

Cselekvéselméleti alapfogalmak

Syi

i@syi.hu

February 6, 2013

Contents

Előszó	1
1 Logikai, módszertani alapok	3
1.1 Logika	4
1.1.1 Relációelmélet	6
1.1.2 Nulladrendű logika	11
1.1.3 Elsőrendű logika	25
1.1.4 Kategoriális logika	27
1.1.5 Modális logika	37
1.1.6 Deontikus logika	40
1.1.7 Preferencialogika	43
1.1.8 Doxasztikus logika	47
1.1.9 Másodrendű logika	47
2 A társadalom modelljei	49
2.1 Ego	49
2.1.1 Szelf	49
2.2 A világ	54
3 Intencionáltság	55
3.1 Intencionalitás	55
3.1.1 Brentano	55
3.1.2 Husserl	56
3.1.3 Ascombe, Dennet, Fodor, Searle	56
3.1.4 egyéb	59
3.1.5 Neurológiai alapok	59
3.2 Reprezentáció	59
3.3 Teleológia vs. teleonómia	60
3.4 Érvényesség	60
3.5 Intenció	61
3.5.1 Bratman előtti korszak: vágy-vélekedés modellek	61
3.5.2 Bratman: BDI-elmélet	61
3.5.3 más elméletek	62
3.6 Vágy	62
3.6.1 Castelfranchi: vélekedés és vágy	63
3.6.2 Cross: diszkrepancia alapú vágyelmélet	64
3.6.3 magasabbrendű vágyak	65

3.7	Másodrendű vágy: érték	65
3.7.1	Osgood	66
3.8	Másodrendű preferencia: aggregálás	66
3.8.1	Archimédeszi preferencia	67
3.8.2	Lexikografikus preferencia	69
3.8.3	Hasznosság-károosság	70
3.9	Akarat	70
3.10	Választás, döntés	71
3.10.1	Ariely, viselkedésgazdaságtan	71
3.10.2	Racionalitásértelmezések	72
3.10.3	döntési szabályok	72
3.10.4	példák	73
3.10.5	“döntési irracionalitások”	74
3.10.6	Közösségi döntések problémái	74
3.11	Értelem, tudat	75
3.11.1	Közös tudás	75
4	Társadalmi cselekvés	77
4.1	Cselekvés	78
4.1.1	Szent Anzelm cselekvéstipológiája	78
4.1.2	Tevés/tevékenység és tartózkodás	79
4.1.3	Magasabbrendű cselekvés	80
4.2	Cselekvő	80
4.2.1	stít-elmélet	80
4.2.2	program-elmélet	81
4.2.3	Davidson	81
4.2.4	Ego és Alter	81
5	A társadalmi cselekvés típusai	83
5.1	Fiske relációs elmélete	83
5.2	Szabályszerűségek	84
5.3	Ego és a világ	84
5.3.1	Célracionalitás	84
5.3.2	Értékracionalitás	84
5.3.3	Érzelem	85
5.3.4	Habitus	88
5.3.5	Miért hamis a mindig racionális ember mítosza?	90
6	Társadalmi kapcsolat	93
6.1	Ágensre irányultság: szociális preferencia	94
6.2	Magasabbrendű beállítódás: büntetni akarás	96
6.3	Társas értékorientáció	97
6.3.1	Egoizmus	104
6.3.2	Altruizmus	106
6.3.3	Kooperáció	108
6.3.4	Versengés	108
6.3.5	Agresszió	108

6.3.6	Autoagresszió	110
6.3.7	Önfeláldozás	111
6.3.8	Destrukció	111
6.4	Dinamikus kapcsolat: reciprocitás	111
6.5	Méltányosságelméletek	111
6.6	Szolidaritás	112
6.7	Dominancia, státus	112
6.8	A társadalmi kapcsolat típusai	112
6.9	Neurológiai folyamatok	115
7	Harcon alapuló társadalmi kapcsolat	117
7.1	Érdek	117
7.2	Hatalom	117
7.2.1	Hatalom az állatok közt	117
7.2.2	Hatalom emberek közt	118
7.3	A kooperáció biológiai mechanizmusai	118
7.3.1	Rokoni altruizmus	118
7.3.2	Zöld szakáll hatás	119
7.3.3	Spite	119
7.3.4	Csoportszelekció	120
7.3.5	A costly signaling és a handicap principle	120
8	Békés társadalmi kapcsolat	121
8.1	Strukturális kényszerek: játékelmélet	121
8.2	Magjátékok	123
8.3	A játékok statikus kiterjesztése	137
8.3.1	Aszimmetrikus játékok	137
8.3.2	Többalternatívás játékok	138
8.3.3	Többszereplős játékok	139
8.3.4	Nem szigorúan rendezett játékok	139
8.4	A játékok magasabbrendű kiterjesztése: büntetés	140
8.4.1	Másodrendű játékok – költséges büntetés	140
8.4.2	Értékelésparadoxon	141
8.4.3	Aszinkron játékok	141
8.4.4	alkujátékok	142
8.5	A játékok dinamikus kiterjesztése: társadalmi emlékezet	143
8.5.1	Reciprocitás	144
8.5.2	Direkt reciprocitás játékok	144
8.6	A játékok közösségi kiterjesztése: társadalmi értékelés	148
8.6.1	Indirekt reciprocitás játékok	148
8.6.2	Értékelési rendszerek	150
8.7	Koordinációs problémák, közös tudás, aszimmetrikus információ	150
8.8	minimal social situation	150
8.9	Kollektív cselekvés, kollektív döntés	151
8.10	Koalícióelmélet	151

9 Társadalmi pozíció	155
9.1 Imperszonális deontika	155
9.2 Pozíciók: jogosultságok és kötelezettségek	155
9.2.1 Hohfeld	155
9.2.2 Kanger-Lindahl	156
9.2.3 post Kanger-Lindahl	156
9.3 Szabadság	157
9.4 Társadalmi szerep	158
9.5 Privacy	158
10 Norma	159
10.1 Legitimitás	159
10.1.1 Normakövetés	159
10.2 Játékelméleti magyarázat a normák keletkezésére	159
10.3 Normatipologizálás	160
10.3.1 Normatípusok a szankció típusa szerint	160
10.3.2 Normatípusok a logikai rend szerint	161
10.3.3 Normatípusok a deontikus modalitás szerint	162
10.3.4 Normatípusok a tartalom szerint	162
10.3.5 Normatípusok a címzett és autoritás viszonya alapján	162
10.3.6 Normatípusok a feltételelesség alapján	162
10.3.7 Normatípusok a címzett alapján	162
10.3.8 Normatípusok az alkalom alapján	162
10.3.9 Normatípusok az autoritás alapján	162
10.4 Rítus, narratíva	162
11 Normavezérelt társadalmi kapcsolat	165
11.1 Uralom	165
11.2 Hatalom és uralom viszonya	165
11.3 Szervezet	165
12 Egyéb	167
12.1 mechanizmusok	167
12.2 Antinómiák, önreferencia	167
12.3 Hibák, javítandók	167

Előszó

“There is only one social science.” (Jack Hirshleifer [190])

A cselekvésselmélet a társadalmi cselekvések leírásával, megértésével, magyarázatával foglalkozik. Ebben a minőségében – definíció szerint – a módszertani individualizmus talaján áll. A cselekvésselmélet 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 cselekvésselmélet fogalmának kiterjesztett jelentését fogadom el, de szentelek egy külön fejezetet a szűkebb értelmezésnek is.

Jack Hirshleifer egy 1985-ös cikkéből származik a oldal tetején olvasható idézet [190], 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.



Chapter 1

Logikai, módszertani alapok

"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. " (Whitehead, An Introduction to Mathematics, 61.o.)

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.

Az egyértelműség és a konzisztencia kialakításának és fenntartásának lehetőségén túl van még egy érv a logikai apparátus alkalmazása mellett. Egyfajta takarékosági, gazdaságossági szempontból is gondolhatunk a logikára, amikor is az elméletépítések során újra és újra alkalmazzuk a logika különböző következtetési szabályait, kalkulusait, szillogizmusait.

¹ A logika iránt érdeklődő olvasóknak magyar nyelven Ruzsa Imre könyveit ajánlom (például ezt: [311]), a matematikai logikai kérdésekben Ferenczi Miklós könyvére támaszkodtam [128].

gizmusait. A következtetési szabályok folyamatos alkalmazása (újrahasznosítása) nagyon fontos hozzáadéka 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, ánya stb.) többszáz tételből álló rendszerét kifejezhetnénk 5-6 alapfogalom segítségével, de nem tesszük ezt³. 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 – a fejezet mottójába foglalt – figyelmeztetésére.

1.1 Logika

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 az adott nyelven. E kettősséggel voltaképpen azt ígéri minden logika, hogy a kalkulussokra 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 gazdaság-

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

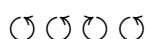
³A rokonsági fogalmakról szóló tanulmányomat meghivatkozhatónak kell tenni.

gos. Említettem már azt a köznapi példát, miszerint hivatkozhatnánk 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 rokonság fogalmát 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 következőket rendszer használhatóságát felülről korlátozza a használt ontológia nyelv kifejezőereje, azaz hogy a nyelven 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 problémakört alkot, a nyelvek használhatósága szempontjából is kiemelt kérdésnek számít, ám könyvem mondanivalója szempontjából elhagyhatónak tűnik, ezért nem tárgyalom.



Az általam használni kívánt logikák egymásra épülését két dimenzióhoz igazodva lehet

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

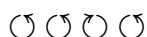
kifejezni. Figyelni lehet arra, hogy i) hanyadrendű logikáról van szó, illetve ii) extenzionális vagy intenzionális logikát vizsgálunk-e [311].

!!!!!!!!!!!!!!!!!!!!

beírni egy mondatot az extenzionális és intenzionális logika meghatározásairól

!!!!!!!!!!!!!!!!!!!!

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.

1.1.1 Relációelmélet

A formális nyelvek által definiált fogalmak mindig valamilyen struktúrát alkotnak, a struktúrák vázát pedig mindig az adott nyelven kifejezett relációk adják. Bármilyen relációs struktúrát építünk is fel, minden szükség van arra, hogy a relációkat magukat is jellemezni tudjuk.

metarelációk, relációtulajdonságok, relációkalkulus

műveletek változói: operandusok

műveletek tulajdonságai:

asszociativitás

kommutativitás

disztributivitás

idempotencia

abszorpció

monotonitás

involúció

affinitás

dualitás

igazságmegőrzés

hamisságmegőrzés (validity)

Lehetne kibontani egy extenzionális relációelméletet, amely a relációk tulajdonságait, a leggyakrabban használt – elemi és összetett – relációtípusokat, a relációk közti függvényeket és műveletek, illetve azok tulajdonságait, valamint a mindezekből felépíthető elemi struktúrákat, kalkulusokat mutatná be. Ezt más helyütt megtettem, itt csak hivatkozom rá.

fogalmi szint és nyelvi szint, fogalmak és terminusok elválasztása

terminológiai körbevezetés

Metarelációk

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 kicsit egyszerűsítve – kétfajta dolgot teszünk. Egyfelől szószerű nyelvi jeleket képzünk, hogy azokkal létező vagy tételezett dolgok halmazára rámutathassunk, 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

ÁLLÍTÁS

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

Az 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. Valamely nyelv állításait változókkal "helyettesítve" ezeket az A, B, C, ..., P, Q stb. jelekkel reprezentálhatjuk. Az állítások logikai értéke igaz vagy hamis lehet, amit szöveges megjelenítésben 'igaz' vagy 'hamis' karakteres formák, táblázatokban a '0' és az '1' értékek feltüntetésével fejeznek 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

TAGADÁS

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. Szinonimájaként szokás a 'negáció' terminust használni. Több jelet is használnak rá, leggyakrabban: $\neg$, $\sim$.

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.

TERMINUS

A világ adott dolgaiból álló halmaz létezését reprezentáló nyelvi megnyilatkozás. A terminus általános fogalmának nincs önálló jele, mint változóra x,y,z stb. jelekkel hivatkozhatunk.

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

KOMPLEMENTER [TERMINUS]

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 szinonim 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 megjelölése, a formális nyelvekben

⁶Itt hivatkozni kell a metarelációs fejezet művelet fogalmára.

nem alkalmaznak önálló jelölést a műveletre. Mivel a későbbiekben szeretném megkülönböztetni a komplementálás és a tagadás műveleteit, ezért a tagadásra egyaránt gyakran használt két jelet ($\sim$, $\neg$) itt eltérő jelentések mentén fogom alkalmazni: a komplementálásra a ' $\sim$ ', a tagadásra a ' $\neg$ ' jelet használom.

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, annyiban szemben áll a komplementálás fogalmával.

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 közülök több is a *közvetlen következtetések* vagyis az egyetlen állításból levonható következtetés 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, amikor is az egyes mondatok mint összetevők már a megnyilatkozás mint egész részeként, 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.

A konverzió már az elemi relációelméletben is fontos szerepet játszik, és mivel ezen a szinten nagyon egyszerűen szemléltethető a művelet, érdemes ezzel kezdeni.

KONVERZIÓ

Ha van egy $R(x, y)$ relációnk, akkor annak a konverze (vagyis a konverzreláció) úgy állítható elő, hogy a relációba kapcsolt elemek, a relátumok sorrendjét felcseréljük egymással:

$R^{+1}(x, y) = R(y, x)$ — javítani: valamelyik latex modul + jelet csinál - helyett a kitevőben

A definícióból adódik, hogy ha a konverziótkétszer egymás után alkalmazzuk, akkor az eredeti relációhoz jutunk vissza. A konverzió jele a relációelméletben általában a felső kitevőbe tett '-1' érték. Ritkábban szokták ugyan használni a fogalom magyar fordítását, de azért elő-előfordul a 'megfordítás' terminus is. 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 (23.o.), amikor a mondatokon belüli összetevők szerkezetét, akkor kategoriális konverzióról (34.o.). Az első esetben a művelet értelmezési tartománya a propozíciókból, a második esetben a kategoriális propozíciók alkotórészeiből áll⁷.

⁷Lásd a 'Cambridge Dictionary of Philosophy' szótár "converse" lexikai egységét [17].

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).

INVERZÍÓ

Ha van egy $R(x, y)$ relációnk, akkor annak az inverze (vagyis a inverzreláció) úgy állítható elő, hogy a relációba kapcsolt elemek, a relátumok komplementereit kapcsoljuk össze:

$$R^{+c}(x, y) = R(\sim x, \sim y) \text{ — minusz jelet csíholni a kitevőbe}$$

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 (35.o.), mivel tárgyalásához szükség van még további fogalmak ismeretére.

Itt elemzem viszont az általános kontrapozíció műveletét, hogy aztán a későbbi fejezetekben kitérjek majd a további típusaira is.

KONTRAPOZÍCIÓ

Ha van egy $R(x, y)$ relációnk, akkor annak a kontrapozíciója (vagyis a reláció kontraponáltja) úgy állítható elő, hogy a relációba kapcsolt elemek, a relátumok komplementereit fordított sorrendben kapcsoljuk össze:

$$R^{+p}(x, y) = R(\sim y, \sim x) \text{ — javítani a kitevőt}$$

A kontrapozíció szinonimájaként szokták olykor használni a ‘transzpozíció’ terminust.

A kontrapozíciót is értelmezhetjük mind a mondatyszerű, mind a szószerű megnyilatkozások szintjén. Összekapcsolt mondatok szerkezetének átalakításakor kondicionális kontrapozícióról (23.o.), mondatokon belüli összetevők átstrukturálásakor kategoriális kontrapozícióról (36.o.) beszélünk. A kontrapozíció definíciójából az is látszik, hogy a kontrapozíció nem más, mint a konverziónak az inverze, illetve az inverzió konverze.

$$\Leftarrow \iff \Rightarrow$$

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 “tárgyasítjuk” 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 hímnemű. 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.⁸ 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 veszünk, hiszen a reifikátum már nem mond semmit a reláció másik relátumáról. A reláció másik relátumát általában a reláció inverze segítségével tudjuk könnyen reifikálni.

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

A reifikáció műveletét tovább absztrahálhatjuk, és tetszőleg absztrakt individuumokat létrehozhatunk, ami a tudásreprezentáció számára nagyon praktikus előnyökkel szolgálhat [54, 41].

REIFIKÁCIÓ

Ha van egy $R(x, y)$ relációnk, akkor annak a reifikációja úgy állítható elő, hogy a relációnak csak az egyik relátumát vesszük figyelembe.

$R^{+r}(x, y) = R(x)$ — jelölést találni rá, javítani a kitevőt

Mivel – bináris reláció esetén – két relátumról beszélhetünk, elméletileg mindig kétféle reifikáció létezhet, egy bal-, illetve egy jobboldali.

mi az, amikor a logikai négyszög elemei közötti kapcsolatot jellemzi (kontadiktórius, kontrárius stb.) Ez az értelmezés különbözik a Berger-Luckman-féle reifikációfelfogástól.

1.1.2 Nulladrendű logika

A logikai elméletek közül elsőként a nulladrendű extenzionális logikát veszem elő. A későbbiekben inkább ‘kijelentéslogikaként’ fogok hivatkozni rá, de ebben a könyvben szinonimaként kezelem még az ‘állításlogika’, illetve a ‘propozicionális logika’ terminusokat is. A kijelentéslogika a mondatyszerű 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.

⁸Ez a művelet olyan, mint a logikai műveletek között a projekció (lásd 14.o.).

A kijelentéslogika alapl műveletei

Amennyiben veszünk állításokat és ezeket valamilyen művelet segítségével összekapcsoljuk, 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.

LOGIKAI MŰVELET

A logikai művelet állításokra vonatkozó művelet. Mivel a fogalom nagyon széles körben használt, rengeteg szinonimája termelődött ki. A fogalomra használt további terminusok: funktor, junktor, logikai kapcsoló, logikai konnektívum, konnektívum, proposicionális kapcsoló, mondatoperátor, mondatkapcsoló, proposicionális operátor.

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 részben 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.⁹

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.

⁹Mivel az alacsonyabb változószámú műveletek mindig besorolhatóak a magasabb változószámú műveletek közé, ezért a nullaváltozós műveletek egyben egy- és kétváltozós műveletnek is tekinthetők.

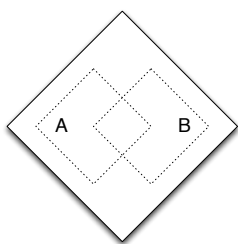
A műveletek jellemzésekor bemutatnom azokat a logikai jeleket, amelyeket valamilyen formális logikai nyelvben elterjedt módon használnak, azokat a – magyar – természetes nyelvű szavakat, kifejezéseket, amelyeket a fontosabb elméletekben gyakrabban felbukkantak, megmutatom az adott művelet ellentétét, 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 figyelmen kívül hagyni. Ilyen esetekben “önkényesen” választottam a versengő nevek közül.

$$\nrightarrow \iff \leftarrow$$

ELLENTMONDÁS



A	B	$\perp$
1	1	0
1	0	0
0	1	0
0	0	0

jelei: $\perp$

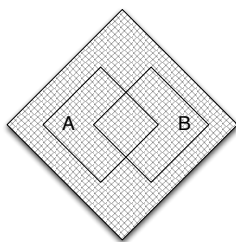
nevei: ellentmondás, kontradikció, sosem igaz, falsum, trivialitás

ellentéte: tautológia

formulái: $(A \wedge \neg A) \wedge (B \wedge \neg B)$

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.

TAUTOLÓGIA

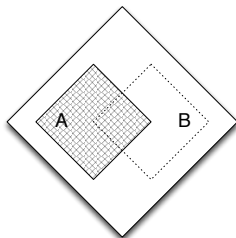


A	B	$\top$
1	1	1
1	0	1
0	1	1
0	0	1

jelei: $\top, \models$
 nevei: tautológia, mindig igaz
 ellentéte: ellentmondás
 formulái: $(A \vee \neg A) \wedge (B \vee \neg B)$

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 mondom, ‘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. Ez a művelet nulláris, konstans.

PROJEKCIÓ

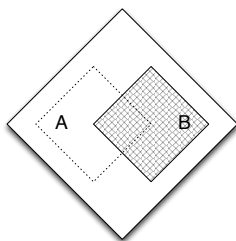


A	B	$A \triangleleft B$
1	1	1
1	0	1
0	1	0
0	0	0

jelei: $\triangleleft$
 nevei: projekció, első projekció, bal művelet
 ellentéte: nem-projekció
 formulái: A

Ritkán szokták a projekció műveletét ebben a kontextusban használni, helyette valamilyik 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.¹⁰ Amennyiben a bal oldali állítást (tehát A-t) projektáljuk, akkor “egyszerű” projekcióról beszélünk. Ez tartalmilag megegyezik a bal oldali propozíció tételezésével.

RETROPROJEKCIÓ

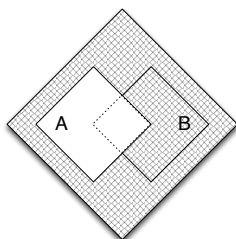


A	B	$A \triangleright B$
1	1	1
1	0	0
0	1	1
0	0	0

jelei: $\triangleright$
 nevei: retroprojekció, második projekció, jobb művelet
 ellentéte: nem-retroprojekció
 formulái: B

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.

KONTRAPROJEKCIÓ



A	B	$A \ntriangleleft B$
1	1	0
1	0	0
0	1	1
0	0	1

jelei: $\ntriangleleft$
 nevei: kontraprojekció
 ellentéte: projekció
 formulái: $A \vee \neg B$

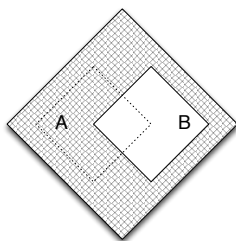
Ez a művelet a projekció ellentéte, tartalmilag megegyezik az A operandusba foglalt kijelentés tagadásával. Azért nevezem kontraprojekciónak, mert a kontra előtaggal valaminek az ellentétét szokás kifejezni, és ebben az esetben pont erről, a projekció tagadásáról van szó.

KONTRA-RETROPROJEKCIÓ

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

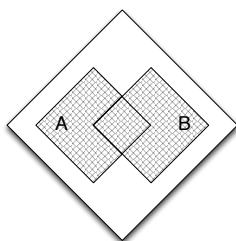
DISZJUNKCIÓ

¹⁰A projekció hasonlít a reifikáció műveletére, hiszen mindkét esetben kiválasztunk egyet két összekapcsolt elemből: logikai műveletek esetén az operandusok, reifikáció esetén a relátumok közül (lásd 11.o.).



A	B	$A \not\supset B$
1	1	0
1	0	1
0	1	0
0	0	1

jelei: $\not\supset$
 nevei: kontra-retroprojekció
 ellentéte: retroprojekció
 formulái: $\neg B$



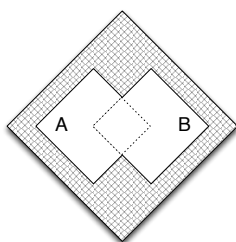
A	B	$A \vee B$
1	1	1
1	0	1
0	1	1
0	0	0

jelei: $\vee, \vee, +$
 nevei: diszjunkció, OR, vagy, alternáció, megengedő vagy, gyenge diszjunkció, inkluzív diszjunkció
 ellentéte: konnegáció
 formulái: $A \vee B$

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 egyiféjű igazságának köszönhető-e. A diszjunkció a hétköznapi kommunikáció során leggyakrabban használt műveletek közé tartozik.

kommutatív, asszociatív, disztributív (konjunkcióra), idempotens, monoton

KONNEGÁCIÓ



A	B	$A \not\supset B$
1	1	0
1	0	0
0	1	0
0	0	1

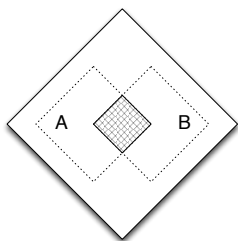
jelei: $\not\supset, \downarrow, \nabla, ||$
 nevei: konnegáció, sem-sem, NOR, not OR, Webb-művelet, egyesített tagadás
 ellentéte: konjunkció
 formulái: $\neg A \wedge \neg B$

A konnegáció a diszjunkció tagadása. Ha a hétköznapi nyelvben is használatos 'sem-sem' formulát nézzük, 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, amit a kommutativitás tételezésével fejezhetünk ki. 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.

‘Sem utódja, sem boldog őse.’

KONJUNKCIÓ

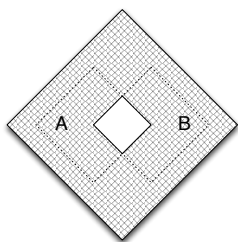


A	B	$A \wedge B$
1	1	1
1	0	0
0	1	0
0	0	0

jelei: $\wedge, \&, \cdot$
 nevei: AND, és
 ellentéte: exklúzió
 formulái: $A \wedge B$

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.

EXKLÚZIÓ



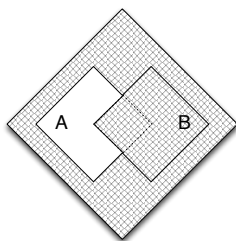
A	B	$A \nabla B$
1	1	0
1	0	1
0	1	1
0	0	0

jelei: $\nabla, \uparrow, |, ||$
 nevei: exklúzió, kizárás, NAND, not AND, nem-és, Sheffer-művelet, vagylagos tagadás, alternatív tagadás, kontrárius
 ellentéte: konjunkció
 formulái: $\neg A \vee \neg B$

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.

mindkettő egyszerre nem lehet igaz, de lehet egyszerre hamis, illetve egyenként igaz lehet barna, lehet zöld (lehet, hogy egyik sem, de egyszerre nem lehet barna is, zöld is)

KONDISZJUNKCIÓ

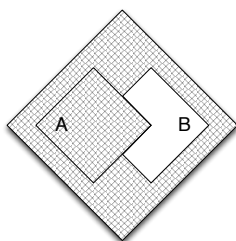


A	B	$A \rightarrow B$
1	1	1
1	0	0
0	1	1
0	0	1

jelei: $\rightarrow, \supset, \Rightarrow$
 nevei: kondicionális, materiális kondicionális, implikáció, materiális implikáció, 'ha, akkor', szubaltern
 ellentéte: szubtrakció
 formulái: $\neg A \vee B$

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

RETROKONDITIONÁLIS



A	B	$A \leftarrow B$
1	1	1
1	0	1
0	1	0
0	0	1

jelei: $\leftarrow, \subset, \Leftarrow$
 nevei: retrokondicionális, konverz kondicionális, konverz implikáció
 ellentéte: retroszubtrakció
 formulái: $A \vee \neg B$

- (R) ha B , akkor A
- (R) nincs B A nélkül
- (R) A , ha B

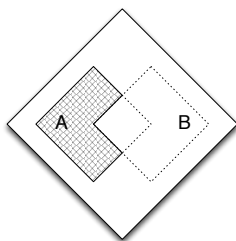
Ez a kondicionális művelet konverze. Tartamilag, formailag mindenben hasonlatos a kondicionálisra, csak éppen minden fordítva, "jobbról-balra van". A megszokás (véltőleg a balról-jobbra olvasás dominanciája) miatt a kondicionálist érezzük természetesnek.

"Nincs karácsony Corvin nélkül."

A csak akkor, ha B

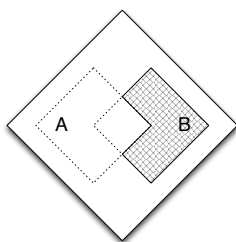
SZUBTRAKCIÓ

A szubtrakció a kondicionális művelet tagadása, amely csak akkor igaz, amikor A "ön-magában" 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.



A	B	$A \nrightarrow B$
1	1	0
1	0	1
0	1	0
0	0	0

jelei: $\nrightarrow, \not\rightarrow$
 nevei: szubtrakció, kivonás, nem-kondicionális, nem-implikáció, abjunkció
 ellentéte: kondicionális
 formulái: $A \vee \neg B$



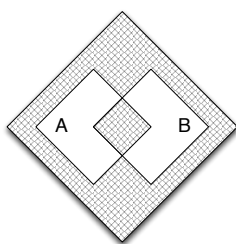
A	B	$A \nleftarrow B$
1	1	0
1	0	0
0	1	1
0	0	0

jelei: $\nleftarrow, \not\leftarrow$
 nevei: retroszubtrakció, konverz szubtrakció, konverz nem-implikáció
 ellentéte: retrokondicionális
 formulái: $\neg A \wedge B$

RETROSZUBTRAKCIÓN

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.

BIKONDITIONÁLIS



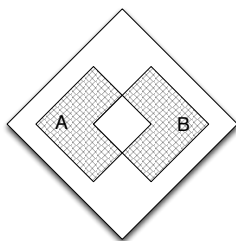
A	B	$A \leftrightarrow B$
1	1	1
1	0	0
0	1	0
0	0	1

jelei: $\leftrightarrow, \Leftrightarrow, \equiv$
 nevei: bikondicionális, ekvivalencia, materiális ekvivalencia, biimplikáció, XNOR, akkor és csak akkor, csakkor, iff
 ellentéte: szubtrakció
 formulái: $(\neg A \vee B) \wedge (A \vee \neg B)$

A bikondicionális művelete a kondicionális és annak konverze, a retrokondicionális egyidejű fennállását jelenti, azaz az implikatív jelleg mindkét irányban érvényes.

BISZUBTRAKCIÓN

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



A	B	$A \nleftrightarrow B$
1	1	0
1	0	1
0	1	1
0	0	0

jelei: $\nleftrightarrow, \leftrightarrow, \oplus, \vee$
 nevei: exkluzív diszjunkció, kizáró diszjunkció, erős diszjunkció, exkluzív OR, XOR, vagy-vagy, kizáró vagy, kontravalencia, antivalencia, kontradiktórius
 ellentéte: bikondicionális
 formulái: $\neg(A \leftrightarrow B)$

Table 1.1: logikai műveletek áttekintő táblázata

művelet	jele	formula	ellentéte	11	10	01	00
ellentmondás	$\perp$	$A \wedge \neg A$	tautológia	0	0	0	0
konnegáció	$\neg, \downarrow, \nabla, $	$\neg A \wedge \neg B$	diszjunkció	0	0	0	1
retroszubtrakció	$\nleftarrow, \leftarrow$	$\neg A \wedge B$	retrokondicionális	0	0	1	0
kontraprojekció	$\nrightarrow$	$\neg A$	projekció	0	0	1	1
szubtrakció	$\dashv, \rightarrow$	$A \wedge \neg B$	kondicionális	0	1	0	0
kontra-retroprojekció	$\nrightarrow$	$\neg B$	retroprojekció	0	1	0	1
biszubtrakció	$\nleftrightarrow, \leftrightarrow, \oplus$	$\neg(A \leftrightarrow B)$	bikondicionális	0	1	1	0
exklúzió	$\nmid, \uparrow, , $	$\neg A \vee \neg B$	konjunkció	0	1	1	1
konjunkció	$\wedge, \&$	$A \wedge B$	exklúzió	1	0	0	0
bikondicionális	$\leftrightarrow, \Leftrightarrow, \equiv$	$A \leftrightarrow B$	biszubtrakció	1	0	0	1
retroprojekció	$\triangleright$	B	kontra-retroprojekció	1	0	1	0
kondicionális	$\rightarrow, \supset, \Rightarrow$	$\neg A \vee B$	szubtrakció	1	0	1	1
projekció	$\triangleleft$	A	kontraprojekció	1	1	0	0
retrokondicionális	$\leftarrow, \subset, \Leftarrow$	$A \vee \neg B$	retroszubtrakció	1	1	0	1
diszjunkció	$\vee, \vee$	$A \vee B$	konnegáció	1	1	1	0
tautológia	$\top, \models$	$A \vee \neg A$	ellentmondás	1	1	1	1

igaz, de úgy, hogy egyszerre sosem lehetnek igazak. Ezt fejezi ki a kizáró vagy exkluzív diszjunkció megnevezés.

“Vagy bolondok vagyunk s elveszünk egy szálig,
 Vagy ez a mi hitünk valóságra válik.”
 csak egy maradhat
 élet vagy halál
 mindent vagy semmit

Áttekintő táblázat

A tizenhat bináris művelet áttekintése után érdemes őket egy táblázatban egyszerre felsorolni, mert így érdekes összefüggések vehetők észre.

A műveletek igazságtáblázatai (illetve a geometriai ábrák) alapján észrevehetünk pár érdekességeket. Nyolc olyan művelet van, amely a két összetevő igazsága esetén igaz, és nyolc olyan, amely ilyenkor hamis. Nyolc művelt másik nyolcnak a tagadása. Az egymás tagadásaként tekinthető 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 (kommutatív), illetve nyolc aszimmetrikus (nem-kommutatív) művelet láthatunk. A nyolc aszimmetrikus műveletből négy a maradék négy művelet konverzánek tekinthető. Nyolc olyan művelet van, amelynél mindkét operandus hamissága esetén a művelet egésze is hamis, míg nyolc másik művelet esetében a művelet egésze ilyenkor igaz.

Fontos szabályszerűség A nulladrendű logika zárt a logikai operátorokra, ami azt jelenti, hogy minden művelet eredménye proposíció, vagyis a műveletekkel képzett új állítások a logika érvényességi körén belül maradnak.

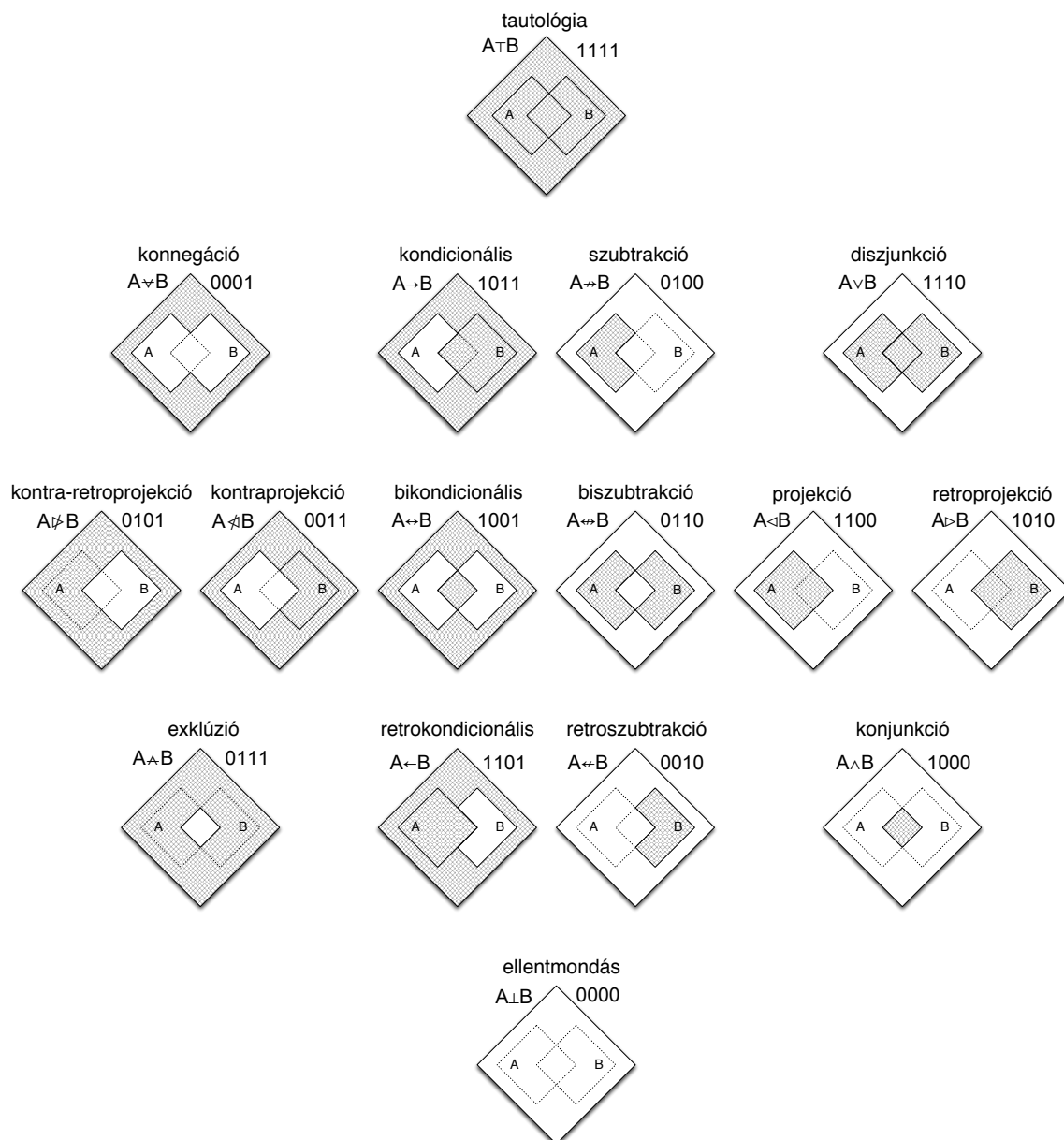


Figure 1.1: logikai műveletek

Table 1.2: a logikai műveletek visszavezetése az exklúzió műveletére

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

Műveletkiküszöbölés

Mind az exklúzió, NAND- vagy Sheffer-művelet ($\downarrow$), mind a konnegáció, NOR- vagy Webb-művelet ($\Downarrow$) azzal az egyedi tulajdonsággal rendelkezik, hogy segítségével minden más logikai művelet kifejezhető. A tagadás – például – a következőképpen írható fel velük:

$$\neg A \Leftrightarrow A \downarrow A, \text{ illetve } \neg A \Leftrightarrow A \Downarrow A$$

Ez azt jelenti, hogy a propozicionális logika minden művelete átalakítható lenne, és minden formulája felírható lenne e két logikai operátor valamelyikével. Az exklúzióra támaszkodva az alábbi módon lehetne a 16 bináris műveletet reprezentálni:

Ez a megoldás egyszerűsítene az adott logikai nyelvet. Ha nem lenne 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 bogyozza 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.

Kijelentéskalkulus

rules of inference

szabályok a tagadásra szabályok a kondicionálisra szabályok a konjunkcióra szabályok a diszjunkcióra szabályok a bikondicionálisra

PC Propositional Calculus kijelentéskalkulus

conjunction elimination zárt
conjunction introduction

Közvetlen következtetések, átalakítási műveletek logikai operátorok esetén

A korábban tárgyalt (9.o.) általános átalakítási műveletek speciális eseteiként definiálhatunk szűkebb értelmezési tartományú, de egyébként ugyanolyan jelentésű fogalmakat. A konverziót, inverziót és kontrapozíciót értelmezhetjük a logikai műveletek egyikére, a kondicionálisra vetítve is. A következő speciális átalakítási műveleteket definiálhatjuk.

KONDITIONÁLIS KONVERZIÓ

A kondicionális konverzió az állításlogikában a 'kondicionális' logikai operátorra alkalmazott konverzió (amikor megfordítjuk a kondicionális két operandusát). A konverzió általános definíciójához képest itt csak annyi változás történt, hogy az értelmezési tartományt leszűkítettük a proposíciók halmazára.

$A \rightarrow B$ állítás konverze: $B \rightarrow A$

Gyakran használt szinonimája az 'implikációs konverzió' kifejezés, én a névadási konzisztencia miatt választottam a 'kondicionális' jelzős szerkezetet.

KONDITIONÁLIS KONTRAPOZÍCIÓ

A kondicionális kontrapozíció az állításlogikában a 'kondicionális' logikai operátorra alkalmazott kontrapozíció (amikor megfordítjuk a kondicionális két operandusának tagadását).

$A \rightarrow B$ állítás kontrapozíciója: $\neg B \rightarrow \neg A$

A terminus szinonimája az implikációs kontrapozíció.

KONDITIONÁLIS INVERZIÓ

A kondicionális konverzió az állításlogikában a 'kondicionális' logikai operátorra alkalmazott inverzió (amikor tagadjuk a kondicionális két operandusát).

$A \rightarrow B$ állítás inverze: $\neg A \rightarrow \neg B$

A három kondicionális átalakítási művelet közt fontos összefüggéseket állapíthatunk meg. Az állításlogika következtetési szabályai közt (24.o.) szerepel a kontrapozíció törvénye, amelynek értelmében ha a kondicionális igaz, akkor a kontrapozíciója is igaz (és fordítva). Ebből viszont könnyen levezethető az is, hogy az állításlogikában az inverzió logikailag megegyezik a konverzióval (egymás kontraponáltjai). Ez azt is jelenti egyben, hogy ha egy kijelentés konverze igaz, akkor igaz a kijelentés inverze is (és fordítva).

p	q	1	2	3
Reductio ad absurdum	negation	introduction	$((A \vdash B) \wedge (A \vdash \neg B)) \vdash \neg A$	
Reductio ad absurdum		related to the law of excluded middle	$((\neg A \vdash B) \wedge (\neg A \vdash \neg B)) \vdash A$	
Noncontradiction	negation	elimination	$((A) \wedge (\neg A)) \vdash B$	
involúció	double negation	elimination	$\neg\neg A \vdash A$	
	double negation	introduction	$A \vdash \neg\neg A$	
Deduction theorem	conditional	introduction	$(A \vdash B) \vdash (A \rightarrow B)$	
Modus ponens	conditional	elimination	$((A) \wedge (A \rightarrow B)) \vdash B$	
Modus tollens			$((\neg B) \wedge (A \rightarrow B)) \vdash \neg A$	
Adjunction	conjunction	introduction	$((A) \wedge (B)) \vdash (A \wedge B)$	
Simplification	conjunction	elimination	$(A \wedge B) \vdash A$	
Addition	disjunction	introduction	$A \vdash (A \vee B)$	
Case analysis				
Disjunctive syllogism			$((A \vee B) \wedge (\neg A)) \vdash B$	
	biconditional introduction		$((A \rightarrow B) \wedge (B \rightarrow A)) \vdash (A \leftrightarrow B)$	
	biconditional elimination		$((A \leftrightarrow B) \wedge A) \vdash B$	
	tautology	introduction	$\vdash \top$	

Állításlogikai következtetési szabályok

logikai ekvivalenciák:

diszjunkció: kommutativitás, asszociativitás, disztributivitás, De Morgan

konjunkció: kommutativitás, asszociativitás, disztributivitás, De Morgan

implikáció kiküszöbölése

elvalencia kiküszöbölése

kontrapozíció

kettős negáció kiküszöbölése

műveletek és szillogizmusok közti kapcsolatok

értékelés, áthelyezés a megfelelő helyre

Szillogizmusok

kategorikus szillogizmus

hipotetikus szillogizmus

diszjunktív szillogizmus

[17] 894-896.

művelet	követk. sz.
ellentmondás	
konnegáció	
retroszubtrakció	
kontraprojekció	
szubtrakció	
kontra-retroprojekció	
biszubtrakció	
exklúzió	
konjunkció	conjunction elimination
bikondicionális	biconditional elimination
retroprojekció	
kondicionális	conditional elimination
projekció	
retrokondicionális	
diszjunkció	disjunction introduction
tautológia	tautology introduction

1.1.3 Elsőrendű logika

A nulladrendű logika a kijelentések, a mondatok szintjén keresi a szabályszerűségeket. A mondat, az állítás ilyenkor bonthatatlan egészset alkot, nem tudjuk, nem tudhatjuk, mi van a proposíciók 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.

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 predikátumlogikai elemzés világába. Ezen a szinten 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 írják le, akkor részben pontosítják, részben igazodnak a formális nyelvelmélet terminológiájához, ezért a neveket logikai konstansokkal (individuumok neveivel) fejezik ki, a predikátumok helyett pedig vagy a relációk vagy a függvények fogalmait használják. A predikátumlogikai elemzések számára be kell vezetnünk az alábbi fogalmakat.

KONSTANS

Értékét nem változtató név, individuumokat azonosító nyelvi megnyilatkozás. szinonima: név

PREDIKÁTUM

A predikátum olyan logikai függvény¹¹, amely egy vagy több név mint bemeneti paraméter alapján állítást képez.

A predikátum szinonimájaként szokták még használni a reláció és/vagy a függvény terminusait. Logikai jele általában 'P'.

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 kevésbé érdekesek a logika számára. Ha csak konstansok és relációk lennének ebben a formális nyelvben, akkor ezt a logikát kevésbé 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 tartalmaz, 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' – $\forall x$ 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 az alábbiakkal:

VÁLTOZÓ

Olyan nyelvi jel, amely egy logikai kifejezésben valamely halmaz elemeit mint lehetséges értékfordozókat helyettesíti. Jele: x,y

¹¹Ruzsa Imre a funktor terminust alkalmazza rá, lásd [311].

EGISZTENCIÁLIS KVANTOR

A kijelentések terjedelmét jelentő előfordulásokra vonatkoztatható operátor, amely legalább egy előfordulás létezését tételezi a logikai kifejezésben. Jele: $\exists$

UNIVERZÁLIS KVANTOR

A kijelentések terjedelmét jelentő előfordulásokra vonatkoztatható operátor, amely az összes előfordulás létezését tételezi a logikai kifejezésben. Jele: $\forall$

1.1.4 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 (mennyiségi-leg vagy minőségileg) módosítani lehet a kategóriákat, de a két kategóriát mindenképpen a kopulával kapcsoljuk össze.

KATEGORIÁLIS SZUBJEKTUM

A kategoriális szubjektum az a kategória, amelyre a kategoriális kijelentés szerkezetén belüli másik kategória, a kategoriális predikátum vonatkozik. Szinonimaként lehetne használni a 'szubjektum, logikai konstans, individuumnév, név' terminusokat, jeleként az 'S' betűt.

KATEGORIÁLIS PREDIKÁTUM

A kategoriális predikátum az a kategória, amellyel a kategoriális kijelentés szerkezetén belüli másik kategóriára, a kategoriális szubjektumra vonatkozóan állítunk valamit. Szinonimaként lehetne használni a 'reláció és a függvény' terminusokat, jeleként a 'P' betűt.

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 – évezredek 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 módon 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 (ahol 'S' a szubjektum, 'P' a predikátum jele):

kijelentéstípus	jele	latin mondat	magyar mondat	modern formula
egyetemes állító	A	Omne S est P.	Minden S [van] P.	$\forall s \in S : P(s)$
egyetemes tagadó	I	Nullum S est P.	Egyetlen S sem P.	$\forall s \in S : \neg P(s)$
partikuláris állító	E	Quoddam S est P.	Néhány S [van] P.	$\exists s \in S : P(s)$
partikuláris tagadó	O	Quoddam S non est P.	Néhány S nem P.	$\exists s \in S : \neg P(s)$

Az évezredek 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. Azt is meg kell itt jegyezzem, hogy akkor is előfordult inkonzisztencia, amikor a kopula tagadása helyett a 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 összemenni 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.

Arisztotelésznél még nem így volt, de a követői a négyféle állítást már egyetlen rendszerbe szervezték szemléletes ábrával reprezentálva az állítások közti összefüggéseket, és innentől kezdve logikai négyszöggént hivatkoztak rájuk. Ezt azonban már a következő fejezetben bontom ki alaposabban.

Logikai négyszög

A kategoriális logika alapjait Arisztotelész (ie.384 - ie.322) fektette le, de ő még nem rajzolt ábrákat. Pár évszázad múlva Apuleius (125 - 180), majd Boethius (480 - 524) volt az, aki a négyféle kijelentés közti kapcsolatokat ábrában kezdte mutatni. Az ilyen az ábrákat így szokták Apuleius négyszöge vagy Boethius-féle négyzet is nevezni, de használják még a logikai négyszög, logikai négyzet, valamint az ellentétek négyszöge vagy az oppozíciós négyszög kifejezéseket is [39][269].¹² Az ábra így néz ki:

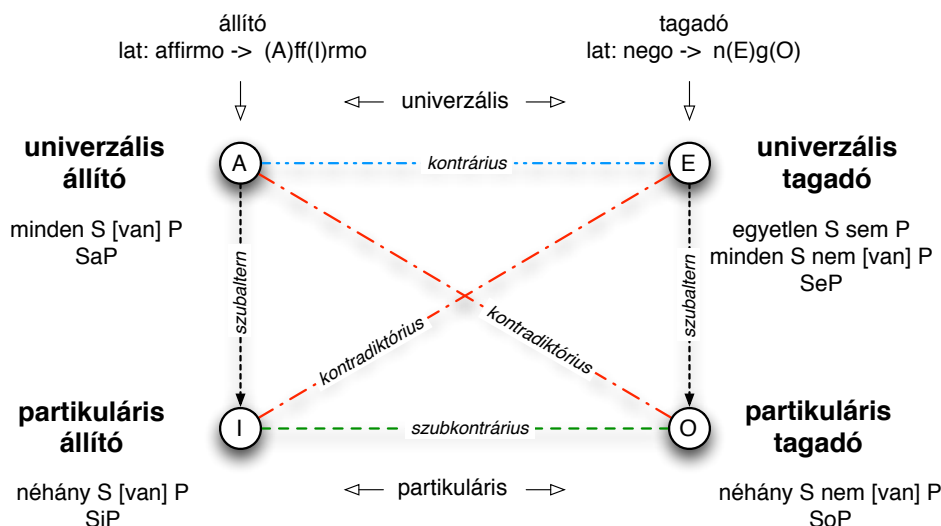


Figure 1.2: logikai négyszög

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 a vizuális megjelenítést, azonban ő volt az, aki alaposan leírta a kijelentések közti viszonyrendszert, és kiderült, a kijelentéstípusok közti kapcsolatok majd mindegyike valamifajta ellentétet fejezett ki. A logikai négyzet elemei közt évszázadokon keresztül három ellentétes viszonyt tételeztek: i) a *kontradiktórus* (ellentmondó), ii) a *kontrárius* (ellentétes), valamint iii) a *szubkontrárius* (alárendelt ellentétes) kapcsolatot. Ezeken felül rögzítettek még egy alárendeltségi viszonyt kifejező *szubaltern* (alárendelt), *superaltern* (fölerendelt) kapcsolatpárt is. Ezen kapcsolatokat, illetve modernkori megfelelőiket az alábbi táblázatban mutatom meg:

¹²Mivel a négyszög továbbfejlesztésekor tartalmilag sokféle irányban elmozdultak, az eredeti négyszöget ma már kvantifikációs négyszögnek is nevezik.

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	$A \nabla B$
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 [283].

A kritikák egyik vonulata úgy szólt [315], 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 [209]. Mivel egy üres terminusról (üres osztályról) minden állítható¹³, 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

¹³Az egyetemes állítás mindig átalakítható a két kategóriát kondicionális művelettel összekapcsoló 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.

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 individuum, előfordulás. Az univerzális kategorikus kijelentések esetében ezt nem tudjuk mindig garantálni, ezért kimondhatjuk a tézist, hogy ezeknek nincs egzisztenciális importja.

??????????

Russel a denotálásról szóló tanulmányában megoldást javasol a denotálás, referálás értelmezésére üres termék, nem létező obejektumok esetében [310] Russel megoldását lehet-e itt alkalmazni?

másról szól Roderick Chisholm cike, de van benne egy használható megoldás az üres termék 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. [73, 147]

* * *
* * *

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

De használták a logikai négyzetet más célokra is. Ami itt a legfontosabb: Canterburyi Szent Anzelm (1033-1109) olyan cselekvéstipológiát dolgozott ki, amely a logikai négyzetre épült. Fontossága miatt az anzelmiánus elméletet később részletesebben bemutatom majd (79.o.)

Logikai hatszög

A logikai négyzettel kapcsolatban felmerültek másfajta problémák is, ezek azonban sokkal inkább a továbbfejlesztési lehetőségek keresésére ösztönözték a kutatókat, mint a négyszöggondolat feladására. A négyszög hatszögesítését Augustin Sesmat és Robert Blanché (1898-1975) egymástól függetlenül, közel egyidőben végezte el az 1950-es években, de közel ötven éven keresztül eredményeiknek nem volt jelentős hatása a tudományos gondolkodásra. Csak a 2000-es években változott ez a helyzet [39].

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 érdekesebb 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 legkülönbözőbb 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, valamint, hogy a szubaltern relációk még több sarokpont között behúzhatóak. A teljes kapcsolatrendszer mutatja az 1.3. ábra.

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 1.4. 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,

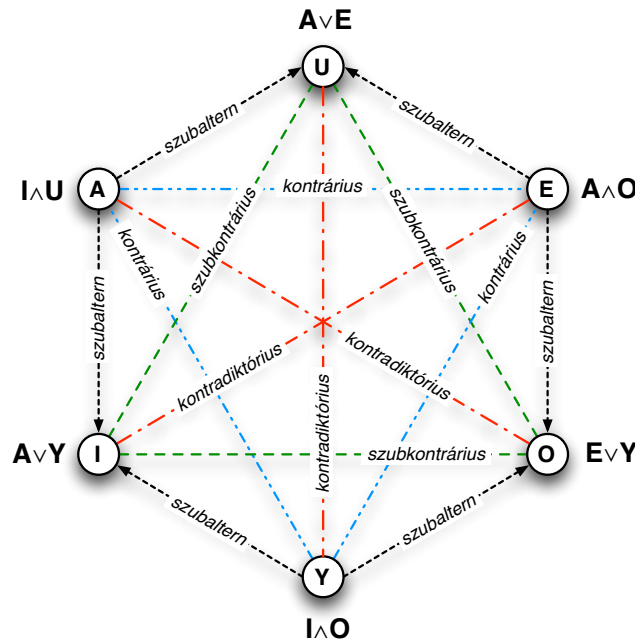


Figure 1.3: logikai hatszög

ezek a “tisztán” másodrendű logikai operátorok, és ezekkel tudta Blanché egy összefüggő logikai szerkezetbe rendezni (1.5. ábra).

A geometriai szemlélet érvényesítését, további hasznosítását azonban itt már nem lehetett leállítani.¹⁴ 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 1.6. ábra mutatja, és hogy könnyebben felismerhetők legyenek a benne “rejtőző” hatszögek, az 1.7. 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 megjelentek azok a cikkek, amelyek logikai n-szögekről és a problémakör geometrizálásáról szóltak [266, 268]. 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 sz-

¹⁴A logikai relációk háromdimenziós megjelenítéséről lásd: [343].

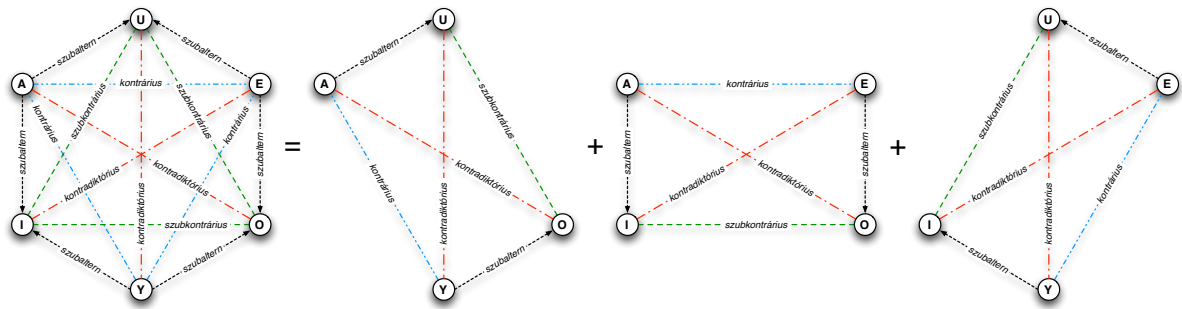


Figure 1.4: logikai hatszög és négyszögek

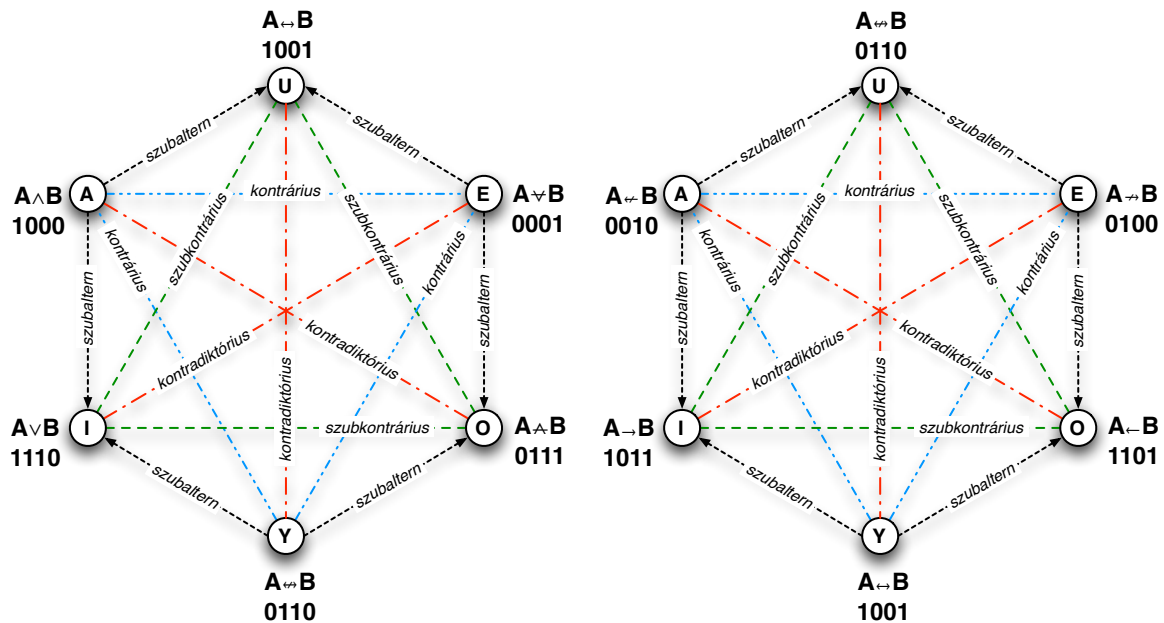


Figure 1.5: Blanché tíz logikai operátora két hatszögbe fogva

abályként is működtek. A már ismertetett átalakítási műveletek meghatározásában (9.o.) csak annyi változásra van itt szükség, hogy az értelmezési tartományt le kell szűkíteni a kategoriális kijelentések halmazára.

KATEGORIÁLIS KONVERZIÓ

Olyan konverzió, amelynek értelmezési tartománya a kategoriális kijelentések halmaza. A kategoriális konverzió praktikusán 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

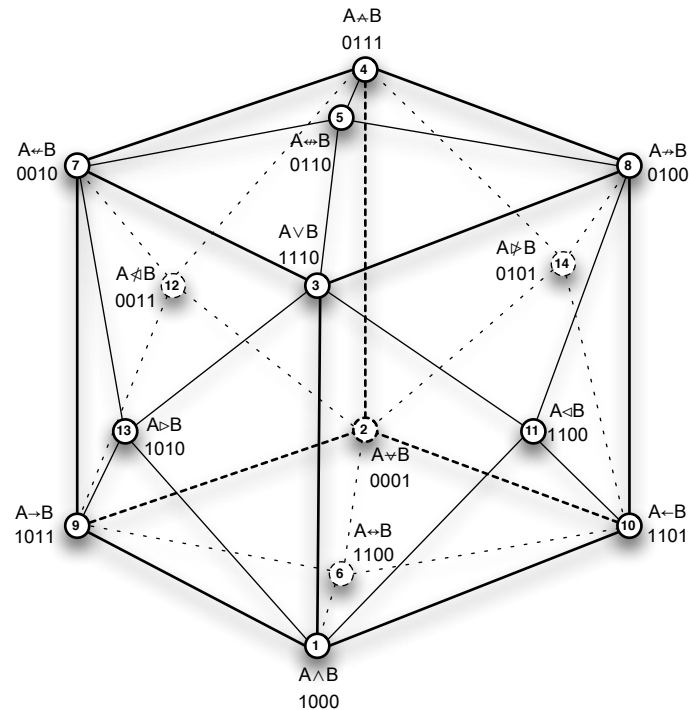


Figure 1.6: logikai operátorok tetrahedronja

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 általános tárgyalásakor nem mutattam be az *obverzió* műveletét, mert ez igazából csak a kategoriális kijelentések esetében értelmezhető (az átalakítási műveletek általános tárgyalásakor viszont még nem volt szó kategoriális kijelentésekről). Ezért itt most új műveletként kell az obverziót definiálnom.

OBVERZIÓ

Olyan átalakítási művelet, amelynek érintetlenül hagyja a kategoriális kijelentés univerzális-partikuláris dimenzióban vett minőségét, de megcseréli a pozitív-negatív tengelyen vett értékét, megtartja a szubjektum és a predikátum sorrendjét, valamint a predikátumterminus helyébe annak komplementumát teszi. A művelet értelmezési tartománya a kategoriális kijelentések halmaza. Szinonímájaként szokták használni az átalakítás terminust is, ami 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 szinonímájaként kell érteni, akkor mégiscsak használni fogom a terminust a továbbiakban.

A kondicionális obverzió műveletét nem használják (bár a kettős negáció törvényét talán annak lenne tekinteni). Vélhetőleg azért nem definiálták ezt a műveletet az állításlogika színterén, mert ott hiányzik a pozitív és negatív (állító és tagadó) kijelentések megkülönböztetése, ami viszont a kategoriális logikában adott.

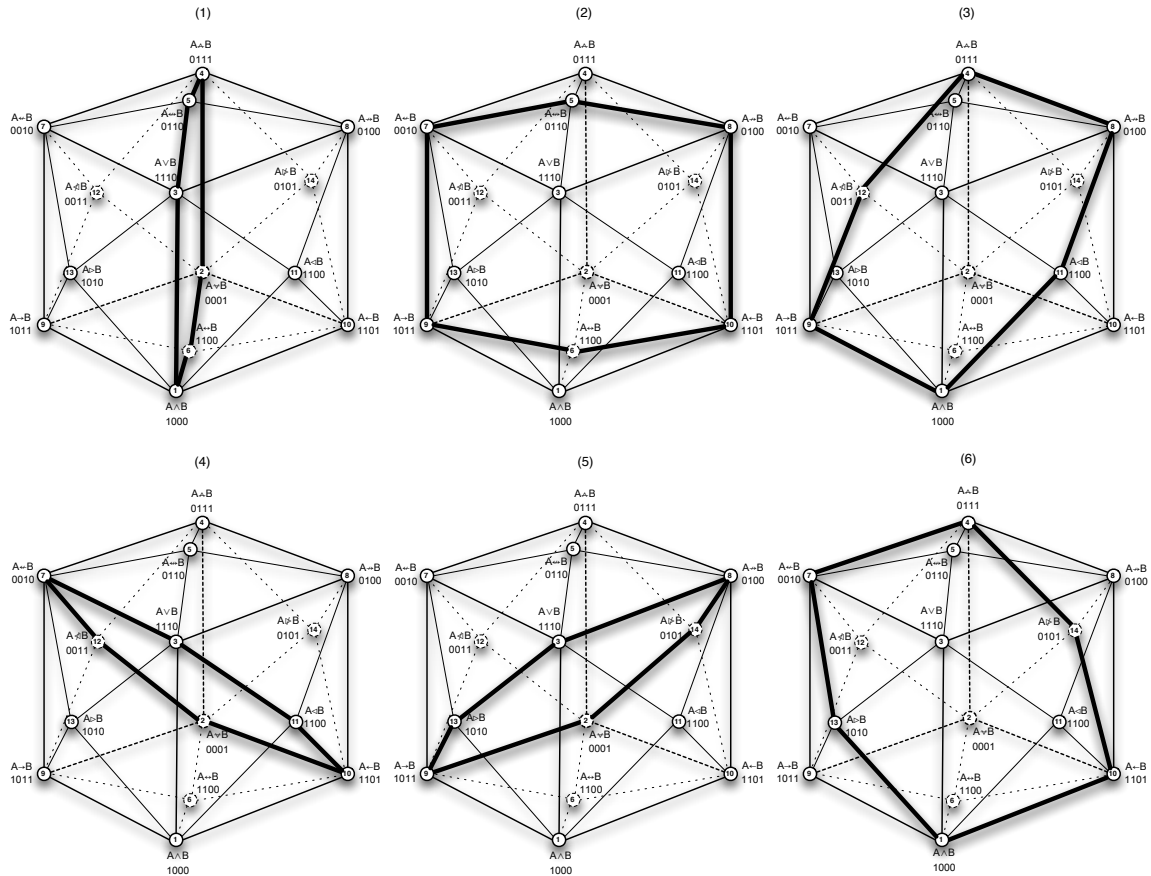


Figure 1.7: hatszögek a tetrahexahedronban

A kategoriális kijelentések színterén is értelmezhető a kontrapozíció művelete, ami itt kifejezhető a konverzió és obverzió segítségével is [17, 895].

TELJES KATEGORIÁLIS KONTRAPÓZÍCIÓ

Olyan kontrapozíció, amelynek az értelmezési tartománya a kategoriális kijelentések halmaza. A kategoriális kontrapozíció praktikusán annyit tesz, hogy megcseréli a kijelentés szubjektumának és predikátumának komplementumait¹⁵. Szinonímája a 'kategoriális kontrapozíció'. A művelet felírható a következő alakban is: kategoriális kontrapozíció = átalakítás + megfordítás + átalakítás.

Ha az (A)-propozíció ("Minden S [van] P.") igaz, akkor annak kategoriális kontraponáltja ("Minden nem-P [van] nem-S.") is igaz, tehát ez az átalakítás használható közvetlen következtetési szabályként. Ugyanez áll az (O)-propozícióra is ("Néhány S nem [van] P."), hiszen a kontraponáltja ("Néhány nem-P nem [van] nem-S.") igaz.

RÉSZLEGES KATEGORIÁLIS KONTRAPÓZÍCIÓ

¹⁵ Az S és P terminusok komplementumait 'nem-S' és 'nem-P' terminusokkal jelölöm.

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

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

KATEGORIÁLIS INVERZIÓ

Az a művelet, amelynek az értelmezési tartománya a kategoriális kijelentések halmaza. A kategoriális inverzió kicseréli a kijelentés két kategóriáját azok negáltjaira, de megtartja a sorrendjüket.

Értékelés az inverzióról. Van-e közvetlen következtetési potenciálja.

Az egyes kategoriális átalakítások tárgyalásakor jeleztem, de most érdemes egy áttekintő táblázatban bemutatni, hogy ezek közül melyek azok az átalakítási műveletek, amelyek közvetlen következtetésekre (vagyis logikai igazságok levonására) adnak lehetőséget (37.o.).

Table 1.3: közvetlen következtetéstípusok egymáshoz való viszonya

premissza	konverzió konverzió	obverzió obverzió	részleges kontrapozíció + konverzió	teljes kontrapozíció + obverzió
<i>SaP</i>		<i>Se¬P</i>	<i>¬PeS</i>	<i>¬Pa¬S</i>
<i>SeP</i>	<i>PeS</i>	<i>Sa¬P</i>	<i>¬PiS</i>	<i>¬Po¬S</i>
<i>SiP</i>	<i>PiS</i>	<i>So¬P</i>		
<i>SoP</i>		<i>Si¬P</i>	<i>¬PiS</i>	<i>¬Po¬S</i>

Értékelés röviden

Kategorikus szillogizmusok

no?

1.1.5 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 szükségszerűen igaz, akkor igaz is-e egyben – ilyen és ezekhez hasonló kérdésekre próbál meg válaszokat találni a modális

logika¹⁶. Mivel sokféleképpen módosíthatjuk az állításainkat, sokféle modalitást kezelhetünk. Egy Q kijelentés kapcsán mondhatjuk azt, hogy ‘tudom, hogy Q ’, ‘azt hiszem, hogy Q ’, ‘gyakran Q ’, ‘biztos, hogy Q ’, ‘tilos Q ’, ‘szükségszerű, hogy Q ’ stb. Ilyen esetekben az állításunkat (Q -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 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. A lehetséges világok közti átmenetekre, elérhetőségekre 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, amely a következő formulába írható fel (ahol $\Box$ a szükségszerűség modális operátora):

$$\Box A \rightarrow A$$

A formulát úgy interpretálhatjuk, hogy ‘ami szükségszerűen igaz, az igaz’. Első “neki-futásra” 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’.

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 érkezhettek 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’.

¹⁶A modális logika két leggyakrabban hivatkozott forrása: [70] és [193]

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 magyarázóképeket és különböző modális logikákat eredményez.

Anélkül, hogy a részletekbe belemennék, inkább csak az érdekesség kedvéért megemlítem, hogy az alábbi relációtulajdonságokkal, az alábbi logikai sémákkal (mindig igaz állításokkal, tehát tautológiákkal) határozzák meg a legfontosabb modális kalkulusokat (modális logikákat):

(D)	$\Box A \rightarrow \Diamond A$	szeriális (balról totális)
(T)	$\Box A \rightarrow A$	reflexív
(B)	$A \rightarrow \Box \Diamond A$	szimmetrikus
(S4)	$\Box A \rightarrow \Box \Box A$	transzitiv
(S5)	$\Diamond A \rightarrow \Box \Diamond A$	euklideszi

Ezek közül a továbbiakban egyedül a (D) séma és a hozzá tartozó modális logika, a deontikus logika lesz fontos, a többivel nem foglalkozom. A deontikus logikát a következő fejezetben elemzem kicsit alaposabban.

Itt kell viszont még kitérni 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:

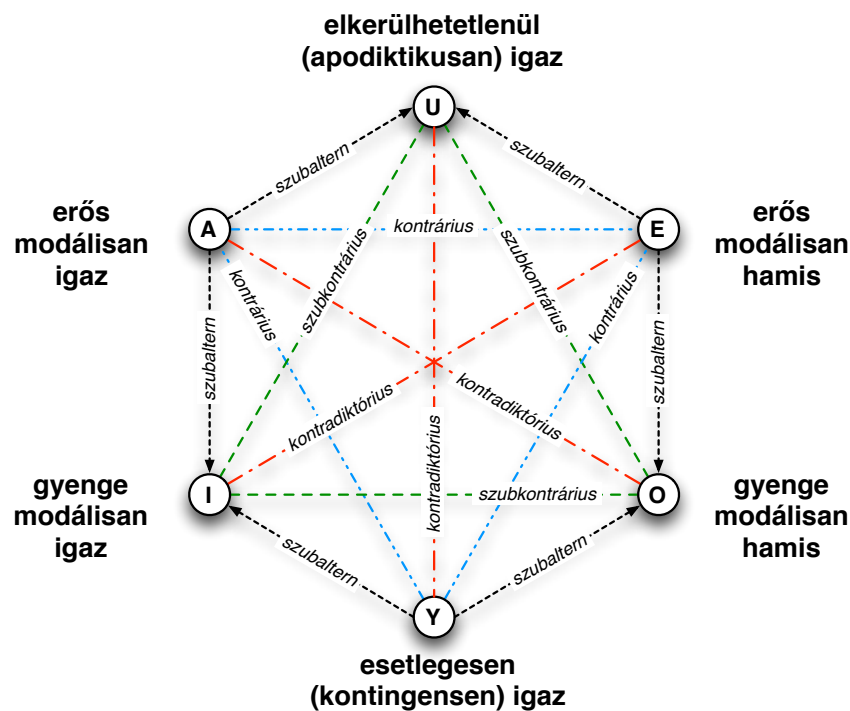


Figure 1.8: modális logikai hatszög

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 [38][39][269]. De, ahogy említettem, a cselekvésemélet felépítése szempontjából csak a deontikus logikát tudom majd hasznosítani, ezért a továbbiakban csak ezt mutatom be kicsit részletesebben.

1.1.6 Deontikus logika

imperatív logika

Alf Ross [305]

von Wright, An Essay in Deontic Logic [389]

Anderson, A Reduction of Deontic Logic to Alethic Modal Logic, [6]

Lemmon [222]

Anderson, [8]

Bengt Hansson [174, 173]

Chellas [68], [69]

Kanger [205]

Seegerberg [324]

Makinson [242]

A különböző logikák közül a cselekvésemélet számára a deontikus logika a legfontosabb. Hogy miért van ez így, azt csak később fejtem ki. A modális logika különböző kalkulusait meghatározó sémák közül a (D) séma az, amit a deontikus logikához kapcsolhatunk.

(D) $\Box A \rightarrow \Diamond A$

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 jövőre vonatkozó – kötelezések és megengedések világáról. Nem tud azonban semmi értelmeset mondani valami nagyon fontos dologról: arról, hogy mikor, miért valósulnak meg a deontikus operátorokkal.

????? Erről bővebben: [352]

Deontikus operátorok

A deontikus logika első elméletében von Wright még négy deontikus operátorról írt [393].
problémák

A ‘megengedett’ fogalmát azonban a hétköznapi kommunikáció során többféle értelemben is használjuk. Erre Joerden hoz szemléletes példát [200, 204].

(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 baleset során nemcsak megengedett, hanem kötelező a segítség. Ez ebben az esetben az is jelenti, hogy itt a kötelezettség maga után vonja a megengedés aktusát is, vagyis 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. Szó sincs ugyanis arról, hogy a jogosítvány megszerzésével kötelességünk lenne autót vezetni. Ebben az esetben a kötelezés már kontradiktórikus viszonyban áll a megengedéssel. Az első példa a (I), a második az (O) esetbe tartozik.

Kálmán László, nyest

"De nemcsak a jövő idő, a tartozás vagy a lehetőség kifejezéséből származhat ?kell? jelentésű ige. A magyar kell például finnugor eredetű szó, és a többi finnugor nyelvben általában (a ?kell? jelentésen kívül) ?kíván, kívánatos, tetszik? jelentésű igék származnak ugyanabból a finnugor tőből (ami a finnugor alapnyelvben *kelke lehetett). A magyarban is vannak nyomai ilyesfajta használatnak: kellemes, kelleti magát."

direktíva

Table 1.4: deontikus operátorok lehetséges megnevezései

kötelező	megparancsolt, kellő, támogatott, obligát, ajánlott, muszáj, szükséges, nem mellőzhető, nem lehet nem megtenni
tilos	tiltott, nem lehetséges, nem szabad, nem megengedett, megengedhetetlen, tűrhetetlen, elfogadhatatlan, kötelező tartózkodni
megengedett	szabad, lehetséges, jogos, nem tilos, nem tiltott, tűrt, elfogadott
mellőzhető	szabad, nem kötelező, elhagyható, mulasztható, jogos, privilégium, kiváltság, immunitás (?), felmentett (?)

A terminológiai gazdagság (és pongolaság) vélhetőleg minden nyelven tetten érhető lenne. Angolul ugyanazt a 'sokszínűséget' tapasztalhatjuk, mint a magyarban. A deontikus operátorok angol névváltozatait az alábbi táblázat mutatja:

obligatory	commanded, duty, compulsory, required, liability (?), subjection (?)
forbidden	impermissible, banned, prohibited, prohibition, disallowed
permitted	permissible, allowed, right, free, claim, power (?)
omissible	non-obligatory, no-duty, duty-free, exemption, privilege, immunity
neutral	indifferent, optional, facultative
prescriptive	non-facultative, non-optional

A logikai hatszög "felfedezése" után G. Kalinowski vezette be a deontikus hatszöget.¹⁷

Új fejlemények: NOT

további kiterjesztés, NOT: Moretti, The Geometry of Standard Deontic Logic [267]

¹⁷G. Kalinowski, La logique des normes. PUF, Paris (1972), idézi: [267].

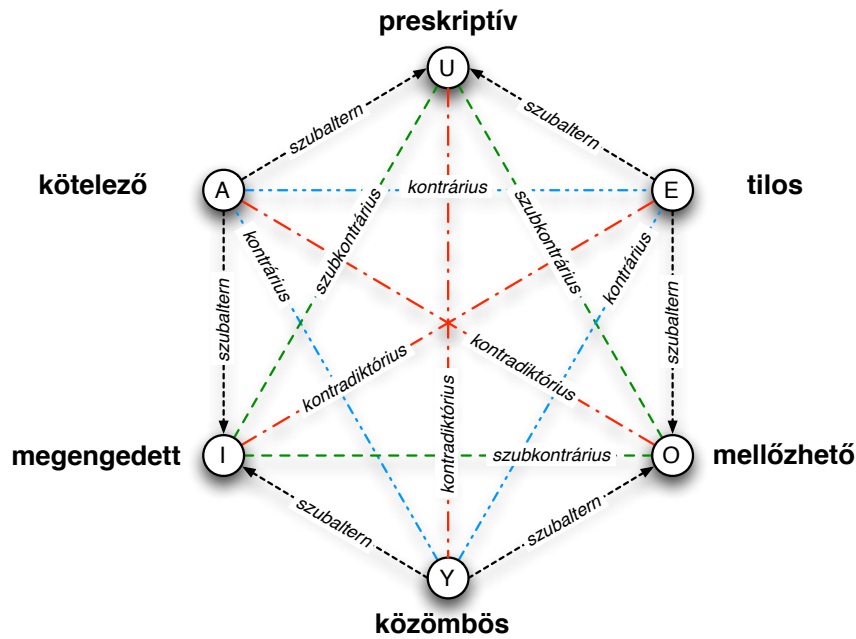


Figure 1.9: deontikus logikai hatszög

supererogation

Urmson problémafelvető tanulmánya: [377] Az elterjedt morálfilozófiai nézet szerint háromféle morális cselekedet lehetséges: a kötelességhez, a megengedéshez és a tiltáshoz igazodó, ám ezzel a fogalomhármassal nem lehet minden morális aktust megragadni. A *szent* és a *hős* cselekvése kilóg ebből a klasszifikációból. 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

az első elméletformálási kísérlet: Chisholm, supererogation and offence [72]

Roderick M. Chisholm és Ernest Sosa: [76]

Roderick M. Chisholm: 1964 [73]

Joel Feinberg: 1961 (annyira nem érdekes) [126]

Paul McNamara: [255]

Jan C. Joerden a túlbuzgóság (Supererogation) kezelhetősége, modellbe emelése érdekében arra tesz javaslatot, hogy a deontikus hatszög helyett inkább tízszöget (Deontological Decagon) kellene használnunk [200].

iszlám protokoll:

megparancsolt	elmulasztásért büntetés jár
ajánlott	elmulasztásért nem jár büntetés, elvégzésért jutalom jár
tiltott	elkövetésért büntetés jár
helytelenített	mellőzésért jutalom jár, gyakorlásért nem jár büntetés
közömbös	sem jutalom, sem büntetés nem jár

operátor	elvégzésekor	mellőzésekor
megparancsolt	nincs semmi	büntetés
ajánlott	jutalom	nincs büntetés
tiltott	büntetés	nincs semmi
helytelenített	nincs büntetés	jutalom
közömbös	nincs semmi	nincs semmi

A supererogation fogalma segíthet tisztázni a jézusi etika legfontosabb tézisének lényegét.

1.1.7 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. A legelemibb rendezésfogalmakat a különféle relációtulajdonságok (aszimmetrikus, antiszimmetrikus, tranzitív, irreflexív stb.) segítségével lehet elkülöníteni egymástól. Így járva el, olyan fogalmakhoz juthatunk, mint szigorú rendezés, részbenrendezés, gyenge rendezés, kvázirendezés, lineáris rendezés, elrendezés, jól-rendezés stb.¹⁸ Ezek a matematikai rendezésfogalmak azonban nagyon általánosak, ami azt jelenti, hogy nagyon 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 részek fogalmáról [337], utóbbi axiómarendszer dolgozott ki a partitív reláció fogalmára, amelyben a három alaptulajdonságon túl még 4-5 további tulajdonságot definiált (Varzi írásai közül lásd [386], a partitív reláció Varzi-féle axiomatizálását magyarul [359]).

Mindez azért érdekes, mert a preferenciarelációra is ki lehet jelenteni, hogy az aszimmetrikus és tranzitív egyben, ám ezzel még biztos nem tudjuk megragadni a preferenciareláció különlegességét, keresni kell tehát még további tulajdonságokat hozzá. Ezt végzik el azok a szerzők, akik a preferencialogika különböző elméleteit dolgozzák ki.

Von Wright preferencialogikája

A preferencia fogalmának első formális elméletét Georg Henrik von Wright adta meg a ‘The Logic of Preference’ című “esszéjében” [391]. Von Wright abból indult ki, hogy

¹⁸Sajnos, a rendezési fogalmak világában sincs terminológiai tisztaság, gyakran találkozhatunk mind a homonimitás, mind a szinonimitás jelenségével. A rendezési fogalmakról bővebben:[358], illetve részben más terminológiát használva: [367].

$$(Pref3) \quad PREF(x,y) \leftrightarrow PREF(x \wedge \neg y, \neg x \wedge y)$$

$$(Pref4) \quad \begin{aligned} PREF(x \vee y, z \vee w) &\leftrightarrow PREF(x \wedge \neg z \wedge \neg w, \neg x \wedge \neg y \wedge z) \wedge \\ &PREF(x \wedge \neg z \wedge \neg w, \neg x \wedge \neg y \wedge w) \wedge PREF(y \wedge \neg z \wedge \neg w, \neg x \wedge \neg y \wedge z) \wedge \\ &PREF(y \wedge \neg z \wedge \neg w, \neg x \wedge \neg y \wedge w) \end{aligned}$$

$$(Pref5) \quad PREF(x,y) \leftrightarrow PREF(x \wedge z, y \wedge z) \wedge PREF(x \wedge \neg z, y \wedge \neg z)$$

Az axiómák rendre a következőket jelentik: a $PREF$ reláció aszimmetrikus (a1), tranzitív (a2), és érvényes a konjunkció (a3), a disztribúció (a4) és az amplifikáció (a5). Az aszimmetrikusság feltételezése azzal jár, hogy nem engedjük meg azt, hogy két 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 propozicionális változókat használnak, ami a [379]

A preferenciareláció axiómái alapján már meghatározható a JÓ (GOOD) és a ROSSZ (BAD) fogalma is. Mivel von Wright a preferenciát a világállapotokra vonatkoztatva definiálta, ezért ebben az elméletben a jó és rossz kategóriája is (világ)állapotokra vonatkozóan érvényes. 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. A meghatározások megfelelő formulái:

$$(Pref6) \quad PREF(GOOD, \neg GOOD)$$

$$(Pref7) \quad PREF(\neg BAD, BAD)$$

A preferenciarelációra támaszkodva von Wright még további két fontos relációt határoz meg: az INDIFFERENCIA (INDIFF), illetve az ÉRTÉKEGYENLŐSÉG (EQVALUE) relációkat. Az indifference reláció (INDIFF) azt fejezi ki, hogy két állapot egyike sem preferált a másikkal szemben, amit az alábbi módon fejezhetünk ki:

$$(Pref8) \quad INDIFF(x,y) \leftrightarrow \neg PREF(x,y) \wedge \neg PREF(y,x)$$

Az értékegyenlőség reláció (EQVALUE) definiálásához – a preferenciareláció meghatározásához hasonlóan – több axiómára is szükség van:

$$(Pref9) \quad EQVALUE(x,y) \rightarrow EQVALUE(y,x) \quad EQVALUE$$

$$(Pref10) \quad EQVALUE(x,y) \wedge EQVALUE(y,z) \rightarrow EQVALUE(x,z) \quad EQVALUE$$

$$(Pref11) \quad EQVALUE(x,y) \leftrightarrow EQVALUE(x \wedge \neg y, \neg x \wedge y)$$

$$(Pref12) \quad \begin{aligned} EQVALUE(x \vee y, z \vee w) &\leftrightarrow EQVALUE(x \wedge \neg z \wedge \neg w, \neg x \wedge \neg y \wedge z) \wedge \\ &EQVALUE(x \wedge \neg z \wedge \neg w, \neg x \wedge \neg y \wedge w) \wedge EQVALUE(y \wedge \neg z \wedge \neg w, \neg x \wedge \neg y \wedge z) \wedge \\ &EQVALUE(y \wedge \neg z \wedge \neg w, \neg x \wedge \neg y \wedge w) \end{aligned}$$

$$(Pref13) \quad EQVALUE(x,y) \leftrightarrow EQVALUE(x \wedge z, y \wedge z) \wedge EQVALUE(x \wedge \neg z, y \wedge \neg z)$$

A fenti formulák az ÉRTÉKEGYENLŐSÉGI reláció szimmetrikusságát (a9), tranzitivitását (a10), illetve a disztribúció (a12), a konjunkció (a11) és az amplifikáció (a13) érvényességét fejezik ki, és a fenti öt formula fennállása – von Wright intuíciója szerint – szükséges és elégséges feltétel a reláció meghatározásához.

Georg Henrik von Wright, 1963-as könyvével kapcsolatban megjelent kritikák, alternatív formalizálási kísérletek után újra elővette a preferenciák problémakörét, és 1972-ben részben újraformálta elméletét.[392] 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 felsorolja azokat a kérdéseket, amelyeket minél pontosabban tisztázni kell egy preferencialogika kidolgozásakor, majd megállapítja, hogy nem igazán arra lát esélyt, hogy ki lehessen dolgozni egy széles körben elfogadott elméletet, mint inkább arra, hogy a különböző intuíciók mentén bár szabatos, de párhuzamos értelmezések alakulhassanak ki. Von Wright újra megpróbálkozik a preferencia fogalmának formalizálásával, de ezt a kísérletet már nem mutatjuk be

Chisholm és Sosa preferencialogikája

Roderick M. Chisholm and Ernest Sosa [76, 77]

Von Wright preferenciák logikájáról szóló elméletére többen (bár nem túl sokan) reagáltak. Más szerzők nem tudtak teljesen azonosulni a von wrighti intuíciókkal, más axiómákat tartottak meghatározónak. Roderick M. Chisholm és Ernest Sosa alternatív formalizmust dolgozott ki von Wright formuláihoz képest:

- (PCS1) $\text{PREF}(x,y) \rightarrow \neg \text{PREF}(y,x)$
- (PCS2) $[\neg \text{PREF}(x,y) \wedge \neg \text{PREF}(y,z)] \rightarrow \neg \text{PREF}(x,z)$
- (PCS3) $[(\neg \text{PREF}(x,\neg x) \wedge \neg \text{PREF}(\neg x,x)) \wedge (\neg \text{PREF}(y,\neg y) \wedge \neg \text{PREF}(\neg y,y))] \rightarrow [\neg \text{PREF}(x,y) \wedge \neg \text{PREF}(y,x)]$
- (PCS4) $[\neg \text{PREF}(y,\neg y) \wedge \neg \text{PREF}(\neg y,y) \rightarrow \text{PREF}(x,y)] \rightarrow \text{PREF}(x,\neg x)$
- (PCS5) $[\neg \text{PREF}(y,\neg y) \wedge \neg \text{PREF}(\neg y,y)] \rightarrow \text{PREF}(y,\neg x) \rightarrow \text{PREF}(x,\neg x)$

A fenti formulákat csak az első két esetben tudjuk a „szokásos módon” értelmezni (a preferenciareláció aszimmetrikus és tranzitív – az 1. és 2. formulák szerint), míg Chisholm és Sosa további három formulájára nincs széles körben elterjedt megnevezés, ami megint csak azt mutatja, hogy az axiomatizálás még mindig nem tudott konszenzus közeli állapotba kerülni a szakterület alapfogalmainak definiálásában.

További elméletek

modern elméletek, mérnökök közt citeFrancescaRossiKristenBrentVenableTobyWalsh
Bengt Hansson [172, 171]

Bengt Hansson – részben – új relációkat bevezetve először megmutatta, hogy von Wright, illetve a Chisholm és Sosa által lefektetett elmélet túl erős, és a preferenciareláció mindennapi használatával kapcsolatos intuícióknak több ponton ellentmondó következményei vannak. Ezért B. 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.

B. Hansson az axiómaiban nem a preferenciarelációt, hanem egy rendezési relációt vesz alapul, melyekre előbb megállapít két rendezési (az 1-es és 2-es állításokat), illetve két disztribúciós axiómát (a 3-as és 4-es állításokat):

- (PBH1) $[P(x,y) \wedge P(y,z)] \rightarrow P(x,z)$
 (PBH2) $P(x,y) \vee P(y,z)$
 (PBH3) $P(x,y) \wedge P(x,z) \rightarrow P(x,y \vee z)$
 (PBH4) $P(x,z) \wedge P(y,z) \rightarrow P(x \vee y, z)$

Kanger is írt egy rövid cikket a preferencialogikáról, amiben Wright elméletét is kritizálta

Svan Ove Hansson, preference logic [176] + S.O. Hansson könyve [177]

From Preference Logics to Preference Languages, and Back, 2010 [42]

Rescher

Reasoning about Preference Dynamics [236]

A von wrighti preferencialogikáról Fenrong Liu: [235]

Preference Change: Approaches from Philosophy, Economics and Psychology [164]

1.1.8 Doxasztikus logika

mentális állapot a tudás és a vélekedés, melyek tartalma egy kijelentés, propozíció.

A vélekedés (belief) logikája.²⁰

Egy vélekedés lehet megalapozatlan (hamis) és megalapozott (igaz).

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.

1.1.9 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 [311]. 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.

²⁰Az angol 'belief' terminus lehetne 'hit'-ként vagy 'hiedelem'-ként fordítani, de a magyarban mindkét terminusnak olyan konnotációja van, ami itt nem kívánatos értelmezést tenne lehetővé, amit szeretnénk elkerülni. A magyar nyelvhasználatban ugyanis mind a hit, mind a hiedelem terminusához valamiféle megalapozatlanságot társítunk. Angolban a 'belief' terminus ebben az értelemben semleges, amit magyarul jobban kifejezhetünk a 'vélekedés' terminusával.

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 [128, 238-239]. A formulák szintjén ez úgy ismerhető fel, hogy az individuumváltozók mellett a predikátumváltozókra is érvényesítjük a formulában használt kvantorokat, valahogy így:

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

Ez a formula a generikus alárendeltje 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 jeleníthetők meg, hogy a 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:

$$\text{tud}(x,p) \quad - \quad x \text{ tudja, hogy } p$$

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

$$p = \text{tud}(y,q) \quad - \quad y \text{ tudja, hogy } q$$

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

$$\text{tud}(x,p) \quad = \quad \text{tud}(x, \text{tud}(y,q)) \quad - \quad x \text{ tudja, hogy } y \text{ tudja, hogy } q$$

Ezen a ponton szépen látszik a predikátum önmagába való beágyazottsága. Ez csak másodrendű logika esetén fordulhat elő, mert itt már megengedjük, hogy a változókat és kvantorokat predikátumra is lehessen alkalmazni.

A predikativitás problémáját megemlíteni, de nem kifejtetni, nem megoldani.

$$\nabla \triangle \nabla \triangle \nabla$$

A másodrendű logika meghatározásának ismeretében már “adja magát” a magasabbrendű logikák értelmezésének lehetősége.

Chapter 2

A társadalom modelljei

2.1 Ego

2.1.1 Szelf

The Emergence of Self Terrence W. Deacon, James W. Haag and Jay Ogilvy

szelf

Ágensek

Ego

Feinberg

Todd E. Feinberg, *Altered Egos. How the Brain Creates the Self*: [127]

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. déjà vu (where the unfamiliar appears familiar)

Capgras szindróma (kicsérélt 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 boundaries 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 mirror 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, background, 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 psychological 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 Doppelgänger 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 autoscopia 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. Autoscopia

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 represent 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 jobbkezevel kanállal akart volna levest enni, miközben a balkeze egy kenyérdarabot rakott volna a szájba 98

hogyan 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 egyrészt 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 tolnáeggyl 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 contribute 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 cannot 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 [371] 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 hydrocephalus.

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 [?]

bontakozó szelf érzése	0-2 hónapos korban
magszef érzése	2-6 hónapos korban
szubjektív szelfézés	7-15 hónapos korban
verbális szelfézés	15 hónapos kortól

Ezek nem egymást váltó szakaszok, megjelenésük után mindegyik továbbra is fennmarad, és az egész élet folyamán “funkcionál”.

Önismeret

[3] de se hitek szelftörténelem önazonosság kijelentések tartalmának Szelfhez kötése
 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

Intencionáltság

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

intencionalitás = valamire vonatkozás (aboutness) (fodor cikke az elemefilozófia könyvben)
irányultság beállítódás motiváltság attitűd diszpozíció

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.” [60, 55]

Brentano szerint az intencionlítás három típusa, osztálya

„Az első alaposztály a legágabb 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

eszközwálasztás legbonyolultabb jelenségéig. Arisztotelész mindezt az *órexisz* fogalmában foglalta össze.” [60, 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.” [60, 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 tárgyasultságnak 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.” [59, 161]

Frederick Chisholm egy lexikonbejegyzésében elemzi Brentano intencionalitás fogalmát. [74] hogyan lehet az intencionális

mondat prefixumokat keres

Mások által már javasolt, de Chisholm által nem megfelelőnek talált kritériumok:

existential generalization

nonexistential occurrence

referential opacity

Megfelelő kritérium:

M intencionális, ha minden p proposícióra 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

3.1.2 Husserl

Husserl cogito cogitatio

3.1.3 Ascombe, Dennet, Fodor, Searle

Ascombe, Dennet, Fodor, Searle

Dennet: az intencionalitás filozófiája, [103]

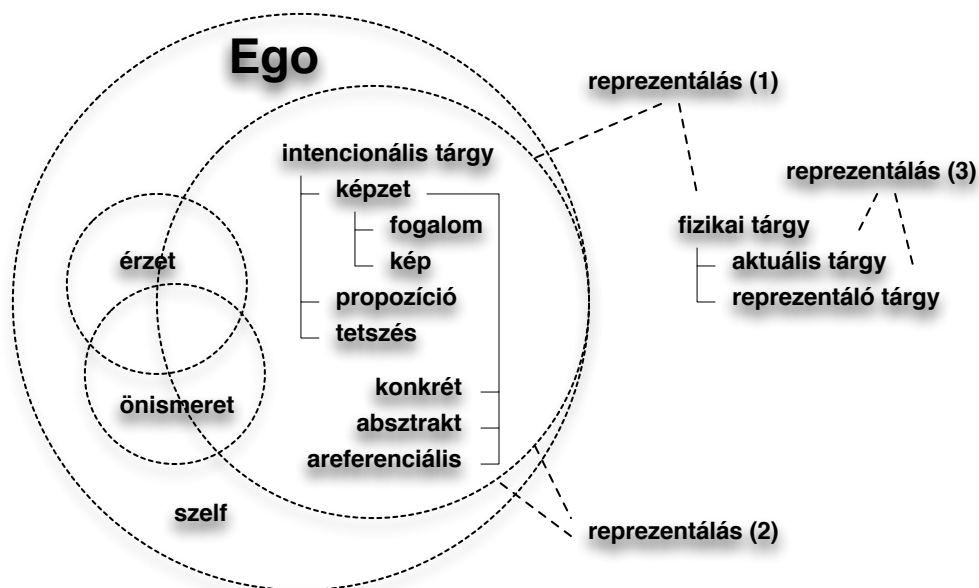


Figure 3.1: Brentano intencionalitás fogalma

Searle

kollektív intencionalitás mi-intencionalitás 121.o. a módszertani individualizmus talaján maradvá 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 interdependes helyzetekbe kerülnek a résztvevők, akkor használható a mi-tudat, kollektív intencionalitás

mindkét helyről hivatkozni ide

A *p* proposicionális tartalom minden esetben ugyanaz

beszédkattípus	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, jóállás, fenyegetés	bocsánatkérés, köszönetnyilvánítás, gratuláció, felköszöntés, résztvényilvánítás	hadüzenet, kinevezés, házastársrá nyilvánítás

értékelés, ítékezés, figyelmeztetés, könyörgés, imádkozás, mentegőzés, panaszkodás, becslés, vélekedés, előrejelzés átok, eskü, áldás vallomás, hazugság/hazudás, titkolódzás/titkolás diplomakihirdetés, fokozatszerzés, kinevezés

Searle [323, 92]

Ruth Millikan: miben más az emberi és nem-emberi reprezentáció? hat különbség van

felszólító reprezentációk, szándék-doboz [263]

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 felszólítás 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.” [262, 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.” [262, 267]

ez nagyon kell a weberi hatszöghöz is

[3]

Altrichter Ferenc az intencionalitás fogalmát úgy értelmezi azon belül a propozicionális attitűdöket

A propozicionális 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, propozicionális attitűdjeink vagyis az intencionális tárgyak rendszere kauzális.

A propozicionális 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 viszonyrendszerrel 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 propozicionális 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.” [3, 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.

zárttság 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.” [3, 140]

A zártság elve megköveteli az önismeret létét. Akkor változtathat valaki az attitűdjén, 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 fogalmilag 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 ekvivalenciával referálni, és hogy képesek legyünk önmagunkat a propozicionális attitűdök szubjektumaként felfogni.” [3, 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. [90] „... 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

3.2 Reprezentáció

tartalom propozicionális tartalom képi tartalom

3.3 Teleológia vs. teleonómia

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 [271] [272]

Mayr 1961: két megközelítés a biológiában: hogyan működik (funkcionális biológia) miért úgy működik (evolúcióbiológia)

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 [248] 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 egyfelől az individuumok (rendszerek) célkövető módon (szándékosan - purposefully) képesek viselkedni, másfelől a genetikai programok evolúciójában vélelmezett célra törekvés között [248, 1504]. Mayr a teleonómia fogalmát így valamely program, információ 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 [249]

„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 [250]

emergencia, Todd Feinberg könyve

Hayne Reese: teleology vs. teleonomy - purposefulness, having an internal determination - purposiveness, serving or effecting a useful function [302]

teleológiai funkció fogalma, Larry Wright [412]

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 korábban f -t csinálta

feladat fogalma, John R. Searle [323]

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 [217]

"mintha" [104]

érvényesség, jog, kényszerítő erő, követés [144] 152. oldal

3.5 Intenció

Az intenció mint szándék az intencionalitásnak csak egy típusa, és vigyázni kell arra, hogy a két terminus alaki közelsége ne okoz 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.” [323, 89]

Searle figyelmeztetése természetesen nemcsak az angol nyelv esetében érvényes, a magyarban (és még nyilván sok más nyelvben) ugyanúgy igaz, hogy az intenció az intencionalitás egyik formája csupán.

!!!!!!!!!!!!!!!!!!!!!!

Ajzen, intencióelmélete

[1] [2]

négyszéki cikk az intencionalitásról, logikáról: [239]

Szécsi Gábor, intencionalitás és okság, [360]

3.5.1 Bratman előtti korszak: vágy-vélekedés modellek

G.E.M. Anscombe, Intention: [12]

Jean Beer Blumenfeld, Goldman's account of intentional action, [46]

jövővezérelt intenció (future-directed intention) és *jelenvezérelt intenció* (present-directed intention)

3.5.2 Bratman: BDI-elmélet

Bratman [57]

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.

[309]

Cohen és Levesque formalizmust dolgozott ki Bratman intencióelméletére [81]

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 plauzibilis következtetés.

3.5.3 más elméletek

Searle, Intentionality, [322]

Douglas Ehring, "Normal" Intentional Action, [116]

Bentzen intencionalitás logikája [31]

Roderick M. Chisholm [75]

Wayne A. Davis [95]

Ullmann, Tamás (2010): A láthatatlan forma. Sematizmus és intencionalitás. Budapest, L'Harmattan

Daniel C. Dennett, Az intencionalitás filozófiája [103] D.C: Dennett: tartalom és tudatosság [102]

Dennet intencionalitás koncepciója [136]

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.

3.6 Vágy

three faces of desire [320] 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 v. hová. A másik a teljesíteni akarok v. mit. 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 [66]

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 inkompatibilitási vélekedések (superfluity)

- preferencia vélekedések (preference beliefs) - érték vélekedések (value beliefs) - sürgősségi vélekedések (urgency beliefs)

- előfeltévesi 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 [86]

incompatibility and incompleteness

state of affairs (A,B,C,D ...)

két típus: G – goal, az Ego célállapotja

B – belief, az Ego hiedelmi állapota¹

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.

$\nabla \triangle \otimes \odot$

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.

¹Előfordul a bizonyossági állapot kifejezés is. mit.bme.hu mesterséges intelligencia könyvre hivatkozni.

inkompatibilitási diszkrepancia:

ΔA is true iff the agent's goal entails A , but her belief state entails $\neg A$

ΔA 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 $\neg A$

∇A is true iff the agent's goal entails A , but neither A nor $\neg A$ is entailed by her belief state. Informally,

∇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 [96]

Kris McDaniel and Ben Bradley, Desires [252]

Eleonore Stump, Sanctification, Hardening of the Heart, and Frankfurt's Concept of Free Will [348]

David Lewis, Desire as Belief, [224] Desire as Belief II, [225]

Timothy Schroeder, Reflection, Reason, and Free Will, [321]

Timothy Schroeder, Three Faces of Desire [320]

Contours of Agency: Essays on Themes from Harry Frankfurt, [64]

John Martin Fischer (ed.), Free Will: Free agency, moral responsibility, and skepticism, Routledge, 2005

Stephen Schiffer, A Paradox of Desire, [318]

Scott Sturgeon, Normative Judgement, [349]

Kenneth A. Taylor, De Re and De Dicto: Against the Conventional Wisdom, [365]

A. L. Herman, A Solution to the Paradox of Desire in Buddhism, [185]

Christine Swanton, Weakness of Will as a Species of Executive Cowardice, [351]

Harry G. Frankfurt, The Importance of What We Care About, [141]

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!"

David Lewis koncepciója: az érték mint másodrendű vágy [345]

Charles R. Pigden, Desiring to Desire: Russell, Moore and David Lewis: [289]

Weber, Cs. Kiss Lajos érték hivatkozása

magyar jogász

Harry Frankfurt, másodlagos vágy, másodlagos akarat

Charles Taylor sources of the self [364] Frankfurt másrendű vágy fogalmára támaszkodik

Reasons from Within. Desires and Values [159]

Why don't you want to be rich? Preference explanation on the basis of causal [165]

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 [105] 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 másdolagod vágy koncepciójára épül gyenge és erős értékelések

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

3.8 Másodrendű preferencia: aggregálás

A preferenciarendezés valamilyen halmaz elemeinek kívánatosság szerinti sorbarendezését jelenti. A preferenciarendezés mint rendezés reláció. Mint minden relációnak, úgy a preferenciarendezésnek is vannak argumentumai, relátumai. Az fontos "technikai" kérdés,

hogy milyen módon lehet a preferenciarendezést végrehajtani több elem (és/vagy több szempont) esetén, erre később még visszatérek. Most csak arra hívnám fel a figyelmet, hogy a preferencia jelenségéről beszélve könnyen átválthatunk, átváltunk reifikált tárgyalásmódba. Amikor a ‘preferencia’ terminust használjuk (például feltesszük a kérdést: “Neked mi a preferenciád?”), akkor könnyen a preferenciarendezés eredményére irányul a figyelmünk, ilyenkor a preferenciarelációt tárgyasítjuk. A preferencia foglalt két

Az angolban ritkán használt, de létező terminus a ‘preferent’

Amikor a “vágy” terminusát használjuk, akkor a preferenciareláció reifikátuma relátum

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.[176] ???példa???

a preferencia fogalmilag feltételezi a diszkrimináció létezését: ha ‘A’ dolgot preferáljuk ‘B’-vel szemben, akkor ezennel ‘B’-t diszkrimináljuk is egyben.

[101]

[15]

[281]

Where the preferences come from? [109]

[108]

alternatívák alternatívák tulajdonságai az ágens motivációs állapota motivációs szem-beszököség, szaliencia (motivationnally salient)

Arrow lehetetlenségi tétele a preferenciákon és azok aggregálásán alapul [16]

✠ * ✠

$\mathcal{L}_{\mathcal{P}}$

Veszünk egy nyelvet.

3.8.1 Archimédész preferencia

archimédész preferencia (nincs végtelen nagy, illetve végtelen kicsi eleme)

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 archimédész preferencia meghatározásakor a preferenciarelációt fogyasztói kosarak-ra (matematikailag vektorokra) alkalmazzák, tehát az összehasonlítást, érték kifejeződést nem egyedi entitásokra, hanem entitások (mégpedig különböző típusú entitások) hal-mazaira vonatkoztatva fejezik ki. Erre a „módosított” preferenciarelációra írják elő az archimédész axióma érvényesülését, amely a következőképpen formalizálható (az archimédész tulajdonsággal rendelkező preferenciarelációt az ‘APREF’, a X és Y termék különböző men-nységeit x_i és y_j szimbólumokkal jelölve):

$$\forall i \forall j \exists k (APREF((x_i, y_i), (x_j, y_j)) \rightarrow APREF((x_j, y_k), (x_i, y_i)))$$

A képlet azt fejezi ki, hogy bárhogyan is állítunk össze két csomagot az X és Y termékek-ből, amelyek közül az egyik csomag (x_i, y_i) többet ér a másiknál (x_j, y_j) , ha az Y jószág

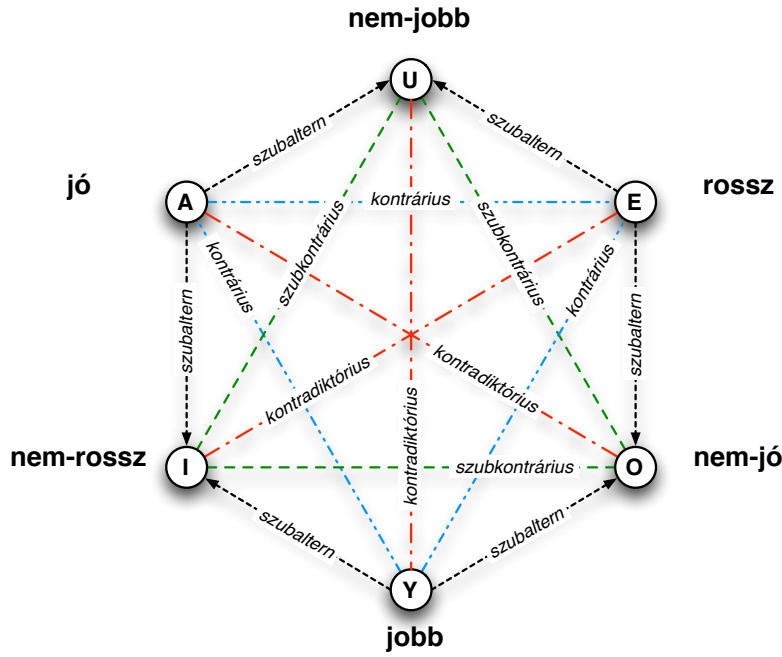


Figure 3.2: értékfogalmak hexagonja

menyiségét 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: $(x_1, x_2, \dots, x_i)$, $(y_1, y_2, \dots, y_j)$, 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):

$$\forall j \exists i (j < i \rightarrow (LPREF(x_j, y_j) \wedge LPREF(y_j, x_j) \wedge LPREF(x_i, y_i)))$$

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 párojaira még lehet döntetlen az állás köztük (ezt fejezi ki az $LPREF(x_j, y_j) \wedge LPREF(y_j, x_j)$ részformula), de az i . elemtől kezdve már az X termék a preferáltabb az Y -nal szemben ($LPREF(x_i, y_i)$). Ezt azt is jelenti egyben, hogy az X termék „összességében” kívánatosabb a másikhöz 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 előre 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 [151], [152], a fogalom elméleti alkalmazásáról, a vele kapcsolatos kísérleti eredményekről lásd: [221], [84], [119].

3.8.2 Lexikografikus preferencia

Natura non facit saltum.

„a természet nem szereti az ugrást”

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, miért korpás lesz a hajam

Formális modellek

a közgazdasági szakirodalomba *N. Georgescu-Roegen* hozta be a fogalmat. *N. Georgescu-Roegen*, *The Pure Theory of Consumer's Behavior* [151] *Choice, Expectations, and Measurability* [152]

A legtöbb formális elméletet *Peter C. Fishburn* állította elő. [130] [131] [132] [129] [133] [220] [219] [56] *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 [11]. *Andrékák* 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.

ez a modell a hierarchia leírása is egyben. A hierarchia fontosságáról, szerepéről én is írtam egy tanulmányt: [357] A hierarchia két fontos típusa: beágyazott és nem-beágyazott

hierarchia az élő rendszerek beágyazott hierarchiával jellemezhetőek [126]

A hierarchia fogalmáról bővebben: [218]

Sven Ove Hansson preferencia változása [175]

prioritizálás = preferenciák aggregálási módszere (Grosz)

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 [198] The Logic of Decision

Self-Control through Second-Order Preferences [273]

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 racionalitásfogalmak megfeleltetése Az ezredvégi nagy borzongás [89]

Lexikografikus döntések egy általános döntési modellben [111]

[110]

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 vétő küszöb

az aggregálási problémát oldják meg

grosz

Preference, priorities and belief [99]

[109] [108]

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

3.8.3 Hasznosság-károosság

hasznosság (utility)

egyik értelmezése szerint függvény, tehát csak az archimédeszi preferenciához rendelhető hozzá

dipólusos fogalom, tehát elméletileg létezik a másik véglete: ez a negatív hasznosság, kellemetlenség, károosság (disutility), például a munkás hasznosságába számít a munkabére, a kellemetlensége pedig a munkavégzés okozta fáradtság, kimerülés

ezzel nem foglalkoznak annyit a szakirodalomban

3.9 Akarat

Schopenhauer az akarat szabadságáról szóló könyvében deklarálta:

„Tehetsz, amit akarsz: de élted minden adott pillanatában csak egy határozott dolgot és szükségképpen csakis ezt az egyet akarhatod, semmi mást. [319, 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 ellentétet, de *akarni* csak az egyiket közülük: még pedig hogy melyiket, azt az öntudatnak is csak a *tett* teszi nyilvánvalóvá.” [319, 19]

elsődrendű akarat, másodrendű akarat

akaratkonfliktus

kényszer

erőszak

Harry Frankfurt az állatok szándékolt cselekvéséről [140]

“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 ... ” [18, 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

3.10 Választás, döntés

a preferenciák konstrukciója: sok helyzetbe úgy megyünk be, hogy nincsenek igazán rögzült előzetesen preferenciánk, hanem maga helyzet alakítja ki azt bennünk [228]

3.10.1 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 [13] a nehéz döntéseknél azt választjuk, amit élénk tesznek (donor felajánlások Európában)

előzetes döntés van a beteg műtétjéről, de “kiderül”, hogy egy vagy két gyógyszer még nem próbált ki az orvos. Mit tesz? Ha egy gyógyszert 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 közepső ajánlat elmaradt, akkor sokkal inkább a print opciót kérték. Amikor nem tudunk dönteni, gyenge a preferenciarendezésünk, akkor az élénk 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 számára helyzetbe kerülünk

előfeszítés, előhangolás jelensége: a vélekedéseink, elfogultságaink befolyásolják az érzéseinket, észleléseinket, döntéseinket, választásainkat prekondíció, előítélet, sztereotípiák hasznos is tud lenni, mert nem lehet mindig, mindent a nulláról indítva megítélni, dönteni

3.10.2 Racionalitásértelmezések

racionalitás és cselekvés

[23]

Robert J. Aumann, Rule-Rationality versus Act-Rationality [20]

P.R. Cohen and H.J. Levesque, Speech acts and rationality, [80]

közgazdaságtanban:

objective function - célfüggvény = preferencia

Harsányi János szerint a racionalitás szabályokhoz való következetes, koherens, konzisztens igazodást jelent [182] [181]

“Rationality is an instrumental concept. In the light of one’s goals (preferences), if the means chosen (actions) are appropriate the individual is rational; if not, irrational. ‘Appropriate’ here refers to method rather than result. Rational behavior is action calculated on the basis of the rules of logic and other norms of validity. Owing to chance, good method may not always lead to good result.” [190]

sok tanulmány van benne a racionalitásról, elköteleződésről [287]

a weberi cselekvéstipizálás alapján saját minősítés (csak praktikus szempontok miatt van rá szükség):

racionalis = cél- és értékracionalis aracionalis = érzelmi és habituális)

3.10.3 döntési szabályok

döntési szempontok

A) ??? módszer (a szempontok egyenrangúak, nem rendezettek):

a) eliminációs eljárások

i) dominancia szabály:

ii) konjunktív szabály (kielégítésre törekvő módszer): minden szemponthoz aspirációs szintet rendelünk, aminek meg kell felelni

iii) diszjunktív szabály : legalább egy szempont szerint kiemelkedő alternatívát választjuk

b) nem tudom, ezek milyen típus alá tartoznak:

i) maximin szabály (pesszimista, biztonságra törekvő, kockázatkerülő módszer, Wald-féle szabály): első körben kiválasztjuk az alternatívák egyes szempontok szerinti leggyengébb értékelését, majd második körben ezek közül a legjobbat választjuk ki a legrosszabbak legjobbjaként Wald kritérium

ii) maximax szabály (optimista módszer): első körben kiválasztjuk az alternatívák egyes szempontok szerinti legerősebb értékelését, majd második körben ezek közül a legjobbat választjuk ki a legjobbak legjobbjaként

iii) minimax szabály: legrosszabb eredmény a legkisebb az elmaradó haszon minimalizálása risk neutral - kockázatsemleges minimum regret - legkisebb megbánás Savage kritérium Savage-Niehans-kritérium

valószínűségi becslést tartalmazó eljárások:

iv) Laplace-Bayes-módszer (realista módszer):

Nem elégséges ok kritérium (Laplace, 1825) Ha nincs "ok", akkor minden esemény bekövetkezésének valószínűsége egyenlő, így a "legjobb" választási alternatíva az, amelyiknek az átlagos nyeresége a legnagyobb Laplace kritérium

E kritérium szerint az lesz az optimális alternatíva, melyhez a környezeti tér különböző lehetséges állapotaihoz tartozó eredmények átlaga maximális értéket ad. $a_i: \max_j \frac{1}{m} \sum_{j=1}^m e_{ij}$ A kritérium az egyes környezeti állapotok bekövetkezését azonos súllyal veszi figyelembe, ezért a semleges (se nem pesszimista, se nem optimista) döntéshozó kritériumának tekinthetjük.

v) Hurwicz-módszer: átmenet a maximin és a minimax elv között, úgynevezett optimizmus koeficiens alkalmaz

Az α pesszimizmus-optimizmus együttható ($0 \leq \alpha \leq 1$) segítségével súlyozza az egyes változatok legnagyobb és legkisebb értékét. Az így számított érték maximumát adó alternatíva adja a kritérium szerint az az adott optimizmus együtthatóhoz tartozó legjobb döntési változatot. $a_i: \max_j (\alpha \max_j e_{ij} + (1-\alpha) \min_j e_{ij})$ $\alpha = 0$ esetén a maximin, $\alpha = 1$ esetén a maximax kritériumot jelenti, míg egyéb α értékeknél a kettő között helyezkedik el az eredmény.

Optimizmuspesszimizmus index (Hurwicz, 1951) A Maximin (pesszimista) és a Maximax (optimista) attitűd túlságosan szélsőséges, az emberek attitűdje ezen két szélsőséges attitűd közé esik. Az optimistapesszimista index minden lehetséges alternatíva esetén súlyozza (egyegyenrejellemző súllyal) a minimális és a maximális nyereségkifizetési értékét és azt az alternatívát választja, amelynél ez a legnagyobb.

B) lexikografikus módszer (a szempontok nem egyenrangúak, rendezettek):

3.10.4 példák

HEMPEL, Carl G. 1965. Aspects of Scientific Explanation. In Aspects of Scientific Explanation and Other Essays in the Philosophy of Science. New York, Free Press. p.466-467



Figure 3.3: döntési szabályok

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.

3.10.5 “döntési irracionalitások”

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 “bejósolására”

3.10.6 Közösségi döntések problémái

[121] [234] [87]

May tétel

univerzalitás

anonimitás

semlegesség

monotonitás

May tétel: az egyéni véleményekből a kollektív döntést előállító F szavazat-összegző függvény akkor és csak akkor tudja kielégíteni az univerzalitási, anonimitási, semlegességi és monotonitási feltételeket, ha többségi szavazás típusú.

Condorcet paradoxon

preferenciakör kialakulása (tranzitív kör)

Arrow tétel

(U) univerzális értelmezési tartomány

(P) gyenge Pareto elv

(L) függetlenség a lényegtelen alternatíváktól

(D) diktátortmentesség

Arrow-tétel: Nem létezik olyan társadalmi jóléti függvény, amely minden preferenciaprofilra helyes rendezést eredményez (tehát egy reflexív, tranzitív és anti-szimmetrikus relációt) és teljesíti az (U), (P), (L) és (D) feltételeket

Diszkurzív paradoxon

doktrinális (bírói) paradoxon

ha bírók (szakértők) logikailag összefüggő állításokról kellj közös döntést hoznia, akkor előfordulhat, hogy – a többségi szavazás elve szerint – a független kijelentésekről

egyenként, egymástól függetlenül képeznek közös döntéseket, illetve a független kijelentésekre vonatkozó egyéni döntésekből logikailag levezethető konklúziókból is közös döntést képeznek, akkor ez a két módon képzett közös döntés ellentmondásba kerül egymással.

List-Pettit tétel

univerzális értelmezési tartomány anonimitás szisztematicitás

2002 Nem létezik olyan véleményösszegző függvény, amely kielégíti az univerzális értelmezési tartomány, anonimitás és szisztematicitás feltételeket, és amely minden lehetséges racionális vélemény-profilra racionális kollektív vélemény-halmazzá összegez.

lehet-e gyöngíteni a feltételeken?

Tétel (Pauly-van Hees 2003). 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 $i \in N$ individuumra.

Véleményaggregálás

áttekintés: List, Synthese [233]

Csáji-Rédei: [87]

List: véleményaggregálás általánosabb, mint a preferenciaaggregálás [107]

3.11 Értelmelem, 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.”

3.11.1 Közös tudás

lásd még a karizmatikus uralomnál a Bourdieu idézetet

ez a fogalom a konvencióhoz is kell

Margaret Gilbert

Szolcsányi Tibor, Davidson háromszögeléséről, [362]

Chwe, rational ritual

Binmore, short introduction to game theory könyvében szó esik a közös tudásról. [44]

Chapter 4

Társadalmi cselekvés

A cselekvéstudomány központi kategóriája a társadalmi cselekvés. Ezt a fogalmat Max Weber is saját elmélete középpontjába helyezte, és Weber meghatározása azóta nem született pontosabb definíció erre a fogalomra [404]. Én is a weberi értelmezést fogom használni a továbbiakban.

A társadalmi cselekvés weberi értelmezését akkor érthetjük meg pontosan, ha követjük azt az logikát, ahogyan Weber összekapcsolja egymással a cselekvéstudományt 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” [404, 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, gondolkodik, 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 cselekvéstudományba, 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 értelmé szerint mások viselkedésére van vonatkoztatva, és menetében is mások viselkedéséhez igazodik” [404, 38]. Sajnos a weberi definícióban a cselekvés “szubjektív értelmé” 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.

4.1 Cselekvés

Bertalan László [36]

Georg Henrik von Wright Norm and Action c. könyvében [390]

K. Segerberg [325]

A Companion to the Philosophy of Action [275]

Michael Thompson, Life and Action. Elementary Structures of Practice and Practical Thought [370]

Daniel Vanderveken, Logic, Thought and Action [385]

Donald Davidson: [94]

Rescher szerkesztésében Davidson, von Wright, Anderson, Simon [303]

Alvin I. Goldman cselekvésemélete [160]

von Wright Norm and Action [390]

Seumas Miller, Social Action. A Teleological Account [261]

4.1.1 Szent Anzelm cselekvéstipológiája

Anzelm cselekvésről szóló fejtegetése

S.L. Uckelman [373]

Douglas N. Walton [402], [400], [401]

Howard L. Dazeley és Wolfgang L. Gombocz [98]

Nuel Belnap [27]

Ricoeur [328]

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, terem, csinál értelemben agere - általános cselekvés értelemben

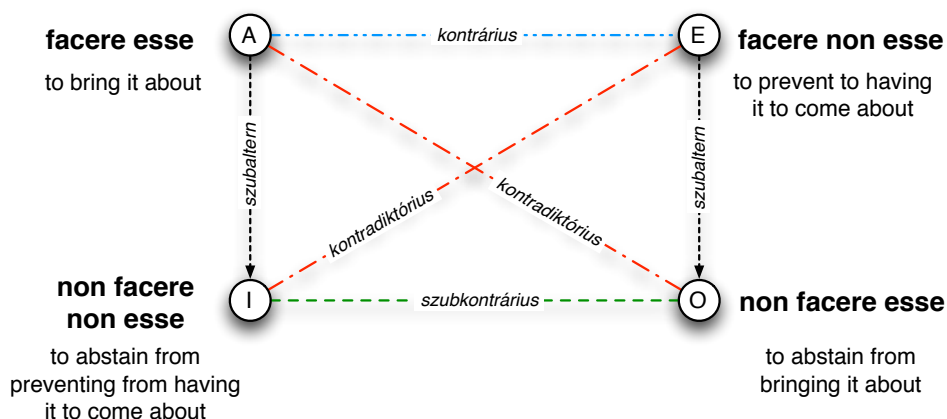


Figure 4.1: cselekvési négyszög

act -agere
 fact - facere
 »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 jól 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 nyomatókat 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” (Boda, i. m. 48. o.). Boda László (1997): A keresztény erkölcs alapkérdései. Szent József Kiadó, Budapest

4.1.2 Tevés/tevékenység és tartózkodás

tevékenység vs. tétlenség
 tétlenül szemlélte, ahogy...

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.” [184, 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.

Danto [93]

Bruce Vermazen négyféle negatív cselekvést különít el [387]

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.” [387] 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

tartózkodik tétlenül szemlél tétlenkedik (??) enged másnak valamit megcsinálni (másodrendű tartózkodás) tűr (enged másnak valamit megcsinálni) hagy valamit megtörténni tagadó forma elmulaszt elfelejt (ebből hiányzik az intencionáltság) kicsit ilyen a tűr visszafogja magát mellőz a kiváltság, a privilégium, az immunitás is jelenthet valamilyen kötelezettség alóli mentességet, ami a tartózkodáshoz is kapcsolható

4.1.3 Magasabbrendű cselekvés

szankcionálás vs. büntetés

4.2 Cselekvő

behozni a pszichológia ‘self’ fogalmát

ágentivitás

Frankfurt: concept of a person [139] The Problem of Action [140]

Gary Watson Agency [403]

Alfred R. Mele, Motivation and Agency [256]

Michael E. Bratman, Structures of Agency [58]

E.J. Lowe, Personal Agency. The Metaphysics of Mind and Action [238]

Christine M. Korsgaard: 2008 [210]

Taking Rights Seriously. Getting it Right: 2006 [142]

4.2.1 stit-elmélet

cselekvés, döntés

He could have chosen to do otherwise, Chisholm

Belnap és Perloff cikke első megjelenés: [28], javított változat: [29]

történeti elemzés: Michael Perloff [285] és Krister Segerberg [325]

két összegző mű:

Nuel Belnap, Michael Perloff, Ming Xu, Facing the Future [30]

John F. Horty Agency and Deontic Logic [192]

4.2.2 program-elmélet

4.2.3 Davidson

Action sentence inkább event sentence viszont nagyon jó javaslát de nem mondja meg, mitől cselekvés a cselekvés

a végén csak mond valamit Davidson: a cselekvésben intencionalitás van [94] 94.o.

'it was intentional of x that p' x a person, p action sentence (uo. 95.o.)

4.2.4 Ego és Alter

A társadalom minimális modelljének két szereplője.

A különböző elméletekben más neveket kaphatnak (donor, recipiens, actor stb.)

Chapter 5

A társadalmi cselekvés típusai

Beállítódás, irányultság, motiváció a tipizálási szempont

5.1 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 is 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 utility 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 játékokra utal, hiszen vagy A van B felett vagy fordítva. (lásd: szervezet és/vagy státus fejezet)

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 5.1: mérési skálák tulajdonságai

5.2 Szabályszerűségek

Ösztön Szokás Habitus Reflex rutin Szokáserkölc szokásnorma Divat Tradíció Erkölcs Morál Illem Etikett Konvenció Minta Séma Alakzat Mintázat mintaszerűség Szabály Szabályszerűség Szabályzat Szabályosság Törvény Jog Jogszabály Törvényszerűség Jogszzerűség rend

Fizikai szint Biológiai szint Társadalmi szint

5.3 Ego és a világ

Motiváció, beállítódás, attitűd

Franz Dietrich és Christian List A reason-based theory of rational choice [108]

propozicionális attitűd olyan mentális állapot, amelyben egy személy valamilyen propozícióval (tehát olyan tartalommal, mely igazságértékkel rendelkezhet) áll relációban.

propozicionális attitűd: vélekedés és vágy a vélekedés doxasztikus attitűd a vágy

konatív attitűd lásd: MEO-szöveg

belief, attitude, proattitude (Davidson)

értékracionális célracionális érzelmi habituális

motiváció Mérő, MacLelan

5.3.1 Célracionális

commitment a kulcs a döntésben (célracionális mérlegelésben) ott van a preferencia, a preferenciában pedig ott van az érzelem lásd: [342]

5.3.2 Értékracionalitás

Zajonc az érzelem, az érték és a preferencia fogalmait összemosza

[417] Zajonc, Feeling and thinking. Preferences need no inferences

[162] The Neural Bases of Cognitive Conflict and Control in Moral Judgment

Manipulations of Emotional Context Shape Moral Judgment [378]

The Roots of Morality [260]

5.3.3 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 it 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 eufórárt 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 szóbjághető 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 bőrének geometriai mintázatáról elmélkedni. Ha a nyúl menekül, ezt jóval az emlékeztetjegyek 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.” ([417] 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

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 predispozíció személyiség szentiment szenvedelem szenvedély (passion) szimpátia temperamentum tendencia

Damasio-kritika

Damasio [91] 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 [188]

50.o. idézet fontos

érdek az ésszerű önzés

Orthmayr, szenvedélymentesség [279] Steiger Kornél válasza, [347]

Susan James, Passion and Action [196]

szenvedély, érzelem, vágy, indulat vs. akarat, értelem, megértés, cselekvés

Susan James, szenvedélyek vágyak cselekvés, [195]

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, adaptívnek bizonyult megoldások, tevékenységek "felidézésével, újrahasznosításával" [117, 46].

TenHouten tisztázza az egymáshoz kapcsolódó "érzelemfogalmak" jelentését [369]: "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 személyiségjegyek, hosszabb ideig (hónapokig, évekig, akár egy egész életen át) áll fent

TenHouten egy táblázatba rendezte [369, 4] a legfontosabb szerzők alapérzelem-listáit¹:

szerző	Proposed pri- mary emo- tions								
Descartes (1647)			joy	sadness					wonder, love, ha- tred, 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	disappoin- tment- sadness					
kínai tradíció	fear	anger	happiness	sadness					
Darwin	fear	anger	happiness	sadness		surprise	disgust		

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 [288]

¹TenHouten táblázatát kiegészítettem két további, tőle származó sorral, a kínai tradíció és Darwin nézeteivel.

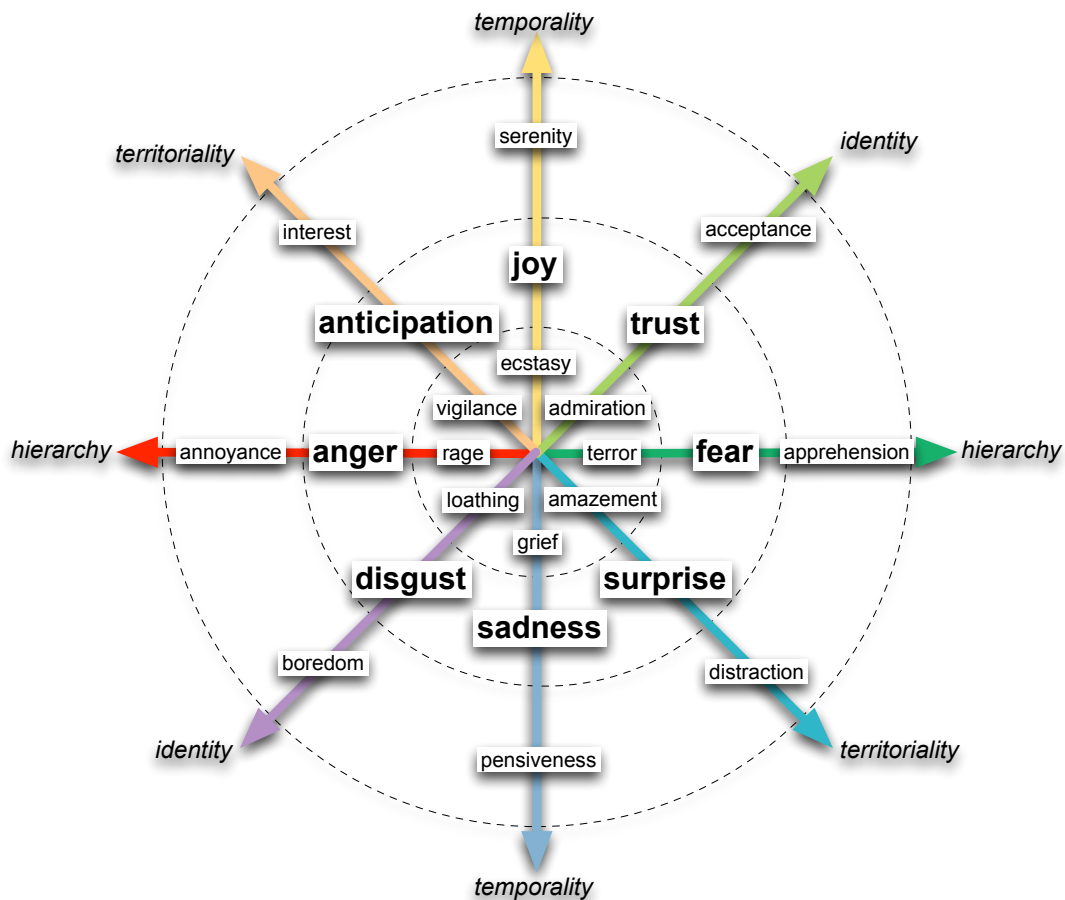


Figure 5.1: Plutchik alapérzelmei

5.3.4 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 észjárás 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 oldálnak ugyanaz a kettőssége, amelyet már Arisztotelésznél is láthatunk, így ez a tudás is egyszer [361]

Polányi Mihály Szívós Mihály [361]

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 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

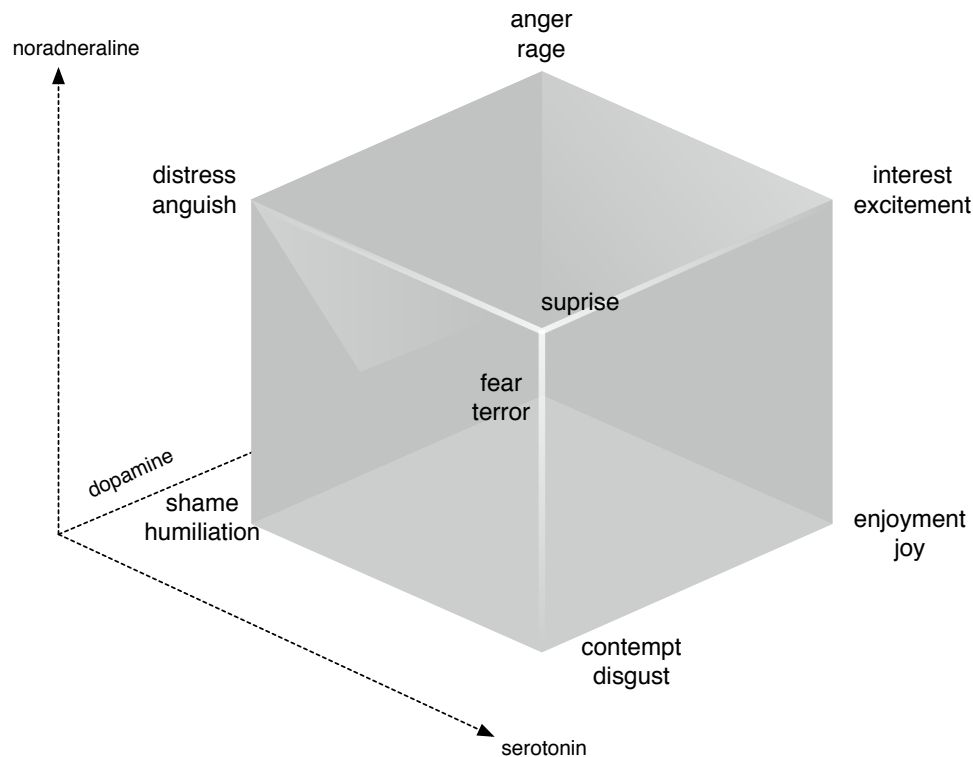


Figure 5.2: L  wheim   rzelemtipol  gi  ja

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: [246]

Pokol B  la, Bourdieu elm  let  nek alapkateg  ri  i, a habitus [292]

“tudattalan  tott rutin”

eg  y  nre zuhan   d  nt  si teher

habituson bel  l  i diszpoz  ci  k

Habitus, Freedom, and Reflexivity [187]

Bourdieu and Wacquant [51]

Lo  c J.D. Wacquant [399]

Pierre Bourdieu habitus  rtelmez  se

“A habituson olyan tart  s   s   tvihet   hajlamok rendszer  t   rtj  k, amely, valamennyi m  lt-beli tapasztalatot be  p  tve, minden pillanatban mint *  szlel  si,   rt  kel  si   s cselekv  si*

minta működik.”[50, 213]

“... a habitus nem más, mint emberi természetté vált történelem, 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.”[50, 215]

Polányi Mihály személyes tudás fogalma

akarat és döntés és neuronok: <http://www.informationphilosopher.com/solutions/scientists/eccles/>

Duhigg

[114]

Eugene Pauli, szokáshurok (habit loop)

Larry Squire,

Claude Hopkins, szokásteremtés "Tudományos reklámozás"

cue-routine-reward + craving

Control and automaticity in social life: [405]

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 gondolkodásunk során is rengetegszer rögzített gondolatmeneteken, hivatkozásokon keresztül haladunk.

Még pontosabban fejezte ki ezt Neumann János, amikor azt mondta, „... a matematikában az ember nem megéri a dolgokat, hanem megszokja.”

5.3.5 Miért hamis a mindig racionális ember mítosza?

Áttekintő ábrán a különböző motivációk, beállítódások egymáshoz való viszonyát a logika hatszög segítségével

“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.)” [190]

Mérő könyve, Hirshleifer cikke a racionalitás, konzisztencia, a logika bukfeceiről
döntési paradoxonok a döntéselmélet területéről Tversky és Kahneman elmélete a korlátozott racionalitásról, kilátásméletről

Herbert Simon korlátozott racionalitás [335]

racionalizálás, kognitív disszonancia “When a person is made aware of a disharmony between his actions and his preferences and beliefs, the economist would expect him to

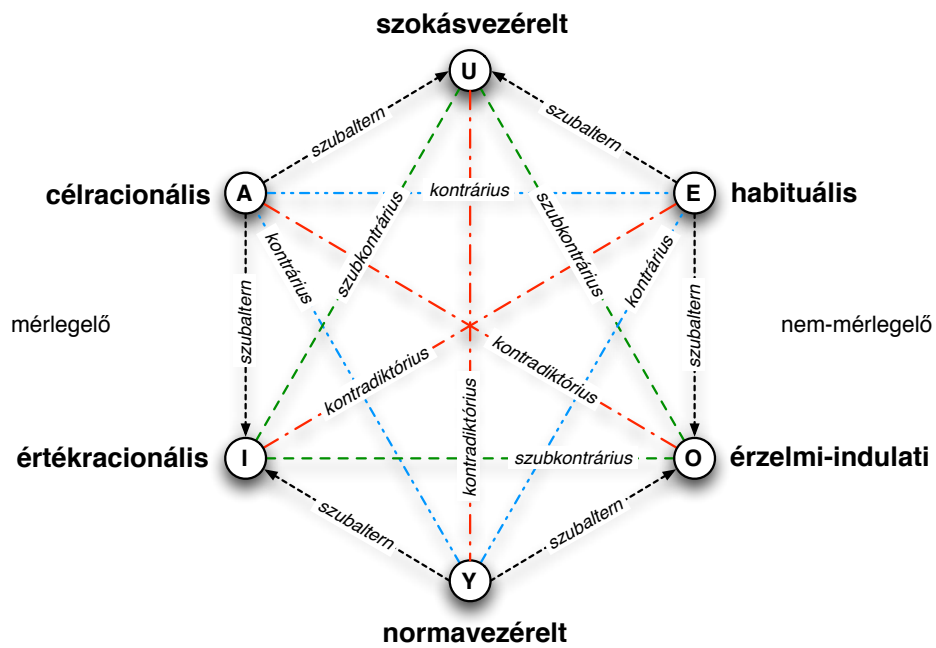


Figure 5.3: cselekvéstípusok beállítódás szerint, Max Weber

revise his choice of action - but the cognitive dissonance theorist predicts that he is likely instead to modify his preferences or beliefs.” [190]

mert habituális vagy érzelmi módon cselekszenek
racionális önbecsapás [366]

Chapter 6

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.” [62]

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 (77.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 rejlik 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.” [404, 54]

A cselekvés és a társadalmi cselekvés fogalmába már bele kellett építeni az ágens, 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.

6.1 Ágensre irányultság: szociális preferencia

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 másokra, 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 másikkal (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 helyzetek-

ben 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.

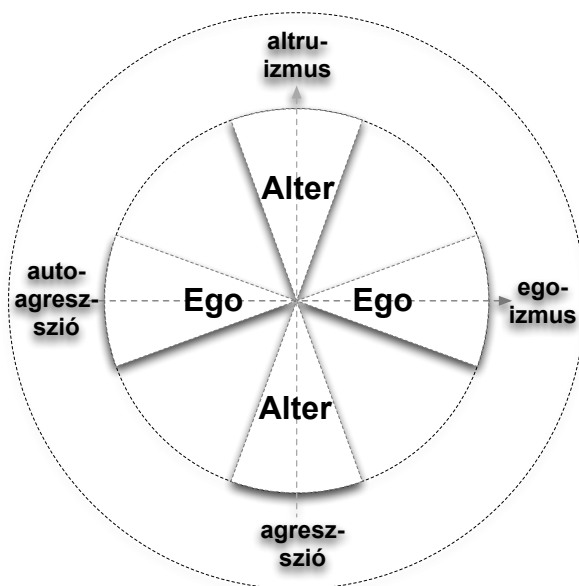


Figure 6.1: magára irányuló, illetve másikra irányuló beállítódások

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 *proself* (pro-self), a másikra irányulást pedig *prosocialis* (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.

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.

6.2 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 jutalmaznak 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.

²A későbbiekben tárgyalandó játékelmélet számára sokkal természetesebb, használhatóbb ez a megközelítés.

...

Hogy lehet a szociális preferencia fogalmát definiálni? Nyilván a PREFERENCIA fogalmából kell kiindulnunk.

Agent szelf Ego Alter

ASD

6.3 Társas értékorientáció

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 [257]. 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 *kifizetések* 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 beállítódást 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			X_{Alter}	Y_{Alter}
Ego kifizetése	c_1	c_2	$\rightarrow$	X_{Ego}	$(c_1 + d_1), (c_1 + d_1)$	$(c_1 + d_2), (c_2 + d_1)$
Alter kifizetése	d_1	d_2		Y_{Ego}	$(c_2 + d_1), (c_1 + d_2)$	$(c_2 + d_2), (c_2 + d_2)$

Az izomorfiaira konkrét példaként a fogolydilemma játék esetét mutatom be. A következő bal oldali SVO-táblázatba beírt négy érték – fenti “kulcs” alapján – pontosan megfeleltethető a jobb oldali táblázatban látható preferenciamátrixnak, ami pedig a jól ismert fogolydilemmához tartozik.

	X	Y			X_{Alter}	Y_{Alter}
Ego kifizetése	1	2	$\rightarrow$	X_{Ego}	3,3	1,4
Alter kifizetése	2	0		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 [251], 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	nyereménykülönbség 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 egybeesnek a játékelmélet világában. Érdekesnek 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 cselekvéseméleti modell legfontosabb ágensszerepéhez, az ‘Ego’-hoz sokkal jobban illeszkedik hangulatában az egoizmus terminusa akkor, amikor a magára irányuló beállítódást akarom kifejezni.³

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 [163], 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 [230]. Az új geometriai keretrendszer, a McClintock-féle motivációs gyűrűt (Ring measure) már

³Még egy indokom a terminuscserére mellett az, hogy az individualizmus fogalma az evolúcióbiológiában széles körben használt, ami szemben áll a társadalomtudományok által preferált egoizmus terminusával. Ebben a tudományszakok közti terminológiai versengésben – ahol csak lehet – inkább saját szakterületem gyakorlatát követem.

lehetővé tette, hogy ne kategoriális, hanem már mértéken alapuló nominális változóként lehessen a vizsgált beállítódásokat kezelni.

DEdededdededdedee

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* (sdomasochism) 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 minimalizálása (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 nyereség minimuma	min(Ego)	(10,50)
destrukció (szadomazochizmus)	együttes nyereség minimuma	min(Ego+Alter)	(20,30)
önfeláldozás (mártíromság)	nyereségkülönbség minimuma	min(Ego-Alter)	(30,80)
agresszió (szadizmus)	másik nyereségének 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 melléklet rendelt, rájuk jellemző kifizetésekkel együtt. A könnyebb összehasonlítás végett a kifizetéseket teszem a táblázat első oszlopába:

(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)

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 másik három negatív beállítódás mérését sokan fölöslegesnek, megalapozatlannak tartották (és tartják). Ezzel szemben a társas értékorientáció elmélet

⁴A kármaximalizálás vs. haszonminimalizálás fogalmai között ekvivalenciát teremthetünk ilyenfajta megfeleltetéseken keresztül:

$$\max(\text{Ego}-\text{Alter}) = \max(\text{Ego}) \ \& \ \min(\text{Alter})$$

$$\max(\text{Alter}-\text{Ego}) = \min(\text{Ego}-\text{Alter}) = \min(\text{Ego}) \ \& \ \min(-\text{Alter}) = \min(\text{Ego}) \ \& \ \max(\text{Alter})$$

egyik fontos 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 törekvő *egalitarianizmust* (egalitarianism) mint a proszociális beállítódás másik megnyilvánulási formáját [382], azonban egy, a munkatársaival írott, 2007-es cikkében a kooperatív és egalitárius irányultságot már összevonja [383]. 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 [384]). A kísérletek tanúsága 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 másoknak. 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 játszhatnak a játékaikat, de proszociálisak (altruisták vagy kooperatívok) 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 leírassuk, értelmezhesük a játékosok (és a kívülálló modellezők, megfigyelők) különböző beállítódásait ugyanazon játékszituációra 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.” [109]

A nagy kérdés itt az, hogy miként akarjuk modellezni a társas preferenciákat. Ha azt

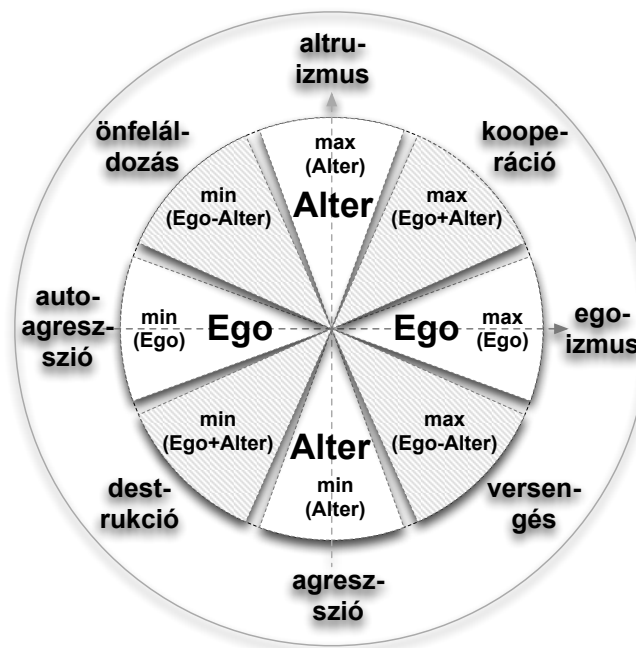


Figure 6.2: a társadalmi értékorientáció típusai

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 másokra 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 más és más kell majd a proszociális irányultságokat leíró formulákba, képletekbe beírni.

A válaszokat nem most próbálom meg kifejtetni, 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 [346]. 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 propozicionális attitűd. Van propozicionális tartalmuk, és viszonyt fejeznek ki az egyén és a propozíció között.

_____ valahova betenni: ⁵

a vágy mindig propozícióra irányul, nem objektumra, példa erre [346, 350].

⁵ Az egoista beállítódás magára irányul, azzal nincs elméleti kiterjesztési probléma. A másokra 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 cselekvésmélet számára ez az értelmezés elegendő. Elméletileg lehetne tágítani ezen az értelelmzésen, 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.

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 és általános beállítódást mint személyiségjegyet határoz meg.

A társas értékorientációk értelmezése

harmónia, mutualizmus (kölsőnyös 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 érték-tartományában helyezkedik 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

“tisztá” egoizmusához 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 = segítség a másiknak (Egonak költség/hátrány, Alternek előny) dezertálás = nem segítség 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.” [270]

önző = saját érdekét elégíti ki, altruista = másik érdekét elégíti ki önérdékkövető vs. másérdékkövető

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) [270]

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 [?] a diakróm időben

6.3.1 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 önérdékkövető. 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 önérdéküket követik. A gazdasági ember (homo oeconomicus) haszonmaximalizáló és kárminimalizáló lény „Economic man is characterized by *self-interested goals and rational choice of means*.” [190]

az egoizmus lehetségesége (The Possibility of Selfishness) [277] 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, inividuális 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.” [120, 61]

harc, szembenállás, konfliktus, érdekellentét lásd: [156] bevezetőjének 2. lábjegyzetében (33.o.)

Hamilton négyes felosztása

mutuality selfish spite altruism

az érdek ésszerű önzés

Albert O. Hirschmann: érdekek és szenvedélyek esztétikájának elemzése [188]
 a szenvedélyeket az érdekek által lehet megzabolázni, az érdek kiszámítható:
 Jhering egoizmus és jog viszonya kapcsán: [146]
 Frank W. Marlowe és szerzőtársai [243]

6.3.2 Altruizmus

fals dichotomy? [213]

kár másik modell a (Gulyás-disszertációból)

Gary Charness és Matthew Rabin [67]

Cox és társai [85]

mi befolyásolja a másakra irányuló döntést: a státus és a reciprocitás

az altruizmusmal kapcsolatos vizsgálatok, kísérletek áttekintése: [212]

Az emberiség különböző (tudományos, vallási, művészeti, erkölcsi) diskurzusaiban mindig kiemelt helyet kapott az önzetlenség, az altruizmus kérdése.

Garret Hardin: diszkriminatív altruizmus típusai [178]

egyre növekvő számosság, egyre csökkenő mértékű altruizmus

– individuum (szerelem, barátság)

– család

– bajtársiság, kronizmus

– tribalizmus, törzsiség

– hazafiság, patriotizmus

áttekintés az altruizmusról: [259] [290]

„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” [372]

többszemélyes altruizmus (multiparty altruism) az altruista rokonával szembeni viszonzás

jóvátételi altruizmus [372]

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, altruism 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.” [265]

Hamilton [168], [169]

6

Gulyás disszertációjából átvenni irodalmat méltányosság

M. Dufwenberg és G. Kirchsteiger 2004 [113]

Az altruizmus fogalmának elemzése: Khalil [208]

1) egoista altruizmus az altruista cselekvést a várható jövőbeni haszon magyarázza (például a direkt reciprocitás esetében) párhuzamba állítható az evolúcióbíológia triversi reciprocity altruizmus fogalmával [372]

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 [125]

illetve

G.E. Bolton és A. Ockenfels modellje theory of equity, reciprocity and competition [47]

a méltányosság szándékcentrikus modellje:

M. Rabin [296]

illetve

Levine altruizmus-rosszindulat modellje [223]

az altruizmus megjelenése a különböző kultúrákban Gintis: [154]

Ulmann-Margalit, Considerateness [376] Family Fairness, [375]

Herbert Simon evolúciós magyarázata az altruizmus kialakulására [334]

kiváló áttekintést ad magyar nyelven Janky Béla [197]

az altruizmus család, vásott könyök modell Gary Becker altruizmusmodellje

Altruism, Egoism, and Genetic Fitness: Economics and Sociobiology [25]

későbbi Becker cikk Altruism in the Family and Selfishness in the Market Place [26]

a vásott könyök modell továbbfejlesztése [34]

a csoportselektációs magyarázatok problémáiról: [314]

barátság kategóriája [145]

etikai egoizmus lehetősége: [147]

H.Simon [333]

Marosán cikk [245], a nők gyűjtögető életmódra berendezkedése miatti munkamegosztás [216]

⁶Az önzetlenség kialakulásáról magyar nyelven ad kiváló áttekintést Bereczkei Tamás [32], a könyvről én is írtam egy recenziót [354].

Az altruizmus modellje

másodrendű fogalom, mert a hasznosságérzetbe be kell venni a másik hasznosságérzetét de ha kettős ontológiát használunk (lásd még méltányosság-lemélet), 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, x -re és y -ra), így nem vette figyelembe a rendezést végrehajtó személyt ($PREF(x,y)$). 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 perszónifikálni kell a fogalmat. Első lépésben fel kell venni a fogalom összetevői közé a preferenciarendezést végrehajtó személyt, valamelyiket az ágensek közül, Egot vagy Altert.

$PREF(Ego, x, y)$ – *Ego preferálja x -t y -nal 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

$x = BENEFIT(Ego) \ \& \ BENEFIT(Alter)$

$PREF(Ego, BENEFIT(Agent), BENEFIT(AlterAgent))$ – *Ego preferálja x -t y -nal szemben*

David K. Levine empirikus kísérletekben vizsgálta az altruizmus jelenlétét Modeling Altruism and Spitefulness in Experiments [223]

R. Sethi és E. Somanathan a preferenciák változásáról (fejlődéséről) és a reciprocitás kapcsolatáról [332] a preferenciák evolúciója: [313]

Becker Rotten kid theorem [197]

[232]

6.3.3 Kooperáció

Competition among cooperators: Altruism and reciprocity [92]

6.3.4 Versengés

6.3.5 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 [340]. Pedig más (például vallási, erkölcsi, művészeti) diskurzusokban mindig is jelen volt a rosszindulat, az agresszió, a gonoszság kérdésköre (akárcsak az altruizmusé). 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 fókusz, amikor vizsgálándónak tartotta a rosszindulat (spite) jelenségét az állatvilágban [170]. 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 figyelték 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 [199]

[298]

[253]

„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.” [63]

a fogorvos fájdalmat okoz a gyökérkezelés során, de nem ez a szándéka, ezért nem agresszió a szodomazo kapcsolatban sincs agresszió, mert az "áldozat" nem akarja elkerülni a fájdalmat

agresszió, autoagresszió (önagresszió) S.A. West és A. Gardner [407]

The 'spiteful' origins of human cooperation [243]

animal conflict limited war? [247]

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. [340]

?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 jelentkezhethet kín, fájdalomkózkodás, gyakran sadizmus 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 konstrukció 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, sadizmus személyiségjegyek is (pszichopátia)

FESHACH (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 than the harm caused to the non-kin individual targeted. [307]

spiteful behaviour, which is costly for the both actor and recipient [201]

“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 [223]

gonoszság, a másik kárára irányuló motiváció

jóindulat vagy rosszindulat (benevolence vs. malevolence) [190]

nem túl gyakori, de azért létezik kétféle értelmezés van a fogalomra

Spite: Hamilton's unproven theory [137]

Wilson könyvében ír a spite és az agresszió jelenségéről [410]

Hamiltonian spite (pozitív) Wilsonian spite (negatív)

[149]

6.3.6 Autoagresszió

önmaga ellen irányuló rosszindulat

mazochizmus

vallási áldozat, áldozat, emberáldozat, Dugonocs Titusz,

Az autoagresszióra hozható legegyszerűbb példa az öngyilkosság jelensége. És nem igaz, hogy az öngyilkosság kizárólag emberi sajátosság volna. Bár nem túl gyakori, de előfordul az állatvilágban belül [202, 113]. A kutatók adaptív öngyilkoságnak nevezik, és a létezését úgy magyarázzák, hogy az egyed a saját génjének sikeres továbbörökítése érdekében pusztítja el magát. Előfordul, hogy a borsótetű (*Acyrtosiphon pisum*) testébe darazsak beszúrnak a tojásaikat, hogy az utódaik parazitaként nevelkedjenek fel a borsótetű

testéből táplálkozva. Ilyen esetekben előfordul, 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 darázslárva 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 [202, 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.

6.3.7 Önfeláldozás

Ez az igazi altruizmus, amikor a cselekvő magának kárt, költséget okoz, a másinak előnyöket

Nem érzem feladatomban, 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 [115].

6.3.8 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. [4]

6.4 Dinamikus kapcsolat: reciprocitás

a reciprocitás altruista/kooperatív és egoista/versengő/destruktív beállítódások időben változó egymásutánisága

pozitív és negatív reciprocitás

6.5 Méltányosságelméletek

A méltányosságelméletekről magyar nyelvű áttekintést ad Gulyás Attila cikke [166].

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 [166]

kettős ontológiát kell használni [109]

erős reciprocitás

Journal of Personality and Social Psychology [78]

Manipulating Allocation Justice: How Framing Effects can Increase the Prevalence of the Talmudic Division Principle "Shnaim Ohazin" [37]

Janky Béla a méltányossági normák két nagy csoportját különíti el [197]. 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 [350]

M. Rabin méltányosságmodellje a szándékra alapoz [296]

ide csatolni Fiske elméletét

igazságosság:

- formális
- szubsztantív
- retributív
- korrektív
- kommutatív
- disztributív

[17] 456.

6.6 Szolidaritás

Lindenberg három cikke

Janky könyve

6.7 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." [190]

Becker tipizálásában szerepel az irigység, gyűlölet, de a cikk húzós [24]

6.8 A társadalmi kapcsolat típusai

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 bemutattem. 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* 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.” [404, 64],

majd kicsivel később pedig megadja a *közösség* és a *társulás* meghatározását:

„*Közösségnek* nevezzük a társadalmi kapcsolatot, ha a társadalmi cselekvést ... a résztvevők szubjektíve átértzett (érzelmi-indulati vagy tradicionális) összetartozásán alapuló beállítódás jellemzi.”

„... *társulásnak* nevezzük a társadalmi kapcsolatot, ha a társadalmi cselekvést racionálisan (érték- vagy célracionálisan) motivált érdekkiegyenlítődesen vagy ugyanígy motivált érdekkapcsolódáson alapuló beállítottság jellemzi.” [404, 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” önértékvényesítés, a *közösséget* és a *társulást* pedig az összetartozás, az érdekkiegyenlítődes, 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 másokra 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

⁷Ezzel a kiegészítéssel korábban másutt már megpróbálkoztam [355], 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

megoldása. Ha a magára irányuló beállítódás főfogalma a harc, akkor ennek párja (ellentéte) az *együttműködés* (kooperáció) lehet a másikra irányuló beállítódás oldalán.

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.

??????? 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 nem-mérlegelő 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 [207] viszály (dischord) harmónia (harmony) kooperáció (cooperation)

Aumann, War and Peace [19]

Rapport a konfliktus három típusa [299]

1) harc, háború, verekedés

2) játék, játszma, önérdékkövetés

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.”

6.9 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 7

Harcon alapuló társadalmi kapcsolat

harc, szembenállás, konfliktus, érdekellentét lásd: [156] bevezetőjének 2. lábjegyzetében (33.o.)

7.1 Érdek

A harc logikailag vagy biológiailag (?) megelőzi a kooperációt ezért a hatalmat előbb kell tárgyalni

- a céltételezés, preferencia függvénye

- nem egyenlő az önzéssel, inkább a beállítódástól függ

7.2 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 '[190]:

- konfliktuskezelés, a konfliktusok okai preferenciák, lehetőségek, percepciók "potenciális megegyezési terület"

- a háború művészete, hatalommérnökség, a harci konfliktus feloldásának, megnyerésének technikái

- a csatateren a nyertes mindent visz (increasing-return), de ez nem terjeszthető ki a legyőzött ország teljes területére (decreasing-returns). 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

7.2.1 Hatalom az állatok közt

biológiai minták, példák

- alfahím, dominancia, csípkedési sorrend

- erőszak mint hatalomforrás

- lehet tudás?

megküzdene a szűkös erőforrásokért, élelemért, területért, szállásért, szexuális partnerért.

7.2.2 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 [404, 77]:

„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.”

A hatalom fogalmának explikációja

1) *Ego és Alter között akaratkonfliktus létezik.*

$\text{exist}(e_2) \wedge \text{want}'(e_2, \text{Ego}, e_1) \wedge \text{fly}'(e_1, \text{Ego}) \wedge \neg \text{exist}(e_1) \wedge \neg \text{want}'(e_3, \text{Alter}, e_1) \wedge \text{exist}(e_3)$
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 $\text{EXIST}(e_i)$ proposíció alapértelmezés szerinti igazságértéke 'hamis', amit nem kell a képletekben megjelenítenünk.

2) *Ego akarata megvalósul.*

$\text{past}(e_2) \wedge \text{want}'(e_2, \text{Ego}, e_1) \wedge \text{fly}'(e_1, \text{Ego}) \wedge \text{exist}(e_1) \wedge \neg \text{want}'(e_3, \text{Alter}, e_1) \wedge \text{past}(e_3)$

$\text{PAST}(e_2) \wedge \text{WANT}'(e_2, \text{Ego}, e_1) \wedge \text{FLY}'(e_1, \text{Ego}) \wedge \text{EXIST}(e_1) \wedge \neg \text{WANT}'(e_3, \text{Alter}, e_1) \wedge \text{PAST}(e_3)$

7.3 A kooperáció biológiai mechanizmusai

Cooperation within and among species [312]

Eusociality: origin and consequences [408]

[406]

[407]

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

7.3.1 Rokoni altruizmus

Hamilton rokon 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

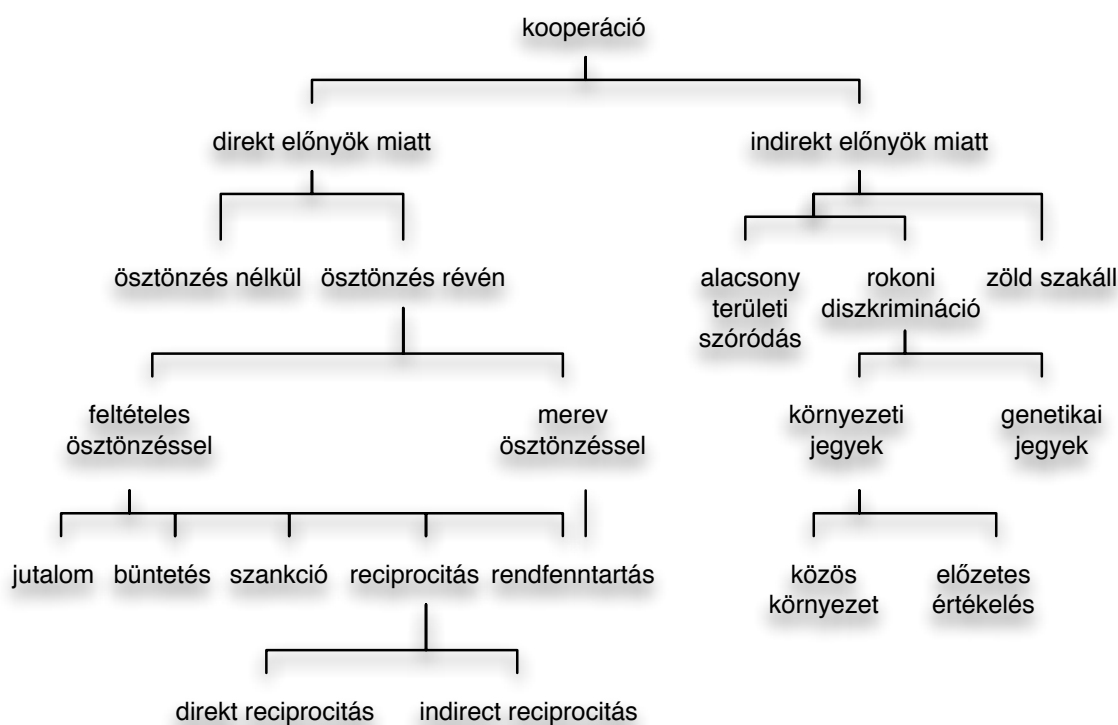


Figure 7.1: a kooperáció típusai – West és tsai ábrája

kritika: Wilson and Wilson: Rethinking the theoretical foundation of sociobiology

„Selfishness beats altruism within single groups. Altruistic groups beat selfish groups.”

[409]

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 [79]

a rokoni altruizmus (kin selection) az inkluzív fitness hamiltoni magyarázósémája helyett lehetséges eredeti evolúciós magyarázatot adni, főleg az inkluzív fitness elmélet The Evolution of Eusociality [274]

7.3.2 Zöld szakáll hatás

zöld szakáll (green beard) [97, 115]

7.3.3 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 átlaghoz 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 (kisebbségre lesz a versenyképessége a recipiensnek). Ezt a jövőbeni versenyképesség-növekedést egyébként ki lehet terjeszteni akár a donorra is.

7.3.4 Csoportszelekció

greenbeard [97] [407]
 skyrms cikke
 Wilsonék elmélete

7.3.5 A costly signaling és a handicap principle

Bereczkei Tamás magyar áttekintése [32]
 The Handicap Principle. A Missing Piece of Darwin's Puzzle [416]
 Zahavi, A. (1975). Mate selection: Selection for a handicap [413]
 Zahavi, A. (2003). Indirect selection and individual selection in sociobiology: My personal views on theories of social behaviour. [414]
 Costly signaling and cooperation H Gintis, EA Smith, S Bowles - Journal of theoretical biology, 2001 [157]
 Zahavi, A. (1995). "Altruism as a handicap – The limitations of kin selection and reciprocity". Avian Biol 26 (1): 1-3. [415]
 Altruism as a Signal: Zahavi's Alternative to Kin Selection and Reciprocity Jonathan Wright [411]
 High-quality male field crickets invest heavily in sexual display but die young [194]

Chapter 8

Békés 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

8.1 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 (ubbi persze sak módjával) 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 indul ki. A legegyszerűbb itt egyszerre sok feltétel teljesülését jelenti. Egyrészt 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 preferenciámatríx sajátosságai alapján végezték az elemzésüket, de nem foglalkoztak komolyabban játékosztályok kialakításával, a játékok tipizálásával, illetve nem adtak neveket sem a játékoknak (sőt, nem is mutatták be őket – például történetek segítségével). A játékok számosságára vonatkozóan Rapoport és Guyer 1966-os cikke alapján közel negyven éven

¹ A játékosok megragadására sokféle jelölési konvenció létezik, amelyek közül olykor használok a 'Sor' és 'Oszlop' játékos megnevezést.

keresztül elfogadott volt az a nézet, hogy 78-féle játékot lehet elkülöníteni egymástól. Ez a szám úgy jött ki, hogy a 2×2 -es preferenciamátrix 2×4 értékét 576 ($4! \times 4!$) féle módon lehet elrendezni, amelyek közül viszont sokat ekvivalensnek lehet tekinteni egymással. Rapoporték úgy gondolták, hogy a mátrix sorait és oszlopot 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 [55].

A közel negyven éves konszenzust rúgta föl 2005-ben David Robinson és David Goforth könyve a 2×2 -es játékok topológiájáról [304]. A két szerző 78 helyett 144 játékra kiterjesztve mutatta be azt, hogy milyen topológiai leírás adható a 2×2 -es játékokra, milyen játékosztályokat lehet elkülöníteni közöttük.

A játékokat több szempont szerint jellemezhetjük. Figyelhetünk arra, hogy:

- hányan játszik 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 nem-kooperatív, “néma” vagy kooperatív, “hangos” a játék²)
- egyszerre döntenek-e a játékosok vagy egymás után, egyik a másik döntésének ismeretében (lehet szinkron, szimultán, statikus vagy aszinkron, szekvenciális, alternáló, dinamikus a játék)

²Meg kell jegyeztem, 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.

a kooperatív játék: [336, 33]?????? a) a játékosok egymástól függetlenül választják stratégiáikat b) a játékosok nem tehetnek a játék megkezdése előtt olyan ígéreteket, amelyeket 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

A játékosok szociális preferenciái

A társas értékorientáció elmélete olyan eszköztárat fejlesztett ki, amely alkalmas a játékelméleti szituációkba került játékosok valós szociális preferenciáinak feltárására (lásd 97.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.

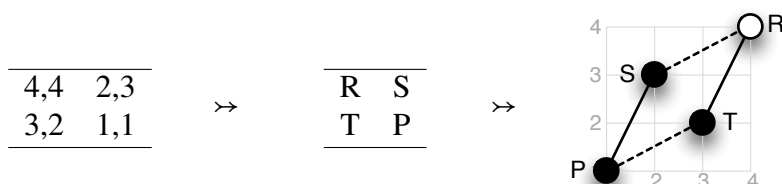
8.2 Magjátékok

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 [353], 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)

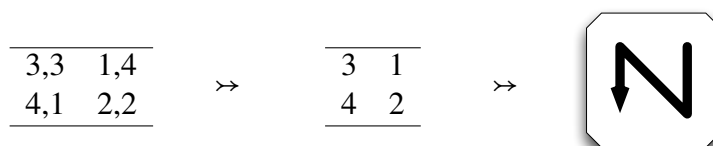
A játékokat a preferenciamátrix alapján jellemezhetjük. A preferenciamátrix a játékosok alternatívákra vonatkozó rendezéseit fejezi ki, amit számokkal jelölünk. A preferenciamátrix segítségével kifejezhető strukturális kapcsolatokat (kényszereket) ábrázolhatjuk grafikusan is. Kétféle geometriai (topológiai) 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 koordináta-rendszerben tüntetik fel. A sor-játékos (Ego) értékeit a vízszintes, az oszlop-játékos (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 [304].



Az ábrában az egyensúlyi kimenet(ek) létezését a fehér középi 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 [112]. Dragicevicnek ezt az egyszerű, de látványos megoldását (még tovább egyszerűsítve, a “végletekig stilizálva”) szintén alkalmazom az egyes játékok jellemzésekor.



A játékok leírásakor feltüntettem 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 jellemzi 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

³Például 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.

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 [353], 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.” [33]

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

———— 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 rukkolt elő arra figyel, hogy “mi lenne ha?” ki tudja megoldani a játékot? Rapoport négyféle játékstípust ír le, és háromféle (????ez biztos????) játékra hivatkozik, miközben kétféle játék egoista és altruista interpretációjáról van szó

a fogolydilemma komplementerét sokféleképpen nevezik (holtpont, PD delight, mártír)

többféle elnevezést érezhetünk intuitívnak: kizsákmányoló 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, közlegelő tragédiája [291], volunteer's dilemma [106]

—— [336, 33] játékok végeredménye, egyensúlyi helyzete,

Pareto-optimalitás: Egy elosztást Pareto-optimálisnak nevezünk, ha nem létezik 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, iteratíván 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

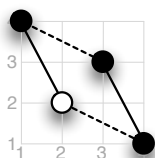
csomagolás hatása (mérő könyv, Simonovits példa:[336, 33])

	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.

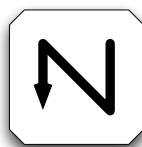
A játék fogolydilemma. Ennek ellenére ezt a játékot a kísérletekben sokkal kooperatívabban játsszák, mint az eredetit, mert nyilvánvaló benne, hogy csak a másiktól jöhet az igazi nyereség.

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 jegyezni, hogy van egy olyan elfogultság ebben a minősítésben, amely a konfliktusos jellegre fakad. Amikor John Platt a társadalmi csapdáról 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 [291]. 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 felvállalni. 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 dilemma jellegét.

A fogolydilemma játék eredeti meséje bár életszerű, de nem mondható túl gyakornak, 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 árucseré helyzetek mélyén mindig is ott rejtőzködő késztetés a partner “átverésére” – olyan példák, amelyek jelzik, hogy a társadalmi cselekvések minden szegmensében felbukkannak fogolydilemma-szerű konfliktusok [353].

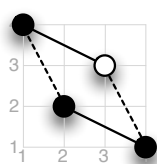
társadalmi csapda, szuboptimális megoldás, társadalmi dilemma

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): $DC > CC > DD > CD$.

Mártírsjáték

A mártírsjáték a fogolydilemma komplementár párja. Szokták holtpontságjátéknak (Deadlock) is nevezni [295], Robinsonék az ‘Anti-PD’ nevet használják, Ken Binmore pedig a ‘Prisoner’s Delight’ nevet húzta rá [44, 131]. A mártírsjáték legfontosabb vonása az, hogy



3,3	4,1
1,4	2,2

SRPT



a játék száma

Rapoport-Guyer:	9
Brams:	9
Robinson-Goforth:	166

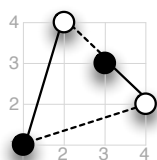
egy fogolydilemma-szerű szituációból képes kihozni a “közösségi optimumot” a játékosok számára. Ezzel egyben arra is példálul szolgál, hogy nemcsak az önzés nem mindig racionális, de előfordulhat az is, hogy az altruizmus nagyobb társadalmi nyereséget biztosít az önzéshez képest [353].

balek, mártír azért mártírajáték, 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 játszottak 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.

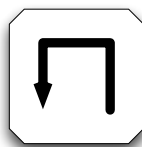
[138]

Gyáva nyúl játék



3,3	2,3
4,2	1,1

TRSP



a játék száma

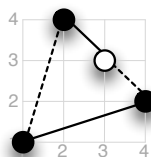
Rapoport-Guyer:	66
Brams:	57
Robinson-Goforth:	122

Ha a fogolydilemmában 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 fenyegetnek, a másikkól 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.).

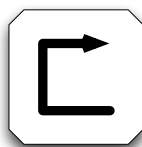
gyáva nyúl (negatív szerep), önkéntes (pozitív)

Önkéntes játék



3,3	4,2
2,4	1,1

SRTTP



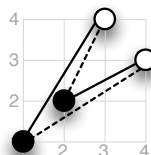
a játék száma

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 mártírijáték esetében: sokak számára a konfliktusmentessége miatt érdektelennek tűnik a játék.

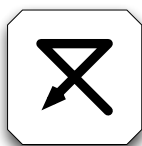
Robinson és Goforth anti-gyávanyúl játéknak (Anti-Chicken) nevezi a játékot. Az általam preferált önkéntesjáték megnevezés arra utal, hogy ezt a játékot a gyáva nyúl játék keresett/remélt önkéntesei játsszák. Amíg azonban a gyáva nyúl játékban az önkéntes magára vállalja a második legrosszabb kimenetet 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

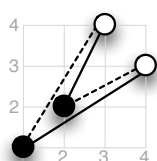
Talán ennél a játéknál (illetve akomplementer párjánál) tapasztalható a legnagyobb névkavarodás a játékelméleti szakirodalomban. A nemek harca megnevezést alkalmazzák erre a játékra és a párjára egyaránt. Bár ez a keveredés önmagában mutatja, hogy valóban értelmes és érdemes a komplementer játékokat 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 [301], 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 valamilyen módon 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 azaz, 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 vezérjátékba kerülnek azok is, akik között megszakad a telefonbeszélgetés, és el kell dönteniük, ki hívja vissza a másikat.

Hősjáték



2,2	4,3
3,4	1,1

STRP



a játék száma

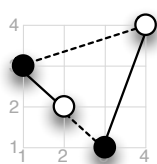
Rapoport-Guyer:	69
Brams:	55
Robinson-Goforth:	144

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. 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 megfigyelni (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 tudósnek 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.

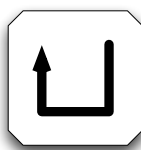
Szarvasvadászat játék

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



4,4	1,3
3,1	2,2

RTPS



a játék száma

Rapoport-Guyer:	61
Brams:	-
Robinson-Goforth:	322

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 szerezhetheti 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 helyzeté válik. Ebben az értelemben és ezen a ponton még mindig a fogolydilemma “hatása” érezhető.

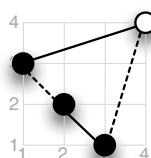
Lehet azonban másfajta értelmezést is adni a szarvasvadászat játék származtatására vonatkozóan. Kiindulhatnánk ugyanis a harmóniajátékból, amelyben ha megcserélnénk a két legkisebb értéket (az 1-t és 2-t), akkor a szarvasvadászat játékhoz jutnánk. Ezt úgy lehetne interpretálni, hogy ha a kölcsönösen 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 matrólázadás 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 [339]. Robinson és Goforth említi, hogy bizalmi dilemmaként (trust dilemma), illetve koordinációs játékként (coordination game) is felfogható a játék [304], Amartya Sen 1967-es cikke óta nevezik olykor biztosítási játéknak (Assurance Game [327]), Sam Perlo-Freeman pedig Deterrier játéknak (Deterrier) nevezi [284].

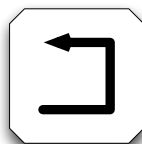
jobb ma egy veréb, mint holnap egy tűzok.

Háztulajdonos játék



4,4	3,1
1,3	2,2

RSPT



a játék száma

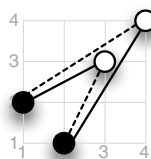
Rapoport-Guyer:	5
Brams:	-
Robinson-Goforth:	355

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 háttérében 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 legkevesbé kíváncsatos lehetőséggé, mivel a szomszéd bordélyház zavarná a háza lakásait bérbe vevő lakosokat, tehát alacsonyabb lakbért tudna tőlük beszélni. 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.

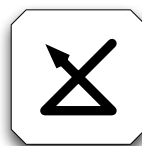
Robinson-Goforth könyv kevert harmóniajátéknak (Mixed Harmony) nevezi

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 kíváncsatosnak, a másikat dezertálásnak, tehát nem kíváncsatosnak tartanánk, mi, a játék megfigyelői. Mégis használom a kétféle minősítést, de csak azért, mert így a már megszokott módon tudok a cellákra, kimenetekre hivatkozni.

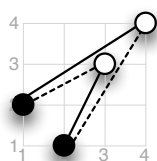
A játék két legmagasabb kifizetése a kétszeres kooperálással vagy kétszeres dezertálással érhető el. Ezért nevezik ezt koordinációs játéknak. Ennek a játéknak is két

egyensúlyi helyzete van. A kétszeres kooperáláshoz szükséges kooperálás választása az-
zal 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 kifizetés-
domináns/kifizetésérzékeny (risk dominant vagy payoff dominant) egyensúly alakul-e ki
[363].

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.

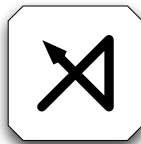
Példa lehet erre a szituációra, amikor két mobilszolgáltató (ATel és BTel) van a piacon
[363]. Az egyik percdíja 10, a másiké 20, ATel felől BTel felé érkező hívások ára 70,
fordítva 40 egység. Ezt a piaci feltételrendszert a tiszta koordinációs játék logikája szerint
rendezik el a vásárlók.

Koordinációs játék: balszélső



4,4	2,1
1,2	3,3

RPST



a játék száma

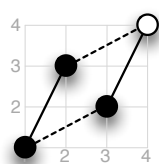
Rapoport-Guyer:	60
Brams:	-
Robinson-Goforth:	344

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 jobb
szél irányába kezdhet el rohanni. Ha mindketten ugyanabba az irányban “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) kompletter párja.

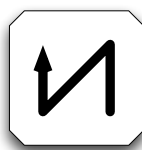
Harmóniajáték: csónakos

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 véltelenül nevezik Robinsonék harmóniajátéknak, és korábban – Olson nyomán –



4,4	2,3
3,2	1,1

RTSP



a játék száma

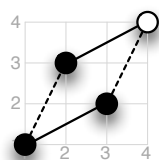
Rapoport-Guyer:	6
Brams:	-
Robinson-Goforth:	311

én is privilégiumjátékként azonosítottam, míg Skryms Prisoner's Delight-ként hivatkozik rá [340]. Kicsit más nézőpontból tekintve a játékra Sam Perlo-Freeman az 'Appeaser' (békéltető, lekenyerező) nevet használja [284]. A

A csónakos játék meséje Brian Skrymstől származik [340]. 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 partnerük 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" [284]. 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 dezertál 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 elmozdulásban van. Azon túl, hogy mindkét fél a kölcsönös kooperálást tartja a legmagasabbnak, 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 [284], míg – más szerzőkkel ellentétben – Amir N. Licht ezt nevezi harmóniajátéknak [227].

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.

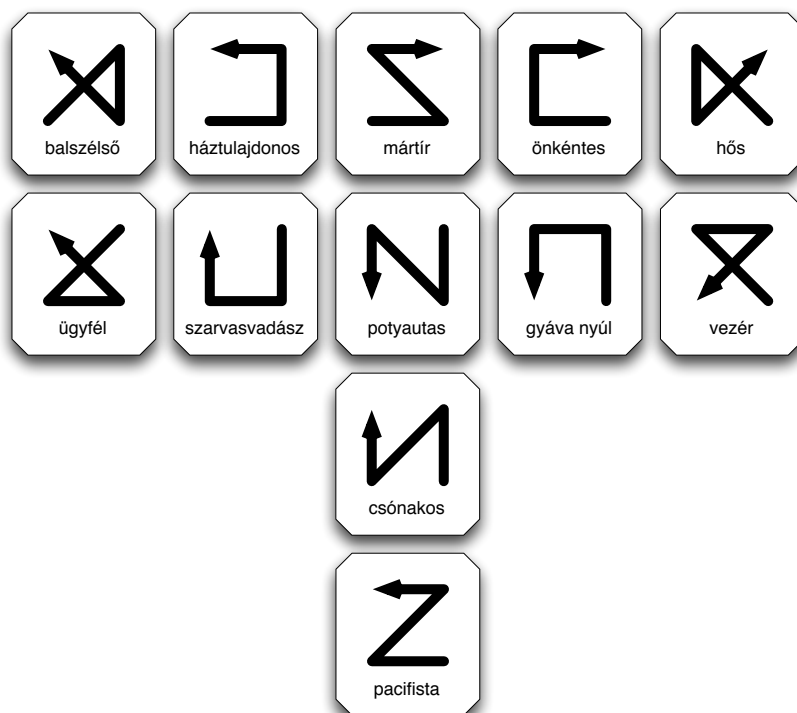


Figure 8.1: magjátékok útjelzői táblái összesítve

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 fogolydilemmából (és komplementer társából, a mártírjátékból) kaphatjuk meg a gyáva nyúl és a szarvasvadász játékokat (illetve ezek komplementereit), majd még egy “lépés téve” jobbra, illetve balra eljuthatunk a kétféle nemek harca, illetve kétféle koordinációs játékhöz.

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 szarvasvadász játékban fordulhat elő az, hogy a szuboptimális egyensúlyi pontot (2,2) választják a résztvevők, egyébként minden más esetben csak magasabb nyereményekkel lehet a játékokból kijönni.

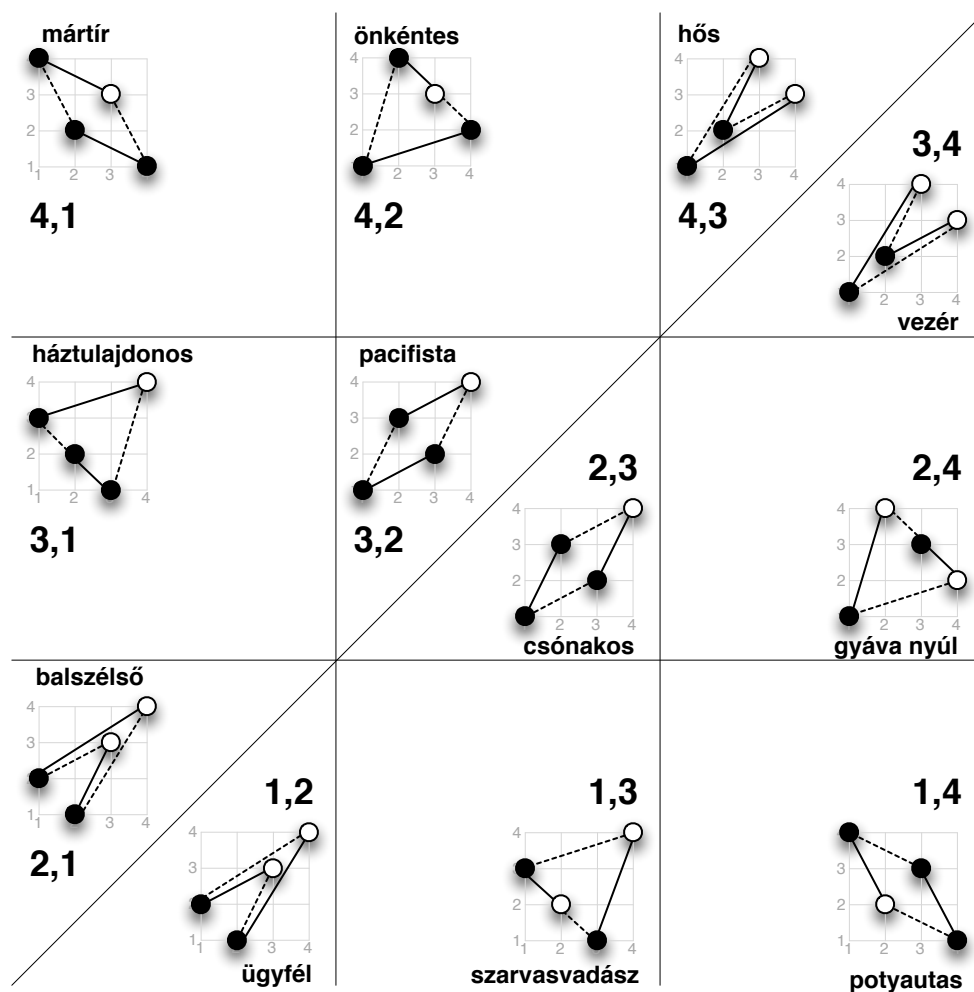


Figure 8.2: magjátékok geometriája – áttekintő ábra

A fenti két ábra érdekes összefüggéseket mutat a magjátékok között, de mindkettő csak egyetlen vonását mutatja be a játékoknak. Érdekes 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 személyiségstípusok megjelenítésére. A zodiákus tizenkét állatövi jegyét, tizenkét személyiségstípusát 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:

⁴Érdekességgént jegyzem csak meg, hogy vannak még további 12-osztatú ábráink: a nap 12 órából, az év 12 hónapból áll.

A magjátékok tipizálása

Ulmann-Margalit [374] fogolydilemma normák koordinációs normák partial/impartial normák (elfogultsági/elfogulatlansági normák)

Robinson-Goforth [304] technikai tipizálás

Wim B.G. Liebrand társadalmi dilemmák osztályozása [229] hármat különít el: fogolydilemma gyáva nyúl bizalmi (szarvasvadászat)

8.3 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

8.3.1 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 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 [304], és ezt fogadom el én is. Ebből következően az alapjátékok száma 144, az aszimmetrikus játékoké pedig 132.

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)

A szimmetrikus játékok ismeretében már könnyebb dolgunk van az aszimmetrikus játékok modellezésekor. Ha ugyanis elváltaszitjuk 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 preferenciámatrixuk, ezért a szétválasztás eltérő magjátékokat eredményez. Azt mondhatjuk tehát, 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

8.3.2 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 kevészer 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 játékokra, 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 összehasonlítással már végső győztes lehet kihozni.

8.3.3 Többszereplős játékok

A többalternatívás játékok modellezése nagyon könnyen 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 játéksituációkban, 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.

közlegelő tragédiája név és problémafelvetés Garret Hardin: [179]

DE: az n-szereplős PD-t belőle Russel Hardin csinálta meg [180]

többszereplős önkéntes dilemma: Andreas Diekmann [106]

+szempont a megváltozott helyzetre: a többszereplős helyzetek gyakorlati szempontból időben szétszúsznak, aszinkronná válnak, így az interdependencia egyidejűségi feltétele nem szorító a modellezésben is könnyít ez

Háromszereplős játékok

„... nekem nem az oroszlánál kell gyorsabban futnom, hanem a nálad”

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

8.3.4 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	kéalternatívás
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 cselekvéseméleti könyvemre + Goforth-ék 2007-es cikke

8.4 A játékok magasabbrendű kiterjesztése: büntetés

8.4.1 Másodrendű játékok – költséges büntetés

másodrendű fogolydilemma játék [278] [280]

büntetések másfajta tipizálása:

2PP (second-party punishment) másod büntetés spite büntetés rosszindulatú ?? kárt okozó ??

3PP (third-party punishment) [122] altruista büntetés harmadik

megkülönböztetik egymástól a csoporton belüli (within-group) és a csoportok közötti (inter-group) altruista büntetést

az evolúcionista magyarázat szerint a csoportok közti versengésben alakulhat ki és szilárdulhat meg az altruizmus jelensége,

parokiális altruizmus (parochial altruism) is megjelenik a közösségen belül [35]

Human punishment is motivated by inequity aversion, not a desire for reciprocity [297]

az altruista viselkedés, beállítódás neurológiai alapjai [100] [264]

Ernst Fehr és Urs Fischbacher (2002) [123] megkülönböztet

- stratégiai büntetés: Fehr-ék 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 háttérében motiváló tényezőként a normatív agresszió jelenik meg, és a cél az együttműködési 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.

8.4.2 Értékelésparadoxon

8.4.3 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 a 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.

Özséb cikk, fbf-stratégia mikor jó [356]

Van, hogy a reciprok helyzeteket szinkron (szimultán) módon hozott döntésekkel jellemezhetjük, de előfordulhatnak olyan esetek, amikor ez feltételezés nem helyénvaló. Ilyenkor az aszinkron (alternáló vagy szekvenciális) játékok használhatók, amelyek során egy adott fordulóban a játékosok nem egyszerre, hanem egymás után hozzák meg a döntéseiket. Az állatvilágból származó reciprok altruizmus egyik leghíresebb példaként említik a vámpírdenevéreket, akik között a sikeres táplálékszerző körútról visszaérkező

magjátékok	
aszimmetrikus játékok	szimmetrikus
többalternatívás játékok	kéaltalternatívás
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)

egyedek megosztják beteg vagy hoppon maradt társaikkal az élelmet, míg egy másik alkalommal ezek a szerepek cserélődnek (Wilkinson 1984). Ebben a példában a reciprocitás időben eltolt módon jelentkezik, amit az aszinkron játékokkal lehet modellezni. Kétfajta modellt is lehet állítani az aszinkron játékokra (Nowak - Sigmund 1994). 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 esetben életszerűbb, hiszen a reciprocitáson alapuló kapcsolatokban sokszor nem lehet biztosítani, hogy a segítségnyújtó és támogatott (donor és recipiens) szerepek pontos rendben kövessék egymást.

Bűnbánó stratégia Ha aszinkron játékokból képezünk játéksorozatot, akkor új sikeres stratégia jelenik meg a színen, amit 'szigorú, de igazságos' (Firm-but-fair – FBF) szabálynak neveztek el (Freen 1994), de szerintem pontosabb a 'bűnbánó stratégia' kifejezés használata. Az FBF stratégia erősen kooperatív, mert csak egy esetben dezertál: ha az előző menetben a játékos kooperált, miközben a partnere dezertált. Nem is annyira a nagylelkűség, mint inkább a "bűnbánás" az 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: (1,0,1,1). Meg kell itt jegyeznem, hogy amikor Freen 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 változatként értelmezem. Érdekes még azt is megjegyezni, hogy ugyanez a cselekvési szabály felbukkan a 'bűnbánó tit-for-tat'/'contrite TFT' [52], valamint a hírnévápoló/'standing?' stratégia (Sugden 1986) név alatt is. Ezekkel azért nem foglalkozom, mert mindkettő stratégia az indirekt reciprocitás körébe tartozik.

8.4.4 alkujátékok

Az aszinkronitásnak nem csak az ismételt játékok esetén lehet értelme. A kísérleti közgazdaságtan, viselkedés-gazdaságtan vizsgál olyan helyzeteket is, amelyek csak egyszer ját-

8.5. A JÁTÉKOK DINAMIKUS KITERJESZTÉSE: TÁRSADALMI EMLÉKEZET143

szott játékoknak felelnek meg, ám a játékosok nem egyszerre, hanem egymás után döntenek, így a későbbi döntéshozók már ismerik a partnereik választásait.

ultimátumjáték diktátorjáték public goods game trust game

Ultimátumjáték

a második játékos és az első játékos egyaránt egy-személyes játékot játszik

... 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. [167]

Selten áruházlánc paradoxona (chain store paradox) [326]

called bluff játék megoldásához is kapcsolni lehet elrettentést néha beváltja valaki, mert további érvek vannak a dezertálás, háború mellett, Falkland háború

simonovits könyve [336, 33]

8.5 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 lehet vizsgálni. Az egyes játékokat fordulónak 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	kéalternatívás
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 stratégia 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 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 játékok 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 (esetleg későbbi) döntésétől, azaz más játékok (más döntések) eredményétől függ az aktuális döntés maga, hanem csak a véletlentől.

8.5.1 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 "felelő", egymásra hivatkozó döntéseket, cselekvéseket köt össze, és ezzel lehetővé teszi azt, hogy modellezni tudjuk a reciprocitás jelenségét. A reciprocitás fogalmát cselekvések időben eltoltsági 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 eltoltsági viszonzási rendszer lényegét (lásd Gouldner híres cikkét a reciprocitás értelmezéséről [161]). Fontosnak tartom azonban megjegyezni, hogy egy szempontból nagyon félrevezető itt reciprocitásról beszélni. A reciprocitás fogalmának gouldneri értelmezése ugyanis normatív szinten zajlik (Gouldner összekapcsolódó jogszerűségekről és kötelezettségekről beszél), míg a játékelméletnek ezen a szintjén még sehol nincsenek normák.

pozitív és negatív reciprocitás

Még jóval a játékelmélet, a kísérleti közgazdaságtan előtt az antropológus Marshall Sahlins is megteszi ezt kettévontást

Fehr és Gächter pozitív és negatív reciprocitást különböztet meg aszerint, hogy valaki kooperál vagy dezertál, büntet [124], Herbert Gintis ugyanerre a kettősségre a gyenge, illetve erős reciprocitás fogalmakat alkalmazza [153].

az altruizmus és reciprocitás kapcsolatát elemzik Gintis és társai Explaining altruistic behavior in humans [155]

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... (Sahlins 1973: 264).

8.5.2 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.

magjátékok	
aszimmetrikus játékok	szimmetrikus
többalternatívás játékok	kéalternatívás
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)

szöveg

ez kis méretű társadalmat ragadja meg

Rapoport már a hatvanas években megmondta [300]

Stratégia

Evolúciósan stabil stratégia Ugyancsak az evolúciós elméletek területéről szivárgott át egy másik fontos terminus a játékelmélet vidékére: az evolúciósan stabil stratégia fogalma. Ezzel azt a minőséget lehet kifejezni, 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 merültek fel. Az evolúciósan stabil stratégiák képesek ellenállni az ilyenfajta “támadásoknak”, vagyis az adott mintázatot tartósan fenn tudják tartani a közösség egészén belül. A tit-for-tat stratégia kezdeti ragyogását kissé beárnyékolta az a tény, amikor kiderült, hogy bár a stratégia kollektíve stabil, de evolúciós értelemben már nem az [53].

John Maynard Smith ESS fogalma [344] +

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 partnerünk 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 itt nem foglalkozom). 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.

Szemet szemért

A direkt reciprocitás fogalmában megbújik a pozitív és negatív reciprocitás fogalma is.

Magyarul is rengeteg közmondás, maxima fejezi ki a direkt reciprocitás szabályát.

Jótett helyébe jót várj.

Ki mint vet, úgy arat.

Szemet szemért, fogat fogért.

Amilyen az adjonisten, olyan 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.

Nagy kérdés, hogy az arany szabály tényleg ide tartozik-e. Mert ha igen, akkor be kell mutatni azt, hogy az összes világvallásban felbukkan, és idézni a megfelelő könyvet.

42. Barátnak barátja légy,
adományt adománnyal
hálálj híven.
Gúnyt fizess meg gúnnal,
s ne maradj fösvény,
fizetvén gonoszért.
43. Barátnak barátja légy,
s barátod barátjának
szintén szívese;
ám aki ellened,
egyetlen hívét
se szíveld soha. (Edda [276])

Legalább Hamurabbi törvénykönyve óta a lex talionis elve az arányos büntetés kívánalmát és gyakorlatát hozta el. Éppen ezért ez az elv, megjelenésekor humánus volt a korábbi állapotokhoz, a korábbi büntetési elvekhez és gyakorlatokhoz képest. A “szemet

8.5. A JÁTEKOK DINAMIKUS KITERJESZTÉSE: TÁRSADALMI EMLÉKEZET147

szemért” elvet úgy kellett érteni, hogy “csak szemet szemért”, nem pedig úgy, ahogy korábban tették, vagyis hogy “életet szemért”.

"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.)

A négy másodrendű stratégia

nincs több,

ismerős mélymintázat [356]

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 (Ridley 1997). 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 (Nowak - Sigmund 1993), valamint a Pavlov megnevezések (Kraines - Kraines 1989, Kraines - Kraines 1993) 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 szimmetrikus cselekvések esetén (CC és DD) kooperál, aszimmetrikus helyzetekben (CD és DC) dezertál, amit így írhatunk fel: (1,0,0,1). 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 kíváncsiságát 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 korábban bevezetett formalizmusnak megfelelően így írhatjuk fel: (1,0,0,1). 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 (Ridley 1997).

Haragtartó stratégia Az elemzésekben időről időre felbukkan a Grim trigger (GT) stratégia [143], amit haragtartó, engesztelhetetlen, zord vagy zordra váltó stratégiának fordíthatnánk (Axelrod 1984-es könyvében ezt Friedman stratégiának nevezi [22]). 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 szimulációs eredmények azonban azt mutatják, hogy ez a stratégia sosem tud igazán jól szerepelni, főleg zajos játékokban, mert nagyon könnyen keveredik megtorlási spirálba [21]. A haragtartó stratégiára a következő mintázat jellemző: (1,0,0,0).

Négy alapstratégia Az eddig bemutatott stratégiák ismeretében felmerülhet pár kérdés. Miért pont ezek a stratégiák szerepeltek eddig? Honnan tudhatjuk, hogy nincsenek mások, esetleg jobbak? Vajon mennyi stratégia lehetséges egyáltalán? Kezdjük az utóbbi kérdés megválaszolásával! A stratégiák itt az előző forduló lépései (a “múlt”) alapján hoznak döntést. Ha eltekintünk a valószínűségi változótól, akkor egy stratégia a múlt mind a négy különböző állapotához (CC-CD-DC-DD) két lehetséges érték valamelyikét rendeli: kooperálást vagy dezertálást (ezeket jelöljük 1-gyel, illetve 0-val). Ezt 16-féleképpen tehetjük meg, tehát ennyi stratégia létezik. 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). Négy olyan szabály van, amely közös abban, hogy mit tesz azután, hogy a játékos kooperált az előző fordulóban: 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). E négy stratégia közt abban van különbség, hogy mit tesznek akkor, amikor a játékos dezertált az előző menetben. Ilyen esetből négyféle lehet, és az az érdekes, hogy pont ez alapján különíthetjük el a korábban bemutatott stratégiáinkat is:

utánzó	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, hogy 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 utánzó (tit-for-tat) stratégia szerint kell eljárunk. Az előző menetbeli dezertálásunk esetén adott négyféle cselekvés lehetőség pedig azt a négy stratégiát írja le, amelyek kooperációs potenciálját a játékelmélet is alátámasztotta, de a mindennapi életben is megtapasztalhattunk évezredek óta.

Társadalmi hiba – zajos játékok

8.6 A játékok közösségi kiterjesztése: társadalmi értékelés

8.6.1 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 jelenségre, az értékelésre, 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 Donor és 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’, mostantól viszont azt vizsgáljuk, hogy ‘miért nyújt segítséget a Donor a Recipiens számára’. A ‘Donor’ terminus aktívabban fejezi ki a kooperálást, segítségnyújtás tényét, elvárását, lehetőségét, míg a ‘Recipiens’ terminus jóval inkább utal a 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.

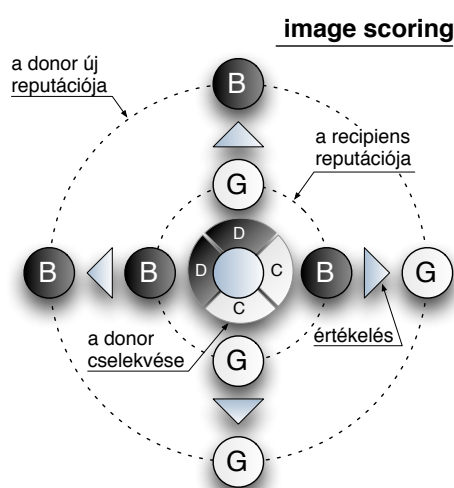
Az indirekt reciprocitás, a reputáció modelljében fel kell vennünk egy teljesen új szereptípust is, az értékelést, a reputációképzést végző közösség fogalmát. Erre a reputációs elméletek kevésbé, de legalábbis nem eléggé reflektáltak, pedig nagyon fontos ez a szerep. Hogy miért is? A direkt reciprocitás 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.

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. Értékelési rendszerek, mechanizmusok

Kép-mutató (image scoring) értékelési rendszer

Szöveg



az

Image scoring and cooperation in a cleaner fish mutualism [61]

simple standing

stern judging

shunning

spite és indirekt reciprocitás

altruizmus és indirekt reciprocitás mellett

agresszió (spite) és indirect reciprocitás kapcsolata: Evolution of spite through indirect reciprocity [201]

8.6.2 Értékelési rendszerek

shame culture vs. guilt culture

Punishment, Cooperation, and Cheater Detection in ‘Noisy’ Social Exchange [48]

8.7 Koordinációs problémák, közös tudás, aszimmetrikus információ

lewisi értelemben vett konvenció

közös tudás

8.8 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 Tabor, 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.

[83]

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).

8.9 Kollektív cselekvés, kollektív döntés

Társadalmi választások elmélete [258]

8.10 Koalícióelmélet

Simmel

Somlai Péter

magjátékok	
aszimmetrikus játékok	szimmetrikus
többalternatívás játékok	kéalternatívás
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)

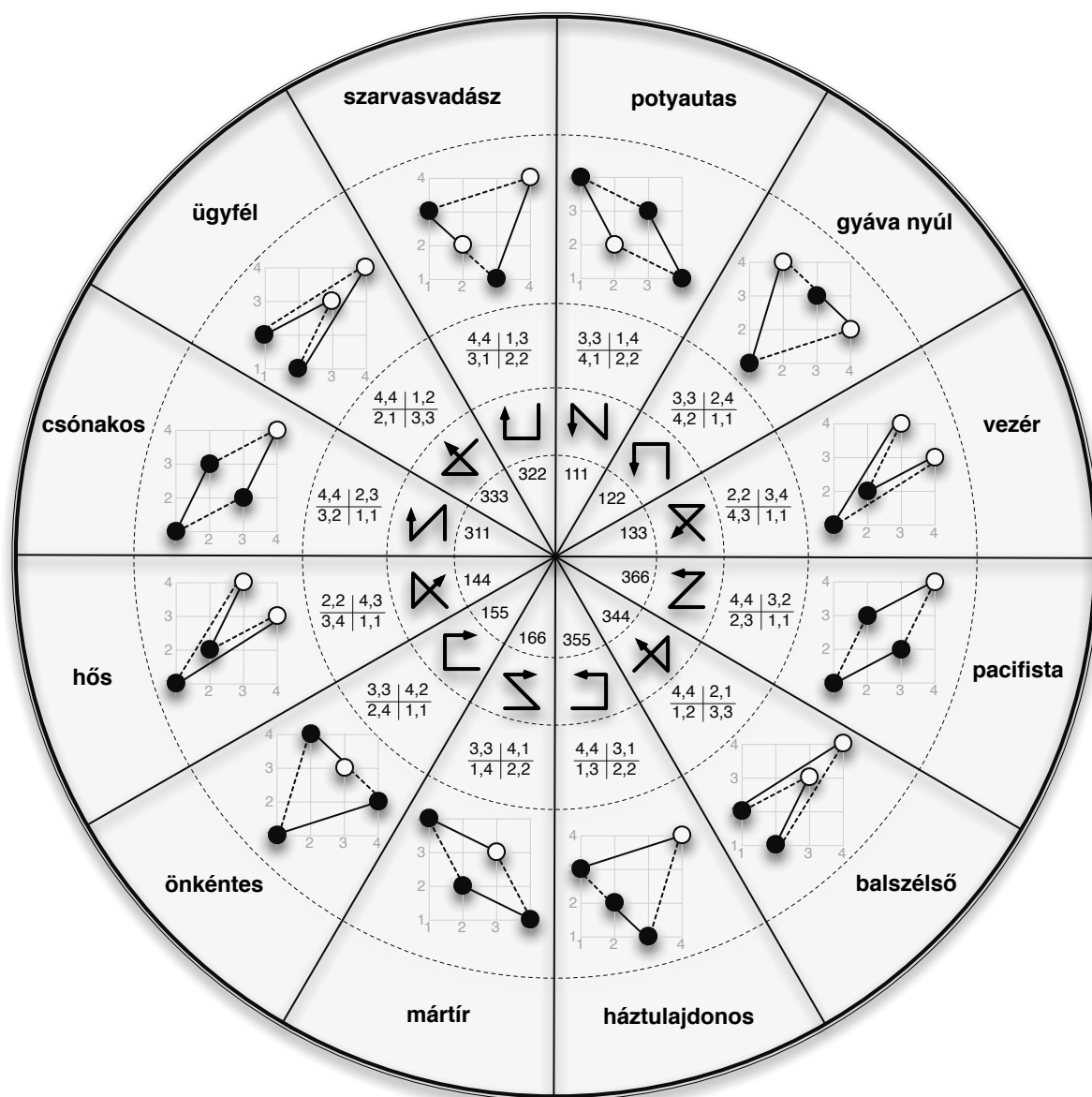


Figure 8.3: a tizenkét magjáték kördiagramja

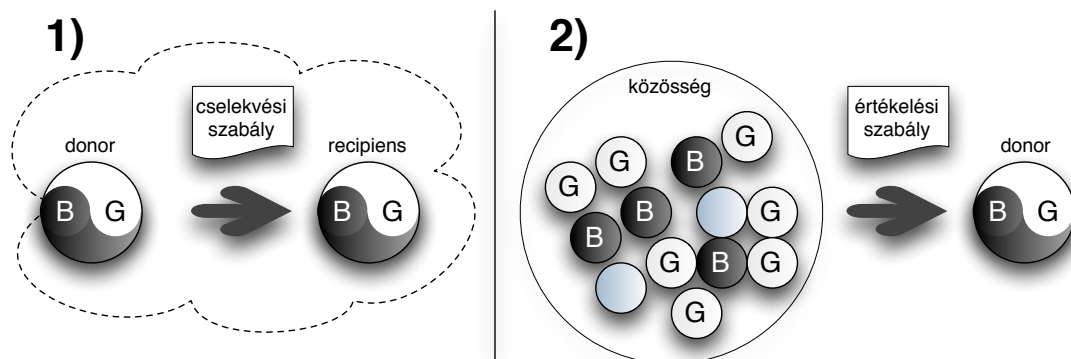


Figure 8.4: a stratégia két összetevője: cselekvési és értékelési szabály

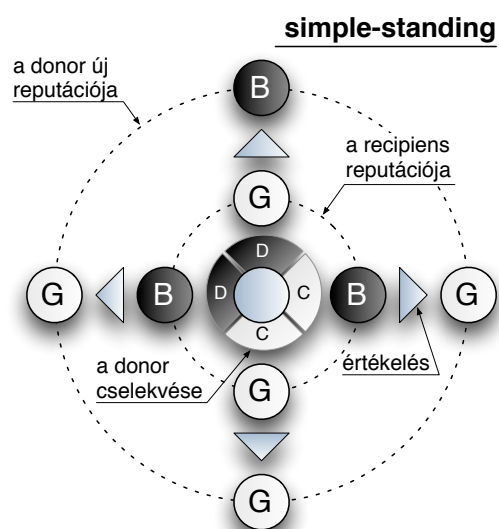


Figure 8.5: a simple standing értékelési szabály

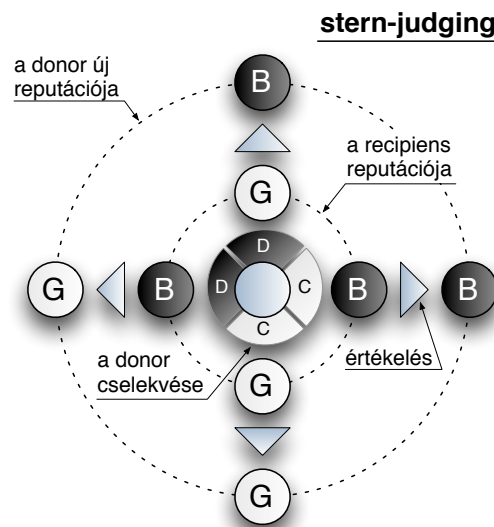


Figure 8.6: a stern judging értékelési szabály

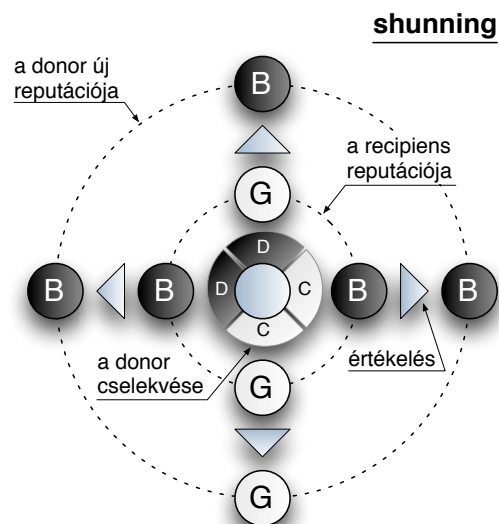


Figure 8.7: a shuning értékelési szabály

Chapter 9

Társadalmi pozíció

9.1 Imperszonális deontika

Hivatkozni az első rész deontikus logikai fejezetére.

A deontikus logika – a modális logika részeként – a deontikus modalitásokkal, a kötelezések és megengedések logikájával foglalkozik.

cselekvésekkel és nem állapotokkal Ez nem mond ellent annak a ténynek, hogy a cselekvések differencia specifikáját megragadni akaró elméletek gyakran az állapotok és az állapotváltozások (az átmenetek, tranzíciók) fogalmain keresztül jutnak el a magához a cselekvés fogalmához.

9.2 Pozíciók: jogosultságok és kötelezettségek

Michael McKinsey a kötelezettség szintjei [254]

Rudolf von Jhering [388]

a jog és kötelezettség összekapcsolódása Jhering kapcsán: [146]

“Jobban kell emlékeznünk a jó tettekre, mint a rosszakra; a kapott javakra jobban, mint az adottakra.[...] Tűrjük el a velünk szemben elkövetett jogtalanságot.” [14] 1374b. 73.

“Ezen altruista mozzanatban az igazságosság olyan erényként való felfogása található meg, amely feltételezi azt, hogy az igazságosságot mindig állandó és kitartó akarattal gyakorolni és persze keresni kell. Jhering azonban explicit módon szakít ezzel a felfogással, s azt mondja: első szabály az, hogy ‘Ne tórd el a jogtalanságot!’ és csak a második az, hogy ‘Ne kövesd el azt!’.” [146]

terminológiai tisztázás:

korrelativitás: Hohfeld

komplementaritás: Sir Ross [306], Gouldner, Parsons [282]

inverz/konverz: Kanger

9.2.1 Hohfeld

A társadalmi cselekvések indoklásai során rengetegszer hivatkozunk a jog, jogosultság fogalmára:

(1) 'Jogom van ezt tenni.'

(2) ':

Wesley Newcomb Hohfeld [191]

Hohfeld kategóriái a következők:

duty	no-right	liability	disability
claim	privilege	power	immunity

A hohfeldi kategóriák magyar fordítása nem minden esetben egyszerű.

kötelezettség	joghiány	beavatkozottságnak kitettség	beavatkozásképtelenség
igény	privilegium	felhatalmazottság	immunitás

A pénzbírságot a férfi nem fizette meg, ezért a bíróság 2012. április 10-én keltezett jogerős végzésével 20 nap elzárásra változtatta át. P. Lászlónak a szabálysértési hatóság felhívása alapján 2012. július 2-án kellett volna jelentkeznie a kijelölt büntetés-végrehajtási intézetben az elzárás letöltésére, de ezt elmulasztotta. A jogi szabályozás szerint viszont az elővezetés elrendelése a jelentkezés kijelölt időpontját követő nyolc nap eredménytelen elteltét követően, ügyési jóváhagyással történhet meg.

Így P. László szabálysértési ügyében az elzárás megkezdésének kikényszerítése érdekében a rendőrség szabálysértési hatóságának az emberölés elkövetésének időpontjáig nem volt intézkedési jogosultsága, illetve kötelezettsége.

9.2.2 Kanger-Lindahl

Stig Kanger

Kanger [205, 206, 204]

Lindahl

Lars Lindahl [231]

F. B. Fitch [134, 135]

Brian F. Chellas [68, 69]

9.2.3 post Kanger-Lindahl

Anderson [5, 7, 9, 10]

Pörn [293, 294]

Krogh és Herrestad [186, 214, 215, 186]

Makinson [240, 241]

Marek Sergot és Jones [330, 203, 329, 331]

Swan Ove Hansson [177]

Vranas [398, 396, 397, 394, 395]

Sartor [317, 316]

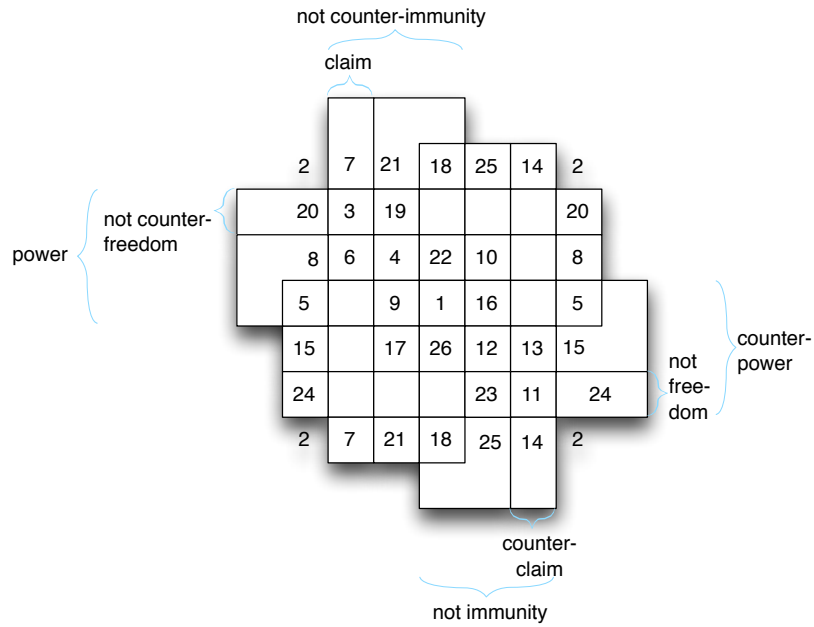


Figure 9.1: Kanger 26 típusa

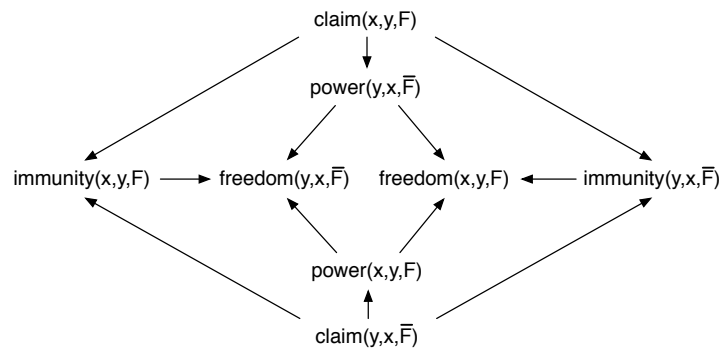


Figure 9.2: logikai összefüggések típusok között

Deontikus logika és az informatikusok

Deontic Logic and Artificial Normative Systems [158]

Deontic Logic in Computer Science, 2004 [237]

Deontic Logic in Computer Science, 2008 [381]

Games, Norms and Reasons. Logic at the Crossroads [380]

9.3 Szabadság

Deutsch és mások szabadságfelfogása

9.4 Társadalmi szerep

szerep, komplementaritás, reciprocitás

Talcott Parsons, Toward a General Theory of Action. Theoretical Foundations for the Social Sciences, [282]

sir David Ross [306]

Alvin W. Gouldner [161]

státus és privilégium

9.5 Privacy

Hirshleifer, Privacy: Its Origin, Function, and Future, három fémszabály, arany-, ezüst-, vasszabály [189]

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 10

Norma

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 [65]

The Normativity of the Natural. Human Goods, Human Virtues, and Human Flourishing [71]

Sartor et al. owl ontology of norms [308] [150]

10.1 Legitimitás

Garai Zsolt, értékek, normák [148]

A társadalmi kapcsolatok és a norma

Max Weber [404]

Eliade, szent és a profán [118]

The Sources of Normativity [211]

Performatives in a rationally based speech act theory, [82]

10.1.1 Normakövetés

A normakövet? egyéni cselekvés magyarázata során - durván - négy lehetséges mechanizmusra hivatkozhatunk: (1) a normától eltér? vagy normasért? (deviáns) viselkedés szankcionálása; (2) az egyéni stratégiaválasztástól függ? nyeremények pozitív interdependenciája (a normakövet? magatartás - pozitív - externális hatása); (3) az egyéni preferenciákba épített konformizmus; és végül (4) ex ante kommunikáció az ex post normakövet? csoport tagjai között. [88]

10.2 Játékelméleti magyarázat a normák keletkezésére

norma emergencia

Dawid Lewis konvenció [226]

Payton Young három cikke: Evolution of Conventions Economics of Conventions The dynamics of social innovation

Bicchieri, Rationality and Coordination [40]

Cristina Bicchieri, The Grammar of Society. The Nature and Dynamics of Social Norms [41]

Skyrms, Evolution of the Social Contracts, [338]

The Stag Hunt and the Evolution of Social Structure [339]

Pete Krisztián a konvenciókról, közös tudásról, intencióról [286]

Ullmann-Margalit, The Emergence of Norms [374] fogolydilemma játékok koordinációs játékok impartiality (pártosság, részrehajlás vs. pártatlanság,)

Binmore a méltányossági normák kialakulásáról: [43]

Binmore a preferenciák és normák különbségeiről, a normák funkcióiról: [45]

10.3 Normatipologizálás

összetevő	norma			
azonosító				
címzett	egyéni	közösségi	állami	
	autonóm	heteronóm		
autoritás	törvényi	rendeleti	bírósági ítéleti	szabályzati
feltétel	hipotetikus	kategorikus		
alkalom	hatályos	hatályon kívüli		
modalitás	kötelező	tiltó	megengedő	
tartalom	koordinációs	dilemmaoldó	partiality	
	elsőrendű	másodrendű	harmadrendű	
szankció	erkölcsi	konvenció	jogi	

Figure 10.1: normatípusok a norma összetevői alapján

10.3.1 Normatípusok a szankció típusa szerint

a norma következménye: szankció vagy joghatás

A normasértés esetén alkalmazott szankció típusa szerint: jog erkölcs konvenció
ez nem a david lewisi, hanem a weberi értelemben vett konvenció [404]

az erkölcs weberi definíciója (hivatkozás??) sem a ma megszokott

XX. század első harmadában már olyan elmélettel is találkozhatunk, mely szerint az állam és a jog azonos (Kelsen), vagy éppen olyannal, ami szerint az állam megelőzi a jogot (Schmitt).

Social Conventions. From Language to Law [244]

Jog

A polgári jog tartalmazza a legfontosabb alapelveket: jóhiszeműség és tisztesség elve, joggal való visszaélés tilalma, együttműködés elve.

A jog jelenti:

- a jogszabályokat, a tárgyi jogot (ius est norma agendi)
- a jogosultságot, az alanyi jogot: bizonyos személynek cselekvési lehetősége van a tárgyi jog adta keretek között (ius est facultas agendi)

bűncselekmény három fogalmi eleme:

- tényállásszerűség/büntetendőség: bűncselekmény az a cselekmény, melyre a törvény büntetés kiszabását rendeli (nullum crimen sine lege, nulla poena sine lege). A büntetendőség arra ad választ, hogy a cselekmény kimerítette-e egy törvényi tényállás kereteit
- társadalomra veszélyesség/jogellenesség: a törvényi tényállásba ütköző cselekmény jogellenes, nincs olyan jogellenességet kizáró ok, mely a cselekményt jogszerűvé tenné.
- bűnösség: az elkövetőt szándékosság vagy gondatlanság jellemezte-e

angolszász jogrendszerekben a bűncselekménynek két fő fogalmi eleme van: mens rea/bűnös tudat actus reus/bűnös tett

in personam jog: meghatározott személlyel vagy személyekkel szembeni, relatív védelmet élvező jog, kötelmi jog

in rem, dologi jog: a személyek számára abszolút jogi uralmat biztosít a dologi javak felett

“A dologi jog Grosschmid Béni szavaival élve: "sűrített negatív kötelelem" ? a dolog tulajdonosával szemben mindenki túrni köteles, a dologi jog abszolút hatályú, mindenkivel szemben fennálló jog, mely a fennálló helyzet változatlanóságát (statika) biztosítja a jogosult javára.”

“Nem a közvetlen ‘erőszak’, hanem a jogi kényszer a pozitív jog szükséges eleme, azonban a jogi előírások kötelező erejét a jogi érvényesség, s nem pedig az erőszak alapozza meg.” [144]

10.3.2 Normatípusok a logikai rend szerint

a norma tartalma és szankciója alapján

Első-, másod- vagy harmadrendű norma

Hart elsőleges és másodlagos szabályai [183]

Elméletileg elképzelhető lenne további magasabbrendű norma

10.3.3 Normatípusok a deontikus modalitás szerint

10.3.4 Normatípusok a tartalom szerint

sokféle tipizálás elképzelhető lenne a norma tartalma alapján.

játékosztályok szerint

Ulmann-Margalit [374]

10.3.5 Normatípusok a címzett és autoritás viszonya alapján

A norma címzettjének és autoritásának egymáshoz való viszonya alapján: Autonóm vs. heteronóm norma

10.3.6 Normatípusok a feltételelesség alapján

Hipotetikus vagy kategorikus norma

kazuisztikus vs. apodiktikus keresztény lexikon

egy félreértés: Jakab András: A norma szerkezetének vizsgálata az occasion és a condition különbsége tényállás és jogkövetkezmény között van egy kondicionális kapcsoló, amit lehet másként is kifejezni, ahogy Jakab mutatja, de a hipotetikusság/kategorikusság ebben az értelemben csak a "tényállásra", még pontosabban arra a diszpozícióra vonatkozik, amely a cselekvésre vonatkozó előírást tartalmazza

példa a hipotetikusság értelmességére (és szükségességére): ha megdobnak kővel, dobd vissza kenyérrel vagy szemet szemért ezek hipotetikus normák dobd meg kenyérrel vagy szúrd ki a szemét kategorikus normák

és érezhető a különbség

10.3.7 Normatípusok a címzett alapján

egyéni: az individuumra érvényes közösségi: valamely közösség virtuális terében minden közösségi tagra érvényes állami: államhatárokon belül mindenre, minden ott tartózkodóra érvényes

10.3.8 Normatípusok az alkalom alapján

hatályos hatályon kívüli

10.3.9 Normatípusok az autoritás alapján

törvény, rendelet, bírósági ítélet, szabályzat törvényhozó szerv mint autoritás: törvény végrehajtó szerv mint autoritás: rendelet bírósági szerv mint autoritás: ítélet nem államigazgatási szerv mint autoritás: szabályzat (szervezeti és működési szabályzat

10.4 Rítus, narratíva

narratív megközelítés

[368]

Chapter 11

Normavezérelt társadalmi kapcsolat

11.1 Uralom

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.” [49, 156]

11.2 Hatalom és uralom viszonya

11.3 Szervezet

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.” [190]

Chapter 12

Egyéb

12.1 mechanizmusok

Gibbard-Satterthwaite tétele: Nincs olyan racionális döntési mechanizmus egy olyan legalább háromszemélyes helyzetben, amely kielégítené az (U) és (SP) feltételeket és diktátormentes lenne.

ha azt akarjuk elérni, hogy a racionális döntés őszinte legyen (megfeleljen az (SP) 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).

McKelvey tétele: Ha legalább három politikai kérdés van napirenden, akkor a döntési folyamat nem diktátormentes ? a napirendet a ?diktátor? fogja irányítani

Ismétléses játszású fogolydilemma ? a kooperáció ?felértékelődésének? esélye: Ha egy vállalat vagy egy üzletember becsapja partnerét, abban az esetben semmi probléma nem áll elő, ha a jövőben nem kíván vele üzletet kötni, azaz egyszeri játszású a játék, azonban ismételt játszás esetén lemondhat arról, hogy partnere szóba áll vele.

12.2 Antinómiák, önreferencia

[341]

12.3 Hibák, javítandók

bibliográfia: harvard stílusú hivatkozás kell a címcímekben most kiveszi a nagybetűket, ezt kikapcsolni von előnevet figyelembe veszi a névrendezésnél (von wright) nem-magyar neveknél családnév+,+keresztnev 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 érvényesíteni 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] 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.
- [2] Icek Ajzen: The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50. évf. (1991), 179–211. p.
- [3] 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.
- [4] Ambrus-Lakatos Loránd – Meszerics Tamás: Az ultimátumjáték elemzéséhez. *Közgazdasági Szemle*, L. évf. (2003. június), 505–518. p.
- [5] Alan Ross Anderson: The logic of norms. *Logique et Analyse*, 2. évf. (1958), 84–91. p.
- [6] Alan Ross Anderson: A reduction of deontic logic to alethic modal logic. *Mind, New Series*, 67. évf. (1958) 265. sz., 100–103. p.
- [7] Alan Ross Anderson: Logic, norms and roles. *Ratio*, 4. évf. (1962), 36–49. p.
- [8] Alan Ross Anderson: Some nasty problems in the formal logic of ethics. *Noûs*, 1. évf. (1967) 4. sz., 345–360. p.
- [9] Alan Ross Anderson: The logic of hohfeldian propositions. *Logique et Analyse*, 1970., 231–242. p.
- [10] Alan Ross Anderson – Omar Khayyam Moore: The formal analysis of normative concepts. *American Sociological Review*, 22. évf. (1957) 1. sz., 9–17. p.
- [11] Hajnal Andr  ka – 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.
- [12] G.E.M. Anscombe: *Intention*. 1963, Harvard University Press.
- [13] Dan Ariely: *Kisz  m  that  an irracion  lis. A racion  lisnak v  lt d  nt  seinket alak  t   rejtett er  kr  l*. 2011, Gabo Kiad  .
- [14] Arisztotel  sz: *R  torika*. Budapest, 1982, Gondolat.

- [15] Gustaf Arrhenius: Superiority in values. *Philosophical Studies*, 123. évf. (2005), 97–114. p.
- [16] Kenneth J. Arrow: A difficulty in the concept of social welfare. *Journal of Political Economy*, 58. évf. (1950) 4. sz., 328–346. p.
- [17] Robert Audi (szerk.): *The Cambridge Dictionary of Philosophy*. 1999, Cambridge University Press.
- [18] Aurelius Augustinus: *Vallomások*. Budapest, 1982, Gondolat Kiadó,.
- [19] Robert J. Aumann: War and peace. nobel prize lecture, 2005. December.
- [20] Robert J. Aumann: Rule-rationality versus act-rationality. Jelentés, 2008, Center for the Study of Rationality at The Hebrew University of Jerusalem.
- [21] Robert Axelrod: On six advances in cooperation theory. *Analyse & Kritik*, 22. évf. (2000) 1. sz., 130–151. p.
- [22] Robert Axelrod: *The Evolution of Cooperation*. 2006 (1984), Basic Book.
- [23] Balkányi Péter: A cselekvések racionalitásának kérdéséről. *Világosság*, 44. évf. (2003) 3-4. sz., 171–181. p.
- [24] Gary S. Becker: A theory of social interactions. *Journal of Political Economy*, 82. évf. (1974) 6. sz., 1063–1093. p.
- [25] 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>.
- [26] 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.
- [27] Nuel Belnap: Backwards and forwards in the modal logic of agency. *Philosophy and Phenomenological Research*, 51. évf. (1991) 4. sz., 777–807. p.
- [28] Nuel Belnap–Michael Perloff: Seeing to it that: A canonical form for agentives. *Theoria*, 54. évf. (1988) 3. sz., 175–199. p.
- [29] 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.
- [30] Nuel Belnap–Michael Perloff–Ming Xu: *Facing the Future. Agents and Choices in Our Indeterminist World*. 2001, Oxford University Press.
- [31] Martin Mose Bentzen: *Stit, Iit, and Deontic Logic for Action Types*. PhD értekezés (Section for Philosophy and Science Studies Roskilde University). 2009.
- [32] Bereczkei Tamás: *Az erény természete*. Budapest, 2009, Typotex.

- [33] 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.
- [34] Theodore Bergstrom: A fresh look at the rotten kid theorem – and other household mysteries. *Journal of Political Economy*, 97. évf. (1989), 1138–1159. p.
- [35] Helen Bernhard–Urs Fischbacher–Ernst Fehr: Parochial altruism in humans. *Nature*, 442. évf. (2006) 7105. sz., 912–915. p.
- [36] Bertalan László: Előadások Max Weberről: Max Weber cselekvésmélete és cselekvéstipológiája. *Szociológiai Szemle*, 11. évf. (2001) 3. sz., 113–136. p.
URL <http://www.szociologia.hu/dynamic/0103bertalan.htm>.
- [37] 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.
- [38] Jean-Yves Béziau: New light on the square of oppositions and its nameless corner. *Logical Investigations*, 10. évf. (2003), 218–233. p.
- [39] Jean-Yves Béziau: The power of the hexagon. *Logica Universalis*, 6. évf. (2012), 1–43. p.
- [40] Cristina Bicchieri: *Rationality and Coordination*. 1993, Cambridge University Press.
- [41] Cristina Bicchieri: *The Grammar of Society. The Nature and Dynamics of Social Norms*. 2006, Cambridge University Press.
- [42] M. Bienvenu–J. Lang–N. Wilson: From preference logics to preference languages, and back. In *Proceedings of the Twelfth International Conference on the Principles of Knowledge Representation and Reasoning (KR 2010)* (konferenciaanyag). 2010.
- [43] Ken Binmore: The evolution of fairness norms. *Rationality and Society*, 10. évf. (1998), 275–301. p.
- [44] Ken Binmore: *Game Theory. A Very Short Introduction*. 2007, Oxford University Press.
- [45] 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.
- [46] 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.

- [47] G. E. Bolton – A. Ockenfels: Erc: a theory of equity, reciprocity and competition. *American Economic Review*, 90. évf. (2000) 1. sz., 166–193. p.
- [48] Gary Bornstein – Ori Weisel: Punishment, cooperation, and cheater detection in “noisy” social exchange. *Games*, 1. évf. (2010) 1. sz., 18–33. p.
- [49] Pierre Bourdieu: *A társadalmi egyenlőtlenségek újratermelődése*. 1978, Gondolat.
- [50] Pierre Bourdieu: *A gyakorlat elméletének vázlatja. Három kabil etnológiai tanulmány*. 2009, Napvilág Kiadó.
- [51] Pierre Bourdieu – Loïc J.D. Wacquant: *An Invitation to Reflexive Sociology*. 2009, Polity Press.
- [52] Robert Boyd: Mistakes Allow Evolutionary Stability in the Repeated Prisoner’s Dilemma Game. *Journal of Theoretical Biology*, 136. évf. (1989), 47–56. p.
- [53] 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.
- [54] Ronald J. Brachman – Hector J. Levesque: *Knowledge Representation and Reasoning*. 2004, Morgan Kaufmann Publishers.
- [55] Steven J. Brams: *Theory of moves*. 1994, Cambridge University Press.
- [56] 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.
- [57] Michael Bratman: *Intention, Plans, and Practical Reason*. Cambridge, MA, 1987, Harvard University Press.
- [58] Michael E. Bratman: *Structures of Agency*. Oxford, 2006, Oxford University Press.
- [59] Franz Brentano: A pszichikai és fizikai jelenségek közötti különbségről. In Csaba Pléh (szerk.): *Pszichológiatörténeti szöveggyűjtemény*. I. köt. 1983, Tankönyvkiadó, 157–163. p.
- [60] Franz Brentano: *Az erkölcsi ismeret eredete*. 1994, Kossuth Könyvkiadó.
- [61] R. Bshary – A.S. Grutter: Image scoring and cooperation in a cleaner fish mutualism. *Nature*, 441. évf. (2006) 22. sz., 975–978. p.
- [62] James N. Buchanan: A törvény és a láthatatlan kéz. In *Piac, állam, alkotmányosság*. 1992, KJK, 39–51. p.
- [63] 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.
- [64] Sarah Buss – Lee Overton (szerk.): *Contours of Agency. Essays on Themes from Harry Frankfurt*. 2002, The MIT Press.

- [65] Georges Canguilhem: *A normális és a kóros*. Budapest, 2004, Gondolat.
- [66] 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.
- [67] Gary Charness – Matthew Rabin: Understanding social preferences with simple tests. *Quarterly Journal of Economics*, 117. évf. (2002), 817–869. p.
- [68] Brian F. Chellas: Imperatives. *Theoria*, 37. évf. (1971) 2. sz., 114–129. p.
- [69] Brian Farrell Chellas: *The logical form of imperatives*. 1968, Dept. of Philosophy, Stanford University.
- [70] Brian Farrell Chellas: *Modal Logic: An Introduction*. 1980, Cambridge University Press. ISBN 0-521-22476-4.
- [71] Mark J. Cherry (szerk.): *The Normativity of the Natural. Human Goods, Human Virtues, and Human Flourishing*. 2009, Springer.
- [72] Roderick M. Chisholm: Supererogation and offence: A conceptual scheme for ethics. *Ratio*, 1963. 5. sz., 1–14. p.
- [73] Roderick M. Chisholm: The ethics of requirement. *American Philosophical Quarterly*, 1. évf. (1964) 2. sz., 147–153. p.
- [74] Roderick M. Chisholm: *The Encyclopedia of Philosophy*. 3 and 4. köt. Intentionality fejezet. 1967, MacMillan, 201–204. p.
- [75] 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.
- [76] Roderick M. Chisholm – Ernest Sosa: Intrinsic preferability and the problem of supererogation. *Synthese*, 16. évf. (1966) 3/4. sz., 321–331. p.
- [77] Roderick M. Chisholm – Ernest Sosa: On the logic of "intrinsically better". *American Philosophical Quarterly*, 1966., 244–249. p.
- [78] 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*.
- [79] Tim Clutton-Brock: Cooperation between non-kin in animal societies. *Nature*, 462. évf. (2009) 5. sz., 51–57. p.
- [80] 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.
- [81] Philip R. Cohen – Hector J. Levesque: Intention is choice with commitment. *Artificial Intelligence*, 42. évf. (1990) 213–261. sz.

- [82] 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.
- [83] 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.
- [84] 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.
- [85] J.C. Cox – D. Friedman – S. Gjerstad: A tractable model of reciprocity and fairness. *Games and Economic Behavior*, 59. évf. (2007), 17–45. p.
- [86] Charles B. Cross: The modal logic of discrepancy. *Journal of Philosophical Logic*, 26. évf. (1997) 2. sz., 143–168. p.
- [87] Csanád Csáji Balázs – Miklós Rédei: A racionális demokratikus véleményösszegzés korlátairól, 2007.
- [88] 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>.
- [89] Csontos László – Király Júlia – László Géza: Az ezredvégi nagy borzongás. *Közgazdasági Szemle*, XLIV. évf. (1997) július-augusztus. sz., 569–596. p.
- [90] A. Damasio: Understanding the mind's will. *The Behavioral and Brain Sciences*, 8. évf. (1985) 4. sz., 589. p.
- [91] Antonio Damasio: *Descartes tévedése. Érzelem, értelem és az emberi agy*. 1996, Aduprint.
- [92] Peter Danielson: Competition among cooperators: Altruism and reciprocity. *PNAS*, 99. évf. (2002), 7237–7242. p.
- [93] Arthur Danto: Freedom and forbearance. In K. Lehrer (szerk.): *Freedom and Determinism*. 1966, Random House, 45–63. p.
- [94] 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.
- [95] Wayne A. Davis: A causal theory of intending. *American Philosophical Quarterly*, 21. évf. (1984) 1. sz., 43–54. p.
- [96] 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.
- [97] Richard Dawkins: *Az önző gén*. 2005, Kossuth Könyvkiadó.

- [98] Howard L. Dazeley – Wolfgang L. Gombocz: Interpreting anselm as logician a. *Synthese*, 40. évf. (1979) 1. sz., 71–96. p.
- [99] D. de Jongh – F. 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.
- [100] 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.
- [101] 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.
- [102] Daniel C. Dennett: *Content and Consciousness*. 1969, Routledge.
- [103] Daniel C. Dennett: *Az intencionalitás filozófiája*. 1998, Osiris.
- [104] Jacques Derrida: Mit jelent a “tanítani” szó? *Iskolakultúra*, 2007. 3. sz., 3–16. p.
- [105] John Dewey: *Experience and Nature*. 1929, George Allan and Unwin.
- [106] Andreas Diekmann: Volunteer’s dilemma. *Journal of Conflict Resolution*, 29. évf. (1985) 4. sz., 605–610. p.
- [107] Franz Dietrich – Christian List: Arrow’s theorem in judgment aggregation. *Social Choice and Welfare*, 29. évf. (2007) 1. sz., 19–33. p.
- [108] Franz Dietrich – Christian List: A reason-based theory of rational choice. *Noûs*, forthcoming. évf. (2012).
- [109] Franz Dietrich – Christian List: Where do preferences come from? *International Journal of Game Theory*, 2012.
- [110] Dombi J. – Vincze N.: A lexikografikus döntési eljárás kapcsolata a hagyományos döntési eljárásokkal. In *XXV. Magyar Operációutatói Konferencia Előadáskivonatok, Debrecen, Hungary* (konferenciaanyag), 32. köt. 2001.
- [111] Dombi J. – Vincze N.: Lexikografikus döntések egy általános döntési modellben. *Sigma*, 36. évf. (2005) 3-4. sz., 149–162. p.
- [112] Pierre Dragicevic: Játékelméleti jelölések.
URL <http://www.lri.fr/~dragice/gameicons/>.
- [113] M. Dufwenberg – G. Kirchsteiger: A theory of sequential reciprocity. *Games and Economic Behavior*, 47. évf. (2004) 2. sz., 268–298. p.
- [114] Charles Duhigg: *A szokás hatalma*. 2012, Casparus Kiadó.
- [115] Émile Durkheim: *Az öngyilkosság. Szociológiai tanulmány*. 1982, KJK.

- [116] Douglas Ehring: “normal” intentional action. *Philosophy and Phenomenological Research*, 46. évf. (1985) 1. sz., 155–157. p.
- [117] 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.
- [118] Mircea Eliade: *A szent és a profán*. Budapest, 1996, Európa Könyvkiadó.
- [119] Jon Elster: *Ulysses and the Sirens*. 1979, Cambridge University Press.
- [120] Jon Elster: *A társadalom fogaskerei*. 1997, Osiris.
- [121] 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.
- [122] Ernst Fehr – Urs Fischbacher: Third-party punishment and social norms. *Evolution and Human Behavior*, 25. évf. (2004), 63–87. p.
- [123] Ernst Fehr – Urs Fischbacher – Simon Gächter: Strong reciprocity, human cooperation and the enforcement of social norms. *Human Nature*, 13. évf. (2002), 1–25. p.
- [124] Ernst Fehr – Simon Gächter: Fairness and Retaliation: The Economics of Reciprocity. *Journal of Economic Perspectives*, 14. évf. (Summer 2000) 3. sz., 159–181. p.
- [125] 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.
- [126] Joel Feinberg: Supererogation and rules. *Ethics*, 71. évf. (1961) 4. sz., 276–288. p.
- [127] Todd E. Feinberg: *Altered Egos. How the Brain Creates the Self*. 2001, Oxford University Press.
- [128] Ferenczi Miklós: *Matematikai logika*. 2002, Műszaki Könyvkiadó.
- [129] 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.
- [130] Peter C. Fishburn: A study of lexicographic expected utility. *Management Science*, 17. évf. (1971. July) 11. sz., 672–678. p.
- [131] Peter C. Fishburn: Lexicographic orders, utilities and decision rules: A survey. *Management Science*, 20. évf. (1974. July) 11. sz., 1442–1471. p.
- [132] Peter C. Fishburn: Axioms for lexicographic preferences. *The Review of Economic Studies*, 42. évf. (1975. July) 3. sz., 415–419. p.
- [133] Peter C. Fishburn – Irving H. LaValle: Subjective expected lexicographic utility: Axioms and assessment. *Annals of Operations Research*, 80. évf. (1998), 183–206. p.

- [134] Frederic B. Fitch: A logical analysis of some value concepts. *The Journal of Symbolic Logic*, 28. évf. (1963) 2. sz., 135–142. p.
- [135] Frederic B. Fitch: A revision of hohfeld's theory of legal concepts. *Logique et Analyse*, 1967., 269–276. p.
- [136] Forrai Gábor: Dennett intencionális hozzáállása. *Replika*, 10. évf. (1999) 36. sz., 59–70. p. URL <http://www.scribd.com/doc/74387228/Dennett+intencionalis+hozzaallasa>.
- [137] K. R. Foster – T. Wenseleers – F. L. W. Ratnieks: Spite: Hamilton's unproven theory. *Ann. Zoo. Fenn.*, 38. évf. (2001), 229–238. p.
- [138] Robert Frank – Thomas Gilovich – Dennis Regan: Does studying economics inhibit cooperation? *Journal of Economic Perspectives*, 7. évf. (1993) 2. sz., 159–171. p.
- [139] 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.
- [140] Harry G. Frankfurt: The problem of action. *American Philosophical Quarterly*, 15. évf. (1978) 2. sz., 157–162. p.
- [141] Harry G. Frankfurt: *The Importance of What We Care About. Philosophical Essays*. 1998, Cambridge University Press.
- [142] Harry G. Frankfurt – Christine M. Korsgaard – Michael E. Bratman – Meir Dan-Cohen: *Taking Rights Seriously. Getting it Right*. 2006, Stanford University Press.
- [143] James W. Friedman: A non-cooperative equilibrium for supergames. *The Review of Economic Studies*, 38. évf. (1971. Jan) 1. sz., 1–12. p.
- [144] Frivaldszky János: *Klasszikus természetjog és jogfilozófia*. 2007, Szent István Társulat.
- [145] Frivaldszky János: Politikai barátság és jogi viszony. *Iustum Aequum Salutare*, V. évf. (2009) 4. sz., 5–52. p.
- [146] 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>.
- [147] Fülöp Endre: Az etikai egoizmus lehetőségéről. *Phronesis*, 1. évf. (2007) 1. sz., 47–54. p.
- [148] Garai Zsolt: Partikularista érvek az értékek realizmusa mellett. *Világosság*, 44. évf. (2003) 5-6. sz., 123–134. p.
- [149] A. Gardner – S.A. West: Spite. *Current Biology*, 16. évf. (2006) 17. sz., R662–R664. p.

- [150] 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.
- [151] N. Georgescu-Roegen: The pure theory of consumer's behavior. *Quarterly Journal of Economics*, 50. évf. (1936) 3. sz., 545–593. p.
- [152] N. Georgescu-Roegen: Choice, expectations, and measurability. *Quarterly Journal of Economics*, 68. évf. (1954) 2. sz., 503–534. p.
- [153] Herbert Gintis: Strong reciprocity and human sociality. *Journal of Theoretical Biology*, 206. évf. (2000) 2. sz., 169–179. p.
- [154] 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.
- [155] Herbert Gintis – Samuel Bowles – Robert Boyd – Ernst Fehr: Explaining altruistic behavior in humans. *Evolution and Human Behavior*, 24. évf. (2003), 153–172. p.
- [156] 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.
- [157] Herbert Gintis – Eric Alden Smith – Samuel Bowles: Costly signaling and cooperation. *J. theor. Biol.*, 213. évf. (2001), 103–119. p.
- [158] Lou Goble – John-Jules Ch. Meyer (szerk.): *Deontic Logic and Artificial Normative Systems*. 2006, Springer.
- [159] Alan H. Goldman: *Reasons from Within. Desires and Values*. 2009, Oxford University Press.
- [160] Alvin I. Goldman: *A Theory of Human Action*. 1970, Prentice-Hall.
- [161] Alvin W. Gouldner: The norm of reciprocity: A preliminary statement. *Americal Sociological Review*, 25. évf. (1960), 161–78. p.
- [162] 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.
- [163] 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.
- [164] T. Grüne-Yanoff – S. O. Hansson: *Preference Change: Approaches from Philosophy, Economics and Psychology*. 2009, Springer.

- [165] 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.
- [166] Gulyás Attila: A méltányosságelmélet alapjai. modellek és nézőpontok. *Közgazdasági Szemle*, LIV. évf. (2007. február), 167–183. p.
- [167] 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.
- [168] W. D. Hamilton: The genetical evolution of social behavior i. *Journal of Theoretical Biology*, 7. évf. (1964), 1–16. p.
- [169] W. D. Hamilton: The genetical evolution of social behavior ii. *Journal of Theoretical Biology*, 7. évf. (1964), 17–52. p.
- [170] W. D. Hamilton: Selfish and spiteful behaviour in an evolutionary model. *Nature*, 288. évf. (1970. 12) 5277. sz., 1218–1220. p.
- [171] Bengt Hansson: Choice structures and preference relations. *Synthese*, 18. évf. (1968) 443–458. sz.
- [172] Bengt Hansson: Fundamental axioms for preference relations. *Synthese*, 18. évf. (1968), 423–442. p.
- [173] Bengt Hansson: d'žžan analysis of some deontic logics. *Noûs*, 3. évf. (1969) 4. sz., 373–398. p.
- [174] Bengt Hansson: Deontic logic and different levels of generality. *Theoria*, 36. évf. (1970. December) 3. sz., 241–248. p.
- [175] Sven Ove Hansson: Changes in preference. *Theory and Decision*, 38. évf. (1995) 1. sz.
- [176] Sven Ove Hansson: Preference logic. In DM. Gabbay – F. Guentner (szerk.): *Handbook of Philosophical Logic*. 4. köt. 2001, Kluwer Academic Publishers, 319–393. p.
- [177] Sven Ove Hansson: *The Structure of Values and Norms*. 2001, Cambridge University Press.
- [178] Garret Hardin: Discriminating Altruism. *Zygon*, 17. évf. (1982) 2. sz., 163–186. p.
- [179] Garrett Hardin: The tragedy of the commons. *Science*, 162. évf. (1968), 1243–1248. p.
- [180] Russel Hardin: Collective action as an agreeable n-prisoners' dilemma. *Behavioral Science*, 16. évf. (1971), 472–481. p.

- [181] John C. Harsanyi: Rational choice models of political behavior vs. functionalist and conformist theories. *World Politics*, 21. évf. (1969), 513–538. p.
- [182] John C. Harsanyi: Advances in understanding rational behavior. In *Essays on Ethics, Social Behavior, and Scientific Explanation*. Dordrecht, 1976, Reidel, 89–117. p.
- [183] H.L.A. Hart: *A jog fogalma*. Budapest, 1995, Osiris.
- [184] Héjjas István: *A keleti vallások filozófiája és világképe*. 2007, ANNO Kiad.
- [185] A. L. Herman: d'žža solution to the paradox of desire in buddhism. *Philosophy East and West*, 29. évf. (1979) 1. sz., 91–94. p.
- [186] Henning Herrestad – Christen Krogh: The right direction.
- [187] Mathieu Hilgers: Habitus, freedom, and reflexivity. *Theory and Psychology*, 19. évf. (2009) 6. sz., 728–755. p.
- [188] Albert O. Hirschmann: *Az érdekek és a szenvedélyek. Politikai érvek a kapitalizmus mellett annak győzelme előtt*. 1998, József Műhely Kiadó.
- [189] Jack Hirshleifer: Privacy: Its origin, function, and future. *The Journal of Legal Studies*, 9. évf. (1980) 4. sz., 649–664. p.
- [190] Jack Hirshleifer: d'žžthe expanding domain of economics. *The American Economic Review*, 75. évf. (1985. Dec) 6. sz., 53–68. p. Centennial Essays and 1985 Survey of Members.
- [191] Wesley Newcomb Hohfeld: Alapvető jogi fogalmak a bírói érvelésben. In Miklós Szabó – Csaba Varga (szerk.): *Jog és nyelv*. 2000, kiadó nélkül, 59–96. p.
- [192] John F. Horty: *Agency and Deontic Logic*. 2001, Oxford University Press.
- [193] G. E. Hughes – M. J. Cresswell: *A New Introduction to Modal Logic*. 1996, Routledge. ISBN 0-415-12599-5.
- [194] 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.
- [195] 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).
- [196] Susan James: *Passion and Action. The Emotions in Seventeenth-Century Philosophy*. 1997, Clarendon Press.
- [197] Janky Béla: *Szolidaritás és jóléti preferenciák*. Philosophiae Doctores sorozat. Budapest, 2005, Akadémiai Kiadó.

- [198] Richard Jeffrey: *The Logic of Decision*. Chicago, 1983, University of Chicago.
- [199] Keith Jensen: Punishment and spite, the dark side of cooperation. *Phil. Trans. R. Soc. B*, 365. évf. (2010), 2635–2650. p.
- [200] Jan C. Joerden: Deontological square, hexagon, and decagon: A deontic framework for supererogation. *Logica Universalis*, 2012. February., 1–16. p.
- [201] R. A. Johnstone – R. Bshary: Evolution of spite through indirect reciprocity. *Proc. R. Soc. Lond. B*, 271. évf. (2004), 1917–1922. p.
- [202] Thomas Joiner: *Why People Die by Suicide*. 2005, Harvard University Press.
- [203] A. J. Jones – M. J. Sergot: A formal characterisation of institutionalised power. *Journal of the IGPL*, 1996. 4. sz., 429–445. p.
- [204] Stig Kanger: New foundations for ethical theory. In R. Hilpinen (szerk.): *Deontic Logic: Introductory and Systematic Readings*. 1971, Reidel, 36–58. p.
- [205] Stig Kanger: Law and logic. *Theoria*, 38. évf. (1972), 105–132. p.
- [206] Stig Kanger – Helle Kanger: Rights and parliamentarism. *Theoria*, 32. évf. (1966), 85–115. p.
- [207] Robert O. Keohane: *After Hegemony. Cooperation and Discord in the World Political Economy*. 1984, Princeton University Press.
- [208] Elias L. Khalil: What is altruism? *Journal of Economic Psychology*, 25. évf. (2004), 97–123. p.
- [209] William Kneale – Martha Kneale: *A logika fejlődése*. Budapest, 1987, Gondolat Kiadó.
- [210] Christine M. Korsgaard: *The Constitution of Agency. Essays on Practical Reason and Moral Psychology*. Christine M. Korsgaard, Oxford University Press.
- [211] Christine M. Korsgaard – G.A. Cohen – Raymond Geuss – Thomas Nagel – Bernard Williams: *The Sources of Normativity*. 1996, Oxford University Press.
- [212] 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.
- [213] Dennis L. Krebs: Altruism and egoism: A false dichotomy? *Psychological Inquiry*, 2. évf. (1991) 2. sz., 137–139. p.
- [214] 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.

- [215] 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.
- [216] 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.
- [217] L. Aczél Petra: Erős szavak. a bizonyítás mint a szöveg tulajdonsága. *Világosság*, 44. évf. (2003) 11-12. sz., 171–178. p.
- [218] David Lane: *Hierarchy in Natural and Social Sciences*. 4. köt. Hierarchy, Complexity, Society fejezet. 2006, Springer, 81–119. p.
- [219] Irving H. Lavalley – Peter C. Fishburn: Lexicographic state-dependent subjective expected utility. *Journal of Risk and Uncertainty*, 4. évf. (1991) 3. sz., 251–269. p.
- [220] Irving H. Lavalley – Peter C. Fishburn: State-independent subjective expected lexicographic utility. *Journal of Risk and Uncertainty*, 5. évf. (1992) 3. sz., 217–240. p.
- [221] Marc Lavoie: Post Keynesian consumer theory: Potential synergies with consumer research and economic psychology. *Journal of Economic Psychology*, 25. évf. (2004), 639–649. p.
- [222] Edward Lemmon: Deontic logic and the logic of imperatives. *Logique et Analyse*, 8. évf. (1965), 39–71. p.
- [223] David K. Levine: Modeling altruism and spitefulness in experiments. *Review of Economic Dynamics*, 1. évf. (1998), 593–622. p.
- [224] David Lewis: Desire as belief. *Mind, New Series*, 97. évf. (1988) 387. sz., 323–332. p.
- [225] David Lewis: Desire as belief ii. *Mind, New Series*, 105. évf. (1996) 418. sz., 303–313. p.
- [226] David K. Lewis: *Convention. A Philosophical Study*. Cambridge (MA), 1969, Harvard University Press.
- [227] Amir N. Licht: Games commissions play: 2x2 games of international securities regulation. *The Yale Law Journal*, 24. évf. (1999) 61. sz., 61–125. p.
- [228] Sarah Lichtenstein – Paul Slovic (szerk.): *The Construction of Preference*. 2006, Cambridge University Press.
- [229] Wim B. G. Liebrand: A classification of social dilemma games. *Simulation & Games*, 14. évf. (1983) 2. sz., 123–138. p.
- [230] 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.

- [231] Lars Lindahl: *Position and Change*. 1977, Springer.
- [232] 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.
- [233] Christian List: The theory of judgment aggregation: An introductory review. *Synthese*, 187. évf. (2012) 1. sz., 179–207. p.
- [234] Christian List: The theory of judgment aggregation: an introductory review. *Synthese*, 187. évf. (2012) 1. sz., 179–207. p.
- [235] Fenrong Liu: Von wright’s “the logic of preference” revisited. *Synthese*, 175. évf. (2010) 1. sz., 69–88. p.
- [236] Fenrong Liu: *Reasoning about Preference Dynamics*. 2011, Springer.
- [237] Alessio Lomuscio – Donald Nute (szerk.): *Deontic Logic in Computer Science*, 2004. 2004, Springer.
- [238] E. J. Lowe: *Personal Agency. The Metaphysics of Mind and Action*. 2008, Oxford University Press.
- [239] John Macnamara – Gonzalo E. Reyes (szerk.): *The logical foundations of cognition*. 1994, Oxford University Press.
- [240] D. Makinson: On the formal representation of rights relations. *Journal of Philosophical Logic*, 15. évf. (1986), 403–425. p.
- [241] David Makinson: On attributing rights to all peoples: Some logical questions. *Law and Philosophy*, 8. évf. (1989) 1. sz., 53–63. p.
- [242] 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.
- [243] 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.
- [244] Andrei Marmor: *Social Conventions. From Language to Law*. 2009, Princeton University Press.
- [245] György Marosán: A proszociális értékek evolúciója a játékelméleti kísérletek tükrében. *Közgazdasági Szemle*, LIV. évf. (2007. augusztus), 716–733. p.
- [246] Marcel Mauss: Techniques of the body. *Economy and Society*, 2. évf. (1973) 1. sz., 70–88. p.
- [247] John Maynard Smith – G. R. Price: The logic of animal conflict. *Nature*, 246. évf. (1973) 5427. sz., 15–18. p.

- [248] Ernst Mayr: Cause and effect in biology. *Science, New Series*, 134. évf. (1961) 3489. sz., 1501–1506. p.
- [249] Ernst Mayr: Teleological and teleonomic, a new analysis. *Boston Studies in the Philosophy of Science*, 14. évf. (1974), 91–117. p.
- [250] Ernst Mayr: The idea of teleology. *Journal of the History of Ideas*, 53. évf. (1992) 1. sz., 117–135. p.
- [251] Charles G. McClintock: Social motivation – a set of propositions. *Behavioral Science*, 17. évf. (1972) 5. sz., 438–454. p.
- [252] Kris McDaniel–Ben Bradley: Desires. *Mind, New Series*, 117. évf. (2008) 466. sz., 267–302. p.
- [253] 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.
- [254] Michael McKinsey: Levels of obligation. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 35. évf. (1979) 4. sz., 385–395. p.
- [255] 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.
- [256] Alfred R. Mele: *Motivation and Agency*. 2003, Oxford University Press.
- [257] David M. Messick–Charles G. McClintock: Motivational bases of choice in experimental games. *Journal of Experimental Social Psychology*, 4. évf. (1968), 1–2. p.
- [258] 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>.
- [259] E. Midlarsky–E. Kahana: Theories and concepts of altruism and helping. In E. Midlarsky–E. Kahana (szerk.): *Altruism in Later Life*. Library of Social Research sorozat, 8. köt. 1994, Sage, 11–40. p.
- [260] Greg Miller: The roots of morality. *Science*, 320. évf. (2008. May), 734–737. p.
- [261] Seumas Miller: *Social Action. A Teleological Account*. 2004, Cambridge University Press.
- [262] Ruth Garret Millikan: Bioszemantika. In Gergely Ambrus–Tamás Deme-ter–Gábor Forrai–János Tőzsér (szerk.): *Elmefilozófia*. 2008, L'Harmattan.
- [263] Ruth Garrett Millikan: Thoughts without laws; cognitive science with content. *The Philosophical Review*, 95. évf. (1986) 1. sz., 47–80. p.

- [264] 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.
- [265] 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.
- [266] Alessio Moretti: *The Geometry of Logical Opposition*. PhD értekezés (University of Neuchâtel). 2009.
- [267] Alessio Moretti: The geometry of standard deontic logic. *Logica Universalis*, 3. évf. (2009), 19–57. p.
- [268] 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.
- [269] Alessio Moretti: Why the logical hexagon? *Logica Universalis*, 6. évf. (2012), 69–107. p.
- [270] Pierre Le Morvan: Selfishness, altruism, and our future selves. *d'žžAustralasian Journal of Philosophy*, 87. évf. (2009) 3. sz., 409–424. p.
- [271] Ernest Nagel: Functional explanations in biology. *The Journal of Philosophy*, 74. évf. (1977) 5. sz., 280–301. p.
- [272] Ernest Nagel: Goal-directed processes in biology. *The Journal of Philosophy*, 74. évf. (1977) 5. sz., 261–279. p.
- [273] Klaus Nehring: Self-control through second-order preferences, 2006.
- [274] Martin A. Nowak – Corina E. Tarnita – Edward O. Wilson: The evolution of eusociality. *Nature*, 466. évf. (2010) 7310. sz., 1057–1062. p.
- [275] Timothy O'Connor – Constantine Sandis: *A Companion to the Philosophy of Action*. 2010, Wiley-Blackwell.
- [276] Óészaki mitológiai és hősi énekek: 2. a fenséges beszéde, 42–43. versszak, 2003. January. URL <http://mek.oszk.hu/00300/00377/>.
- [277] Andrew Oldenquist: The possibility of selfishness. *American Philosophical Quarterly*, 17. évf. (1980) 1. sz., 25–33. p.
- [278] Pamela Oliver: Rewards and Punishments as Selective Incentives for Collective Action: Theoretical Investigations. *The American Journal of Sociology*, 85. évf. (1980. May) 6. sz., 1356–1375. p.
- [279] Orthmayr Imre: A sztoikus szenvtelenség. *Századvég*, 4. évf. (1999) 13. sz. URL <http://www.c3.hu/scripta/szazadveg/13/orthmayr.htm>.
- [280] 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>.

- [281] 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.
- [282] 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.
- [283] Terence Parsons: The traditional square of opposition. in stanford encyclopedia of philosophy. URL <http://plato.stanford.edu/entries/square/>.
- [284] Sam Perlo-Freeman: The topology of conflict and co-operation, 2006. May.
- [285] Michael Perloff: Stit and the language of agency. *Synthese*, 86. évf. (1991) 3. sz., 379–408. p.
- [286] Pete Krisztián: A társadalmi konvenciók természetéről. *Infotárs*, 2008., 73–99. p.
- [287] Fabienne Peter – Hans Bernhard Schmid: *Rationality and Commitment*. 2007, Oxford University Press.
- [288] R. W. Picard: Affective computing. No. 321. Jelentés, 1995, M.I.T Media Laboratory Perceptual Computing Section Technical Report.
- [289] Charles R. Pigden: Desiring to desire: Russell, moore and david lewis.
- [290] Jane Allyn Piliavin – Hong-Wen Charng: Altruism: A review of recent theory and research. *Annual Review of Sociology*, 16. évf. (1990), 27–65. p.
- [291] John Platt: Social traps. *American Psychologist*, 28. évf. (1973) 8. sz., 641–651. p.
- [292] Pokol Béla: *Modern francia szociológiaelméletek*. Miskolc, 1996, Bíbor Kiadó. URL <http://mek.oszk.hu/02000/02027/02027.pdf>.
- [293] Ingmar Pörn: *The Logic of Power*. 1970, Blackwells.
- [294] Ingmar Pörn: *Action theory and social science: some formal models*. 1977, D. Reidel.
- [295] William Poundstone: *Prisoner's Dilemma*. 1992, Anchor.
- [296] Matthew Rabin: Incorporating fairness into game theory and economics. *The American Economic Review*, 83. évf. (1993. Dec) 5. sz., 1281–1302. p.
- [297] N. J. Raihani – K. McAuliffe: Human punishment is motivated by inequity aversion, not a desire for reciprocity, 2012.
- [298] 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.
- [299] Anatol Rapoport: Three modes of conflict. *Management Science*, 7. évf. (1961) 3. sz., 210–218. p.

- [300] Anatol Rapoport: Optimal policies for the prisoner's dilemma. *Psychological Review*, 74. évf. (1967) 2. sz., 136–148. p.
- [301] 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.
- [302] Hyne W. Reese: Teleology and teleonomy in behavior analysis. *The Behavior Analyst*, 17. évf. (1994) 1. sz., 75–91. p.
- [303] Nicholas Rescher (szerk.): *The Logic of Decision and Action*. 1966, University of Pittsburgh Press.
- [304] David Robinson – David Goforth: *The Topology of the 2x2 Games*. 2005, Routledge.
- [305] Alf Ross: Imperatives and logic. *Philosophy of Science*, 11. évf. (1944) 1. sz., 30–46. p.
- [306] David Ross: *The Right and the Good*. Oxford, 2002 (1930), Clarendon Press.
- [307] 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.
- [308] 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.
- [309] Stuart J. Russel – Devika Subramanian: Provably bounded-optimal agents. *Journal of Artificial Intelligence Research*, 2. évf. (1995), 575–609. p.
- [310] Bertrand Russell: A denotálásról. *Világosság*, 46. évf. (2005) 12. sz., 5–16. p.
- [311] Ruzsa Imre: *Bevezetés a modern logikába*. 2000, Osiris Kiadó.
- [312] J. Sachs: Cooperation within and among species. *J. Evol. Biol.*, 19. évf. (2006) 5. sz., 1415–1418. p.
- [313] Larry Samuelson: Introduction to the evolution of preferences. *Journal of Economic Theory*, 97. évf. (2001), 225–230. p.
- [314] 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.
- [315] David H. Sanford: Contraries and subcontraries. *Noûs*, 2. évf. (1968) 1. sz., 95–96. p.
- [316] Giovanni Sartor: Fundamental legal concepts: A teleological characterisation, 2005.

- [317] Giovanni Sartor: Fundamental legal concepts: A formal and teleological characterisation. Jelentés, 2006, European University Institute Department of Law.
- [318] Stephen Schiffer: A paradox of desire. *American Philosophical Quarterly*, 13. évf. (1976) 3. sz., 195–203. p.
- [319] Arthur Schopenhauer: *Az akarat szabadságáról*. 1913, Franklin Társulat.
- [320] Timothy Schroeder: *Three Faces of Desire*. New York, 2004, Oxford University Press.
- [321] Timothy Schroeder: Reflection, reason, and free will. *Philosophical Explorations*, 10. évf. (2007) 1. sz., 77–84. p.
- [322] John R. Searle: *Intentionality*. 1983, Cambridge University Press.
- [323] John R. Searle: *Elme, nyelv, társadalom. A való világ filozófiája*. Budapest, 2000, Vince Kiadó.
- [324] Krister Segerberg: Validity and satisfaction in imperative logic. *Notre Dame Journal of Formal Logic*, 31. évf. (1990), 203–221. p.
- [325] 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.
- [326] Reinhard Selten: The chain store paradox. *Theory and Decision*, 9. évf. (1978) 2. sz., 127–159. p.
- [327] 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.
- [328] Seregi Tamás: Paul ricoeur cselekvésontológiája. *Világosság*, 48. évf. (2007) 1. sz., 43–64. p.
- [329] Marek J. Sergot: Normative positions. In P. McNamara–H. Prakken (szerk.): *Norms, Logics and Information Systems*. 1999, IOS, 289–308. p.
- [330] Marek J. Sergot: A computational theory of normative positions. *ACM Transactions on Computational Logic*, 2. évf. (2001) 4. sz., 581–622. p.
- [331] 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.
- [332] R. Sethi–E. Somanathan: Preference evolution and reciprocity. *Journal of Economic Theory*, 97. évf. (2001) 2. sz., 273–297. p.
- [333] H.A. Simon: Altruizmus és közgazdaságtan. *Magyar Tudomány*, 40 (102). évf. (1995) 1. sz., 22–29. p.
- [334] Herbert Simon: A mechanism for social selection and successful altruism. *Science*, 250. évf. (1990), 1665–1668. p.

- [335] Herbert A. Simon: d'zztheories of decision-making in economics and behavioral science. *The American Economic Review*, 49. évf. (1959) 3. sz., 253–283. p.
- [336] András Simonovits: Bevezetés a játékelméletbe: vázlat, 2007.
- [337] Peter Simons: *Parts: A Study in Ontology*. 2000 (1987), Oxford University Press.
- [338] Brian Skyrms: *Evolution of the Social Contracts*. New York, 1996, Cambridge University Press.
- [339] Brian Skyrms: *The Stag Hunt and the Evolution of Social Structure*. 2004, Cambridge University Press.
- [340] Brian Skyrms: *Evolution and the social contract*, 2007.
- [341] Barry Hartley Slater: Logical paradoxes. *Internet Encyclopedia of Philosophy*, 2005.
- [342] 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>.
- [343] Hans Smessaert: On the 3d visualisation of logical relations. *Logica Universalis*, 3. évf. (2009), 303–332. p.
- [344] John Maynard Smith: *Evolution and the Theory of Games*. 1982, Cambridge University Press.
- [345] 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.
- [346] Elliott Sober – David Sloan Wilson: *Unto Others: The Evolution and Psychology of Unselfish Behavior*. 1998, Harvard University Press.
- [347] Steiger Kornél: Megjegyzések orthmayr imre “a sztoikus szenvtelenség” c. tanulmányához. *Századvég*, 1999.
URL <http://www.c3.hu/scripta/szazadveg/13/steiger.htm>.
- [348] 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.
- [349] Scott Sturgeon: Normative judgement. *Noûs, Supplement: Philosophical Perspectives*, 41. évf. (2007) 21. sz., 569–587. p.
- [350] Robert Sugden: Reciprocity: The supply of public goods through voluntary contributions. *Economic Journal*, 94. évf. (1984), 772–787. p.
- [351] Christine Swanton: Weakness of will as a species of executive cowardice. *Canadian Journal of Philosophy*, 21. évf. (1991) 2. sz., 123–140. p.

- [352] Syi: Akarat-erő-logika. In Katalin S. Nagy – Annamária Orbán (szerk.): *Értékek és normák interdiszciplináris megközelítésben*. Budapest, 2008, Gondolat Kiadó, 50–68. p.
- [353] Syi: *Cselekvésemélet dióhéjban*. 2008, Typotex.
- [354] Syi: Önzetlenségelméletek. *BUKSZ*, 21. évf. (2009) 4. sz., 226–233. p.
- [355] Syi: Max weber cselekvéseméleti ontológiája. *Szociológiai Szemle*, 21. évf. (2011) 2. sz., 87–110. p.
- [356] 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.
- [357] Szakadát István: Hierarchia. In Helmich Dezső – Szántó Zoltán (szerk.): *Metodológia, társadalom, gazdaság. Im memoriam Bertalan László*. 2004, Közgazdasági Szemle Alapítvány, 139–166. p.
- [358] Szakadát István: Reláció, szintaktika, szemantika. *Tudományos és Műszaki Tájékoztatás*, 51. évf. (2004) 12. sz., 531–540. p.
- [359] Szakadát István: Réteges struktúra, alaprelációk. In Csendes Lilla – Alexin Zoltán (szerk.): *III. Számológépes nyelvészeti konferencia* (konferenciaanyag). Szeged, 2005, SZTE.
- [360] Szécsi Gábor: Okság és intencionalitás. *Magyar Filozófiai Szemle*, 39. évf. (1995) 3-4. sz., 529–561. p.
- [361] Szívós Mihály: Az elsődleges és másodlagos habitus szociológiai és tudománysszociológiai fogalma a személyes és a hallgatólagos tudás elméletének szemszögéből. *Polanyiana*, 18. évf. (2009) 1-2. sz., 29–52. p.
- [362] Szolcsányi Tibor: A davidsoni háromszögelés információ-elméleti elemzése. In *Igazság és kommunikáció. Tanulmányok Donald Davidson filozófiájáról*. 2004, Brambauer Kiadó, 145–166. p.
- [363] Takács Károly: Hálózati kísérletek. *Közgazdasági Szemle*, LVII. évf. (2010), 958–979. p.
- [364] Charles Taylor: *Sources of the Self. The Making of the Modern Identity*. 1989, Harvard University Press.
- [365] 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.
- [366] Rafael Di Tella – Ricardo Pérez-Truglia: Conveniently upset: Avoiding altruism by distorting beliefs about others. Jelentés, 2010, National Bureau of Economic Research.

- [367] Temesi József: *A döntéelmélet alapjai*. 2002, Aula Kiadó.
- [368] Tengelyi László: Tapasztalat, cselekvés és elbeszélte történet. *Világosság*, 50. évf. (2009) tavasz. sz., 87–99. p.
- [369] Warren D. TenHouten: *A General Theory of Emotions and Social Life*. 2007, Routledge.
- [370] Michael Thompson: *Life and Action. Elementary Structures of Practice and Practical Thought*. 2008, Harvard University Press.
- [371] Giulio Tononi: An information integration theory of consciousness. *BMC Neuroscience*, 5. évf. (2004), 42–66. p.
- [372] Robert L. Trivers: A reciprok altruizmus evolúciója. In Zsuzsanna Kulcsár (szerk.): *Morális fejlődés, empátia és altruizmus*. 1999, ELTE Eötvös Kiadó, 245–272. p.
- [373] Sara L. Uckelman: Anselm's logic of agency, 2012. February.
- [374] Edna Ullmann-Margalit: *The Emergence of Norms*. 1977, Clarendon Press.
- [375] Edna Ullmann-Margalit: Family fairness. *Social Research*, 73. évf. (2006), 575–596. p.
- [376] Edna Ullmann-Margalit: Considerateness. *Iyyun, The Jerusalem Philosophical Quarterly*, 60. évf. (2011. July), 205–244. p.
- [377] J. Urmson: Saints and heroes. In A.J. Melden (szerk.): *Essays in Moral Philosophy*. Seattle, 1958, University of Washington Press, 198–216. p.
- [378] Piercarlo Valdesolo – David DeSteno: Manipulations of emotional context shape moral judgment. *Psychological Science*, 17. évf. (2006) 6. sz., 476–477. p.
- [379] 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), 83–125. p.
- [380] Johan van Benthem – Amitabha Gupta – Eric Pacuit (szerk.): *Games, Norms and Reasons. Logic at the Crossroads*. 2011, Springer.
- [381] Ron van der Meyden – Leendert van der Torre: *Deontic Logic in Computer Science*, 2008. 2008, Springer.
- [382] 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.
- [383] 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. Higgins (szerk.): *Social Psychology: Handbook of Basic Principles*. 2nd edition. kiad. 2007, Guilford, 540–561. p.

- [384] 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.
- [385] Daniel Vanderveken (szerk.): *Logic, Thought and Action*. 2005, Springer.
- [386] A.C. Varzi: Parts, wholes, and part-whole relations: The prospects of mereotopology. *Data and Knowledge Engineering*, 20. évf. (1996), 177–198. p.
- [387] Bruce Vermazen: Negative acts. In Bruce Vermazen – Merril B. Hintikka (szerk.): *Essays on Davidson. Actions and Events*. 1985, Oxford University Press, 93–104. p.
- [388] Rudolf von Jhering: Küzdelem a jogért. In Csaba Varga (szerk.): *Jog és jogfilozófia*. 2001, Szent István Társulat, 1–25. p.
- [389] Georg Henrik von Wright: *An Essay in Deontic Logic*. 1951, North-Holland Publishing Company.
- [390] Georg Henrik von Wright: *Norm and Action*. London, 1963, Routledge and Kegan Paul.
- [391] Georg Henrik von Wright: *The Logic of Preference*. Edinburgh, 1963, Edinburgh University Press.
- [392] Georg Henrik von Wright: The logic of preference reconsidered. *Theory and Decision*, 3. évf. (1972) 3. sz., 140–169. p.
- [393] 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.
- [394] Peter B.M. Vranas: New foundations for imperative logic ii: Pure imperative inference.
- [395] Peter B.M. Vranas: New foundations for imperative logic iii: A general definition of argument validity.
- [396] Peter B.M. Vranas: New foundations for deontic logic: A preliminary sketch, 2002.
- [397] Peter B.M. Vranas: New foundations for imperative logic i: Logical connectives, consistency, and quantifiers. *Noûs*, 42. évf. (2008), 529– 572. p.
- [398] Peter B.M. Vranas: Imperatives, logic of. In *The International Encyclopedia of Ethics*. 2013, John Wiley and Sons.
- [399] Loïc J.D. Wacquant: *Body and Soul. Notebooks of and Apprentice Boxer*. 2004, Oxford University Press.
- [400] Douglas N. Walton: Logical form and agency. *Philosophical Studies*, 29. évf. (1976), 75–89. p.

- [401] Douglas N. Walton: St. anselm and the logical syntax of agency. *Franciscan Studies*, 36. évf. (1976), 298–312. p.
- [402] Douglas N. Walton: On the logical form of some commomplace action expression. *Grazer Philosophische Studien*, 10. évf. (1980), 141–148. p.
- [403] Gary Watson: *Agency and Answerability*. 2004, Oxford University Press.
- [404] Max Weber: *Gazdaság és társadalom I. A megértő szociológia alapvonalai*. 1987, KJK.
- [405] 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.
- [406] S. A. West–A. S. Griffin–A. Gardner: Evolutionary explanations for cooperation. *Current Biology*, 17. évf. (2007), R661–R672. p.
- [407] S.A. West–A. Gardner: Altruism, spite and greenbeards. *Science*, 327. évf. (2010), 1341–1344. p.
- [408] Edward O. Wilson–Bert Hölldobler: Eusociality: origin and consequences. *Proc. Natl. Acad. Sci. U.S.A.*, 102. évf. (2005) 38. sz., 13367–13371. p.
- [409] Edward O. Wilson–David Sloan Wilson: Rethinking the theoretical foundation of sociobiology. *Q Rev Biol*, 82. évf. (2007) 4. sz., 327–348. p.
- [410] Edward.O. Wilson: *Sociobiology: the new synthesis*. 1975, Harvard University Press.
- [411] 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.
- [412] Larry Wright: Functions. *The Philosophical Review*, 82. évf. (1973) 2. sz., 139–168. p.
- [413] A. Zahavi: Mate selection: Selection for a handicap. *Journal of Theoretical Biology*, 53. évf. (1975) 1. sz., 205–214. p.
- [414] 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.
- [415] 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.
- [416] Amotz Zahavi–Avishag Zahavi: *The Handicap Principle. A Missing Piece of Darwin’s Puzzle*. 1997, Oxford University Press.
- [417] R.B. Zajonc: Érzelem és gondolkodás (a preferenciák nem igényelnek következtetéseket). In György Hunyady (szerk.): *Szociálpszichológiai tanulmányok*. Budapest, 1991, Tankönyvkiadó, 7–55. p.