A filozófia alapkérdései természettudományos aspektusból

E. Szabó László

MTA-ELTE Elméleti Fizika Kutatócsoport

ELTE Tudománytörténet és Tudományfilozófia Tanszék

http://philosophy.elte.hu/leszabo

leszabo@philosophy.elte.hu

Tartalomjegyzék

1. előadás

Az idő

(Mi a filozófia?)

1.0.0.1. A hétköznapi gondolkodás szerint a létezés szoros összefüggésben áll az idővel. Gondoljuk csak el, mennyire természetesnek találjuk a következő két gondolatot:

> Minden, ami létezik, a jelenben létezik. A múltbeli dolgok már nem léteznek, a jövőbeli dolgok még nem léteznek.
>
> Minden, ami létezik, időben létezik. Az idő múlásával egyszer csak nem létező dolgok létezővé válnak, majd, az idő múlásával, nem létezővé lesznek.

A hétköznapi szemlélet számára tehát, a múltbeli illetve jövőbeli események *ontológiai státusza* különbözik a jelenbeli események ontológiai státuszától.

A filozófiai gondolkodás számára azonban felmerül a kérdés: valóban így van-e. Önmagában az a tény, hogy a fizikai események az időre vonatkozó szubjektív élményünk szerint *bekövetkeznek*, még nem garantálja, hogy ennek a „bekövetkezésnek” bármiféle, tudatunktól független státusza legyen. Igaz-e tehát, hogy a jövőbeli dolgok, események bekövetkezése, „létrejövetele” a dolgok egy, a fizika szintjén is megnyilvánuló, és tudatunktól független tulajdonsága, mint ahogyan azt a köznapi gondolkodás feltételezi? És képesek-e a fizikai elméletek alapot szolgáltatni e kérdés megválaszolásához? – mint ahogyan azt Reichenbach feltételezi:

> Más lehetőség nincs, mint hogy a fizika révén oldjuk meg az idő problémáját. A fizika minden más tudománynál többet törődik az idő természetével. Ha az idő objektív, a fizikusnak e tényt fel kell fedeznie, ha van Bekövetkezés (Becoming), a fizikusnak tudnia kell erről; ám, ha az idő csupán szubjektív és a Létezés időtlen (timeless), akkor a fizikusnak képesnek kell lennie arra, hogy az időt

ignorálja mindabból, ahogyan a valóságot megkonstruálja, hogy a világot az idő segítsége nélkül írja le.[1]

1.0.0.2. Az időről folytatott diskurzusnak alapvetően két típusát szokás megkülönböztetni a filozófiai irodalomban. E két típus megkülönböztetése az idővel kapcsolatos hétköznapi nyelvhasználatban is megjelenő különbségen alapszik: Ha eseményeket időbeliségük alapján jellemezni akarunk, akkor olyan szavakat szoktunk használni, mint *előbb–később*, vagy *jövőbeli–jelenbeli–múltbeli*. Az események e két lehetséges temporális jellemzése között az az alapvető különbség, hogy az egyik „időtlen”, míg a másik nem. Ha egy X esemény *előbbi*, mint Y, akkor ez a viszony minden időpillanatban változatlanul fennáll. Ezzel szemben a *jövő–jelen–múlt*-típusú temporalitás azt tételezi, hogy az eseményeknek van egy tulajdonsága, amely változik: jövőbeli események az idő múlásával jelenbelivé válnak, majd múltbeliek lesznek. Más szóval, hogy az idő „folyik”, hogy az események „bekövetkeznek”, hogy az idő múlása az események ontológiai státuszában valamiféle változást eredményez. A McTaggart-tól származó[2] terminológiában, események időbeli sorozatát, az első esetben *B-sorozatnak*, a második esetben *A-sorozatnak* nevezzük. E tradicionálisan használt, és nem túl szerencsés terminológia, valójában az idő lényegének két különböző megközelítését tükrözi.

1.0.0.3. Érdemes röviden megismerkednünk azzal, ahogyan McTaggart eredetileg használta ezeket a fogalmakat, továbbá azzal az argumentummal, mellyel azt kívánta bebizonyítani, hogy az időnek nincs semmiféle realitása. McTaggart érvelése két lépésből áll: az első lépésben azt mutatja meg, hogy az idő fogalma csak az A-sorozat segítségével ragadható meg. Vagyis, hogy A-sorozat nélkül nincs idő. Második lépésben pedig azt, hogy A-sorozat a valóságban nem létezik. Abból – a senki által nem vitatott megállapításból indul ki, mely

[1] Reichenbach 1956, 16. o.

[2] McTaggart 1908, 1993.

szerint az idő elválaszthatatlan a változás fogalmától. Nem lehetne időről beszélnünk akkor, ha semmi nem változna a világban. McTaggart szerint azonban csak egyetlen esetben lehet változásról beszélni, nevezetesen, ha a világ eseményeinek van valamilyen tulajdonsága, amely megváltozik:

> Nem születik változás abból, ha egy esemény megszűnik esemény lenni, abból sem, ha egyik eseményt egy másik váltja fel. Milyen más módja van tehát a változásnak? Ha egy esemény jellemzői megváltoznak, akkor bizonyosan van változás! De milyen jellemzői változhatnak meg egy eseménynek? Úgy gondolom, a jellemzőknek egyetlen ilyen osztálya van csupán. És ez az osztály a szóban forgó esemény azon meghatározottságaiból áll, melyeket az A-sorozat terminusaival ragadhatunk meg. Vegyünk bármilyen eseményt – Anna Királynő halálát, például – és vizsgáljuk meg, milyen változások történhetnek ennek az eseménynek a jellemzőiben. Az, hogy ez egy halál, hogy ez Stuart Anna halála, hogy ilyen és ilyen okai voltak, hogy ilyen és ilyen hatása volt – ezek a jellemzők mind olyanok, melyek sohasem változnak meg. ... Egyetlen vonatkozásban történik változás: Kezdetben egy olyan esemény volt, amely a távoli jövőben van, majd egyre közelebbi esemény lett. Egyszer csak jelenbelivé vált. Aztán múltbelivé, s mindörökké az is marad, minden pillanattal az egyre távolabbi múltba kerül. Ez az egyetlen jellemzője az eseménynek, amely megváltozhat. És ennélfogva, ha van változás, az csakis az A-sorozatban kereshető. Ha nincs A-sorozat, akkor változás sincs.[3]

1.0.0.4. Az argumentáció második részében McTaggart azt bizonyítja be, hogy A-sorozatok nincsenek. McTaggart bizonyítása a követ-

[3] Uo. 26. o.

kező:

> A múlt, a jelen és a jövő, inkompatibilis meghatározottságok. Minden eseményre igaz valamelyik, de csak az egyik. ... E tulajdonságok tehát inkompatibilisek. Ugyanakkor, mindegyik esemény rendelkezik mindhárom tulajdonsággal. Ha egy *M* esemény jövőbeli, akkor jelenbeli és múltbeli lesz. Ha jelenbeli, akkor jövőbeli volt, és múltbeli lesz. Vagyis minden eseményhez mindhárom karakterisztika hozzátartozik. Hogy egyeztethető ez össze azzal, hogy inkompatibilisek? Úgy tűnhet, ezt könnyen megmagyarázhatjuk. ... Az, hogy *M* jelenbeli, múltbeli lesz, jövőbeli volt, azt jelenti, hogy jelenbeli a jelen pillanatban, hogy múltbeli valamilyen jövőbeli pillanatban, és, hogy jövőbeli valamilyen múltbeli pillanatban. De minden pillanat, éppúgy mint minden esemény, egyaránt múltbeli, jelenbeli és jövőbeli. Így tehát hasonló problémával találjuk magunkat szemben. Ha *M* jelenbeli, akkor nincs olyan múltbeli pillanat, amikor múltbeli. De a jövőbeli pillanatok, melyekben *M* múltbeli, éppannyira múltbeliek is, melyekben viszont *M* nem lehet múltbeli. ... Tehát megint ellentmondásra jutunk, hiszen azok a pillanatok, melyekben *M* az A-sorozatbeli valamelyik meghatározottsággal bír, egyben olyan pillanatok is, melyekben nem rendelkezhet a szóban forgó meghatározottsággal. Ha most úgy próbáljuk feloldani az ellentmondást, hogy ugyanazt mondjuk ezekről a pillanatokról, amit az előbb mondtunk magáról *M*-ről, hogy tudniillik egy bizonyos pillanat, például jövőbeli, jelenbeli, valamint múltbeli lesz, akkor az így használt igeidőknek ugyanazt a jelentést tulajdonítjuk, mint az előbb. Vagyis, hogy a szóban forgó pillanat jövőbeli valamilyen jelenbeli pillanatban, és jelenbeli illetve múltbeli lesz valamilyen jövőbeli

pillanatokban. Ezzel persze ugyanaz a probléma kezdődik elölről, s ez folytatódik a végtelenségig.[4]

1.0.0.5. Nem véletlen, hogy a bizonyítást szó szerint idéztem. Meggyőződésem ugyanis, hogy azoknak van igazuk, akik szerint a fenti bizonyítás ebben a formában nem elfogadható. Ezzel nem azt kívánom állítani, hogy A-sorozatok léteznek, hanem csupán azt, hogy létezésüket a fenti gondolatmenettel nem zárhatjuk ki. Ugyanis

1. Ha a *múlt*, a *jelen*, és a *jövő* kategóriákat hallgatólagosan mindig egy meghatározott időpillanatra, pontosabban a sorozat egy eseményére vonatkoztatva értelmezzük, akkor ez azt jelenti, hogy e kategóriáknak van egy „elhallgatott” indexe. Helyesen azt kellene írnunk, hogy $\textit{múlt}_X$, $\textit{jelen}_X$ és $\textit{jövő}_X$, ahol X a sorozat egy tetszőleges eseménye. Az viszont, hogy egy M esemény, például, $\textit{múlt}_X$-beli, M-nek egy olyan tulajdonsága, amely nem változik. Ebben az esetben tehát a sorozat nem is A-sorozat. Semmiféle ellentmondás nincs.

2. Tegyük most fel, hogy a sorozat tényleg A-sorozat, és a „múltbeli eseménynek lenni”, „jelenbeli eseménynek lenni” és „jövőbeli eseménynek lenni” – így, indextelenül – az eseményeknek olyan tulajdonságai, amelyek időben változnak. Természetesen egyetérthetünk McTaggarttal, hogy kizárólag ezek a tulajdonságok lesznek az események olyan tulajdonságai, amelyek változnak. Nem abban az értelemben változnak, hogy relatívak lennének egy időpillanatra vonatkoztatva, mint az előző esetben, hanem ahogyan bárminek bármilyen tulajdonsága időben változik, ha változik. Az alma előbb *éretlen*, aztán *ehető*, majd végül *rothadt*. E három kategória az alma időben változó tulajdonságait tükrözi. Hogy van-e A-sorozat, azaz, hogy az eseményeknek van-e ilyen időben változó tulajdonsága, hogy a múlt-,

[4] McTaggart 1993, 32-33. o.

> jelen- és jövőbeliség köthető-e az ontológiai státuszuk változásának fázisaihoz, hogy a jövőbeli események „megérnek-e” mint az alma, és „bekövetkeznek-e”, ezek mind súlyos filozófiai kérdések, és egyáltalán nem zárom ki, hogy a válasz ezekre a kérdésekre nemleges. De pusztán az a tény, hogy az alma az idő múlásával átmegy az *éretlenség*, az *ehetőség* és a *rothadtság* fázisain, nem fordítható le arra a logikai formulára, mely szerint az alma *éretlen* & *ehető* & *rothadt*. Tehát semmiféle ilyen logikai ellentmondás az A-sorozat létezéséből nem vezethető le.

1.0.0.6. Meggyőződésem szerint van még egy pontja az érvelésnek, ahol nem kell McTaggarttal egyetértenünk. Nevezetesen, hogy nincsen változás A-sorozat nélkül. Igaz, hogy „ha egy esemény jellemzői megváltoznak, akkor bizonyosan van változás”, de ebből nem következik, hogy a változás fogalmát kizárólag az események jellemzőinek megváltozása révén lehet filozófiailag megragadni. A B-teoretikusok szerint az idő lényege, hogy a világ állapotai egy „megelőzi” relációval lineárisan rendezhetők, és a változás egyszerűen annyit jelent, hogy a dolgok állása e sorozat egyik, t-vel indexelt helyén különbözik a dolgok állásától egy másik, t' helyen. Mint Russell írja:

> A változás nem más, mint az igazság és hamisság tekintetében fennálló különbség két kijelentés között, melyek egyike egy bizonyos entitásra és egy T időpillanatra, a másik pedig ugyanarra az entitásra és egy T' időpillanatra vonatkozik, feltéve, hogy a két állítás csupán abban tér el, hogy az egyikben T, a másikban T' szerepel.[5]

Vagyis, továbbra is Russell egyik példájánál maradva, változás történik, ha <A piszkavas a T időpontban forró> állítás *igaz*, és <A piszkavas a T' időpontban forró> állítás *hamis*. Russellel polemizálva McTaggart a következőt írja:

[5] McTaggart 1993, 27. o.

> Vegyünk egy másik sorozatot. A greenwich-i meridián szélességi fokok egész sorozatát szeli át. És találhatunk két olyan S és S' pontot ebben a sorozatban, melyekre fennáll, hogy az <*S* pontban a greenwich-i meridián az Egyesült Királyság területére esik> kijelentés igaz, míg az <*S′* pontban a greenwich-i meridián az Egyesült Királyság területére esik> kijelentés hamis. De senkinek se jutna eszébe azt mondani, hogy ez valamiféle változást jelent. Akkor miért mondanánk ezt egy másik sorozat esetében?[6]

Tovább gondolva McTaggart példáját, tegyük fel, hogy van egy műszerünk, amely valamilyen τ mennyiség éppen aktuális értékét mutatja. Készítsünk a világ állapotairól felvételeket úgy, hogy a kép sarkában ott látható a műszer mutatójának állása. Tegyük fel, hogy τ paraméter értékei lineárisan rendezhetők, és ennek megfelelően ezeket a „fényképeket” egy sorozatba rendezzük. Vagyis a világ állapotainak egy sorozatát – ha tetszik, események egy sorozatát – kaptuk. Tükröz-e e sorozat bármiféle változást a világban? McTaggart fenti példáját alapul véve, azt kell mondanunk, hogy általában nem. Kivéve, ha a szóban forgó műszer egy óra! De mi tesz egy mérőműszert órává? S mi teszi az általa mért τ paramétert idővé? Elég-e egyszerűen a hasunkra ütnünk és a világról készült felvételeken ezt és ezt a részletet kiragadva, ezt mondanunk: mostantól, azt az „óramutató állásának” nevezzük, és az így definiált paramétert „időnek”, a felvételek e szerinti rendezését pedig „időrendezésnek”, „idősorozatnak”? Más megfogalmazásban:

> Az A- és B-időskálák közötti elsőbbség, s ezáltal az idő alapvető természete, egyetlen kérdésre adott helyes válaszon múlik, nevezetesen, hogy mi tesz egy B-pillanatot jelenné? [7]

[6] Uo. 28. o.

[7] Mellor 1998, 11. o.

Például, a relativitáselméletben egyszerűen úgy döntünk, hogy a deformált órákkal mért mennyiséget időnek tekintjük, és a világnak így is egy konzisztens leírásához jutunk. Ez persze még nem olyan radikális lépés, a deformált órák nem fogják átrendezni az idősorozatokat, hiszen a deformált óra által mért mennyiség szigorúan monoton növekvő függvénye a deformálatlan óra által mért időnek. Megengedhető-e radikálisabb módosítása az idő fogalmának? Jogos-e bármiféle szabad választás, filozófiai szempontból? S ha igen, hol van ennek a szabadságnak a határa? Ezekre a kérdésekre különböző válaszokat szokás adni a filozófiában. A korábbi fejezetekben általam szorgalmazott konvencionalista felfogás szerint a rendezés alapjául szolgáló lehetséges paraméterek közül nincs egy kitüntetett, amelyik ténylegesen tükrözné a világ változásait, míg a többi nem. Az, hogy a hétköznapi gondolkodásunk szerint, valamint pszichikailag, ennek ellenkezőjéről vagyunk meggyőződve, csupán azzal magyarázható, hogy létezik az időrendezésnek egy „kitüntetett” módja, amely az evolúció során mélyen belénk ivódott. Nevezetesen az, amely a nappalok és éjjelek váltakozására és más csillagászati/meteorológiai jelenségekre épül. Az a tény, hogy a világ fizikai leírásában is, *konvencionálisan*, ezt a természetesnek érzett „csillagászati időt” használjuk, csak fokozta azt a meggyőződésünket, hogy a világ változásainak ez az „igazi” leírása, s hogy az így kialakult időfogalmunk valamiféle metafizikai kitüntetettséget élvez. Más filozófiai irányzatok szerint a kiválasztás nem csupán konvención alapszik, s nem is csupán biológiailag, vagy *a priori* (mint Kantnál) adott. Egyesek szerint a kauzalitás alapvetőbb fogalom, mint az idő, és az időrendezés módját a kauzális rendezés választja ki.[8] Vannak, akik szerint az időrendezésnek a termodinamikai entrópia fogalmára kell épülnie, míg pl. Swinburne szerint az idő struktúrája, és ezáltal a helyes időrendezés pusztán logikai szükségszerűségekből levezethető.[9] Az

---

[8]Reichenbach 1956; Grünbaum 1974; Mellor 1981, 1995.

[9]Swinburne 1968.

A-teoretikus Prior, és a B-teoretikus Belnap egyaránt úgy gondolják, hogy az idősorozatoknak tükrözniük kell a világ modális szerkezetét.[10]

1.0.0.7. Ezen a ponton meg kell említenünk egy másik nehézséget. Bárhogyan fogjuk is fel a változás fogalmának lényegét, ahhoz, hogy megmondjuk valamiről, hogy változik-e, és hogy milyen mértékben és milyen módon, előzetesen meg kell tudnunk mondani azt, milyen esetben lenne változatlan. Nevezhetnénk ezt a „filozófiai párhuzamos eltolás problémájának”. Ahhoz, hogy két geometriai alakzat egymástól különbözzön, előzetesen definiálnunk kell az egybevágóság fogalmát, vagyis a kongruencia esetét. Ehhez az kell, hogy előzetesen definiáljuk, mit nevezünk *párhuzamos eltolásnak*. Ahhoz, hogy azt mondhassuk, a test mozgásállapota valamely kölcsönhatás következtében megváltozik, előzetesen definiálnunk kell, mi számít *inerciális mozgásnak*,[11] ahhoz, hogy egy fizikai mező két lokális konfigurációját a téridőn összehasonlítsuk, s hogy azt mondhassuk, hogy valamilyen más mezőkkel való kölcsönhatás következtében e konfiguráció megváltozott, előzetesen szükségünk van egy *konnekció*-fogalomra,[12] és így tovább. Vagyis a változás fogalma nem kevésbé konvencionális, mint a téré és az időé.

1.0.0.8. Már az is problematikus, pontosabban csupán konvenció kérdésének tűnik, hogy mi az az entitás, amelynek a változásáról van szó. Mi változik meg a valóságban, ha az <én autóm> 1999-ben *Fiat*, és 2000-ben *Suzuki*? Ebben az esetben „tudjuk” a választ: semmi, <az én autóm> 1999-ben nem ugyanazt az entitást jelöli, mint

[10]Prior 1993; Belnap 1992.

[11]Ez történik az általános relativitáselméletben: a geodetikus mozgástól történő deviációhoz előbb szükség van a geodetikus fogalmára, melyet a Riemann-konnekcióval definiálunk (Hawking és Ellis 1973).

[12]Figyelemre méltó, hogy a részecskefizika ma standardnak számító *gauge*-elméletei pontosan ebből a megfontolásból születtek meg (Yang és Mills 1954), és persze nem véletlen, hogy a matematikai leírásukra szolgáló fogalom éppen egy konnekció, hiszen a differenciálgeometriában a konnekció fogalma mögött ugyanaz az intuitív gondolat húzódik meg.

<az én autóm> 2000-ben. De nem ugyanez-e a helyzet bármilyen más entitással? Vajon <én> ugyanaz vagyok-e 1999-ben, mint 2000-ben? Például nem pontosan ugyanazokból a sejtekből állok! Vajon a személyiségem ugyanaz-e 1999-ben, mint 2000-ben, kérdezhetnénk Derek Parfittel?[13] Egy társasház lakóközössége ugyanaz a lakóközösség-e, ha közben a lakók kicserélődtek? Egy foton ugyanaz a foton-e 10^{-9} másodperccel később? Ezek mind nagyon is releváns metafizikai kérdések. Csupán azt kívántam érzékeltetni, hogy egyfelől igaza van McTaggartnak abban, hogy az idő és a változás fogalma szorosan összefonódik, és abban is, hogy a változás Russell-féle értelmezése nem kielégítő, ugyanakkor nem gondolom, hogy ezeket a filozófiai kérdéseket akár az angol igeidők elemzésével, akár az **1.0.0.4.** pontban bemutatott, logikai bűvészmutatvánnyal lehetséges volna megválaszolni.

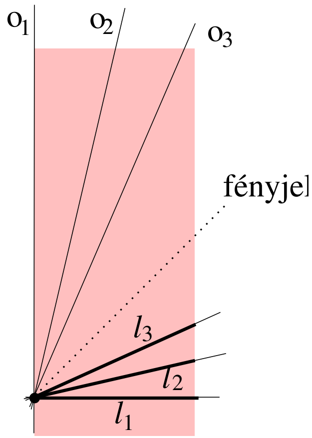
1.0.0.9. Úgy tűnik, hogy a relativitáselmélet téridő fogalma összeegyeztethetetlen az indeterminizmus gondolatával. A „nyitott jövőhöz” mindenek előtt „jövő” kell, de mint Einstein írja egyik levelében: *„Számunkra, akik hiszünk a fizikában, a múlt, a jelen és a jövő közötti szeparáció csupán illúzió, nagyon makacs illúzió.”*[14] Ha a világ négydimenziós entitásokból áll, mint ahogyan azt a relativitáselmélet[15] ma elfogadott, Minkowski-féle felfogása állítja, akkor az idődimenziónak nincs különösebben kitüntetett szerepe a többi (térszerű) dimenziókkal szemben.

1.0.0.10. Érdemes felidéznünk például a Lorentz-kontrakciót szemléltető téridő diagramot (1. ábra). A relativitáselmélet tanítása szerint, ontológiai státuszukat illetően, a rúdnak különböző vonatkoztatási rendszerben értelmezett „hosszúságai” egyenértékűek. A különböző megfigyelők rendszerében, találkozásuk pillanatában, maga a „rúd” különböző események összességét jelenti. A tanulság, amit

[13] Parfit 1987.

[14] Wang 1995.

[15] „Relativitáselmélet” alatt mindig a speciális relativitáselméletet fogjuk érteni.



1. ábra. *A „rúd” más és más események összességét jelenti a három megfigyelő számára találkozásuk pillanatában. Ezek a „pillanatnyi rudak” azonos ontológiai státusszal rendelkeznek*

ebből a téridő diagramból levonhatunk, hogy – a relativitáselmélet szerint – a „rúd” valójában egy négydimenziós objektum; a különböző megfigyelők egyetlen négydimenziós entitás más és más térszerű metszeteinek a kiterjedését tekintik a rúd hosszának. Nincs tehát különbség a különböző térszerű metszetekhez tartozó események ontológiai státuszában.

1.0.0.11. Egyes szerzők[16] arra a konklúzióra jutottak, hogy a relativitáselmélet 4-dimenziós megközelítésmódja[17] nem összeegyeztethető az objektív indeterminizmussal. E szerzők eltérő formában fogalmazták meg érveiket, melyeket nem kívánok itt mind felidézni. Helyette a problémának egy olyan absztrakt megfogalmazását adjuk meg, amelyhez nem szükséges, hogy előzetesen tisztázzunk olyan fogalmakat, mint *létező*, *valóságos*, *determinált*, stb. Az argumentum ilyen absztrakt megfogalmazása, remélhetően, sűrítve tartalmazza a fent említett szerzők megfigyeléseinek lényegét.

Jelölje $(\mathcal{M},\eta)$ a Minkowski-téridőt. Egy inerciális megfigyelő világvonala egy jövőbe mutató, időszerű γ egyenes. Jelölje $p \sim_\gamma q$ azt a relációt, amely abban áll, hogy két esemény, p és q, egyidejű a γ megfigyelő rendszerében. Használni fogjuk továbbá a szokásos $J^+(p)$, $J^-(p)$, $I^+(p)$ és $I^-(p)$ jelöléseket a $p \in \mathcal{M}$ téridőpont ka-

[16] Rietdijk 1966; 1976; Maxwell 1985; Putnam 1967.

[17] Vö. Bennett 1988, 114. o.

uzális jövőjére, múltjára, illetve kronológiai jövőjére és múltjára.[18] Vezessük be a $C(p) = \mathcal{M} \setminus (J^+(p) \cup J^-(p))$ jelölést. E jelöléseket kiterjeszthetjük a téridő tetszőleges $X \subseteq \mathcal{M}$ részhalmazára:

$$\begin{aligned} J^{\pm}(X) &= \bigcup_{p \in X} J^{\pm}(p) \\ I^{\pm}(X) &= \bigcup_{p \in X} I^{\pm}(p) \\ C(X) &= \bigcup_{p \in X} C(p) \end{aligned}$$

1. Tétel (Putnam). *Legyen A egy nem üres részhalmaza $\mathcal{M}$-nek, melyre fennáll, hogy*

$$(\forall p, q) \left[\left(p \in A \,\&\, (\exists \gamma) \left[q \sim_\gamma p \right] \right) \Rightarrow q \in A \right] \tag{1}$$

Ekkor $A = \mathcal{M}$.

Bizonyítás. Ha $p \in A$, akkor $C(p) \subseteq A$. Valóban, ha $r \in C(p)$, akkor p és r térszerűen szeparáltak. Következésképpen létezik egy olyan γ inerciális megfigyelő, hogy $r \sim_\gamma p$, amely maga után vonja, hogy $r \in A$. Hasonló megfontolásból, $C(C(p)) \subseteq A$. Másfelől azonban, a Minkowski-téridőben, $(\forall p) [C(C(p)) = \mathcal{M}]$, tehát $A = \mathcal{M}$.

1.0.0.12. A tétel következménye nyilvánvaló: Mindegy, hogy mit értünk pontosan „ontológiai státusz” alatt. Ha ez a „státusz” úgy van hozzárendelve a téridő eseményeihez, hogy eleget tesz a (1) követelménynek, akkor, a tétel állítása szerint, a téridő minden eseménye azonos „ontológiai státuszú”.

1.0.0.13. Ha az objektív modalitás ideáját összhangba akarjuk hozni a relativisztikus téridő modellünkkel, akkor a következők valamelyikét fel kell adnunk:

[18] Wald 1984, 188. o.

(A) a determináltság-indetermináltság eredeti, vonatkoztatási rendszertől független koncepcióját,

(B) hogy bármely vonatkoztatási rendszerben kapcsolat legyen az események determináltsága–indetermináltsága és a múlt–jelen–jövő alapú temporális klasszifikációja[19] között,

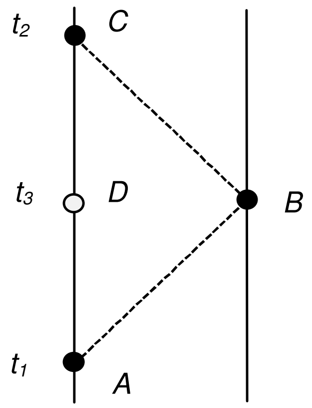
(C) a Minkowski-téridőt, hogy elkerüljük a $C(C(p)) = \mathcal{M}$ egyenlőséget, melyet a tétel bizonyításában kihasználtunk.

1.0.0.14. Régóta folyik a vita a relativitáselmélettel kapcsolatos filozófiai irodalomban arról a kérdésről, vajon a távoli események egyidejűségének fogalma, ahogyan azt az elmélet definiálja, csupán konvenció, vagy a természet egy ténye. Túl azon a már említett *triviális szemantikai konvencionalizmuson*, hogy semmiféle tudományos terminus nem vezethető be valamifajta definíció nélkül, a gondosabb analízis azt mutatja, hogy az egyidejűség fogalma elkerülhetetlenül tartalmaz egy nem triviális konvencionális elemet.

1.0.0.15. Idézzük fel az egyidejűség standard definícióját. Jelölje A azt az eseményt, amikor az O_1 megfigyelő egy fényjelet indít el az O_2 megfigyelő felé. Legyen B az az esemény, amikor a fényjel megérkezik O_2-höz, és O_2 azonnal egy másik jelet indít visszafelé O_1-hez, akihez ez a C téridőpontban érkezik meg (2. ábra). A kérdés az, hogy az O_1 megfigyelő világvonalán melyik az az esemény, amelyik egyidejű B-vel. Eléggé kézenfekvő, hogy ez az esemény valahol az A és C között van a $t_3 = t_1 + \varepsilon(t_2 - t_1)$ időpontban, ahol $\varepsilon \in (0,1)$. Azt gondolhatnánk, hogy ε értékét a fény oda- és visszafelé történő terjedésének sebessége meghatározza. Kiderül azonban, hogy ez nincs így. A

$$\frac{\overrightarrow{c}}{\overleftarrow{c}} = \frac{\varepsilon}{1-\varepsilon}$$

[19] Melyet – a McTaggarttól származó elnevezést használva – A-determinációnak neveznek az irodalomban. (Shimony 1993b, 272. o.)



2. ábra. *A azt az eseményt jelöli, amikor az O_1 megfigyelő egy fényjelet indít el az O_2 megfigyelő felé. Legyen B az az esemény, amikor a fényjel megérkezik O_2-höz, és azonnal egy másik jelet indít visszafelé O_1-hez, akihez ez a téridő C pontjában érkezik meg. Tradicionálisan az A és C esemény közötti felezőpontot szokás a B eseménnyel egyidejű eseménynek tekinteni*

hányadosból ε nem határozható meg, mert ennek a hányadosnak az értéke nem kevésbé konvencionális, mint magának az ε-nak az értéke. Nem lehetséges ugyanis a fény terjedésének egyirányú sebességét megmérni az egyidejűség előzetes fogalma nélkül. Ez egy egyszerű logikai tény, ha direkt módon a fény egyirányú terjedési sebességét akarnánk megmérni, hiszen ehhez a fényjel elindulásának és megérkezésének idejét kellene összevetnünk. Számos olyan javaslat is volt az irodalomban, amely különböző indirekt módszerekre vonatkozott. E javaslatok részletesebb analízise kimutatta,[20] hogy nincs olyan trükkös mérési eljárás, amellyel lehetséges lenne a fény egyirányú terjedési sebességét meghatározni anélkül, hogy az eljárás maga ne tételezné fel az egyidejűség előzetes definícióját.[21]

1.0.0.16. Nincs tehát a természetnek olyan ténye, amellyel ellentmondásba kerülnénk, ha az egyidejűség definíciójában ε értékét minden inerciarendszerben $\frac{1}{2}$-től különbözőnek választanánk. Ezzel ter-

[20]Salmon 1977.

[21]Érdemes talán megjegyeznem, hogy a Michelson–Morley-kísérlet csupán a fény oda-vissza terjedésének átlagsebességére vonatkozó izotrópiát igazol.

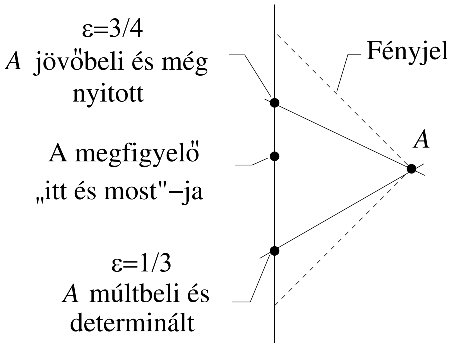
mészetesen a fény (egyirányú) terjedési sebessége c-től különböző lenne, kettős értelemben is, különböző az egyes inerciarendszerekben, és különböző a különböző irányokban. Érdemes megjegyeznünk, hogy a relativitáselmélet megfogalmazható anélkül is, hogy az ε értékét rögzítenénk. Mint Winnie (1970) megmutatta, meghagyhatjuk ε-t szabad „gauge”-paraméternek az elméletben. Minden olyan döntő kísérleti tény, amely a relativitáselméleten belül levezethető az $\varepsilon = \frac{1}{2}$ feltevés mellett, levezethető a relativitáselmélet többi posztulátumából, melyek nem tartalmaznak az egyidejűségre illetve a fény egyirányú sebességére vonatkozó utalást.

1.0.0.17. Azt hihetnénk, hogy az egyidejűség definíciója legalább a pre-relativisztikus fizikában, az éterhez rögzített vonatkoztatási rendszerben egyértelmű, hiszen ott a fényterjedés sebességét a Maxwell-egyenletek határozzák meg. Vegyük észre, hogy ezt az argumentumot a relativitáselméletben is használhatnánk, azonban egyik elmélet keretében sem tartható. Valóban, ha sikerülne empirikusan megállapítanunk a Maxwell-egyenletekben szereplő, *kritikus sebesség*[22] egyirányú értékét, akkor ebből megtudhatnánk a fényterjedés egyirányú sebességét. A kritikus sebesség meghatározásához azonban arra lenne szükség, hogy kimérjük, mekkora erőt gyakorol a mágneses mező egy adott irányban adott sebességgel mozgó töltésre. Tehát előzetesen meg kellene határoznunk a töltés egyirányú sebességét egy adott inerciarendszerben, és ezt, mint fentebb kimutattuk, nem tehetjük meg az egyidejűség előzetes definiálása nélkül.

1.0.0.18. Ami minket az egyidejűség fogalmának fenti analíziséből elsősorban érdekel, az a következő. Egy pillanatra fogadjuk el, hogy a „determináltság”, éppúgy, mint a „múlt”, „jelen” és „jövő”, olyan fogalmak, melyeket csak egy-egy megfigyelő vonatkoztatási rendszeréhez kötve értelmezhetünk. A fenti analízisből azonban az derül ki,

[22] Továbbra is a Maxwell-egyenletekben szereplő c természeti állandóról van szó, azonban most úgy gondolunk rá, mint egy sebességdimenziójú állandóra, amely a mágneses jelenségekre vonatkozó formuláinkban szerepel.

hogy a helyzet ennél is rosszabb. Egyetlen vonatkoztatási rendszeren belül maradva is azt kell mondanunk, hogy ezek a fogalmak semmi olyat nem fejeznek ki, ami az objektív világra vonatkozna. A „determináltság”, a „múlt”, a „jelen” és a „jövő” mind *gauge*-függő fogalmak. Az A esemény determinált az $\varepsilon = \frac{1}{3}$ választás szerint (3. ábra), de nem determinált, ha történetesen az ε értékét $\frac{3}{4}$-nek választjuk.



3. ábra. *Az A esemény determinált az* $\varepsilon = \frac{1}{3}$ *választás szerint, de nem determinált az* $\varepsilon = \frac{3}{4}$ *választás mellett*

1.0.0.19. Mindezt figyelembe véve, az 1. Tételt erősebb formában ismételhetjük meg, úgy, hogy egyetlen vonatkoztatási rendszerben maradunk, és nem hivatkozunk a különböző megfigyelők vonatkoztatási rendszereire. Ennek az a jelentősége, hogy a tétel a Lorentz-elmélet keretei között is érvényes lesz.

2. Tétel. *Legyen* γ *egy rögzített megfigyelő. Jelölje* $q \sim_\varepsilon p$ *azt a relációt, hogy* q *és* p *egyidejű a* γ *megfigyelő rendszerében egy adott* ε *mellett. Legyen* A *egy nem üres részhalmaza* $\mathcal{M}$ *téridőnek, melyre fennáll, hogy*

$$(\forall p, q)\left[\left(p \in A \,\&\, (\exists \varepsilon)\left[q \sim_\varepsilon p\right]\right) \Rightarrow q \in A\right] \tag{2}$$

Ekkor $A = \mathcal{M}$.

Bizonyítás. Ha $p \in A$, akkor $C(p) \subseteq A$. Valóban, ha $r \in C(p)$, akkor p és r térszerűen szeparáltak. Következésképpen létezik olyan ε, hogy $r \sim_\varepsilon p$, amely maga után vonja, hogy $r \in A$. Hasonló megfontolásból, $C(C(p)) \subseteq A$. Másfelől azonban, a Minkowski-téridőben, $C(C(p)) = \mathcal{M}$, tehát $A = \mathcal{M}$.

1.0.0.20. Nem a teljes relativitáselmélet és még csak nem is az inerciális vonatkoztatási rendszerek relatív mozgása az, amely maga után vonja az egyidejűség fogalmának konvencionális jellegét, hanem *az az egyszerű természeti tény, hogy van határsebesség.* Néhány kvantumjelenségtől eltekintve (EPR- és GHZ-kísérletek), melyekre még a későbbi fejezetekben visszatérünk, erős kísérleti alátámasztást nyert az a tény, hogy semmiféle fizikai hatás nem terjedhet a fénysebességnél nagyobb sebességgel. Az egyidejűség konvenció jellegéből a következő konklúziót vonhatjuk le: A „múlt- jelen- és jövőidejűség", valamint a „determináltság–indetermináltság" – amennyiben azt a múlt–jelen–jövő-típusú temporális tulajdonságokhoz kötjük –, konvencionális, gauge-függő fogalmak, vagyis *nincs a valóságban az eseményeknek olyan objektív tulajdonsága, amelyet ezek a fogalmak fejeznek ki*. Hangsúlyoznunk kell, hogy lehetne ez az objektív tulajdonság akár egy vonatkoztatási rendszerhez kötött tulajdonság is. Egy test sebességének nagysága, például, egy olyan objektív tulajdonság, amelyet egy adott vonatkoztatási rendszerhez kötött módon értelmezünk. A múlt-, jelen- és jövőidejűség, valamint a determináltság azonban olyan fogalmak, mint például a test potenciális energiájának értéke. A potenciális energia tipikusan gauge-függő mennyiség, ezért nem fejez ki semmit sem. Nincs olyan fizikai, empirikus tény, amelyből ez a számérték következne.

Mi tehát a filozófia?

A filozófia [. . .] olyan szellemi tevékenység, mely pontosan arra irányul, hogy ilyen végső kérdésekre választ adjunk, méghozzá nem meggondolatlan és dogmatikus módon, ahogy a mindennapi életben, sőt még a tudományokban is szokásunk, hanem kritikai alapon: miután átvizsgáltunk mindent, ami ezeket a kérdéseket bonyolulttá teszi, és ráeszméltünk arra a rengeteg határozatlanságra és zavarosságra, ami közönséges gondolatainkat jellemzi.[23]

[A] filozófiát nem a tudományhoz szükséges *a priori* előtanulmánynak vagy alapozásnak tekintem, hanem a tudománnyal egységben álló vállalkozásnak. A filozófiát és a tudományt egyazon hajó utasainak tekintem, mégpedig egy olyan hajóénak, amelyet, hogy – szokásom szerint – felidézzem Neurath hasonlatát, csak a nyílt tengeren építhetünk át, miközben benne utazunk. Nincs külső kilátópont vagy első filozófia. Minden tudományos eredmény és minden tudományos sejtés, amely jelenleg kézenfekvőnek tűnik, nézetem szerint éppolyan üdvözlendő a filozófiában, mint másutt.[24]

[23] Russell 1996, p. 15.
[24] Quine 1999.

2. előadás

Ontológia

3. előadás

Nyelv, logika, matematika

Szintetikus – analitikus, *a priori – a posteriori*

A nyelv

Szintaxis és szemantika

A jelentés problémája

- Referencia-elmélet
- Mentalizmus
- Nyelvhasználat
- Következtetés-alapú elmélet
- Verifikáció-elmélet
- Igazságfeltétel-elmélet
- Eliminatív elmélet

Továbbára is kérdés: Mit jelent „igaznak” lenni?

Szintaxis és a szemantika viszonya – kompozicionalitás elve

Logika/Matematika

Mi a logika?

Tudományszociológiai értelemben a logika a matematika egyik ága ÉS a filozófia egyik ága. (A világ nagy egyetemein pl. matematika és filozófia tanszékeken is szokás logikával foglalkozni.)

Egy logika általában a következőkből áll:

- Formális nyelv

- deduktív (következtetési) rendszer
- modell-elméleti szemantika (mi mit jelent, mi mikor igaz vagy hamis, stb.)

Ezek tipikusan matematikai fogalmak.

Filozófiai értelemben—azt szokás mondani—a logika a helyes gondolkodás/következtetés tudománya. A következtetés episztemikus (a megismeréssel összefüggő) mentális aktivitás. Milyen filozófiai relevanciája van tehát a logika matematikai aspektusainak? Szokásos válaszok:

- a logika a helyes gondolkodás mélystruktúrája
- a természetes nyelvet, elégtelenségei miatt, egy formalizált nyelvvel és a formalizált következtetési szabályokkal kell helyettesíteni
- a logika a természetes nyelv matematikai modellje

Az igazi kérdés tehát az, hogy

Mi teszi a logika következtetési szabályait „helyessé”?

Alapvetően az IGAZSÁG-MEGŐRZŐ TULAJDONSÁGA, vagyis, hogy igaz premisszákból igaz konklúziókra vezetnek.

Bár áttételesen beépül a racionális gondolkodás és érvelés társadalmilag/történetileg kialakult normáiba, mindenekelőtt a nyelv használatával összefüggő társadalmi normákba, s ezért úgy tűnhet, hogy semmiféle tapasztalásra nincs szükség egy következtetés helyességének megítéléséhez, ez a tulajdonság alapvetően EMPIRIKUSAN tesztelhető.

$$\begin{array}{ccc}\text{ha a premisszák igazak} & \Rightarrow & \text{a következtetések igazak}\\ \updownarrow & & \updownarrow\\ \text{világ tényei} & & \text{világ tényei}\end{array}$$

A logikai következtetés helyességének kérdése ott tűnik problematikusnak, ahol ezt a legkevésbé várnánk: A MATEMATIKÁBAN! Mi teszi helyessé azt a következtetést, hogy

$$\text{ha az Euklideszi axiómák igazak} \Rightarrow \text{igaz, hogy } a^2+b^2=c^2$$

Honnan tudjuk ugyanis, hogy $a^2+b^2=c^2$ igaz?!

Mi tesz egy matematikai állítást igazzá?

Realizmus, platonizmus, intuicionizmus

A REALIZMUS szerint (pl. J. S. Mill) *a matematikai állítások akkor igazak, ha megfelelnek a minket körülvevő fizikai valóságnak.* Más szóval, a matematika empirikus tudomány: a matematikai állítások a fizikai világ legáltalánosabb tulajdonságait fejezik ki. E felfogás fontos szerepet töltött be a matematika történetében, manapság azonban senki sem gondolja komolyan, hiszen a matematika fogalmai nincsenek közvetlen megfelelésben a valóság elemeivel, például a végtelen fogalmának semmi sem felel meg a külső (a matematikán kívüli) világban.

A MATEMATIKAI PLATONIZMUS a matematika klasszikus fogalmainak *önálló létezést tulajdonít*, függetlenül attól, gondoljuk-e azokat vagy nem, s úgy véli, a matematikai állítások igazságát pusztán e fogalmak analízisével, logikai úton fedezhetjük fel.

AZ INTUICIONISTÁK tagadják a matematikai objektumoknak – az értelemszerűen véges – *konstrukciójuktól független* létezését, ám helyette „saját istenük" (Curry kifejezése[25]), az Intuíció létezésében hisznek, vagyis valami olyasmiben, ami az egyetemes emberi értelem számára *a priori adott*, garantálva ezzel a matematika *objektivitását és használhatóságát.*

[25]Haskell B. Curry: *Outlines of a Formalist Philosophy of Mathematics*, North-Holand, Amsterdam 1951.

Realisták, platonisták és intuicionisták mind hisznek azonban abban, hogy a matematikai állításoknak *jelentésük* van, s ha – a Hilbert-programot követve – formalizáljuk is a matematika nyelvezetét, azt azért tesszük, hogy e jelentést precízebben és tömörebben adhassuk vissza.

A matematika formalista felfogása

szerint az igazság ezzel szemben az, hogy *a matematikai objektumoknak nincs jelentése*. A matematika a formális rendszerek tudománya: *Jeleket* definiálunk és *szabályokat*, melyek alapján e jeleket kombinálhatjuk. Ahogy Hilbert mondta „A matematika egy játék, melyet a papírlapra írt, jelentés nélküli szimbólumokkal játszunk, egyszerű szabályok szerint.” „Pont, egyenes és sík helyett folyamatosan mondhatnánk, asztalt, széket és söröskorsót” – mondta egy másik alkalommal az euklideszi geometriára utalva.

A matematikának semmi köze nincs a végtelen metafizikai fogalmához, és közömbös a térre, időre, valószínűségre vagy a folytonosságra vonatkozó intuíciónkkal szemben. A matematika nem produkál, és nem old meg Zénón-paradoxonokat! „Leírhatok egy jelet, mondjuk α-t, és elnevezhetem az egész számok kardinalitásának. Aztán rögzíthetem a rá vonatkozó manipulációs szabályokat”, mondja Dieudonné.[26] Az egész finitista próbálkozás felesleges. Ha a papírra azt írom $10^{10^{10}}$, ez éppúgy csak egy jel, amellyel manipulálhatok, mint bármelyik más. A matematika jelenlegi gyakorlata azt mutatja, hogy minél precízebben látjuk be valamely matematikai állítás igazságát, annál nyilvánvalóbb, hogy őt kizárólag az teszi igazzá, hogy levezethető az rendszer axiómáiból a rendszerben érvényes következtetési szabályok segítségével – függetlenül attól, hogy egyébként milyen filozófiai nézeteket vall egy matematikus. Jól jellemzi a helyzetet Jean Dieudonné-nek, a francia Bourbaki csoport egyik ve-

[26]Lásd Arend Heyting: *Intuitionism: an Introduction*, North-Holland, Amsterdam 1956.

zéralakjának sokat idézett mondása : „In everyday life, we speak as Platonists, treating the objects of our study as real things that exist independently of human thought. If challenged on this, however, we retreat to some sort of formalism, arguing that in fact we are just pushing symbols around without making any metaphysical claims. Most of all, however, we want to do mathematics rather than argue about what it actually is. We're content to leave that to the philosophers."

Tehát,

1. A formalizmus lényege, hogy egy állítás bizonyításának/levezetésének létezése nem más, mint a szóban forgó állítás *igazságfeltétele*.

2. Az axiómák sem azért „igazak", mert valamiféle referenciájuk van a valóságos (vagy valamiféle platóni) világra, hanem mert (triviálisan) levezethetők (tudniillik az axiómákból), más szóval definíció szerint igazak.

3. A matematikában az igazság fogalma általában értelmetlen, csak egy adott axiómarendszerre nézve értelmes (ahol az axiómarendszerbe a következtetési szabályokat is beleértjük). Annak a kijelentésnek, hogy „a háromszög szögeinek összege 180 fok" az igazságáról nincs értelme anélkül beszélnünk, hogy ne specifikálnánk, hogy melyik axiómarendszerben (tehát melyik geometriában) van értve.

4. A matematika története ebben a vonatkozásban nem egységes. A matematika reális interpretációja például szinte kihalt a nemeuklideszi geometriák megszületése után. Korábbi korokban elfogadottnak tekintett bizonyításokat ma nem tekintünk elfogadható, precíz formális bizonyításnak. Mint – kissé sarkítva – Russell írja Boole *Laws of Thought*-ja (1854) volt „az első könyv, amelyet matematikáról írtak".

Matematikai elmélet mint formális rendszer

Általában tehát egy matematikai elmélet egy formális nyelv, amely szimbólumokat tartalmaz, szintaktikai szabályokat arra nézve, hogy ezekből a szimbólumokból hogyan lehet összetettebb un. formulákat és formula-sorozatokat előállítani, és logikai szabályokat, amelyek következtetési szabályokat mondanak ki bizonyos formulák „átalakítására", egyikről a másikra való „áttérésre".

Példa (Paul Lorenzen)

Jelek: Olyan stringek, amelyek két betűből állnak, a és b.

Axiómák:

$$L = \begin{cases} a & \\ X \vdash Xb & \text{(Rule 1)} \\ X \vdash aXa & \text{(Rule 2)} \end{cases}$$

Például,

Tétel: $aababb$

Bizonyítás:

$$a \underset{(1)}{\vdash} ab \underset{(2)}{\vdash} aaba \underset{(1)}{\vdash} aabab \underset{(1)}{\vdash} aababb$$

Elsőrendű formális nyelv

Ábécéje

- individuum változók halmaza: $x_1, x_2, x_3, \dots$
- individuum konstansok (esetleg nincs): $a_1, a_2, a_3, \dots$
- függvény-jelek (esetleg nincs): f_i^n
- egy- vagy többváltozós predikátum-jelek (esetleg nincs): P_i^n
- két logikai konnektív: $\neg$ (nem) $\rightarrow$ (ha...akkor)

- egy kvantifikátor: $\forall$ (minden, univerzális kvantor)
- mellékszimbólumok: $($, , és $)$ (a bal zárójel, a vessző és a jobb zárójel)

NB

- A „nem (negáció)", „ha...akkor (implikáció)", valamint „minden" szavak csupán a szimbólumok elnevezései (matematikai terminusai), nem szabad e szimbólumokra úgy gondolni mint amiknek ilyen *jelentése* van. Ezzel szemben a „halmaz" szó nem halmazelméleti terminusként van használva (hiszen még nincs halmazeléletünk!), hanem abban a hétköznapi értelemben mint szimbólumoknak a sokasága. Éppen ezért, ezen a ponton, kerüljük az olyan állításokat, mint hogy „megszámlálhatóan végtelen individuum változónk van", stb.
- „Elsőrendű" arra utal, hogy van benne kvantifikálás (nem nulladrendű) viszont csak individuum változókra vonatkoznak (nincsenek predikátum változók és azokra történő kvantifikálás, stb.)
- A függvény-jelekre nem szabad itt úgy gondolnunk, mint (a naiv halmazelméletben, más szóval, korábbi tanulmányaikban megszokott) „függvényre", vagyis „hozzárendelésre". Csak egy jel, egy szintaktikai egység, melynek segítségével lehet olyat írni, hogy $f^n(t_1, t_2, \ldots t_n)$.

Terminus (term)

A *terminus* fogalmát a következő definícióval adjuk meg:

1. az individuum változók és az individuum konstansok terminusok.

2. Ha f^n egy függvény-jel, és $t_1, t_2, \ldots t_n$ terminusok, akkor $f^n(t_1, t_2, \ldots t_n)$ is terminus.

3. Más nincs.

Helyesen képzett formula (well-formed formula, wf)

(a) Ha $t_1, t_2, \ldots t_n$ terminusok, akkor $P^n(t_1, t_2, \ldots t_n)$ egy wf. (Az ilyet atomi formulának hívjuk.)

(b) Ha ϕ, ψ tetszőleges két wf, akkor $(\phi \to \psi)$ is és $\neg\psi$ is az.

(c) Ha x egy individuum változó és ϕ egy wf, akkor $\forall x\phi$ is wf.

(d) Más nincs.

Rövidítések

A következő *rövidítéseket* definiáljuk:

$\phi \vee \psi$ (vagy)	arra, hogy $(\neg\phi \to \psi)$
$\phi \wedge \psi$ (és)	arra, hogy $\neg(\phi \to \neg\psi)$
$\phi \leftrightarrow \psi$ (akkor és csak akkor)	arra, hogy $(\phi \to \psi) \wedge (\psi \to \phi)$
$\exists x\phi$ (létezik, exisztenciális kvantor)	arra, hogy $\neg(\forall x\neg\phi)$

NBA „vagy (diszjunkció)”, „és (konjunkció)”, stb. elnevezések is csupán matematikai szakkifejezések. Nem kell hozzájuk a hétköznapi nyelvhasználat szerinti jelentést társítanunk.

HFMutassuk meg, hogy a $\{\neg, \rightarrow\}$ konnektívek helyett használhatnánk a $\{\neg, \wedge\}$ vagy $\{\neg, \vee\}$ párokat is a rendszer definíciójában! Hogy pl. $\phi \wedge \psi$ értelmezhető úgy mint $\neg(\neg\phi \vee \neg\psi)$ rövidítése (magát a formulát De Morgan-azonosságnak hívjuk), etc. Hasonlóképpen, $\forall$ helyett kezdhettük volna $\exists$-kel.

Kötött és szabad változó

Egy változót *kötött változónak* nevezünk, ha egy kvantifikátor vonatkozik rá. Egyébként *szabad változónak* nevezzük.

Például:

- A $\exists x P(x,y)$ formulában (röviden formulának fogjuk nevezni a wf-t) x kétszer kötött változóként van jelen, y szabad.
- A $\forall x \forall y (P(x,y) \rightarrow Q(y))$ formulában x és y minden előfordulása kötött. A $\forall x$ kvantifikálás hatóköre a $\forall y(P(x,y) \rightarrow Q(y))$ részformula. A $\forall y$ hatóköre a $P(x,y) \rightarrow Q(y)$ részformula.
- A $\forall x (P(x,y) \rightarrow \forall y Q(y))$ formulában az x kétszer kötött, az y egyszer szabad és két helyen kötött.

Egy ϕ formulában a *t terminus szabad az x változóra nézve*, ha a t-ben előforduló egyetlen y változó esetében sincs x-nek szabad előfodulása valamely $\forall y$ kvantifikáció hatókörében. Más szóval, t terminust büntetlenül behelyettesíthetjük x minden ϕ-beli szabad előfordulásába, anélkül hogy összetütközésbe kerülnénk a ϕ-ben előforduló kvantifikációkkal. Például, tekintsük a

$$\forall x P(x,y) \rightarrow \forall z Q(z,y)$$

formulát. Itt például egy $f^2(x,v)$ terminus nem szabad y változóra nézve. Ezzel szemben például $g^2(y,z)$ szabad x-re nézve, vagy y szabad x-re nézve.

MondatEgy formulát *mondatnak* (vagy *zárt* formulának) nevezünk, ha nem tartalmaz szabad változót.

Prenex formátum Egy formulát *prenex formátumúnak* mondunk, ha a következő alakú:

$$(K_1x_1)\,(K_2x_2)\dots(K_nx_n)\,\phi$$

ahol minden K_i vagy $\forall$ vagy $\exists$, ϕ pedig egy olyan formula, amelyben nincs kvantifikáció. (Az olyan formulát, amelyben egyáltalán nincs kvantifikálás prenex formátumúnak tekintjük.)

A predikátum kalkulus (PC)

A PC axiómái és a következtetési szabályok

A PC egy, a fenti értelemben vett formális nyelv +

Axiómák (Axióma sémák)

A következőkben, ϕ, ψ, χ formulák, $x, y, y_1, y_2, \dots y_n, \dots$ változók, és jelölje $\phi(y)$ az a formulát, melyet úgy kapunk, hogy a $\phi(x)$ formulában az x változót, annak minden szabad előfordulása esetében y-nal helyettesítjük.

(PC1) $(\phi \rightarrow (\psi \rightarrow \phi))$

(PC2) $((\phi \rightarrow (\psi \rightarrow \chi)) \rightarrow (\phi \rightarrow \psi) \rightarrow (\phi \rightarrow \chi))$

(PC3) $((\neg\phi \rightarrow \neg\psi) \rightarrow (\psi \rightarrow \phi))$

(PC4) $(\forall x\,(\phi \rightarrow \psi) \rightarrow (\phi \rightarrow \forall x\psi))$ ha x nem fordul elő szabadon ϕ-ben.

(PC5) $(\forall x\phi \rightarrow \phi)$ ha x nem fordul elő szabadon ϕ-ben.

(PC6) $(\forall x\phi(x) \rightarrow \phi(t))$ feltéve, hogy a t terminus szabad x-re nézve $\phi(x)$-ben.

Következtetési szabályok

(MP) ϕ-ből és $(\phi \to \psi)$-ből következik ψ (modus ponens)

(G) ϕ-ből következik $\forall x \phi$ (generalization)

NB

- Az axiómák tehát egyszerűen a nyelv kiválasztott formulái. („Alapigazságok”, stb. csak verbális dekoráció).
- Egy formális nyelv + néhány axióma + a következtetési szabályok együttesét általában *formális rendszernek* hívjuk.

PC egy tétele Ha a PC egy ϕ formulája véges számú lépésben levezethető az axiómákból a következtetési szabályok alkalmazásával, akkor a ϕ-t *tételnek* nevezzük és azt írjuk, hogy $\vdash \phi$.

BizonyításEgy *bizonyítás* formuláknak egy (véges) sorozata, úgy, hogy mindegyik formula vagy axióma, vagy a sorozatban szereplő korábbi formulából van levezetve a következtetési szabályok valamelyikének alkalmazásával. A sorozat utolsó formulája nyilvánvalóan egy tétel.

$\Sigma \vdash \phi$Gyakran extra axiómákat adunk a rendszerhez és a bővebb rendszerben konstruálunk bizonyításokat. Ha Σ ilyen extra axiómák halmaza, akkor azt írjuk, hogy $\Sigma \vdash \phi$, ha ϕ levezethető abban a bővebb rendszerben, melyet úgy kapunk, hogy a Σ-ba tartozó formulákat mint axiómákat hozzáadjuk az eredeti PC axiómákhoz.

KonzisztenciaFormulák egy Σ halmazáról azt mondjuk, hogy *konzisztens*, ha nem létezik olyan ϕ formula, melyre egyszerre fennállna, hogy $\Sigma \vdash \phi$ és $\Sigma \vdash \neg\phi$.

PC(=) (predikátum kalkulus identitással)

Az előzőekben megismert predikátum kalkulust egy további predikátum-jellel egészítjük ki. Legyen E („ugyanaz mint”, „egyenlő”) egy kétváltozós predikátum. E tulajdonságait a következő axiómák hozzáadásával rögzítjük:

Az egyenlőség axiómái

(E1) $E(x,x)$

(E2) $E(t,s) \to E\left(f^n(u_1,u_2,\ldots,t,\ldots u_n), f^n(u_1,u_2,\ldots,s,\ldots u_n)\right)$

(E3) $E(t,s) \to (\phi(u_1,u_2,\ldots,t,\ldots u_n) \to \phi(u_1,u_2,\ldots,s,\ldots u_n))$

Kényelmesebb jelölés: $x = y \equiv E(x,y)$

HF. Mutassuk meg, hogy E tranzitív és szimmetrikus.

Aritmetika

(A1) $\neg(0 = sx)$

(A2) $(sx = sy) \to (x = y)$

(A3) $x + 0 = x$

(A4) $x + sy = s(x + y)$

(A5) $x \cdot 0 = 0$

(A6) $x \cdot sy = (x \cdot y) + x$

(A7) $(P(0) \wedge \forall x\,(P(x) \to P(sx))) \to \forall x P(x)$

ahol természetesen $x = y$, $x + y$ illetve $x \cdot y$ az $E(x,y)$, $+(x,y)$ illetve $\cdot(x,y)$ helyett áll, ahol E az egyenlőség predikátum, $+$ és $\cdot$ pedig függvények. s szemléletes jelentése az „hozzáadunk egyet” művelet, P pedig egy tetszőleges egyváltozós predikátum. (A7)-et az *indukció axiómasémájának* is szokás nevezni.

Vezessük be a $1, 2, 3, \ldots$ jeleket a következő individuális konstansok jelölésére:

1: $s0$

2: $ss0$

$\vdots$

k: $\underbrace{s \ldots ss}_{k \text{ darab}}\ 0$

$\vdots$

NB.

A jelölésekben használjuk a számokat és írunk olyat, hogy „k-darab”, stb. Vegyük észre, hogy ezek csak kényelmi, tipográfiai eszközök, és nem történik lényegi hivatkozás valamilyen „előzetesen ismert aritmetikára”.

Igaz-e, hogy $2+2=4$?

3. Tétel. $\{aritmetika\} \vdash 2+2=4$ *a PC(=)-ben.*

BizonyításA jelölések definícióját alapul véve tehát azt kell bizonyítanunk, hogy $ss0+ss0=ssss0$:

1. $ss0+ss0=s(ss0+s0)$ [(A4)-ből]
2. $s(ss0+s0)=ss(ss0+0)$ [(A4)-ből]
3. $ss0+ss0=ss(ss0+0)$ [1. és 2. alapján (E3) és (MP) felhasználásával]
4. $ss0+0=ss0$ [(A3)-ból]
5. $ss0+ss0=ssss0$ [3. és 4. alapján (E3) és (MP) felhasználásával]

NB.

- Gyakran olvashatunk az irodalomban olyan gondolatmeneteket, amelyek a „szándékolt interpretációról” szólnak. Természetesen, lehet valamilyen intuíciónk előzetesen arról, hogy az axiomatikusan felépítendő matematikai struktúrától mit várunk. De ennek szigorú, elméleti, matematikai értelemben nyilván nem lehet semmiféle jelentősége. (A matematikában egyébként is számtalanszor elfogadunk formális elméleti gondolatmenetek útján nyert konklúziókat, melyek esetleg ellentmondanak a „józan észnek”, vagy az előzetes intuitív várakozásainknak. Gondoljunk például arra, hogy intuitíve több racionális számnak kellene lennie, mint egész számnak, mégis elfogadjuk a formális bizonyítást, hogy a két halmaz számossága azonos.) Összegezve tehát, az aritmetika *az*, amit most axiomatikusan felépítünk!

- Természetesen lehet arról beszélni, hogy egy axiomatikusan felépített aritmetika hasznos matematikai struktúra-e számunkra, abban az értelemben, hogy használható-e a világ leírásában, vagyis a fizikai elméletekben. Tehát az aritmetika axiomatikus felépítése során lehet az a szándékunk, hogy egy olyan struktúrát hozzunk létre, amely majd alkalmas lesz — egy megfelelő fizikai elmélet részeként — annak leírására, hogy hogyan működik a pénztárgép, vagy alkalmazható lesz abban a fizikai elméletben, amelyet egy juhász használ a nyájba tartozó juhok nyilvántartására, stb.

- Félreértések elkerülése érdekében felhívjuk a figyelmet arra, hogy az itt alkalmazott jelölések eltérnek a tankönyvekben szokásos jelölésektől. Az itt számokkal jelölt $1, 2, \ldots$, és számoknak, tehát „egynek", „kettőnek", stb. nevezett individuum konstansok rendszerint valamilyen megkülönböztető jelölést kapnak, $\overline{1}, \overline{2}, \overline{3}, \ldots$ (lásd Crossley), vagy $0^{(1)}, 0^{(2)}, 0^{(3)}$, (lásd Hamilton), stb. És rendszerint nem is nevezik őket számoknak, hanem „számjegyeknek", „számneveknek" (numerals, numeral terms), megkülönböztetésül az „igazi" számoktól, azaz valamilyen értelemben már előzetesen létező számelmélet számfogalmától, melyeknek a fenti értelemben vett axiomatikus aritmetika valamiféle „axiomatizált" elmélete. Az itt szorgalmazott felfogás szerint azonban az aritmetika az, amit itt axiomatikusan megadunk. Nincsenek „aritmetikai igazságok" mások, mint amiket az axiomatikus aritmetikában az axiómákból levezethetünk. Semmi okunk tehát arra, hogy éppen azt jelöljük valami mással, ami van, és azt jelöljük $1, 2, 3, \ldots$-mal, ami nincs!

The physicalist ontology of formal systems

It is a common belief that philosophy of mathematics must take account of our impression that mathematical truth is a reflection of fact. As Hardy expresses this constraint,

> [N]o philosophy can possibly be sympathetic to a mathematician which does not admit, in one manner or the other, the immutable and unconditional validity of mathematical truth. Mathematical theorems are true or false; their truth or falsity is absolute and independent of our knowledge of them. In *some* sense, mathematical truth is a part of objective reality.

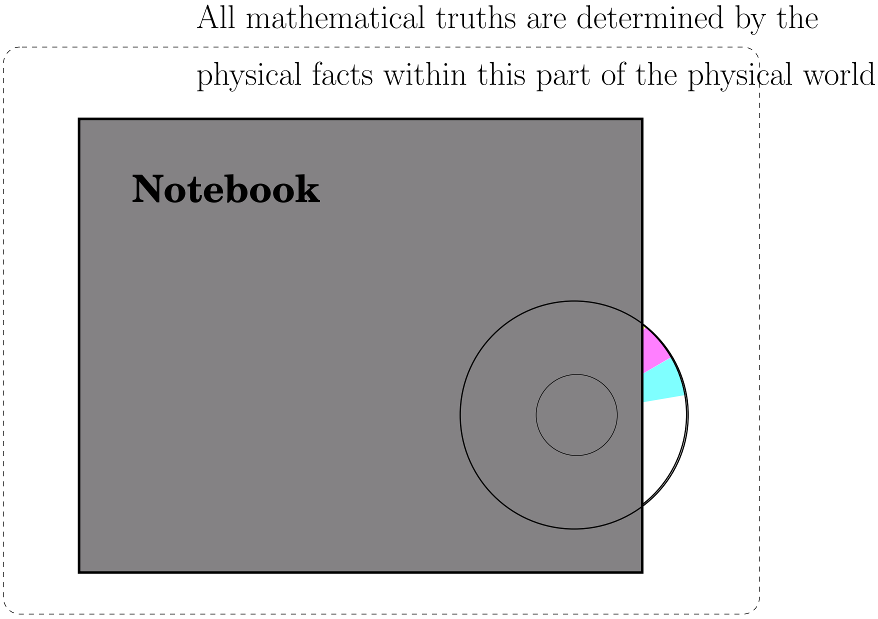
Now we determine what this objective reality actually is.

Thesis:

The objective fact expressed by a mathematical proposition is a fact of a particular part of the physical world: it is a fact of the formal system itself, that is, a fact about the physical system consisting of the signs and the mechanical rules according to which the signs can be combined.

Arguments

- Taking into account that the only means of obtaining reliable knowledge about this fact is mathematical proof, ***it must be a fact of the realm inside of the scope of formal derivations.***



- Sometimes it is argued that symbolism is merely a "convenient shorthand writing" to register the results obtained by *thinking*. However, in the contemporary mathematics there are derivations which are ***not surveyable by the human mind***. (the proof of the four-color theorem) Sometimes even ***the theorem obtained through the derivation process is not surveyable***. (symbolic computer language manipulations)

- Sometimes one executes simple formal derivations also **in the head**. However, *from the point of view of the physicalist interpretation of mind* this case of formal manipulation **does not principally differ** from any other cases of derivation processes.

$\boxed{\text{‘}3+2=5\text{’}}$	This *is not* a linguistic object!
actually means that	*the usual formalist step*
‘formula $3+2=5$ derives from the formulas called the axioms of arithmetics’	This *is* a linguistic object!
which is nothing but	*the physicalist step*
the assertion that there exists a proof-process in the formal system called arithmetic, the result of which is the formula $3+2=5$	This is a usual scientific assertion, just as the assertion of the chemist about the existence of the process $$2H_2+O_2 \longrightarrow 2H_2O$$

In this way, a mathematical truth **has contingent factual content, as any similar scientific assertion**. It is

- expressing **objective fact** of the physical world
- **synthetic**
- ***a posteriori***
- **not necessary** and **not certain**
- **true before anybody can prove it**

4. előadás

A természet törvényei

Mit ír az IEP erről: http://www.iep.utm.edu/l/lawofnat.htm#H1

The essential difference between mathematical truth and a semantical truth in a scientific theory describing something in the world

A physical theory P is a formal system L + a semantics S pointing to the empirical world. In the construction of the formal system L one can *employ* previously prepared formal systems which come from mathematics and/or logic. That is, in general, L is a (first-order) system with

- some logical axioms and the derivation rules (usually the first-order predicate calculus with identity)
- the axioms of certain mathematical theories
- some physical axioms.

A sentence A in physical theory P can be true in *two different* senses:

Truth$_1$: A is a theorem of L, that is, $L \vdash A$ (which is a mathematical truth within the formal system L, a fact of the formal system L).

Truth$_2$: According to the semantics S, A refers to an empirical fact (about the physical system described by P).

For example, 'The electric field strength of a point charge is $\frac{kQ}{r^2}$' is a theorem of Maxwell's electrodynamics—one can derive the corresponding formal expression from the Maxwell equations. (This is a fact of the formal system L.) On the other hand, according to the semantics relating the symbols of the Maxwell theory to the empirical terms, this sentence corresponds to an empirical fact (about the point charges).

Truth$_1$ and Truth$_2$ are *independent concepts*, in the sense that one does not automatically imply the other. Moreover, assume that Γ is a set of true$_2$ sentences in L, i.e., each sentence in Γ refers to an empirical fact, and also assume that $\Gamma \vdash A$ in L. It does not automatically follow that A is true$_2$. Whether A is true$_2$ is again an empirical question. If so, then it is new empirically obtained information about the world, confirming the validity of the whole physical theory $P = L + S$. In **Kevin J. Davey** words:

> The world is built in such a way that from certain bodies of knowledge, certain types of valid mathematical deductions take us from true claims about reality to other true claims about reality. But not all such otherwise valid mathematical deductions do this. This is a fact about the world that physicists must struggle to get their hands around — to learn which deductions are good in which contexts, and which are not.

If it turns out that A is not true$_2$, then this information disconfirms the physical theory, *as a whole*. That is to say, one has to think about *revising one of the constituents* of P, the physical axioms, the semantics S, the mathematical axioms, or the axioms of logic or the derivation rules we applied in the derivation of A—probably in this order. (Here we can see how Quine's *semantic* holism works. The unit of meaning is not the single sentence, but systems of sentences or even the whole of language).

We usually change one entire mathematical/logical theory for some other. For example, when we learned new empirical facts about physical space(-time), we replaced the whole Euclidean geometry with another one. (Unimportant sociological fact.)

The empirical disconfirmation of a physical theory, in which the Euclidean geometry is applied, can disconfirm the *applicability* of

Euclidean geometry in the physical theory in question, *but it leaves Euclidean geometry itself intact*. In this way, one cannot disconfirm mathematical truths like ‘formula $a^2 + b^2 = c^2$ derives from the formulas called Euclidean axioms’. (Here we can observe how Quine’s *confirmational* holism fails. It is not the case that if an empirical finding disconfirms a physical theory which employs a given mathematical theory then it also disconfirms the mathematical theory.)

5. előadás

Okság I.

Az előző előadásban közel kerültünk a kauzalitás kérdésköréhez, vagyis ahhoz a metafizikai problémához, hogy mely események (tények, jelenségek, dolgok) állnak egymással ok-okozati viszonyban, és hogy miben is áll ez a viszony. E rövid fejezet kereteit mindenképpen meghaladná az okságra vonatkozó akárcsak lényegesebb filozófiai álláspontok áttekintése.[27] Csupán arra vállalkozhatunk, hogy röviden utaljunk azokra az irányzatokra, melyek említése elkerülhetetlen az itt képviselt álláspont megfelelő kontextusba helyezéséhez.

A hétköznapi gondolkodásban éppúgy, mint a tudományban az oksági magyarázat, az okokra való hivatkozás centrális szerepet játszik. Hume felbecsülhetetlen érdeme, hogy a kauzalitásra vonatkozó kritikai elemzését követően „az okság a metafizikában *magyarázó* fogalomból fokozatosan *magyarázandó* fogalommá vált".[28] Hume legfontosabb felismerése az oksággal kapcsolatban, hogy a jelenségek közötti ok-okozati viszony nem figyelhető meg. A Tanulmány az emberi értelemről Absztraktjában ezt a következő példán mutatja meg:

> Itt egy biliárdgolyó, és egy másik golyó egy bizonyos sebességgel közelít hozzá. Aztán összeütköznek; az a golyó, amelyik eddig nyugalomban volt, most mozgásba jön. Ez éppoly tökéletes példája az ok-okozati viszonynak, mint bármelyik más, amelyről érzékszerveink vagy értelmünk által tudomást nyerhetünk. Érdemes tehát közelebbről megvizsgálnunk. Nyilvánvaló, hogy a két golyó megérintette egymást, mielőtt a mozgás átadódott volna, és nem telt el idő a golyó meglökése és a mozgásbajövés között. Térben és időben való *szomszédosság* tehát az egyik szükséges feltétele annak, hogy egy ok kifejthesse

[27] Az okság problémakörét összefoglaló, és a modern irányzatokat is bemutató magyar nyelvű irodalom: Huoranszki 2001, III. fejezet. A kauzalitásról szóló bővebb, és sokat idézett monográfia: Mackie 1974. Az itt kifejtett állásponthoz közel álló „ontológiai" megközelítést olvashatjuk Wesley Salmonnál (1984).

[28] Huoranszki 2001, 86. o.

> hatását. Hasonlóképpen nyilvánvaló, hogy az a mozgás, amelyik az ok szerepét játszotta időben megelőzte azt a mozgást, amelyik az okozat szerepét tölti be. Következésképpen, az *időbeli elsőbbség* egy másik szükséges rekvizituma annak, hogy valamit oknak tekinthessünk. De ez még nem minden! Próbáljuk ki másik ugyanilyen golyókkal, megismételve az egész szituációt, és mindig azt tapasztaljuk, hogy az egyik lendülete mozgásba hozza a másikat. Ezzel megtaláltuk a *harmadik* feltételt, nevezetesen az ok és az okozat *állandó együttjárását*. Minden az okhoz hasonló dolog az okozathoz hasonló dolgot hoz létre. E három mozzanaton, tehát a szomszédosságon, az időbeli elsőbbségen és az állandó együttjáráson kívül nincs semmi, amit az ok eme példájában felfedezhetnénk.[29]

Amit közvetlenül tapasztalhatunk, az a szomszédosság térben és időben, az időbeli elsőbbség és az állandó együttjárás. Nem tapasztaljuk azonban az ok-okozati viszonyt. A tudományok, jelesül a fizika látványosan kerüli is az ok fogalmának használatát. Bertrand Russell „Az ok fogalmáról” c. tanulmányában a következőket írja:

> Minden filozófus, bármelyik iskolához tartozzék is, úgy képzeli, hogy az okozatiság a tudomány egyik alapvető axiómája vagy posztulátuma; viszont az olyan fejlett tudományokban, mint például az égi mechanika, az „ok” szó furcsa módon soha nem fordul elő. A „Naturalizmus és agnoszticizmus” c. munkájában dr. James Ward panaszt is emel ezen az alapon a fizika ellen: láthatólag úgy gondolja, hogy azoknak, akik a világra vonatkozó végső igazságról akarnak megbizonyosodni, az okok felfedezésével kellene foglalkozniuk, a fizika viszont még csak

[29] Salmon 1984, 136. o.

> nem is keresi az okokat. Nekem úgy tűnik, hogy a filozófiának nem szabadna ilyen törvényhozói funkciókat magára vállalnia, és hogy a fizika azért nem kutat többé az okok után, mert valójában egyáltalán nincsenek ilyesféle dolgok. Úgy hiszem, az okság törvénye, mint annyi minden más is, egy elmúlt kor emléke, s a monarchiához hasonlóan csak azért él tovább, mert – tévesen – ártalmatlannak vélik.[30]

Az igazsághoz hozzá tartozik, hogy Hume szövegei nem egyértelműek, és nincs egyetértés a filozófiatörténészek között azt illetően, hogy Hume szkepszise azt jelenti-e, hogy van a világban kauzalitás, de lehetetlen az ok-okozati viszonyt közvetlenül tapasztalnunk, vagy pedig hogy tagadja a kauzalitás létezését ontológiai értelemben is. Másrészt, Russell későbbi írásaiban megváltoztatta véleményét, és elismerte, hogy a kauzalitás fundamentális szerepet játszik a fizikában. E két idézetből inkább csak az a fontos számunkra, hogy lássuk, egyáltalán nem magától értetődő, hogy van-e kauzalitás a világban, és hogy hogy mi az.

Mélyebb metafizikai tartalmát illetően, a kauzalitás három különböző felfogását különböztetjük meg: a kauzalitás *episztemikus*, *modális* és *ontológiai* értelmezését.

Episztemikus értelmezés

Az általam kifejtett álláspont, mint majd látni fogjuk, tagadja, hogy a kauzalitás feltétlenül összefüggne a természeti törvény és a tudományos magyarázat fogalmával, abban az értelemben, hogy a kauzalitást a másik két fogalomra vezethetnénk vissza. Fordítva, az okság és például a tudományos magyarázat között fennáll bizonyos összefüggés, amennyiben – Wesley Salmon (1984, p. 19) értelmezése szerint –

[30] Russell 1976, 291. o.

tudományos magyarázatnak azt tekinthetjük, ha a megmagyarázandó jelenséget képesek vagyunk a világ jelenségeinek (eseményeinek) kauzális rendjébe beilleszteni. Mindenesetre, az okságról folytatott filozófiai diskurzusban gyakran keverednek a természeti törvényre, a tudományos magyarázatra és a kauzalitásra vonatkozó megállapítások. Az episztemikus értelmezés szerint ezek nem is elválasztható problémakörök.

Az episztemikus felfogás szerint egy A esemény oka a B eseménynek, ha létezik olyan T természeti törvény (vagy esetleg törvények egy rendszere), hogy A-ból és a T törvényből logikailag következik B. Tehát, ha az asztal egyik végén fekvő mágnesrúd elmozdul, akkor egy nanoszekundum múlva az asztal másik végénél addig nyugalomban álló iránytű is megmozdul. Ez logikai következménye az elektrodinamika egyenleteinek és a <mágnesrúd elmozdul> eseménynek. Az okság ezen értelmezésének episztemikus jellege abban nyilvánul meg, hogy például abból a tényből (kezdeti adatból), hogy a másgnesrúd elmozdult, figyelembe véve a Maxwell-egyenleteket, meg tudjuk jósolni, hogy az iránytű meg fog mozdulni.

Világos, hogy a kauzalitásnak ez az értelmezése sokféle problémát vet fel. Az egyik probléma, hogy mi számít eseménynek. Nem igaz ugyanis, hogy minden kijelentés olyan eseményt takar, amelyre nézve a kauzalitási reláció értelmes. Ha Kovács úr lánya New Yorkban gyereket szül (A), akkor Kovács úr Budapesten (azonnal!) nagypapává változik (B). (Az ilyen, nem fizikai eseményt hívják „Cambridge event”-nek a filozófiában.) B ugyan logikai következménye A-nak és valami „családtan” elméletnek (T), de mégsem tekintenénk ezt a viszonyt kauzális kapcsolatnak.

Továbbá, Hume szerint, az okozat nem lehet szükségszerű következménye az oknak, *logikai* értelemben. Tehát nem tekinthető ok-okozati viszonynak az, hogy ha egy autó sebessége kétszeresére változik (A), akkor a sebességének négyzete a négyszeresére változik (B). Mint Hume mondja, lehetségesnek kell lennie, hogy az ok és az okozat egymástól függetlenül fennálljanak.

Az sem mindegy, hogy milyen jellegű a szóban forgó T törvény. Vannak olyan természeti törvények, amelyekről úgy gondoljuk, hogy nem fejeznek ki kauzális kapcsolatot. Az impulzusmomentum megmaradásából tudjuk, hogy ha egy szabadon keringő műhold távolodik a Földtől (A), akkor lecsökken a sebessége (B), mégsem mondjuk, hogy az egyik jelenség a másiknak oka.

Vagyis egy törvényszerű kapcsolat eredményezheti két dolog együttjárását, két fajta esemény korrelációját. Az együttjárásban megnyilvánuló szimmetrikus viszony azonban nem feltétlenül jelent kauzális kapcsolatot, abban az aszimmetrikus értelemben, hogy az egyik jelenség oka lenne a másiknak. (Később látni fogjuk, hogy ilyen esetben is mindig van valamilyen kauzális kapcsolat a jelenségek között, nevezetesen, az, hogy a korrelációt egy közös ok magyarázza.)

Ha a kauzális kapcsolatnak szükséges feltétele, hogy az ok és az okozat között törvényszerű összefüggés álljon fenn, akkor az okság fogalma máris megterhelődik mindazokkal a metafizikai nehézségekkel, amelyek a természeti törvénnyel kapcsolatban felvetődnek. Itt most arra a nehézségre kell gondolnunk elsősorban, hogy a törvényszerűségek regularitásokhoz kötődnek. A kauzális kapcsolat ezek szerint nem partikuláris események, hanem eseménytípusok közötti viszonyt jelentene. Partikuláris esemény alatt a téridő egy meghatározott tartományában végbement eseményt értjük, vagyis az univerzum történetének azt az epizódját, amely a téridő adott tartományában történik meg.

Igaza van Hume-nak, hogy a kauzális kapcsolatot a regularitások észlelésén keresztül ismerjük fel, a regularitás azonban a kauzális kapcsolat *felismerésének* a szükséges feltétele, és nem magának a kauzális kapcsolatnak. Kérdés azonban, hogy miként lehet a partikuláris események közötti kauzális kapcsolatot értelmeznünk.

Modális értelmezés

A kauzalitás modális értelmezésének lényege a következő: A téridő egy adott tartományában végbement *A* partikuláris esemény oka a téridő egy adott tartományában végbement partikuláris *B* eseménynek, ha igaz a következő kontrafaktuális[31] állítás:

Ha *A* nem következett volna be, akkor *B* nem következett volna be.

Az ok szükségességének ilyen kontrafaktuális megfogalmazása szintén Hume-tól ered. Hume azonban az okság kontrafaktuális megfogalmazását a regularitáselmélet átfogalmazásának gondolta, miközben az alkalmasnak látszik a partikuláris események közötti oksági viszony kifejezésére is – és ez motiválta a kontrafaktuális kauzális relációk elméletének megalkotóját, David Lewist is.[32]

A kauzalitás kontrafaktuális analízise nagymértékben függ magának a kontrafaktuálisnak az értelmezésétől. Hogyan kell értenünk a tényellentétes kondicionálisokat? Vizsgáljuk meg Lewis híres példáját a „Ha a kengurunak nem lenne farka, felbukfencezne.” mondatot. Lewis szerint ezt a mondatot úgy kell értenünk, hogy minden olyan lehetséges világban, amelyikben nincs a kengurunak farka, de minden más vonatkozásban olyan, mint az aktuális világ, a kenguruk felbukfenceznek. Lewis kontrafaktuálisokról szóló művének[33] kiinduló gondolata ez. Ám éppen ez a kezdeti elgondolás az, amely magában hordozza Lewis kontrafaktuális analízisének folyamatosan, újból és újból felmerülő problémáját. Azt tudniillik, hogy milyen mértékben kell hasonlítania egy olyan világnak az aktuális világra, amelyikben nincs a kengurunak farka? Sokféle ilyen alternatív világot képzelhetünk el. Nyilván arra kell gondolnunk, hogy vannak releváns és irreleváns különbségek. Mert nyilvánvaló (talán?), hogy lehetett volna nekem a harmadik elemiben írásból továbbra is hármasom, min-

[31] tényellentétes
[32] Lewis 1986.
[33] Lewis 1973.

den olyan világban, amelyikben a kenguruknak nincs farkuk, de a melbourne-i cirkusz kengurujának farkán nem lehet csokornyakkendő, ha nincs azt mire rákötni. Lewis maga is készséggel elismeri, hogy ezen a ponton a kontrafaktuális analízis homályos, vagyis hogy nem lehetséges világos definíciót adni arra nézve, mit is jelent az, hogy „az alternatív világ olyan, mint az aktuális, csak éppen ...”. Lewis szerint azonban e homályosság természetes, és nem érinti a fogalom használhatóságát.

Hangsúlyoznunk kell, hogy a kauzalitás kontrafaktuális értelmezése független attól, hogy egyébként milyen álláspontot foglalunk el az ún. *modális realizmus* ügyében, vagyis azt illetően, hogy léteznek-e az alternatív világok, s ha igen, milyen értelemben. Témánk szempontjából tehát közömbös, hogy a legszélsőségesebb modális realisták szerint[34] a lehetséges világok ontológiai státusza semmiben sem különbözik az aktuális világ ontológiai státuszától.

Kérdés az is, hogy milyen értelemben „lehetségesek” ezek a lehetséges világok. Lewis – megint csak hume-i hagyományokat követve – bevezeti az ún. „újrakombinálhatóság elvét”[35] vagyis azt az elvet, mely szerint – tömör egyszerűséggel kifejezve – az aktuális világunkban lévő dolgokat szabadon összekuszálhatjuk, és mindezt elszállásolhatjuk egy lehetséges világban.

> A lehetséges világok teljességének (plenitude) érzékeltetése céljából elfogadom az újrakombinálhatóság elvét, ami szerint különböző lehetséges világok részeinek összetoldozásaiból újabb lehetséges világok jönnek létre. Durván azt mondhatnánk, hogy az elv szerint bármi együtt létezhet bármi mással, feltéve, hogy más helyet foglal el a téridőben. Hasonlóképpen, bármi létezhet bármi más nélkül is.[36]

[34]Belnap és Green 1994; Vö. Lewis 1973, 84. o.

[35]Huoranszki 2001, 161. o.

[36]Huoranszki 2001, 161. o. (A szövegrész angol eredetije: Lewis 1986, 87-88. o.)

Ebben az értelemben tehát minden lehetséges. Lewis azonban korlátozza ezt az elvet, mégpedig a következőképpen:

> [A]zt hiszem, hogy léteznek olyan világok, melyekben a fizika különbözik a mi világunk fizikájától, de nem léteznek olyanok, melyekben a logika és az aritmetika más, mint a mi világunk logikája és aritmetikája. Ezzel nem mondok többet, mint hogy szisztematikusan kifejtem azt a naiv, filozófia előtti véleményemet, mely szerint a fizika lehet más, de a logika és az aritmetika nem.[37]

Nem célunk itt Lewis metafizikai álláspontját megvitatni, csak zárójelben megjegyezzük, hogy 1) teljesen indokolatlan a matematika és a logika e kitüntetett státuszba helyezése a természet törvényeivel szemben,[38] és megkérdezzük, 2) *melyik* logika és *melyik* aritmetika, és például *melyik* geometria, stb. *az* aktuális világ „logikája”, „aritmetikája” és „geometriája”, melyet e kitüntetett státuszba emelünk. Lewis „lehetséges világ”-fogalma nem mentes tehát egyfajta platonizmustól.

A kauzalitás kontrafaktuális értelmezését tekintve azonban nem közömbös, hogy hol húzza meg Lewis az újrakombinálhatóság elvének határait. Első közelítésben tekintsük azt az esetet, amikor az elvet nem korlátozzuk semmivel, tehát nem vagyunk tekintettel a Lewis által respektálni ajánlott logikai és matematikai igazságokra sem. Ebben az esetben nyilván az lesz az aktuális világhoz legközelebbi olyan lehetséges világ, amelyben A nem történik meg, amelyet úgy kapunk, hogy az aktuális világból elvesszük A-t, és kész (azaz benne hagyjuk B-t), minden további változtatás nélkül. Ebben az esetben tehát *nincs kauzalitás*. Nincs két olyan esemény, melyekre a lewisi definíció alkalmazható lenne.

---

[37] Lewis 1973, 88. o.

[38] Bővebben erről: E. Szabó 2002.

Második lépésben kapcsoljuk be a Lewis által előírt korlátot. Vagyis az újrakombinálhatóságnak egyedüli korlátja a logika és a matematika legyen. Ebben az esetben már nem feltétlenül igaz, hogy létezik olyan lehetséges világ, amelynek nem része A, de része B (minden mást fixen hagyva). Mikor nem lehetséges ilyen világ? Akkor, ha A és B között logikai/matematikai kapcsolat van, pontosan, ha $\neg A\&B$ logikai ellentmondás! Más szóval ez azt jelenti, hogy a lewisi értelmezés szerint az ok és az okozat egymás logikai/matematikai következményei, vagyis azzal az esettel van dolgunk, amit Hume indítványára nem tekintettünk kauzális kapcsolatnak.

Ha valódi, kontingens kauzális kapcsolatot kívánunk visszakapni a lewisi értelmezés segítségével, akkor a lehetséges világok családját kiválasztó újrakombinálhatóság elvének máshol kell meghúzni a határait. Például szokás azt mondani, hogy azok a lehetséges világok, amelyek respektálják a természet törvényeit. Ezzel azonban nagyon bizonytalan talajra léptünk. Nehéz ugyanis megmondani, hogy a természet törvényei alapján milyen az a lehetséges világ, amelyikben nem történik meg egy adott (tehát a téridő egy adott helyén végbement) A esemény, és melyik ezek közül az aktuális világhoz legközelebbi. S még nehezebb arra válaszolni, hogy egy ilyen világban bekövetkezik-e B. Mert például az evolúció során, ha nincs farok a kengurun, akkor másképpen alakul a testének súlyeloszlása (mivel, ha mindig felbukfencezett volna, akkor most nem lenne kenguru az élővilágban). Minél gazdagabb képünk van az aktuális világban érvényesülő törvényszerűségekről, annál valószínűtlenebb, hogy az univerzum történetének bármelyik porcikáját elhagyhatjuk, a történet egészének átírása nélkül.

Egyáltalán, vizsgáljuk meg a lewisi definíció logikai szerkezetét! Tegyük fel, hogy a lehetséges világoknak egy T elméletet kell kielégíteniük. Vitathatatlan – ha mégoly homályos is ennek a „távolságnak” az értelme –, hogy az aktuális világhoz legközelebbi A nélküli világ az lenne, amelyben – minden mást fixen hagyva – nincs A, de van *B*. Mikor nem található ott ez a világ a lehetséges világok családjában? Akkor, ha a $\neg A$-ból és a *T* elméletből következik $\neg B$. Tehát, *a kauzális viszony lewisi értelmezése praktikusan nem különbözik az episztemikus értelmezéstől*. Igaz ez abban az értelemben is, hogy a kontrafaktuális értelmezés sem mellőzheti a regularitásra való hivatkozást, vagyis az azzal történő érvelést, mely szerint *A* és *B* események partikuláris esetei bizonyos eseménytípusoknak, melyekről a *T* elmélet azt állítja, hogy az egyikből következik a másik.

Végül megemlítünk néhány olyan problémát, amely a lewisi definícióval kapcsolatban felvetődik, a fenti, részletesebb elemzés nélkül is. Először is tisztáznunk kell, hogy a lewisi definíció csupán a kauzális faktorok értelmezésére alkalmas, és nem az ok definiálására, abban az értelemben, hogy az ok (önmagában) kiváltaná az okozatot. Nyilvánvaló, hogy a szükséges feltétel nem feltétlenül elégséges is. Az ismert orosz népmesében az óriásira nőtt répa kihúzásának szükséges feltétele a sor végén az Egér húzóereje, de nem tekinthetjük úgy, hogy az óriási répa földből történő kimozdulásának oka az Egér által kifejtett húzóerő. Az Egér húzása, csak egy a kauzális faktorok közül, ugyanúgy, mint a Medve húzása, és a többieké. Más szóval, a lewisi értelemben definiált okok konjunkciója tekinthető olyan értelemben oknak, melyek már maguk után vonják (a természeti törvények értelmében, például) az okozatot. Ez nem csupán a kontrafaktuális értelmezés problémája, hanem általános nehézsége a kauzalitásról való elmélkedésnek, s majd csak az általunk javasolt ontológiai értelmezés keretében fog megoldódni.

Megengedve tehát, hogy a Lewis-féle definícióval csupán egy releváns kauzális faktort értelmezünk, újabb nehézség merül fel: Tekintsük például azt az állítást, hogy „Ha Mari locsolta volna a virágaimat, nem hervadtak volna el.” Első közelítésben valóban úgy tűnik, ha ez a kontrafaktuális mondat igaz, akkor a virágok elhervadásának oka Mari nemlocsolása. Úgy értjük, sok minden más mellett, ez egy releváns kauzális faktor. A helyzet azonban ettől bonyolultabb, hiszen nyilván igaz a következő mondat is: „Ha George Bush locsolta volna a virágaimat, akkor nem hervadtak volna el.”[39] Mégsem gondolnánk, hogy virágaim elhervadásának oka az Egyesült Államok elnökének nem-locsolása. Mégcsak azt sem gondoljuk, hogy az elnök nem-locsolása egyike a releváns kauzális faktoroknak.

6. előadás

Valószínűség

[39] A találó példa Jaegwon Kimtől származik.

„Valószínűségelmélet” a matematikában

Fontos: **Az esemény és a valószínűség itt sugallt „jelentései” csupán didaktikai szerepet játszanak!**

A valószínűségelmélet kiinduló (alap)fogalma az *esemény*. Az események egy *eseménystruktúrát* alkotnak. Jelölje Σ az események halmazát. Σ-n a következő „logikai” műveleteket vezetjük be:

$$\begin{aligned} \text{„vagy”}\ \vee &: (A,B)\in\Sigma\times\Sigma\mapsto A\vee B\in\Sigma \\ \text{„és”}\ \wedge &: (A,B)\in\Sigma\times\Sigma\mapsto A\wedge B\in\Sigma \\ \text{„nem”}\ \neg &: A\in\Sigma\mapsto\neg A\in\Sigma \end{aligned}$$

Tulajdonságok:

$$\left.\begin{aligned} (A\vee B)\vee C &= A\vee(B\vee C) \\ (A\wedge B)\wedge C &= A\wedge(B\wedge C) \\ A\vee B &= B\vee A \\ A\wedge B &= B\wedge A \\ A\wedge(A\vee B) &= A \\ A\vee(A\wedge B) &= A \end{aligned}\right\}\ \text{háló} \tag{3}$$

$$\left.\begin{aligned} A\wedge(B\vee C) &= (A\wedge B)\vee(A\wedge C) \\ A\vee(B\wedge C) &= (A\vee B)\wedge(A\vee C) \end{aligned}\right\}\ \text{disztributív} \tag{4}$$

$$\left.\begin{array}{c} \exists\emptyset,1\in\Sigma \text{ olyan, hogy } (\forall A\in\Sigma)\text{-ra} \\ A\vee\emptyset=A \text{ és } A\wedge\emptyset=\emptyset \\ A\vee 1=1 \text{ és } A\wedge 1=A \end{array}\right\}\ \begin{array}{l}\text{nullelemes és}\\ \text{egységelemes}\end{array} \tag{5}$$

$$\exists\neg:\Sigma\to\Sigma \text{ olyan, hogy } \left.\begin{array}{c} A\vee\neg A=1 \\ A\wedge\neg A=\emptyset \end{array}\right\}\ \text{komplementumos} \tag{6}$$

$$\left.(\forall I)\text{-re } \bigvee_{i\in I}A_i\in\Sigma \text{ és } \bigwedge_{i\in I}A_i\in\Sigma\right\}\ \text{teljes} \tag{7}$$

Továbbá bevezetjük a $A\leq B \Leftrightarrow A\wedge B=A$ parciális rendezést. A (3)–(7) tulajdonságokat kielégítő algebrai struktúrákat *Boole-hálónak* nevezzük. A Boole-háló fogalmát szemléletessé tevő fontos tétel (Stone), hogy minden Boole-háló izomorf egy megfelelő Ω halmaz részhalmazainak Boole-hálója egy részhálójával, ahol a részhalmazokon értelmezett hálóműveletek, természetesen, az unió, a metszet és a komplementer képzés.

A *valószínűség* egy $p : \Sigma \to [0,1]$ leképezés, a következő tulajdonságokkal:

$$p(1) = 1 \tag{8}$$

$$(\forall A_1, A_2, \ldots) \left[(\forall i \neq j) \left[A_i \leq \neg A_j \right] \Rightarrow \left[p\left(\bigvee_{k=1,2,\ldots} A_k \right) = \sum_{k=1,2,\ldots} p\left(A_k\right) \right] \right] \tag{9}$$

A (3)–(9) tulajdonságokra a későbbiekben mint *Kolmogorov-axiómákra* fogunk hivatkozni, a (Σ, p) párt pedig *Kolmogorov-féle valószínűségi modellnek* fogjuk nevezni.

6.0.0.21. Két további fontos fogalmat kell bevezetnünk: Legyen $p(B) > 0$. Az A esemény B eseményre vett *feltételes (kondicionális-kondicionális) valószínűsége*:

$$p\left(A|B\right) \stackrel{def}{=} \frac{p\left(A \wedge B\right)}{p\left(B\right)} \tag{10}$$

A definíciós egyenlőség hangsúlyozása lényeges! A (10) formulát „Bayes-szabálynak” szokás nevezni, de ez az elnevezés meglehetősen félrevezető, mintha nem definícióról lenne szó, hanem arról, hogy a kondicionális valószínűséget valahonnan tudnánk, hogy mi, és mennyi az értéke, és ez a „szabály” azt állítaná, hogy ez egyenlő a (10) jobb oldalával. A kondicionális valószínűséggel kapcsolatban egyébként is sok félreértés van forgalomban, melyekkel bővebben a **6.0.0.21.** pontban fogunk foglalkozni.

A és B események *korrelációjának* nevezzük a

$$\Delta(A,B) = p(A \wedge B) - p(A)p(B)$$

mennyiséget. Ha $\Delta(A,B)=0$, akkor A és B eseményt *függetleneknek* mondjuk.[40] Világosan kell látnunk, hogy két esemény korreláltsága egy szimmetrikus tulajdonságuk. Erről sokszor megfeledkeznek. Mint a **8.0.0.24.** pontban meg fogjuk mutatni, a korreláltságot kifejezhetjük olyan módon is, hogy az látszólag aszimmetrikus:

$$\begin{aligned}
\Delta(A,B) > 0 \;\Leftrightarrow\; p(A|B) > p(A) \Leftrightarrow p(A|B) > p(A|\neg B)\\
\Delta(A,B) < 0 \;\Leftrightarrow\; p(A|B) < p(A) \Leftrightarrow p(A|B) < p(A|\neg B)
\end{aligned}$$

ez azonban csak látszólag aszimmetrikus, hiszen ugyanígy fennáll, hogy

$$\begin{aligned}
\Delta(A,B) > 0 \;\Leftrightarrow\; p(B|A) > p(B) \Leftrightarrow p(B|A) > p(B|\neg A)\\
\Delta(A,B) < 0 \;\Leftrightarrow\; p(B|A) < p(B) \Leftrightarrow p(B|A) < p(B|\neg A)
\end{aligned}$$

Helytelen tehát olyasmit mondanunk, hogy B esemény (sztochasztikus értelemben) oka A-nak, mert $p(A|B) > p(A)$, vagyis, mert „B bekövetkezése megnöveli A valószínűségét”. E megfogalmazás egyébként a kondicionális valószínűség fogalmának félreértését is tükrözi.

Végül legyen itt megemlítve az úgynevezett *teljes valószínűség tétel*: Legyen $\{E_i\}_{i=1,2,\ldots,n}$ egy egységpartíció, tehát

$$\bigvee_{i=1}^{n} E_i = 1$$

$$(\forall i \neq j)\,[E_i \wedge E_j = \emptyset]$$

Ekkor tetszőleges$A \in \Sigma$-ra fennáll, hogy

$$p(A) = \sum_{i=1}^{n} p\left(A|E_i\right) p\left(E_i\right) \tag{11}$$

[40] A „korreláció” és a „függetlenség” csak matematikai definíciók. Bármennyire is sugallják ezek az elnevezések, egyelőre semmi közük a kauzalitáshoz!

A valószínűség értelmezései

Consider the following scientific assertions:

$$\mathbf{E}(\mathbf{r}) = q\frac{\mathbf{r}-\mathbf{r}_q}{\left|\mathbf{r}-\mathbf{r}_q\right|^3}$$

$$t(A) = 43\text{s}$$

On the left hand side of we have a well known, empirically defined physical quantities.

Compare these assertions with typical probabilistic assertions:

$$\text{probability of } A = \ldots$$

For example,

$$\begin{aligned} p(a) &= \operatorname{tr}\left(\hat{P}_a\hat{W}\right) \\ p\left(\{N_i\}_{i=1,2,\ldots}\right) &= \frac{\left(\sum_i N_i\right)!}{\prod_i N_i!} \\ p\,(\text{raining}) &= 0.8 \\ p(\text{H}) &= \frac{1}{2} \end{aligned}$$

What quantities are on the left hand sides of these formulas? “**Interpretation of probability**” ought to mean the answering *this* question.

Megjegyzések

1. A valószínűségszámítás, mint a matematika része, nem a valóság leírására vonatkozik. Ha tehát a valószínűségelmélet fogalmait a reális világra kívánjuk vonatkoztatni, ez – hasonlóan a geometriához – nem tehető meg közvetlenül, hanem csak úgy, hogy a valószínűségszámítás részét képezi a világ leírását szolgáló – például fizikai – elméletnek. A szóban forgó elmélet

lehet a tárgyát képező jelenségek egészen primitív, köznyelvi leírása is. Például egy kockadobást egészen egyszerű nyelven is leírhatunk. Nevezzük a valószínűségelmélet ilyen, a világ leírására szolgáló nyelvben történő interpretációját *reális interpretációnak*.

The right epistemic order:

(a) We have to define—in empirical terms—what "probability" means on the left hand side.

(b) From the *a posteriori* knowledge of "$p(a) = \operatorname{tr}\left(\hat{P}_a \hat{W}\right)$", etc. ... we determine the basic laws satisfied by the quantity we previously **defined and called** "probability".

(c) Finally, we may conclude that probability can be conveniently **described** by the mathematical concept called "probability" in, say, Kolmogorov's "probability theory".

2. A valószínűségszámítást gyakran olyan feladatok megoldásában alkalmazzuk, ahol a feladat maga sem a reális világra vonatkozik, hanem matematikai. Tipikus példa erre a később tárgyalt Bertrand-paradoxon. Ezekben az esetekben az „interpretáció" azt jelenti, hogy a valószínűségszámításnak, mint formális nyelvnek, egy másik formális nyelven belüli modelljét adjuk meg, anélkül, hogy ezt az interpretációt a valóság tényei a legcsekélyebb mértékben is korlátoznák. Nevezzük ezt *matematikai interpretációnak*.

3. Ami minket filozófiai szempontból érdekel, az nyilván az a kérdés, hogy mi az a valószínűség. És erre a kérdésre nem kaphatunk választ másból, csak a reális interpretációkból, vagyis abból, ahogyan ezt a fogalmat a valóság leírásában alkalmazzuk.

4. Nem csak a valószínűség értelmezésének kérdése merül fel, hanem az is, hogy mi az esemény fogalma. Megint csak világo-

san kell látni, hogy erre a kérdésre is csak a reális interpretációk keretében kaphatunk választ. Hiszen filozófiai szempontból teljesen irreleváns, hogy esetleg felmutatunk valamilyen matematikai elméletet, amelyben egy megfelelő Boole-háló elemeit „eseményeknek” fogjuk nevezni.
Az esemény fogalmát illetően sokat vitatott kérdés, hogy a valószínűséget egy egyedi eseményhez, vagy egy eseménytípushoz köthetjük.

5. Végül felmerül az is, hogy valószínűség – feltéve, hogy nem 0 vagy 1 az értéke – értelmezhető-e egy determinisztikus világban, vagy pedig kizárólag egy nem determinisztikus világ tartozéka.

Klasszikus interpretáció

A *klasszikus interpretáció* Laplace-tól ered. Eszerint

> *Egy esemény valószínûsége nem más, mint a kedvezõ esetek számának és az összes, egyformán valószínû esetek számának aránya*.

Laplace egy további elvet vezetett be, amely azt hivatott eldönteni, hogy miket tekintsünk „egyformán valószínû” eseteknek. Az ún. *pártatlanság elve* szerint

> *Két lehetõséget akkor tekintünk egyformán valószínûnek, ha nincs semmi okunk egyiket a másikkal szemben preferálnunk.*

Általános felfogás szerint Laplace valószínûség-értelmezésének gyenge pontja az „egyformán valószínû esetetek” fogalma, illetve a pártatlanság elve. A helyzet ennél azonban bonyolultabb. Élesen különbséget kell tennünk aközött, hogy a laplace-i definíciót reális vagy matematikai interpretációnak tekintjük. Mint látni fogjuk, ha a laplace-i definíciót a valószínûség matematikai interpretációjaként használjuk, akkor semmiféle probléma nincs, az irodalomból ismert „paradoxonok" teljesen félrevezetõek. Ezzel szemben, a Laplace-féle klasszikus értelmezés, mint a valószínûség reális interpretációja sokkal alapvetõbb problémákat vet fel, és végsõ soron teljesen tarthatatlan.

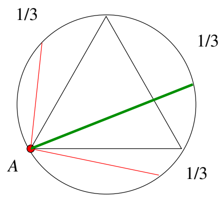
A laplace-i értelmezés matematikai interpretációként való felfogására példa a sokat idézett Bertrand-paradoxon. A vizsgált probléma a következõ:

> *Adott egy kör. Határozzuk meg mi a valószínûsége annak, hogy a kör egy véletlenszerûen kiválasztott húrja hosszabb, mint a körbe írt szabályos háromszög oldala.*

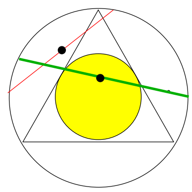
Az állítólagos „paradoxon", abban áll, hogy a feladatnak több különbözõ megoldása van, s azok eltérõ eredményre vezetnek. Nézzünk meg ezek közül kettõt:

ELSÕ MEGOLDÁS Véletlenszerûen ki kell választanunk két pontot a körön és vennünk kell az õket összekötõ húr hosszát. Tehát, az egyetlen fontos dolog a második pont helyzete a körön. Rögzítsük ezért az elsõ pontot, mondjuk *A*-ban (4. ábra). A második pont a kör három egyforma hosszúságú ívére eshet, melybõl kettõ olyan, hogy az oda befutó húr hossza rövidebb, mint a háromszög oldala. A kérdezett valószínûség tehát $\frac{1}{3}$.

MÁSODIK MEGOLDÁS A húrt egyértelmûen meghatározza középpontjának helye a kör belsejében (5. ábra). Akkor hosszabb, mint a beírt szabályos háromszög oldala, ha a középpontja a kisebbik kör belsejében van. A kisebbik kör sugara a nagy kör sugarának fele, területe ezért a nagy kör területének negyede. Következésképpen a kérdezett valószínûség $\frac{1}{4}$.



4. Figure. *Rögzítsük az elsõ pontot A-ban. A második pont a kör három egyforma hosszúságú ívére eshet, melybõl kettõ olyan, hogy az oda befutó húr hossza rövidebb, mint a háromszög oldala*



5. Figure. *A húrt egyértelmûen meghatározza középpontjának helye a kör belsejében. Akkor hosszabb, mint a beírt szabályos háromszög oldala, ha a középpontja a kisebbik kör belsejében van*

Hasonló „paradoxonok” tucatjait ismeri a valószínûségszámításról szóló irodalom. Ezek természetesen nem igazi paradoxonok, tehát nem tükröznek tényleges ellentmondásokat. Tévesen azt szokás mondani, hogy ezekben a példákban – paradox módon – egy esemény valószínûsége „függ attól, hogy a feladatot hogyan paraméterezzük”. Az ilyen vélekedés azonban súlyos konfúziót tükröz úgy a „paraméterezés” mibenlétét, mint a matematika és a valóság viszonyát illetõen. A „paraméterezés” ugyanis a következõt jelenti: Ugyanazon az eseményalgebrán különbözõ valószínûségi mértéket definiálhatunk, s ezzel különbözõ Kolmogorov-féle valószínûségi modelleket kapunk. Tekintsünk most egy, az algebrán értelmezett valószínûségi változót. A valószínûségi mérték egyértelmûen indukál egy valószínûségeloszlást az adott változóra nézve. A feladat egy „paraméterezése” nem mást jelent, mint egy olyan megszorítást a valószínûségi mértékre, hogy az indukált valószínûségeloszlás az adott változóra nézve legyen egyenletes. Különbözõ „paraméterezések” tehát különbözõ valószínûségi mértéket engedhetnek meg (esetleg fixálnak) ugyanazon az eseményalgebrán. Vagyis nem ugyanazon valószínûségi modellen belül lesz ugyanannak az eseménynek különbözõ a valószínûsége.

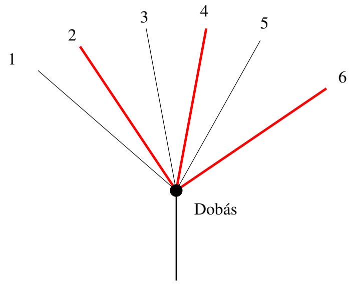
Mármost a kérdés az, baj-e, hogy a pártatlanság elvének alkalmazásával több, különbözõ valószínûségi modellhez jutunk. Válaszunk az, hogy nem. Fontos azonban látni, hogy itt a valószínûségelmélet fogalmainak egy matematikai interpretációjáról van szó. Mert hogyan kell értenünk azt az eseményt, hogy <a kör húrjának hossza nagyobb, mint a beírt szabályos háromszög oldalának hossza>? Ezek nem valóságos események. Nincs a világnak olyan állapota, a dolgoknak olyan állása, amelyet úgy interpretálhatnánk, hogy ez az „esemény” bekövetkezik. Ha ez így van, és itt csupán „matematikai eseményekrõl” van szó, akkor ezt pontosan a következõképpen kell értenünk: Adott egy matematikai elmélet M_1, egy formális nyelv, amely tartalmaz olyan szavakat, mint <háromszög>, <kör>, <húr>, <hosszúság>, stb. Adott egy másik matematikai elmélet, M_2, egy másik formális nyelv, amely olyan szavakat tartalmaz, mint <esemény>, <valószínûség>, és így tovább. E két elméletnek egymáshoz semmi köze nincs. Nem értelmes kérdés az euklideszi geometriában, hogy <mi a valószínûsége annak, hogy...?>, mint ahogyan nem lehet tudni a valószínûségelméletben, hogy mi az a <háromszög>. A feladat kitûzõi tehát arra szólítanak fel, hogy *alkossunk meg egy olyan új M_3 matematikai elméletet*, egy új formális nyelvet, amely M_1-nek is és M_2-nek is tartalmazza egy modelljét, és minden mást, amely szükséges ahhoz, hogy a feladatban szereplõ mondatok értelmesek legyenek. A Bertrand-féle feladatban ezt úgy tehetjük meg legegyszerûbben, hogy megadjuk M_2-nek egy modelljét M_1-ben, például a következõképpen: *Legyen* Ω az adott kör húrjai középpontjainak halmaza. A Σ eseményalgebra *legyen* Ω olyan részhalmazainak σ-algebrája, melyek mérhetõek a síkon értelmezett Lebesgue-mérték szerint. A valószínûségi mérték pedig *legyen* a

Lebesgue-mérték maga, normálva a kör területével. Természetesen, rengeteg más M_3 is megalkotható. A Laplace-féle „pártatlanságelv” legjobb esetben is csak úgy értelmezhetõ, mint jó tanács ilyen elméletek megalkotásához, és az, hogy alkalmazásával nem jutunk egyértelmû M_3 konstrukciókhoz, semmiféle problémát nem jelent.

Problémák egész sorát veti fel azonban, ha a laplace-i definíciót reális interpretációként értelmezzük. A valószínûség klasszikus értelmezésének prototípusaként szokás említeni a „szimmetrikus dobókocka” esetét: „6 egyformán valószínû kimenetel lehetséges. Ebbõl 3 esetben valósul meg a <páros dobás> esete, tehát a <páros dobás> valószínûsége $\frac{3}{6}$.”

Reális interpretációról lévén szó, a valószínûségelméleti fogalmakat a valóság leírását szolgáló, tehát empirikus nyelvre redukálható fogalmakkal kell reprezentálnunk. De hogyan is lehet ezt az empirikus nyelvre történõ redukciót végigvinni? Úgy tûnik, sehogy. Mert vagy olyan valószínûségfogalomhoz jutunk, amely érzéketlen a valóság tényeire, vagy pedig burkoltan hivatkoznunk kell a valószínûség más, például frekventista interpretációjára.

Vizsgáljuk meg közelebbrõl a „szimmetrikus dobókocka” esetét. A klasszikus definíció szerint tehát az a lényeg, hogy a kocka feldobásának pillanatában 6 különbözõ dolog történhet. Vagyis a feldobás pillanatában az univerzum története 6 különbözõ ágon folytatódhat (6. ábra). Pontosabban a lehetséges történeteket 6 különbözõ osztályba sorolhatjuk. (Nem feltétlenül gondolunk itt objektív modalitásra. Az ábrán látható elágazás szimbolizálhat egy episztemikus modalitást is.) Ezek közül három ágon valósul meg a <páros dobás> esete. A valószínûség tehát ezeknek az lehetõségeknek a számaránya. Semmilyen egyéb részlete a világ tényeinek nem számít. Ha például a kocka cinkelt, akkor is 6 lehetséges kimenetele van a dobásnak és ebbõl három vezet páros számhoz, tehát a páros dobás valószínûsége – ezek szerint – továbbra is $\frac{3}{6}$.



6. Figure. *A kockadobás pillanatában az univerzum története 6 különbözõ ágon folytatódhat*

De hiszen – mondhatná erre valaki – a cinkelt kocka már nem szimmetrikus!

Miért ne lenne – válaszolhatnánk. Továbbra is mindegyik szám dobásának valószínûsége – a laplace-i definíció értelmében – $\frac{1}{6}$!

Igen – hangzana máris a válasz –, de a kocka egyéb tulajdonságaiban, pl. a tömegeloszlásában, nem szimmetrikus.

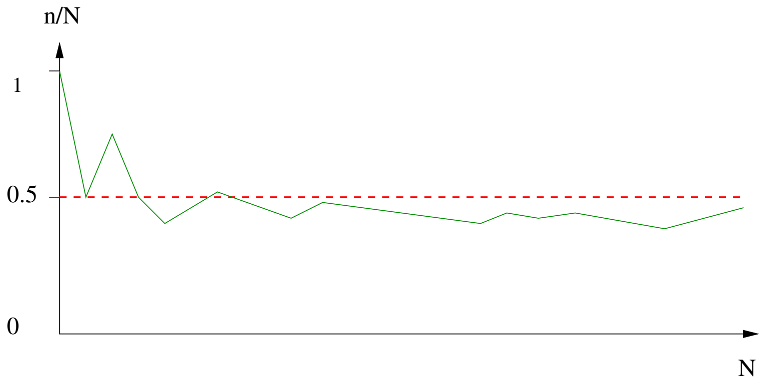
Hát persze, hogy nem – válaszolnánk. A nem cinkelt kocka sem teljesen szimmetrikus minden tulajdonságában, hiszen akkor nem is lehetne tudni, mikor melyik oldalára esik. Például más és más szám van a különbözõ oldalaira festve. Vagyis vannak releváns, és vannak irreleváns aszimmetriák?

Igen! – jönne a válasz. Olyan aszimmetria számít csak, amelyik befolyásolja a hat lehetséges kimenetel valószínûségét.

Milyen „valószínûségét”? – kérdeznénk ekkor, és képzeletbeli vitapartnerünk nem tehetne mást, mint hogy valami olyasmire hivatkozik, hogy „empirikusan megfigyelhetõ tény, hogy a cinkelt kocka gyakrabban esik egyik oldalára, mint a másikra”, vagyis a „valószínûség” szónak valamely nyilvánvalóan nem laplace-i jelentésére utalna.

Relatív gyakoriság interpretáció

Arisztotelész nevéhez szokás kötni, noha minden bizonnyal sokkal régebbi, az embernek azt a hétköznapi nyelvhasználatban tükrözõdõ valószínûség-értelmezését, hogy tudniillik „az valószínû, ami gyakran megtörténik”. A valószínûség relatív gyakoriság interpretációjának elsõ matematikailag is precíznek tekinthetõ megfogalmazását az angol logicista John Venn adta meg (1866). Alapgondolata a következõ volt. Dobjunk fel egy érmét sokszor egymás után. Az eredményeket mutatja a 7. ábra. A függõleges tengelyen a <Fej> esemény *relatív gyakorisága* van feltüntetve, vagyis az $\frac{n_N}{N}$ hányados, ahol n_N a <Fej> esemény bekövetkezésének a számát, N pedig az érme feldobásának a számát jelöli.



7. Figure. *A <Fej> esemény relatív gyakorisága a feldobások számának függvényében*

A relatív gyakoriság interpretáció feltételezi, hogy a relatív gyakoriságok $\left\{\frac{n_N}{N}\right\}_{N=1,2,\ldots}$ sorozatának létezik limesze, és ez a limesz definíció szerint a szóban forgó esemény valószínûsége.

Általánosabb megfogalmazásban, tekintsük eseményeknek egy $\mathcal{A}$ Boole-hálóját. *Klasszikus igazságérték függvénynek* nevezünk egy olyan $u : \mathcal{A} \rightarrow \{0,1\}$ leképezést, amely eleget tesz a következõknek:

$$\begin{aligned} u(\emptyset) &= 0 \\ u(\neg A) &= 1 - u(A) \\ u(A \wedge B) &= u(A)\, u(B) \end{aligned}$$

Az igazságérték függvény intuitív jelentése nyilvánvaló: 0 értéket vesz fel, ha a szóban forgó esemény nem következik be, és 1 értéket vesz fel, ha az esemény bekövetkezik.

Végezzünk el egy kísérletet sokszor egymás után. A kísérlet minden egyes elvégzésekor az $\mathcal{A}$ eseményalgebra elemei vagy bekövetkeznek, vagy nem, vagyis, ebbõl a szempontból, a dolgok állása egy alkalmas klasszikus igazságérték függvénnyel írható le. Legyenek az egymást követõ kísérletekhez tartozó igazságérték függvények

$$u_1, u_2, \ldots u_N, \ldots \tag{12}$$

Vezessük be az elsõ N kísérlet alapján kiszámolt relatív gyakoriság függvényt a következõképpen:

$$\nu_N : A \in \mathcal{A} \mapsto \nu_N(A) = \frac{1}{N}\sum_{i=1}^{N} u_i(A) \in [0,1] \tag{13}$$

Könnyen belátható, hogy ha a

$$\nu_1, \nu_2, \ldots, \nu_N, \ldots \tag{14}$$

függvénysorozat pontonként konvergens, akkor a $p = \lim_{N\to\infty} \nu_N$ egy, a kolmogorovi axiómákat kielégítõ valószínûség $\mathcal{A}$ eseményalgebrán.

A fenti konstrukcióval elégedettek lehetünk, ha az volt a célunk, hogy a valószínûség fogalmát más matematikai fogalmakra redukáljuk, vagyis a valószínûség egy matematikai reprezentációját kívántuk megadni.

Nehézségek akkor merülnek fel, ha a fenti konstrukciót reális interpretációnak tekintjük. Ha a (12) sorozatot valóságos kísérletsorozat kimeneteleit leíró igazságérték sorozatnak gondoljuk el, akkor ezt a sorozatot, közvetlenül vagy közvetve, valamilyen valóságos folyamat kimenetele determinálja (vagy legalábbis megszorítja), és semmi sem garantálja, hogy a megfelelõ (14) sorozat konvergens. A pénzfeldobás példájánál maradva, hajlamosak vagyunk azt gondolni, hogy a feldobás kimenetelének véletlenszerûsége garantálja azt, hogy a relatív gyakoriság konvergál. Könnyû azonban olyan random sorozatot generálni, melyre a relatív frekvencia nem konvergens. Vegyük a <0>-ból és <1>-bõl álló sorozatot, melyet a következõképpen generálunk: Feldobunk két érmét. Ha az eredmény <Fej> &

<Fej>, akkor írjunk egy <1>-et, minden más esetben írjunk <0>-t. Ismételjük ezt addig, amíg az <1> relatív frekvenciája nem kisebb, mint 0.4. Ekkor cseréljük fel a <0> és <1> szerepét, vagyis <Fej> & <Fej> esetén <0>-t, más esetben <1>-et írjunk. Egészen addig, amíg a relatív frekvencia nagyobb nem lesz 0.6-nál. Ekkor ismét cserélünk, és így tovább. Világos, hogy az így nyert sorozatban a <0> és az <1> relatív frekvenciája oszcillálni fog 0.4 és 0.6 között, tehát egyik sem konvergál. Ha tehát a valószínûség nem más, mint a relatív frekvencia határértéke, akkor ez azt jelenti, hogy ebben a kísérletben a $p(<1>)$ valószínûség nem létezik? Intuíciónk szerint nem ez a helyzet, hanem egyszerûen csak arról van szó, hogy $p(<1>)$ értéke a folyamat során egyszer 0.25 majd egy ponton 0.75-re változik, aztán megint 0.25 lesz, és így tovább.

A fenti példából is kitûnik a relatív gyakoriság interpretáció egyik sokat kritizált vonása, hogy tudniillik nem képes minden esetben számot adni az individuális események valószínûségérõl. Elõzõ példánkban minden egyes kísérletben egy bizonyos valószínûséget tulajdoníthatunk annak, hogy az adott kísérletben az <1> esemény bekövetkezik, míg az <1> esemény relatív gyakoriságának nem is létezik limesze a kísérlet sokszori megismétlése során. Ez az „egy bizonyos valószínûséget tulajdoníthatunk” természetesen csak intuitíve értendõ, hiszen a relatív gyakoriság értelmezés alapján ez a valószínûség nem definiált.

A relatív gyakoriság interpretáció reális interpretációként való értelmezésének egy másik nehézsége az a matematikai tény, hogy egy végtelen sorozat limesze teljesen független a sorozat tetszõleges, de véges hosszúságú elejétõl. Ez azt jelenti, hogy – szemben a mindennapos gyakorlattal – egy véges mintán leszámolt relatív gyakoriság semmilyen logikai összefüggésben nem áll a szóban forgó esemény valószínûségével. A valószínûség mibenlétének feltárásában nem nélkülözhetjük annak megértését, hogyan lehetséges mégis, hogy a véges mintákon vett relatív gyakoriságok jó közelítéssel megegyeznek a megfelelõ valószínûségekkel.

Propensity interpretáció

A relatív gyakoriság interpretáció tehát ismétlõdõ események egy végtelen sorozatához rendel valószínûséget, és nem egy individuális eseményhez. Popper[41] megkísérelte a valószínûség egy olyan értelmezését megadni, amely értelmessé tenné, hogy egy-egy partikuláris esemény valószínûségérõl is beszélhessünk. Ez a *propensity* (hajlam) interpretáció. A propensity interpretáció lényege, hogy feltételezi, amikor a dobókockát feldobjuk, akkor a kockának és az egész valóságos/fizikai szituációnak együtt van egy objektív tulajdonsága, nevezetesen az arra való hajlamának mértéke, hogy 6-os legyen az eredmény. Ennek a hajlamnak a számszerû mértékét fejezi ki, amikor azt mondjuk, hogy a $\frac{1}{6}$ a valószínûsége annak, hogy az eredmény 6-os lesz.

A propensity interpretációval szemben általában azt az ellenérvet szokás felhozni, hogy nem tekinthetõ a valószínûség teljeskörû értelmezésének. Vannak olyan értelmesnek tekintett valószínûségek, amelyekhez semmiféle propensity nem társítható. Tekintsük a következõ példát:[42] Egy frisbee-gyárnak van két gépe, egy régi és egy új. Az új gép napi 800, a régi 200 frisbeet gyárt. Az új gépen gyártott termékek 1%-a, a régi gépen gyártottak 2%-a selejt. Találunk egy hibás frisbeet. Mi annak a valószínûsége, hogy ez a frisbee az új gépen lett gyártva? A Bayes-szabály alkalmazásával ezt könnyen kiszámíthatjuk:

$$\begin{aligned} p(U) &= 0.8 \\ p(\neg U) &= 0.2 \\ p(S|U) &= 0.01 \end{aligned}$$

[41] Popper 1960.

[42] Earman és Salmon 1992.

$$
\begin{aligned}
p(S|\neg U) &= 0.02 \qquad (15)\\
p(U|S) &= \frac{p(S|U)p(U)}{p(S|U)p(U)+p(S|\neg U)p(\neg U)}\\
&= \frac{0.01\times 0.8}{0.01\times 0.8+0.02\times 0.2} = 0.66
\end{aligned}
$$

Mármost ennek a valószínûségnek van értelme, azonban nem értelmezhetõ propensityként. Miféle hajlama lehet a hibás frisbeenek arra, hogy õ fél nappal ezelõtt az új gépen legyen volt gyártva? – hangzik a szokásos ellenvetés.

Vegyük azonban észre, hogy a propensity interpretációval szembeni fenti méltatlankodás nem egészen helytálló. Hiszen nyilvánvaló, hogy a $p(U|S)$ valószínûség, éppúgy, mint a (15) formulában szereplõ többbi valószínûség, nem a selejtes frisbeet magát, hanem az õ gyártásának egész folyamatát jellemzi, és hogy mi történik azzal a darab mûanyag masszával, amely a technológiai lánc elején bemegy, azt elsõsorban nem az õ hajlamai, hanem a gyártási folyamatban résztvevõ berendezések hajlamai határozzák meg. A kondicionális valószínûségnek pedig bármely más interpretációt véve alapul sincs több jelentése, mint a (10) formulában szereplõ hányados. Ezt azért kell hangsúlyoznunk, mert a fenti kifogást a következõképpen is meg szokták fogalmazni: Tekintsünk két egymással kauzális kapcsolatban álló eseményt, A-t és B-t. A kauzális kapcsolat következtében, mondjuk, $p(A|B) > p(A)$. Mármost, hangzik az érv, ezt a kondicionális valószínûséget gond nélkül lehet propensityként értelmezni, hiszen „van értelme az ok abbéli hajlamáról beszélnünk, hogy az okozatot létrehozza”. Ám ezzel szemben – mondják – $p(B|A)$ nem értelmezhetõ propensityként, hiszen „értelmetlen dolog az okozat abbéli hajlamáról beszélnünk, hogy õt ilyen vagy olyan ok létrehozza”.

Nyilvánvaló, hogy ebben az érvelésben a kondicionális valószínûség fogalma olyan plusz tartalommal van megterhelve, mellyel az nem rendelkezik – a valószínûség fogalmának bármelyik interpretációját is vesszük alapul. A kondicionális valószínûség fogalmát sokan félreértik. Érdemes ezekkel a félreértésekkel egy rövid kitérõ erejéig bõvebben foglalkozunk.

A $p(A|B)$ kondicionális valószínûség fogalma sem többet sem kevesebbet nem jelent, mint amit a definíciója állít, vagyis a $\frac{p(A \wedge B)}{p(B)}$ hányadost. Tehát, hogy hogyan aránylik az A és B események együttes bekövetkezésének valószínûsége az egyik, történetesen a B esemény bekövetkezésének valószínûségéhez. Ezzel szemben,

1. *$p(A|B)$ nem jelenti azt, hogy mekkorára változik az A esemény valószínûsége, akkor, amikor B esemény bekövetkezik.* A dobókockával dobunk a t_0 pillanatban. Egy másodperccel késõbb a t_1 pillanatban megszületik az eredmény: a kettest dobtuk, tehát egyidejûleg bekövetkezett a <páros> esemény is. Az eldobás elõtti pillanatban $p_{t_0}(<4>) = \frac{1}{6}$ és $p_{t_0}(<\text{páros}>) = \frac{1}{2}$. Kérdés, mekkora a négyest dobás valószínûsége a t_1 pillanatban, vagyis, amikor a <páros> esemény bekövetkezett? A válasz – függetlenül attól, hogy melyik interpretációról van szó – $p_{t_1}(<4>) = 0$. És ez nem egyenlõ a $p_{t_0}(<4> \mid <\text{páros}>) = \frac{1}{3}$ kondicionális valószínûséggel.

2. *$p(A|B)$ nem jelenti azt, hogy mekkora az A esemény valószínûsége akkor, ha a rendszer (a világ) olyan módon van preparálva (olyan állapotban van), hogy az garantálja, hogy a B esemény egy valószínûséggel bekövetkezik.* Ha ez lenne a kondicionális valószínûség értelme, akkor nem is lenne egyértelmûen definiálva, és nem is lenne feltétlenül egyenlõ $\frac{p(A \wedge B)}{p(B)}$-val! A dobókocka példájánál maradva, preparáljuk a kockát úgy, hogy $p(<2>) = 1$ legyen. Ezzel teljesítettük a $p(<\text{páros}>) = 1$ feltételt. Ilyenkor $p(<4>) = 0$, tehát ezek szerint, „$p(<4> \mid <\text{páros}>) = 0$”. Most preparáljuk a kockát úgy, hogy $p(<4>) = 1$. Ekkor is teljesül a $p(<\text{páros}>) = 1$ feltételt, és most „$p(<4> \mid <\text{páros}>) = 1$”.

Az ilyen és hasonló tévedések alapját gyakran annak az egyszerû *matematikai* ténynek a félreértése képezi, hogy tudniillik tetszõleges $(\mathcal{A}, p)$ valószínûségi modell esetében, minden $p(Y) \neq 0$ esetén, a

$$p' : X \in \mathcal{A} \mapsto p'(X) = p(X|Y)$$

egy, a Kolmogorov-axiómákat kielégítõ *új* p' valószínûségi mértéket definiál az $\mathcal{A}$ eseményalgebrán, és így egy *új* $(\mathcal{A}, p')$ valószínûségi modellhez jutunk, melyben $p'(Y) = 1$.

A félreértések másik forrása, hogy figyelmen kívül hagyják, a $p(A|B) > p(A)$ korreláció csupán a szükséges, de nem elégséges feltétele annak, hogy az események között a „*B* oka *A*-nak” típusú kauzális kapcsolat álljon fenn (lásd a 10. fejezetet).

Ehhez a kérdéskörhöz kapcsolódik még a **6.0.0.21.** és a **6.0.0.21.** pont, melyekben a Bayes-szabály és a szubjektív valószínûség kapcsolatát fogjuk megvizsgálni.

A propensity interpretációval szemben felhozott szokásos kifogások tehát nem mondhatók megalapozottaknak. Ezzel szemben nem szokták említeni a propensity fogalmával kapcsolatban felmerülõ lényegesebb problémát. Nevezetesen, hogy a propensity nem egy másodlagos, származtatott tulajdonsága a vizsgált objektumoknak, vagyis a propensity interpretáció a valószínûség fogalmát nem valamilyen már ismert fogalomhoz köti, nem a világ objektumainak már értelmezett, és empirikusan is megragadott tulajdonságaiból vezeti le, hanem egy új tulajdonság létezését állítja, olyan tulajdonságét, amely egy számértékkel fejezhetõ ki. De hogy ez a számérték pontosan hogyan határozódik meg, hogyan kötõdik más, mérhetõ fogalmakhoz, arra nézve semmiféle magyarázatunk nincs. Ad absurdum, semmiféle fogódzónk nincs arra nézve, hogyan tesztelhetnénk empirikusan azt a kijelentést, hogy a feldobott egyforintosnak $propensity = \frac{1}{2}$-e van <Fej>-jel felfelé landolni az asztalon.

Szubjektív interpretáció

A szubjektív interpretáció úgy értelmezi a valószínûséget, mint annak a mértékét, amellyel egy egyén hisz valamely esemény bekövetkezésében. Azt gondolhatnánk, hogy ennélfogva a valószínûségek értéke bármi lehet, és az események valószínûsége semmilyen törvényszerûséget nem kell, hogy mutasson. Látni fogjuk, hogy ez tévedés. Tárgyalásunkat azonban megint két részre kell bontanunk, hiszen a szubjektív valószínûség elmélete a valószínûség reális, illetve matematikai interpretációjaként is értelmezhetõ.

Alapjában véve a szubjektív interpretáció – meglepõ módon – egy reális interpretáció, hiszen a valóságos világban bekövetkezõ eseményekre vonatkozóan vezeti be a szubjektív valószínûség fogalmát, egy új mennyiséget, melynek számértéke valóságos személyek hitének mértékét fejezi ki valóságos események bekövetkezésével kapcsolatban. Vagyis, amikor arról beszélünk, hogy Kovács úr a Rózsi nevû kanca gyõzelmében 0.9 mértékben hisz, akkor egy valóságos lóról, annak a célfotón látható valóságos gyõzelmérõl és egy valóságos személy valamely valóságos tulajdonságáról van szó. Legyen az a szubjektív interpretációt propagálók gondja, hogy valami bõvebbet mondjanak arról – feltehetõen Kovács úr pszichológiai analízise alapján –, hogy mi határozza meg ennek a szubjektív valószínûségnek a számértékét. Mindenesetre, nem mondhatjuk, hogy ezek a szubjektív valószínûségek tetszõleges értéket felvehetnek, mint ahogyan azt sem tudhatjuk, vajon kielégítik-e a kolmogorovi-axiómákat.

A szubjektív valószínûség, mint egy valóságos személy egy valóságos esemény bekövetkezésében való hitének számszerûsített mértéke, elvben, empirikusan megragadható fogalmakhoz kötõdik. Hogy hogyan, arra vonatkozóan nem tudunk semmit. Szokás azonban azzal a feltételezéssel élni, hogy egy racionálisan gondolkodó ember fogadásaiban olyan arányban fogad egy esemény bekövetkezésére, mint amennyi a szóban forgó eseményre vonatkozó szubjektív valószínûsége. Vagyis, ha én 2/3 mértékben hiszek egy esemény bekövetkezésében, akkor 3 a 2-höz arányban vagyok hajlandó fogadni arra, hogy az esemény bekövetkezik. *E feltételezés mellett*[43] érdekes eredményt sikerült bizonyítani: A fogadások világában *Dutch book*nak nevezik fogadásoknak egy olyan együttesét, amely arra vezet, hogy a fogadó mindenképpen veszít, bárhogyan alakuljon is a szóban forgó játék eredménye. Megmutatható (Dutch book-tétel), hogy egy racionális fogadó fogadásai akkor és csak akkor nem alkotnak Dutch bookot, ha szubjektív valószínûségei kielégítik a Kolmogorov-axiómákat. Az elégségesség bizonyítása nem egyszerû, a szükségesség belátása azonban triviális. Vegyük például a Kolmogorov-axiómák azon egyszerû következményét, hogy $p(A) + p(\neg A) = 1$. Tegyük fel, hogy valaki figyelmen kívül hagyja ezt a szabályt, és 2/3 valószínûséget tulajdonít annak, hogy az érme feldobásának kimenetele <Fej> lesz, és 2/3 valószínûséget tulajdonít annak is, hogy <Írás> lesz. Ennek megfelelõen az egyik ablaknál 2:1 arányban fogadást köt a <Fej>-re, és a másik ablaknál 2:1 arányban fogad az <Írás>–ra is. Ezek a fogadások egy Dutch bookot képeznek, ugyanis, ha az eredmény <Fej>, akkor a fogadó nyer 1 forintot az elsõ fogadása alapján és veszít 2 forintot a másik fogadása miatt. Ha az eredmény <Írás>, akkor veszít 2 forintot és nyer 1 forintot. Tehát mindenképpen veszít.

A Dutch book argumentumban felsejlik a szubjektív interpretáció matematikai interpretációként való felfogása is. Egy matematikai interpretáció esetében, a valószínûségelméleti fogalmakat valamely más matematikai elmélet fogalmaival reprezentáljuk. A szubjektív interpretáció esetében ez az elmélet a fentiekben leírt fogadási szituációt absztrakt módon modellezõ játékelmélet. A Dutch book-tétel értelemszerûen ekkor is érvényben van, tehát a szubjektív

[43]És természetesen a világ mûködésére vonatkozó néhány triviálisnak gondolt, kimondatlan feltételezés mellett...

valószínûségek ilyenkor is kielégítik a Kolmogorov-axiómákat.

A szubjektív interpretációval kapcsolatban érdemes néhány szót szólnunk az úgynevezett *bayesianizmusról*. A (10) Bayes-szabályból, valamint a (11) teljes valószínûség tételbõl azonnal levezethetõ a következõ matematikai összefüggés:

$$p(A|B) = \frac{p(B|A)p(A)}{p(B|A)p(A) + p(B|\neg A)p(\neg A)} \tag{16}$$

Tekintsük a következõ példát: a szomszéd szobában valaki egy kockát dobál és jelenti nekünk az eredményt. Nem tudjuk, hogy a kocka egy szabályos dobókocka-e, vagy úgy van cinkelve, hogy csak páros számot lehet vele dobni (de az egyszerûség kedvéért tegyük fel, hogy a kettõ közül valamelyik igaz). Jelöljük H-val azt a hipotézist, hogy a kocka cinkelt. Legyen $p_{t_0}(H) \in (0,1)$ egy személy szubjektív valószínûsége a t_0 pillanatban arra vonatkozóan, hogy a H hipotézis igaz, és általában jelöljük p_{t_0}-val az adott személy szubjektív valószínûségeit a t_0 pillanatban. Nyilvánvaló,[44] hogy $p_{t_0}(<2>|H) = p_{t_0}(<4>|H) = p_{t_0}(<6>|H) = \frac{1}{3}$, $p_{t_0}(<5>|H) = 0$ és $p_{t_0}(<i>|\neg H) = \frac{1}{6}\; i = 1,2,\ldots 6$. Most tegyük fel, hogy azt a jelentést kapjuk, hogy a kockadobás eredménye <6>. Mennyi a H hipotézis valószínûsége ezen új információ birtokában az ezt követõ t_1 pillanatban? A (16) összefüggés felhasználásával,

$$\begin{aligned} p_{t_1}(H) &= p_{t_0}(H|<6>) = \frac{p_{t_0}(<6>|H)p_{t_0}(H)}{p_{t_0}(<6>|H)p_{t_0}(H) + p_{t_0}(<6>|\neg H)p_{t_0}} \\ &= \frac{\frac{1}{3}p_{t_0}(H)}{\frac{1}{3}p_{t_0}(H) + \frac{1}{6}(1 - p_{t_0}(H))} \end{aligned}$$

Hasonlóan, ha arról értesülünk, hogy az eredmény <5>,

$$p_{t_1}(H) = p_{t_0}(H|<5>) = \frac{p_{t_0}(<5>|H)p_{t_0}(H)}{p_{t_0}(<5>|H)p_{t_0}(H) + p_{t_0}(<5>|\neg H)p_{t_0}}$$

[44]Pontosabban, egyáltalán nem nyilvánvaló. Lásd a bayesianizmussal kapcsolatos pontban kifejtett kritikai megjegyzéseket. Az itteni gondolatmenetet tekintsük úgy, mint szabad idézését egy, a téma irodalmából vett tipikus kifejtésnek.

$$= \frac{0}{0+\frac{1}{6}(1-p_{t_0}(H))} = 0$$

vagyis a hipotézisnek ellentmondó esemény észlelése (Popper nagy megelégedésére[45]) a hipotézist azonnal nulla valószínûségûvé teszi.

Tegyük most fel, hogy a kocka valóban cinkelt. Ha például Eszter szubjektív valószínûsége a H hipotézist illetõen 0.01, akkor a <6> eseményrõl értesülve ez a szubjektív valószínûség a következõ lesz:

$$p_{t_1}(H) = p_{t_0}(H| < 6 >) = \frac{\frac{1}{3}0.01}{\frac{1}{3}0.01+\frac{1}{6}(1-0.01)} \approx 0.02$$

Ha most ismét a hipotézissel összhangban álló, mondjuk <2> eseményrõl kapunk jelentést, akkor a (17) formula ismételt alkalmazásával Eszter hitének mértéke (szubjektív valószínûsége) a H hipotézis igazságában a következõre változik:

$$p_{t_2}(H) = p_{t_1}(H| < 2 >) = \frac{\frac{1}{3}0.02}{\frac{1}{3}0.02+\frac{1}{6}(1-0.02)} \approx 0.04$$

és így tovább. A századik páros szám után ez a valószínûség nagy pontossággal 1.

Ha most például Iván szubjektív valószínûsége a H hipotézist illetõen 0.7, akkor az elsõ <6> esemény után ez $\frac{\frac{1}{3}0.7}{\frac{1}{3}0.7+\frac{1}{6}(1-0.7)} \approx 0.82$-re változik, és a századik páros szám után ez is sok jegy pontossággal 1. Vagyis, függetlenül attól, hogy kinek milyen elõzetes várakozásai vannak, a bekövetkezõ események, az empirikus tények, a hipotézist maximálisan konfirmálják.

Érdemes itt két dolgot megfigyelnünk, melynek nagy jelentõsége van az induktív „következtetés” Hempel-féle kritikája,[46] és általában a tudományos hipotézisek empirikus megalapozhatóságára

[45] Popper 1963.

[46] Hempel 1965.

vonatkozó bírálatok, és az ezekre adott bayesianus válasz[47] tekintetében:

1. A $\frac{p(H|<6>)}{p(H)}$ hányados jó mérõszáma annak, hogy milyen mértékben nyer a H hipotézis megerõsítést a <6> evidencia észlelése által. Vegyük észre, hogy ez a hányados az egymást követõ megerõsítések során fokozatosan csökken, és egyhez tart: a kilencvenkilencedik megerõsítéshez képest a századik már nem jelent lényeges emelkedést a H hipotézis valószínûségében.

2. Tegyük fel, hogy egy E esemény (evidencia) következik a H hipotézisbõl, tehát $p(E|H) = 1$. Ekkor a (16) összefüggést a következõképpen is olvashatjuk:

$$\frac{p(H|E)}{p(H)} = \frac{1}{p(E)}$$

   Vagyis, a megerõsítés mértéke fordítottan arányos az E eseménynek a hipotézis feltételezése nélküli valószínûségével. Más szóval, egy hipotézist akkor tudunk jelentõsen megerõsíteni, ha a van egy olyan következménye, amely egyébként nagyon valószínûtlen volna, és ezt a következményt sikerül megfigyelnünk.

Ahogy mondani szokás, ez túl szép ahhoz, hogy igaz legyen! Az elsõ, és nagyon is lényeges probléma, amivel szembe találjuk magunkat – és a valószínûség interpretációja szempontjából is ez a fontos most számunkra –, hogy miként kell az olyan valószínûségeket értelmeznünk, mint $p(< 6 > |H)$, $p(H)$, stb., és honnan tudjuk azok értékét, valamint a változásukra vonatkozó törvényeket. Bár a fenti kifejtésben szubjektív valószínûségekrõl volt szó, a bayesianizmus, mint ismeretelméleti, a világról szóló (tudományos) teóriák empirikus konfirmációjára vonatkozó tan, ezeket a valószínûségeket

[47] Grünbaum 1976a,b.

általánosabb érvényûnek tûnik tekinteni, mint egyetlen partikuláris személy ilyen vagy olyan állítás igazságába vetett hitének mértékét. De legalábbis a szubjektív valószínûségek változására felírt (17) összefüggés egy mindenkire nézve egyformán érvényes törvényszerûség. Hanem akkor kinek a szubjektív valószínûségeirõl van szó? Az „emberé”, általában? Tehát az Úristen olyannak teremtette volna az embert, hogy koffeint itatva vele, felmegy a vérnyomása, és felmutatva neki egy evidenciát a (17) formula szerint megváltozik a szubjektív valószínûsége? Egyáltalán, honnan tudnak hirtelen a bayesianusok ilyen sokat az emberi elme viselkedésének általános, mindenkire egyformán érvényes törvényszerûségeirõl? És egyáltalán, mi köze van ezeknek a pszichológiai felfedezéseknek az episztemológiához?

Azt hiszem, itt valami egészen másról van szó. A bayesianizmus sohasem hivatkozik pszichológiai kísérletekre. Valamiért azonban – a pszichológiai ismereteinktõl függetlenül – *felteszi*, hogy egy személy szubjektív valószínûsége egy új evidencia megismerésével a (17) formula szerint változik meg. Mire alapozza a bayesianizmus a (17) összefüggést? Arra a feltevésre, hogy a szubjektív valószínûség bizonyos tulajdonságaiban ugyanúgy viselkedik, mint „más valószínûségek”, tehát, mint a relatív gyakoriság, vagy a propensity, vagy a klasszikus interpretáció szerint értelmezett valószínûség. Vagyis abból a feltételezésbõl indul ki, hogy a valószínûség többi interpretációjában teljesül, hogy egy A esemény valószínûsége egy B esemény bekövetkezése pillanatában $p(A)$-ról $p(A|B)$-re változik. Ha így lenne, akkor – a feltételezésnek megfelelõen – teljesülne a szubjektív valószínûségekre is, és akkor valóban fennállna a (17) összefüggés. Mint a **6.0.0.21.** pontban láttuk, szó sincs azonban arról, hogy $p(A)$-ról $p(A|B)$re változna az A valószínûsége B bekövetkezésének hatására. Tehát, *a bayesianizmus a Bayes-szabály félreértésére épül.*[48]

[48]Fontos ezt világosan látnunk például annak érdekében is, hogy helyesen értékelhessük a kortárs

Mielõtt befejeznénk a szubjektív interpretáció áttekintését, érdemes itt még egy kérdést tisztázni. Szokás a filozófiai irodalomban *objektív* és *episztemikus* valószínûségekrõl beszélni. Az egyik elnevezés arra utal, hogy a valószínûséggel jellemzett esemény objektíve indeterminisztikus, a másik arra, hogy csak számunkra tûnik annak, vagyis, hogy a szóban forgó jelenség számunkra megnyilvánuló valószínûségi jellege csupán tudásunk hiányából fakad. Fontos azonban, hogy ennek a felosztásnak semmi köze a fenti interpretációk szerinti „felosztáshoz”. Vagyis például a szubjektív valószínûség nem feltétlenül episztemikus. Éppúgy lehet beszélni arról, hogy milyen mértékben hisz egy személy egy esemény bekövetkezésében, ha a szóban forgó esemény objektíve nem determinált, mint ha determinált, csak nem tudjuk, hogyan. Másfelõl pedig, a többi interpretáció nem feltétlenül objektív valószínûséget ír le.

Összegezve tehát:

There is a consensus that none of the standard interpretations can provide an ultimate definition in empirical terms of what probability is:

- Probability is **not** the ratio of cases favourable to the event in question over the total number of (equally possible) possibilities.
- Probability is **not** relative frequency on a finite sample.
- Probability is **not** limiting relative frequency.
- Probability is **not** propensity.
- Probability is **not** degree of belief.

teizmus „design-argumentumait” (Craig 1988), illetve annak a bayesiánus konfirmációelméletre épített változatait (Swinburne 1990, 1998). Ezek az argumentumok a kozmológiában is felbukkannak a különbözõ antropikus elvek formájában.

Then what is probability?

And how is it possible then that physics and other empirical sciences apply formal (mathematical) theory of probability, without noticing problem arising from this unanswered fundamental question?

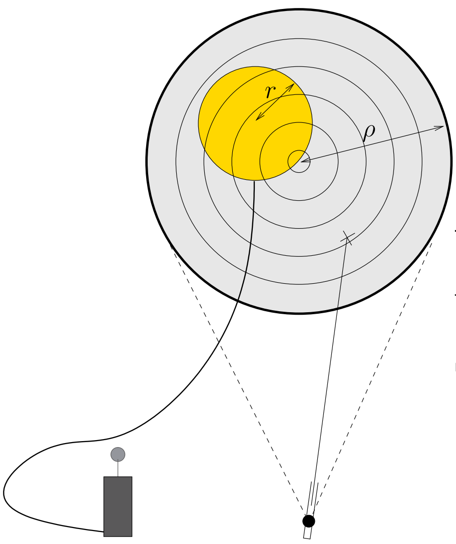
‘No-probability’ Interpretation of Probability

The key idea of my proposal: **there is no such property of an event as its “probability”.** “Probability” is merely a collective term. This is why standard interpretations fail to give a sound definition of probability, and this is why empirical sciences like physics can manage without such a definition.

My argument will be based on the following two general principles:

I shall use **verficationist theory of meaning** in the following very weak sense: In physics, and in other empirical sciences, the meaning of a term standing for a quantity which is supposed to characterise an *objective feature of (physical) reality* is determined by *the empirical operations with which the value of the quantity in question can be ascertained.*

Indispensability argument claims that we ought to have ontological commitment to all and only the entities that are indispensable to our best scientific theories. *Mutatis mutandis*, we ought to have ontological commitment to all and only the features of reality that are indispensable to our best scientific theories.



What is the probability that the balloon will be burst (event A)?
The physicist’s standard answer:

$$p(A) = \frac{\pi r^2}{\pi \rho^2}$$

The right hand side is meaningful. It is an expression consisting of known physical quantities. On the left hand side, however, $p(A)$ is not a known quantity which could be contingently equal to $\frac{\pi r^2}{\pi \rho^2}$.

So the only possible interpretation of this equation is that it is a definition of $p(A)$:

$$p(A) \overset{def}{=} \frac{\pi r^2}{\pi \rho^2}$$

Note that physical quantity $\frac{\text{area of} \ldots}{\pi \rho^2}$ happens to be a **probability like** quantity. It is a dimensionless normalised measure, satisfying the Kolmogorov axioms. Similarly:

$$p(a) \overset{def}{=} \operatorname{tr}\left(\hat{P}_a \hat{W}\right)$$

$$p\left(\{N_i\}_{i=1,2,\ldots}\right) \overset{def}{=} \frac{(\sum_i N_i)!}{\prod_i N_i!}$$

Thesis 1 *There is no such property of an event as its "probability." What we call probability is always a physical quantity characterising the state of affairs corresponding to the event in question. "Probability" can be used only as a collective term: it means different dimensionless $[0,1]$-valued physical quantities, or more precisely, different dimensionless normalised measures composed by different physical quantities in the various specific situations.*

In some sense it is classical

In some sense it reflects the ratio of favourable cases to the number of equally possible cases.

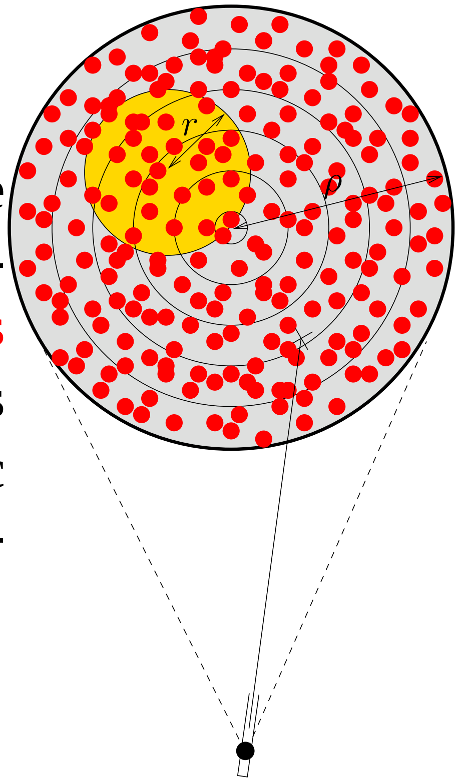
It is like propensity

- The physical quantity $\frac{\pi r^2}{\pi\rho^2}$ is meaningful and has a certain value in every single run of the experiment.

- If we change the size of the balloon during the sequential repetitions of the experiment, such that the sequence of relative frequencies does not converge to a limiting value, then $\frac{\pi r^2}{\pi\rho^2}$ has nothing to do with the relative frequency of event A.

Sometimes like relative frequency

If, however, the value of $\frac{\pi r^2}{\pi\rho^2}$ is constant and the uniform distribution of shots at the target is ensured, then the relative frequency of event A is approximately equal to "probability" $\frac{\pi r^2}{\pi\rho^2}$. This fact has nothing to do with probability-theoretic considerations. It is a simple result of elementary kinematics.



Thesis 2 *The physical quantity identified with "probability" is not the limiting value of relative frequency, and not even necessarily related to the notion of frequency. In some cases, the conditions of the sequential repetitions of a particular situation are such, however, that the physical quantity called "probability" in the given particular context is approximately equal to the relative frequency of the event in question.*

Determinism–indeterminism

- Its existence and value are independent of whether the laws of nature governing the gun firing and the path of the bullets are deterministic or not.

- Moreover, the relationship between $\frac{\pi r^2}{\pi \rho^2}$ and the relative frequency of A (if there is such a relationship at all) is not influenced by the deterministic or indeterministic character of the physical processes in question.

- The relative frequency will be (approximately) equal to $\frac{\pi r^2}{\pi \rho^2}$ independently of whether the uniform distribution of the shots is ensured by means of a deterministic ergodic process, or by means of an objectively random firing.

Lack of knowledge

- Nothing can influence the value of $\frac{\pi r^2}{\pi \rho^2}$, which would be related to our knowledge about the details of the process.

- If the uniform distribution of shots condition is satisfied, $\frac{\pi r^2}{\pi \rho^2} \approx \frac{N(A)}{N}$, independently of whether we know the direction of the subsequent shot, or not.

- The real process actually determines whether the distribution of the shots is uniform or not. *A priori* we must not suppose that it is uniform, merely because we have no information about how the direction of the consecutive shots is determined.

Thesis 3 *The value of the physical quantity identified with "probability" is not influenced by the fact whether the process in question is indeterministic or not. (This value is not necessarily* 0 *or* 1*, merely because the process is deterministic.)*

The value of the physical quantity identified with "probability" is not influenced by the extent of our knowledge about the details of the process.

Neither the value of the physical quantity identified with "probability," nor the existence of the conditions under which this value and the relative frequency of the corresponding event are approximately equal can be knowable a priori.

So far so good. But ...

$$p(a) \stackrel{def}{=} \operatorname{tr}\left(\hat{P}_a \hat{W}\right)$$

$$p\left(\{N_i\}_{i=1,2,\ldots}\right) \stackrel{def}{=} \frac{(\sum_i N_i)!}{\prod_i N_i!}$$

$$p\,(\text{raining}) \stackrel{?}{=} 0.8$$

$$p(\text{H}) \stackrel{?}{=} \frac{1}{2}$$

Consider our example: Assume we know that $r = \frac{1}{2}\rho$, therefore

$$p(A) \stackrel{def}{=} \frac{\pi r^2}{\pi \rho^2} = \frac{1}{4} \tag{18}$$

In brief,

$$p(A) = \frac{1}{4} \tag{19}$$

Statement (19) in itself is completely meaningless. It is just an incomplete formulation of (18).

Similarly, we must assume that there exists a physical quantity X corresponding to "probability" in the given context, such that

$$p(\text{H}) \stackrel{def}{=} X = \frac{1}{2}$$

Although the system in question, the coin together with its environment, is a very complex physical system, we may have a good intuition about what kind of physical quantity X might be. On this intuitive level, we have enough reason to apply some symmetry principles, and conclude that $X = \frac{1}{2}$.

The meaning of $p(\text{raining}) = 0.8$ is much less clear. We have to assume that the meteorologist has a (physical) theory on which this assertion is based. That is,

$$p(\text{raining}) \overset{def}{=} X = 0.8$$

for some physical quantity X.

(In practice, however, the meteorologist does not necessarily have such a theory with such an X, but (s)he simply asserts that the observed relative frequency of raining in similar situations was 0.8. In this case, it is not an assertion about "probability" of raining.

A problem

Let us consider again the quantum mechanical expression

$$p(a) \overset{def}{=} \operatorname{tr}\left(\hat{P}_a \hat{W}\right)$$

First sight it is a correct definition of the "probability" on the left hand side, just like in cases of $p(A) \overset{def}{=} \frac{\pi r^2}{\pi \rho^2}$ and $p\left(\{N_i\}_{i=1,2,\ldots}\right) \overset{def}{=} \frac{(\sum_i N_i)!}{\prod_i N_i!}$. But, $\operatorname{tr}\left(\hat{P}_a \hat{W}\right)$ itself is not a well defined physical quantity! We have no information about $\hat{W}$ independent of "$p(a)$" (more exactly, various "$p(a)$"-s).

Reflections

- We could ascertain the state operator $\hat{W}$ by measuring many different "probabilities" like $p(a)$. But, what is "probability" here?

- We can measure relative frequencies, and in this way determine “probabilities”, even if we do not know what “probability” exactly is, if we assume that relative frequency $\approx$ “probability”.

- But, in QM we have no justification for such an assumption whatsoever!

Perhaps

- Perhaps “probability” is an inappropriate concept for QM, and the statistical rules of QM simply connects relative frequencies (on finite samples) relative to different situations.

- This is, perhaps, the case not only in QM, but “probability” is a concept completely eliminable from scientific discourse.

7. előadás

Okság II.

8. A kauzalitás valószínűségi elmélete

8.0.0.22. Hume úgy gondolta, hogy az ok nem csak szükséges feltétele az okozat bekövetkezésének, hanem egyben elégséges feltétele is: „Az okot úgy definiálhatjuk, mint *egy olyan objektumot, amelyet egy másik követ, és amelyekre fennáll, hogy minden, az elsőhöz hasonló objektumot követ egy, a másodikhoz hasonló objektum.*”[49]

Későbbi filozófusok finomították ezt a képet. A legismertebb Mackie *inus*-elve:[50] az ok *szükséges, de nem elégséges feltétel a feltételek egy olyan rendszerében, amely elégséges, de nem szükséges* az okozat bekövetkezéséhez. A lényeg azonban, hogy az ok szükséges és elégséges volta nem összeegyeztethető az indeterminizmussal: Ha egy B esemény nincs arra determinálva, hogy megtörténjen, akkor nem lehet egy másik A esemény része kondíciók egy olyan rendszerének, amely elégséges ahhoz, hogy B bekövetkezzen. Az indeterminizmus tételezése mellett, a kauzalitásra vonatkozó elgondolásainkat módosítanunk kell.

8.0.0.23. A valószínűségi kauzalitás alapgondolata a következő: Az okozat bekövetkezhet az ok nélkül is, és fordítva, az is megtörténhet, hogy az ok bekövetkezése ellenére az okozat nem történik meg. Az A és B események közötti ok-okozati viszony abban áll, hogy *az ok bekövetkezése megnöveli az okozat bekövetkezésének valószínűségét*. Ezt a következő formulával szokás kifejezni:

$$p\left(B|A\right) > p\left(B|\neg A\right) \tag{20}$$

A tipikus példa, melyet az sztochasztikus kauzalitásról szóló irodalomban olvashatunk a következő: „A dohányzás tüdőrákot okoz.” Nem jelenti ez azt, hogy egy adott személy dohányzása szükségszerűen rákot okoz, mint ahogy nem dohányzónak is lehet tüdőrákja. A

[49] Hume 1748, VII. fejezet.

[50] Insufficient but Non-redundant part of an Unnecessary but Sufficient condition (Mackie 1974).

kauzális kapcsolat abban áll, hogy a dohányzás megnöveli a tüdőrák valószínűségét.

8.0.0.24. A kauzalitás valószínűségi értelmezésével kapcsolatban számos problémát szokás felvetni, melyek közül a három legfontosabbat említjük meg:

1. A (20) egyenlőtlenség megtévesztő abban az értelemben, hogy azt sugallja, hogy A és B között egy aszimmetrikus viszony áll fenn. Ugyanakkor (20) a következőképpen alakítható tovább:

$$\begin{aligned}
\frac{p(A \wedge B)}{p(A)} &> \frac{p(B) - p(A \wedge B)}{1 - p(A)} \\
p(A \wedge B)(1 - p(A)) &> (p(B) - p(A \wedge B))\, p(A) \\
p(A \wedge B) &> p(A)\, p(B)
\end{aligned}$$

   Vagyis (20) egyszerűen azzal ekvivalens, hogy a két esemény között pozitív korreláció van, ami egy szimmetrikus viszony A és B között, és akár úgy is kifejezhető, hogy

$$p(A|B) > p(A|\neg B)$$

2. A korreláció ténye önmagában nem feltétlenül jelenti azt, hogy az egyik esemény a másiknak oka. A korreláció tudniillik származhat közös okból is. Hogy megint egy tipikus példát említsünk a téma irodalmából, korreláció van aközött, hogy valakinek elsárgultak az ujjai és hogy tüdőrákja van, mégsem a sárga ujjak okozzák a tüdőrákot. A korrelációt a dohányzás mint közös ok magyarázza. A korrelációk közös okkal történő magyarázatával részletesen fogunk foglalkozni a 10. fejezetben.

3. Az irodalomban gyakran érvelnek úgy, hogy vannak olyan korrelációk, amelyek mögött sem direkt, sem közös ok típusú kauzális kapcsolat nincs. Tekintsük Eliot Sober ismert példáját:[51]

[51] Sober 1988.

> A kenyér ára Londonban az elmúlt néhány évszázadban folyamatosan emelkedik, és a víz szintje Velencében szintén folyamatosan emelkedik az elmúlt néhány évszázadban. Létezik tehát egy (szimultán) korreláció a velencei vízszint és a londoni kenyérár között – mondja Sober. Ugyanakkor joggal feltételezhetjük, hogy sem közvetlen kauzális kapcsolat, sem közös ok nem létezik a korreláció mögött.

E problémákkal kapcsolatban a következőket érdemes megjegyeznünk. Sem a aszimmetria kérdése, sem a közös ok problémája nem a sztochasztikus kauzalitás sajátossága. Minden eddig említett értelmezés szenved ugyanezektől a nehézségektől. Mint ahogyan a valószínűségi értelmezés is szembe találja magát olyan további kérdésekkel, hogy a kauzális viszony szinguláris események között értelmezett viszony-e, vagy csak eseménytípusokra vonatkozik. (Ha tudniillik szinguláris eseményekre vonatkozik, akkor természetesen felmerül az a kérdés, hogyan van értelmezve a szinguláris események valószínűsége.)

A harmadik problémát illetően meg kell jegyeznünk, hogy Sober példája – amennyiben helytálló – ellentmond a 10. fejezetben tárgyalt Reichenbach-féle közösok-elvnek. Az elv röviden azt állítja, hogy nem létezik korreláció valamilyen kauzális magyarázat nélkül. Tehát, ha két esemény között korreláció van, akkor a két esemény vagy direkt kauzális kapcsolatban áll egymással, vagy létezik a korrelációt magyarázó közös ok. Más szóval, az elv azt állítja, hogy nincsenek a valóságban „véletlen regularitások”.

8.0.0.25. Szokás a Sober-féle ellenpélda és a Reichenbach-féle közösok-elv közötti ellentmondást a következő érveléssel „feloldani”:[52] A kenyér árát Londonban bizonyos dolgok határozzák meg, például a kenyér korábbi ára. Hasonlóan, a víz szintjét Velencében

[52] Arntzenius 1997.

meghatározza a tengervíz szintje egy korábbi időpillanatban, és esetleg más lokális dolgok. Tegyük fel, hogy az időfejlődés mindkét esetben determinisztikus, továbbá, a példa szerint, olyan, hogy mindkét mennyiség értéke folyamatosan növekszik. Tehát a kenyér árának növekedését t időpontban, X_{t+}-t valamilyen korábbi, lokális (londoni) esemény határozza meg, X_{t-}. Hasonlóan, a víz szintjének emelkedését, Y_{t+}-t egy korábbi lokális Y_{t-} esemény determinálja. Determinisztikus esetben $p(X_{t+}|X_{t-}) = p(Y_{t+}|Y_{t-}) = 1$. Nyilvánvaló, hogy, ha X_{t+} és Y_{t+} között maximális korreláció van, akkor maximális korreláció van X_{t-} és Y_{t+} között is. Ha azonban ez igaz – hangzik az érv –, akkor ez egyben megmagyarázza az X_{t+} és Y_{t+} közötti korrelációt.

Ezzel azonban nem tettünk mást, mint a két távoli esemény közötti, eddig magyarázatra szoruló korrelációt visszavezettük korábbi két, szintén szeparált esemény közötti korrelációra, amely azonban ugyanúgy magyarázatra szorul, hacsak nem akarjuk ezt a korrelációt egy korábbi események közötti korrelációval magyarázni, és így tovább, az Ősrobbanásig. Ez tehát nem a megfelelő út a Sober által felvetett ellenpélda kezelésére.

8.0.0.26. A Sober-féle példa ugyanis – mint minden más hasonló példa a filozófiai irodalomban, amelyik „véletlen együttjárásról” szól – egyszerűen nem helytálló. Ugyanis nincs korreláció! Ha elfogadjuk a példa feltételezését, hogy a kenyér árának növekedése is és a vízszint emelkedése is évről évre biztosan bekövetkezik, akkor tehát két egyvalószínűségű eseményről van szó, melyek között a korreláció zérus, azaz a két esemény statisztikailag független:

$$\Delta(X,Y) = p(X \wedge Y) - p(X)\,p(Y) = 1 - 1 \cdot 1 = 0$$

8.0.0.27. Persze megpróbálhat valaki azzal érvelni, hogy talán mégsem egy valószínűségű eseményekről van szó, hanem csak, mondjuk, 0.99 mindkét esemény valószínűsége, azaz mindegyikre igaz, hogy egy évszázadban várhatóan egyszer nem következik be. Akkor már lehetséges közöttük korreláció, és ha mégoly gyenge is ez a korreláció, nincs rá kauzális magyarázat.[53] Az érv azért nem elfogadható, mert a „lehetséges, hogy van” és a „van” között igen nagy a különbség! Mert ha $p(X) = p(Y) = 1$, akkor valóban tudjuk $p(X \wedge Y)$ értékét ($= 1$). De ha $0 < p(X), p(Y) < 1$, akkor a **??.** pont **(??)** alapján a konjunkció valószínűsége a

$$\min\{p(X), p(Y)\} \geq p(X \wedge Y) \geq p(X) + p(Y) - 1$$

határok között bármi lehet, beleértve a függetlenséget jelentő $p(X \wedge Y) = p(X)\, p(Y)$ értéket is. És hogy ezen az intervallumon belül mekkora a konjunkció valószínűsége, az kizárólag empirikusan eldönthető kérdés. Ha például $p(X) = p(Y) = 0.99$, akkor a koincidencia (tehát a konjunkció) valószínűsége a $0.98 \leq p(X \wedge Y) \leq 0.99$ intervallumba esik. Ezen a szűk intervallumon belül, a $p(X)p(Y) = 0.9801$ érték felel meg a függetlenség esetének. Ahhoz tehát, hogy azt mondhassuk, hogy valóban van korreláció a londoni kenyérár növekedése és a velencei tengerszint növekedése között, a szóban forgó valószínűségeket 10^{-4} pontossággal ismernünk kellene, amihez soksok ezer év megfigyelésére lenne szükség. Ilyen empirikus adatokra való referencia nélkül nem lehet korrelációról beszélni.

8.0.0.28. A kauzalitás valószínűségi elméletének kiterjedt és fontos irodalma van.[54] Nem célunk most e téma további részleteit kifejteni. A Reichenbach-féle közösok-elv tárgyalásához visszatérünk a 10. fejezetben, és további részleteket vizsgálunk meg.

Összefoglalóan azt kell hangsúlyoznunk, hogy a valószínűségi értelmezés alapfogalma, a statisztikus korreláció, tökéletes indikátora a mögöttes kauzális mechanizmusoknak. Egész pontosan azt állítjuk, hogy ha korrelációt látunk, amögött mindig – direkt vagy közösoktípusú – kauzális kapcsolat van. Még talán arra is van mód – a való-

[53] Vö. Arntzenius 1997. Hasonló érveket fogalmazott meg J. Berkovitz „On the Relation Between Correlation and Causation in Deterministic Models” c. előadásában, *International Interdisciplinary Workshop on Determinism*, Ringberg, (Németország), 2001 június 4-8.

[54] Suppes 1970; Salmon 1980, 1984; Menzies 1987; Mellor 1995.

színűség megfelelő értelmezésével –, hogy a korrelációkat és a mögöttes kauzális kapcsolatokat szinguláris eseményekre vonatkozóan is értelmezzük. *A kauzális viszony azonban nem merül ki, nem azonos a statisztikus korrelációval. A korreláció a kauzális kapcsolat következménye, de nem azonos a két fogalom*. Mindenekelőtt azért nem, mert a valószínűségi leírás nem tükrözi a téridőbeli viszonyokat.

9. A kauzalitás ontológiai elmélete

9.0.0.29. A kauzalitás most kifejteni kívánt felfogását elsősorban Russell, Reichenbach és Salmon kauzális folyamatokra vonatkozó nézetei, valamint Salmon *„at-at” theory of causal propagation* néven ismert felfogása motiválta, és bizonyos vonásaiban megegyezik ezekkel az elképzelésekkel. A kauzalitás ontológiai elméletének alapvető nézőpontja ugyanaz a fizikalista szemléletmód, amely a valószínűség **??.** pontban kifejtett fizikalista interpretációját jellemzi.

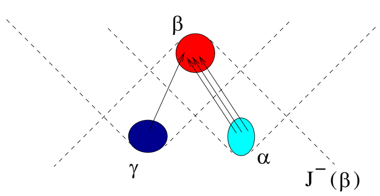
Fizikalista szemmel áttekintve a kauzalitásról eddig elmondottakat, a következő megjegyzéseket kell tennünk:

- Feltűnő, hogy a téma irodalmában használt példákban olyan események közötti kauzális kapcsolatokat analizálnak, mint <a virág locsolása>, <a virág elhervadása>, <a repülőgép lezuhanása>, <Brunó leesése a lépcsőn>, <a fegyver helyszínen felejtése>, <a rabló lebukása>, <árvíz>, <a ház összeomlása>, <dohányzás>, <a tüdőrák kialakulása>, stb. Ezek olyan „események”, amelyek elemi(bb) fizikai események milliárdjainak összessége, komplex eseménysorozatok, kiterjedt folyamatok. Amikor a kauzalitás végső mibenlétét kutatjuk, akkor feltételezhetően az elemi történések közötti kauzális kapcsolatot kell analizálnunk. Az említett komplex folyamatok „kauzális magyarázatának” azt kellene tekintenünk, ha értelmezni tudjuk azokat a megfelelő kauzális rendbe illesztett elemi történések összességeként.

- Mármost, hogy milyen beszámolót adunk ezekről az elemi fizikai eseményekről és azok kauzális viszonyairól, erősen függ attól az ontológiai képtől, amelyet a világról előzetesen kialakítottunk. Tény, hogy ennek az ontológiai képnek a megalkotása során a fent említett komplex jelenségek közötti korrelációk megfigyeléséből indulunk ki. Ez nem jelent azonban semmiféle cirkularitást.

- Az elemi fizikai történések mindig partikulárisak. Kauzális viszonyt csak partikuláris események között tételezhetünk fel, vagyis az univerzum életének olyan elemi történései között, amelyeknek definit téridőbeli locusa van. Amikor eseménytípusok között regularitást tapasztalunk, akkor mindig arról van szó, hogy az egyik típusba tartozó partikuláris esemény áll kauzális kapcsolatban egy másik típusba tartozó partikuláris eseménnyel. A partikuláris események közötti kauzális kapcsolat szüli a típusok között észlelt regularitást, nem fordítva.

- A kauzalitás ontológiai szempontból tartalmas fogalom. A kauzális kapcsolatot mindig a négy ismert fizikai kölcsönhatás valósítja meg. Minden további állításunk a kauzális kapcsolatokról úgy értendő, hogy azzal e négy fizikai kölcsönhatás valamilyen tulajdonságát jellemezzük.



8. ábra. *A téridő két tartományában történő esemény közötti kauzális kapcsolat*

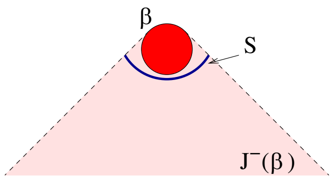
9.0.0.30. A kauzális kapcsolatot tehát a téridő adott tartományához tartozó, partikuláris események között értelmezzük (8. ábra). Mint tudjuk, egyetlen kölcsönhatást közvetítő mező terjedési sebessége sem nagyobb a fényterjedés sebességénél, tehát ahhoz, hogy az α esemény hatással lehessen a β eseményre, α-nak benne kell lennie a β múlt-fénykúpjában, $\alpha \subseteq J^{-}(\beta)$. Természetesen, nem csak az α esemény, vagyis nem csak a téridő α tartományában történtek lehetnek hatással a téridő β tartományában történtekre, hanem minden olyan γ tartományban végbement történés is, amelyre $\gamma \subseteq J^{-}(\beta)$.

9.0.0.31. Mivel nem csupán egy kauzális faktort kívánunk értelmezni, hanem a partikuláris β esemény okát, abban a hume-i értelemben, hogy az ok legyen szükséges és elégséges feltétele a β esemény bekövetkezésének, nem mondhatjuk azt, hogy β oka az α esemény. Hogy miért nem, ahhoz a következőket kell előzetesen belátnunk:

- A β eseményt a maga totális partikularitásában fogjuk fel, tehát az univerzum történetének azt a darabját értjük alatta, amelyik a β téridőtartományban megy végbe. Lehet, hogy *számunkra* ennek a tartománynak a tartalma lényeges és lényegtelen elemekre bontható, most azonban nincsenek lényegtelen elemek. Ha a β történés beleesik a <lámpa felgyullad> eseménytípusba, az lehet számunkra praktikus megfontolásból lényeges, de a β eseménynek, mint partikuláris eseménynek a kauzális viszonyait illetően semmi jelentősége.

- A téridő minden tartományában történik valami.

- Bármi is történik a téridő egy X tartományában, annak valamilyen hatása van a tartomány $J^{+}(X)$-szel jelölt jövő-fénykúpjának *egészére*. Ha más nem, a legkisebb átrendeződése a tömegenergia-eloszlásnak megváltoztatja a gravitációs tér tulajdonságait $J^{+}(X)$-ben. Hasonlóképpen, tetszőleges X téridőtartományra igaz, hogy bármi is történt $J^{-}(X)$-ben, annak valamilyen hatása van az X tartományban történtekre.



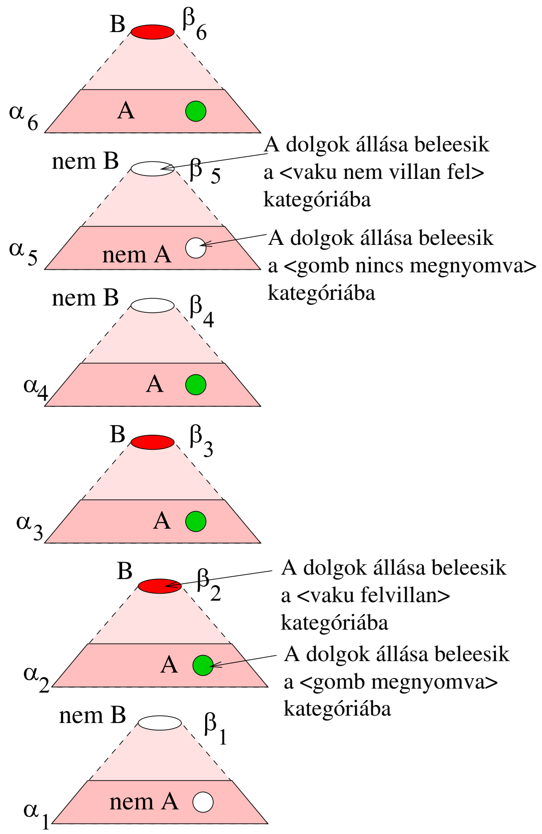
9. ábra. *A* β *esemény oka az egész* $J^{-}(\beta)$ *téridőtartomány. A markovitást is figyelembe véve,* β *oka az S hiperfelülethez tartozó események összessége*

Ha tehát azt kérdezzük, mi az oka a β eseménynek, és β alatt a β téridőtartományban történt partikuláris eseményt értjük, akkor azt kell válaszolnunk, hogy β oka minden, ami a $J^{-}(\beta)$ téridőtartományban történik (9. ábra).

Feltételezve a fizikai folyamatok markovitását – nem látunk példát ugyanis az ellenkezőjére –, $J^{-}(\beta)$-nak a β tartományban történtekre való hatása ugyanaz, mint az *S* Cauchy-felület mentén történtek hatása a β tartományban történtekre. Azt is mondhatjuk tehát, hogy a β partikuláris esemény oka az *S* hiperfelület mentén történt események összessége.

9.0.0.32. Érdemes még egyszer hangsúlyosan átgondolnunk a partikuláris események közötti kauzális kapcsolatok viszonyát az eseménytípusok közötti regularitáshoz, illetve a regularitáselmélet szerinti értelemben vett „kauzális” kapcsolatokhoz.

Vizsgáljuk meg egy vaku „flash” gombjának megnyomása és a vaku felvillanása közötti kapcsolatot. A regularitáselmélet, valamint a kauzalitás valószínűségi elmélete abból a megfigyelésből indul ki, hogy valahányszor megnyomom a gombot, a vaku felvillan, vagy legalábbis $p(B|A) > p(B|\neg A)$, ahol *A* a gomb megnyomását, *B* a vaku felvillanását jelöli. *A* és *B* itt azonban eseménytípusokat jelöl, eseménytípusok közötti reguláris kapcsolatról van szó. Ez a reguláris kapcsolat azonban nem kauzális kapcsolat ontológiai értelemben. Ontológiai értelemben kauzális kapcsolat csak partikuláris események között van. Viszont a megfigyelt eseménytípusok, és a közöttük megfigyelt reguláris kapcsolat jól értelmezhető, és beilleszthető a vi-



10. ábra. *A „flash" gomb megnyomása és a vaku felvillanása közötti reguláris kapcsolat a megfelelő téridőtartományok közötti ontológiai kauzális kapcsolaton nyugszik. Az A eseménytípus akkor történik meg, ha az adott téridőtartományban a dolgok állása beleesik <a gomb megnyomva> kategóriába. Hasonlóan, a B esemény annak felel meg, hogy a dolgok állása a későbbi téridőtartományban beleesik <a vaku felvillan> kategóriába. Az eseménytípusok közötti regularitás az ontológiai kauzális kapcsolatok következménye:* $p(B|A) > p(B|\neg A)$

lág – ontológiai értelemben vett – kauzális szerkezetébe. A 10. ábrán a vaku gombjának megnyomása és a vaku felvillanása közötti kapcsolatra vonatkozó kísérletsorozat téridő diagramját látjuk. A gomb megnyomása egy eseménytípust jelent, vagyis azt, hogy a dolgok állása az adott téridőtartományban (például az α_2 tartományban) beleesik <a gomb megnyomva> kategóriába. Hasonlóan, a *B* esemény annak felel meg, hogy a dolgok állása a későbbi (például a β_2) téridőtartományban beleesik <a vaku felvillan> kategóriába. Kauzális kapcsolat a partikuláris események, a téridő α_i és β_i tartományaiban történtek között van, függetlenül attól, hogy az univerzum történetének milyen epizódjai történnek ezekben a tartományokban. Ezeknek a történéseknek és a köztük lévő kauzális kapcsolatoknak a *tulajdonsága*, hogy éppen olyanok, hogy az $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ és α_7 tartományban történtek beleesnek <a gomb megnyomva> kategóriába, valamint, hogy a $\beta_1, \beta_2, \beta_4$ és β_7 tartományban történtek beleesnek <a vaku felvillan> osztályba, míg történetesen β_3 az univerzum történetének egy olyan darabkája, amelyik <a vaku nem villan fel> kategóriába tartozik, stb., s hogy mindezekből következően az eseményosztályokra történetesen fennáll, hogy $p(B|A) > p(B|\neg A)$. *A korreláció tehát következménye a partikuláris események közötti kauzális kapcsolatoknak, pontosabban a partikuláris események és azok kauzális kapcsolatainak egy tulajdonsága.* Lehetnének ezek a kauzális kapcsolatok olyanok is, hogy a szóban forgó eseményosztályok között nincs korreláció. Ez nem jelentené azt, hogy nincs kauzális kapcsolat a megfelelő (α_i, β_i) partikuláris eseménypárok között. Más szóval, *a partikuláris események szintjén létezik egy mélyebb kauzális ontológia, és ehhez képest esetleges, hogy ez milyen regularitásokat produkál a különböző eseménykategóriák között.* Ezt illusztrálandó, gondoljuk el a következő példát.

9.0.0.33. Legyen a vaku, amellyel kísérletezünk, olyan, hogy változtatja a színét, minden felvillanáskor mondjuk pirosan vagy zölden villan fel. (Könnyű lenne ilyet készíteni.) A vaku nyomógombja legyen hibás, és hol valóban kapcsol, hol nem. A kapcsolásnak legyen valamilyen jellemzője, mondjuk, néha erősen kattan, néha halkan. Induljunk ki abból, hogy a gombot mindig megnyomjuk. E kondíció utáni valószínűségek legyenek a következők:

$$\begin{aligned} p(<\text{erős}>) &= 0.5 \\ p(<\text{halk}>) &= 0.2 \\ p(<\text{zöld}> \mid <\text{erős}>) &= 0.2 \\ p(<\text{zöld}> \mid <\text{halk}>) &= 0.5 \\ p(<\text{piros}> \mid <\text{erős}>) &= 0.8 \\ p(<\text{piros}> \mid <\text{halk}>) &= 0.5 \end{aligned}$$

Minden további valószínűség ezekből könnyen kiszámítható. Ugyanazok a partikuláris események és ugyanazok a kauzális kapcsolatok <a kapcsoló kattan> és <a vaku felvillan> eseményosztályok között szoros korrelációt jelent:

$$\begin{aligned} & p(<\text{villan}> \mid <\text{kattan}> \\ = p(<\text{zöld}> \vee <\text{piros}> \mid <\text{erős}> \vee <\text{halk}>) &= 1 \\ p(<\text{villan}> \mid <\text{nem kattan}> &= 0 \end{aligned}$$

Ezzel szemben <a gomb erősen kattan> és <a vaku zölden felvillan> eseményosztályok függetlenek:

$$p(<\text{zöld}> \wedge <\text{erős}>) = 0.1 = p(<\text{zöld}>)p(<\text{erős}>) = 0.2 \times 0.5$$

Ugyanakkor, tehát még mindig ugyanazon kauzális ontológia jegyében, a <zölden villan> és a <halkan kattan> eseménytípusok között van korreláció.

E példa is azt a konklúziónkat erősíti meg, hogy a világ kauzális struktúrája nem keverendő össze az eseménytípusok közötti regularitásokkal.

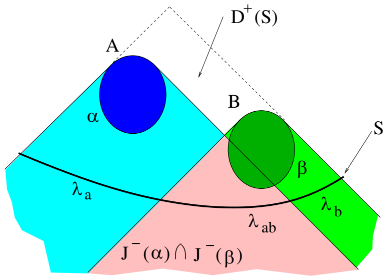
10. Nincs korreláció kauzalitás nélkül

10.0.0.34. Az előzőekben azt hangsúlyoztuk, hogy a világ kauzális struktúrája nem keverendő össze az eseménytípusok közötti regularitásokkal. Nem állítjuk azonban, hogy a két dolog között nincs semmilyen összefüggés. A kauzalitás, mint alapvetőbb ontológiai struktúra *produkálja* a regularitásokat – ha vannak. Nem létezik regularitás, eseménytípusok közötti korreláció kauzalitás nélkül! Pontosabban, ha két eseményosztály között korrelációt látunk, akkor van a két eseményosztály összetartozó partikuláris eseményeinek kauzális múltjában valami közös (közös ok), amely megmagyarázza a korreláció tényét.

Egy irodaházban a főnöki telefon fejhallgatójában folyamatosan ugyanaz hallatszik, mint a másik szobában hallgatózó titkár telefonkagylójában (11. ábra). Ha ezeket a hangjeleket összehoznánk egy oszcilloszkóp ernyőjére, nagyon szép koincidenciát látnánk, azaz a két jel között erős korreláció van: Valahányszor a főnöknő férjének hangján elhangzó „drágám”-nak megfelelő mintázat jelenik meg a főnöki membrán rezgéseiben (*A* esemény), a „drágám”-nak megfelelő mintázat észlelhető a titkár fülére tapasztott membrán rezgéseiben is (*B* esemény).

10.0.0.35. Szándékosan olyan példát választunk, amelyben a két esemény nem kauzálisan szeparált, de nem is olyan, hogy az egyik esemény a másiknak kauzális faktorát képezné. A telefondrót először a titkár íróasztalához megy, onnan tovább a főnökhöz. Tehát a bejövő jel, a drótban terjedő elektromágneses hullám először a titkár telefonjára van hatással, de a titkár telefonjának rezgésbe jövő membránja, a telefon áramkörein keresztül nyilván vissza is hat a drótban terjedő elektromágneses hullámra, és ennek a visszahatásnak még van módja módosítani a főnöki telefon membránjának rezgéseit.

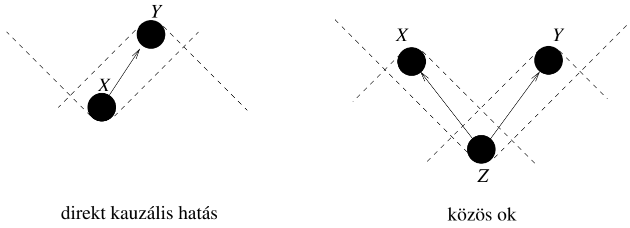
Szokás az irodalomban a közös ok fogalmát arra az esetre szűkíteni, amikor a két esemény „nincs direkt kauzális kapcsolatban”. E



11. ábra. *A főnöki telefon fejhallgatójában folyamatosan ugyanaz hallatszik, mint a másik szobában hallgatózó titkár telefonkagylójában: Valahányszor a főnöknő férjének hangján elhangzó „drágám”-nak megfelelő mintázat jelenik meg a főnöki membrán rezgéseiben (A esemény), a „drágám”-nak megfelelő mintázat észlelhető a titkár fülére tapasztott membrán rezgéseiben is (B esemény). Ha a világ LDM tulajdonságú, akkor a $D^+(S)$ tartományban történteket egyértelműen meghatározza a dolgok állása az S Cauchy-felületen*

megfogalmazás mögött az a dichotómia áll, mely szerint két esemény közötti korrelációt látva két lehetőséget tudunk elképzelni (12. ábra). Vagy az egyik esemény direkt kauzális hatással van a másik eseményre, vagy létezik egy közös ok, amelyik mindkettőre hatással van, és ez eredményezi a korrelációt.[55] Mint példánk mutatja, ez nem mindig van így. A direkt és közösok-típusú kauzális séma esetei nem mindig válnak szét ilyen tiszta formában. Mondandónk lényegét tekintve azonban az a fontos, hogy ennek nincs is jelentősége. A lényeg ugyanis az, hogy mindkét esetben van a két esemény kauzális múltjában valami közös, amely megmagyarázza a korrelációt.

[55] Reichenbach (1956) és Salmon (1984) a direkt és közösok-típusú kauzális séma, más szóval a valóságos és pszeudo-folyamatok megkülönböztetésére az ún. „jel átvivési kritériumot” (mark criterion) alkalmazza. Egy elforduló reflektor által a falon létrehozott mozgó fényfolt egy pszeudo-folyamat, mert ha egy ponton megváltoztatjuk, például megszínezzük egy színes üveggel, akkor ez az elszíneződés nem halad tovább, a fényfolt a fal másik végénél ugyanolyan lesz, mintha semmit sem csináltunk volna. A fény terjedése a reflektortól a falig viszont valóságos folyamat, mert ha a reflektor előtt a fényt „megfestjük”, akkor a falon keletkező fényfolt is „megfestődik”. A jel átvivés kritériuma nem minden esetben alkalmazható. Nem alkalmazható például az EPR-kísérlet random eseményeivel kapcsolatban. Mint a **??.** pontban bemutatott távíró példáján láthatjuk, lehetséges olyan kauzális folyamat, amelyik nem alkalmas jelek továbbítására. Fontos azt is tudatosítanunk, hogy a kritérium a szabad akaratra vonatkozó előzetes metafizikai feltevésekre épül.



12. ábra. *Ha két esemény között korrelációt látunk, két lehetőséget tudunk elképzelni. Vagy az egyik esemény direkt kauzális hatással van a másik eseményre, vagy létezik egy közös ok, amelyik mindkettőre hatással van, és ez eredményezi a korrelációt. Mint a* 11. *ábrán bemutatott példán láthatjuk, a direkt és közösok-típusú kauzális séma esetei nem mindig válnak szét ilyen tiszta formában*

10.0.0.36. De mit jelent pontosan az, hogy valami „megmagyarázza" a korrelációt? Vegyük észre, hogy ennek a kifejezésnek nincs eleve adott jelentése. Ellenkezőleg, most adunk ennek értelmet. Induljunk ki abból, hogy ha feltesszük, hogy a világ LDM tulajdonságú, akkor a $D^{+}(S)$ tartományban történteket egyértelműen meghatározza a dolgok állása az S Cauchy-felületen. A statisztikai sokaság 11. ábrán látható mintázat egymást követő ismétlődéseiből áll, a Cauchy-adatok valamilyen statisztika szerint szóródó értékével. Az A és B eseménytípusok közötti korrelációt mindenképpen megértettük az adott statisztikai sokaságban, ha 1) értjük, hogy az egymást követő esetekben a $(\lambda_a, \lambda_{ab}, \lambda_b), (\lambda'_a, \lambda'_{ab}, \lambda'_b), (\lambda''_a, \lambda''_{ab}, \lambda''_b), \ldots$ Cauchy-adatok értéke miért pont annyi, amennyi, továbbá 2) hogy hogyan determinálják ezek az értékek a megfelelő eseménytípusok bekövetkezését. Ez azonban a korreláció létrejöttének túl ambiciózus magyarázata. Annak megértéséhez, hogy a példánkban miért hallatszik ugyanaz a szöveg a két telefonkagylóban, elégséges azt látnunk, hogy a közös bejövő telefonkábelre van csatlakoztatva mindkét telefon, és értenünk a telefonok működését. Nem kell azonban tudnunk/értenünk azt is, hogy a betelefonáló férj miért mondja pont azt, amit mond. Vagyis a korreláció megértéséből elhagyhatjuk az 1) pontot.

10.0.0.37. Foglalkozzunk tehát a 2) ponttal. A klasszikus fizikai képnek megfelelően tehát az S hiperfelület mentén a Cauchy-adatok értéke egyértelműen meghatározza, hogy mi történik $D^{+}(S)$-ben, így az, hogy az A és B események bekövetkeznek-e vagy sem, a $(\lambda_a, \lambda_{ab}, \lambda_b)$ adatoktól függ. λ_a a Cauchy-adatoknak az a része, melyek az S hiperfelület azon részére esnek, amelyik beleesik az α partikuláris esemény hátrafénykúpjába, de kívül van β hátrafénykúpján, λ_b jelentése hasonló a fordított esetre vonatkozóan, míg λ_{ab} az adatoknak az a része, amelyik a Cauchy-felületnek a két fénykúp metszetébe eső részére esik.

Egy X eseménytípus bekövetkezése azt jelenti, hogy a megfelelő $D^{+}(S)$ tartományban a dolgok állása beleesik az X kategóriába. Hogy mely eseménytípusok következnek be és melyek nem, elvben kifejezhető a következő függvényekkel:

$$u^{X}(\lambda_a, \lambda_{ab}, \lambda_b) = \begin{cases} 1 & \text{ha a } D^{+}(S) \text{ beleesik az } X\text{-típusba} \\ 0 & \text{ha nem} \end{cases} \tag{21}$$

Figyelembe véve, hogy egy eseményre nem lehet hatással olyan esemény, amely a fénykúpján kívül esik,

$$\begin{aligned} u^{A}(\lambda_a, \lambda_{ab}, \lambda_b) &= u^{A}(\lambda_a, \lambda_{ab}) \\ u^{B}(\lambda_a, \lambda_{ab}, \lambda_b) &= u^{B}(\lambda_{ab}, \lambda_b) \end{aligned} \tag{22}$$

A statisztikus sokaságot alkotó ismétlődő szituációkban a $\lambda_a, \lambda_{ab}, \lambda_b$ paraméterek mindenkori értékei – a (21) függvényeknek megfelelően – egyértelműen determinálják, hogy mi történik az adott esetben. Az egymást követő téridő mintázatokon – képzeletben – leszámolhatjuk a különböző $(\lambda_a, \lambda_{ab}, \lambda_b)$-kombinációk relatív gyakoriságát. Így adottnak vesszük a $p(\lambda_a), p(\lambda_{ab}), p(\lambda_b), p(\lambda_a \wedge \lambda_{ab}), \ldots p(\lambda_a \wedge \lambda_{ab} \wedge \lambda_b)$ valószínűségeket. Ezek segítségével, felhasználva a (22) függvényeket, az események valószínűsége a következőképpen reprodukálható:

$$p(A) = \sum_{\lambda_a, \lambda_{ab}} u^{A}(\lambda_a, \lambda_{ab})\, p(\lambda_a \wedge \lambda_{ab}) \tag{23}$$

$$p(B) = \sum_{\lambda_{ab},\lambda_b} u^B(\lambda_{ab},\lambda_b)\, p(\lambda_{ab}\wedge\lambda_b) \tag{24}$$

$$p(A\wedge B) = \sum_{\lambda_a,\lambda_{ab},\lambda_b} u^A(\lambda_a,\lambda_{ab})\, u^B(\lambda_{ab},\lambda_b)\, p(\lambda_a\wedge\lambda_{ab}\wedge\lambda_b) \tag{25}$$

És természetesen a valószínűségek reprodukálásával együtt reprodukáltuk az A és B eseménytípus közötti korrelációt is. Ha tehát elfogadjuk Wesley Salmon meghatározását, hogy egy jelenség megértése annyit tesz, hogy képesek vagyunk azt elhelyezni a világ jelenségeinek kauzális rendjében, akkor azt mondhatjuk, hogy megértettük, megmagyaráztuk az A és B közötti korrelációt.

10.0.0.38. λ_a, λ_{ab} és λ_b általában nem függetlenek statisztikailag. Tehát az A és B eseménytípusok közötti korreláció (23)–(25) egyenletekben megnyilvánuló „magyarázata” annyit jelent, hogy alkalmas u^A és u^B függvények mellett, ha van korreláció a térszerűen szeparált λ_a, λ_{ab} és λ_b értékek között, akkor van korreláció A és B között is. Más szóval, az A és B közötti korrelációt más, korábbi korrelációkra vezettük vissza.

Fontosabb azonban, hogy a (23)–(25) összefüggések akkor is eredményezhetnek korrelációt, amikor λ_a, λ_{ab} és λ_b értékek között nincs statisztikus korreláció, vagyis az A és B közötti korreláció mintegy a semmiből keletkezik. Tulajdonképpen ezt az esetet kell a korreláció igazi magyarázatának tekintenünk! Most megmutatjuk, hogy ennek szükséges feltétele, hogy a két fénykúp metszetébe eső λ_{ab} paraméter minden lehetséges értéke kielégítse az ún. *árnyékolási (screening off) feltételt*:

$$p(A\wedge B|\lambda_{ab}) = p(A|\lambda_{ab})\, p(B|\lambda_{ab}) \tag{26}$$

vagyis a λ_{ab} rögzített értéke mellett az A és B közötti korrelációnak el kell tűnnie. Ha ugyanis feltételezésünk szerint

$$p(\lambda_a\wedge\lambda_{ab}\wedge\lambda_b) = p(\lambda_a)\, p(\lambda_{ab})\, p(\lambda_b) \tag{27}$$

akkor ezt behelyettesítve a (23)–(25) egyenletekbe, a következőt kapjuk:

$$\begin{aligned} p(A\wedge B|\lambda_{ab}) &= \sum_{\lambda_a,\lambda_b} u^A(\lambda_a,\lambda_{ab})\, u^B(\lambda_{ab},\lambda_b)\, p(\lambda_a)\, p(\lambda_b) \\ &= \sum_{\lambda_a} u^A(\lambda_a,\lambda_{ab})\, p(\lambda_a) \sum_{\lambda_b} u^B(\lambda_{ab},\lambda_b)\, p(\lambda_b) \\ &= p(A|\lambda_{ab})\, p(B|\lambda_{ab}) \end{aligned}$$

Ez tehát azt jelenti, hogy – eltekintve attól a triviális esettől, amikor más, korábbi események közötti korrelációra vezetjük vissza – *az A és B eseménytípusok közötti korrelációt akkor lehet a kauzális ontológia szintjén maradéktalanul megértenünk, ha a két eseménytípus eseteit megvalósító partikuláris események hátrafénykúpjainak metszetében, pontosabban annak egy Cauchy-felülettel való szelésén, a dolgok állása klasszifikálható úgy, hogy minden egyes osztályra teljesüljön az árnyékolási feltétel*, vagyis ha az osztályokat egy $\lambda_{ab},\lambda'_{ab},\ldots$ paraméterezéssel adjuk meg, akkor minden λ_{ab} paraméterérték teljesítse a (26) feltételt.

10.0.0.39. Érdemes megvizsgálni, hogy mi annak a feltétele, hogy több esemény között fellépő korreláció-rendszert elhelyezzünk egy LDM világ kauzális rendjében? Tekintsük azt az egyszerű esetet, amikor három különböző eseménytípus közötti három különböző korrelációról van szó. A fentiekhez hasonlóan,

$$\begin{aligned} p(A) &= \sum_{\substack{\lambda_a,\lambda_{ab}\\ \lambda_{ac},\lambda_{abc}}} u^A(\lambda_a,\lambda_{ab},\lambda_{ac},\lambda_{abc})\, p(\lambda_a)\, p(\lambda_{ab})\, p(\lambda_{ac})\, p(\lambda_{abc}) \\ p(B) &= \sum_{\substack{\lambda_b,\lambda_{ab}\\ \lambda_{bc},\lambda_{abc}}} u^B(\lambda_b,\lambda_{ab},\lambda_{bc},\lambda_{abc})\, p(\lambda_b)\, p(\lambda_{ab})\, p(\lambda_{bc})\, p(\lambda_{abc}) \end{aligned}$$

$$p(C) = \sum_{\substack{\lambda_c,\lambda_{ac}\\ \lambda_{bc},\lambda_{abc}}} u^C(\lambda_c,\lambda_{ac},\lambda_{bc},\lambda_{abc})\, p(\lambda_c)\, p(\lambda_{ac})\, p(\lambda_{bc})\, p(\lambda_{abc})$$

$$\begin{aligned}
p(A\wedge B) &= \sum_{\substack{\lambda_a,\lambda_b,\lambda_{ab}\\ \lambda_{ac},\lambda_{bc},\lambda_{abc}}} u^A(\lambda_a,\lambda_{ab},\lambda_{ac},\lambda_{abc})\, u^B(\lambda_b,\lambda_{ab},\lambda_{bc},\lambda_{abc}) \\
&\times\ p(\lambda_a)\, p(\lambda_b)\, p(\lambda_{ab})\, p(\lambda_{ac})\, p(\lambda_{bc})\, p(\lambda_{abc}) \\
p(A\wedge C) &= \sum_{\substack{\lambda_a,\lambda_b,\lambda_{ab}\\ \lambda_{ac},\lambda_{bc},\lambda_{abc}}} u^A(\lambda_a,\lambda_{ab},\lambda_{ac},\lambda_{abc})\, u^C(\lambda_c,\lambda_{ac},\lambda_{bc},\lambda_{abc}) \\
&\times\ p(\lambda_a)\, p(\lambda_c)\, p(\lambda_{ab})\, p(\lambda_{ac})\, p(\lambda_{bc})\, p(\lambda_{abc}) \\
p(B\wedge C) &= \sum_{\substack{\lambda_a,\lambda_b,\lambda_{ab}\\ \lambda_{ac},\lambda_{bc},\lambda_{abc}}} u^B(\lambda_b,\lambda_{ab},\lambda_{bc},\lambda_{abc})\, u^C(\lambda_c,\lambda_{ac},\lambda_{bc},\lambda_{abc}) \\
&\times\ p(\lambda_b)\, p(\lambda_c)\, p(\lambda_{ab})\, p(\lambda_{ac})\, p(\lambda_{bc})\, p(\lambda_{abc})
\end{aligned}$$

Vegyük észre, hogy most a három fénykúp metszetére vonatkozó λ_{abc} paraméter általában nem elégíti ki a három korrelációra vonatkozó árnyékolási feltétel mindegyikét (sőt, általában egyiket sem). Ezzel szemben a

$$\lambda_{ab}\wedge\lambda_{abc}$$
$$\lambda_{ac}\wedge\lambda_{abc}$$
$$\lambda_{bc}\wedge\lambda_{abc}$$

paraméterek külön-külön igen, például

$$\begin{aligned}
p(A\wedge B|\lambda_{ab}\wedge\lambda_{abc}) &= \sum_{\lambda_a,\lambda_b,\lambda_{ac},\lambda_{bc}} u^A(\lambda_a,\lambda_{ab},\lambda_{ac},\lambda_{abc}) \\
&\times\ u^B(\lambda_b,\lambda_{ab},\lambda_{bc},\lambda_{abc})\, p(\lambda_a)\, p(\lambda_b)\, p(\lambda_{ac})\, p(\lambda_{bc}) \\
&= \sum_{\lambda_a,\lambda_{ac}} u^A(\lambda_a,\lambda_{ab},\lambda_{ac},\lambda_{abc})\, p(\lambda_a)\, p(\lambda_{ac})
\end{aligned}$$

$$
\begin{aligned}
&\times \sum_{\lambda_b,\lambda_{bc}} u^B(\lambda_b,\lambda_{ab},\lambda_{bc},\lambda_{abc})\, p(\lambda_b)\, p(\lambda_{bc}) \\
&= p(A|\lambda_{ab}\wedge\lambda_{abc})\, p(B|\lambda_{ab}\wedge\lambda_{abc})
\end{aligned}
$$

10.0.0.40. Az a gondolat, hogy két esemény közötti korreláció magyarázata mindig a két esemény közös kauzális múltjában, mint „közös okban" keresendő, Reichenbachtól származik.[56] Reichenbach a $\Delta(A,B) \neq 0$ korrelációt magyarázó közös okot egy olyan C eseményként definiálta, amely kielégíti az alábbi feltételeket:

$$p(A\wedge B|C) = p(A|C)\, p(B|C) \tag{28}$$

$$p(A\wedge B|\neg C) = p(A|\neg C)\, p(B|\neg C) \tag{29}$$

Vegyük észre, hogy a (28)–(29) egyenletek ugyanazok, mint a (26) árnyékolási feltétel, arra a speciális esetre vonatkozóan, amikor a λ_{ab} paraméter két lehetséges értéket vehet fel, vagyis, amikor a „dolgok állását" egyszerűen két lehetséges osztályba soroljuk, történetesen a „C" és „nem C" osztályokba.[57]

Reichenbach a közös ok fogalmának fenti definícióját intuitív példákra alapozta. Egyik ilyen példa – kis módosítással – a következő: Képzeljünk el egy dobozt, amelybe be van szerelve két hagyományos izzólámpa. A lámpák időnként kiégnek, és megfigyeljük, hogy ez gyakrabban történik meg egyszerre, mint az a statisztikus függetlenség esetén várható lenne, vagyis $p(A\wedge B) > p(A)\, p(B)$, ahol A és B a két lámpa kiégését jelöli, mondjuk egy adott órában. A korrelációt az magyarázza meg, hogy időnként az elektromos hálózatban valami zavar támad, és a feszültség hirtelen megnő, s ez a zavar egyszerre megnöveli mindkét lámpa kiégésének valószínűségét. A korreláció magyarázatának valami olyasmit kell megmagyaráznia, hogy „honnan tudja az egyik lámpa, hogy most a másik lámpa nagy valószínűséggel kiég, ezért lehetőleg ő is kiég", ha feltevésünk szerint

[56] Reichenbach 1956, 19. fejezet.

[57] Természetesen, szemben a λ_{ab} paraméterrel a Reichenbach-féle közös ok „nem tud" a téridőbeli viszonyokról, a közös ok fogalmát csak valószínűségi szempontból ragadja meg.

az egyik lámpa nincs közvetlen hatással a másikra. Ha el akarjuk dönteni, hogy valóban a feszültségingadozás az a C esemény, ami a korrelációt okozza, akkor azt tehetjük, hogy a statisztikus sokaságot azokra az esetekre szűkítjük le, amikor például soha sincs áramingadozás. Ilyenkor már – egymástól függetlenül – csak a vak véletlenen múlik, hogy kiég-e az egyik vagy a másik égő. Ez tehát azt jelenti, hogy ezen a részsokaságon a korrelációnak el kell tűnnie. Teljesen hasonló eredményre jutunk, ha azt a részsokaságot vizsgáljuk, amikor mindig van áramingadozás. A korrelációnak ilyenkor is el kell tűnnie. Vagyis a C és $\neg C$ eseményekre vett kondicionális valószínűségekre nézve az A és B eseményeknek függetleneknek kell lenniük, azaz pontosan a (28)–(29) árnyékolási feltételeknek kell teljesülniük.

Reichenbach szerint, a definíció helyességét az intuitív példákon túl a következő (triviális) tétel is alátámasztja:

4. Tétel. *Legyen A,B és C három tetszőleges olyan esemény, melyekre teljesülnek a (28)–(29) feltételek. Ekkor*

$$\left.\begin{array}{l} p(A\wedge C) > p(A)\,p(C) \;\;\&\;\; p(B\wedge C) > p(A)\,p(C) \\ p(A\wedge C) < p(A)\,p(C) \;\;\&\;\; p(B\wedge C) < p(A)\,p(C) \end{array}\right. \Rightarrow \Delta(A,B) > 0$$

$$\left.\begin{array}{l} p(A\wedge C) > p(A)\,p(C) \;\;\&\;\; p(B\wedge C) < p(A)\,p(C) \\ p(A\wedge C) < p(A)\,p(C) \;\;\&\;\; p(B\wedge C) > p(A)\,p(C) \end{array}\right. \Rightarrow \Delta(A,B) < 0$$

10.0.0.41. A közös ok fenti fogalmát felhasználva Reichenbach „közösok-elv” néven a következő metafizikai tézist fogalmazta meg: *Bármely két korreláló eseményhez létezik a világban olyan esemény, amelyik a korrelációt megmagyarázó közös ok, a fenti értelemben.*

10.0.0.42. Hogy bizonyos korrelációk megmagyarázhatók-e közös okkal, vagy sem, az EPR–Bell, illetve a GHZ-kísérletek kapcsán kruciális kérdéssé vált a kvantummechanikában (lásd a **??**. és a **??**. fejezetet). Első pillantásra úgy tűnhet, hogy a közös ok fogalmának Reichenbach-féle valószínűségi leírása alkalmas arra, hogy a világban tapasztalt korrelációk közül kiragadjuk azokat, amelyek közös okkal megmagyarázhatók. Így a reichenbachi közös ok fogalom elemzése – explicite vagy implicite – a nyolcvanas és kilencvenes évek EPR–Bell-irodalmának középpontjába került.[58] Nem célunk itt ezeknek az eredményeknek az áttekintése. Elsősorban azért nem, mert a számos, sokszor messze nem triviális részletkérdés tisztázása után az derült ki, hogy a reichenbachi közös ok általános filozófiai szempontból éppúgy, mint a spin-korrelációs kísérletek leírása szempontjából alkalmatlan fogalom. Mindezt megvilágítandó, a következő rövid megjegyzéseket tesszük.

1. Nancy Cartwright helyesen mutat rá[59], hogy a közös ok definíciójában megkövetelt tulajdonságokat, jelesül a (28)–(29) árnyékolási feltételeket kizárólag a determinisztikus világból vett példákból olvastuk ki, vagyis olyan példákból, amelyekben a valószínűségek episztemikusan értelmezhetők. Semmiféle megalapozott ismeretünk nincs azt illetően, hogy mit kell tudnia a közös ok fogalmának egy objektíve indeterminisztikus világban.

2. A közös ok definíciója természetesen csak olyan szükséges feltételeket tartalmaz, melyeket a valóságos közös oknak ki kell elégítenie. Több, akár kontinuum sok olyan esemény létezhet, amelyek kielégítik ezeket a feltételeket (lásd a 4. megjegyzésben adott példát).

3. Egy közös okot valószínűségi szempontból öt adattal jellemezhetünk (a **10.0.0.40.** pont jelöléseit használva): $p(A|C)$, $p(A|\neg C)$, $p(B|C)$, $p(B|\neg C)$ és $p(C)$, melyek közül kettő független. Előfordulhat, hogy a vizsgált jelenségkört leíró valószínűségi modell nem tartalmaz olyan eseményt, amely egy

[58]Van Fraassen 1977, 1982, 1989; Salmon 1978, 1980, 1984; Skyrms 1984; Cartwright 1987; Butterfield 1989; Suppes 1990; E. Szabó 1993, 2000a; Hofer-Szabó, Rédei és E. Szabó 1999, 2000a, 2002.; Placek 2000; Rédei és Summers 2002; Rédei 2002; Gyenis és Rédei 2002.

[59]Uo.

adott korreláció közös oka lenne, vagyis nem tartalmaz a modell olyan eseményt, amelyik eleget tesz a Reichenbach-féle kritériumoknak és amelyre nézve a megfelelő valószínűségek a valamilyen más megfontolás alapján előírt $p(A|C)$, $p(A|\neg C)$, $p(B|C)$, $p(B|\neg C)$ és $p(C)$ értékekkel egyeznek meg. Bebizonyítható, hogy ilyen esetben a valószínűségi modell mindig kibővíthető úgy, hogy a bővebb modell már tartalmazza a megfelelő tulajdonságú közös okot. Sőt, a kibővítés korreláló eseménypárok tetszőleges véges

$$\{(A_i, B_i)\}_{i=1,2,\ldots N}$$

halmazára és tetszőleges

$$\{(p(A_i|C_i), p(A_i|\neg C_i), p(B_i|C_i), p(B_i|\neg C_i), p(C_i))\}_{i=1,2,\ldots N}$$

típusokra elvégezhető úgy, hogy a kibővítés mindegyik korrelációra nézve tartalmazzon egy megfelelő típusú közös okot.[60] Ez azt jelenti tehát, hogy nincs valószínűségelméleti akadálya annak, hogy tetszőleges korrelációkat közös okkal magyarázzunk meg.[61]

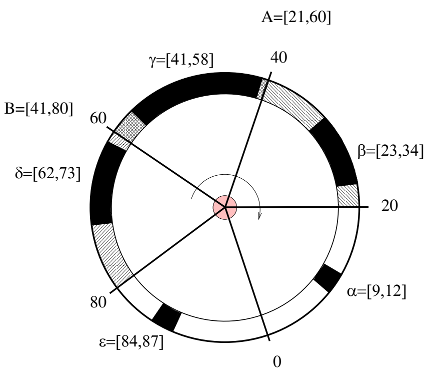
4. Sokkal problematikusabb azonban a Reichenbach-féle közös ok jelentése. Vizsgáljuk meg a következő egyszerű példát. Mari és Kati szeretnek rulettezni és gyakran játszanak. Mindkettőjüknek nagyon egyszerű stratégiája van: Mari minden alkalommal az $A = [21, 60]$, Kati pedig a $B = [41, 80]$ intervallumra tesz (13. ábra). Sokáig játszva, észreveszik, hogy nyeréseik között pozitív korreláció van:

$$p(A \wedge B) - p(A)\,p(B) = 0.2 - 0.4 \times 0.4 = 0.04$$

[60] Hofer-Szabó, Rédei és E. Szabó 1999.

[61] Ezzel nem állítjuk azt, hogy a releváns kauzális modelljeink közösok-zárttá tehetők, abban az értelemben, hogy minden korrelációnak lenne benne közös oka. (A részletekről lásd Gyenis és Rédei 2002.)

$$p(C) = \frac{4+4+12+12+18}{100} = 0.5$$



$$C = \alpha \cup \beta \cup \gamma \cup \delta \cup \varepsilon$$

$$\begin{aligned}
p(A \wedge B|C) &= \frac{18/100}{0.5} = 0.36 \\
&= \left(\frac{(18+12)/100}{0.5}\right)^2 \\
&= p(A|C)\,p(B|C) \\
p(A \wedge B|\neg C) &= \frac{2/100}{0.5} = 0.04 \\
&= \left(\frac{(2+8)/100}{0.5}\right)^2 \\
&= p(A|\neg C)\,p(B|\neg C)
\end{aligned}$$

13. ábra. *Mari és Kati ruletteznek. Mari minden alkalommal az* $A = [21, 60]$, *Kati pedig a* $B = [41, 80]$ *intervallumra fogad. Nyeréseik korrelálnak. Mint a fenti számolás mutatja, a feketével jelölt tartomány teljesíti a Reichenbach-féle feltételeket*

Izgatja őket, hogy mi ennek a magyarázata, és megkérdezik Reichenbach legjobb tanítványát, aki a következő választ adja: Azért van korreláció a nyeréseitek között, mert a rulett golyó időnként (0.5 valószínűséggel) az ábrán feketével jelölt

$$C = [9,12] \cup [23,34] \cup [41,58] \cup [62,73] \cup [84,87]$$

tartományban áll meg.

El lehet képzelni a két lány arckifejezését, amikor ezt a választ hallják! Nem csak, és elsősorban nem az fogja okozni megrökönyödésüket, hogy a Reichenbach-tanítvány kontinuum sok különböző ilyen halmazt jelölhetett volna meg (az öt fekete tartományt tetszés szerint eltolhatjuk az egyes szektoron belül), hanem hogy egy ilyen C egy „Cambridge event”, aminek nyilvánvalóan semmi szerepe nincs a világ kauzális rendjében, s aligha lehet őt értelmesen felhasználni a korreláció létrejöttének kauzális magyarázatában.

A korreláció létrejöttének kauzális magyarázata sokkal inkább abban áll, hogy a golyó bizonyos valószínűséggel ezen vagy azon a számon áll meg (történetesen minden szám valószínűsége $\frac{1}{100}$), és ha egy számon megállt, az egyértelműen meghatározza, hogy melyik lány nyer és melyik nem. Vegyük észre azonban, hogy ezeknek a <0>, <1>,... <99> eseményeknek egyike sem elégíti ki a Reichenbach által előírt feltételeket, hiszen csak (28) teljesül, (29) nem. A <0>, <1>,... <99> események az egységelemnek olyan diszjunkt partícióját alkotják viszont, melynek minden eleme teljesíti a (28) árnyékolási feltételt.

5. Gondolatmenetünket folytatva, vegyünk egy másik esetet. A **10.0.0.40.** pontban használt példát módosítsuk úgy, hogy a lámpák kiégése közötti korrelációt nem a feszültségingadozás (C) okozza – mert garantáltan állandó a feszültség –, hanem mondjuk az, hogy a lámpákat tartalmazó dobozt időnként kala-

pácsütés éri (D). D tökéletes reichenbachi közös ok, s kielégíti a (28)–(29) feltételeket. Mi a helyzet azonban akkor, ha mindkét esemény, C is és D is megtörténhet? Világos, hogy mindkét esemény egy-egy kauzális faktor a korreláció létrejöttében. Ugyanakkor, jól értett okokból egyik sem fogja kielégíteni az árnyékolási feltételeket: ha például a statisztikai sokaságot leszűkítjük azokra az esetekre, amikor C nem következik be, akkor most nem kell a korrelációnak eltűnnie, hiszen D bekövetkezése vagy be nem következése még eredményezhet korrelációt A és B között. El kell tűnnie azonban a korrelációnak a $C \wedge D$, $C \wedge \neg D$, $\neg C \wedge D$ és $\neg C \wedge \neg D$ kondíciókra nézve. Vagyis megint, $C \wedge D$, $C \wedge \neg D$, $\neg C \wedge D$, $\neg C \wedge \neg D$ egy olyan egységpartíciót képez, melynek minden eleme teljesíti a (28) árnyékolási feltételt.

6. Azt látjuk tehát, hogy az eredeti Reichenbach-féle koncepció – noha bizonyos esetekben jól alkalmazható – általában alkalmatlan a korrelációk eredetének kauzális magyarázatára, s hogy helyette a reichenbachi közös ok fogalom egy általánosítására van szükség. Nevezzük ezt az általánosított fogalmat *közösok-rendszernek*, amely tehát eseményeknek egy olyan $\{C_i\}_{i=1,2,\ldots}$ rendszere, amely eleget tesz a következő feltételeknek:

$$\bigcup_i C_i = \mathbf{1}$$

$$(\forall i \neq j) \quad [C_i \wedge C_j = \emptyset]$$

$$(\forall i) \quad [p(A \wedge B|C_i) = p(A|C_i)\, p(B|C_i)]$$

7. A Reichenbach-féle közösok-elvet úgy illő tehát módosítanunk, hogy *bármely két korreláló eseményhez létezik a világban olyan eseményrendszer, amelyik a korrelációt, a fenti értelemben megmagyarázó közösok-rendszert alkot.*

8. Csakhogy az elv ebben a formájában metafizikai értelemben üres. Könnyen belátható ugyanis, hogy

5. Tétel. *Tetszőleges* $(\mathcal{A}, p)$ *valószínűségi modellben, korreláló párok tetszőleges véges halmazához létezik közös közösok-rendszer.*

Bizonyítás. Tekintsük ugyanis a korreláló párokban előforduló események halmazát, s vegyük az ezeket tartalmazó részhálót $\mathcal{A}$-ban. A végesség miatt ez a részháló garantáltan atomos, s az atomok halmaza olyan közösok-rendszert alkot, amely egyszerre közösok-rendszere az összes korrelációnak.

Azt mondani tehát, hogy a $p(A \wedge B) \neq p(A)\, p(B)$ korrelációnak létezik közös oka (értsd: közösok-rendszere) a reichenbachi értelemben, az nem egy szintetikus ítélet, hanem analitikus. Egyszerű logikai következménye annak a ténynek, hogy $p(A)$, $p(B)$ és $p(A \wedge B)$ valószínűségeket jelölnek.

10.0.0.43. Nem meglepő, hogy a közös ok fogalmának Reichenbach-féle értelmezése semmitmondónak bizonyul, hiszen semmit sem ragad meg a partikuláris események közötti, ontológiai értelemben vett kauzális viszonyokból. Nem lenne azonban helyes, ha a közös ok Reichenbach-féle statisztikus értelmezésének elvetésével együtt elvetnénk a Reichenbach által megfogalmazott közösok-elvet is, abban az általános metafizikai értelemben, ahogyan azt a **10.0.0.34.** pontban megfogalmaztuk, hogy tehát nincs korreláció kauzalitás nélkül, vagyis bármely két esemény között tapasztalt korreláció a két esemény kauzális múltjának közös részéből vezethető le, annak alapján érthető meg.

11. előadás

Determinizmus–indeterminizmus

12. Mi a determinizmus?

Megkíséreljük összefoglalni, mit is értünk determinizmus alatt. Kiindulásképpen álljon itt két jelentős filozófus víziója, melyek jól szemléltetik, hogyan szokás egy determinisztikus világot elképzelni. William James így ír:

> Mit tanít a determinizmus? Azt tanítja, hogy az univerzumnak a már lejátszódott része teljes egészében meghatározza, hogy a maradék része milyen lesz. A jövő nem rejteget a méhében különféle lehetőségeket: az a rész, amit jelennek hívunk, csak egyetlen totalitással kompatibilis. Az örökkévalóság által kiválasztott egyetlen jövőbeli komplementeren kívül minden más lehetetlen.[62]

Karl Popper szemléletes képe pedig a következő:

> A determinizmus intuitív fogalmát úgy foglalhatjuk össze, hogy a világ olyan, mint egy mozgófilm: az éppen kivetített kocka a *jelen*. A filmnek az a része, amelyet már levetítettek, a *múlt*. Az a része pedig, amelyiket még nem láthattuk, a *jövő*.[63]

Mindkét vízió a múlt, jelen és jövő fogalmával operál, s mint az előző fejezetekben láttuk e fogalmak meglehetősen problematikusak. Azután további tisztázásra szorul, mit is jelent Jamesnél az, hogy a jelen csak egyetlen totalitással „kompatibilis”. Hogy e „kompatibilitást” értelmezzük, nem mást kell tennünk, mint definiálnunk, hogy mi a determinizmus.

---

[62] Earman 1986, 4. o.

[63] Popper 1988, 5. o.

12.0.0.44. De Popper film-víziója sem kevésbé problematikus.[64] Hiszen a kérdés éppen az, hogy determinálja-e valami azt, hogy a filmen mi van. Vagy, hogy a film eleje meghatározza-e a film folytatását. Mert, hogy a filmen valami van, hogy egy ilyen film létezik, az nyilvánvaló. *Che Sera, Sera.*[65] – idézik ilyenkor a jól ismert olasz dalt. És ez a filozófusok nagy táborának elégséges ahhoz, hogy a világot determinisztikusnak gondolja:

> Bár a „determináltnak” az az értelme, mely szerint a jövő determinált pusztán azért, mert az lesz, ami lesz, elegendő (legalábbis én így vélem) a determinizmus egyes ellenfeleinek – nevezetesen Bergsonnak és a pragmatistáknak – a megcáfolásához, a legtöbb ember mégsem erre az értelemre gondol, amikor úgy beszél a jövőről, mint ami determinált. Ilyenkor voltaképpen egy formulára gondolnak, amelynek segítségével a jövőt úgy lehet ábrázolni – s legalábbis elvileg ki lehet számítani – mint a múlt függvényét. E ponton azonban nagy nehézségekkel találjuk magunkat szembe...[66]

Mielőtt megismerkednénk azzal, hogy Russell milyen nehézségekre gondol, néhány idézet erejéig tekintsük át, hogy a determinizmus fogalmának milyen főbb megközelítésmódjai ismertek.

Megjósolhatóság, kiszámíthatóság

A determinizmus lényege a jövő eseményeinek megjósolhatósága, kiszámíthatósága. Itt a terminológiát tovább kell finomítanunk. Hiszen nem mindegy, hogy azt állítjuk, hogy a világ olyan, hogy a jövő eseményei *elvileg* megjósolhatók, vagy pedig azt, hogy *megjósolhatók.* Az előzőre példa Laplace determinizmus-felfogása:

[64] Popper maga is csak egy felvezető, intuitív gondolatnak szánta.

[65] Ahogy lesz, úgy lesz.

[66] Russell 1976, 329. o.

> ... a világmindenség jelen állapota az előző állapot okozatának és az eljövendő állapot okának tekintendő. Az olyan értelem, mely egy bizonyos pillanatban a természet összes erőit és az azt összetevő egységek helyzetét ismerné, mely továbbá eléggé mélyreható volna ezen adatok elemzésére, egyazon képletbe foglalhatná a világ legnagyobb testének és legkönnyebb atomjának mozgását. Semmi sem volna bizonytalan előtte; a jelen módjára látná a jövőt, éppúgy, mint a múltat. Az emberi értelem a csillagászatban elért tökéletességgel halvány vázául szolgálhat egy ilyen értelemnek... Az igazság keresése érdekében kifejtett minden erőlködése arra irányul, hogy minél jobban megközelítse a fent elképzelt értelmet.[67]

A praktikus, tudományos gyakorlat értelmében vett megjósolhatóságra példa Karl Popper determinizmus-definíciója:

> (A tudományos determinizmus) az a doktrína, mely szerint egy zárt fizikai rendszer tetszőleges jövő pillanatban vett állapotát, tetszőleges pontossággal meg lehet jósolni, még akár a rendszeren belülről is, oly módon, hogy a jóslatainkat tudományos elméletekből vezetjük le, a kezdeti feltételek ismeretében, melyek szükséges pontossága mindig kiszámolható, ... ha tudjuk, hogy milyen jóslási feladatról van szó. A démon, csakúgy, mint egy tudós ember, nem kell, hogy abszolút matematikai precizitással ismerje a kezdeti feltételeket; a tudóshoz hasonlóan meg kell elégednie véges pontossággal.[68]

---

[67]Eddington 1935, 75. o.

[68]Earman 1986, 8. o.

Funkcionális kapcsolat az időszeletek között

Grünbaum megfogalmazásában a determinizmus a következőt jelenti:

> Mi a különbség e kétféle [determinisztikus és indeterminisztikus] világ között a jövőbeli események determináltsága szempontjából? A különbség abban a *funkcionális kapcsolatban* [Kiemelés tőlem.] van, amely összeköti a jövőbeli események attribútumait a jelenbeli, illetve múltbeli események attribútumaival.[69]

Russell ezt a funkcionális kapcsolatot a következőképpen írja le:

> Egy rendszerről akkor mondjuk azt, hogy „determinisztikus", ha – feltéve, hogy adva vannak bizonyos $t_1, t_2, t_3, \ldots t_n$ időpontban e rendszerre vonatkozó $e_1, e_2, e_3, \ldots e_n$ adatok, és E_t a rendszer állapota valamely tetszőleges t időpontban – van egy olyan funkcionális viszony, mely szerint
>
> $$E_t = f(e_1, t_1, e_2, t_2, \ldots e_n, t_n, t)$$
>
> A rendszer „egy egész időszak folyamán determinisztikus" lesz, ha a fenti formulában szereplő t bármely időpont lehet ezen időszakon belül... Ha a világegyetem, mint egész, egy ilyen rendszert alkot, a determinizmus igaz a világegyetemre nézve; ha nem alkot ilyen rendszert, akkor nem igaz.[70]

Primitív modális szerkezet

Nuel Belnap a következőképpen írja le az objektív modalitást:

---

[69] Grünbaum 1974, 324. o.

[70] Russell 1976, 322. o.

> A világ történetének egy adott pillanatában többféle lehetőség létezik arra, hogy a dolgok hogyan folytatódjanak. Mielőtt feldobjuk az érmét két lehetséges dolog is történhet, Fej vagy Írás. E két lehetőség nem csupán episztemikus értelemben létezik, hanem a valóságban.[71]

Ennek megfelelően, a világ akkor determinisztikus, ha a modális szerkezete a lehető legprimitívebb, vagyis egyetlen szálon fut. Elágazási pontok nincsenek.

12.0.0.45. E meghatározások mindegyike különböző problémákat vet fel. A nehézség, amelyre Russell utal a következő:

> Ha a megengedett formulák bonyolultságának foka tetszőlegesen nagy lehet, akkor – úgy tetszik – minden rendszernek, melynek egy adott pillanatban vett állapota bizonyos mérhető mennyiségek függvénye, determinisztikusnak kell lennie. Illusztrációként tekintsünk egyetlen anyagi részecskét, amelynek koordinátái t időpontban legyenek x_t, y_t, és z_t. Ekkor bárhogy mozog is a részecske, elméletileg kell lennie olyan f_1, f_2, f_3 függvényeknek, hogy
>
> $$\begin{aligned} x_1 &= f_1(t) \\ x_2 &= f_2(t) \\ x_3 &= f_3(t) \end{aligned}$$
>
> Következésképpen elvileg lehetségesnek kell lennie annak, hogy az anyagi világegyetem t időpontban vett egész állapotát t függvényeként ábrázoljuk. Így világegyetemünk a fentebb definiált értelemben determinisztikus lesz. Viszont ha ez igaz, akkor semmiféle információt nem adunk a világegyetemről, amikor azt állítjuk róla,

[71]Belnap 1992.

> hogy determinisztikus. Igaz, a szóban forgó formulák szigorúan végtelen bonyolultságúak lehetnek, tehát gyakorlatilag nem lehet sem leírni, sem felfogni őket. Ez azonban, ha eltekintünk tudásunk szempontjától, részletkérdésnek tetszhet: az anyagi világegyetemnek – amennyiben a fenti megfontolások helyesek – önmagában véve determinisztikusnak kell lennie, törvényeknek kell engedelmeskednie.[72]

A determinizmus Laplace- illetve Popper-féle értelmezése szempontjából a kérdés az, hogy ezeknek a függvényeknek a bonyolultsága milyen fokú lehet ahhoz, hogy a világ jövőbeli állapotai megjósolhatóak, vagy legalább elvileg megjósolhatóak legyenek. Az egyik lehetséges válasz erre a kérdésre, hogy a szóban forgó függvények legyenek kiszámíthatóak, azaz létezzen olyan algoritmus (hogy most ne kelljen belemennünk az algoritmusok elméletének részleteibe, a szemléletesség kedvéért mondjuk azt, hogy olyan véges idő alatt lefuttatható számítógépprogram), melynek input adatai a világ állapotát jellemző adatok összessége egy korábbi pillanatban, outputja pedig ezeknek az adatoknak az értéke egy későbbi pillanatban. Mint az ilyen algoritmusok elméletéből kiderül azonban, ez a követelmény túlságosan erős. Ebben az értelemben például nem lennének determinisztikusak az olyan fizikai elméletek, mint a hidrodinamika, vagy az elektrodinamika (általában a parciális differenciálegyenletekkel leírt dinamikai rendszerek[73]). A determinizmusnak primitív modális szerkezetként való értelmezése ezzel szemben annyira általános, hogy az általánosságnak ezen a szintjén a világ determinisztikus volta sohasem zárható ki. Ugyanis minden empirikus információnk a világról olyan, hogy kizárólag csak az éppen aktuális történethez tartozik, és ezeknek az empirikus adatoknak az alapján sem kizárni, sem megerősíteni nem tudjuk, hogy az aktuális történetnek létezik-e al-

[72] Uo. 329. o.

[73] Earman 1986.

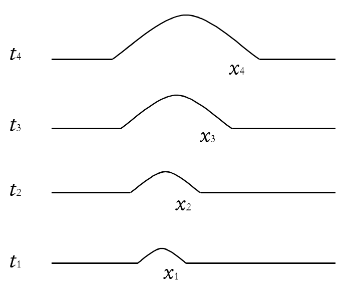
ternatívája. A determinizmus fogalmának fenti megfogalmazásaival kapcsolatban felmerül az a nehézség is, hogy egy részük olyan fogalmakkal operál, mint „idő”, „időszelet”, „a világ állapota egy adott pillanatban”. Ezek a fogalmak, mint korábban láttuk, meglehetősen problematikusak.

A determinizmusnak talán használhatóbb fogalmát kapjuk, ha azt definiáljuk először, hogy egy elméletet mikor nevezünk determinisztikusnak. Végső soron minden fizikai – vagy bármilyen más, a világ leírását célzó elmélet – a következő struktúrájú. Feltételez – gyakran csak implicite – egy $\mathcal{W}$ halmazt, a lehetséges világok halmazát, és megad egy Γ sablont, melynek az a funkciója, hogy a lehetséges világok közül kiválassza az aktuálisat. Az elméletet akkor fogjuk determinisztikusnak nevezni, ha csak egyetlen olyan $W \in \mathcal{W}$ lehetséges világ van, amelyik a Γ sablonnal összeegyeztethető. Például az elektrodinamika egy olyan elmélet, amelyben a lehetséges világok halmaza az összes lehetséges $\vec{E}(\vec{x},t)$ és $\vec{B}(\vec{x},t)$ térkonfigurációkkal van paraméterezve, a sablon pedig nem más, mint a Maxwell-egyenleteknek és a Cauchy-adatoknak a rendszere. Az elektrodinamika tehát determinisztikus, mert az egyenleteknek egyetlen megoldása létezik, amelyik a Cauchy-adatokkal is kompatibilis. Ha egy jelenségkörnek ismerünk determinisztikus elmélettel történő leírását, akkor a szóban forgó jelenségkör objektíve is determinisztikus. Ha nem, akkor csak annyit mondhatunk, hogy nem biztos, hogy determinisztikus, hiszen nem zárható ki, hogy létezik a jelenségkörnek egy determinisztikus leírása. Mint majd látni fogjuk a kvantummechanikáról szóló fejezetekben, a kvantummechanika önmagában nem kérdőjelezi meg a determinizmust, csupán azért, mert a világról egy nem determinisztikus, valószínűségi leírást ad. A determinizmus elleni komoly kihívás a kvantummechanikával kapcsolatban levezetett, úgynevezett *no go* tételekkel kezdődik, melyek azt állítják, hogy a kvantummechanikával leírt jelenségkörnek – általában – *nem létezhet* determinisztikus (pontosabban determinisztikus és lokális, lásd a **??.**, **??.** és **13.0.0.46.** pontokat) elmélete.

13. Determinizmus és lokalitás

A fizikai hatások véges sebessége

A „klasszikus” fizika világképének nem csak a determinizmus, hanem a *lokalitás* is természetes tartozéka, vagyis az az idea, mely szerint nincsenek távolhatások. Itt természetesen nem a newtoni fizika fogalmaira, hanem a 19. század végére kialakult, tehát az úgynevezett „modern” fizikát közvetlenül megelőző állapotra gondolok. A 19. század végén a fizikai elmélet prototípusa az elektrodinamika volt. Az elektrodinamika nem csak determinisztikus fejlődését írja le az elektromágneses mezőnek, hanem lokális is, abban az értelemben, hogy a mező egy adott tértartományban vett konfigurációját „megváltoztatva” a Maxwell-egyenletek megoldása úgy változik meg, hogy ez a változás kezdetben csak e tartomány kis környezetében jelenik meg, majd véges sebességgel (történetesen a fény sebességével) terjed szét (14. ábra). A tényt, hogy tapasztalatunk szerint *nincs olyan*



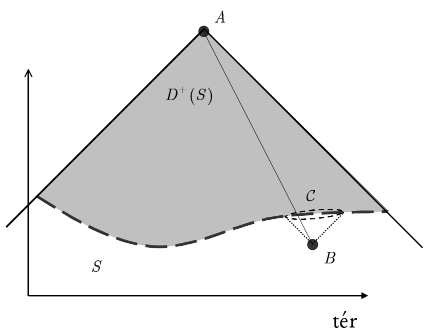
14. ábra. *Az elektromágneses mező konfigurációját egy kis tartományban az x_1 helyen megváltoztatjuk a t_1 pillanathoz tartozó Cauchy-felületen. A megváltozott Cauchy-adatokhoz egy új megoldása tartozik a téregyenleteknek. Ez az új megoldás olyan, hogy az eredeti megoldástól való eltérés időszeletről időszeletre fokozatosan nagyobb tartományban mutatkozik*

fizikai hatás, amelyik a fénynél gyorsabban terjedne – a kvantummechanikában majd felmerül, hogy látunk-e ilyet –, univerzálisan érvényes törvénynek feltételezzük. Világosan kell látni, hogy ez a törvény nem a relativitáselmélet kizárólagos sajátja, és logikailag független is a téridő struktúrájáról tett kijelentéseinktől. Hogy semmilyen fizikai hatás nem terjedhet a fénynél gyorsabban a Lorentz-elméletben természetesen nem fordítható le arra, hogy a „fénysebesség” nevű, c-vel jelölt természeti állandó *határsebesség*. Amint azt a **??.** pontban megmutattuk, a Lorentz-elmélet és a relativitáselmélet nem ugyanazt a fizikai mennyiséget nevezi sebességnek: a relativitáselmélet által használt „deformált sebesség” fogalomra természetesen – a Lorentz-elmélet szerint is – fennáll, hogy nem lehet nagyobb c-nél, míg a Lorentz-elmélet által sebességnek nevezett hagyományos sebességfogalomra, a (**??**) összeadási formula értelmében, – a Lorentz-elmélet szerint és a relativitáselmélet szerint egyaránt – nem.

Markovitás

A lokalitás fogalma a klasszikus fizikában kialakult intuíciónk szerint nem csak a hatások véges terjedését feltételezi, hanem magában foglal egy ettől logikailag független másik feltevést is. Az előbbi elektrodinamikai példánknál maradva, nem lehetséges ugyanis, hogy egy adott (t_1, x_1) helyen megváltoztatva a mező konfigurációját (14. ábra), e változás a téregyenletek megoldásának olyan módosulását eredményezze egy távoli (t_4, x_4) téridő-tartományban, hogy közben, e téridő-tartományokat szétválasztó Cauchy-felületek közül, akár csak egyiken is, a mező-konfiguráció változatlan maradjon. Szemléletesebben szólva, a múltbeli mező-konfigurációk a jövőbelieket csak a jelenbeli mező-konfigurációkon keresztül determinálhatják. Vagyis a múltbeli egész történet – a jővőbeli történet determinációja szempontjából – a jelenbeli Cauchy-adatok összességében reprezentálódik. A rendszer a múltjának emlékét a jelenbeli állapotában reprezentálva őrzi. Az ilyen rendszereket szokás Markov-féle rendszereknek nevezni. Természetesen ez nem azt jelenti, hogy nincs olyan rendszer, amelyiknek memóriája van, de – például – az emberi agy is a jelenbeli állapotában *reprezentálva* őrzi a múltbeli állapotainak emlékét. A valóság leírásában időnként alkalmazunk nem markovi modelleket. A lokalitás elve szerint azonban, ha a szóban forgó jelenség mélyére nézünk, és egy részletesebb leírását adjuk meg, akkor annak markovinak kell lennie.

13.0.0.46. A lokalitásnak e két aspektusát együttesen a 15. ábrán látható téridő diagramon[74] szemléltethetjük: Az S felület $D^+(S)$ (jövő-



15. ábra. *A téridő $D^+(S)$ tartományában mindent meghatároznak az S Cauchy-felületen korábban fixálódott adatok. $D^+(S)$-t az S felület (jövő-) dependencia tartományának nevezzük*

)dependencia tartományában[75] mindent meghatároznak az S Cauchy-felületen korábban fixálódott adatok. Az A eseményre egy távoli és múltbeli B esemény hatással lehet. Ez a hatás azonban kifejeződik azokban a Cauchy-adatokban, amelyek a $\mathcal{C}$ tartományra vonatkoznak.

[74] Mostantól – hacsak külön nem jelzem – minden téridő diagramot olyannak gondolok, hogy az a Lorentz-elméletre is és a relativitáselméletre is érvényes.

[75] Hawking és Ellis 1973, Wald 1984

14. előadás

Szabad akarat

A szabad akarat problémájának kontextusa

Karl Poppert idézve már utaltunk rá, hogy a determinizmus kérdése szorosan kapcsolódik a szabad akarat problémájához, vagyis ahhoz a szerteágazó metafizikai problémához, hogy ha egy ember valamilyen körülmények között, egy adott pillanatban így vagy úgy dönt, ezt vagy azt cselekszi, vagy gondolja, mindezt „szabadon” teszi-e, vagy a világ és benne az ember olyan, hogy ezt a döntést, cselekedetet vagy gondolatot valamilyen módon valami determinálja. Más szóval, gondolhatna-e mást, dönthetne-e másképpen, cselekedhetne-e másként, mint ahogyan azt teszi?

Nyilvánvaló, hogy az akaratszabadságot illető metafizikai meggyőződés kiinduló pontja lehet számos morális, jogfilozófiai vagy akár esztétikai megfontolásnak, s e megfontolások visszahatnak a szabad akaratra vonatkozó metafizikai gondolkodásra. Minthogy tárgyalásunk elsődleges célja a szabad akarat és a determinizmus-indeterminizmus probléma viszonyának elemzése, anélkül, hogy tagadnánk e tágabb kontextus metafizikai jelentőségét, igyekszünk olyan példákat tekinteni, amelyben a morális felelősség kérdése nem játszik szerepet. A tudat működésének determinisztikus vagy nem determinisztikus jellege, mutat rá Ted Honderich, nem múlhat azon, hogy mi e működés morális kontextusa. Searle-lel polemizálva[76] a következőket írja:

> Searle egyfajta szimultán kapcsolatot tételez fel az agy és az elme neurális, illetve mentális állapotai között. Tehát (1) egy csinos nő látványa, mint a *perceptuális* tudat egy eleme, együtt jár bizonyos szimultán neurális állapottal, csakúgy, mint (2) a *reflektív* tudat ezt követő eleme, az arra való emlékezés, hogy már nős vagy. Searle szerint e korábbi állapotok valamilyen módon kapcsolódnak a későbbi állapotokhoz – mondjuk (3) az *affektív* tudat azon

[76] Searle 2000.

eleméhez, ahhoz a mentális eseményhez, hogy úgy döntesz, meghívod a hölgyet egy italra.

Figyelembe véve az agykutatás bizonyos eredményeit, na meg egy sereg filozófiai megfontolás alapján, Searle megengedi, hogy a neurális állapotok és a velük szimultán tudati állapotok, illetve események között standard kauzális kapcsolat álljon fenn. A neurális állapot okozza a szimultán tudati állapotot. Vagyis létezik egy lentről felfelé irányuló kauzalitás.

Nagyjából hasonló okok miatt – agykutatási eredmények meg a többi – a *perceptuális* és a *reflektív* tudat vonatkozásában a standard kauzális mechanizmusok működnek. A csinos nő látványának a tudatban való megjelenése egy standard okozat, éppúgy, mint az a gondolat, hogy nős vagy. Mellőzve a további részleteket, van tehát egy bal-jobb irányú standard kauzalitás is.

De amint egy *döntés* eredetéről, egyáltalán, bármiféle döntésről van szó, az affektív tudatban, *nincs* standard kauzalitás a döntés neurális megfelelőjét illetően. Nincs semmi, aminek az okozata lett volna az a neurális állapot, amely a döntéssel járt együtt, hogy meghívtad egy italra. Ebben az esetben egy, a kvantumelmélet szokásos interpretációja szerint feltételezett, véletlenszerű kapcsolatról van szó.

Ez borzalmas!

Le-föl kauzalitás mindenütt, de véletlen egyes bal-jobb kapcsolatokban.

Bal-jobb kauzalitás a perceptuális és a reflektív tudattal kapcsolatban, de bal-jobb véletlenszerűség a döntéseket illetően.

... más szóval, az agy következetesen egy gép lentről felfelé, de nem mindig viselkedik gépként balról jobbra.

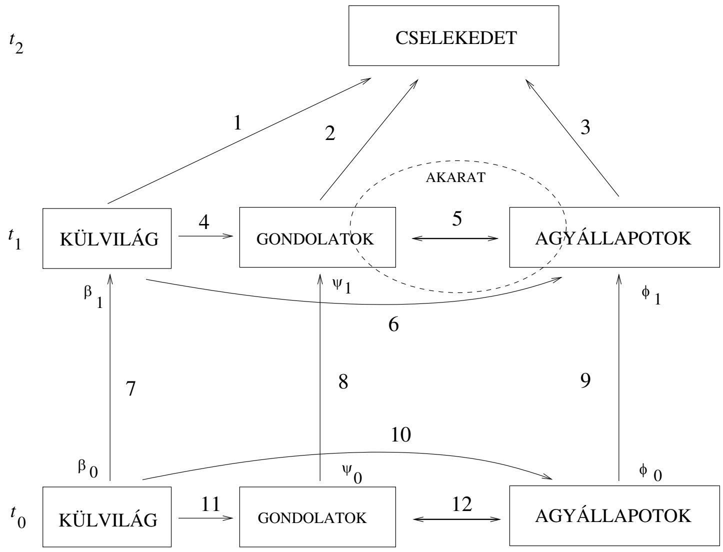
> És ha csak a bal-jobb működést tekintjük, és figyelembe vesszük a perceptuális, a reflektív és az affektív tudatot, az agy egyszer gép, másszor meg nem az. A tények és a tapasztalatok tükrében ez számomra teljesen abszurdnak tűnik, melyet az agykutatás eredményei a legcsekélyebb mértékben sem támasztanak alá.[77]

Mindennek tükrében a probléma felvetéséhez tekintsük a következő egyszerű példát: Egy kísérleti személyt az elé a döntés elé állítunk, hogy vagy a piros, vagy a kék gombot nyomja meg. Döntésének nincs semmi különösebb következménye. Ha a piros gombot nyomja meg, akkor a piros, ha a kéket, akkor a kék lámpa villan fel. Kísérleti alanyunk a piros gombot nyomta meg. Kérdés, szabadon döntött-e? Más szóval, dönthetett volna-e úgy, hogy a kéket nyomja meg?

14.0.0.47. A kérdés ilyen megfogalmazásával már állást is foglaltunk azt illetően, hogy körülbelül mit értünk szabad akarat alatt. Ez azért fontos, mert – noha egy-egy dolog metafizikai elemzése természetesen magában foglalja az arról való elmélkedést is, hogy miben is áll a hétköznapi nyelvhasználatban így és így nevezett, az ember mindennapos életében így és így megélt, megtapasztalt jelenség – az egymással polemizáló nézetek összevetése során pontosan kell értenünk az egyes irányzatok koncepcionális és terminológiai különbségeit.

E terminológiai különbségek tisztázását szolgálja a 16. ábra. Meg fogjuk különböztetni a külvilág, az agy – és ha szükséges – a tudat állapotait a t_0 illetve a t_1 időpillanatban, illetve a vizsgált személy cselekvését a t_2 pillanatban. A számozott nyilak mindegyike valamilyen fajta determinációt, illetve időfejlődést szimbolizál. Nyilvánvaló, hogy az ágens t_2-ben bekövetkező cselekvését valamilyen módon a külvilág, az agy és a tudat azt közvetlenül megelőző t_1 pillanatban vett állapota határozza meg. Hogy hogyan, az a

[77] Honderich 2002.



16. ábra. *A szabad akarat problémájának kontextusa*

cselekvés szabadságának kérdése, melyet gyakran kevernek össze az *akarat* szabadságának kérdésével. Nem kétséges, hogy a cselekvés szabadságának problémája is fontos, és az is összefüggésbe hozható a determinizmus-indeterminizmus ügyével, s hogy amit az akarat szabadságával kapcsolatban mondani fogunk, talán mind elmondható lenne a cselekvés szabadságáról is, azzal a nyilvánvaló különbséggel, hogy a 2. és 3. nyíllal jelzett kapcsolatok nem annyira az agy és az elme, mint inkább a motorikus idegrendszer működésére és más fiziológiai tényezőkre utalnak. Nem témánk továbbá a szabadság problémája abban az 1. nyíllal kifejezett értelemben, hogy tudniillik autonóm módon, saját elhatározásunkból, a külső körülmények akadályozó vagy kényszerítő hatásától függetlenül cselekedhetünk-e. A külvilág tényei természetesen befolyásolják cselekvésünket a 4. és 6. kapcsolaton keresztül. Valahogy úgy, mint a Csipkerózsika szakácsát ama 100 évre felfüggesztett pofon lekeverése előtti pillanatban. A szándék kialakulásában, hogy nyakon vágja a kuktát, nyilván szerepe volt a külvilág olyan közvetlen mentális üzeneteinek, mint az odakozmált rántás, saját inaséveinek emléke, vagy a királyi udvar vele szembeni elvárásai (4. nyíl), valamint az agyára ható olyan külső körülménynek, mint a hőség a konyhában (6. nyíl). Ha a szakács cselekvése szabad, autonóm cselekvés lett volna, a kialakult szándékot rögvest a motorikus kielégülés követi. A külvilág azonban a pofonra lendülő kart megbénító varázslat (1. nyíl) formájában hatással volt a cselekedetre.

A 2. nyíl feltételezhetően nem létezik, hiszen még egy test–elme dualizmus vagy paralelizmus esetén is feltételezhető, hogy a mentális kizárólag az agy közvetítésével hat az idegrendszernek a cselekvést végrehajtó motorikus részeire.

Metafizikai szempontból a cselekvés szabadsága helyett sokkal izgalmasabb kérdés tehát az akarat szabadsága. Akarat alatt a cselekvést közvetlenül megelőző, t_1 időpillanathoz tartozó mentális állapotot (pontosabban talán – bár ennek nincs különösebb jelentősége – ennek a mentális állapotnak a cselekvésre irányuló – ha tetszik, intencionális – komponensét) értjük, illetve az ennek megfelelő agyállapotot. A kérdés az, determinálja-e valami ezt a t_1 időpillanatbani mentális/agyi állapotot, és ha igen, akkor mi és hogyan. A külvilág hatással lehet erre az állapotra (4. és 6. nyíl), ám amikor azt firtatjuk, gondolhatta-e, akarhatta-e valaki másképpen, akkor ezt *ceteris paribus* értjük, vagyis – függetlenül attól, hogy a külvilág időfejlődése (7. nyíl) determinisztikus vagy nem – a külvilág hatását nem kell figyelembe vennünk. Mert például a morális felelősség szempontjából egyetlen libertariánus[78] sem érezné megnyugtatónak, ha az lenne a helyzet, hogy a gyilkos az adott külső körülmények hatására nem cselekedhetett ugyan másképpen, mint hogy megöli áldozatát, ám mégis elítéljük, mondván, cselekedhetett volna másképpen, ha aznap a légköri folyamatok másképpen alakulnak, és nincs akkora hőség. Vagyis ha az akarat szabadsága csupán abban a modalitásban merülne ki, hogy a döntést egyébként teljesen determináló külső hatások lehettek volna másmilyenek is.

Konklúziónk tehát az, hogy az akarat szabadsága azon áll vagy bukik, hogy a mentális illetve agyi állapotok 8. és 9. nyíllal jelölt

[78]Libertarianizmus az a filozófiai irányzat, amely feltételezi, hogy az embernek van szabad akarata abban az értelemben, hogy „akarhatott volna mást, cselekedhetett volna másképpen, mint ahogy tette”.

időfejlődése determinisztikus-e vagy sem, abban a legáltalánosabb értelemben, hogy a t_0 pillanathoz tartozó mentális/agyi állapotról a t_1 pillanatbani mentális/agyi állapotra lépés az objektív modalitás esete-e vagy sem. Nem létezik tehát szabad akarat, ha a világ – mindenekelőtt az agyi/mentális folyamat – determinisztikus, vagyis ha a külvilág adott $\beta_0 \rightarrow \beta_1$ állapotfejlődése mellett, a t_0 pillanatbani ψ_0 illetve ϕ_0 állapotok csak egyetlen ψ_1 és ϕ_1 állapotot engednek meg a későbbi t_1 pillanatban.[79]

Vizsgáljuk most meg ezeknek a ψ és φ állapotoknak a viszonyát. E viszony megítélése alapvetően függ a test–elme kérdésben elfoglalt metafizikai álláspontunktól, vagyis hogy miben is áll az 5. és 12. nyíllal reprezentált kapcsolat, illetve hogy egyáltalán szükséges-e fenntartanunk a mentális állapotoknak és az agyállapotoknak az ábrán jelzett kettősségét. Legyen az a mentalizmus különböző iskoláinak problémája, hogy milyen tapasztalatok alapján és mit állít az autonóm ψ_t állapotok időbeli változásának törvényszerűségeiről, ha vannak egyáltalán szerintük ilyen törvényszerűségek. Mi a továbbiakban élni fogunk azzal a fizikalista feltevéssel, hogy a mentális állapotok lokálisan[80] ráépülnek az agy (fizikai/neurofiziológiai) állapotaira. A fizikalista felfogásból sem következik azonban, hogy nincs szükség erre a ψ – φ kettős nyelvezetre. Például a termodinamikai állapotjelzők értelmes és használható fogalmak maradnak akkor is, ha képesek vagyunk őket a statisztikus fizikában a rendszer mikroszkopikus jellemzőiből származtatni.

A szimultán ψ_t és φ_t állapotok közötti megfelelés nem kölcsönösen egyértelmű. Mert nyilvánvaló, hogy különböző φ_t és φ_t' fizikai/agyi állapotokhoz tartozhat ugyanaz a mentális ψ_t állapot. Ebből következően, mint Grünbaum rámutatott,[81] nem zárható ki, hogy míg az agy állapotfejlődése – például bizonyos kvantumeffektusok miatt – indeterminisztikus, a ráépülő mentális folyamat determinisz-

[79] A szabad akarat ilyen értelmezését Campbellnek (1976) szokás tulajdonítani.

[80] Vö. Chalmers 1996, 33-34. o.

[81] Grünbaum 1972.

tikus. Hasonlóan ahhoz, ahogyan különböző mikroszkopikus állapotokhoz tartozhat ugyanaz a makroszkopikus, termodinamikai állapota a makroszkopikus rendszernek, s a mikroszkopikus állapotok – tegyük fel – indeterminisztikus fejlődése eredményezheti a termodinamikai állapothatározók determinisztikus változását. Biztos azonban, hogy ezt nem fordíthatjuk meg. Ha a neurofiziológiai folyamatok determinisztikusak, akkor a mentális folyamatok is azok.

Éppen ezért rendkívül fontos, hogy a neurális folyamatok indeterminisztikusak-e vagy sem. A libertarianizmus joggal vél megerősítést minden olyan neurális folyamatban, amelyet valószínűségi törvények írnak le. Ám, mint Grünbaum rámutat, a libertariánus szabadságot nem garantálja önmagában az a tény, hogy a döntési folyamatok valószínűségi törvényeknek engedelmeskednek. Tegyük fel – írja – hogy egy populációra érvényesek bizonyos valószínűségi törvények, melyekből az következik, hogy – hosszú távon – a lakosság 80%-a elkövet egy bizonyos bűncselekményt. A közösség egy olyan tagja, aki elkövette a bűncselekményt – a libertariánus álláspont szerint – csak akkor vonható morálisan felelősségre, ha az illető cselekedhetett volna másképpen. A törvény valószínűségi jellege azonban nem jogosít fel bennünket arra, hogy azt mondjuk, az adott személy cselekedhetett volna másképpen. Annyi biztos, hogy a valószínűségi törvény alapján nem tudjuk megmondani, hogy a közösség melyik tagja fogja elkövetni a bűncselekményt. De ez a korlátozás nem jelenti azt, hogy az adott körülmények között, az adott pillanatban, amikor az illető elkövette a cselekményt, akkor cselekedhetett volna másképpen is.[82]

Arthur Fine helyesen világít rá azonban, hogy a Grünbaum-féle argumentum csak akkor áll, ha a szóban forgó valószínűségi modell olyan, hogy elvben létezhet rejtettparaméteres elmélete.[83] Tudjuk, hogy a klasszikus valószínűségi modellek ilyenek, „de mint megta-

[82]Grünbaum 1972.

[83]Fine 1993.

nulhattuk a kvantumelmélet alapjaival kapcsolatos kutatásokból, éppen az ilyen kontrafaktuális distinkcióknak lehetnek váratlan, ugyanakkor tesztelhető következményei" – írja. Majd az antilibertarianizmus és a kvantummechanika összeférhetetlenségével kapcsolatban a következő konklúzióra jut:

> Ha feltesszük, hogy a kvantumelmélet korrekt statisztikus predikciókat tesz, és ésszerű módon tartjuk magunkat a távolhatásnak a lokalitás-elvben megnyilvánuló tilalmához, akkor arra a következtetésre jutunk, hogy a kvantumelmélet statisztikus törvényeinek nem létezik antilibertariánus interpretációja. ... Úgy tűnik tehát, hogy – szemben azzal, amit Grünbaum mond – a libertariánusok „csinálhatta volna másképpen"-je valóban megerősítésre talál az indeterminizmusban, feltéve, hogy az indeterminisztikus törvények olyan típusúak, mint amilyet a kvantumelméletben találunk.[84]

Szabad akarat és a kvantummechanika

Az utóbbi években széles körben elterjedt az az elképzelés, hogy a libertariánus szabadsághoz nélkülözhetetlen indeterminizmus gyökerét az agy működésében fellépő kvantummechanikai jelenségekben kell keresnünk.[85] Túl azon a metafizikai infantilizmuson, hogy „a tudat egy misztérium" és „a kvantummechanika egy misztérium", nosza, kapcsoljuk őket össze, a tudat minden különösebb argumentáció nélküli összekapcsolása a kvantummechanikával hosszú múltra tekint vissza (lásd a Wigner-idézetet a **??.** pontban). Anélkül, hogy állást foglalnánk abban a vitában, vajon a tudat működésének megértéséhez elegendő-e az agy neurális hálójának rendkívül komplex

[84]Uo. 555-556. o.

[85]Penrose 1993, 1994, 1997; Lockwood 1989; Stapp 1993.

dinamikája, vagy pedig más, nem neurális elméletekre van szükség, mint pl. a Penrose–Hameroff-féle mikrotubuláris kvantumjelenségek elmélete,[86] az a gondolat, hogy a kvantummechanika által leírt jelenségek az agy működésének bizonyos részleteiben szerepet játszhatnak, elég plauzibilisnek tűnik.[87] Eddigi vizsgálódásaink alapján viszont határozottan állíthatjuk, hogy a kvantummechanika a tudat „misztériumának" megértésében – legalábbis a szabad akarat problémáját illetően – semmi olyan újat nem szolgáltat, amit a klasszikus elméletektől – elvben – ne kaphatnánk meg. Állításunkat a következő argumentummal támasztjuk alá.

A kvantummechanikának az volna a szerepe az agy működésének leírásában, s mindenekelőtt a szabad akarat melletti érvelésben, hogy olyan irreducibilisen indeterminisztikus jelenségeket produkáljon, melyek nem írhatók le klasszikus eszközökkel, vagyis amelyek olyan speciális statisztikát mutatnak, ami nem enged meg determinisztikus, lokális rejtettparaméteres elméletet. Vagyis az argumentum előfeltételezi, hogy maga a kvantummechanika ilyen, tehát feltételezi, hogy a kvantummechanika szokásos *no go* tételei igazak, abban az értelemben, hogy valóban azt bizonyítják, hogy a kvantummechanika törvényszerűségei nem redukálhatók, nem vezethetők vissza valamilyen klasszikus, determinisztikus rejtettparaméteres elméletre. Mint megmutattuk, a *no go* tételek közül csak kettőről, az EPR-tételről és a GHZ-tételről mondható el, hogy kihívást jelentenek a determinizmus híveinek, és mindkét tétel csak további két feltétellek teljesülése mellett volt bizonyítható. Az is megmutatható, hogy e feltételek egyikének sérülése esetén létezik determinisztikus rejtettparaméteres modellje a szóban forgó kvantummechanikai rendszernek.

Önmagában az a tény, hogy bizonyos kémiai szinapszisokat és

[86]Churchland 1998; Hameroff 1998.

[87]Nem feltétlenül a Penrose–Hameroff-elméletre kell gondolnunk, hiszen az tartalmaz olyan, a hullámfüggvény objektív redukciójára és a kvantumgravitációra vonatkozó hipotéziseket, melyek nem tekinthetők a fizikában széles körben elfogadott elméleteknek.

más neurális membrán aktivitásokat kvantummechanikailag írunk le, még nem jelenti azt, hogy ezeknek a történéseknek ne létezhetne determinisztikus rejtettparaméteres modellje, még akkor sem, ha a kvantummechanikai leírásban a rendszert makroszkopikus koherens állapottal[88] jellemezhetjük. Vagyis mindaddig, amíg a kvantummechanika alkalmazása abban áll, hogy az agy működése során *ténylegesen bekövetkező* események valószínűségeit a kvantummechanikából származtatjuk, nincs okunk feltételezni, hogy a szóban forgó események relatív gyakoriságát ne lehetne determinisztikus rejtettparaméteres elméletből, episztemikus valószínűségként származtatni (vö. **??.** pont). Nem látunk azonban példát arra, hogy az agy működésének (kvantummechanikai) leírása során megvalósulna olyan szcenárió, amelyre akár az EPR-, akár a GHZ-tétel alkalmazható lenne.

Megjegyzendő továbbá, hogy az EPR- és a GHZ-kísérletekre vonatkozó tételekkel kapcsolatban még nem mondtuk ki az utolsó szót (lásd a **??**. fejezetet).

Newcomb-paradoxon

A szabad akarat kérdésében a kvantummechanika nem ad különösebb alátámasztást az indeterminizmus számára. Mielőtt azonban feladnánk az objektív modalitást és vele együtt az akarat szabadságát, ismerkedjünk meg egy paradoxonnal, amely nagyon világosan mutat rá, mennyire mélyen él bennünk az akarat szabadságának élménye, és mennyire nehéz a szabad akarat tagadását összhangba hoznunk más metafizikai meggyőződéseinkkel. A következő paradoxont Nozick publikálta először,[89] s a fizikus William Newcomb nevéhez fűződik:

> Az erdőben járva egyszer csak két dobozt látunk ma-

[88]Fröhlich 1968.

[89]Nozick 1969.

> gunk előtt. Mellettük áll egy kísérletvezető, és a következőket közli: A bal oldali dobozban garantáltan van 1000 $. A jobb oldali doboz vagy üres, vagy 1000000 $-t tartalmaz. Mindez attól függ, hogy egy mindent tudó Lény, aki képes arra, hogy a Te jövőbeli gondolataidat megjósolja, hogyan döntött. Ha tegnap úgy látta, hogy – mohó módon – mindkét dobozt magaddal viszed, akkor a jobb oldali dobozba nem tett semmit. Ha úgy látta, hogy „szerényen” csak a jobb oldali dobozt viszed magaddal, akkor beletett 1000000 $-t. A kísérletvezető felszólít, hogy állj háttal a dobozoknak, és felnyitja a dobozokat. Majd felszólít, hogy válassz, mindkettőt elviszed, vagy csak a jobb oldalit. Mielőtt az utasítását követnéd, még azt is elmeséli, hogy eddig 4010 turistával végezték el ezt a kísérletet, de senkinek sem sikerült még elvinnie 1001000 $-t. A kérdés: Hogyan döntsünk?

A példa paradox jellege abban áll, hogy mindkét lehetséges döntés mellett erős érveket lehet felsorakoztatni. Eléggé kézenfekvőnek tűnik ugyanis az a feltételezés, hogy az említett 4010 esetben a Lény nem véletlenül találta el, hogy mi lesz valakinek a jövőbeli döntése, ennek valószínűsége ugyanis $\frac{1}{2^{4010}}$, vagyis praktikusan nulla.[90] Bele kell tehát törődnünk, hogy a világ, beleértve a mi gondolkodásunkat is, determinisztikus, s hogy a Lény valóban tudhatta, hogy mi fog történni. Ennek megfelelően tehát a helyes döntés, hogy csak a jobb oldali dobozt választjuk.

Másfelől azonban képzeljük el azt a pillanatot, amikor ott állunk háttal a két doboznak. A dobozok tartalmát, ha igaz az egész történet, a Lény már tegnap bekövetkezett cselekvése meghatározta. Itt és most a dobozokkal fizikailag már nem történhet semmi. Ha a jobb oldali dobozban nincs ott az 1000000 $, akkor nincs ott, és semmit

[90] Legalábbis ezt szokás mondani. Ha az olvasó már olvasta a **??**. fejezetet, remélem, egyetért velem, hogy ennek az *a priori* valószínűségi állításnak semmi értelme nincs.

sem vesztünk, ha mindkét dobozt felvesszük. Ha ott van, akkor az a *fizikai realitás*, hogy ott van, a kísérletvezető látja is, hogy ott van, s ez nem változhat meg anélkül, hogy a dobozt valamilyen fizikai hatás ne érné. Minthogy ilyen hatás már nincs, megint csak indokolatlan lenne a bal oldali dobozt otthagynunk. A helyes döntés tehát, hogy mindkét dobozt el kell vinnünk.

Szokás a Newcomb-paradoxont úgy értelmezni, hogy az a determinizmus elleni argumentum, tudniillik hogy a determinizmus ellentmondásosságát jelenti. Mint Ted Honderich helyesen mutat rá, ez nem igaz.[91] Semmiféle logikai ellentmondást nem jelent, ha a világ determinisztikus: A Lény képes az én jövőbeli gondolataimat megjósolni, szabad akarat, a mi általunk használt értelemben nincs, agyunk a determinált módon dönt, s az előre látható döntésnek megfelelően ott lesz a dobozban 1000000 $, vagy nem. Az ellentmondás nem logikai, hanem kizárólag arról van szó, hogy a gondolatkísérletben vázolt szituáció az intuíciónkkal ellentétes, ellentmond a szabad akarat szubjektív élményének. Nyilvánvaló, hogy nehéz elfogadni azt, hogy a kísérletvezető ott áll, a dobozok tartalma alapján már tudja, hogy mit fogok dönteni, én, aki még akkor háttal a dobozoknak töprengek, és végül kibököm a kísérletvezető számára – feltéve, hogy a Lény helyesen jósolt – már előre tudott választ. S mindeközben én ezt úgy élem meg, hogy teljesen szabadon döntök.

A szabad akarat fenomenológiája

A kompatibilizmus[92] persze analitikus képtelenség, ha a szabad akarat általunk adott definíciójához ragaszkodunk. Mindazonáltal a Newcomb-paradoxon példáján láthatjuk mennyire fontos annak megértése is, hogy hogyan értelmezhetjük egy determinisztikus világban

[91] Honderich 1993, 74. o.

[92] Kompatibilizmus az a filozófiai irányzat, mely szerint a szabad akarat létezése összeegyeztethető a determinizmussal.

a szabad akaratra vonatkozó objektív tapasztalatainkat, illetve az akarat szabadságának szubjektív élményét, amely persze, ha bonyolultabban is, de – mint bármely más pszichikai jelenség – elvben tárgyát képezheti az objektív tapasztalásnak.

Az első és legfontosabb kérdés persze az, hogy tisztázzuk, miben is áll ez a szubjektív élmény. Egyes értelmezések szerint a szabad akarat szubjektív érzése nem más, mint annak retrospektív érzése, hogy „gondolhattuk volna másképpen is”. Grünbaum szerint[93] ilyen szubjektív érzés nem létezik. Nem ismeretes, hogy lenne bármiféle pszicho-neurológiai megfelelője egy ilyen érzésnek. Amit a gondolat szabadságának szubjektív élményeként átélünk, az tulajdonképpen a *cselekvés* szabadságának szubjektív élménye. Annak élménye, hogy „cselekedhettünk volna másképpen, ha másképpen akartunk volna cselekedni, azaz, ha másképpen gondoltuk volna”. Ez azonban nem azonos azzal az (állítólagos) élménnyel, hogy „gondolhattuk volna másképpen”. Egyáltalán nem magától értetődő, hogy van-e szabadságunk azt illetően, hogy mikor mit gondolunk. Egy elakadt liftben ránk törő klausztrofóbikus gondolatok helyett szeretnénk mást gondolni, de nem megy! Az agykutatás bizonyos kísérleti eredményei[94] is arra utalnak, hogy egyszerű döntési szituációkban agyunk csak néhány tized másodperc késéssel, utólag „értesít” bennünket döntéseiről.

Egyes értelmezések szerint a szabad akarat szubjektív élményének forrása az a tapasztalat, hogy a jövőre vonatkozó döntéseinket/akaratunkat bármikor visszavonhatjuk. Ha ma úgy gondolom, hogy holnap Miskolcra utazom, akkor bármikor visszaléphetek ettől az elhatározásomtól. A Newcomb-paradoxonban, a döntésemet szabadnak érzem, mert kimondása előtti utolsó pillanatig megváltoztathatom azt. Vegyük észre azonban, hogy ez nem a szabad akarat közvetlen megélése. Hiszen a következőről van szó: a t pillanatban

[93] Grünbaum 1972.

[94] Libet *et al.* 1979.

úgy gondoljuk, hogy a $t+\Delta t$ pillanatban $X^{t+\Delta t}(t)$ gondolatunk, akaratunk lesz. És ezt megváltoztathatjuk, vagyis a $t+\Delta t$ pillanatban akarhatunk mást, mint amit a t pillanatban gondoltunk, hogy akarni fogunk a $t+\Delta t$ pillanatban, azaz annak a t pillanatbani (pontosabban, ha ragaszkodunk a közvetlen tapasztaláshoz, akkor a $t+\Delta t$ pillanatbani) megéléséről van szó, hogy $X^{t+\Delta t}(t) \neq X^{t+\Delta t}(t+\Delta t)$. És ez nem ugyanaz, mint annak az állítólagos élménye, hogy „akarhattuk volna másképpen is”, hiszen az annak a t pillanatbani megélését jelentené, hogy egy korábbi $t-\Delta t$ pillanathoz tartozó $X^{t-\Delta t}(t-\Delta t)$ gondolatunk lehetett volna más, valamilyen $\widetilde{X}^{t-\Delta t}(t-\Delta t) \neq X^{t-\Delta t}(t-\Delta t)$.

A fenti elemzésben ez „a jövőre vonatkozó döntéseinket/akaratunkat bármikor visszavonhatjuk/megváltoztathatjuk” egy erősen libertariánus megfogalmazása annak az egyszerű ténynek, hogy a jövőbeli $X^{t+\Delta t}(t+\Delta t)$ gondolatunk lehet más, mint $X^{t+\Delta t}(t)$, vagyis mint amilyennek a korábbi t pillanatban feltételeztük hogy lesz. Felmerül a kérdés, miért van különbség $X^{t+\Delta t}(t+\Delta t)$ és $X^{t+\Delta t}(t)$ között, ha – mint ahogyan ezt most feltételezzük – a világ determinisztikus, tehát az $X^{t+\Delta t}(t+\Delta t)$ gondolatunk a t pillanatban már teljesen determinált. A válasz nyilván az, hogy nem vagyunk képesek mindig helyesen megjósolni a t pillanatban, hogy mit fogunk gondolni a $t+\Delta t$ pillanatban. Grünbaum tovább megy, és a gondolat szabadságának szubjektív élményét éppen úgy értelmezi, mint annak hiányát, hogy az önmagára reflektáló szubjektum képes lenne saját jövőbeli gondolatait megjósolni. Grünbaum MacKay egyik tanulmányára[95] támaszkodik, aki Poppernek az „önjóslás” lehetetlenségéről szóló levezetésére[96] építve kimutatja, hogy az akaratszabadság szubjektív élménye kompatibilis egy szigorúan mechanisztikus agyműködéssel is. Mint ismeretes, Popper azt mutatta meg, hogy egy Turing-gép nem képes saját maga jövőbeli állapotait kiszámítani. Tehát ha az agy egy Turing-gép determinisztikusságával működik, akkor sem

[95]MacKay 1967.

[96]Popper 1988, 68. o.

vagyunk képesek saját jövőbeli mentális állapotainkat megjósolni, s ezt az objektív tényt – szubjektíve – az akaratunk, illetve gondolkodásunk szabadságaként éljük meg.

A szabad akarat fenomenológiája tehát tökéletesen értelmezhető egy determinisztikus világban. Vegyük azonban észre, hogy mindez elmondható lett volna egy indeterminisztikus világban is. Más szóval, az akarat szabadságának fenomenológiája tökéletesen érzéketlen arra nézve, vajon a világ determinisztikus-e vagy sem.

Bibliográfia

Belnap, N. (1992): Branching space-time, *Synthese* **92**, 385.

Bennett, J. (1988): *Events and their Names*, Hackett Publishing Company, Indianapolis–Cambridge.

Grünbaum, A. (1974): *Philosophical Problems of Space and Time*, Boston Studies in the Philosophy of Science, Vol. XII. (R. S. Cohen and M. W. Wartofsky, eds.) D. Reidel, Dordrecht.

Hawking, S. W. és Ellis, G. F. R. (1973): *The Large Scale Structure of Space-Time*, Cambridge Univrsity Press, Cambridge.

Maxwell, N. (1985): Are probabilism and special relativity incompatible?, *Philosophy of Science* **52**, 23.

McTaggart, J. M. E. (1908): The Unreality of Time, *Mind* **17**, 457. (Magyarul: http://www.szv.hu/cikkek/miert-nem-valosagos-az-ido)

McTaggart, J. M. E. (1993): The Unreality of Time, in: *The Philosophy of Time* (Oxford Readings in Philosophy), R. Le Poidevin, M. MacBeath (eds.), Oxford University Press, Oxford. (Eredeti mű: *The Nature of Existence*, 33. fejezet, Cambridge University Press, Cambridge 1927.) (Magyarul: http://www.szv.hu/cikkek/miert-nem-valosagos-az-ido)

Mellor, D. H. (1981): *Real Time*, Cambridge University Press, Cambridge & New York.

Mellor, D. H. (1995): *The Facts of causation*, Routledge, London

Mellor, D. H. (1998): *Real Time II.*, Routledge, London.

Parfit, D. (1987): *Reasons and Persons*, Oxford University Press, Oxford.

Prior, A. N. (1993): Change in Events and Change in Things, in: *The Philosophy of Time* (Oxford Readings in Philosophy), R. Le Poidevin, M. MacBeath (eds.), Oxford University Press, Oxford. (Eredeti mű, in: *Papers on Time and Tense*, Clarendon Press, Oxford.)

Putnam, H. (1967): Time and physical geometry, *The Journal of Philosophy* **64**, 240.

Quine, W. O. (1999): Természeti fajták, in: *Tudományfilozófia – szöveggyűjtemény*, Forrai G. és Szegedi P. (szerk.), Áron Kiadó, Budapest. (http://nyitottegyetem.phil-inst.hu/tudfil/ktar/forr_ed/forr_ed.htm)

Reichenbach, H. (1956): *The Direction of Time*, University of California Press, Berkeley.

Rietdijk, C. W. (1966): A rigorous proof of determinism derived from the special theory of relativity, *Philosophy of Science* **33**, 341.

Rietdijk, C. W. (1976): Special relativity and determinism, *Philosophy of Science* **43**, 598.

Russell, B. (1976): *Miszticizmus és logika és egyéb tanulmányok*, Magyar Helikon, Budapest.

Russell, B. (1996): *A filozófia alapproblémái*, Kossuth Könyvkiadó, Budapest.

Salmon, W. C. (1977): The Philosophical Significance of the One-Way Speed of Light, *Noûs* **11**, 253.

Swinburne, R. (1968): *Space and Time*, Macmillan, London.

Swinburne, R. (1990): Argument from the fine-tuning of the universe, in: *Physical cosmology and philosophy*, J. Leslie (Ed.), Collier Macmillan, New York.

Swinburne, R. (1998): *Van Isten?*, Kossuth Kiadó, Budapest.

Wald, R. M. (1984): *General Relativity*, University of Chicago Press, Chicago and London.

Wang, H. (1995): Time in philosophy and physics: from Kant and Einstein to Gödel, *Synthese* **102**, 215.

Yang, C. N. és Mills, R. L. (1954): Conservation of Isotopic Spin and Isotopic Gauge Invariance, *Phys. Rev.* **96**, 191.