

Kenéz László:
*Logikai és fizikai konstrukció**

Bevezetés

Az alábbiak során azt vizsgálom meg, hogy Russell konstrukciós eljárása felhasználható-e az elme megismerésére a mesterséges intelligencia kutatásában. A probléma ismert: amennyiben az elme bizonyos tulajdonságai kizárólag az első személyű perspektíva felől ismerhetők meg, mit kezdhetnénk az elmével a gépek felől, amely esetében legfeljebb a harmadik személyű perspektíva viselkedési adataira hagyatkozhatunk. Nos, az alapgondolat igen egyszerű: ha a *logikai konstrukció* a megismerés eszköze, miért ne lehetne egy *fizikai konstrukció*, vagyis *maga a gép* is az? Tudjuk ugyanis, *Russell* a konstrukció segítségével érzetadatokkal váltja ki az elvont fizikai fogalmakat, a fizikát ily módon a közvetlen megismerés fundamentumára helyezve (Russell 1914/1976). Vajon nem lehetne egy nagy fordulattal ugyanilyen módon, de más kiindulópontot választva, a gépek megfigyelésiadataival kiváltani az elmét, bevonva ezzel a harmadik személyű perspektíva adataira alapozott tudás körébe?

Ki fog derülni. Elsőként inkább arról érdemes néhány szót ejteni, miért is olyan fontos ez. Mint ismeretes, manapság egyre többen vélik úgy, hogy a természettudomány, s szűkebben a kognitív tudomány, és még szűkebben a pszichológiai igényű mesterséges intelligencia kutatás *nem* boldogulhat az elmével. Mivel nem mindenki gondolja azt, hogy ez így van, azt is meg kell világítani, miért gondolják sokan (velem együtt) úgy, hogy mégis. A baj (már ha van) mellel *episztemológiai* gyökerű, mint igyekszem megmutatni, így tehát

* A szöveg a 2005 szeptember 29-30 között, az ELTE-n rendezett *Leírás, Jelentés, Denotáció* c. Russell-konferencián hangzott el.

episztemológiai gyógymódra van szükség. Márpedig Russell eljárása éppen ezzel szolgál. Érdemes tehát alaposabban megnézni, hátha valóban segíthet.

Az alábbi útvonalat ajánlom mindehhez. Kezdő lépésként (1) felvázolom Russell konstrukciós eljárását, majd (2) kiemelek egy olyan vázat, egészen pontosan formális feltételrendszert, amely aztán lehetővé teszi, hogy a konstrukciót bármilyen tetszőleges tartalommal feltölthessük, már amennyiben a terület, ahol alkalmazni kívánjuk, eleget tesz e feltételrendszer kívánalmainak. Röviden, általánosítjuk a russell-i módszert, hogy ne csak érzetadatokkal, hanem más adatokkal is dolgozhassunk. Ezután veszem szemügyre (3) az elme helyzetét a természettudományban, azért, hogy megvilágítsam a problémát, amelyre Russell-től ihletve remélem a megoldást. Ekkor vizsgáljuk csak meg (4), hogy alkalmazható-e egyáltalán a konstrukciós eljárás e problématerületen. Másképp, arra leszünk kíváncsiak, hogy az elme természettudományos megközelítése eleget tesz-e a konstrukció a második fejezetben kiemelt feltételeinek, a gondolatmenetet a mesterséges intelligencia kutatás közegére szűkítve. Ha igen, akkor a gépek értelmezhetőek úgy, mint fizikai konstrukciók, s ezáltal eszközként szolgálhatnak az elme megismerésére, ha nem, akkor nem. Mivel pedig csakis e vizsgálat közben érünk el arra a pontra, ahonnan majd észrevehető lesz, miféle zsákutcába is tévedtünk, egy merész fordulattal csak ekkor vezénylem a menetet más irányba (elárulom, a kontinentális fenomenológia felé). Végül pedig (5) egy olyan konstruktív elmetudomány képét villantom fel, amelyben lehetővé válik, hogy a gépeket az elmemegismerés eszközeinek tekintsük, azt is sugallva, hogy a mesterséges intelligencia kutatás, bár jelenleg az elmefilozófiával jár jegyben, egy hosszú távú, kölcsönösen előnyös, kiegyensúlyozott kapcsolatot mégis más partnertől (a kontinentális fenomenológiától) remélhet.

(1) Logikai konstrukció

Jól tudjuk, amikor csak igény támadt arra, hogy a filozófia tudományá váljon, új módszer született. Nem csoda hát, ha, mint minden nagyra törő gondolkodó

az ókortól napjainkig, Russell is megfogalmazza a „tudományos filozofálás legfőbb maximájá”-t, amely a következőképpen szól:

Ahol csak lehetséges, logikai konstrukciókkal kell helyettesíteni a levezetett [inferred] entitásokat. (Russell 1914/1976, 252)

Fejtsük ki, mit jelent ez. A kulcsszavak rendre a következők: logikai konstrukció, helyettesítés, levezetett entitás. Vegyük először ez utóbbi kifejezést.

Mint ismeretes, Russell egész munkálkodásában döntő szerephez jut egy különbségtétel, amely tartalmában ugyan változásokat mutat, intenciójában azonban nem. Russell szerint ugyanis a dolgokat, még pontosabban a különféle *entitásokat*, két alapvető osztályba sorolhatjuk aszerint, ahogyan *megismerjük* őket. Az első osztályba a *közvetlenül* megismert entitások kerülnek, míg a másodikba azok, amelyeket csak *közvetve* ismerünk meg. Russell a múlt század elején, az *On Denoting*-ban így tesz különbséget „ismeret” és „valamiről való tudás” között (Russell 1905/1985,144), öt évvel később pedig az ismertség [acquaintance] és leírás [description] révén szerzett tudás között (Russell 1910/1976).

Ezek a megkülönböztetések azonban rögtön plasztikusabbá válnak, ha a korrelatív tárgyakra, vagyis a különféleképpen megismert entitásokra is példát hozunk. Míg az *érzetadatokról* [sense-data] közvetlen ismeretet nyerhetünk, állítja Russell 1914-ben, addig az *anyagról* közvetlen ismeretünk nem lehet, csak tudomásunk, így például csakis denotáló kifejezések révén tudhatunk róla, ahogy ez már az *On Denoting*-ban is (Russell 1905/1985, 165) áll. Következésképp, az érzetadatok a *fellelhető* entitások közé sorolandók, az anyag pedig a lényegi természetüknél fogva *fel nem lelhető* entitások közé tartozik (Russell 1914/1976, 236).

Miért fontos ez? Egy közvetlenül ismert entitás egészen más *ismeretelméleti* státusszal bír, mint azok, amelyekről csak közvetett tapasztalattal rendelkezünk. Egyszerűen szólva, az előbbi entitások léte bizonyos, míg az utóbbiaké nem az.

Ami az érzetadatokat illeti, tudjuk, hogy léteznek, amíg adatok a számunkra, és hogy ez képezi a külső különös entitásokra vonatkozó minden tudásunk ismeretelméleti alapját. (Russell 1914/1976, 239)

Az ismeretelméletileg kívánatos állapot így nyilvánvalóan az lenne, ha a közvetve megismerhető „külső” entitások létét valamiképp a közvetlen tapasztalat számára hozzáférhető entitások, vagyis az idézetben megnevezett érzetadatok segítségével lehetne bebiztosítani.

A kérdés tehát az, miképp? Nézzük elsőként, mi jellemzi a közvetett ismeretet. Russell szerint a közvetve megismert entitásokra csak *következtetünk avagy levezetjük őket* [infer]. Megvan tehát az első keresett fogalom, ám haladjunk rögvest tovább, mert az a kérdés, miért nem optimális ez. Nos, a következtetés avagy levezetés nem képes olyan megfelelő kapcsot képezni a kiindulópont és a végpont között, ami ismeretelméleti szempontból elégséges lenne, mivel „kétséges” eljárás (Russell 1914/1976, 252). Egy levezetés ugyanis mindenkor csak állít valamely létet, amelyről azonban közvetlenül nem tudunk meggyőződni. A levezetett entitások nem adottak, hanem feltételezettek, létük nem bizonyos, hanem hipotetikus. Márpedig ily módon ez a lét csak „óhajtólagos”, következésképp megfelelő metafizikai aládúcolást igényel. Vagy lehetőség van valamilyen más megoldásra is? Csakhogy ha nem kívánunk metafizikai posztulátumokkal élni, mi mást tehetünk még?

Russell megoldása az, hogy a levezetett vagy hiposztazált entitásokat közvetlenül megismerhető (de legalábbis kevésbé hipotetikus) entitásokkal *helyettesíti*. Ha ugyanis helyettesítünk, akkor nem merészkedünk el a kiindulóponttól: mindvégig fenntartjuk a közvetlen ismertséget azokkal az elemekkel, amelyeket a megismerendő levezetett entitás helyébe állítunk.

Megvan tehát a második kulcsfogalmunk is, bár még nem világos, hogyan lehet ezt a szubsztitúciót végrehajtani.

Legyen az érzetadatok halmaza az episztemológiailag kitüntetett bázis. Amint azt ki-kí maga is megállapíthatja, a sorjázó érzetadatok (vagyis nekem most az a barna folt, és az attól jobbra elhelyezkedő zöldes-sárgás folt, hadd ne folytassam) *illetve* az anyag között még óriási a távolság. Amikor a fizika a tömegvonzásról beszél, biztos, hogy nem az én érzetadataimra gondol. Ez pedig azt jelenti, hogy az érzetadatok úgy, ahogyan elem kerülnek, nem alkalmasak az anyag kifejezéssel illetett hipotetikus entitás helyettesítésére. Valamilyen *műveletre* van szükségünk, ám mégis olyanra, amelynek során nem megy veszendőbe az érzetadatok kiemelt episztemológiai minősítettsége.

Nos, pontosan ezzel szolgál a *logikai konstrukció* eljárása. S ezzel megvan a harmadik kifejezés is, amelynek kapcsán rögvést említést érdemel az eljárást kiötlő *Whitehead* neve is, hiszen Russell tőle származtatja a módszert s magának csak az alkalmazás névumát tartja fenn (Russell 1914/1976, 254). Nézzük meg azonban alaposabban, hogyan képes sajátos teljesítményére a konstrukció. Russell így írja le az eljárást:

Ha adva van egy sor kijelentés, amely névlegesen a feltételezett levezetett entításokról szól, megfigyeljük azokat a tulajdonságokat, amelyeket meg kell követelnünk a feltételezett entításoktól, hogy kijelentéseink igazak legyenek. Azután – kis logikai ötletességgel – megkonstruáljuk bizonyos kevésbé hipotetikus entításoknak egy olyan logikai függvényét, amely rendelkezik a megkövetelt tulajdonságokkal. A [levezetett] entításokat helyettesítjük az így megkonstruált függvénnyel, s ily módon a kérdéses kijelentéshalmaz új, kevésbé kétséges értelmezését kapjuk meg. (Russell 1914/1976, 253)

Jól látható tehát, az eljárás ismereti hozadéka nem megvetésre méltó. Mielőtt azonban tovább méltatnánk, szenteljük figyelmünket az idézet legfurcsább

kifejezésének, mellesleg az egész gondolatmenet lelkének, a *logikai függvény* fogalmának. Mivel később ez fontos lesz, pontosan látnunk kell, mi is ez.

Nos, legyen y egy hipotetikus entitás, amelyet helyettesíteni akarunk kevésbé hipotetikus entításokkal, mondjuk x_1 -gyel és x_2 -vel. Ahhoz, hogy y -t helyettesíthessük, az előnyben részesített x_1 -nek és x_2 -nek olyan tulajdonságokat kell mutatnia, mint y -nak. Ha pedig egyenként erre nem képesek, s egyenként így helyettesíteni sem képesek y -t, akkor oly módon kell *egymáshoz rendelniük* őket, hogy az így képződő komplex egység képes legyen a szóban forgó tulajdonság implementálására. Röviden, egy f *függvényt* kell előállítanunk, amely a kívánt módon rendeli egymáshoz x_1 -et és x_2 -t. Ez a „metafizikai valósággal” nem rendelkező „megkonstruált fikció” (Russell 1914/1976, 253), vagyis maga a függvény (amely *csakis* x -eket tartalmaz) lesz az, amellyel helyettesíthetjük y -t. Vegyük tehát észre, nem metafizikai, hanem szigorúan ismeretelméleti vállalkozásról van szó. Később, a neutrális monista Russell már nem lesz ennyire óvatos, amikor konstruál (vö. Russell 1927), ez azonban most mellékes.

Ha tehát valóban sikerül megszerkeszteni egy megfelelő függvényt, akkor ennek alkotórészei egytől-egyig kevésbé hipotetikus entítások lesznek. Russellnek ez azért nem sikerül: két levezetett entitást, nevezetesen, mások érzetadatait, valamint az úgynevezett *szenzibiléket* fel kell tennie (Russell 1914/1976, 255). Végző soron azonban mégiscsak eléri célját. Márpedig ily módon a tudomány óriásit lép előre, hiszen a bizonyosság talaján maradván éri el a bizonytalant. A konstrukció tehát egy adott kijelentéshalmaz ismeretelméletileg előnyösebb *értelmezéséhez* segít hozzá. Nem tudásunk mennyisége gyarapszik, hanem, ha szabad így kifejezni, *minősége* javul. A későbbiekben ez ugyancsak fontos lesz.

De ennyi elég is. Mivel távolabbi célom az, hogy ezt az eljárást más területen használjam fel, tehát ne arra alkalmazzam, hogy érzetadatokkal helyettesítsem például a fizika fogalmait, a konkrét részletektől eltekintek, és nem követem tovább Russellt. Az idézet alapján inkább azt határozom meg, milyen *formális feltételek* között használható a fenti eljárás. Másképp, megkísérlem meghatározni

azokat a feltételeket, amelyeknek teljesülnie kell ahhoz, hogy a logikai konstrukció russell-i módszerét egészen más területen is munkába állítsuk.

(2) Konstruktív feltételek

Lássuk tehát újra az idézetet. Kizárólag arra leszünk tehát kíváncsiak, milyen általános kívánalmak támaszthatók a konstrukció körülményeivel szemben. Négy feltételt fogalmaztam meg az eljárás leírása alapján:

- (i) Adva kell legyen egy megkülönböztetés hipotetikus és kevésbé hipotetikus entitások között.
- (ii) Rendelkezésre kell, hogy álljanak bizonyos kijelentések a hipotetikus entitásokról.
- (iii) A hipotetikus entitások megfigyelhető tulajdonságokkal kell, hogy rendelkezzenek.
- (iv) Lehetséges kell legyen, hogy a kevésbé hipotetikus entitások is rendelkezhessenek a megfigyelt tulajdonságokkal.

Ez az a formális feltételrendszer, amelynek teljesülnie kell azon a területen, ahol a konstrukciót alkalmazni kívánjuk. Lássuk tehát, miért alkalmazható ez a fizika esetén.

Az érzetadatok, mint láttuk, Russell szerint nemhogy kevésbé hipotetikus entitások, mondjuk az anyagnál, hanem ismeretelméleti szempontból épphogy azok a közvetlenül megismert entitások, amelyekre a tudásnak támaszkodnia kell. Megvan tehát a kétféle entitás (i). A fizika számos kijelentést tesz az anyagról, így (ii) is rendben van. Az anyagnak bizonyos tulajdonságokkal kell rendelkeznie ahhoz, hogy a rá vonatkozó kijelentések igazak legyenek. Ilyen tulajdonság például a kiterjedés (iii). Érzetadataim folytonos változások közepette különféleképpen tagolják látóteremet, s ilyeneként, kiterjedtek.

Rendelkeznek tehát azzal a tulajdonsággal, amelyet az anyagtól követelünk meg ahhoz, hogy a rá vonatkozó kijelentés igaz legyen (iv). Mivel a feltételek teljesültek (igaz, egyelőre csak egyetlen kijelentésre), kezdetét veheti a konstrukció. Akit érdekel, hogyan hajtja végre ezt Russell, még hozzá átfogó érvennyel, nem kevés logikai ötletességről tanúbizonyságot téve, az lapozza fel a megfelelő részeket (Russell 1914/1976, 256 skk.). Számunkra ennek most nincs jelentősége, egyvalami azonban külön figyelmet érdemel.

A feltételek kiemelése e pontról ugyanis már jól érzékelhetővé teszi, hogy a konstrukció döntő jelentőségű a különféle kijelentések *verifikációjában* is. Lehetővé teszi ugyanis, hogy olyan kijelentéseket, amelyek feltételezett entitásokról szólnak, olyan adatok segítségével igazoljunk, amelyeket közvetlenül (vagy általánosabban, ismeretelméletileg a kívánatos módon) ismerünk meg. Ez annyit tesz, hogy a hipotetikus entitásokról szóló kijelentéseket kevésbé hipotetikus entítások segítségével verifikálhatjuk. Világos ennek jelentősége a fizika területén, ahol Russell szerint, a fizika tárgyait „csakis az érzéki adatokhoz való viszonyukon keresztül” verifikálhatjuk, hiszen

[a] fizikát nem tekinthetjük úgy, mint ami érvényes módon tapasztalati adatokon alapul, amíg [például] a hullámokat nem sikerült kifejeznünk a színek és már érzéki adatok függvényeként. (Russell 1914/1976, 236)

Talán nem szükséges kifejteni, hogy ez a mozzanat szintén általánosítható. Később azt is látni fogjuk, miért fontos ez, most azonban megtudva azt, amit kívántunk, foglaljuk össze az eddigieket.

Russell egy új módszer általános alkalmazását javasolja a tudományos filozofálásban, amely oly módon javítja a tudást, hogy a megismerendő entitást az ismeretelméletileg preferált entítások függvényeként mutatja be. E függvények képzése konstrukcióval történik, képzésük célja pedig az, hogy a függvény helyettesítse a feltételezett entitást s így módon a róla tett kijelentéseket

a függvényt alkotó, ismeretelméletileg előnyben részesített entitások segítségével verifikálhassuk.

Rendkívül vonzó tehát a gondolat, hogy ezt a módszert olyan területen is próbára tegyük, ahol a körülmények formálisan megfelelőnek tűnnek, annak ellenére, hogy tartalmilag igencsak különböznek az eddigiektől. Különösképp akkor, ha olyan területről van szó, ahol a megismerendő tárgy már régóta makacsul ellenszegül minden szabályos megismerési kísérletnek.

(3) Elme és természettudomány

A kortárs természettudomány, amennyiben a manapság elfogadott és érvényben lévő keretek között tör a megismerésre egy materialista metafizika égisze alatt, empirikus eszközökre hagyatkozva kísérli megmagyarázni az elme jelenségét. Vagyis, nem nyugodhat addig, amíg meg nem mutatja, hogy az elme miként jön létre és működik *fizikai* folyamatok bázisán. Márpedig ez igencsak bajos ügy, ám korántsem amiatt, mert egyelőre nem sikeres. Nagyon úgy fest ugyanis, hogy elvi akadályok nehezítik a munka bevezetését.

Mivel elvi okokkal a filozófia foglalkozik, lássuk, mit szól a filozófia a tudomány elmével kapcsolatos ügyködéseihez. Nézzük meg tehát, mi jellemzi azt az episztemológiai miliőt, amelyben az elme tudományos megismerése a cél.

Manapság az elmével kapcsolatban két adathalmazra szokás hivatkozni. Ahhoz, hogy a megfelelő kontraszthatást elérjük, e két típusú ismeretet kellő élességgel el kell választanunk egymástól. Röviden, idealizálunk. Lássuk, hogyan és miért.

Az elmefilozófia terminológiáját használva, *fenomenológiai* nézőpontból megállapíthatom például, hogy mentális állapotaim sajátosan minősítettek. Ilyen adatok esetén az *én* saját elmémről van szó, ezért *első személyű perspektíváról* is beszélhetünk. Fenomenológiai adat például ennek a fájdalomnak a

minősítettsége [quale, kválél], de ilyen a piros szín látásának vagy a tea illatának élménye is. Ez tehát az egyik adathalmaz.

Az elmét azonban nemcsak ily módon közelíthetjük meg. A tudomány, az *empirikus* tudomány kihívása éppen az, hogy *harmadik személyű perspektívából*, amikor mi, a tudományos megfigyelők fóruma, az *ő* tudatát tesszük vizsgálatunk tárgyává, mit kezdhetünk az elmével. Mivel a *megfigyelés* nézőpontjából az elme fenomenológiai sajátosságaihoz nem férhetünk hozzá, elvileg olyan adottságokra kell, hogy hagyatkozzunk, amelyek mindannyiunk számára hozzáférhetőek. Csakis tisztán a megfigyelés segítségével nyerhető, harmadik személyű *viselkedési* adatok segítségével dolgozunk. Ez utóbbiak alkotják tehát a másik adathalmazt.

Természetesen, ahogy már említettem, mindez sarkítás. Nem igaz, hogy a tudomány teljesen nélkülözni lenne kénytelen az elme első személyű perspektívájából nyert adatokat. Itt van például a mikroökonómia. Megfelelő mennyiségű adatot begyűjtve arról, hogy hányadik pohár joghurt mennyire ízlett, olyan általánosítható statisztikai adatokhoz jutunk, amelyekből levonhatók bizonyos következtetések. Ilyen esetekben az adatközlő éppen arról tájékoztat, *milyen is* neki az akárhányadik joghurt. Ez azonban *nem* magyarázat, a magyarázathoz azonban kénytelenek vagyunk karanténba zárni a magyarázandót, hiszen nem vonhatjuk be ennek folyamatába. Ha tehát a tudomány az elme *magyarázatára* törekszik, akkor csakis viselkedési adatokra hagyatkozva kell megmagyaráznia, hogy miért érzem a fájdalmat olyannak, amilyen. Már abban az esetben, ha nem kívánja letagadni, vagy figyelmen kívül hagyni a teljes, mindannyiunk által folyamatosan átélt emberi valóságot (hogy *ez milyen is*, azt mindenki maga érzi). És itt kezdődnek a bajok. Lássuk azonban még egyszer a gondolatmenetet, tömören.

A tudomány egy materialista metafizika keretében arra törekszik, hogy magyarázatot adjon az elmére, mint fizikai folyamatok által létrehozott jelenségre, mivel azonban az elme jelenségét fenomenológiai tulajdonságok is jellemzik, a tudománynak így meg kell magyaráznia, miképpen tesz szert az

elme ezekre. Ahhoz pedig, hogy ez teljesüljön, a harmadik személyű perspektívából valamiképp *el kell érnie* az én elmémhez.

Sokan gondolják úgy, hogy ez a vállalkozás nem járhat sikerrel. *Joseph Levine* egyenesen amellet érvel, hogy a tudomány téved, ha úgy véli, magyarázataival képes az első és a harmadik személyű perspektíva közötti szakadékot áthidalni (Levine, 1983). A megfigyelés bázisára alapozott magyarázatok és a fenomenológiai adatok között *magyarázati szakadék* [explanatory gap] tátong, mivel a magyarázó és magyarázott között nem lehet szükségszerű összefüggést kimutatni oly módon, ahogyan a víz állapotváltozásai és molekuláris szerkezete magyarázati kapcsolatba hozható. Semmi szükségszerű nincs abban, hogy *azok* a bizonyos fiziológiai állapotok, amelyek jeleit azon a képernyőn észlelem, miért éppen *ezekkel* a most átélt fenomenológiai sajátosságokkal járnak együtt. Ez pedig annyit tesz, hogy fizikai állapotokkal magyarázni az elmét *episztemológiai* értelemben melléfogás (vö. Tőzsér 2001, 184 skk.).

Nos, *feltéve* és *meg is engedve*, hogy ez az érv konkluzív, s hogy a támadott felfogás égisze alatt dolgozó tudományra különösképp érvényes, teljesen természetesnek tűnik a kérdés: hogyan tovább? Kézenfekvő megoldásnak tűnik például, már ha pozitívan viszonyulunk e problémához, hogy (ahol csak lehetséges), az első személyű perspektíva adatainak bevonásával orvosoljuk a problémát. Sokan valóban ezt a megoldást választják (Varela 1996, Chalmers 1999, vö. még Velmans 1999, Varela és Shear 1999, Metzinger 2000). A tudományos munka így nem magyarázati kísérlet lesz, hanem a korreláció adatolása. Igen ám, csak hogy itt van a *mesterséges intelligencia kutatás*, amely erre nem képes.

Mint a kognitív tudomány integráns része, az elme pszichológiai igényű gépi modellezése immár tiszteletreméltó múltra tekinthet vissza, s ilyenként mindenképpen megérdemli a figyelmet. Sokkal fontosabb azonban az, hogy ami az emberrel foglalkozó tudományokban talán magától értetődő megoldás, vagyis az első személyű perspektíva adatainak bevonása, itt kivitelezhetetlen. Amíg tehát például a pszichológiában vagy az idegtudományban lehetséges olyan komplex számadás készítése, amely egyaránt tartalmazza a megfigyelési

és az introspektív adatokat, addig a mesterséges intelligencia kutatása csakis harmadik személyű adatokat vehet figyelembe, röviden, kizárólag a *viselkedést*. Másként, a kognitív viselkedést mutató gépek tudománya biztosan nem alakítható át egy olyan korreláció-kutatássá, amely az első és a harmadik személyű perspektíva adatait közé nem magyarázati hidat kíván ácsolni, hanem kétoldali adatgyűjtést javasol, illetve filozofálva e korrelációkutatás módszertani buktatóit igyekszik elkerülni (vö. Roy et al. 1999). Tulajdonképpen azt is mondhatnánk tehát, hogy a mesterséges intelligencia kutatás pszichológiai igényű modellezése a kizárólag harmadik személyű adatokra hagyatkozó *tudományoszmény* utolsó letéteményese. S ilyenként megint csak kivételes figyelmet érdemel. Mit tehetünk tehát?

Javaslatom a következő. Próbálkozzunk meg a *konstrukció* fogalmával. Episztemológiai bajt gyógyítsunk episztemológiai orvossággal. Ne az elme magyarázatát várjuk tehát a gépektől, hanem *magukat a gépeket* vegyük olyan konstrukcióknak, amelyek segítségével preferált adatokra hagyatkozva ismerhetjük meg az elmét. Másképp, tekintsünk a gépekre úgy, hogy bennük olyan egymáshoz rendelt fizikai elemeket lássunk, amelyek külön-külön csakis fizikai tulajdonságokkal rendelkeznek, ám megfelelő *fizikai függvény*ként már képesek az elmétől megkövetelt tulajdonságokat mutatni. Jól látható, a russell-i eljárást proponáljuk, csak éppen nem érzetadatokat ajánlva előnyben részesítendő adatokként, hanem a gépekről nyerhető harmadik személyű adatokat. Az talán jól érzékelhető, hogy egy halom különféle kábel és alkatrész külön-külön legalább olyan messze van az intelligenciától, mint az anyag az én érzetadataimtól. Az első fejezetben azonban láttuk, miképpen működhet a konstrukció érzetadatokkal, próbáljuk meg most az eljárást a tudomány által preferált megfigyelhető entitásokkal, nevezetesen egy gép felépítésére alkalmas alkatrészekkel megvalósítani. Ehhez azonban először az kell, hogy teljesüljenek azok a feltételek, amelyeket a második fejezetben fogalmaztunk meg.

(4) Fizikai konstrukció

Ám valóban megvalósítható ez a szemléletváltás? Vajon tényleg eszközölhető a konstrukció eljárása a harmadik és az első személyű perspektíva által meghatározott episztemológiai mezőn? Ezek tesztelhető kérdések: egyszerűen azon áll vagy bukik a dolog, hogy teljesülnek-e a konstrukció feltételei vagy sem. Lássuk tehát, vajon eleget tesznek-e (i)-(iv)-nek azok a körülmények, amelyek között a mesterséges intelligencia kutatása számára a konstrukció módszerét javasoljuk az elme megismerésére. Haladjunk tehát sorjában, minden egyes feltétel kapcsán megvizsgálva, vajon a mesterséges intelligencia kutatásának közegében érvényesíthető-e vagy sem. Lássuk tehát az elsőt.

(i) Adva kell legyen egy megkülönböztetés hipotetikus és kevésbé hipotetikus entitások között.

Belegondolva megállapíthatjuk, hogy mint minden modellezés, az elmével foglalkozó kutatás is legalább két fajta dolgot követel meg: egyrészt a mintát, másrészt azokat az elemeket, amelyekből a modellt összeállítjuk. Nos, az elme a minta, a fizikai építőelemek pedig a játékkockák. Talán nem kell külön bizonygatni, hogy harmadik személyű perspektívából az elme *hipotetikus* entitás, hiszen az, ami adott, a viselkedés. Episztemológiai értelemben mások elméjéhez nem férünk hozzá oly módon, ahogyan testüket és tágabb értelemben vett viselkedésüket megfigyeljük. Másfelől, a fizikai építőkövek, ez a processzor itt, azok a vezetékek ott, nemhogy az elménél kevésbé hipotetikusan létező, hanem épphogy az igazán valóságos, sőt, a preferált módon megismerhető entitások (legalábbis jelen kontextusban az általunk elfoglalt materialista álláspont szerint). Összegezve, az első feltétel teljesül. Lássuk a másodikat.

(ii) Rendelkezésre kell, hogy álljanak bizonyos kijelentések a hipotetikus entitásokról.

Vagyis, az elmére vonatkozó kijelentéseket után kell néznünk. Nos, vegyünk csak *egyetlen* kijelentést, ám legyünk nagyon céltudatosak:

(1) Elmém sajátosan minősített mentális állapotokkal rendelkezik.

Természetesen a fentebb már említett „kválé”-król van szó. Tudjuk, következő lépésben majd olyan tulajdonságokat kell keresnünk, amelyeket meg kell követelnünk a hipotetikus entitástól ahhoz, hogy a (ii) feltétel teljesítéséhez szükséges kijelentések igazak lehessen. Magyarul, olyan tulajdonságokat kellene elősorolnunk, amelyeket megkövetelve a hipotetikus entitástól, az elmétől, a (1) kijelentést igazként értékelhetjük.

Csakhogy már látható is, hogy *baj van*, hiszen az (1) kijelentés akkor lesz igaz, ha elmém állapotai sajátosan minősítettek. Ez azonban rövidzárlat, akárhogy fogalmazzuk is át. Nem fogunk tudni ugyanis olyan megfigyelhető *fizikai* tulajdonságokat előhúzni, amelyek igazzá tehetik (1)-et. Persze, hiszen olyan tulajdonságokra lenne szükségünk, amelyen osztozhatnak a hipotetikus és a kevésbé hipotetikus entitások. Egy ilyen *common property* híján azonban nem tudjuk megkezdeni a fizikai függvény (a gép) építését, lévén, hogy még csak el se tudunk képzelni olyan tulajdonságot, amellyel rendelkezve fizikai entitások rendszere igazzá tehetné (1)-et.

Ne adjuk fel azért ilyen könnyen. Gondoljunk csak bele, talán még jó is lenne, ha a konstrukciós eljárás üzemelne, hiszen így egyszerűen a konstruktőrökre testálhatnánk a tudományos filozofálást. A logikai konstrukció ugyan konceptuális ügynek tűnik, ám nem azért filozofálás, mert fogalmi, hanem azért, mert konstrukció. Márpedig ily módon a fizikai elemekből konstruálók is *filozófusok*. Sőt, megkockáztatom, hogy *ez* lenne a vérbeli materializmus. Nietzsche tehát még csak kalapáccsal filozofált, a fizikai konstrukció módszeré

viszont már intelligens gépekkel is lehetne. Akinek pedig ez a kilátás nem túl vonzó, lássa be, hogy legalább egyértelműen gazdaságos. Nem tehetjük meg, hogy nem járunk igazán utána csak azért, mert nem szimpatikus, még ha talán a saját bőrünket kockáztatjuk is.

Gondoljuk tehát végig még egyszer. Az eddigiek során lassan már kirajzolódni látszottak egy russell-i elveket követő, ám attól tartalmában alapvetően eltérő konstrukció lépései. Amennyiben tehát a logikai konstrukció lelke az a *logikai* függvény, amely ugyanolyan tulajdonságokat mutat, mint az általa helyettesített entitás, úgy a fenti körülmények között alkalmazható konstrukciót *fizikai konstrukciónak* nevezhetnénk, eredményét pedig *fizikai függvénynek*, ha sikerrel járt volna a vállalkozás. E fizikai függvénynek ugyanis, mint megkonstruált, harmadik személyű perspektívabeli adatokkal leírható elemek egymáshoz rendelt rendszerének, *ugyanolyan* tulajdonságokat kellene mutatnia, mint amelyeket a feltételezett entitástól követelünk meg azért, hogy a róla szóló kijelentéseket igazként értékelhessük. S itt volt a gond. Annyi eredménnyel mindenestre mégiscsak járt e kísérlet, hogy megállapíthatjuk, nemcsak magyarázati, hanem *konstrukciós szakadék* is tátong az első és a harmadik személyű perspektíva adatai között. Ám derítsük ki végre, hogy pontosan mi lehet a probléma gyökere.

(5) Konstruktív elmetudomány

Nos, *szerintem* azon múlik a dolog, hogy milyen kijelentéseket tekintünk irányadónak. Ezek miatt kellett ugyanis olyan tulajdonságokat követelnünk az elmétől, amelyekkel a preferált entitások nem rendelkezhetnek. Mi lenne tehát, ha kibúvóként egészen más tulajdonságokat követelnénk meg?

Idézzük csak fel a gépekkel filozofáló konstruktőrök nem igazán filozófiabarát képét, de csakis azért, hogy végleg *megszabaduljunk tőle*. Világos ugyanis, hogy bármiféle konstrukció első lépése mégiscsak az lesz, hogy meghatározzuk, milyen tulajdonságokkal rendelkezik a megkonstruálandó entitás. A

konstruktőrök csak építkeznek, de hogy mit, azt elvileg más fórum, többnyire a *filozófia* mondja meg nekik, akár explicite, akár hallgatólagos intuíciók útján. Nincs kutatási program filozófiai döntések nélkül. Márpedig, ha a konstruktőrök nem csak gépekkel filozofálnak (a fenti szigorú értelemben), hanem például elgondolkodnak és vitatkoznak afelől, hogy milyen tulajdonságokkal is bír az elme, akkor már ugyanazt teszik, mint mi. Egyáltalán nem állunk olyan rosszul tehát: a fizikai konstrukció működhet, marad hely a filozófiának. Menjünk tehát tovább.

Azt már láttuk, hogy a mindennapi tapasztalat intuíciója szerint milyen tulajdonságok illetik az elmét, ha ki-ki első személyű perspektívából nyer adatokat róla. Azt is láthattuk, hogyan párolható le mindez az elmefilozófiában szofisztikált ellenérvvé. Csakhogy az elméről nemcsak a fent bemutatott kétféle úton nyerhetünk ismereteket. A *kontinentális fenomenológia* ugyanis nem introspekció, de nem is empirikus megfigyelés, mégis az emberi elme, az emberi tapasztalás és emberi valóság megismerését tűzi célul maga elé. És nézzük csak meg, milyen kijelentéseket is tesz az elméről. Tekintsük például az alábbiakat:

- (2) Az elme megtestesült.
- (3) A megtestesült elme környezetébe ágyazott.
- (4) A megtestesült és környezetébe ágyazott elme a világban szituált.

Ami döntő jelentőségű és talán már most látható: (2), (3) és (4)-gyel *semmi olyasmit* nem követelünk az elmétől, amelyet egy *pusztán viselkedő* ágenstől ne kérhetnénk számon. Ezek a kijelentések, ha nem is közvetlenül, de a kontinentális fenomenológiából származtathatók, nevezetesen *Maurice Merleau-Ponty*-tól. Hogy miként, az fellelhető (Varela et al. 1991, Dreyfus 1996, Braddock, 2001).

Ha tehát olyan kijelentéseket tekintünk irányadónak az elméről, amelyeket a kontinentális fenomenológia egy markáns iránya tesz róla, akkor remény

mutatkozik arra, hogy teljesüljenek a konstrukció feltételei. Ellenőrizzük azonban, hogy valóban így van-e. Közelítsünk a gépek felől ismét. Az első feltételt korábban már megvizsgáltuk, és (2), (3), illetve (4), íme, rendelkezésünkre áll. Láttuk, (1)-gyel bajba kerültünk. Kérdés, mi történik most. Milyen tulajdonságokat kell tehát megkövetelnünk az elmétől ahhoz, hogy ezeket a kijelentéseket igazként értékelhessük? Testesüljön meg, és megtestesültként ágyazódjon környezetébe, hogy ily módon szituált lehessen a világban. Vegyük tehát a harmadik feltételt.

(iii) A hipotetikus entitások megfigyelhető tulajdonságokkal kell, hogy rendelkezzenek.

Nézzük azt az embert ott. Van teste, s tudjuk, milyen tulajdonságokat kell mutatnia ahhoz, hogy kijelenthessük, fizikai környezete változásaira reagál, s környezete állapotaira visszahat. Azt is tudjuk, miképpen kell viselkednie ahhoz, hogy a mi emberi világunk részévé váljon, tehát megfelelően kiismerje magát, és tanúságot adjon róla, képes eligazodni a világban.

Vegyük észre, ez most már tisztán harmadik személyű perspektíva, de éppen ez a lényeg. A kontinentális fenomenológia kijelentései az elméről, vagyis (2), (3) és (4), olyan tulajdonságokat követelnek meg, amelyek megfigyelhetők. *Nincs* tehát konstrukciós szakadék. Van viszont egy halom *common properties*. S ezzel már teljesült az utolsó feltétel is, hiszen ilyen tulajdonságokat gépek, nevezetesen bizonyos *robotok* is mutatnak.

(iv) Lehetséges kell legyen, hogy a kevésbé hipotetikus entitások is rendelkezhessenek a megfigyelt tulajdonságokkal.

Az elme megismerése fizikai konstrukció segítségével így a következő: építs egy gépet, egy ágenst, amelynek teste környezetébe illeszkedik, s amely szituációba lép a világ eseményeivel (vö. Brooks 1987 és 1991). Ha sikerült, nézd meg,

hogyan épül fel a masina. Azok a fizikai adatok, amelyek így rendelkezésre állnak az adott fizikai függvényben egymáshoz rendelve (ez maga a gép) russell-i értelemben véve *helyettesíthetik* az elmét. Mivel pedig a természettudományban azokat az adatokat preferáljuk, amelyek maradéktalanul hozzáférhetőek, ezekkel helyettesítve a harmadik személyű pozícióból hipotetikusnak tekintett elmét, *megismertük azt*. Legalábbis russell-i értelemben. Vagyis, a tudomány szempontjából nézve javítottunk bizonyos, az elméről tett kijelentések ismeretelméleti minőségén. A konstrukció segítségével ugyanis az elméről tett fenti kijelentések egy olyan új értelmezését kaptuk meg, amely kizárólag a tudományban preferált, harmadik személyű perspektívabeli adatokat tartalmaz.

Talán már jól kivehető, mindezzel egy olyan *konstruktív elmetudomány* körvonalait szeretném itt vizionálni, amelynek feladata nem az, hogy megmagyarázza az elmét, hanem az, hogy ismert elemek segítségével valamely fizikai hordozón megkonstruálja. Ha elfogadjuk a konstrukció elvét, az elme tudományos megismerése nem magyarázat, hanem helyettesítés lesz, mégpedig fizikai konstrukcióval. Ez ugyan korlátozásnak tűnik, de jobb, ha egyszerű szemléletváltásként értjük: *e_x és e_y* a tudomány dolga, nem pedig *a_z*.

A konstruktív elmetudomány tehát kéttagú. Egyfelől működik benne a fenomenológia, amely így nem elszigetelődik a természettudománytól, hanem munkát ad neki, és működik benne, másfelől egy technikai részleg, amely a fenomenológia által rendelkezésre bocsátottak szerint építi az elmét. Ez azt jelenti, hogy a tényleges konstrukcióval a fenomenológusnak nem is kell foglalkoznia (éppen úgy, ahogy a konkrét russell-i konstrukcióval sem foglalkoztunk), a konstruktőrök pedig tulajdonképpen azt csinálhatják, amit mindig is akartak: építkezhetnek, még hozzá anélkül, hogy állandóan azt kelljen hallgatniuk, mi hiányzik gépeikből. S ha a konstrukció sikerül, akkor kijelenthetik, elérték azt, amit *kizárólag* harmadik személyű perspektívából csak elérni lehet.

Ám a csattanó még csak most jön, mert e ponton fordítunk egyet a nézőponton. Idézzük csak fel, mire is volt még alkalmas az érzetadatokból

kiinduló konstrukció? Hát persze, a hipotetikus entitásokról szóló kijelentések *verifikálására*. A (2), (3) és (4) kijelentés a kontinentális fenomenológiából ered. Következésképp, a fizikai konstrukció segítségével harmadik személyű adatokkal verifikálhatunk kontinentális fenomenológiai kijelentéseket, ami pedig nem kis eredmény. Egyrészt azért, mert indirekt érveléssel ezáltal megszabadíthatjuk végre a fenomenológiát attól, hogy introspekcióként félreértelmezzék: az introspekció adatairól tett kijelentések nem igazolhatóak megfigyelési adatokkal, a fenomenológia kijelentései igen, tehát a fenomenológia nem introspekció. Másrészt, mivel a fenomenológia, saját módszeréből adódóan elvileg nem szorul rá, hogy kijelentéseit empirikusan *is* igazolja, igencsak elgondolkodtató helyzetbe kerül akkor, ha kiderül, hogy ez a lehetőség valamiképp mégiscsak fennáll.

Összefoglalás

Lássuk még egyszer az egészet, mert nagy utat jártunk be, éles kanyarokkal. Kísérletünk alap gondolata az volt, hogy kihasználva Russell konstrukciós eljárásának előnyeit, megpróbáljuk a módszert más területen is munkába állítani. Ennek első lépése az eljárás ismertetése volt, második lépése pedig egy olyan feltételrendszer megfogalmazása, amelyeknek teljesülnie kell bármely területen, ahol csak a konstrukciót alkalmazni kívánjuk. Miután felvázoltuk, miben is látjuk az elme és a természettudomány problematikus viszonyának okát, felvetettük, megkísérelhetjük-e a konstrukció alkalmazását e területen. Megvizsgáltuk ennek lehetőségét, nevezetesen azt, vajon teljesülnek-e a konstrukció feltételei az elme első és harmadik személyű perspektívából nyert adatainak területén. Úgy találtuk, hogy nem. Ennek okát azokban a kijelentésekben véltük felfedezni, amelyekkel az elmét az elmefilozófia nyomán vált szokássá illetni. Kibúvóként tehát más kiinduló kijelentésekre tettünk javaslatot, olyanokra még hozzá, amelyeket a kontinentális fenomenológia nyomán fogalmazhatunk meg az elméről, s meglepő módon arra jutottunk, hogy a feltételek így teljesülnek. Ezen felbuzdulva vetettük fel egy konstruktív

elmetudomány gondolatát. Végül pedig, egyértelműen kerítési szándékkal azt is megpendítettük, hogy a fenomenológia és a mesterséges intelligencia kutatás miképpen munkálkodhat oly módon együtt, hogy az mindkét fél számára kölcsönös előnyökkel járjon. S ennél messzebb egyelőre nem is merészkedünk.

Hivatkozások:

BRADDOCK, G.: Beyond Reflection in Naturalized Phenomenology. *Journal of Consciousness Studies: November* (2001)

BROOKS, R. A.: Intelligence without Representation. *Artificial Intelligence: 47* (1987)

BROOKS, R. A.: New Approaches to Robotics. *Science: 253* (1991)

CHALMERS, D.: First-Person Methods in the Science of Consciousness. *Consciousness Bulletin: Fall* (1999)

DREYFUS, H.: The Current Relevance of Merleau-Ponty's Phenomenology of Embodiment. In: Haber, H., Weiss, G. (szerk.): *Perspectives on Embodiment*. New York/London, Routledge (1996)

LEVINE, J.: Materialism and Qualia: The explanatory gap. *Pacific Philosophical Quarterly: 64* (1983)

METZINGER, T.: *Neural Correlates of Consciousness: Empirical and Conceptual Questions*. Cambridge, MIT Press (2000)

ROY, J.-M., PETITOT, J., PACHOUD, B., VARELA, J. F.: Beyond the Gap: An Introduction to Naturalizing Phenomenology. In: J. Petitot, F. J. Varela, B. Pachoud, J.-M. Roy. (szerk.): *Naturalizing Phenomenology. Issues in Contemporary Phenomenology and Cognitive Science*. (1999) Stanford, Stanford University Press

RUSSELL, B.: *The Analysis of Matter*. London (1927)

RUSSELL, B.: On denoting. *Mind: 14* (1905) és uő.: A denotálásról. In: Copi, I. M., Gould, J. A. (szerk.): *Kortárs tanulmányok a logikaelmélet kérdéseiről*. Budapest, Gondolat (1985) 143 – 166.

RUSSELL, B.: Knowledge by Acquaintance and Knowledge by Description. *Proceedings of the Aristotelian Society: 11*. (1910) és uő.: Tudás ismertség révén és tudás leírás révén. In: *Miszticizmus és logika*. Budapest, Helikon (1976) 339 – 376.

RUSSELL, B.: The Relation of Sense-Data to Physics. *Scientia* (1914) 1.szám és uő.: Az érzéki adatok viszonya a fizikához. In: *Miszticizmus és logika*. (1976) Budapest, Helikon 235 – 290.

TÓZSÉR, J.: Négy érv a fizikalizmus ellen. *Játékok és nyelvjátékok*. (2001) Budapest, Kávé kiadó

VARELA, F. J., THOMPSON, E., ROSCH, E.: *The Embodied Mind. Cognitive Science and Human Experience*. Cambridge, MIT Press (1991)

VARELA, F. J.: Neurophenomenology: A Methodological Remedy for the Hard Problem. In: Shear, J. (szerk.): *Special Issues on the Hard Problem, Journal of Consciousness Studies: June* (1996) és uő.: Neurofenomenológia. Ford.: Sutyák Tibor. *Vulgo* (2005/1)

VARELA, F. J., SHEAR, J.: *The View from Within: First Person Methodologies in the Study of Consciousness*. London, Imprint Academic (1999)

VELMANS, M.: *Investigating Phenomenal Consciousness*. Amsterdam, Benjamin (1999)