

Lehmann Miklós

A racionalitás innátista felfogása

A filozófiatörténet jelentős vonulata az embernek a racionális gondolkodásra való képességét e képesség veleszületettségével magyarázza. Szemben a tanulási folyamat során megszereshető készségekkel és képességekkel, ez a felfogás a racionalitást egy olyan mély és eredeti tulajdonságnak tartja, amely más képességek elsajátításának is alapjául szolgálhat. Ennek megfelelően, mint veleszületett képesség szorosan hozzátartozik az ember természetéhez; az emberi mivolt egyik ismérve, hogy bár gondolkodását tekintve az emberiség nem mondható egységesnek, a gondolkodás racionális aspektusa azonban közös, azaz minden emberben azonos formában fellelhető. Mindemellett – és talán éppen ezért – a racionalitás fogalma egyaránt lényeges szerepet tölt be a megismerés, a tudomány és a társas együttélés, az emberek közötti viszonyok és az etika terén is; egyes korokban olyan tekintélyre tett szert, amely önmagában elegendő kritériummá avatta a viták – és vitás élethelyzetek – megoldásában. Mint közös, veleszületett emberi képesség, így egyben a viták általános (mindenkiben fellelhető és mindenki számára kötelező érvényű) alapját és döntési mércéjét nyújtja. A veleszületettség mint magyarázat tehát közvetlenül az ember természetében mutat ki olyan tulajdonságokat, amelyek tudományos értelemben is képesek az emberek közös gondolkodási formáinak magyarázatára. A következőkben megkísérlem áttekinteni e felfogás gyökereit és alapvető forrásait, majd ezek hatását a kognitív tudomány modern nézeteire; végül a bemutatott érvek fényében a veleszületettség értékelését nyújtom.

I.

A görög filozófia klasszikus korában Szókratész, Platón és Arisztotelész megegyeznek abban, hogy az embert veleszületetten racionálisan gondolkodó lénynek tartják; a racionalizmus gondolata tehát eleve összekapcsolódik a veleszületettséggel, mivel a kor filozófiai szemlélete számára kézenfekvőnek tűnt, hogy az ésszerű gondolkodás eredete az emberi természetben keresendő. Ez a racionalitás-tulajdonítás egyben definíciószerűen különbözteti meg az emberi lényeket az állatoktól, valamint a természet egyéb alkotórészeitől. Az ésszerűen működő értelem egyben közös emberi tulajdonságnak is számít, amely közel azonos mértékben van elosztva az egyes emberekben.¹

Az ismert példa Platón *Menón* című dialógusában található. Az anamnézisz-elmélet bizonyítékaként Szókratész itt bemutatja, hogy egy tanulatlan rabszolga is képes egyetemes geometriai összefüggések felismerésére. A feladat során a rabszolga ráébred egy négyzet oldalhosszúságának és területének növelése közötti összefüggésre. A megoldásban lassan halad előre: eleinte a négyzet területének megkétszerezé-

¹ E definíción belül ugyanakkor találhatók eltérések is, például abból a szempontból, hogy vajon minden ember (szabad polgár, rabszolga, férfi, nő vagy barbár) azonos mértékben tekinthető-e racionálisnak.

se alapján úgy véli, ehhez oldalának hosszát is kétszerezni kell; majd a rajzolt ábra alapján belátja, hogy ebben az esetben a négyzet területe négyszeresére növekedne. Ekkor megpróbálkozik egy köztes megoldással, az oldal másfélszeres növelésével. Minthogy ekkor sem jár sikerrel, teljesen elbizonytalanodik, tanácsstalanná válik (Szókratész mindezt úgy magyarázza, hogy bár eleinte a rabszolga azt hitte, tudja a megoldást, most ráébredt, hogy nem tudja – ez pedig megadja a lehetőséget arra is, hogy ráébredjen az igazi megoldásra, mivel azt korábbi hite elhomályosította). Ekkor Szókratész rávezeti, hogy figyeljen fel az átlók és az oldalak összefüggésére: a négyzet átlójának hossza egyenlő lesz egy kétszer akkora területű négyzet oldalának hosszával. A filozófus ezt úgy szemlélteti, hogy az átlójával két egyenlő háromszögre osztott négyzetet ezen háromszögek segítségével kétszerezi meg. A rabszolga végül nem számszerűen ismeri fel a kétszer akkora területű négyzet oldalának hosszát, hanem ráébred a terület és az oldalhossz, valamint az átló közötti arányokra – vagy másképpen, a négyzetgyök fogalmára.

Miután a rabszolga képes mindezt határozottan kimondani, Szókratész ebből arra következtet, hogy valójában mindig is tudta, csupán elő kellett hívni belőle:

„Szókratész: Tehát a nem tudóban benne vannak a helyes vélemények azokról a dolgokról, amikről nem tud?

Menón: Úgy látszik.” (PLATÓN 1984. I. köt., 688.)

A történettel Platón az anamnézisz-elméletet kívánja demonstrálni, ám általános értelemben mindez a közös emberi tudásról szól: arról, amely alapját képezi minden további ismeretnek. Szókratész ráadásul a rávezetés során feltételezi, hogy azok a logikai lépések, amelyeken keresztül a demonstráció során halad, minden ember számára világos és magától értetődő gondolatmenetet képeznek: a rabszolgában fel sem mérül, hogy megcáfolja Szókratész okfejtésének egyes lépéseit. A racionális gondolkodás általános és közös alapjai eszerint megkérdőjelezhetetlennek tűnnek.

A történet ugyanakkor arra is rávilágít, hogy az emberi megismerőképesség két alapvető, minden emberben azonosképp fellelhető és veleszületett forrással rendelkezik: egyrészt bizonyos alaptételekkel, másrészt a gondolkodás szabályaival. Ezt a kettősséget a görög *logosz* és *nousz* fogalma is tükrözi. A logosz azt a képességet jelöli, ami által az ember képes állítások megformálására és azok ellenőrzésére (ennek paradigmátikus arisztotelészi példája a szillogizmus), míg a nousz a kiinduló premisszákról gondoskodik, tekintettel arra, hogy igaz következtetésekhez igaz premisszákra van szükség. Önmagában a logosz elégtelennek bizonyul, mivel az első premisszák nem igazolhatók szillogizmussal; az pedig, ha ezeket a premisszákat még inkább alapvető, korábbi premisszákra kívánnánk visszavezetni, végtelen regresszushoz vezetne. Platón ezeket az alapelveket az intuíció és az anamnézisz-teória segítségével magyarázza, Arisztotelész pedig az indukció bevezetésével.²

² Logosz és nousz, értelem és ész viszonya meglehetősen összetett problémát jelent a filozófiatörténet során, e helyen nem részletezem a kérdéskört. Hierarchikus egymás fölé rendelésük, illetve egymásra utaltságuk azonban minden ismeretelméleti teóriában felmerül. Vázlatos megoldást jelenthet Wolfgang Welsch elgondolása, aki a kettő különbségét eltérő irányultságukban próbálja megragadni: a logosz tárgyra irányuló reflexió, amely a megismerő szubjektumot háttérbe szorítja; a nousz pedig önmagára irányuló reflexió, amely kellő alapot biztosít a logosz számára (WELSCH 1995).

II.

A következő évszázadokon keresztül a platóni–arisztotelészi hagyományokat folytató európai filozófia e tekintetben is követi a görög elődöket. A felfogás jól összeegyeztethetőnek tűnt a keresztény filozófia vallási elveivel, ahol is az ember kitüntetettsége, test és szellem viszonya, valamint az ember isteni adományként kapott értelme fontos szerephez jutott. Mi több, az újkor hajnalán e szemlélet még inkább kiszélesedett: az értelem erejébe és az ész mindenhatóságába vetett hit a vallási háttér nélkül is életképes maradt, így az értelem végső soron a megismerés alapvető kritériumaként jelenhetett meg.

Descartes mintha a korábbi logosz fogalmát visszhangozná az *Értekezés a módszerről* egyik ismert passzusában, mellyel értekezését indítja:

„A józan ész az a dolog, amely a legjobban oszlik meg az emberek között, mert mindenki azt hiszi, hogy annyit kapott belőle, hogy még azok sem szoktak maguknak többet kívánni, mint amennyijük van, akiket minden más dologban csak igen nehéz kielégíteni. Az nem valószínű, hogy ebben mindenki téved; ez inkább azt bizonyítja, hogy az a képesség, amelynél fogva helyesen ítélünk és az igazat megkülönböztetjük a hamistól – s tulajdonképpen ez az, amit józan értelemnek vagy észnek nevezünk –, természettől fogva egyenlő minden emberben; úgyhogy véleményeink nem azért különböznek, mert némelyek eszesebbek másoknál, hanem azért, mert gondolatainkat különböző utakon vezetjük, s nem ugyanazokat a dolgokat nézzük.” (DESCARTES 1992, 15.)³

Descartes ezt az értelmet kizárólagosan fenntartja az emberek számára (így ez egyben az ember és a többi élőlény szigorú elválasztását is szolgálja), a gondolkodást az isteni eredetű lélek lényegének tartva. Ez az értelmes lélek egyben kiindulópontot képez a megismerés számára: a világ alapvetően a racionális gondolkodás által ismerhető meg. Hangsúlyt kap tehát, hogy a valóság már a gondolkodás tiszta elveiből kifürkészhető; sőt, a descartes-i racionalitás előfeltételezi, hogy maga a valóság is racionális elvek alapján szerveződik – így a valóság megismerhetőségét az ember és a világ racionalitása közötti korrespondencia szavatolja.⁴

Hasonló gondolatmenetet folytat később Kant is, ha lehet, még radikálisabban. A *tiszta ész kritikájában* a racionalitást a megismerés konstitutív alapelveinek forrásaként írja le és a kategóriákon keresztül szemlélteti (KANT 1995). Jóllehet a realitás és a megismerés közötti korrespondencia korántsem egyértelmű, a racionalitás jelenti a kiindulópontot; egyben pedig korlátozza is az értelem használatát – hiszen ezen korlátok nélkül az értelem antinómiákba bonyolódik. Ilyen keretek közt az értelemre egy „le-fokozott” szerep hárul, melyben legfeljebb regulatív elvekkel szolgálhat.

³ Maga a descartes-i módszer a közös, általános értelem számára szolgál útmutatóként, amely útmutatót az értelem azonosságából adódóan bárki képes követni, illetve annak igazságáról meggyőződni.

⁴ Descartes maga a gondolkodás egyik első igazságának tartja, hogy létezik egy, az embernél (gondolkodásában is) tökéletesebb lény, egy tisztán gondolkodó szubsztancia, a világ teremtője. A teremtés egységén keresztül így a világ ésszerű strukturáltsága érintkezik az ember értelmével (vö. DESCARTES 1992, 46. sk.).

III.

A fent vázolt filozófiai alapok nyújtottak hallgatólagos előfeltevéseket a mesterséges intelligenciával foglalkozó kutatások megkezdésekor is. A racionalitás szempontjából ezek a következő tételekben jelentkezhetnek:

- Ha a racionalitás paradigmátikus példája a szillogizmus, akkor ahhoz, hogy értelmes rendszerek jöjjenek létre, következtetőgépeket kell megalkotni;
- ha a következtetések automatizálhatók és algoritmizálhatók (márpedig a szillogizmusok szisztémája éppen ezt sugallja), akkor ezek a gépek hasonló – vagy még inkább: azonos – módon működnek, mint az emberi gondolkodás; vagy másképpen: mindannyiunk agyában ott csikorog egy gép;
- ha sikerül racionális – avagy következtető- – gépeket építeni, azok szükségképpen értelmesek lesznek.⁵

Miközben azonban evidenciának veszik a gondolkodás racionális jellegét, amely az ember lényegi tulajdonságát képezi, figyelmen kívül hagyják azt a problémát, amely a feltételezett veleszületettségéből származik:

Ha akár evolúciós folyamat során, akár valamiféle teremtés során került az ember veleszületett képességei közé a racionalitás, úgy azt géppel legfeljebb imitálni lehetséges (lásd például Searle kínai szoba-érvét, SEARLE 1996)⁶, vagy más szavakkal: a mesterséges rendszerek csak szoftver szinten képesek a racionalitás alacsonyabb-magasabb szintű utánzására, márpedig a problémát hardver szinten, nem pedig szoftver szinten kell megoldani.

A kételkedők pedig a mesterségesintelligencia-kutatások számbavételével könnyen kijelenthetik, hogy ilyen előfeltevések alapján nem csoda, hogy a kutatók más eredményt kaptak, mint várták: alapvetően „buta” gépek jöttek létre, sőt, még csak racionálisoknak sem lehet őket nevezni. Kielezve, a kudarc több szinten is jelentkezik, mivel a gépek nem képesek sem a racionalitás, sem a tágabb értelemben vett emberi gondolkodás tulajdonságait tükrözni. A hiba így több gyökérrel is rendelkezhet:

- Az ember sem teljesen racionális lény (mint azt ma már számtalan pszichológiai kutatás bizonyította, döntéseiben a racionalitás csupán az egyik tényezőnek tekinthető);
- a racionalitás sem pontosan úgy működik, mint azt eredetileg leegyszerűsítve elképzelték (például már egységes racionalitásfogalomról is nehéz beszélnünk, helyette többféle racionalitásról, illetve stratégiákról lehet beszámolni).

A klasszikusnak is nevezhető racionalitásfelfogást utóbb több nézőpontból is cáfolták. Egyrészt, a korábban élesen szembeállított racionalitás és emóció dichotómiát az érzelmeknek a döntéshozatalban és cselekvési stratégiákban betöltött szerepére hi-

⁵ Az értelem gépiesítése vagy gépiesíthetősége ellen persze Descartes vehemensen tiltakozna: szerinte az ember lényegénél fogva gépként elképzelhetetlen lenne, míg más élőlények viselkedése arra utal, nem többek egyszerű gépeknél. Figyelemre méltó azonban, hogy módszere megalkotásában a matematikát tartotta példaértékűnek; valamint az, hogy nem vette észre, a közös emberi racionalitás-karakter és a csalhatatlan módszer miként vezethet a gépi értelem alappondolatához.

⁶ Searle érve – akárcsak más, a mesterséges intelligenciát megkérdőjelező érvek – némiképp kényes területre téved, mivel szubjektív, minőségi kifogást emel. Az érv egyébként meglepő párhuzamot mutat Descartes azon gondolatmenetével, melyben az ember és az állat (valamint a gépek) különbségét taglalja. Két érvet említ ekkor: (1) az ember nyelvhasználó lény, amelyre nincs példa az állatvilágban és minden bizonnyal erre gépek sem lehetnek képesek; (2) az ember tudatosan cselekvő lény, amelyre szintén nem találunk más példát (DESCARTES 1992, 64. skk.).

vatkozva haladták túl. Értelemszerűen, ezen a téren elsősorban az etikai vonzatok kerültek előtérbe. A kérdés hangsúlyos kísérleti vizsgálatát az utóbbi évtizedek neurológiai kutatásai szolgáltatták, melyek során például agysérült betegeken elemezték a racionális döntéshozatal képességét (ld. pl. DAMASIO 1996, 131–145; 190–194).

Másrészt, számos pszichológiai kísérlet világított rá arra, hogy a racionalitás tekintében nem lehet figyelmen kívül hagyni a kontextualizáltság kérdését. A racionális gondolkodást és következtetést nagyban meghatározza az adott probléma környezete; e környezet alapján pedig az egyes személyek jobban, illetve gyengébben oldják meg ugyanazt a feladatot (olyan ismert kísérletekre szokás itt hivatkozni, melyek elsősorban arra világítanak rá, hogy konkrét tartalommal megtöltött problémákat könnyebb megoldani, mint absztrakt problémákat). Az evolúciós pszichológia erre a problémára olyan választ kínál, amely szerint az egyének logikus következtetéseit nagyban befolyásolja az, hogy a környezet miatt milyen diszpozíciókkal viszonyulnak egy adott helyzethez. A tisztán logikai (elvont) problémák megoldása lényegesen nagyobb nehézséget jelent mindenki számára, mint azok, amelyek magát az egyént is érintik: Cosmides szerint az ember racionális gondolkodása egy, a csalók leleplezésére irányuló mechanizmussal van összekötve, amely az egyén túlélését segíti azzal, hogy az együttélésből adódó nehézségeket és esetleges hátrányokat megpróbálja kiszűrni. Az önzetlenség és az együttműködés kényszere csak akkor szolgálja sikeresen a fennmaradást, ha az egyén képes megakadályozni a visszaéléseket – így a képesség, hogy a csalást leleplezze, szorosan összekapcsolódik az önzetlenséggel. Cosmides egyenesen úgy gondolja, hogy ez a döntési mechanizmus fölérendelt viszonyban van a racionális vagy logikus gondolkodással, amivel megmagyarázható számos olyan eset, amikor az egyének téves – de számukra racionálisnak tűnő – döntéseket és ítéleteket hoznak (ld. COSMIDES 1989; COSMIDES–TOOBY 1992). Ezek a döntések olykor egybeesnek a logikusnak tűnő megoldással, olykor pedig nem.

Szintén e kutatások világítanak rá arra is, hogy a racionalitás kritériumai sok esetben kultúrafüggők. Ennek egyik ismertebb példája a *kpelle* libériai törzs körében elvégzett következtetési vizsgálat, melynek során a kísérleti személyek meglepő, az európai gondolkodás számára teljességgel irracionálisnak tűnő válaszokat adtak egy egyszerű szillogizmus konklúziójaként (ld. COLE 1971).⁷

Végül pedig több vizsgálat azt is kimutatta, hogy a racionalitás számos aspektusa csak a szocializálódás során szilárdul meg, így azok nem lehetnek veleszületettek. Ez a folyamat egyben arra is rávilágít, miért lehetséges, hogy a racionális gondolkodás tekintetében lényeges különbségek figyelhetők meg az azonos kultúrában élő egyének között is. Összességében, nem lehet különösebben elcsodálkozni azon, hogy az egyének számtalan esetben a racionális gondolkodással ellentétes ítéletet alkotnak, miközben maguk ésszerűnek tartják azokat. Ez azonban mégsem jelenti, hogy az ember ne lenne racionális lény; sokkal inkább azt jelzi, hogy a racionalitás csupán egy az

⁷ A kutatók és a kísérleti személyek beszélgetésének egy érdekes részlete:

Kérdés: Flumo és Yakpalo mindig együtt isznak nádlevet. Flumo nádlevet iszik. Iszik-e Yakpalo is nádlevet?

Válasz: Flumo és Yakpalo együtt isznak nádlevet, de amikor Flumo először ivott, Yakpalo aznap nem volt ott.

Kérdés: De azt mondtam, hogy Flumo és Yakpalo mindig együtt isznak nádlevet. Az egyik nap Flumo nádlevet ivott. Ivott-e Yakpalo is nádlevet?

Válasz: Azon a napon, amikor Flumo nádlevet ivott, Yakpalo nem volt ott.

Kérdés: Miért?

Válasz: Mert Yakpalo aznap a farmjára ment, Flumo pedig a városban maradt.

ember veleszületett gondolkodási képességei közül (és hierarchikusan talán nem is a legfontosabb), amely a döntési stratégiák egy lehetőségét képviseli.

Mindezek után világosnak tűnik, hogy a racionalitás helyes értelmezéséhez a mesterséges intelligenciával kapcsolatos kutatásokban is túl kell lépni a klasszikus felfogáson.

Ennek egy példáját a szakértői rendszerek különböző elméletei szolgáltatathatják. Ezek olyan próbálkozások a mesterséges intelligencia terén, melyek manapság is sikeresen működnek egy-egy jól körülhatárolható szakterületen, mint amilyen a diagnosztika vagy a mérnöki modellezés. A szakértői rendszerek két fő részre oszthatóak: (1) tartalmaznak egy adatbázist, amely lényegében a szakértő tudásának reprezentációjaként fogható fel (ezeket a „valódi” szakértők megfigyelésével, beszámoltatásával és rendszerezésével gyűjtik össze, majd programozzák a gépekbe); (2) tartalmazzák azokat a következtetési eljárásokat, amelyek alapján egy szakértő meghozza döntését vagy felállítja diagnózisát (szigorú értelemben véve ezek feleltethetők meg a korábban tárgyalt szillogizmusoknak).

A szakértői rendszerek megalkotói azonban a rendszerek finomítása és pontosítása során hamarosan abba a problémába ütköznek, hogy azok lassanként áttekinthetlenné válnak. A jobb és pontosabb eredmény érdekében egyre aprólékosabb reprezentációkra és következtetésekre van szükség. Ez az összetettségi probléma előbb-utóbb ahhoz vezethet, hogy a rendszer meglehetősen labilissá válik. Könnyű belátni, hogy ez valamely korábbi, hibás előfeltevésre vezethető vissza, hiszen az eredmény ellentétes azzal, amit a „valódi” szakértők produkálnak, akik a tapasztalat bővülésével egyre pontosabbakká válnak.⁸ Azaz, megint a korábbi racionalitás-tulajdonítással van probléma: a reprezentációk és a következtetési sémák bővülése ellentétes hatást ér el, ha azok elszigetelten, felsorolásszerűen kerülnek a rendszerbe.

Így a szakértői rendszerek építésének más stratégiát kell követnie. Ez a módszer pedig azon az egyszerű megfigyelésen alapszik, miként válik valakiből szakértő (szemben a korábbi stratégiával, amely arra törekedett, hogy minél tapasztaltabb szakértők kész tudását térképezze fel).

A kezdő lépéseket a tanítás adja: ekkor a dekompozíció során nyert kontextusfüggetlen vonások elsajátítása folyik, instruktor segítségével. Valójában a tanulásnak ehhez a kezdeti szakaszához nem feltétlenül szükséges a közvetlen tapasztalat. Ez a folyamat még gond nélkül megfeleltethető a számítógépszerű programkövetésnek, ám igazi szakértőről ekkor még nem beszélhetünk.

A haladó jelölt átlép egy másik tanulási fázisba, amelyben saját tapasztalatai, valamint újabb instrukciók alapján kiemeli a szakterület lényeges aspektusait. Ezt a szakaszt a példák alapján történő tanulás jellemzi, és megkívánja a jelölt aktív részvételét is.

A hozzáértő, de még kezdő szakember már képes a releváns vonások hatalmas mennyiségének felismerésére, ez azonban gyakran még kezelhetetlennek tűnik számára, mivel egyelőre nem tudja teljes bizonyossággal kiemelni a leglényegesebbeket. A tanulásnak/tapasztalásnak ebben a szakaszában az jellemző, hogy bár a szakértői döntések többsége végül helyesnek bizonyul, a szakértő módszere nem nevezhető hatékonynak – és természetesen nem elhanyagolható hibaszázalékkal dolgozik.

Az utolsó előtti szakaszban megjelenik az, amit a köznyelv a „profi” fogalommal jelöl: a szakértő könnyedén és kellő biztonsággal oldja meg a nehezebb problémákat is.

⁸ A mesterséges intelligencia szakértői rendszereken keresztül történő bírálatához ld. DREYFUS 1972; PENROSE 1991.

A professzionális szakértő döntéshozatalában tetten érhető az intuitív viselkedés, melynek során a szakértő a korábbi pozitív és negatív példák hatására kialakult, nem teoretizált ismeretek alapján kezd dolgozni.⁹

Végül pedig az igazi szakértő még a profinál is finomabb, nem teoretizált különbségtételekre és osztályozásra képes. Az alapvető fejlődés ott ragadható meg, hogy míg a profi a probléma felismerésében kiválóan teljesít, utána azonban tudatos mérlegelés során kell döntenie, mit tegyen, addig a szakértő számára már maga a cselekvés is egyértelmű – mondhatni, döntéshozatalában teljesen intuícióját követi, nem pedig explicit szabályokat.¹⁰

Érthető tehát, hogy az egyre finomabb reprezentációkkal és következtetési stratégiákkal felruházott rendszerek nem működtek megfelelően: egyszerűen, a valódi szakértők nem ezeken az alapokon hozzák meg döntéseiket. Helyettük inkább olyan rendszerekre lenne szükség, amelyek képesek a szakértő gondolkodását utánozni. Az a racionalitás, amelyet a reprezentációkon és következtetési módokon keresztül ültettek beléjük, a valódi világban felmondja a szolgálatot; sőt, nem egy esetben irracionálisnak és értelmetlennek tűnhet.¹¹

Maga a probléma végső soron nem is szakértői szinten jelentkezik, hiszen hasonló folyamatok tapasztalhatók a mindennapi élet terén is. Ismeretes, hogy a mesterséges intelligencia kutatásának egyik jelentős problémája abból adódik, mennyire nehézkesen viselkednek a programok a mindennapi élet számunkra egyszerű világában. Valójában itt is arról van szó, hogy az egyének a szocializálódás során – társadalmi szinten – észrevétlenül szakértőkké válnak. Az emberi együttélés során minden bizonytalanságot a *common sense* terén is.

IV.

Mit lehet mindebből a racionalitás innátista felfogása és a mesterséges intelligencia kapcsolatára nézve kikövetkeztetni? Könnyűnek látszik az a válasz, hogy a veleszületett racionalitás – legyen szó akár kreacionista, akár evolúciós felfogásról – az ember jellegzetessége, és mint ilyen, nem tulajdonítható gépeknek (ezt látszik alátámasztani az a tény is, hogy valódi intelligens rendszert eddig még nem sikerült létrehozni, és a szakértői rendszerek sem képesek maradéktalanul utánozni az emberi döntéshozatalt). Másik oldalról vizsgálva azonban, a mesterséges intelligencia pártján állók is felhasználhatják az innatizmus érvét arra, hogy a gépi intelligencia lehetőségét – legalább elvben – igazolják: hiszen az emberben készen fellelhető struktúrák megfeleltethetők kell, hogy legyenek a gépekbe készen beültetendő struktúráknak, avagy programoknak.

⁹ Megfigyelhető az is, hogy a professzionális szakember időnként nem képes pontosan beszámolni arról, miért, milyen információk alapján dönt egy adott stratégia mellett.

¹⁰ Természetesen ez a megállapítás sarkítottnak tűnhet; nem véletlen azonban, hogy a szakértői tudás feltérképezése során a legtapasztaltabb egyének esetében a kutatók gyakran olyan válaszokba ütköznek, mint: „Evidens, hogy így kell cselekedni”; „Ebben az esetben csak így lehet a problémát megoldani”.

¹¹ A szakértői rendszerek egyik kiemelkedő példája a sakk. Többen is feltérképezték már, miként válik egy jelöltből nagymester, valamint azt, hogyan hozza meg a nagymester egyes döntéseit. Ezek átültetése a mesterséges rendszerekbe azonban csak nehézkesen valósul meg. A szenzációnak számító Deep Blue, amely néhány éve legyőzte Garri Kaszparovot, az akkori sakkvilágbajnokot, sikerét leginkább hatalmas számítási kapacitásának köszönhette, amellyel a másodperc tört része alatt lépésvariációk tömegét volt képes kiszámítani. Ezzel szemben a nagymester esetenként csupán néhány lehetséges lépést vesz számba, a lehetőségek repertoárjának túlnyomó részét eleve, intuitív módon kizárja.

A kérdés tehát arra a teoretikus problémára vezethető vissza, hogy vajon meddig terjed a racionalitás: csakis a pontos következtetési stratégiáig, a szillogizmusokig és az algoritmizálható feladatokig, vagy esetleg túllép azon? Szigorú értelemben véve, amennyiben az előbbi lehetőség igaz, akkor a veleszületettség feltételezése egyfajta következetes programhasználatot jelentene. Amennyiben azonban a racionalitás nem merül ki ebben, és a szakértői intuíció vagy az emocionális alapokon meghozott döntés is racionálisnak nevezhető, a veleszületettség nem lehet teljes mértékben igaz. Márpedig a szakértői rendszerek példája (de hivatkozhatnánk itt akár a mindennapi életben való tájékozódásra is) éppen ezt sugallja.

A veleszületettség tehát csak olyan mértékben lehet igaz, amennyiben egyfajta racionális előstruktúrákat vagy diszpozíciókat érint. Az, amit teljes joggal nevezhetünk racionális gondolkodásnak, nem nélkülözheti a későbbi tanulást, tapasztalást és szocializálódást. A mesterséges intelligencia kérdésére visszavetítve ez azt jelenti, hogy egy intelligens rendszernek képesnek kell lennie arra, hogy a szakértőkhöz hasonló „fejlődési perióduson” menjen végig; a készen kapott program nem lehet intelligens. Ehhez az eredmények visszacsatolására, heurisztikák önálló kialakítására és az explicit módon nem reprezentálható struktúrák felismerésére van szükség.

Nem lehetséges azonban, hogy teljesen lemondjunk a veleszületettség feltételezéséről. A mesterségesintelligencia-kutatásban ez természetesnek tűnik, mert minden rendszer alapját valamely program adja meg. Ugyanez elmondható azonban az emberre is. Az alapvető következtetési stratégiák és a megismerés néhány egyszerű eleme (mint például a kategóriaalkotás és -megértés) veleszületett jellegzetességeket mutatnak.¹² A fentiekben vázolt korlátozott veleszületettség tehát azt jelenti, hogy a program – vagy az elme – fejlődni képes alapstruktúrákkal rendelkezik; ezek a struktúrák azonban egyáltalán nem tekinthetők véglegesen lezártaknak és kötelező érvényűeknek.

IRODALOM

- COLE, Michael et al. 1971. *The Cultural Context of Learning and Thinking*. New York, Basic Books.
- COSMIDES, Leda 1989. The Logic of Social Exchange: Has Natural Selection Shaped How Humans Reason? *Studies with the Wason Selection Task*. *Cognition*, 31. 187–276.
- COSMIDES, Leda – TOOBY, John 1992. Cognitive Adaptations for Social Exchange. In Barkow, Jerome H. – Cosmides, Leda – Tooby, John (eds.): *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*. New York, Oxford University Press.
- DAMASIO, Antonio R. 1996. *Descartes tévedése*. Ford. Pléh Csaba. Budapest, Aduprint.
- DESCARTES 1992. *Értekezés a módszerről*. Ford. Szemere Samu – Boros Gábor. Budapest, Ikon.
- DREYFUS, Hubert L. 1972. *What Computers Can't Do: A Critique of Artificial Reason*. New York, Harper & Row.
- KANT 1995. *A tiszta ész kritikája*. Ford. Kis János. Szeged, Ictus.
- LENAT, Douglas B. – GUHA, R. V. 1990. *Building Large Knowledge-Based Systems*. Reading, Mass., Addison-Wesley.
- PENROSE, Roger 1991. *Computerdenken: Die Debatte um Künstliche Intelligenz, Bewußtsein und die Gesetze der Physik*. Heidelberg, Spektrum der Wissenschaft.
- PLATÓN 1984. *Összes művei*. Ford. Szabó Miklós. Budapest, Európa.
- SEARLE, John R. 1996. Az elme, az agy és a programok világa. Ford. Thuma Orsolya. In Pléh Csaba (szerk.): *Kognitív tudomány*. Budapest, Osiris–Láthatatlan Kollégium.
- WELSCH, Wolfgang 1995. *Vernunft: Die zeitgenössische Vernunftkritik und das Konzept der transversalen Vernunft*. Frankfurt am Main, Suhrkamp.
- WELSCH, Wolfgang 1999. Reason: Traditional and Contemporary or Why Should We Still Speak of Reason after All? <http://www2.uni-jena.de/welsch/>

¹² Olyan egyszerű következtetési szabályokra lehet itt gondolni, mint pl. „egy tárgy nem lehet egyszerre két helyen”. A következtetési szabályok és tudásbázis veleszületett-ellátottságára a mesterséges intelligencia-rendszerekben ld. LENAT–GUHA 1990.