

Ambrus-Lakatos Loránd

MORALITÁS ÉS RACIONALITÁS A DÖNTÉSELMÉLETBEN*

(Huoranszki Ferenc: „Racionalitás és erkölcsi normák”)[§]

1. Ha a racionális döntések elmélete igényt tart arra, hogy az emberi cselekvés magyarázatában érdekelt doktrínák között privilegizált vagy éppenséggel elsődleges helyet foglaljon el, akkor számot kell adjon a moralitáshoz való viszonyáról is. Ehhez Huoranszki Ferenc szerint az elméletet módosítani kell, mégpedig úgy, hogy a speciálisan morális motivációkat számba vehesse a cselekedetek magyarázatában, illetve, hogy helyüket a racionális megfontolások között kijelölhesse. Mindez azért fontos és sürgős, mert nincs alkalmasabb elméletünk ezen megfontolások stuktúrájának vizsgálatára (7).

A javasolt módosítás kifejtéséhez elengedhetetlen a bevett elmélet felvázolása (14–18). Így lokalizálhatók azon kitételei, amelyek a moralitásra vonatkozó elfogadhatatlan korolláriumokért elsősorban felelősek. Ezek között kulcsfontosságú az az álláspont, mely szerint, ha a rendelkezésünkre álló döntések között van olyan, amelyik minden lehetséges esetben („világállapotban”) többet hoz a konyhára, mint bármely másik, akkor ezt kell választanunk. Ez a dominancia elve. – Gondoljunk a híres fogolydilemmára, amely éppen egy olyan szituációt mutat be, ahol a szokványos döntésemélet morálisan elfogadhatatlan választ ad a „mit tegyünk ilyenkor?” kérdésre – mégpedig pontosan azért, mert a nem morális akció (a fenti technikai értelemben) dominálja a morálisat.

A sztenderd döntésemélet szerint cselekedeteink következményei rangsorba állíthatók, és többnyire megengedi azt is, hogy a rangsort kardinális skálán is kifejezzük. Huoranszki kulcsfontosságú érve szerint azonban semmi sem akadályozhat meg bennünket abban, hogy a következményeken kívül a cselekedetekhez magukhoz is ilyen értékszámokat rendeljünk. Hanem a racionális döntések így módosított elméletében előfordulhat, hogy – *ceteris paribus* – az optimális választás (a dilemma opciói közül) a morálisan megfelelő akcióra esik – mert a nem morális akció dominanciája feloldódik. Hogy ez így lesz-e, azon múlik, hogy a morális tett önértéke felülmúlja-e következményeinek értékét. Ez esetről esetre és személyről személyre változhat (28). Mindenesetre ez a módosítás helyet talál a morális motivációknak a formális döntésemélet keretében.

* A kommentár első változata elhangzott 1998. november 28-án, a „Cselekvésemélet és társadalomkutatás” című, Csontos László emlékének szentelt konferencián.

§ A zárójelben lévő számok a „Racionalitás és erkölcsi normák” (Szociológiai Szemle 1999/1.) oldalszámaira vonatkoznak.

2. Ez a gondolatmenet azonban aligha állja meg a helyét. Először is, láttuk, hogy a szerző szerint a deliberáció logikai-matematikai struktúrájának legmegfelelőbb reprezentációja a döntéelmélet, és ez az elmélet kész arra, hogy az értékeket számokkal fejezze ki. Tehát, ha az egyes cselekedeteknek is önálló értéket tulajdonítunk – cselekedetek és következményeik egy bizonyos kitüntetett értelemben egy szintre kerülnek.

Azonban, ha a cselekedetek értékeit hozzáadjuk a kifizető (következmény) mátrix megfelelő elemeihez, akkor az eredetihez képest új játékot definiálunk. A lehetséges stratégiák és kifizetések konstellációja már nem ugyanaz, mint az eredeti problémában. Így hát ebben a módosított játékban, akár az is lehet a végkifejlet, hogy a felek az erkölcs által preferált lépést teszik meg; de mindez nem befolyásolja az *eredeti* fogolydilemma megoldását, sem a kritizált dominanciaelvet.

A problémák gyökerét abban látom, hogy racionális választások elmélete nem lép fel avval az igénnyel, hogy a »természeti állapotok–következmények–cselekedetek–preferenciák« négyesének elemeihez szubsztantív értelmezést is csatoljon. Ha egy vizsgálandó szituációban a lehetséges cselekedeteket azonosítottuk, akkor az összes értékelést lehetséges kimenetelekhez tartozó értékelésekbe sűrítethetjük. Vagy másképpen fogalmazva: az alapmodell formális értelemben véve invariáns arra nézve, hogy pontosan mit is értékel az elemzendő döntéshozó.¹ És ez éppen a logikai-matematikai „megjelenítés” követelményéből következik.

Így hát nehezen védhető az az álláspont, hogy a racionális választások sztenderd elméletének a dolgozatban javasolt módosítása jelentősen hozzájárulhat ahhoz, hogy az elméletet alkalmassá tegye a morális motivációk kezelésére – legfeljebb annyit mondhatunk, hogy a javaslat rámutat arra, hogy a kérdéses elmélet képtelen rá, hogy a cselekedetek értékelését megfelelően elválassa a következmények értékelésétől. Olyannyira, hogy kérdéses az is, hogy az elmélet igazából képes-e célok és eszközök megkülönböztetésére: jól ismert, hogy a formális döntéelmélet megannyi bevett és kurrens változatában nem is különül el igazán cselekedet és következmény.² A morális normakövetés és a racionális döntések elméletének viszonya a javasolt módosítás után is problema-

¹ Az érvelés kifejthető a dinamikus (intertemporális) döntések elméletére való hivatkozással is. A dinamikus döntések extenzív formában való megjelentetésénél az összes kifizetés a döntési fa végpontjaiban történik. Ez a modell, megint csak, invariáns arra, hogy a döntéshozó az egyedi cselekedeteket, vagy csak az adott végpontba kerülést értékeli. De mivel az esszé a fogolydilemma egyszer-játszott (szimultán) változatát tárgyalja, a döntési fára való ezen utalás csak illusztrációként szolgálhat.

² Ezekben a változatokban a döntéshozó csupán arra koncentrál, hogy a következményeket megfelelően (preferenciái szerint) rendezze. Ha bizonytalanság melletti döntésről van szó, ugyanő (megfelelően értelmezett) szerencsejátékokat rendez sorba, majd a sor tetején állót választja – itt a cselekedetek és a végeredmény között csak a véletlen közvetít. (Ebben a kontextusban persze a „véletlen” értelmezése is speciális.) Ez olyannyira igaz, hogy Martin Osborne és Ariel Rubinstein 1994-es tankönyvében (*A Course in Game Theory*, MIT Press) külön érvelni kényszerül amellett, hogy a társadalomtudományoknak a döntés- és játékelmélet olyan modelljére van újból szüksége, amelyik a cselekedeteket alapvető elemként kezeli. Egy másik vonatkozó példa: John Harsanyi szerint a racionális döntések cél-eszköz modellje általában elvetendő, és a preferencia-modellt kell követni („Does Reason Tell Us What Moral Code to Follow and, Indeed, to Follow Any Moral Code at All?”, *Ethics* 96.1 [1985]: 42–43.).

tikus marad: az, hogy hová számoljuk el a cselekedeteknek tulajdonított értékeket nem dönthető el az elmélet berkein belül.

Vessük össze a dolgozat fő érvét egy egyszerűbb és jól ismert megoldási javaslattal. E szerint: mondjuk azt, hogy vannak olyan egyének, döntéshozók, akik altruista preferenciákkal vannak ellátva egy fogolydilemma-szituációban, és hogy ezek a preferenciák számszerűsíthetők. Ekkor, mivel az altruista cselekedetekhez rendelt értékeket hozzáadhatjuk a következmények (az alaptörténetben eltöltendő „börtönhónapok”) értékeléseihez, kijöhet az is, hogy az egyensúlyban mindkét fél kooperálni fog. De megint csak: ez az állítás már egy, az eredeti fogolydilemmához képest módosított, új játékra lesz csak érvényes.³

3. Figyelemre méltó és érdekes, hogy a szerző a racionális választások sztenderd elméletének felvázolásakor Savage alapvető, de ma már ritkán forgatott művében kifejtett modellből indul ki. Ebben két alapfeltevést tart különösen kifogásolhatónak. Az egyik az, hogy a lehetséges cselekedetek és a releváns „természeti állapotok” bekövetkezésének valószínűségei egymástól függetlenként vannak meghatározva. Hanem ez a feltételezés korántsem tartható fenn mindig – cselekedeteink nagyon gyakran befolyásolják a kérdéses valószínűségeket. Ez az észrevétel vezet át a Richard Jeffrey döntésméletét bemutató részhez (20–26), ez utóbbi elmélet képes a fenti probléma hiteles kezelésére.

A másik korrigálandó feltételezést már ismertettük. Ez az, hogy a cselekedeteket magukat nem értékeljük, csak azok következményeit. Azonban, amikor a morális motivációk döntésméleti tárgyalására tér, a szerző lényegében csak a második módosítást használja fel fő érvének kifejtése során.⁴ Vegyük figyelembe azt is, hogy nem egyértelmű, hogy Jeffrey maga úgy gondolta-e, hogy elméletével megoldotta a fogoly dilemmájában rejlő paradoxont.⁵ De a Huoranszki által elővezetett megoldáshoz nem is kell Jeffrey elmélete.⁶

4. A dolgozattal kapcsolatos legkomolyabb probléma azonban nem technikai jellegű megfontolásokra vezethető vissza. Huoranszki kiemelt helyen hangsúlyoz-

³ Még két technikai megjegyzést szeretnék tenni. (i) A dohányzásról való leszokás problémájának modelljében (20–23) nem feltétlenül kell olyan, a kimenetekhez kapcsolódó, preferenciákat elképzelnünk, amelyek szerint a dohányzás folytatása domináns stratégia lesz. Akkor viszont az esszében adott példa nem ad elég alapot a savage-i elmélet *across-the-board* elvetéséhez. Csupán annyit mondhatunk, hogy egy bizonyos preferenciakonstelláció problematikus következtetésekre vezet. De akkor annyi azért még felvethető, hogy akinek a kérdéses preferenciái olyanok, amelyek a dohányzás folytatását domináns stratégiává teszik, az egyrészt magára vessen, másrészt még mindig racionális a savage-i értelemben. (ii) A fogolydilemma kerettörténetének módosításaira való hivatkozás nem hatásos: az esszében szereplő mindkét javaslat megint csak egy-egy új játékot definiál (25–26). A megismételt fogolydilemma megoldása közismerten más lehet, mint az egyszer-játszotté.

⁴ Ennek, úgy látom, az a felismerés az egyik oka, hogy a cselekedetek és a kimenetek függetlensége egy játékelméleti situációban nem is jöhet szóba. Így hát ebben a kontextusban a fenti első feltevés kritikája nem sok munkát végezne.

⁵ Vö. Richard Jeffrey ([1965] 1983): *The Logic of Decision*, University of Chicago Press, 16–20.

⁶ Ahhoz viszont, hogy a savage-i elmélet itt elsőként említett bajos háttérfeltételét (a cselekedetek és a kimenetek valószínűségeinek függetlenségét) feloldjuk, hasznos lehet Jeffrey elképzeléseinek igénybevétele.

za, hogy a moralitás mely aspektusát kell a racionális döntések elméletének kezelnie: nem a morális normák igazolásáról, hanem motivációs erejükről van szó (4–7). Azonban úgy látom, hogy a javasolt fogalmi módosítás a szerző szándékai ellenére sem kezeli megfelelően az erkölcsi motivációk természetével kapcsolatos nehézségeket. Elsőrendű jelentőségű ugyanis az, hogy mit ért a szerző „morális cselekedetek” alatt: csakis olyan cselekedeteket tekintünk morális indíttatásúnak, amely a cselekvést magát is értékeli (26). Kiemelendő még az az állítás, hogy eddig a formális döntésemélet hívei nem tudtak mit kezdeni a morális normák szerepével; a morális döntéseket vagy az önérdék vezérelte cselekedetek egy válfajának állították be (9–12), vagy az embereknek két külön, egy egoista és egy altruista, motivációs bázist tulajdonítottak (12–14). A dolgot ezen fő premisszáinak tükrében nehéz elszámolni avval az eredménnyel, miszerint az, hogy a fogoly-dilemma-szerű szituációkban az erkölcs által sugallt akciót hajtuk-e végre, az erkölcsi normák *relatív* súlyától függ (28).

Ezek a normák és értékük egyénről egyénre változnak (uo.). – Tegyük fel, hogy valaki a következő dilemmával áll szemben: elérhetővé válik számára egy olyan kimenetel, amelyet nagyon nagyra értékel. Azonban ahhoz, hogy ez a kimenetel ki is teljesedjék, egy sor immorális lépést kell megtennie: mondjuk azt, hogy az áhított állapot útjában vastag hazugságok állnak. Na már most a szerző által módosított döntésemélet azt javasolja, hogy vessük össze az áldásos végkifejlet értéket az okos húzások morális költségével: ha az előbbi felülmúlja az utóbbit, megindulhat a célirányos prudens gyakorlat. Azonban lehet, hogy tájékozott és morálisan nem retardált külső megfigyelők szerint ekkor a morális normák radikális áthágásáról van szó. „Fejüket csóválják, hogy bizony ez csúnya látvány”. – De a jelen elmélet szerint a cselekvő hitelesen védekezhet avval, hogy az ő saját értékelése szerint az elérendő prudenciális haszon felülmúlja az erkölcsi költségeket.⁷

5. De mi is a tétje annak, hogy maga a módosítás működik-e vagy sem? – Sokan azok közül, akik eddig az erkölcs szerepét a döntésemélet keretein belül óhajtották tisztázni, felvetették álláspontjuk filozófiai előzményeinek kérdését is. Harsanyi – például – szükségét látja hangsúlyozni, hogy munkája a gyakorlati filozófia kanti hagyományát folytatja; ő ráadásul Hume örökösének is tartja magát. Ez pont idevág: az esszé szerint a döntésemélet Hume moralitásról alkotott nézetein és Kant metafizikájának szekularizálásán alapuló konszenzus, akkor is, ha a módszertani individualizmus társadalomtudósai ezt általában nem is ismerik fel (18–20). Ezt a véleményt avval a már felemlített kulcspremisszával együtt kell értékelnünk, hogy a racionális választások elmélete a döntést megelőző megfontolások logikai-matematikai modelljét adja, ezen megfontolások legjobb elméletének szá-

⁷ Harsanyi szintén hangsúlyozza a morális követelmények viszonylagosságát, hiszen ezek esetleg túlságosan is költséges akciókat írhatnak elő. Úgy tartja, hogy a „morális” levezethető a „racionális”-ból, igazából az előbbi az utóbbi egy aloslata. Nem azért, mert a morális motivációk önérdékkövetést rejtjenének, hanem pontosan fordítva, az erkölcsőt figyelembe vevő cselekedet racionális igazolásakor mintegy feladjuk az egyéni preferenciáinkat. Így ő is számot ad a morális normák tartalmáról a racionális választások elméletében, és Harsanyi és Huoranszki megoldásának morális receptje hasonló. Azonban Harsanyinak eszébe se jutna a morális cselekedetek önmagukban való értékelése (lásd „Does Reason Tell Us”, 52.).

mít, és ezért az erkölcsi normák kontextusában sem szabad elvetni. Szerintem éppen ezen utóbbi meggyőződés vállalását kellene újra jól megfontolni.

Láttuk, hogy Huoranszki módosítása, tudatosan, az eredeti elmélet keretein belül kíván maradni; bár fentebb amellet érveltem, hogy az eszközlendő változtatás hatálya kétséges. Ha ez így van, akkor miért nem kerülheti el azt sem, hogy a döntésmélet és moralitás viszonyát a konszenzus terminusai szerint jelölje ki, hiszen erre csak két mód adódik (hívjuk ezt „kényszerű dichotómiának”⁸): a moralitás vagy az önérdékkövetésből vezethető le (9–12), vagy a racionalitás határán kívül eső motivációra vonatkozik (12–14)? – *Be this as it may*⁹, az az (itt motiválatlan) alapállítás – nevezzük a dolgot az alapján szókratészinek –, hogy a morális cselekvés szükséges feltétele, hogy annak a cselekvő önértéket tulajdonítson (26) manifeszt ellentmond annak, hogy a jelen módosítás mégiscsak a fenti két lehetséges megoldás egyikét fogalmazza újra (20).¹⁰

Melyik premissza erősebb: a szókratészi, vagy az, hogy a racionális döntések elmélete a cselekvést megelőző megfontolások (deliberáció) legjobb formális megjelenítése?

Na már most, ha a szókratészi alaptételt vesszük elsődlegesnek, akkor talán úgy áll a dolog, hogy a módosításban foglalt nézet mégiscsak kikerül a konszenzus hatálya alól, ekkor azonban az egyéni morális költségekre és a formális modell prioritására vonatkozó eszmefuttatás szofisztikusnak tűnik – ezek igazából csak a tisztán prudenciális cselekedetekre érvényesek.

De, másfelől, kiindulhatunk abból a már ismertetett érvelésből is, miszerint a döntésmélet nem látja egy cselekedethez és annak következményeihez rendelhető értékek között a különbséget. Ha emellé (második premisszaként) tesszük azt az álláspontot, hogy a döntésmélet a deliberációról adható legjobb számadás, akkor az egyetlen következtetés az lesz, hogy a szókratészi premissza nem lehet elsődleges. Sőt: el kell vetni.

Bármit teszel, okosan tedd?

⁸ Meglehet, hogy úgy léphetünk csak tovább, hogy tüzetesebb vizsgálat alá vesszük azt az érvelést, amely szerint a döntésmélet valójában csak kétféleképpen viszonyulhat az erkölcs-höz. Észre kell vennünk, hogy a „racionalitás és moralitás kérdésében elfoglalt logikailag lehetséges álláspontok” (8) talán még sincsenek kimerítően ábrázolva. A következő lehetséges nézet nem is szerepel: az önérdékkövető racionális döntések „tétélezése” másra jó, mint a morális cselekedetek tárgyalása. A formális döntésmélet számot tud arról adni, hogy bizonyos helyzetekben és szcenáriókban mit várhatunk az emberektől. De ezen helyzetek tanulmányozásakor főleg a következmények (főleg a társadalmi, politikai, gazdasági stb. következmények) érdekelnek bennünket. A moralitás azonban „másról szól”, azaz másról is szól; és persze nem feltétlenül az altruista cselekedetekről.

És: jócskán akad még olyan figyelemre méltó nézet a problémával kapcsolatban, amelyik nem esik a dichotómia hatálya alá. Lásd például John Griffin (1986): *Well-being*, Oxford University Press; John Harsanyi ([1977] 1982): „Morality and the Theory of Rational Behavior”, in.: A. Sen–B. Williams (szerk.) *Utilitarianism and Beyond*, Cambridge University Press; Derek Parfit (1984): *Reasons and Persons*, Oxford University Press.

⁹ További tisztázásra váró kérdés a „kényszerű dichotómia” kapcsolata a hume-i/posztkantian konszenzussal.

¹⁰ Vö. avval a megfigyeléssel, hogy ahogyan a szerző a „kényszerű dichotómiát” és a hume-i/kantian konszenzust tárgyalja, nyilvánvaló kritikái álláspontot tükröz.

6. Mindazonáltal, mélységesen egyetértek a dolgozat szándékaival és alapállásával. Valóban, racionalitás és moralitás viszonya elsőrendű kérdés, bár ezt a kérdést a racionális döntések sztenderd elmélete nem tudja megfelelően kezelni. És nem lehet elfogadni a fenti viszony deproblematizálását a moralitás önérdekkövetésből való levezetése révén, és szintúgy nem adhatjuk be a derekunkat a kettős motiváció elméletének sem. Azután: racionalitás és moralitás viszonyát a morális követelmények szubsztantív igazolása nélkül is lehet tárgyalni (és nagyon is rendjénvaló így tárgyalni), *érvelésünk során* minimális koncessziót adva csak az erkölcsi kérdés-feltevés jogosságának.

Valamint: elsőrendű jelentőséget kell annak a meggyőződésnek tulajdonítani, hogy a morális cselekedeteket önmagukban is értékeseknek kell tartanunk. Ennek a követelménynek az *igazolásához* a racionális döntések elmélete hozzá sem tud szólni. Végezetül, rendkívül fontos annak a hatásnak a megvilágítása, amelyet Hume és Kant erkölcsfilozófiája a döntéselmélet világnézetére (és így kedvenc modelljeinek kialakulására) gyakorolt.

És az is igaz, hogy a racionális döntések elmélete minden cselekvés mozgatórugójáról számot óhajt adni, köztük az általában morálisnak tartottakról is. De ezt az igényt nem kell olyan nagyon komolyan venni (másfelől viszont Hume és Kant erkölcsfilozófiáját nagyon komolyan kell venni).

Ezért az erkölcsi normák tárgyalásánál nem kell feltétlenül elfogadni a racionális döntések elméletével szövetséges „logikai-matematikai” modell keretét sem, doktrinális prioritása tételezése még káros is lehet. Tehát az esszé által követett út, az alapmodell kiterjesztése vagy módosítása: bár fegyelmezett, de azért mégiscsak félrevezető módszert jelöl ki.