Alden Smith – Samuel Bowles: Costly signaling and cooperation. *J. theor. Biol.*, 213. évf. (2001), 103–119. p.

[297] Patrick Girard: *Modal logic for belief and preference change*. Illc dissertation series ds-2008-04 (Institute for Logic, Language and Computation Universiteit van Amsterdam). 2008.

[298] U. Gneezy – A. Rustichini: A fine is a price. *The Journal of Legal Studies*, 29. évf. (2000) 1. sz., 1–18. p.

[299] Lou Goble – John-Jules Ch. Meyer (szerk.). *Deontic Logic and Artificial Normative Systems. 8th International Workshop on Deontic Logic in Computer Science, DEON 2006* (konferenciaanyag). Springer, 2006.

[300] Alan H. Goldman: The paradox of punishment. *Philosophy and Public Affairs*, 9. évf. (1979) 1. sz., 42–58. p.

[301] Alan H. Goldman: *Reasons from Within. Desires and Values*. 2009, Oxford University Press.

[302] Alvin I. Goldman: *A Theory of Human Action*. 1970, Prentice-Hall.

[303] Jonathan Lamb Gorman: A problem in the justification of democracy. *Analysis*, 39. évf. (1978) 1. sz., 46–50. p.

[304] Alvin W. Gouldner: The norm of reciprocity: A preliminary statement. *American Sociological Review*, 25. évf. (1960), 161–78. p.

[305] Umberto Grandi – Ulle Endriss: First-order logic formalisation of arrow's theorem. In *Logic, Rationality, and Interaction*. 2009, Springer, 133–146. p.

[306] Umberto Grandi – Ulle Endriss: First-order logic formalisation of impossibility theorems in preference aggregation. *Journal of Philosophical Logic*, 42. évf. (2013), 595–618. p.

[307] Joshua D. Greene – Leigh E. Nystrom – Andrew D. Engell – John M. Darley – Jonathan D. Cohen: The neural bases of cognitive conflict and control in moral judgment. *Neuron*, 44. évf. (2004. October), 389–400. p.

[308] P. S. Greenspan: Conditional oughts and hypothetical imperatives. *The Journal of Philosophy*, 22. évf. (1975) 10. sz., 259–276. p.

[309] Donald W. Griesinger – James W. Livingston: Toward a model of interpersonal motivation in experimental games. *Behavioral Science*, 18. évf. (1973) 3. sz., 173–188. p.

[310] T. Grüne-Yanoff – S. O. Hansson: *Preference Change: Approaches from Philosophy, Economics and Psychology*. 2009, Springer.

[311] Till Grüne-Yanoff: Why don't you want to be rich? preference explanation on the basis of causal structure. In J. K. Campbell – M. O'Rourke – Silverstein (szerk.): *Topics in Contemporary Philosophy: Explanation and Causation*. 2007, MIT Press, 217–240. p.

[312] Gulyás Attila: A méltányosságelmélet alapjai. modellek és nézõpontok. *Közgazdasági Szemle*, LIV. évf. (2007. február), 167–183. p.

[313] Gulyás Attila: Diktátor a barátom? a barátság és a méltányosság kapcsolata. *Közgazdasági Szemle*, LVIII. évf. (2011), 430–444. p.

[314] Johan E. Gustafsson: A money-pump for acyclic intransitive preferences. *Dialectica*, 64. évf. (2010) 2. sz., 251–257. p.

[315] Werner Güth – Rolf Schmittberger – Bernd Schwarze: An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior and Organization*, 3. évf. (1982) 4. sz., 367–388. p.

[316] Jürgen Habermas: *A kommunikatív cselekvés elmélete*. Budapest, 2011, Gondolat Kiadó.

[317] Ian Hacking: *Rewriting the Soul. Multiple Personality and the Sciences of Memory*. 1995, Princeton University Press.

[318] Ian Hacking: Kinds of people: Moving targets. In *Proceedings of the British Academy* (konferenciaanyag), 151. köt. 2007, 285–318. p.

[319] Sören Halldén: *On the Logic of 'Better'*. Lund, 1957, Cwk Gleerup.

[320] J. Y. Halpern: Should knowledge entail belief? *Journal of Philosophical Logic*, 25. évf. (1996), 483–494. p.

[321] J. Y. Halpern – R. Pucella: Dealing with logical omniscience: Expressiveness and pragmatics. *Artificial Intelligence*, 175. évf. (2011), 220–235. p.

[322] W. D. Hamilton: The genetical evolution of social behavior i. *Journal of Theoretical Biology*, 7. évf. (1964), 1–16. p.

[323] W. D. Hamilton: The genetical evolution of social behavior ii. *Journal of Theoretical Biology*, 7. évf. (1964), 17–52. p.

[324] W. D. Hamilton: Selfish and spiteful behaviour in an evolutionary model. *Nature*, 288. évf. (1970. 12) 5277. sz., 1218–1220. p.

[325] *Hammurabi törvényei*. 1911, Erdélyi Múzeum-Egyesület Jog- és Társadalomtudományi Szakosztálya.

[326] J. Hansen – G. Pigozzi – L. van der Torre: Ten philosophical problems in deontic logic. In G. Boella – L. van der Torre – H. Verhagen (szerk.): *Normative Multiagent Systems* (konferenciaanyag), 07122. köt. 2007.

[327] Jörg Hansen: *Imperatives and Deontic Logic. On the Semantic Foundations of Deontic Logic*. PhD értekezés (Der Fakultat für Sozialwissenschaften und Philosophie der Universitat Leipzig). 2008.

[328] Bengt Hansson: Choice structures and preference relations. *Synthese*, 18. évf. (1968) 443-458. sz.

[329] Bengt Hansson: Fundamental axioms for preference relations. *Synthese*, 18. évf. (1968), 423–442. p.

[330] Bengt Hansson: d'ẑźan Analysis of Some Deontic Logics. *Noûs*, 3. évf. (1969) 4. sz., 373–398. p.

[331] Bengt Hansson: Deontic logic and different levels of generality. *Theoria*, 36. évf. (1970. December) 3. sz., 241–248. p.

[332] S.O. Hansson: Defining 'good' and 'bad' in terms of 'better'. *Notre Dame of Journal of Formal Logic*, 31. évf. (1990), 136–149. p.

[333] Sven Ove Hansson: A new semantical approach to the logic of preference. *Erkenntnis (1975-)*, 31. évf. (1989) 1. sz., 1–42. p.

[334] Sven Ove Hansson: The revenger's paradox. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 61. évf. (1991) 3. sz., 301–305. p.

[335] Sven Ove Hansson: Changes in preference. *Theory and Decision*, 38. évf. (1995) 1. sz.

[336] Sven Ove Hansson: Preference logic. In DM. Gabbay – F. Guenthner (szerk.): *Handbook of Philosophical Logic*. 4. köt. 2001, Kluwer Academic Publishers, 319–393. p.

[337] Sven Ove Hansson: *The Structure of Values and Norms*. 2001, Cambridge University Press.

[338] Garret Hardin: Discriminating Altruism. *Zygon*, 17. évf. (1982) 2. sz., 163–186. p.

[339] Garrett Hardin: The tragedy of the commons. *Science*, 162. évf. (1968), 1243–1248. p.

[340] Russel Hardin: Collective action as an agreeable n-prisoners' dilemma. *Behavioral Science*, 16. évf. (1971), 472–481. p.

[341] Gilbert Harman: *Explaining Value and Other Essays in Moral Philosophy*. 2000, Clarendon Press.

[342] John C. Harsanyi: Rational choice models of political behavior vs. funcionalist and conformist theories. *World Politics*, 21. évf. (1969), 513–538. p.

[343] John C. Harsanyi: Advances in understanding rational behavior. In *Essays on Ethics, Social Behavior, and Scientific Explanation*. Dordrecht, 1976, Reidel, 89–117. p.

[344] H. L. A. Hart: Definition and theory in jurisprudence. In *Essays in Jurisprudence and Philosophy*. 1983, Clarendon Press, 21–48. p.

[345] H.L.A. Hart: Bentham on legal rights. In A.W.B. Simpson (szerk.): *Oxford Essays in Jurisprudence, Second Series*. Oxford, 1973, Clarendon Press, 171–201. p.

[346] H.L.A. Hart: *A jog fogalma*. Budapest, 1995, Osiris.

[347] H.L.A. Hart: Felelősség. In Zsolt Krokovay (szerk.): *Felelősség*. 2006, L'Harmattan, 51–70. p.

[348] Hegyi beszéd, 2013. URL `http://szentiras.hu/SZIT/Mt5`.

[349] Héjjas István: *A keleti vallások filozófiája és világképe*. 2007, ANNO Kiad.

[350] Carl G. Hempel: Aspects of scientific explanation. In *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. New York, 1965, Free Press, 466–467. p.

[351] Joseph Henrich – Robert Boyd – Samuel Bowles – Colin Camerer – Ernst Fehr – Herbert Gintis (szerk.): *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies*. 2004, Oxford University Press.

[352] A. L. Herman: A solution to the paradox of desire in buddhism. *Philosophy East and West*, 29. évf. (1979) 1. sz., 91–94. p.

[353] Henning Herrestad – Christen Krogh: The right direction.

[354] Mathieu Hilgers: Habitus, freedom, and reflexivity. *Theory and Psychology*, 19. évf. (2009) 6. sz., 728–755. p.

[355] F.A. Hindriks: Constitutive rules, language, and ontology. *Erkenntnis*, 71. évf. (2009), 253–275. p.

[356] Albert O. Hirschmann: *Az érdekek és a szenvedélyek. Politikai érvek a kapitalizmus mellett annak győzelme előtt*. 1998, Jószöveg Műhely Kiadó.

[357] Jack Hirshleifer: Privacy: Its origin, function, and future. *The Journal of Legal Studies*, 9. évf. (1980) 4. sz., 649–664. p.

[358] Jack Hirshleifer: The expanding domain of economics. *The American Economic Review*, 75. évf. (1985. Dec) 6. sz., 53–68. p. Centennial Essays and 1985 Survey of Members.

[359] Jerry R. Hobbs: Ontological promiscuity. In *Proceedings, 23rd Annual Meeting of the Association for Computational Linguistics* (konferenciaanyag), 23. köt. Chicago, Illinois, 1985, Association for Computational Linguistics, 61–69. p. URL `www.isi.edu/~hobbs/op+acl85.pdf`.

[360] Wesley Newcomb Hohfeld: Alapvető jogi fogalmak a bírói érvelésben. In Miklós Szabó – Csaba Varga (szerk.): *Jog és nyelv*. 2000, kiadó nélkül, 59–96. p.

[361] Wesley H. Holliday: Epistemic logic and epistemology. In S.O. Hansson – V.F. Hendricks (szerk.): *Handbook of Formal Philosophy*. 2013, Springer.

[362] Martin Hollis: Penny pinching and backward induction. *The Journal of Philosophy*, 88. évf. (1991) 9. sz., 473–488. p.

[363] John F. Horty: *Agency and Deontoic Logic*. 2001, Oxford University Press.

[364] John F. Horty – Nuel Belnap: The deliberative stit: A study of action, omission, ability, and obligation. *Journal of Philosophical Logic*, 24. évf. (1995), 583–644. p.

[365] Harold Hotelling: Stability in competition. *The Economic Journal*, 39. évf. (1929), 41–57. p.

[366] H. S. Houthakker: Revealed preference and the utility function. *Economica, New Series*, 17. évf. (1950. May) 66. sz., 159–174. p.

[367] G. E. Hughes – M. J. Cresswell: *A New Introduction to Modal Logic*. 1996, Routledge. ISBN 0-415-12599-5.

[368] J. Hunt – R. Brooks – M. Jennions – M. Smith – C. Bentsen – L. Bussière: High-quality male field crickets invest heavily in sexual display but die young. *Nature*, 432. évf. (2004) 7020. sz., 1024–1027. p.

[369] Ken ichi Inada: A note on the simple majority decision rule. *Econometrica*, 32. évf. (1964) 4. sz., 525–531. p.

[370] Ken ichi Inada: The simple majority decision rule. *Econometrica*, 37. évf. (1969) 3. sz., 490–506. p.

[371] Ken-Ichi Inada: Alternative incompatible conditions for a social welfare function. *Econometrica*, 23. évf. (1955) 4. sz., 396–399. p.

[372] Index.hu: Szándékos öngól a koreai ligában. URL `http://sportgeza.hu/futball/2013/07/05/koreai_ongol/`.

[373] Jakab András: A norma szerkezetének vizsgálata. *JESZ*, 2001. 4. sz. URL `http://jesz.ajk.elte.hu/jakab8.html`.

[374] Susan James: Szenvedélyek, vágyak és a cselekvés magyarázata. *Magyar Filozófiai Szemle*, 39. évf. (1995) 5-6. sz., 743–757. p. 'Desires, Passions and the Explanation of Action' in S. Gaukroger (ed.), The Soft Underbelly of Reason (Routledge, 1998).

[375] Susan James: *Passion and Action. The Emotions in Seventeenth-Century Philosophy*. 1997, Clarendon Press.

[376] Janky Béla: *Szolidaritás és jóléti preferenciák*. Philosophiae Doctores sorozat. Budapest, 2005, Akadémiai Kiadó.

[377] Richard Jeffrey: *The Logic of Decision*. Chicago, 1983, University of Chicago.

[378] R. E. Jennings: Can there be a natural deontic logic? *Synthese*, 65. évf. (1985) 2. sz., 257–273. p.

[379] Ray E. Jennings: Preference and choice as logical correlates. *Mind*, 76. évf. (1967) 304. sz., 556–567. p.

[380] Keith Jensen: Punishment and spite, the dark side of cooperation. *Phil. Trans. R. Soc. B*, 365. évf. (2010), 2635–2650. p.

[381] Jan C. Joerden: Deontological square, hexagon, and decagon: A deontic framework for supererogation. *Logica Universalis*, 2012. February., 1–16. p.

[382] David B. Johnson: *A közösségi döntések elmélete*. 1999, Osiris.

[383] R. A. Johnstone – R. Bshary: Evolution of spite through indirect reciprocity. *Proc. R. Soc. Lond. B*, 271. évf. (2004), 1917–1922. p.

[384] Thomas Joiner: *Why People Die by Suicide*. 2005, Harvard University Press.

[385] A. J. Jones – M. J. Sergot: A formal characterisation of institutionalised power. *Journal of the IGPL*, 1996. 4. sz., 429–445. p.

[386] A.J.I. Jones – M.J. Sergot: On the characterization of law and computer systems: The normative systems perspective. In J-J.Ch. Meyer – R.J. Wieringa (szerk.): *Deontic Logic in Computer Science*. 1993, Wiley.

[387] Andrew J.I. Jones – Ingmar Pörn: Ideality, sub-ideality and deontic logic. *Synthese*, 65. évf. (1985), 275–290. p.

[388] Jörgen Jörgensen: Imperatives and logic. *Erkenntnis*, 7. évf. (1937/193) 8. sz., 288–296. p.

[389] Dan Kahan – Ellen Peters – Erica Cantrell Dawson – Paul Slovic: Motivated numeracy and enlightened self-government. Jelentés, 2013, SSRN.

[390] Walter J. Kamba: Legal theory and hohfeld's analysis of a legal right. *Judicial Review*, 19. évf. (1974), 249–262. p.

[391] Michihiro Kandori: Social Norms and Community Enforcement. *The Review of Economic Studies*, 59. évf. (1992. January) 1. sz., 63–80. p.
URL `http://www.jstor.org/pss/2297925`.

[392] Stig Kanger: New foundations for ethical theory. In R. Hilpinen (szerk.): *Deontic Logic: Introductory and Systematic Readings*. 1971, Reidel, 36–58. p.

[393] Stig Kanger: Law and logic. *Theoria*, 38. évf. (1972), 105–132. p.

[394] Stig Kanger – Helle Kanger: Rights and parliamentarism. *Theoria*, 32. évf. (1966), 85–115. p.

[395] Hans Kelsen: *Tiszta jogtan*. Budapest, 2001, Rejtjel Kiadó.

[396] Robert O. Keohane: *After Hegemony. Cooperation and Discord in the World Political Economy*. 1984, Princeton University Press.

[397] Elias L. Khalil: What is altruism? *Journal of Economic Psychology*, 25. évf. (2004), 97–123. p.

[398] Kiss Balázs: Michel foucault hatalomfelfogásáról. *Politikatudományi Szemle*, 1994. 1. sz., 43–68. p.

[399] Bettina Klaus: Population-monotonicity and separability for economies with single-dipped preferences and the assignment of an indivisible object. *Economic Theory*, 17. évf. (2001), 675–692. p.

[400] Bettina Klaus – Hans Peters – Ton Storcken: Strategy-proof division of a private good when preferences are single-dipped. *Economics Letters*, 55. évf. (1997), 339–346. p.

[401] Stephen Cole Kleene: *Mathematical Logic*. New York, 2002, Dover.

[402] William Kneale – Martha Kneale: *A logika fejlődése*. Budapest, 1987, Gondolat Kiadó.

[403] Joshua Knobe: Intentional action and side effects in ordinary language. *Analysis*, 63. évf. (2003), 190–193. p.

[404] Albert Kocourek: The hohfeld system of fundamental legal concepts. *Illinois Law Review*, 15. évf. (1920) 1. sz., 24–39. p.

[405] Kóczy László Á.: Központi felvételi rendszerek. taktikázás és stabilitás. *Közgazdasági Szemle*, 56. évf. (2009) 5. sz., 422–442. p.

[406] Kóczy László Á.: A magyarországi felvételi rendszerek sajátosságai. *Közgazdasági Szemle*, 57. évf. (2010) 2. sz., 142–164. p.

[407] David H. Koehler: Vote trading and the voting paradox: A proof of logical equivalence. *The American Political Science Review*, 69. évf. (1975) 3. sz., 954–960. p.

[408] S.-C. Kolm: The buddhist theory of "no-self". In Jon Elster (szerk.): *The Multiple Self*. 1986, Cambridge University Press, 233–265. p.

[409] S. S. Komorita – David D. Meek: Generality and validity of some theories of coalition formation. *Journal of Personality and Social Psychology*, 36. évf. (1978) 4. sz., 392–404. p.

[410] L. A. Kornhauser – L. G. Sager: The one and the many: Adjudication in collegial courts. *California Law Review*, 81. évf. (1993), 1–59. p.

[411] Körösényi András: Lehangolótudomány, vagy a politikaművészete? vezetés, manipuláció és demokrácia. *Politikatudományi Szemle*, XVIII. évf. (2009) 3. sz., 33–56. p.

[412] Christine M. Korsgaard: *The Constitution of Agency. Essays on Practical Reason and Moral Psychology*. Christine M. Korsgaard, Oxford University Press.

[413] Christine M. Korsgaard – G.A. Cohen – Raymond Geuss – Thomas Nagel – Bernard Williams: *The Sources of Normativity*. 1996, Oxford University Press.

[414] Máté Kovács: Kinyilvánított preferencia és racionalitás. *Közgazdasági Szemle*, LVI. évf. (2009) 6. sz., 546–562. p.

[415] David Kraines – Vivian Kraines: Pavlov and the Prisoner's Dilemma. *Theory and Decision*, 26. évf. (1989), 47–79. p.

[416] David Kraines – Vivian Kraines: Learning to cooperate with pavlov an adaptive strategy for the iterated prisoner's dilemma with noise. *Theory and Decision*, 35. évf. (September, 1993) 2. sz., 107–150. p.

[417] Gerald Kramer: Strategic voting. *Journal of Mathematical Sociology*, 6. évf. (1972).

[418] Gerald H. Kramer: On a class of equilibrium conditions for majority rule. *Econometrica*, 41. évf. (1973) 2. sz., 285–297. p.

[419] Matthew H. Kramer: Rights without trimmings. In M. Kramer – N. Simmonds – H. Steiner (szerk.): *A Debate over Rights. Philosophical Enquiries*. 2000, Oxford University Press, 7–111. p.

[420] Matthew H. Kramer – N.E. Simmonds – Hillel Steiner: *A Debate over Rights. Philosophical Enquiries*. 2000, Oxford University Press.

[421] Angelika Kratzer: What ‘must’ and ‘can’ must and can mean. *Linguistics and Philosophy*, 1. évf. (1977), 337–355. p.

[422] D. Krebs – D. Miller: Altruism and aggression. In G. Lindzey – E. Aronson (szerk.): *Handbook of Social Psychology*. 1985, Random House, 1–71. p.

[423] Dennis L. Krebs: Altruism: An examination of the concept and a review of the literature. *Psychological Bulletin*, 73. évf. (1970) 4. sz., 258–302. p.

[424] Dennis L. Krebs: Altruism and egoism: A false dichotomy? *Psychological Inquiry*, 2. évf. (1991) 2. sz., 137–139. p.

[425] D. Kreps – R. Wilson: Reputation and imperfect information. *Journal of Economic Theory*, 27. évf. (1982), 253–279. p.

[426] Christen Krogh – Henning Herrestad: Getting personal. some notes on the relationship between personal and impersonal obligation. In Mark A. Brown – Jose Carmo (szerk.): *Deontic Logic, Agency and Normative Systems. International Workshop on Deontic Logic in Computer Science*. 1996, Springer, 134–153. p.

[427] Christen Krogh – Henning Herrestad: Hohfeld in cyberspace and other applications of normative reasoning in agent technology. *Artificial Intelligence and Law March*, 7. évf. (1999) 1. sz., 81–96. p.

[428] S. L. Kuhn – M. C. Stiner: What’s a mother to do? the division of labor among neandertals and modern humans in eurasia. *Current Anthropology*, 47. évf. (2006) 6. sz., 953–963. p.

[429] L. Aczél Petra: Erős szavak. a bizonyítás mint a szöveg tulajdonsága. *Világosság*, 44. évf. (2003) 11-12. sz., 171–178. p.

[430] David Lane: *Hierarchy in Natural and Social Sciences*. 4. köt. Hierarchy, Complexity, Society fejezet. 2006, Springer, 81–119. p.

[431] Maria Knight Lapinski – Rajiv N. Rimal: An explication of social norms. *Communication Theory*, 15. évf. (2005) 2. sz., 127–147. p.

[432] Irving H. Lavalle – Peter C. Fishburn: Lexicographic state-dependent subjective expected utility. *Journal of Risk and Uncertainty*, 4. évf. (1991) 3. sz., 251–269. p.

[433] Irving H. Lavalle – Peter C. Fishburn: State-independent subjective expected lexicographic utility. *Journal of Risk and Uncertainty*, 5. évf. (1992) 3. sz., 217–240. p.

[434] Marc Lavoie: Post Keynesian consumer theory: Potential synergies with consumer research and economic psychology. *Journal of Economic Psychology*, 25. évf. (2004), 639–649. p.

[435] Keith Lehrer – Carl Wagner: Intransitive indifference: The semi-order problem. *Synthese*, 65. évf. (1985) 2. sz., 249–256. p.

[436] E. J. Lemmon: Moral dilemmas. *The Philosophical Review*, 71. évf. (1962) 2. sz., 139–158. p.

[437] Edward Lemmon: Deontic logic and the logic of imperatives. *Logique et Analyse*, 8. évf. (1965), 39–71. p.

[438] David K. Levine: Modeling altruism and spitefulness in experiments. *Review of Economic Dynamics*, 1. évf. (1998), 593–622. p.

[439] Michael E. Levine – Charles R. Plott: Agenda influence and its implications. *Virginia Law Review*, 63. évf. (1977) 4. sz., 561–604. p.

[440] D. Lewis: A problem about permission. In E. Saarine (szerk.): *Essays in Honour of Jaakko Hintikka*. Dordrecht, 1979, D. Reidel, 163–175. p.

[441] David Lewis: Desire as belief. *Mind, New Series*, 97. évf. (1988) 387. sz., 323–332. p.

[442] David Lewis: Desire as belief ii. *Mind, New Series*, 105. évf. (1996) 418. sz., 303–313. p.

[443] David Lewis: Semantic analyses for dyadic deontic logic. In *Papers in Ethics and Social Philosophy*. 2000, Cambridge University Press, 5–19. p.

[444] David K. Lewis: *Convention. A Philosophical Study*. Cambridge (MA), 1969, Harvard University Press.

[445] Amir N. Licht: Games commissions play: 2x2 games of international securities regulation. *The Yale Law Journal*, 24. évf. (1999) 61. sz., 61–125. p.

[446] Sarah Lichtenstein – Paul Slovic (szerk.): *The Construction of Preference*. 2006, Cambridge University Press.

[447] Wim B. G. Liebrand: A classification of social dilemma games. *Simulation & Games*, 14. évf. (1983) 2. sz., 123–138. p.

[448] Wim B. G. Liebrand: The effect of social motives, communication and group size on behaviour in an n-person multi-stage mixed-motive game. *European Journal of Social Psychology*, 14. évf. (1984), 239–264. p.

[449] Lars Lindahl: *Position and Change*. 1977, Springer.

[450] Siegwart Lindenberg – Bruno S. Frey: Alternatives, frames, and relative prices: A broader view of rational choice theory. *Acta Sociologica*, 36. évf. (1993) 3. sz., 191–205. p.

[451] Christian List: The theory of judgment aggregation: An introductory review. *Synthese*, 187. évf. (2012) 1. sz., 179–207. p.

[452] Christian List – Philip Petit: Aggregating sets of judgments : an impossibility result. *Economics and philosophy*, 18. évf. (2002) 1. sz., 89–110. p.

[453] Christian List – Philip Pettit: Aggregating sets of judgement: Two impossibility results compared. *Synthese*, 140. évf. (2004), 207–235. p.

[454] Fenrong Liu: Von wright's "the logic of preference" revisited. *Synthese*, 175. évf. (2010) 1. sz., 69–88. p.

[455] Fenrong Liu: *Reasoning about Preference Dynamics*. 2011, Springer.

[456] Barry Loewer – Marvin Belzer: Dyadic Deontic Detachment. *Synthese*, 54. évf. (1983), 295–318. p.

[457] Alessio Lomuscio – Donald Nute (szerk.): *Deontic Logic in Computer Science, 2004*. 2004, Springer.

[458] Graham Loomes – Robert Sugden: Regret theory: An alternative of rational choice under uncertainty. *The Economic Journal*, 92. évf. (1982) 368. sz., 805–824. p.

[459] E. J. Lowe: *Personal Agency. The Metaphysics of Mind and Action*. 2008, Oxford University Press.

[460] R. Duncan Luce: Semiorders and a theory of utility discrimination. *Econometrica*, 24. évf. (1956) 2. sz., 178–191. p.

[461] David Lyons: The correlativity of rights and duties. *Noûs*, 4. évf. (1970) 1. sz., 45–55. p.

[462] John Macnamara – Gonzalo E. Reyes (szerk.): *The logical foundations of cognition*. 1994, Oxford University Press.

[463] Magyarkúti Gyula: *A racionalitás fogalmának axiomatikus megközel´ése*. PhD értekezés (BKÁE Közgazdaságtani PhD-program). 2001.

[464] D. Makinson: On the formal representation of rights relations. *Journal of Philosophical Logic*, 15. évf. (1986), 403–425. p.

[465] David Makinson: Stenius' approach to disjunctive permission. *Theoria*, 50. évf. (1984) 2-3. sz., 138–147. p.

[466] David Makinson: On attributing rights to all peoples: Some logical questions. *Law and Philosophy*, 8. évf. (1989) 1. sz., 53–63. p.

[467] David Makinson: On a fundamental problem of deontic logic. In Paul McNamara – Henry Prakken (szerk.): *Norms, Logics and Information Systems. New Studies in Deontic Logic and Computer Science*. 1999, IOS Press, 29–53. p.

[468] Mala József: Arrow-típusú lehetetlenségi tételek. *Szigma*, XXXVIII. évf. (2007) 3-4. sz., 77–110. p.

[469] Radu Mardare – Corrado Priami: Dynamic-epistemic spatial logic. DIT-06-010. Jelentés, 2006, DEPARTMENT OF INFORMATION AND COMMUNICATION TECHNOLOGY UNIVERSITY OF TRENTO.

[470] E. Mares – P. McNamara: Supererogation in deontic logic: Metatheory for dwe and some close neighbours. *Studia Logica*, 57. évf. (1997), 397–415. p.

[471] Howard Margolis: Pivotal voting. *Journal of Theoretical Politics*, 13. évf. (2001) 1. sz., 111–116. p.

[472] Howard Margolis: d'żźpivotal voting and the emperor's new clothes. *Social Choice and Welfare*, 19, No. 1 (), pp. évf. (2002) 1. sz., 95–111. p.

[473] Frank W. Marlowe – J. Colette Berbesque – Clark Barrett – Alexander Bolyanatz – Michael Gurven – David Tracer: The 'spiteful' origins of human cooperation. *Proc. R. Soc. B*, 278. évf. (2011), 2159–2164. p.

[474] Marosán György: A proszociális értékek evolúciója a játékelméleti kísérletek tükrében. *Közgazdasági Szemle*, LIV. évf. (2007. augusztus), 716–733. p.

[475] Marcel Mauss: Techniques of the body. *Economy and Society*, 2. évf. (1973) 1. sz., 70–88. p.

[476] Kenneth O. May: A set of independent necessary and sufficient conditions for majority decision. *Econometrica*, 20. évf. (1952) 4. sz., 680–684. p.

[477] John Maynard Smith – G. R. Price: The logic of animal conflict. *Nature*, 246. évf. (1973) 5427. sz., 15–18. p.

[478] Ernst Mayr: Cause and effect in biology. *Science, New Series*, 134. évf. (1961) 3489. sz., 1501–1506. p.

[479] Ernst Mayr: Teleological and teleonomic, a new analysis. *Boston Studies in the Philosophy of Science*, 14. évf. (1974), 91–117. p.

[480] Ernst Mayr: The idea of teleology. *Journal of the History of Ideas*, 53. évf. (1992) 1. sz., 117–135. p.

[481] R. P. McArthur: Anderson's deontic logic and relevant implication. *Notre Dame Journal of Formal Logic*, 22. évf. (1981), 145–154. p.

[482] Charles G. McClintock: Social motivation – a set of propositions. *Behavioral Science*, 17. évf. (1972) 5. sz., 438–454. p.

[483] Kris McDaniel – Ben Bradley: Desires. *Mind, New Series*, 117. évf. (2008) 466. sz., 267–302. p.

[484] Joseph E. McElliskem: Affective and predatory violence: a bimodal classification system of human aggression and violence. *Aggression and Violent Behavior*, 10. évf. (2004), 1–30. p.

[485] R.D. McKelvey: Intransitivities in multidimensional voting models and some implications for agenda control. *Journal of Economic Theory*, 12. évf. (1976), 472–482. p.

[486] Richard D. McKelvey: General conditions for global intransitivities in formal voting models. *Econometrica*, 47. évf. (1979) 5. sz., 1086–1112. p.

[487] Richard D. McKelvey – Richard G. Niemi: A multistage game representation of sophisticated voting for binary procedures. *Journal of Economic Theory*, 18. évf. (1978), 1–22. p.

[488] Richard D. McKelvey – Peter C. Ordeshook: A decade of experimental research on spatial models of elections and committees. In James N. Enelow – Melvin J. Hinich (szerk.): *Advances in the spatial theory of voting*. 1990, Cambridge University Press, 99–144. p.

[489] Michael McKinsey: Levels of obligation. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 35. évf. (1979) 4. sz., 385–395. p.

[490] P. McNamara: Making room for going beyond the call. *Mind*, 105. évf. (1996) 419. sz.

[491] Paul McNamara: Doing well enough: Toward a logic for common-sense morality. *Studia Logica: An International Journal for Symbolic Logic*, 57. évf. (1996) 1. sz., 167–192. p.

[492] Alfred R. Mele: *Motivation and Agency*. 2003, Oxford University Press.

[493] David M. Messick – Charles G. McClintock: Motivational bases of choice in experimental games. *Journal of Experimental Social Psychology*, 4. évf. (1968), 1–2. p.

[494] Mészáros József – Szakadát István: *Választási eljárások – választási rendszerek*. A BME Szociológia Tanszék kiadványai sorozat, 2. köt. 1993, BME Szociológia Tanszék. URL `http://mycite.omikk.bme.hu/doc/88355.pdf`.

[495] Mészáros József – Szakadát István: *Magyarország Politikai Atlasza*. Budapest, 1995, Konrad Adenauer Stiftung.

[496] Mészáros József – Szakadát István: *Parlamenti képviselői helyek megoszlásának becslése közvélemény-kutatási adatok alapján*. Tárki társadalompolitikai tanulmányok sorozat, 6. köt. Budapest, 1998, TÁRKI.

[497] Mészáros József – Szakadát István: Választási eredmények területi megoszlása. In Kolosi Tamás – Tóth István György – Vukovich Györgyi (szerk.): *Társadalmi Riport*. Budapest, 1998, TÁRKI, 412–433. p.

[498] Mészáros József – Szakadát István: *Magyarország Politikai Atlasza 1998*. Budapest, 1999, Osiris Kiadó.

[499] Mészáros József – Szakadát István: Territorial distribution of election results. In Kolosi Tamás – Tóth István György – Vukovich Györgyi (szerk.): *Social Report 1998*. Budapest, 1999, TÁRKI, 409–430. p.

[500] Mészáros József – Szakadát István: A választási rendszer elemei. In Dezső Márta – Kukorelli István (szerk.): *Választástudományi tanulmányok*. 1999, OVB, 217–242. p.

[501] József Mészáros – István Szakadát: *Magyarország Politikai Atlasza 2004*. Budapest, 2005, Gondolat Kiadó.

[502] Meszerics Tamás: Stratégiai viselkedés és bizottsági döntés. *Közgazdasági Szemle*, XLIV. évf. (1997) 7-8. sz., 687–697. p.

[503] Meszerics Tamás: A manipuláció és a mögöttes normák. *Beszélő*, 8. évf. (2003) 11. sz., 18–25. p.

[504] J. J. C. Meyer: A different approach to deontic logic: Deontic logic viewed as a variant of dynamic logic. *Notre Dame Journal of Formal Logic*, 29. évf. (1988) 1. sz., 109–136. p.

[505] J.-J.Ch. Meyer – Wiebe van der Hoek: *Epistemic Logic for Computer Science and Artificial Intelligence*. Cambridge Tracts in Theoretical Computer Science sorozat, 41. köt. Cambridge, 1995, Cambridge University Press.

[506] E. Midlarsky – E. Kahana: Theories and concepts of altruism and helping. In E. Midlarsky – E. Kahana (szerk.): *Alruism in Later Life*. Library of Social Research sorozat, 8. köt. 1994, Sage, 11–40. p.

[507] Greg Miller: The roots of morality. *Science*, 320. évf. (2008. May), 734–737. p.

[508] Seumas Miller: *Social Action. A Teleological Account*. 2004, Cambridge University Press.

[509] Ruth Garrett Millikan: Thoughts without laws; cognitive science with content. *The Philosophical Review*, 95. évf. (1986) 1. sz., 47–80. p.

[510] Ruth Garrett Millikan: Bioszemantika. In Gergely Ambrus – Tamás Demeter – Gábor Forrai – János Tőzsér (szerk.): *Elmefilozófia*. 2008, L'Harmattan, 252–270. p.

[511] Jorge Moll – Frank Krueger – Roland Zahn – Matteo Pardini – Ricardo de Oliveira-Souza – Jordan Grafman: Human fronto–mesolimbic networks guide decisions about charitable donation. *PNAS*, 103. évf. (2006) 42. sz., 15623–15628. p.

[512] Dov Monderer – Dov Samet: Approximating common knowledge with common belief. *Games and Economic Behavior*, 1. évf. (1989), 170–190. p.

[513] Philippe Mongin: The doctrinal paradox, the discursive dilemma, and logical aggregation theory. *Theory and Decision*, 73. évf. (2012) 3. sz., 315–355. p.

[514] Kristen Renwick Monroe: A fat lady in a corset: Altruism and social theory. *American Journal of Political Science*, 38. évf. (1994) 4. sz., 861–893. p.

[515] Douglas G. Mook: Why can't altruism be selfish? *Psychological Inquiry*, 2. évf. (1991) 2. sz., 139–141. p.

[516] Alessio Moretti: Correcting mcnamara's "deontic dodecagon for dwe": through not, 2009.

[517] Alessio Moretti: *The Geometry of Logical Opposition*. PhD értekezés (University of Neuchâtel). 2009.

[518] Alessio Moretti: The geometry of standard deontic logic. *Logica Universalis*, 3. évf. (2009), 19–57. p.

[519] Alessio Moretti: The critics of paraconsistency and of many-valuedness and the geometry of oppositions. *Logic and Logical Philosophy*, 19. évf. (2010), 63–94. p.

[520] Alessio Moretti: Why the logical hexagon? *Logica Universalis*, 6. évf. (2012), 69–107. p.

[521] Pierre Le Morvan: Selfishness, altruism, and our future selves. *d'żźAustralasian Journal of Philosophy*, 87. évf. (2009) 3. sz., 409–424. p.

[522] Hervé Moulin: On strategy-proofness and single peakedness. *Public Choice*, 35. évf. (1980), 437–455. p.

[523] Hervé Moulin: Condorcet's principle implies the no-show paradox. *Journal of Economic Theory*, 45. évf. (1988), 53–64. p.

[524] Roger B. Myerson: Axiomatic derivation of scoring rules without the ordering assumption. *Social Choice and Welfare*, 12. évf. (1995), 59–74. p.

[525] Ernest Nagel: Functional explanations in biology. *The Journal of Philosophy*, 74. évf. (1977) 5. sz., 280–301. p.

[526] Ernest Nagel: Goal-directed processes in biology. *The Journal of Philosophy*, 74. évf. (1977) 5. sz., 261–279. p.

[527] Thomas Nagel: *The Possibility of Altruism*. 1978, Princeton University Press.

[528] Jan Narveson: *The Libertarian Idea*. 2001, Broadview Press.

[529] Héctor-Neri Casta neda: Intentions and the structure of intending. *The Journal of Philosophy*, 68. évf. (1971) 15. sz., 453–466. p.

[530] Klaus Nehring: Self-control through second-order preferences, 2006.

[531] Klaus Nehringa – Clemens Puppe: Abstract arrowian aggregation. *Journal of Economic Theory*, 145. évf. (2010), 467–494. p.

[532] Némedi Dénes: *Émile Durkheim*. Budapest, 1996, Áron Kiadó.

[533] Manfred Nermuth: Two-stage discrete aggregation: the ostrogorski paradox and related phenomena. *Social Choice and Welfare*, 9. évf. (1992) 2. sz., 99–116. p.

[534] Yew-kwang Ng: Sub-semiorder: A model of multidimensional choice with preference intransitivity. *Journal of Mathematical Psychology*, 16. évf. (1977) 1. sz., 51–59. p.

[535] M. Nowak – K. Sigmund: A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game. *Nature*, 1993. 364. sz., 56–58. p.

[536] M. A. Nowak – K. Sigmund: The alternating prisoner's dilemma. *Journal of Theoretical Biology*, 168. évf. (1994), 219–226. p.

[537] M. A. Nowak – K. Sigmund: Evolution of indirect reciprocity by image scoring. *Nature*, 1998. 393. sz., 573–577. p.

[538] Martin A. Nowak: Evolutionary dynamics of cooperation. In *Proceedings of the International Congress of Mathematicians* (konferenciaanyag). 2006, 1523–1540. p.

[539] Martin A. Nowak – Karl Sigmund: Tit for tat in heterogeneous populations. *Nature*, 355. évf. (1992. January), 250–253. p.

[540] Martin A. Nowak – Corina E. Tarnita – Edward O. Wilson: The evolution of eusociality. *Nature*, 466. évf. (2010) 7310. sz., 1057–1062. p.

[541] Robert Nozick: *Philosophical Explanations*. 1981, Harvard University Press.

[542] Hannu Nurmi: Compound majority paradoxes and proportional representation. *European Journal of Political Economy*, 13. évf. (1997), 443–4454. p.

[543] Brendan Nyhan – Jason Reifler: When corrections fail: The persistence of political misperceptions. *Political Behavior*, 32. évf. (2010) 2. sz., 303–330. p.

[544] Timothy O'Connor – Constantine Sandis: *A Companion to the Philosophy of Action*. 2010, Wiley-Blackwell.

[545] Óészaki mitológiai és hősi énekek: 2. a fenséges beszéde, 42–43. versszak, 2003. January. URL `http://mek.oszk.hu/00300/00377/`.

[546] Hisashi Ohtsuki – Yoh Iwasa: How should we define goodness? – reputation dynamics in indirect reciprocity. *Journal of Theoretical Biology*, 231. évf. (2004. November) 1. sz., 107–120. p.

[547] Hisashi Ohtsuki – Yoh Iwasa: The leading eight: Social norms that can maintain cooperation by indirect reciprocity. *Journal of Theoretical Biology*, 239. évf. (2006. April) 4. sz., 435–444. p.

[548] Hisashi Ohtsuki – Yoh Iwasa: Global analyses of evolutionary dynamics and exhaustive search for social norms that maintain cooperation by reputation. *Journal of Theoretical Biology*, 244. évf. (2007) 3. sz., 518 – 531. p. ISSN 0022-5193. URL `http://www.sciencedirect.com/science/article/pii/S0022519306003900`.

[549] Andrew Oldenquist: The possibility of selfishness. *American Philosophical Quarterly*, 17. évf. (1980) 1. sz., 25–33. p.

[550] Pamela Oliver: Rewards and Punishments as Selective Incentives. An Apex Game. *The Journal of Conflict Resolution*, 28. évf. (1980. May) 1. sz., 123–148. p.

[551] Pamela Oliver: Rewards and Punishments as Selective Incentives for Collective Action: Theoretical Ivestigations. *The American Journal of Sociology*, 85. évf. (1980. May) 6. sz., 1356–1375. p.

[552] Peter C. Ordeshook: *Game theory and political theory. An introduction*. 1986, Cambridge University Press.

[553] Orthmayr Imre: A sztoikus szenvtelenség. *Századvég*, 4. évf. (1999) 13. sz. URL http://www.c3.hu/scripta/szazadveg/13/orthmayr.htm.

[554] Orthmayr Imre: Erkölcsi normák. *Világosság*, 44. évf. (2003) 5-6. sz., 149–161. p.

[555] Orthmayr Imre: A társadalmi normák döntéselméleti és evolúciós magyarázata. *Szociológiai Szemle*, 14. évf. (2004) 3. sz., 3–22. p.

[556] Orthmayr Imre: Konfliktusok. *Világosság*, 48. évf. (2007) 4. sz., 71–89. p.

[557] Daniel Osherson – Scott Weinstein: Preference based on reasons. *The Review of Symbolic Logic*, 5. évf. (2012) 1. sz.

[558] Jorge M. Pacheco – Francisco C. Santos – Fabio A.C.C. Chalub: Stern-Judging: a Simple, Successful Norm which Promotes Cooperation under Indirect Reciprocity. *PLoS Computational Biology*, 2. évf. (2006. december) 12. sz., 1634–1638. p.

[559] Karthik Panchanathan – Rob Boyd: Human cooperation: Second-order free-riding problem solved? (reply). *Nature*, 437. évf. (2005. 09) 7058. sz., E8–E9. p. URL http://dx.doi.org/10.1038/nature04202.

[560] Karthik Panchanathan – Robert Boyd: Indirect reciprocity can stabilize cooperation without the second-order free rider problem. *Nature*, 2004. November 25. 432. sz., 499–502. p.

[561] Derek Parfit: Overpopulation and the quality of life. In J. Ryberg – T. Tännsjö (szerk.): *The Repugnant Conclusion. Essays on Population Ethics*. 2004, Springer, 7–22. p.

[562] Talcott Parsons – Edward A. Shils: *Toward a General Theory of Action. Theoretical Foundations for the Social Sciences*. New Brunswick (USA) and London (UK), 2007 (1951), Transactions Publisher.

[563] Terence Parsons: The traditional square of opposition. in stanford encyclopedia of philosophy. URL http://plato.stanford.edu/entries/square/.

[564] M. Pauly – M. van Hees: Logical constraints on judgment aggregation. *Journal of Philosophical Logic*, 35. évf. (2006) 6. sz., 569–585. p.

[565] A. Pepitone: Toward a normative and comparative biocultural social psychology. *Journal of Personality and Social Psychology*, 34. évf. (1976) 4. sz., 641–653. p.

[566] Sam Perlo-Freeman: The topology of conflict and co-operation, 2006. May.

[567] Michael Perloff: Stit and the language of agency. *Synthese*, 86. évf. (1991) 3. sz., 379–408. p.

[568] Thomas D. Perry: A paradigm of philosophy: Hohfeld on legal rights. *American Philosophical Quarterly*, 14. évf. (1977) 1. sz., 41–50. p.

[569] Pete Krisztián: A társadalmi konvenciók természetéről. *Infotárs*, 2008., 73–99. p.

[570] Fabienne Peter – Hans Bernhard Schmid: *Rationality and Commitment*. 2007, Oxford University Press.

[571] R. W. Picard: Affective computing. No. 321. Jelentés, 1995, M.I.T Media Laboratory Perceptual Computing Section Technical Report.

[572] Charles R. Pigden: Desiring to desire: Russell, moore and david lewis.

[573] G. Pigozzi: Two aggregation paradoxes in social decision making: the ostrogorski paradox and the discursive dilemma. *Episteme: A Journal of Social Epistemology*, 2(2):. évf. (2005) 2. sz., 33–42. p.

[574] Jane Allyn Piliavin – Hong-Wen Charng: Altruism: A review of recent theory and research. *Annual Review of Sociology*, 16. évf. (1990), 27–65. p.

[575] John Platt: Social traps. *American Psychologist*, 28. évf. (1973) 8. sz., 641–651. p.

[576] Ifjabb Plinius: *Levelek*. Budapest, 1981, Európa Könyvkiadó.

[577] Charles R. Plott – Michael E. Levine: A model of agenda influence on committee decisions. *American Economic Review*, 68. évf. (1978) 1. sz., 146–70. p.

[578] Pokol Béla: *Modern francia szociológiaelméletek*. Miskolc, 1996, Bíbor Kiadó. URL `http://mek.oszk.hu/02000/02027/02027.pdf`.

[579] Ingmar Pörn: *The Logic of Power*. 1970, Blackwells.

[580] Ingmar Pörn: *Action theory and social science: some formal models*. 1977, D. Reidel.

[581] William Poundstone: *Prisoner's Dilemma*. 1992, Anchor.

[582] A. N. Prior: The paradoxes of derived obligation. *Mind, New Series*, 63. évf. (1954) 249. sz., 64–65. p.

[583] A. N. Prior: Escapism: The logical basis of ethics. In A.I. Melden (szerk.): *Essays In Moral Philosophy*. Seattle, 1958, University of Washington Press, 135–146. p.

[584] Warren S. Quinn: Theories of intrinsic value. *American Philosophical Quarterly*, 11. évf. (1974), 123–132. p.

[585] Warren S. Quinn: The puzzle of the self-torturer. *Philosophical Studies*, 59. évf. (1990) 1. sz., 79–90. p.

[586] Jacob M. Rabbie – Hein F.M. Lodewijkx: Conflict and aggression: An individual-group continuum. In *Advances in Group Processes*. 11. köt. 1994, JAI Press, 139–174. p.

[587] Matthew Rabin: Incorporating fairness into game theory and economics. *The American Economic Review*, 83. évf. (1993. Dec) 5. sz., 1281–1302. p.

[588] Douglas W. Rae – Hans Daudt: The ostrogorski paradox: a peculiarity of compound majority decision. *European Journal of Political Research*, 4. évf. (1976) 4. sz., 391–399. p.

[589] N. J. Raihani – K. McAuliffe: Human punishment is motivated by inequity aversion, not a desire for reciprocity, 2012.

[590] J. M. Ramírez – J. M. Andreu: Aggression, and some related psychological constructs (anger, hostility, and impulsivity). some comments from a research project. *Neurosci Biobehav Rev*, 30. évf. (2006) 3. sz., 276–291. p.

[591] Joseph Ransdell: Constitutive rules and speech-act analysis. *The Journal of Philosophy*, 68. évf. (1971) 13. sz., 385–400. p.

[592] Anatol Rapoport: Three modes of conflict. *Management Science*, 7. évf. (1961) 3. sz., 210–218. p.

[593] Anatol Rapoport: Optimal policies for the prisoner's dilemma. *Psychological Review*, 74. évf. (1967) 2. sz., 136–148. p.

[594] Anatol Rapoport: A kizsákmányoló, a vezér, a hős és a mártír: a 2 x 2-es játszma négy archetípusa. In László Csontos (szerk.): *A racionális döntések elmélete*. Budapest, 1998, Osiris Kiadó – Láthatatlan Kollégium, 129–134. p.

[595] John Rawls: Kétfajta szabályfogalom. In Mária Lónyai (szerk.): *Tények és értékek. A modern angolszász etika irodalmából*. Budapest, 1981, Gondolat, 364–416. p.

[596] Paramesh Ray: Independence of irrelevant alternatives. *Econometrica*, 41. évf. (1973) 5. sz., 987–991. p.

[597] Joseph Raz: *The Morality of Freedom*. 1989, Clarendon Press.

[598] Hyne W. Reese: Teleology and teleonomy in behavior analysis. *The Behavior Analyst*, 17. évf. (1994) 1. sz., 75–91. p.

[599] Raymond R. Reno – Robert B. Cialdini – Carl A. Kallgren: The transsituational influence od social norms. *Journal of Personality and Social Psychology*, 64. évf. (1993) 1. sz., 104–112. p.

[600] Philip J. Reny: Arrow's theorem and the gibbard-satterthwaite theorem: A unified approach. *Economics Letters*, 70. évf. (2001) 1. sz., 99–105. p.

[601] Nicholas Rescher (szerk.): *The Logic of Decision and Action*. 1966, University of Pittsburgh Press.

[602] Nicholas Rescher: Semantic foundations for the logic of preference. In Nicholas Rescher (szerk.): *The Logic of Decision and Action*. Pittsburgh, 1966, University of Pittsburgh Press, 37–80. p.

[603] Anna Réz: Makacs érzelmek és morális felelősség. *Magyar Filozófiai Szemle*, 55. évf. (2011) 1. sz., 23–33. p.

[604] Marcel K. Richter: Revealed preference theory. *Econometrica*, 34. évf. (1966) 3. sz., 635–645. p.

[605] Matt Ridley: *The Origins of Virtue*. 1997, Softback Preview.

[606] William H. Riker: *The Art of Political Manipulation*. 1986, Yale University Press.

[607] Kevin W.S. Roberts.: Voting over income tax schedules. *Journal of Public Economics*, 8. évf. (1977) 3. sz., 329–340. p.

[608] David Robinson – David Goforth: *The Topology of the 2x2 Games*. 2005, Routledge.

[609] Ronald Rogowski: Representation in political theory and in law. *Ethics*, 91. évf. (1981) 3. sz., 395–430. p.

[610] Alejandro Rosas: Evolutionary game theory meets social science: Is there a unifying rule for human cooperation? *Journal of Theoretical Biology*, 264. évf. (2010), 450–456. p.

[611] Alf Ross: Imperatives and logic. *Philosophx of Science*, 11. évf. (1944) 1. sz., 30–46. p.

[612] David Ross: *The Right and the Good*. Oxford, 2002 (1930), Clarendon Press.

[613] Francesca Rossi – Kristen Brent Venable – Toby Walsh: *A Short Introduction to Preferences. Between Artificial Intelligence and Social Choice*. 2011, Morgan & Claypool.

[614] A. E. Roth: Deferred acceptance algorithms: History, theory, practice, and open questions. *International Journal of Game Theory, Special Issue in Honor of David Gale on his 85th birthday*, 36. évf. (2008) 3-4. sz., 537–569. p.

[615] Alvin E. Roth: The evolution of the labor market for medical interns and residents: A case study in game theory. *Journal of Political Economy*, 92. évf. (1984) 6. sz., 991–1016. p.

[616] P. Rothstein: Order restricted preferences and majority rule. *Social Choice and Welfare*, 7. évf. (1990) 4. sz., 331–342. p.

[617] P. Rothstein: Representative voter theorems. *Public Choice*, 72. évf. (1991), 193–212. p.

[618] M. Roubens – Ph. Vincke: *Preference modelling*. Lecture Notes in Economics and Mathematical Systems sorozat, 250. köt. Berlin, 1985, Springer.

[619] N. Rouland: Az emberi jogok antropológiai alapjai. In István H. Szilágyi (szerk.): *Jog és antropológia*. 2005, Szent István Társulat, 275–294. p.

[620] Lajos Rózsa: The motivation for biological aggression is an inherent and common aspect of the human behavioural repertoire. *Medical Hypotheses*, 72. évf. (2009), 217–219. p.

[621] David-Hillel Ruben: John searle's the construction of social reality. *Philosophy and Phenomenological Research*, 57. évf. (1997) 2. sz., 443–447. p.

[622] Rossella Rubino – Antonino Rotolo – Giovanni Sartor: An OWL Ontology of Norms and Normative Judgements. In Carlo Biagioli – Enrico Francesconi – Giovanni Sartor (szerk.): *Proceedings of the V Legislative XML Workshop* (konferenciaanyag). 2007, 173–187. p.

[623] A. Rubinstein – P. C. Fishburn: Algebraic aggregation theory. *Journal of Economic Theory*, 38. évf. (1986), 63–77. p.

[624] D.W.P. Ruiter: A basic classification of legal institutions. *Ratio Juris*, 10. évf. (1997) 4. sz., 357–371. p.

[625] Stuart J. Russel – Devika Subramanian: Provably bounded-optimal agents. *Journal of Artificial Intelligence Research*, 2. évf. (1995), 575–609. p.

[626] Bertrand Russell: A denotálásról. *Világosság*, 46. évf. (2005) 12. sz., 5–16. p.

[627] Ruzsa Imre: *Bevezetés a modern logikába*. 2000, Osiris Kiadó.

[628] Donald G. Saari: *Decisions and Elections: Explaining the Unexpected*. Cambridge, 2001, Cambridge University Press.

[629] J. Sachs: Cooperation within and among species. *J. Evol. Biol.*, 19. évf. (2006) 5. sz., 1415–1418. p.

[630] Marshall D. Sahlins: *Törzsek*. Budapest, 1973, Kossuth Könyvkiadó, 135–315. p.

[631] Ton Sales: Logic as general rationality: A survey. In Dov M. Gabbay – F. Guenthner (szerk.): *Handbook of Philosophical Logic*. 9. köt. 2nd edition. kiad. 2002, Kluwer Academic Publishers, 321–366. p.

[632] Larry Samuelson: Introduction to the evolution of preferences. *Journal of Economic Theory*, 97. évf. (2001), 225–230. p.

[633] Paul Samuelson: Altruism as a problem involving group versus individual selection in economics and biology. *The American Economic Review*, 83. évf. (1993), 143–148. p.

[634] John T. Sanders: Why the numbers should sometimes count. *Philosophy and Public Affairs*, 17. évf. (1988) 1. sz., 3–14. p.

[635] David H. Sanford: Contraries and subcontraries. *Noûs*, 2. évf. (1968) 1. sz., 95–96. p.

[636] Alejandro Saporiti – Fernando Tohmé: Single-crossing, strategic voting and the median choice rule. *Social Choice and Welfare*, 26. évf. (2006), 363–383. p.

[637] Giovanni Sartor: Fundamental Legal Concepts: A Teleological Characterisation, 2005.

[638] Giovanni Sartor: Fundamental Legal Concepts: A Formal and Teleological Characterisation. Jelentés, 2006, European University Institute Department of Law.

[639] Mark Allen Satterthwaite: Strategy-proofness and arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory*, 10. évf. (1975), 187–216. p.

[640] Frederick Schauer: A szabályok és a jog uralma. In *A Hart utáni jogelmélet alapproblémái*. 2004, Bíbor Kiadó, 407–457. p.

[641] Stephen Schiffer: A paradox of desire. *American Philosophical Quarterly*, 13. évf. (1976) 3. sz., 195–203. p.

[642] Wijnand J. Schoenmakers: A problem in knowledge acquisition. *SIGART Newsletter*, 1986. 95. sz., 56–57. p.

[643] Arthur Schopenhauer: *Az akarat szabadságáról*. 1913, Franklin Társulat.

[644] Timothy Schroeder: *Three Faces of Desire*. New York, 2004, Oxford University Press.

[645] Timothy Schroeder: Reflection, reason, and free will. *Philosophical Explorations*, 10. évf. (2007) 1. sz., 77–84. p.

[646] P. Wesley Schultz – Jessica M. Nolan – Robert B. Cialdini – Noah J. Goldstein – Vladas Griskevicius: The constructive, destructive, and reconstructive power of social norms. *Psychological Science*, 18. évf. (2007) 5. sz., 429–434. p.

[647] George Schwab: Enemy vagy foe: a modern politika konfliktusa. In Márton Szabó (szerk.): *Az ellenség neve*. 1998, Jószöveg Műhely, 39–55. p.

[648] Thomas Schwartz: Rationality and the myth of the maximum. *Noûs*, 6. évf. (1972) 2. sz., 97–117. p.

[649] Thomas Schwartz: The paradox of representation. *The Journal of Politics*, 57. évf. (1995) 2. sz., 309–323. p.

[650] John R. Searle: *Intentionality*. 1983, Cambridge University Press.

[651] John R. Searle: *The Construction of Social Reality*. 1995, The Free Press.

[652] John R. Searle: *Elme, nyelv, társadalom. A való világ filozófiája*. Budapest, 2000, Vince Kiadó.

[653] John R. Searle: *Beszédaktusok*. 2009, Alkalmazott Kommunikációtudományi Intézet – Gondolat Kiadó.

[654] Krister Segerberg: Validity and satisfaction in imperative logic. *Notre Dame Journal of Formal Logic*, 31. évf. (1990), 203–221. p.

[655] Krister Segerberg: Getting started: Beginnings in the logic of action. *Studia Logica: An International Journal for Symbolic Logic*, 51. évf. (1992) 3/4. sz., 347–378. p.

[656] Reinhard Selten: The chain store paradox. *Theory and Decision*, 9. évf. (1978) 2. sz., 127–159. p.

[657] A.K. Sen: Goals, commitment, and identity. *Journal of Law, Economics, and Organization*, 1. évf. (1985) 2. sz., 341–355. p.

[658] A.K. Sen: Internal consistency of choice. *Econometrica*, 61. évf. (1993) 3. sz., 495–521. p.

[659] Amartya K. Sen: A possibility theorem on majority decisions. *Econometrica*, 34. évf. (1966) 2. sz., 491–499. p.

[660] Amartya K. Sen: Isolation, assurance and the social rate of discount. *The Quarterly Journal of Economics*, 81. évf. (1967. Feb) 1. sz., 112–124. p.

[661] Amartya K. Sen: Quasi-transitivity, rational choice and collective decisions. *The Review of Economic Studies*, 36. évf. (1969. July) 3. sz., 381–393. p.

[662] Amartya K. Sen: *Collective Choice and Social Welfare*. 1970, Holden-Day.

[663] Amartya K. Sen: The impossibility of a paretian liberal. *Journal of Political Economy*, 78. évf. (1970. Feb. Jan..) 1. sz., 152–157. p.

[664] Amartya K. Sen: Choice functions and revealed preference. *Review of Economic Studies*, 38. évf. (1971) 3. sz., 307–317. p.

[665] Amartya K. Sen: Behaviour and the concept of preference. *Economica. New Series*, 40. évf. (1973) 159. sz., 241–259. p.

[666] Amartya K. Sen: Social choice theory: A re-examination. *Econometrica*, 45. évf. (1977) 1. sz., 53–89. p.

[667] Amartya K. Sen: Rationality and social choice. *American Economic Review*, 85. évf. (1995) 1. sz., 1–24. p.

[668] Amartya K. Sen: Maximization and the act of choice. *Econometrica*, 65. évf. (1997) 4. sz., 745–779. p.

[669] Amartya K. Sen – Prasanta K. Pattanaik: Necessary and sufficient conditions for rational choice under majority decision. *Journal of Economic Theory*, 1. évf. (1970), 178–202. p.

[670] Seregi Tamás: Paul ricoeur cselekvésontológiája. *Világosság*, 48. évf. (2007) 1. sz., 43–64. p.

[671] Marek J. Sergot: Normative positions. In P. McNamara – H. Prakken (szerk.): *Norms, Logics and Information Systems*. 1999, IOS, 289–308. p.

[672] Marek J. Sergot: A computational theory of normative positions. *ACM Transactions on Computational Logic*, 2. évf. (2001) 4. sz., 581–622. p.

[673] Marek J. Sergot – F.M. Richards: On the representation of action and agency in the theory of normative positions. *Fundamenta Informaticae*, 48. évf. (2001. November) 2-3. sz., 273–293. p.

[674] R. Sethi – E. Somanathan: Preference evolution and reciprocity. *Journal of Economic Theory*, 97. évf. (2001) 2. sz., 273–297. p.

[675] Kieran Setiya: Knowledge of intention. In Anton Ford – Jennifer Hornsby – Frederick Stoutland (szerk.): *Essays on Anscombe's Intention*. 2011, Harvard University Press, 170–197. p.

[676] Leigh S. Shaffer: Toward pepitone's vision of a normative social psychology: What is a social norm? *Journal of Mind and Behavior*, 4. évf. (1983) 2. sz., 275–294. p.

[677] Stewart Shapiro: *Foundation without Foundationalism. A Case for Second-order Logic*. 1991, Clarendon Press.

[678] Lloyd S. Shapley – Martin Shubik: On the core of an economic system with externalities. *The American Economic Review*, 59. évf. (1969) 4. sz., 678–684. p.

[679] Georg Simmel: *The Sociology of Georg Simmel*. The Triad fejezet. 1950, The Free Press, 145–169. p.

[680] H.A. Simon: Altruizmus és közgazdaságtan. *Magyar Tudomány*, 40 (102). évf. (1995) 1. sz., 22–29. p.

[681] Herbert Simon: A racionalitás mint folyamat és gondolkodási produktum. In *Korlátozott racionalitás*. Budapest, 1982, KJK, 59–86. p.

[682] Herbert Simon: A mechanism for social selection and successful altruism. *Science*, 250. évf. (1990), 1665–1668. p.

[683] Herbert A. Simon: d'żźtheories of decision-making in economics and behavioral science. *The American Economic Review*, 49. évf. (1959) 3. sz., 253–283. p.

[684] Simonovits András: Bevezetés a játékelméletbe: vázlat, 2007.

[685] Peter Simons: *Parts: A Study in Ontology*. 2000 (1987), Oxford University Press.

[686] Munindar P. Singh: An ontology for commitments in multiagent systems: Toward a unification of normative concepts. *Artificial Intelligence and Law*, 7. évf. (1999), 97–113. p.

[687] Brian Skyrms: *Evolution of the Social Contracts*. New York, 1996, Cambridge University Press.

[688] Brian Skyrms: *The Stag Hunt and the Evolution of Social Structure*. 2004, Cambridge University Press.

[689] Brian Skyrms: Evolution and the social contract, 2007.

[690] Barry Hartley Slater: Logical paradoxes. *Internet Encyclopedia of Philosophy*, 2005.

[691] Paul Slovic – Melissa L. Finucane – Ellen Peters – Donald G. MacGregor: Rational actors or rational fools? implications of the affect heuristic for behavioral economics, 2002.
URL http://www.decisionresearch.org/pdf/dr498v2.pdf.

[692] Hans Smessaert: On the 3d visualisation of logical relations. *Logica Universalis*, 3. évf. (2009), 303–332. p.

[693] Barry Smith: Searle and De Soto: The New Ontology of the Social World. In *The Mystery of Capital and the Construction of Social Reality*. 2008, Open Court, 35–51. p.

[694] Barry Smith – John Searle: The construction of social reality: An exchange. *American Journal of Economics and Sociology*, 62. évf. (2003) 2. sz., 285–309. p.

[695] John H. Smith: Aggregation of preferences with variable electorate. *Econometrica*, 41. évf. (1973) 6. sz., 1027–1041. p.

[696] John Maynard Smith: *Evolution and the Theory of Games*. 1982, Cambridge University Press.

[697] Michael Smith – David Lewis – Mark Johnston: Dispositional theories of value. *Proceedings of the Aristotelian Society, Supplementary Volumes*, 63. évf. (1989), 89 –111+113–137+139–174. p.

[698] Elliott Sober – David Sloan Wilson: *Unto Others: The Evolution and Psychology of Unselfish Behavior*. 1998, Harvard University Press.

[699] Arend Soeteman: *Logic in Law. Remarks on Logic and Rationality in Normative Reasoning, Especially in Law*. 1989, Kluwer Academic Publishers.

[700] Solt Kornél: *Jogi logika I-II.* 1996, MTA Állam- és Jogtudományi Intézete and Seneca.

[701] Somlai Péter: Triádok, közvetítők, koalíciók. konfliktusok és egyensúlyok kettőnél több tagú szociális rendszerekben. *Szociológiai Szemle*, 2. évf. (1992) 4. sz., 3–14. p.
URL http://www.szociologia.hu/dynamic/9204somlai.htm.

[702] Dan Sperber – Deirdre Wilson: Kölcsönös tudás és relevancia a megértés elméletében. In István Síklaki (szerk.): *Szóbeli befolyásolás*. 2008, Typotex Kiadó, 171–194. p.

[703] Wolfgang Spohn: An analysis of hansson's dyadic deontic logic. *Journal of Philosophical Logic*, 4. évf. (1975. May) 2. sz., 237–252. p.

[704] J. A. Srejder: *Egyenlőség, hasonlóság, rendezés*. Budapest, 1975, Gondolat.

[705] R. Stalnaker: Knowledge, belief and counterfactual reasoning in games. In C. Bicchieri – R. C. Jeffrey – B. Skyrms (szerk.): *The Logic of Strategy*. 1999, Oxford University Press, 3–38. p.

[706] Robert Stalnaker: Belief revision in games: forward and backward induction. *Mathematical Social Sciences*, 36. évf. (1998), 31–56. p.

[707] Robert Stalnaker: Extensive and strategic forms: Games and models for games. *Research in Economics*, 53. évf. (1999) 3. sz., 293–319. p.

[708] Robert Stalnaker: On logics of knowledge and belief. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 128. évf. (2006) 1. sz., 169–199. p.

[709] Robert Stalnaker: Iterated belief revision. *Erkenntnis*, 70. évf. (2009) 2. sz., 189–209. p.

[710] Steiger Kornél: Megjegyzések Orthmayr Imre "a sztoikus szenvtelenség" c. tanulmányához. *Századvég*, 1999.
URL `http://www.c3.hu/scripta/szazadveg/13/steiger.htm`.

[711] Daniel N. Stern: *A csecsemő személyközi világa*. 2002, Animula.

[712] Eleonore Stump: Sanctification, hardening of the heart, and frankfurt's concept of free will. *The Journal of Philosophy*, 85. évf. (1988. Aug) 8. sz., 395–420. p.

[713] Scott Sturgeon: Normative judgement. *Noûs, Supplement: Philosophical Perspectives*, 41. évf. (2007) 21. sz., 569–587. p.

[714] Robert Sugden: Reciprocity: The supply of public goods through voluntary contributions. *Economic Journal*, 94. évf. (1984), 772–787. p.

[715] Robert Sugden: Spontaneous order. *The Journal of Economic Perspectives*, 3. évf. (1989) 4. sz., 85–97. p.

[716] Robert Sugden: *The Economics of Rights, Co-operation and Welfare*. 2004 (1986), Palgrave MacMillan.

[717] Cass R. Sunstein: A pozitív jogok ellen. *Fundamentum*, 1998. 3. sz., 11–14. p.

[718] Kotaro Suzumura: *Rational choice, collective decisions, and social welfare*. 1983, Cambridge University Press.

[719] Lars-Gunnar Svensson: The proof of the gibbard-satterthwaite theorem revisited, 2000.

[720] Christine Swanton: Weakness of will as a species of executive cowardice. *Canadian Journal of Philosophy*, 21. évf. (1991) 2. sz., 123–140. p.

[721] Syi: Akarat-erő-logika. In Katalin S. Nagy – Annamária Orbán (szerk.): *Értékek és normák interdiszciplináris megközel´ésben*. Budapest, 2008, Gondolat Kiadó, 50–68. p.

[722] Syi: *Cselekvéselmélet dióhéjban*. 2008, Typotex.

[723] Syi: Önzetlenségelméletek. *BUKSZ*, 21. évf. (2009) 4. sz., 226–233. p.

[724] Syi: Max weber cselekvéselméleti ontológiája. *Szociológiai Szemle*, 21. évf. (2011) 2. sz., 87–110. p.

[725] Syi: Feléd fordítom, újra. In Bajnok Andrea – Korpics Márta – Milován Andrea – Pólya Tamás – Szabó Levente (szerk.): *A kommunikatív állapot: Diszciplináris rekonstrukciók*. 2012, Typotex Kiadó, 122–134. p.

[726] Syi: A racionális aracionális. *BUKSZ*, 24. évf. (2012) 2. sz., 118–125. p.

[727] Szakadát István: *A kommunista uralmi rend és a nómenklatúraelv*. PhD értekezés (MTA szociológiai tudományok kandidátusa). 1999.

[728] Szakadát István: Hierarchia. In Helmich Dezső – Szántó Zoltán (szerk.): *Metodológia, társadalom, gazdaság. Im memoriam Bertalan László*. 2004, Közgazdasági Szemle Alapítvány, 139–166. p.

[729] Szakadát István: Reláció, szintaktika, szemantika. *Tudományos és Műszaki Tájékoztatás*, 51. évf. (2004) 12. sz., 531–540. p.

[730] Szakadát István: Réteges struktúra, alaprelációk. In Csendes Lilla – Alexin Zoltán (szerk.): *III. Szám´ógépes nyelvészeti konferencia* (konferenciaanyag). Szeged, 2005, SZTE.

[731] Szakadát István: A politikai cselekvések szabályszerűségei. In S.Nagy Katalin (szerk.): *Szociológia*. 2006, Typotex Kiadó, 99–114. p.

[732] Szakadát István: Politikai gazdaságtan. In S.Nagy Katalin (szerk.): *Szociológia közgazdászoknak*. 2007, Typotex Kiadó, 213–231. p.

[733] Szécsi Gábor: Okság és intencionalitás. *Magyar Filozófiai Szemle*, 39. évf. (1995) 3-4. sz., 529–561. p.

[734] Szily László: Az apostoli magyarok leváltották schmittet.
URL `http://index.hu/belfold/2012/03/01/schmitt_palt_levaltottak_lazar_es_tarsai_haditorvenyszek_elott/`.

[735] Szívós Mihály: Az elsődleges és másodlagos habitus szociológiai és tudományszociológiai fogalma a személyes és a hallgatólagos tudás elméletének szemszögéből. *Polanyiana*, 18. évf. (2009) 1-2. sz., 29–52. p.

[736] Szolcsányi Tibor: A davidsoni háromszögelés információ-elméleti elemzése. In *Igazság és kommunikáció. Tanulmányok Donald Davidson filozófiájáról*. 2004, Brambauer Kiadó, 145–166. p.

[737] Takács Károly: Hálózati kísérletek. *Közgazdasági Szemle*, LVII. évf. (2010), 958–979. p.

[738] N. Takahashi – R. Mashima: The emergence of indirect reciprocity: Is the standing strategy the answer?, 2003. Aug.
URL http://www.allacademic.com/meta/p107391_index.html. Paper presented at the annual meeting of the American Sociological Association, Atlanta Hilton Hotel, Atlanta, GA Online <.PDF> Retrieved 2006-10-05.

[739] John M. Taurek: Should the numbers count? *Philosophy and Public Affairs*, 6. évf. (1977) 4. sz.

[740] Alan D. Taylor: *Social Choice and the Mathematics of Manipulation*. Cambridge, 2005, Cambridge University Press.

[741] Charles Taylor: *Sources of the Self. The Making of the Modern Identity*. 1989, Harvard University Press.

[742] Kenneth A. Taylor: De re and de dicto: Against the conventional wisdom. *Noûs, Supplement: Philosophical Perspectives, Language and Mind*, 36. évf. (2002) 16. sz., 225–265. p.

[743] Rafael Di Tella – Ricardo Pérez-Truglia: Conveniently upset: Avoiding altruism by distorting beliefs about others. Jelentés, 2010, National Bureau of Economic Research.

[744] Temesi József: *A döntéselmélet alapjai*. 2002, Aula Kiadó.

[745] Tengelyi László: Tapasztalat, cselekvés és elbeszélt történet. *Világosság*, 50. évf. (2009) tavasz. sz., 87–99. p.

[746] Warren D. TenHouten: *A General Theory of Emotions and Social Life*. 2007, Routledge.

[747] Michael Thompson: *Life and Action. Elementary Structures of Practice and Practical Thought*. 2008, Harvard University Press.

[748] Tibor Szolcsányi: *A feljogos´ó következtetések logikája*. PhD értekezés (Pécsi Tudományegyetem Bölcsészettudományi Kar Filozófia Intézet). 2009.

[749] Tokics Imre: *A Ius Talionis és az asylum városok jogi, vallási és kultúratörténeti jelentősége a tanakban*. PhD értekezés (Országos Rabbiképző – Zsidó Egyetem Zsidó Vallástudományi Doktori Iskola). 2011.

[750] James E. Tomberlin: Good samaritans and castañeda's system of deontic logic. In James E. Tomberlin (szerk.): *Héctor-Neri Castañeda*. 6. köt. 1986, Springer, 255–272. p.

[751] Giulio Tononi: An information integration theory of consciousness. *BMC Neuroscience*, 5. évf. (2004), 42–66. p.

[752] Tőzsér János: Az önismeret határa. *Magyar Filozófiai Szemle*, 55. évf. (2011) 1. sz., 9–22. p.

[753] Robert L. Trivers: A reciprok altruizmus evolúciója. In Zsuzsanna Kulcsár (szerk.): *Morális fejlődés, empátia és altruizmus*. 1999, ELTE Eötvös Kiadó, 245–272. p.

[754] Gordon Tullock: Problems of majority voting. *Journal of Political Economy*, 67. évf. (1959) 6. sz., 571–579. p.

[755] L. Tummolini: Convention: an interdisciplinary study. *Topoi. An International Review of Philosophy*, 27. évf. (2008) 1-2. sz.

[756] L. Tummolini: Social conventions. In B. Kaldis (szerk.): *Encyclopedia of Philosophy and the Social Sciences*. 2013, SAGE Publications, 900 – 902. p.

[757] L. Tummolini – G. Andrighetto – C. Castelfranchi – R. Conte: A convention or (tacit) agreement betwixt us: on reliance and its normative consequences. *Synthese*, 190. évf. (2013) 4. sz., 585 – 618. p.

[758] L. Tummolini – C. Castelfranchi: The cognitive and behavioral mediation of institutions: Towards an account of institutional actions. *Cognitive Systems Research*, 7. évf. (2006) 2.3. sz., 307 – 323. p.

[759] Amos Tversky: Intransitivity of preferences. *Psychological Review*, 76. évf. (1969) 1. sz., 31–48. p.

[760] Luis Ubeda: Neutrality in arrow and other impossibility theorems. *Economic Theory*, 23. évf. (2004), 195–204. p.

[761] Sara L. Uckelman: Anselm's logic of agency, 2012. February.

[762] Edna Ullmann-Margalit: *The Emergence of Norms*. 1977, Clarendon Press.

[763] Edna Ullmann-Margalit: Family fairness. *Social Research*, 73. évf. (2006), 575–596. p.

[764] Edna Ullmann-Margalit: Considerateness. *Iyyun, The Jerusalem Philosophical Quarterly*, 60. évf. (2011. July), 205–244. p.

[765] James O. Urmson: Saints and heroes. In A.J. Melden (szerk.): *Essays in Moral Philosophy*. Seattle, 1958, University of Washington Press, 198–216. p.

[766] James O. Urmson: Szentek és hősök. *Café Bábel*, 13. évf. (1994) 3. sz., 53–62. p.

[767] Piercarlo Valdesolo – David DeSteno: Manipulations of emotional context shape moral judgment. *Psychological Science*, 17. évf. (2006) 6. sz., 476–477. p.

[768] J. van Benthem – S. van Otterloo – O. Roy: Preference logic, conditionals, and solution concepts in games. In H. Lagerlund – S. Lindström – R. Sliwinski (szerk.): *Modality matters: Twenty-five essays in honour of Krister Segerberg*. Uppsala Philosophical Studies sorozat, 53. köt. Uppsala, 2006, Uppsala Universitet, 61–77. p.

[769] Johan van Benthem: Rational dynamic and epistemic logic in games. In S. Vanucci (szerk.): *Logic, Game Theory and Social Choice III.* 2003, University of Siena, 19–23. p.

[770] Johan van Benthem: For better or for worse: Dynamic logics of preference. In T. Grüne-Yanoff – S. O. Hansson (szerk.): *Preference change. Approaches from Philosophy, Economics and Psychology*. Dordrecht, 2008, Springer, 57–84. p.

[771] Johan van Benthem – Patrick Girard – Olivier Roy: Everything else being equal: A modal logic for ceteris paribus preferences. *Journal of Philosophical Logic*, 38. évf. (2009) 1. sz., 83–125. p.

[772] Johan van Benthem – Davide Grossi – Fenrong Liu: On the two faces of deontics: Semantic betterness and syntactic priority, 2011.

[773] Johan van Benthem – Davide Grossi – Fenrong Liu: Priority structures in deontic logic, 2012.

[774] Johan van Benthem – Amitabha Gupta – Eric Pacuit (szerk.): *Games, Norms and Reasons. Logic at the Crossroads*. 2011, Springer.

[775] Johan van Benthem – Fenrong Liu: Dynamic logic of preference upgrade. *Journal of Applied Non-Classical Logics*, 17. évf. (2007) 2. sz., 157–182. p.

[776] Ron van der Meyden – Leendert van der Torre (szerk.): *Deontic Logic in Computer Science, 2008*. 2008, Springer.

[777] L.W.N. van der Torre – Y.-H. Tan: The many faces of defeasibility. In Donald Nute (szerk.): *Defeasible Deontic Logic*. Dordrecht/Boston/London, 1997, Kluwer Academic Publishers, 79–121. p.

[778] H. van Ditmarsch – B. Kooi – W. van der Hoek: *Dynamic epistemic logic*. Synthese Library sorozat, 337. köt. Heidelberg, 2007, Springer.

[779] Hans van Ditmarsch – Tim French – Fernando R. Velazquez-Quesada: Action models for knowledge and awareness. In Wiebe van der Hoek – Lin Padgham – Vincent Conitzer – Michael Winikoff (szerk.): *Proceedings of 11th AAMAS* (konferenciaanyag). 2012, IFAAMAS, 1091–1098. p.

[780] Hans van Ditmarsch – Andreas Herzig – Tiago de Lima.: Public announcements, public assignments and the complexity of their logic. *Journal of Applied Non-Classical Logics*, 22. évf. (2012) 3. sz., 249–273. p.

[781] Hans van Ditmarsch – W. van der Hoek – P. Iliev: Everything is knowable – how to get to know whether a proposition is true. *Theoria*, 78. évf. (2012) 2. sz., 93–114. p.

[782] Paul A. M. van Lange: The pursuit of joint outcomes and equality in outcomes: An integrative model of social value orientation. *Journal of Personality and Social Psychology*, 77. évf. (1999) 2. sz., 337–349. p.

[783] Paul A. M. van Lange – D. De Cremer – E. Van Dijk – M. Van Vugt: Self-interest and beyond: Basic principles of social interaction. In A. W. Kruglanski – E. T. Higgings (szerk.): *Social Psychology: Handbook of Basic Principles*. 2nd edition. kiad. 2007, Guilford, 540–561. p.

[784] Paul A. M. van Lange – W. Otten – E. M. N. De Bruin – J. A. Joireman: Development of prosocial, individualistic, and competitive orientations: Theory and preliminary evidence. *Journal of Personality and Social Psychology*, 73. évf. (1997) 4. sz., 733–746. p.

[785] Daniel Vanderveken (szerk.): *Logic, Thought and Action*. 2005, Springer.

[786] A.C. Varzi: Parts, wholes, and part-whole relations: The prospects of mereotopology. *Data and Knowledge Engineering*, 20. évf. (1996), 177–198. p.

[787] Bruce Vermazen: Negative acts. In Bruce Vermazen – Merril B. Hintikka (szerk.): *Essays on Davidson. Actions and Events*. 1985, Oxford University Press, 93–104. p.

[788] William Vickrey: Utility, strategy, and social decision rules. *The Quarterly Journal of Economics*, 74. évf. (1960) 4. sz., 507–535. p.

[789] G. H. von Wright: A note on deontic logic and derived obligation. *Mind, New Series*, 65. évf. (1956) 260. sz., 507–509. p.

[790] G. H. von Wright: On so-called practical inference. *Acta Sociologica*, 15. évf. (1972) 1. sz., 39–53. p.

[791] Georg Henrik von Wright: *An Essay in Deontic Logic*. 1951, North-Holland Publishing Company.

[792] Georg Henrik von Wright: *Norm and Action*. London, 1963, Routledge and Kegan Paul.

[793] Georg Henrik von Wright: Practical inference. *Philosophical Review*, 72. évf. (1963) 2. sz., 159–179. p.

[794] Georg Henrik von Wright: *The Logic of Preference*. Edinburgh, 1963, Edinburgh University Press.

[795] Georg Henrik von Wright: The logic of preference reconsidered. *Theory and Decision*, 3. évf. (1972) 3. sz., 140–169. p.

[796] Georg Henrik von Wright: Deontikus logika. In Irving M. Copi – James A. Gould (szerk.): *Kortárs-tanulmányok a logikaelmélet kérdéseiről*. Budapest, 1985, Gondolat, 531–554. p.

[797] Peter B.M. Vranas: New foundations for imperative logic ii: Pure imperative inference.

[798] Peter B.M. Vranas: New foundations for imperative logic iii: A general definition of argument validity.

[799] Peter B.M. Vranas: New foundations for deontic logic: A preliminary sketch, 2002.

[800] Peter B.M. Vranas: New foundations for imperative logic i: Logical connectives, consistency, and quantifiers. *Noûs*, 42. évf. (2008), 529– 572. p.

[801] Peter B.M. Vranas: Imperatives, logic of. In *The International Encyclopedia of Ethics*. 2013, John Wiley and Sons.

[802] Loïc J.D. Wacquant: *Body and Soul. Notebooks of and Apprentice Boxer*. 2004, Oxford University Press.

[803] Carl Wagner: Anscombe's paradox and the rule of three-fourths. *Theory and Decision*, 15. évf. (1983) 3. sz., 303–308. p.

[804] Carl Wagner: Avoiding anscombe's paradox. *Theory and Decision*, 16. évf. (1984) 3. sz., 233–238. p.

[805] Douglas N. Walton: Logical form and agency. *Philosophical Studies*, 29. évf. (1976), 75–89. p.

[806] Douglas N. Walton: St. anselm and the logical syntax of agency. *Franciscan Studies*, 36. évf. (1976), 298–312. p.

[807] Douglas N. Walton: On the logical form of some commomplace action expression. *Grazer Philosophische Studien*, 10. évf. (1980), 141–148. p.

[808] Peng-Hsiang Wang: Anderson's reduction and kelsen's normativism. In Arend Soeteman (szerk.): *Pluralism and Law* (konferenciaanyag). 2004, 104–110. p.

[809] Gary Watson: *Agency and Answerability*. 2004, Oxford University Press.

[810] Max Weber: *Gazdaság és társadalom I. A megértő szociológia alapvonalai*. 1987, KJK.

[811] D. M. Wegner – J. A. Bargh: Control and automaticity in social life. In D. Gilbert – S. T. Fiske – G. Lindzey (szerk.): *Handbook of social psychology*. 1. köt. 4th. kiad. 1998, McGraw-Hill, 446–496. p.

[812] Leif Wenar: Rights, 2005 (2011).
URL `http://plato.stanford.edu/entries/rights/`.

[813] S. A. West – A. S. Griffin – A. Gardner: Evolutionary explanations for cooperation. *Current Biology*, 17. évf. (2007), R661–R672. p.

[814] S.A. West – A. Gardner: Altruism, spite and greenbeards. *Science*, 327. évf. (2010), 1341–1344. p.

[815] A.N. Whitehead: *An Introduction to Mathematics*. 1911, Henry Holt and Company.

[816] David Wilkinson: *Cohesion and Conflict. Lessons from the Study of Three-Party Interaction*. 1976, Frances Pinter.

[817] G.S. Wilkinson: Reciprocal food-sharing in the vampire bat. *Nature*, 308. évf. (1984), 181–184. p.

[818] Bernard Williams: Deciding to believe. In *Problems of the Self*. 1973, Cambridge University Press, 136–151. p.

[819] Timothy Williamson: Converse relations. *The Philosophical Review*, 94. évf. (1985) 2. sz., 249–262. p.

[820] Timothy Williamson: First-order logics for comparative similarity. *Notre Dame Journal of Formal Logic*, 29. évf. (1988) 4. sz., 457–481. p.

[821] Timothy Williamson: Rational failures of the kk principle. In C. Bicchieri – R. C. Jeffrey – B. Skyrms (szerk.): *The Logic of Strategy*. 1999, Oxford University Press, 101–118. p.

[822] Timothy Williamson: *Knowledge and Its Limits*. 2000, Oxford University Press.

[823] Timothy Williamson: Gettier cases in epistemic logic. *Inquiry*, 56. évf. (2013) 1. sz., 1–14. p.

[824] Edward O. Wilson – Bert Hölldobler: Eusociality: origin and consequences. *Proc. Natl. Acad. Sci. U.S.A.*, 102. évf. (2005) 38. sz., 13367–13371. p.

[825] Edward O. Wilson – David Sloan Wilson: Rethinking the theoretical foundation of sociobiology. *Q Rev Biol*, 82. évf. (2007) 4. sz., 327–348. p.

[826] Edward.O. Wilson: *Sociobiology: the new synthesis*. 1975, Harvard University Press.

[827] Robert Wilson: An axiomatic model of logrolling. *The American Economic Review*, 59. évf. (1969) 3. sz., 331–341. p.

[828] Robert Wilson: On the theory of aggregation. *Journal of Economic Theory*, 10. évf. (1975), 89–99. p.

[829] Jonathan Wright: Altruism as a signal: Zahavi's alternative to kin selection and reciprocity. *Journal of Avian Biology*, 30. évf. (1999) 1. sz., 108–115. p.

[830] Larry Wright: Functions. *The Philosophical Review*, 82. évf. (1973) 2. sz., 139–168. p.

[831] H. P. Young: Social choice scoring functions. *SIAM Journal of Applied Mathematics*, 28. évf. (1975) 4. sz., 824–838. p.

[832] H. Peyton Young: The dynamics of social innovation.

[833] H. Peyton Young: An axiomatization of borda's rule. *Journal of Economic Theory*, 9. évf. (1974), 43–52. p.

[834] H. Peyton Young: The evolution of conventions. *Econometrica*, 61. évf. (1993. Jan) 1. sz., 57–84. p.

[835] H. Peyton Young: The economics of convention. *The Journal of Economic Perspectives*, 10. évf. (1996) 2. sz., 105–122. p.

[836] A. Zahavi: Mate selection: Selection for a handicap. *Journal of Theoretical Biology*, 53. évf. (1975) 1. sz., 205–214. p.

[837] A. Zahavi: Indirect selection and individual selection in sociobiology: My personal views on theories of social behaviour. *Animal Behaviour*, 65. évf. (2003), 859–863. p.

[838] Amotz Zahavi: Altruism as a handicap: The limitations of kin selection and reciprocity. *Journal of Avian Biology*, 26. évf. (1995) 1. sz., 1–3. p.

[839] Amotz Zahavi – Avishag Zahavi: *The Handicap Principle. A Missing Piece of Darwin's Puzzle*. 1997, Oxford University Press.

[840] R.B. Zajonc: Érzelem és gondolkodás (a preferenciák nem igényelnek következtetéseket). In György Hunyady (szerk.): *Szociálpszichológiai tanulmányok*. Budapest, 1991, Tankönyvkiadó, 7–55. p.

[841] Richard Zeckhauser: Voting systems, honest preferences, and pareto optimality. *The Americal Political Science Review*, 67. évf. (1973) 3. sz., 934–946. p.

[842] Jing Zhu: Intention and volition. *Canadian Journal of Philosophy*, 34. évf. (2004) 2. sz., 175–194. p.

[843] Gary Zukav: *The Dancing Wu Li Masters. An Overview of the New Physics*. 1984, Bantam Books.