

A tudatos rendszerek elmélete

Marosán Bence

Bevezetés

A jelen szövegben elme és test viszonyáról a következő hipotézist szeretném megfogalmazni: a tudatosság nem más, mint anyagi rendszerek által végzett információ-feldolgozás egy speciális módja, ahol az információt egy szigorúan természettudományos fogalomként értem; (abban az értelemben, ahogyan azt az információ-elmélet, illetve a kibernetikai klaszrikus szerzői az 1940-es és '50-es években megfogalmazták. Pl. Claude Shannon, 1948, Shannon-Weaver, 1949, Ross Ashby, 1956/1972, Norbert Wiener, 1974). Azt állítom, hogy önfenntartó és önszabályozó biológiai rendszerek a komplexitás egy bizonyos fokán működésük strukturális és funkcionális elveiből fakadóan *szükségszerűen* tudatos (átélt) módon dolgozzák fel a külső és belső környezetükre vonatkozó információkat. A rendszer működési elveiből levezethető, hogy a rendszernek, mégpedig apriori okoknál fogva, szükségszerűen tudattal kell rendelkeznie; nem képzelhető el, hogy ne rendelkezzen tudattal. Minden olyan rendszer, ami egy bizonyítottan tudattal rendelkező rendszerrel funkcionálisan izomorf, tudattal kell, hogy rendelkezzen. Vagyis: ha tudunk olyan mesterséges gépet készíteni, ami tökéletesen úgy működik, mint egy tudatossággal rendelkező biológiai rendszer (funkcionálisan teljes mértékben ekvivalens vele), akkor annak a gépezetnek is mentális tulajdonságokkal és állapotokkal kell bírnia. Mindaz, amit idáig elmondtam, teljesen megfelelni látszik az elmefilozófiában *funkcionalizmus* néven forgalomban lévő irányzatnak. Az alábbiakban jelezni fogom, hogy a standard funkcionalizmushoz képest vannak eltérések az általam megfogalmazott elmélethez képest; noha nincs kifogásom az ellen, hogy azt a funkcionalizmus egy *megszorított* változatának nevezzük. Többek között arról van szó, hogy a funkcionalizmust leginkább azért szokás támadni, mert nem tud magyarázatot adni a kváléra, illetve annak létrejöttére, (Tózsér, 2008: 63). Az alábbiakban körvonalazott elképzelés egyebek mellett pontosan azzal az igénnyel lép föl, hogy a kváléra vagy kváliára, tehát a tudatosan átélt érzéki minőségre, annak keletkezésére és mibenlétére magyarázatot adjon.

Az elmélet egy kulcsfontosságú pontja, hogy elhatárolja egymástól a tudatos és nem-tudatos rendszereket; (illetve, ezzel összefüggésben, annak közelebbi tanulmányozása, hogy milyen indokok alapján tulajdonítunk tudatosságot egy anyagi rendszernek). Ehhez elkerülhetetlenül egy homályzónába kell majd merészkednünk: a növények és a gerincesek közötti állatok birodalmába; illetve választ kell találnunk arra a kérdésre, hogy tulajdoníthatunk-e legalább fenomenális (szenzoros) tudatosságot (tehát érzéki minőséget megjelenítő átélt tapasza-

latot) az olyan alacsonyabb rendű állatoknak, mint a csalánozók (pl. medúza), a férgek, a rovarok, tüskésbőrűek, illetve bizonyos alacsonyabb rendű puhatestűek.

Az alábbi szövegben még nem fogok választ adni a fentebb jelzett kérdésekre, és nem szolgálok még az elmélet teljes kifejtésével; csupán annak bizonyos főbb elemeit szeretném lefektetni. Reményeim szerint ezt a felvetést az elkövetkezendő 5-6 évben fogom tudni szisztematikusan kidolgozni. A jelen írás csak az út első kövét szeretné elhelyezni.

Az írás két nagyobb részre oszlik. Az elsőben a főbb elmefilozófiai álláspontokat szeretném ismertetni, hogy azután az itt bemutatott felvetéshez képest meghatározhassuk a hasonlóságokat és a különbségeket. A második rész magának az elméletnek a részleteivel foglalkozik.

I. Elmefilozófiai elméletek

Öt elméletet kell kiemelnünk a jelen összefüggésben test és elme viszonyát illetően: 1) eliminativizmus, 2) reduktív fizikalizmus, 3) nem-reduktív fizikalizmus, 4) funkcionalizmus, 5) szubsztancia-dualizmus. Lentebb még más elméleteket is érintek (pánpszichizmus, fikcionalizmus), első körben azonban au imént említetteket szeretném kiemelni.

1) Eliminativizmus: nincs semmiféle elme. Csupán anyagi szubsztanciák léteznek, anyagi tulajdonságokkal; még az azonosság-elmélet is hamis (mely szerint a mentális entitások fizikai folyamatokkal, állapotokkal és tulajdonságokkal azonosak); teljességgel illegitim és tudománytalan mentális entitások feltételezése. Épp ezért anyagi folyamatokra, állapotokra és tulajdonságokra *semmilyen* szempontból sem lehet úgy hivatkozni, mint mentálisakra. (Emiatt inkonzisztens az eliminativizmus a reduktív fizikalizmussal). Mentális épp oly kevésbé létezik, ahogy nem léteznek tündérek, boszorkányok, manók vagy Jupiter isten, (Patricia Churchland, 1986, Paul Churchland, 1981/1991).

Ha a jelen elméletnek van ellenfele, akkor az eliminativizmus az. Az alább bemutatott elmélet egyik alapvető feladata az eliminativizmus kiküszöbölése vagy eliminálása. Ezen a helyen nem szeretnék részletesen érvelni az eliminativizmus ellen, ezt megtették már sokan mások, (Tózsér, 2008: 73skk). Itt csak néhány pontot szeretnék megemlíteni. Meggyőződésem szerint a kválé olyan elemi adottság, amely semmilyen logikai bűvészmutatvánnyal vagy elegáns érveléssel sem tüntethető el; hanem végső alapjául szolgál minden elméletalkotásnak. Az eliminativisták elutasítanak azt az ellenérvet, hogy ha nem létezik kválé, akkor talán nyugodtan alávethetnék magukat egy érzéstelenítés nélküli gyökérkezelésnek; mondván, hogy továbbra sem tagadják azt, hogy az organizmus bizonyos neurális állapotok elkerülésére törekszik. Én azonban nagyon is megengednék egy ilyen ellenérvet. Az érzéki minőségek minden eleven tapasztalatban kinyilvánítják létük kétségbevonhatatlanságát. Egy kávé elfogyasztásakor a forróság, az édeskés és sajátosan kávé *íz* kétségbevonhatatlan; amikor kezünkkel a for-

ró tűzhelyhez érünk, a *fájdalom* abszolút és kétségbevonhatatlan adottság, stb. Az eliminativista csak úgy képes fenntartani elméletét, ha elvonatkoztat saját első személyű tapasztalatától. Továbbá az elmélet rendszeres kifejtése során mégis kénytelen újra és újra mentális terminusok használatához folyamodni, ami legalábbis kétségessé teszi az elmélet hitelességét.

A második részben körvonalazott elmélet csak azokkal az elméletekkel kompatibilis, amelyek elismerik a kválé, illetve kvália létezését. Az alább tárgyalt négy elmélet mindegyike ennek a kritériumnak eleget tesz.

2) Reduktív-fizikalizmus, (pl. John Jamieson Smart, 1959/1981, David Lewis, 1966, David Armstrong, 1968, Anthony Quinton, 1973): léteznek mentális entitások, amelyek azonban fizikai folyamatokkal, állapotokkal és tulajdonságokkal azonosak. A mentális a fizikaira redukálódik, pontosan ugyanúgy, ahogyan a tudománytörténetben a víz a H₂O-ra redukálódott. Világosan kell látni az eliminativizmus és a reduktív fizikalizmus közti különbséget: az eliminativista nem a víz és az oxigén alapján világítja meg elme és test viszonyát, hanem az égés flogisztion- és oxigénelmélete alapján. A flogisztionelmélet *nem* redukálódott az oxigénelméletre, hanem *kiküszöbölődött* az utóbbi javára. A kettő ontológiája teljesen más: az eliminativizmusban a mentális entitások halmaza üres halmaz; a reduktív fizikalizmusban egy nagyon is gazdag halmaz, mely azonban a fizikai entitások egy *részalmazát* alkotja. A mentális események, állapotok, folyamatok, stb. *azonosak* a fizikai eseményekkel, állapotokkal, folyamatokkal (neurális állapotokkal és folyamatokkal), de attól még léteznek, és értelmes dolog róluk beszélni. Ahogy Smart fogalmaz: az azonossági tézis azt állítja, „hogy az a valami, melyről az érzet-mondatok beszámolnak, valójában agyfolyamat. Az érzetek nem mások, mint agyfolyamatok”, (Smart, 1959/1981: 163). A reduktív fizikalizmus azt állítja, hogy a vágyak, hiedelmek, érzések, érzetek, stb. neurális állapotok és folyamatok, csupán *más szempontból* tekintve. Hogyan kell ezt érteni? Valahogy így:



A fentebbi képen az egyik szempontból egy fiatal, a másiktól pedig egy idős hölgy portréja válik hozzáférhetővé. Ugyanaz az ábra, két különböző szempontból tekintve. Ugyanilyen módon fogja fel a reduktív fizikalista a mentális és a fizikai azonosságát. *Külső* szempontból az érzet egy neurális állapot, illetve esemény. *Belső* szempontból pedig az átélt tapasztalat. Az érzet, általában a mentális egy anyagi rendszer belső nézőpontját tükrözi; a rendszer „első személyű” hozzáférése saját neurális állapotaihoz. Minden ember a saját első személyű perspektívájában tulajdon neurális állapotainak és eseményeinek amatőr neurológusa.

A reduktív fizikalizmus egyik legfontosabb erőssége, hogy megoldja a *mentális okozás* kérdését; azt a kérdést tehát, hogy hogyan lehetséges mentális és fizikai közti kölcsönhatás. Mivel a mentális okozás nem más, mint anyagi és anyagi közti kölcsönhatás, így a reduktív fizikalizmus egyfelől biztosítja a *szabad akarat* lehetőségét (azt, hogy egy rendszer ténylegesen önvezérlő legyen), másfelől nem kényszerül arra, hogy *természetfölötti magyarázatokhoz* forduljon, tehát biztosítja *az anyagi világ oksági zártságát*.

3) A nem-reduktív fizikalizmus szerint csak anyagi szubsztanciák (testek) léteznek, melyek azonban rendelkeznek nem-anyagi, mentális tulajdonságokkal is rendelkeznek. Ezek a tulajdonságok nem redukálhatók anyagi tulajdonságokra. A kválé vagy kvália pontosan ilyen tulajdonság. A nem-reduktív fizikalista szerint nincs olyan nézőpont, amely felől nézve fizikai és mentális azonosnak bizonyulna; mivel a kettő egészen egyszerűen nem azonos, a kettő között *exkluzív* viszony áll fenn. Az elmélet egyik jelentős képviselője, David Chalmers ezt az ún. „zombi-érvvel” próbálja megvilágítani a két szféra lényegi különbségét, (David Chalmers, 1995/2004). Eszerint elgondolhatunk két teljesen embernek kinéző lényt, amelyek közül az egyik viselkedésében megkülönböztethetetlen bármely másik embertől, ugyanúgy viselkedik, beszél, reagál, mint minden normális ember, eltekintve attól, hogy *nem rendelkezik tudatos*

tapasztalattal. Csupán anyagi folyamatok komplex együttese, amelyhez semmiféle belső, átélt tapasztalat nem tartozik; egy nagyon bonyolult anyagi *automata*. Mint írja: „minden egyes azonosított fizikai folyamat kapcsán megválaszolatlan marad az a kérdés, hogy az illető folyamat miért vezet tapasztalathoz. Bármelyik folyamat esetében fogalmilag koherens módon feltehető, hogy ugyanez a folyamat lejátszódhatna tapasztalat hiányában is”, (David Chalmers, 1995/2004: 23). Ebből Chalmers a két szféra nem-azonosságára, és egymásra való redukálhatatlanságára következtet.

A fizikai szubsztancia és nem-fizikai tulajdonságai közti viszonyt többféle módon is szokták ábrázolni: részben a *szupervenienencia* (ráépülési) relációval (eszerint a mentális tulajdonságok a fizikai tulajdonságokon szupervenienálnak, rájuk épülnek; ha az utóbbi fennáll, akkor az előbbinek is szükségszerűen fenn kell állnia), illetve a *konstitúciós* relációval, (egy agyi állapot úgy konstituál egy mentális állapotot, ahogy egy márványtömb konstituál egy szobrot. A nézet képviselői szerint a konstitúciós reláció a szupervenienenciánál erősebb, az azonossági relációnál viszont gyengébb; vö. Tózsér, 2008: 45).

A nem-reduktív fizikalizmus talán leggyengébb pontja, hogy állandóan fenyegeti az epifenomenalizmus veszélye; az a veszély tehát, hogy a szabad akarat kiiktatásához vezet. Az autonóm cselekvés illúzió, csupán elszenvetői vagyunk a testünkben végbemenő oksági folyamatoknak. Az elmét ebben a leírásban továbbá a *feleslegesség* veszélye terheli; az tehát, hogy igazából nincs is rá szükségünk, hogy megmagyarázzuk a világban végbemenő eseményeket; tudat nélkül teljesen jól el tudunk boldogulni, a tudat semmit sem tesz hozzá a világra vonatkozó magyarázatainkhoz. Harmadik nehézségként (az epifenomenalizmus, a feleslegesség mellett) megemlíthetjük még az *ontológiai pesszimizmus* lehetőségét. Nevezetesen: egy természettudományos jellegű elmélettől szeretjük elvárni azt, hogy a természeti jelenségekre egységes és szakadásmentes magyarázatot nyújtson. A nem-reduktív fizikalizmus azonban, éppen az elme és a test közötti áthidalhatatlan különbség hangsúlyozásával, egy ponton szükségszerűen szakadás áll be: az elme megjelenik a magyarázat egy bizonyos fázisában, mintegy előbukkan a semmiből, és pontosan a testtől való különbsége miatt nem látszik a magyarázati alap, hogy miért is jön létre a tudat, mi teszi szükségszerűvé. Éppen azért, mivel az elme léte a nem-reduktív fizikalista szerint nem szükségszerű, hanem *kontingens*.

A nem-reduktív fizikalizmus képviselői különböző módokon próbálnak megküzdeni az imént említett nehézségekkel; egyik megoldás sem bizonyult minden kétséget eloszlatóan megnyugtatónak.

4) Funkcionalizmus: talán a legnépszerűbb elmélet, mely szerint „a mentális állapotok funkcionális állapotok. Mentális állapotainkat az határozza meg, hogy milyen funkciót vagy

szerepet töltenek be a mentális életünkben (milyen oksági viszonyban állnak környezetünkkel és más mentális állapotainkkal), és nem az, hogy milyen fizikai (neurofiziológiai) állapotokkal azonosak”, (Tózsér, 2008: 56sk). Ami a mentális állapotokat igazán érdekessé és vizsgálatra érdemessé teszi, az nem a mibenlétük vagy a kauzális eredetük, hanem az, hogy milyen *okági szerepet* töltenek be az élőlény (vagy az adott élőlényvel funkcionálisan ekvivalens bármely anyagi rendszer) életében. A funkcionalista szerint az alapvető séma így néz ki: érzékszervi bemenet → belső folyamatok → viselkedéses kimenet. A struktúrát mindig lehet (és kell is) finomítani, de alapvetően az oksági és funkcionális folyamatok az érdekesek, és a mentális mibenlétére vagy születésére vonatkozó kérdéseket zárójelbe is tehetjük. A funkcionalizmus kompatibilis a fizikalizmus különböző változataival.

A funkcionalizmust az általunk kifejtett elmélet szempontjából a következő okok miatt nem fogadhatjuk el minden megszorítás nélkül: I) a funkcionalizmus funkcionalizálja a kválét, vagyis annak csak oksági szerepére kíváncsi. Alább a kválé születésére kívánunk magyarázatot adni; vagyis lényeges szerepet játszik az a kérdés, hogy hogyan születik meg az agyban, illetve az idegrendszerben az érzéki minőség fenomenális tudata; tehát egy – a kauzális genezis értelmében vett – genetikus magyarázatról van szó. II) A funkcionalizmus elvileg bármiféle anyagot alkalmasnak tart, hogy bizonyos funkcionális állapotokat megvalósítson, akár egy kupac kavicsot is, (vö. Jonathan Lowe, 2000: 48-51). Az alábbi (a második részben közölt) elmélet szerint nem minden anyag képes arra, hogy bizonyos funkcionális állapotokat realizáljon. III) A funkcionalizmus a mentális realizációját nem köti komplexitási fokhoz. A tudatos rendszerek elméletének egyik, talán legfontosabb állítása azonban az, hogy az anyagi komplexitás egy bizonyos foka az, ahol megjelenik a tudatosság, és a funkcionális működések komplexitásának köszönhető a tudatosság létrejötte; mely a legalsóbb szinten szenzoros tudat (az érzet; a kválé).

5) Szubsztancia-dualizmus: két szubsztancia van, egy anyagi és egy nem-anyagi, és a kettő kölcsönhatása alkotja az individuum életét. Ez a hagyományos platóni, descartes-i álláspont, valamint a legtöbb vallás álláspontja. Értelemszerűen ez az elképzelés felmondja a fizikai világ oksági zártságának elvét, és ebben az értelemben „természetfölötti” magyarázathoz folyamodik; amennyiben az elmét kívül helyezi a természeti valóságon.

A szubsztancia-dualizmussal nem kívánok foglalkozni. Úgy vélem, az inkább a teológia, mint a természettudományos orientációjú filozófia területére tartozik, és ebben a formában nem is kívánok konfrontálódni vele. A fentebb körvonalazott elméletek közül csupán eggyel kényszerülök konfrontációra: az eliminatív materializmussal. Az eliminativizmus, úgy gondolom, úgy véli megoldani test és elme viszonyának problémáját, hogy nem-létezőnek nyil-

vánítja a kérdést; felmondja egyik legalapvetőbb (ha nem *a* legalapvetőbb) intuíciónkat, nevezetesen a mentális létezésébe vetett mindennapi hitünket, s ezzel meggyőződésem szerint több problémát generál, mint amennyit megoldani vél.

*

Daniel Dennett álláspontja

A fentebbi elméletekhez képest meglehetősen furcsa helyet foglal el Dennett. Dennettet gyakran fikcionalistának vagy instrumentalistának szokták tartani, (pl. Tózsér, 2008: 75, Ravenscroft, 2005: 72-75). Ez azt jelenti, hogy szerinte sem létezik elme vagy mentális, de ebből nem következik az, hogy a mentális terminusokat is teljesen ki kellene küszöbölnünk a beszédhasználatunkból. Dennett szerint az „intencionális stratégia” a legjobb eszköz, hogy bejósoljuk valamely, a környezetével aktív viszonyt kialakító, teleologikusan viselkedő anyagi rendszer viselkedését, (Dennett, 1998a). Dennett mindenesetre, mint mondja, „meglehetősen boldogtalan”, hogy őt fikcionalistának vagy instrumentalistának nevezik, ezt ő magáról kifejezetten tagadja. Ő magát a mentális állapotok (és a népi pszichológia) vonatkozásában „realistának” mondja, igaz, nem nagy R-rel, hanem kis r-rel, (Dennett, 1998a: 81).

Amit Dennetről határozottan kijelenthetünk, hogy ő materialista és funkcionalista. Ezt annál is inkább megtehetjük, mivel ezt ő maga ismeri el, (Dennett, 2004: 38). Hogy a materializmus melyik válfaját képviseli, ez úgy tűnik, számára is kérdéses. (Egy konferencián megkérdezték tőle, hogy ő minek tartja magát, és a válaszából az derült ki, hogy ezt ő sem döntötte el teljesen). Ami a dolgot messzemenően bizonytalanná teszi az az, hogy Dennett számos olyan kijelentést tesz, ami a fikcionalizmus, illetőleg az eliminativizmus álláspontjával konzisztens. Így például Dennettet kvália-eliminativistának szokták mondani, mivel Dennett szerint a kváléra, illetve kváliára vonatkozó alapvető hiteink inkonzisztensek egymással, (Ambrus Gergely, 2008: 314skk, Dennett, 1988, 1991). Dennett ezen nézetével kapcsolatban csak annyit mondhatunk, hogy ha a kváliára vonatkozó hiedelmeink inkonzisztensek egymással, akkor a hibát nem a kvália létében, hanem inkább a rá vonatkozó hiedelmeinkben kell keresnünk. Az alább nyújtott elméletben a kváliát nem egy elmélet által posztulált vagy feltételezett, bizonytalan státusú entitásnak tartjuk, hanem a legalapvetőbb adottságnak, amely minden elméletalkotás lehetőség-feltételeként szolgál, és amelynek kétségbevonása csak a szavak szintjén lehetséges; szigorú vizsgálat alatt azonban merő agyrémnek bizonyul.

Dennettet mindazonáltal úgy értelmezem, hogy neki a leginkább a kvália *privát* jellegével van baja; azzal az elképzeléssel tehát, hogy kinek-kinek a maga fenomenális tudatosságához szigorúan első személyű, tehát *privát* és *kizárólagos* hozzáférése van. Úgy gondolom, hogy

Dennett valójában a kvália privát természete ellen szeretne érvelni, (valamint az ellen, hogy a kváliára anti-materialista érveket lehessen alapozni), és nem akarja letagadni az érzéki minőségek létét. Mindannyian tudjuk (tudhatjuk), hogy milyen egy éles, hasogató fájdalmat átélni, hogy milyen a kávé íze, hogy milyen a langyos, tavaszi szellő érzése – ezek nem privát élmények, hanem mindenki számára hozzáférhető, ilyen értelemben nagyon is nyilvános tapasztalatok. Ezen alapul egyáltalán az *empátia* lehetősége. Nekem tehát számos szöveghely alapján úgy tűnik, hogy Dennett inkább a tudatos tapasztalatok *nyilvánossága* mellett szeretne érvelni.

II. A tudatos rendszerek elmélete

Az „intencionális rendszer” kifejezést először Charles Taylor használja (majd az ő nyomán Dennett); az övével szoros összefüggésben használjuk mi is a „tudatos rendszer” megfogalmazást; (Charles Taylor, 1964: 62). Anyagi rendszerekről van szó, amelyek közül egyesek rendelkeznek tudattal, és a tudat a viselkedésüket lényegileg meghatározó viszonyulási mód; más rendszerek viszont nem rendelkeznek tudattal. „Viszonyulási módnak” mondtuk a tudatot: olyan viszonyulási módról van szó nevezetesen, amelynek révén egy anyagi rendszer külső, illetve belső környezetéhez viszonyul. Az intencionalitás, a viszonyulási mód itt a *reprezentáció*, valamint az átélés fogalmaival áll közvetlen összefüggésben. Egyes rendszerek képesek arra, hogy átéléssel (tudatossággal, fenomenális tudatossággal) társuló reprezentációkat alkossanak külső, illetve belső környezetükről, és ez a reprezentáció alapvető módon határozza meg, illetve befolyásolja a rendszer viselkedését. A reprezentációk sorában a legalapvetőbb az átélt, szenzoros tudat. Ez az érzéki minőség fenomenális tudata, a *kválé* vagy *kvália*.

A tudatos rendszerek elméletével továbbra is követni szeretném az általam eddig is folytatott *fenomenológiai* irányultságú utat; csupán az eddigiekhez képest más elemeket és szerzőket is szeretnék bevonni vizsgálódásaimba. A tudatfilozófiai irányultságú (Edmund Husserl) tudat-tárgy, illetve tudat-világ korreláció helyett most a test-környezet korrelációra szeretnék összpontosítani; nem azért, hogy az előbbit az utóbbival helyettesítsem, hanem inkább azért, hogy az előbbit (a tudat-tárgy/tudat-világ korrelációt) az utóbbi révén világítsam meg, illetve tegyem hozzáférhetővé. Amikor test-környezet korrelációról beszélek, akkor a fenomenológiai antropológiai tradíció olyan szerzőire kell gondolni, mint Helmuth Plessner, Arnold Gehlen, Max Scheler, Maurice Merleau-Ponty, illetve az alapvetően természettudós, biológus, a biokibernetika alapítóatyja, Jakob von Uexküll.

Az elmélet kiindulópontja az, hogy a tudat zárójelezése mellett, pusztán az anyagi folyamatok elemzésével kell eljutnunk a leírásnak arra a szintjére, amikor már nem tartható fenn ez a zárójel, hanem az anyagi rendszer viselkedéséhez szükségszerűen fel kell oldanunk. Az elmélet alapvetően három lehetőséget szeretne kiküszöbölni: 1) a „filozófiai zombi” lehetősé-

gét; vagyis a tudat megjelenése nem lehet pusztán esetleges, hanem apriori szükségszerűséggel kell, hogy következzen a rendszer működésének formális elveiből; 2) a pánpszichizmust; azt az elméletet tehát, amely minden anyagi rendszernek és egységnek tudatosságot tulajdonít, ami egyáltalán információ-feldolgozást végez. Ez az elmélet tehát azonosítja a tudatosságot az információ-feldolgozás minden fajtájával (William Seager, 1995/2004), amit viszont én kontra-intuitívnek tartok. Végül 3) elutasítom az „erős mesterséges intelligencia”-felfogást, amely szerint minden rendszer tudatos, amely a környezettel visszacsatolásszerű viszonyban áll; így pl. egy termosztát is. (Ezt „termosztát-paradoxonként” fogom emlegetni).

II.1. A tudatosság kritériumai

Milyen érvek alapján tulajdonítunk egy másik embernek tudatosságot? Mindenekelőtt: az illető többé-kevésbé hasonlít hozzánk. Rendelkezőnk *belső* tapasztalattal saját élményeinkről, tudjuk, milyen az fájdalmat, örömet, undort, kétségbeesést, stb. átélni, és rendelkezünk annak külső tapasztalatával, hogy a *mi* esetünkben az ilyen élményekhez milyen *külső*, viselkedéses reakciók szoktak társulni. Ha valaki hasonlóan reagál, mint mi, arról feltételezzük, hogy hozzánk hasonló élményeket él át. Továbbá a másik *kommunikáció* útján tudósíthat bennünket belső állapotairól. („Fáj a fejem”, „Most nagyon fáradt vagyok”, „Szeretem a frissen főzött kávé”). Van ezen kívül egy *közvetett* természettudományos hozzáférési mód is a tudatossághoz: az emberek hozzávetőlegesen azonos *fiziológiai* felépítéssel rendelkeznek. Az emberi szervezet, mindenekelőtt az *idegrendszer*, más és más élményekre különböző módokon reagál. Ezek (a funkcionális fizikai reakciók) különböző eljárásokkal és műszerekkel egyre pontosabban megfigyelhetőek, és a megfigyelt alany (a legtöbb esetben) kommunikáció révén tájékoztatni tudja a megfigyelést végző tudóst, hogy mi megy végbe benne. *Okunk* van feltételezni, hogy az idegrendszeri működések és a tudati működések között szoros korreláció áll fenn. *Okkal* feltételezzük továbbá, hogy mindenütt, ahol izomorf funkcionális bázissal (tehát strukturálisan lényegileg ekvivalens, megfigyelhető testi működésekkel), izomorf viselkedéses struktúrákkal (a kommunikációt is beleértve) találkozunk, ott tudatossággal van dolgunk; egyszerűen azért, mert a mi esetünkben, ahol első személyű hozzáférésünk van saját mentális állapotainkhoz és eseményeinkhez, ez a helyzet, és a korrelációról közvetlen tapasztalattal rendelkezünk. Az következik ebből, hogy semmilyen olyan lénytől nincs jogunk elvitatni a tudatosságot, amelyik (vagy aki) velünk funkcionálisan ekvivalens bázissal rendelkezik (egy normálisan működő idegrendszerrel), és ugyanígy viselkedik, ugyanígy reagál, mint mi. Az elvi kételkedés lehetősége persze nyitva áll, egy ilyen elvi kétely azonban soha nem lesz több, mint merő *absztrakció*.

A kérdés az, hogy nem-emberi lényekkel, állatokkal mi a helyzet. Ezek analóg módon működnek az emberi szervezettel; vagyis az emberi és állati szervezetek között különböző mértékű funkcionális analógiát találunk. (A kibernetika szóhasználatában az állati rendszerek az emberi rendszerekkel *homomorfak*; ahol a homomorfia a *hasonlóságot* jelöli, szemben az izomorfíával, ami a kibernetikában „a legszigorúbb értelemben vett” strukturális *azonosságot* jelöli. Ross Ashby, 1956/1972: 125sk). A viselkedéses reakciók is részben analógok az embernek a megfelelő élményekre, tapasztalatokra adott viselkedéses reakcióival. Minél távolabbi az evolúciós létrán való elhelyezkedés, minél jobban különböznek tőlünk szervezettségben és működésben, annál távolabbi az analógia az ember és az egyéb élőlények között. Ebből az a további kérdés adódik, hogy hol húzzuk meg a határvonalat a tudatos és nem-tudatos rendszerek között. Kell-e egyáltalán ilyen határvonalat húznunk? Dennett a következőt írja az állati tudatossággal kapcsolatban: „a tudat nem egy fekete-fehér, mindent-vagy-semmit típusú jelenség, szemben azzal, ahogyan azt gyakran feltételezik”, (Dennett, 1995: 691). Könnyen előfordulhat, hogy Dennettnak igaza van. De akkor mégis: hogyan kell elképzelnünk a fizikai és a mentális között „félúton” lévő entitást? Milyen lenne egy félig fizikai, félig mentális képzet, intencionalitás vagy reprezentáció? Mi volna egy ilyen „hibrid” entitásnak az értelme? Ennek a kérdésnek a megválaszolásához a tudattalan problémája fontos támpontot adhat; ezzel összefüggésben a „szubliminális” vagy „szubkritikus intencionalitás” fogalmát fogom használni.

Sokan kétségbe vonják, hogy anyagi jelenségek elemzése révén eljuthatnánk valami szellemihez vagy tudatoshoz, illetve, hogy anyagi működésekből bármilyen módon létrejöhetnének szellemi entitások, vagy anyagi működésekből ilyen entitásokra legitim módon tudnánk következtetni. (Pl. Fichténél a materializmussal szembeni ellenérv: anyagi mechanizmus csakis anyagi mechanizmust eredményezhet, és soha nem következhet belőle valami nem-anyagi. Fichte, 1797/1981: 37sk). A kérdés az, hogy ez valóban így van-e? Tényleg nem következtethetünk anyagi működésekből a nem-anyagira?

A fizika ismeri a „*fázisátalakulás*” fogalmát: azt a fogalmat tehát, amely arra utal, hogy bizonyos körülmények között az anyag radikálisan másképpen kezd el viselkedni, mint korábban. Például: ha a jeget melegíteni kezdjük, akkor ezt nem tehetjük meg tetszőleges hőfokig, a szilárd halmazállapot megtartása mellett. Tehát a -50 Celsius-fokos jeget hevítjük fel, akkor nem kapunk +10, +20, +30, stb. fokos jeget. 0 foknál a jég megolvad, és vízzé válik. Ugyanígy: ha melegítjük a folyékony vizet, akkor ez a halmazállapot sem marad meg mindig: nincsen +120, +130, +500, stb. fokos víz, hanem gőzzé válik. Az anyag átalakul valami radikálisan mássá, mint korábban volt. Ez a fázisátalakulás. A most megfogalmazott feltevés ér-

telmében a tudatosság születése nem más, mint az anyagi rendszer információ-feldolgozásában bekövetkező fázisátalakulás. Ehhez a tudattalan, illetve a szubliminális vagy szubkritikus intencionalitás fogalma úgy kapcsolódik, hogy már a „fázisátalakulás” határának közelében (de még a határ alatt) sem teljesen úgy zajlik az információ-feldolgozás, mint mélyen a küszöb alatt.

Milyen élőlényeknek tulajdonítanak a természettudományos kutatásokban szubjektív értelemben vett tudatot; (a legalsóbb szinten: érzékszervi érzékelést, szenzoros tudatot, tehát szubjektíve átélt érzetet)? A gerinchúrosoknak, gerinceseknek, és attól felfelé. Itt nincs vita a természettudósok között. (Nos, az eliminativista biológusok erről mást mondanának, de itt most a döntő többségről beszélünk). Abban sincs (említésre méltó) vita, hogy az egysejtűek vagy a növények nem érzékelnek a szó szigorú értelmében véve. (A „növényi érzékenység”, „plant sensitivity” elképzelése az ezotéria körébe tartozik. A „növényi neurológia”, „plant neurology” hívei közül többen szintén hajlanak rá, hogy pl. a növények fototropizmusát a szó szigorú értelmében vett fényérzékelésként értelmezzék; de itt megint csak egy nagyon szűk kisebbségről van szó). Az a mód, ahogyan egysejtűek, alacsonyabb rendű többsejtűek, növények, gombák, stb. dolgozzák fel a külvilág ingereit, illetve szabályozzák saját állapotaikat, teljesen anyagi jellegű mechanizmus. Ahogyan egy zöldszemes ostoros (Euglena), illetve egy növény érzékeli a fényt, ontológiailag egy kategóriába tartozik azzal, ahogyan a fényérzékeny lemez „reagál” a fényre, vagy ahogyan a fotocella indít be egy gépi mechanizmust fény hatására. Egyik esetben sem beszélhetünk szenzoros tudatosságról (érzetről). (Itt a pánpszichizmus hívei mondanának mást; de a pánpszichizmus ebben a történetben egy olyan álláspont, amelyet hosszú távon éppenséggel ki szeretnénk küszöbölni). A nagy kérdés – ebben az összefüggésben a legnagyobb – az, hogy a gerinchúrosok, illetve a gerincesek (akik a funkcionális bázis, és a viselkedéses struktúrák egyes fontos elemeit tekintve velünk, emberekkel szoros rokonságot mutatnak), valamint a növények közötti állatvilágban kik rendelkeznek, és kik nem rendelkeznek tudatossággal. A felsorolás kedvéért, olyan lények tartoznak ide (a fejlettség, a szervezettség sorrendjében), mint 1) szivacsok, 2) csalánozók, 3) férgek, 4) puhatestűek, 5) ízeltlábúak és 6) tüskésbőrűek. A kérdés az, hogy ezek közül kiknek tulajdoníthatunk tudatot kellő tudományos alappal, és kiktől kell ezt megtagadnunk.

Úgy tűnik, a döntő különbséget az élővilágban az idegrendszer megjelenése jelenti. Elméletem értelmében az idegrendszer nem önmagában, mint ilyen és ilyen anyagi sajátosságokkal bíró szerveződés érdekes, hanem speciálisan mint egy *információ-feldolgozó rendszer vagy hálózat*. (Az idegrendszer és az idegi hálózat a biológiában nem egyjelentésű fogalmak. Az idegrendszer fokozottabb központosulást és hierarchizálódást feltételez; az információ-

feldolgozás központjait és sajátosan elkülönülő szintjeit. Az élő szervezet az ilyen központosulással és hierarchizálódással jóval energiatakarékosabban és hatékonyabban tudja feldolgozni az információkat. A hálózat alacsony fokú központosulást jelent, vagy egyenesen mindenféle központosulás hiányát, ahol a hálózat elemei között szimmetrikus vagy csak kevésbé aszimmetrikus viszony van. Amikor az alábbiakban nem szinonim értelemben használom a „rendszer” és „hálózat” fogalmait, akkor kifejezetten erre a különbségtételre utalok). A növényi szervezeteknél nincsen specializált, állandó, szilárd struktúrával rendelkező információfeldolgozó hálózat. (Az ún. „növényi neurológia” képviselői éppenséggel a növényeknél megtalálható hálózatos jellegű információ-fogadásra és -feldolgozásra hivatkoznak. Vö. pl. Stefano Mancuso, 2010. Ezek azonban még nem specializált feldolgozó-rendszerek, illetve -hálózatok, hanem még csupán pseudo-hálózatok, amelyek evolúciós módon előremutatnak a magasabb szerveződési szintek felé, de még nem képesek olyan módon fogadni, kezelni és „kiértékelni” az információkat, mint egy tényleges idegrendszer, vagy akár csak egy idegi hálózat. Jelenlegi biológiai ismereteinkben kataklizmaszerű változásokra volna szükség, hogy a növények esetében ugyanolyan információ-feldolgozási módokról tudjunk beszélni, mint az idegrendszerrel rendelkező állatok esetében. Azonban még maguk a „növényi neurológia” hívei is elismerik, hogy itt, a növények esetében „neurológiáról” vagy „neurobiológiáról” csak *analógiás* vagy *metaforikus* értelemben van szó. Vö. Anthony Trewavas, 2007, Eric Brenner et al., 2007). A növények esetében tehát sem a funkcionális alap, sem a megfelelő viselkedéses reakció nem elég komplex ahhoz, hogy tudatos információ-feldolgozást feltételezzünk. Az állatok esetében sem rendelkezik azonban mindegyik állat a megfelelő fejlettségi funkcionális apparátussal, valamint – ezzel összefüggésben – nem képes ilyen jellegű reakciókat produkálni, amelyek alapján joggal következtethetnénk tudatosságra.

A többsejtű állati szerveződéseknél az evolúciós létra alján a szivacsokat (porifera) találjuk, ahol még nincsenek idegsejtek, sem a szó szigorú értelmében (a magasabb komplexitási szinteken azonosítható), kifejlett érzékszervek; jóllehet, egyes szivacsoknál vannak olyan evolúciós kezdemények, amelyek az idegsejtek és az érzékszervek felé mutatnak, (vö. Danielle A. Ludemann et al., 2014). Ez azonban még csak az alap, amelyre ráépülhettek, illetve amelyből kifejlődhettek a fejlődés magasabb szintjein az információ-feldolgozás komplexebb struktúrái. Idegsejtek, és egyelőre még kevésbé specializált idegi hálózat az evolúciós fejlődés során először a csalánozóknál (cnidaria) jelenik meg, amelyek (akik?) közé többek között számos medúzafaj is tartozik. Ahogy haladunk felfelé a fejlettségi létrán a funkcionális bázis és a viselkedéses reakciók alapján úgy lesz egyre bizonyosabban megállapítható szenzoros tudat (a fenomenális tudatosság, átélt érzet) megléte. Így az ízeltlábúak (Arthropoda) törzsébe tar-

tozó sziklavízi garnélarák (*Palaemon elegans*, rockpool prawn) esetében Robert Elwood és társai teljes bizonyossággal megállapítani vélték a fájdalomérzetek tényleges jelenlétét, (Robert Elwood és Mirjam Appel, 2009, Robert Elwood és Barry Magee, 2013). Ugyanakkor a fejlődés nem töretlen; vannak bizonyos „visszaesések”, amelyek arra utalnak, hogy a mozgás és az érzékelés képessége evolúciósan összekapcsolódik. Mindenütt, ahol helyváltoztató mozgás van egy többsejtű élőlény esetében, ott az idegrendszer és az érzékszervek is „robbanás-szerű” fejlődésnek indulnak; ahol viszont nincs mozgás, ott az idegrendszer is visszafejlődik, az érzékszervek elcsökevényesednek. Ez a helyzet például a tüskésbőrűek (Echinodermata) törzsének (újszájúak – Notoneuralia – törzscsoportja) néhány bizonyos fajtájának esetében. Evolúciósan a fejlődés előrehaladottabb stádiumában létrejött törzsről, illetve fajtákról van szó. A tüskésbőrűek általában sugaras szimmetriájúak, viszont lárváik háromosztátú, illetve kétosztátú szimmetriát mutatnak, jelezve a rokonságot az evolúciós elődökkel, (Vogel-Angermann, 1999: 192, 579). Viszont egyes fajták helytűlők, nem változtatják a helyüket, (tengeri liliumok, Crinoidea, egyes tengeri csillagok, Asteroidea). A környezetből feléjük sodródó szerves anyagokból táplálkoznak, szabad külső megtermékenyítéssel szaporodnak; (vagyis a tengerbe ürítik a hímivarsejteket és a petesejteket, ezek „félúton” találkoznak, megtermékenyítik egymást, és így jönnek létre az utódok). Ezeknél a helytűlő állatoknál az érzékszervek, valamint az ideg-rendszer szenzoros alrendszerei is visszafejlődött állapotban vannak, (i.m. 136).

Nem mindenről tudunk azonban, ami az idegrendszerben lejátszódik. Nem érzékeljük például az autonóm (vegetatív) idegrendszerben lezajló folyamatokat, amelyek a szervezet belső környezetét felügyelik és vezérlik. A jelen tanulmányban speciálisan az idegrendszernek azokról a funkcióiról van szó, amelyek létrehozzák a szenzoros tudatosság eseményeit (érzetek), illetve az utóbbiak alapján a tudatosság magasabb rendű formáit (perceptuális tudat, érzelmek, szimbolikus tudat – nyelv – és kognitív tudat – gondolkodás). Az érzet az idegrendszer szenzoros alrend-szereiben jön létre. (Legalsóbb szinten az elsőrendű feladat az érzet kialakulásának mint egy információ-feldolgozási folyamatnak a magyarázata; illetve az érzetképződés elhatárolása a pusztán anyagi jellegű ingerfeldolgozástól, ami például a növényeknél és az egysejtűeknél megy végbe). Ami bizonyosan megállapítható, hogy az érzetképződéshez kellőképp differenciálódott érzékszervre van szükség, egy megfelelően komplex idegrendszerre vagy idegi hálózatra, valamint a kettő közötti teljes kommunikációs folyamatra. Az érzékszerv, az idegpálya és a lokális adatfeldolgozó központ között végbemenő teljes információ-feldolgozási folyamat felel meg egy szenzoros, fenomenálisan is átélt eseménynek; vagyis a szó szigorú értelmében vett érzékelési aktusnak.

Az a kérdés, hogy az alacsonyabb rendű állatok is érzékelnek-e, azért is vethető fel értelmesen, mivel a nem-gerinces lények idegrendszere lényegi különbséget mutat a gerincesekéhez (illetve gerinchúrosokéhoz) képest. A csalánozóktól kezdve a tüskésbőrűekig bezárólag (1) csalánozók, 2) férgek, 3) ízeltlábúak, 4) puhatestűek, 5) tüskésbőrűek) az idegrendszer fejlődése csupán a dúcosodásig jut el. Nincs agy. Az agy először a gerinceseknél jelenik meg. (Van egy átmeneti forma: a gerinchúrosok, amelyeknél azonban a csőidegrendszer már a fejlettebb csontos gerincoszlop kialakulása felé mutat. Vö. Vogel-Angermann, 1999: 102skk, 138skk). A kérdés az, hogy érzetek képződnek-e az idegdúcokban is? Érzékel-e például egy méh vagy légy? Amikor azt olvassuk, hogy a méhek képesek *lát*ni a fény ultraibolya-tartományát is (i.m. 351), akkor ezt átvitt vagy szó szerinti értelemben kell-e értenünk? Descartes például úgy gondolta, hogy egyedül az emberek rendelkeznek a tudat képességével, és az állatok csak afféle nagyon bonyolult gépezetek; mondhatni feltuningolt óragépek. Vajon például a méh esetében az volna a helyzet, hogy az egész élőlény a beérkező fényrészecskékkel csupán (ontológiailag) ugyanolyan típusú információ-feldolgozást végez, mint egy növény? A méh szeme ugyanolyan volna, mint egy komplex robotot vezérlő kamera, ahol a robot ténylegesen (az átélt tudatosság értelmében) nem lát? Csupán egy nagyon bonyolult sejt-automatával volna dolgunk a méh (valamint az agy kifejlődéséig megtalálható összes, ideg-rendszerrel egyébként rendelkező állat) esetében? A jelenlegi természettudományos kutatások nem ezt mutatják. Mindaz alapján, amit ma tudunk az állatokról, meglehetősen „sovén” pozíciónak mondható az az álláspont, amely a tudatosságot csak az aggyal rendelkező élőlények számára tartja fenn, és az agyszerv fejlettségi szintjét el nem érő összes állattól szeretné a tudatosságot megtagadni. Utaltam Robert Elwood és társai kutatásaira, akik éveken át tartó kísérletekkel meglehetősen bizonyossággal kimutatták, hogy pl. garnélarákok egész biztosan élnek át fájdalmat. Márpedig az ő esetükben még csupán idegdúccal rendelkező, ízeltlábú élőlényekről van szó, (vö. pl. i.m. 130sk).

A puhatestűek esetében néhány élőlény már gyakorlatilag eléri a felsőbb szinteken található állatok idegrendszeri fejlettségét. Egyes lábasfejűek (Cephalopoda) esetében az idegrendszer az agydúccig fejlődik, ami csaknem eléri a gerincesek szintjén jelenlévő agy komplexitási fokát, (vö. i.m. 134). Ez a helyzet pl. a nyolckarú polipok esetében is, (Octopodiformes). Ezeknél az élőlényeknél az érzékszervek is rendkívül magas módon fejlettek. Az idegi hálózattól tehát, amelyet először a csalánozóknál találunk meg, az aggyal rendelkező gerincesekig valóban folyamatos az átmenet. A további kérdés az, hogy az idegrendszer fejlődésén belül hol jelöljük meg a határpontot a tudatos és a nem-tudatos információ-feldolgozás között? A válaszom az, hogy a kellőképp differenciált érzékszerv és (a hozzárendelt, szintén kellőképp

differenciált) idegrendszer megléte önmagában elég ahhoz, hogy tudatosságot állapítsunk meg az állatvilág egy bizonyos szintjétől felfele. Ez a komplexitás pedig nem kell, hogy szükségszerűen elérje az agy komplexitásának szintjét. Azt állítom, hogy érzet képződhet az idegdúcban is (a hozzákapcsolt megfelelő érzékszervvel végzett érzékelési folyamat során). (Már egyes medúzák szeme, valamint a vele összekapcsolt szenzoros alrendszer összetettsége is elég magas ahhoz, hogy feltételezzük, az a medúza a szó szoros értelmében látja a fényt, nem pedig csupán arról van szó, hogy fényrészecskék pusztán anyagi folyamatokat váltanak ki az élőlénybe. Vö. i.m. 350/D). *Feltételezésem szerint érzékszerv-idegpálya-idegdúc/agydúc alrendszerben zajló folyamat az alacsonyabb rendű állatoknál funkcionálisan izomorf vagy ekvivalens a magasabb rendű állatok érzékszerv-idegpálya-agy rendszerében lezajló érzékelési folyamattal. Sejtésem szerint ez a funkcionális izomorfia kimutatható lesz a megfelelő rendszerek formális modellezése révén. Azt az állítást szeretném tehát megfogalmazni, hogy a kutatás egy későbbi pontján, az idegrendszerrel rendelkező élőlényekben végbemenő információ-feldolgozási folyamatok formális modellezése révén apriori bizonyossággal kimutatható lesz, hogy amennyiben a magasabb rendű idegrendszerrel rendelkező állatok tudatossággal bírnak, úgy az alacsonyabb idegrendszerű állatok is szükségszerűen kell, hogy rendelkezzenek bizonyos egyszerűbb érzetekkel.*

II.2. A hipotézis

Ezután megfogalmazhatjuk a tanulmány legfontosabb feltevését. Ennek értelmében pusztán önszabályozó és önfenntartó anyagi rendszerek által végzett információ-feldolgozás megfigyelésével és leírásával szükségszerűen el kell érkeznünk egy bizonyos ponthoz, amelyen túl azt kell feltételeznünk, hogy a rendszer radikálisan másképpen dolgozza fel az önmagára és környezetére vonatkozó információkat, mint az alsóbb komplexitási szinteken. Egyszerűen nem tudjuk másképp koherensen leírni a rendszer működését, csak úgy, ha feltételezzük, hogy az információ-feldolgozás módja alapvetően más lett, mint korábban. Az információ fogalmát a biológián belül kétféle értelemben is szokták használni: 1) a szaporodás során a DNS-ben tárolt genetikai információra vonatkoztatva, 2) valamint arra, ahogyan egy biológiai rendszer, egy élőlény értelmezi a külvilág és a belső környezet jelzéseit, (vö. pl. Luciano Floridi, 2010: 73skk). Itt értelemszerűen mindenekelőtt az utóbbiról van szó. Amikor anyagi hatásokat egy önfenntartó, önszabályozó rendszer jelzéseként fogad, mégpedig olyan jelzéseként, amelyhez működését valamilyen módon igazítani kell, akkor ott információról beszélünk. Az információ lényegileg egységekben értelmezhető, attól függően, hogy a szóban forgó anyagi hatás elérte-e a rendszer ingerküszöbét, vagy sem, (0 vagy 1 értékű információ-egység).

A tudat valamely kielégítő elméletének egyszerre kell apriorinak lennie (azaz olyan filozófiai jellegű magyarázatnak, amely a tapasztalat esetleges jellegétől független, abszolút szükségszerűséggel bíró érvényességű elveket követ), másfelől támaszkodnia kell a természettudományos kutatás legfrissebb eredményeire is; vagyis nem lehet pusztán „karosszék-filozófia” (Armchair-philosophy). Úgy vélem, az itt körvonalazott elképzelés mindkét, általam alapvetőnek tartott kritériumnak megfelel. Egyfelől arról van szó, hogy élő rendszerek kibernetikai, tisztán formális modelljét elkészítsük. Mivel a működés tisztán formális modelljeiről van szó, ezek a leírások apriorinak tekinthetők. Másfelől ezeket a formális modelleket annak alapján készítjük el, amit jelenleg tudunk a vizsgált élőlényekről. Ugyanakkor a tudományos későbbi haladása folytán mégsem képzelhető el az, hogy a tudatosság ilyen elméletének radikálisan át kellene alakulnia; hanem csupán arról lehet szó, hogy egyik vagy másik élőlényt, amelyik pontosan a határzónában helyezkedik el, át kell helyeznünk a tudatos rendszerek csoportjából a nem-tudatos rendszerek csoportjába, és viszont. Az önszabályozó kibernetikai rendszerekben végbemenő információs folyamatok kellőképp kifinomult leírásának választ kell tudni adnia arra a kérdésre, hogy mely rendszerek tekinthetők tudatosnak, és melyek nem.

A helyzet érdekességét az adja, hogy már egy egysejtű is elképesztően bonyolult tud lenni. A papucsállatkában (Paramecium) több szinten zajló információ-feldolgozási folyamatot állapítottak meg, lokális központokkal, ami analóg a többsejtű állatok idegrendszerében zajló információ-feldolgozással. (Vogel-Angermann, 1999: 71. „Az együttműködés [a sejtszervecskék és folyamatok közötti együttműködés – M.B.P.] magas foka az úszásban, a táplálkozásban és a védekezésben *koordinációs mechanizmust* sejtet, amely funkcionálisan a magasabb rendű állatok idegrendszeréhez hasonlít”). És ez csak egy egysejtű. A növényeknél az információ-feldolgozás még bonyolultabb módon megy végbe. Ilyen szervezetek formális működési modelljét elkészíteni meglehetősen összetett feladat, mégis, meggyőződésem szerint az ilyen formális elemzések adják a kezünkbe a kulcsot annak megértéséhez, hogy az egyik szervezet miért szükségszerűen tudatos, és a másik nem.

Azt kell látni, hogy az idegsejt az idegrendszerben nem csupán egy ilyen és ilyen anyagi sajátosságokkal bíró képződmény, mely hozzájárul az élőlény normális működéséhez és önfenntartásához. Nem. *Az idegsejt az idegrendszerben vagy idegi hálózatban nem egy idegsejt, hanem egy információ-feldolgozó egység.* Az idegrendszer nem (vagy nem pusztán) idegrendszer, hanem információ-feldolgozó rendszer, amely információ-feldolgozó egységekből és részrendszerekből épül fel. A (szenzoros, perceptuális és ennél magasabb rendű) idegrendszer egy olyan specializált információ-feldolgozó apparátus, amelynek segítségével az élőlény begyűjti és feldolgozza az önmagára és környezetére vonatkozó információkat. Az idegrend-

szer speciálisan abban a tekintetben érdekes számunkra, amennyiben ilyen információ-be-gyűjtő és –feldolgozó apparátusként működik. Ez az apparátus az információ-feldolgozás egy meghatározott komplexitási fokával bír. Állításunk szerint minden olyan anyagi rendszer, amely megvalósítja az információ-feldolgozás adott komplexitású fokát, és funkcionálisan ek-
vivalens vagy izomorf a megfelelő fejlettségű idegrendszerrel, tudatossággal bír. (A legalsóbb szinten szenzoros tudatossággal, tehát az érzéki minőségek – a kvália – tudatos átélésével, majd a magasabb tudatos működési módokkal). Amit itt előadtam, az leginkább a *szerep-funkcionalizmus*hoz (role-functionalism) áll közel; de megfelelő értelmezés esetén éppígy konzisztens lehet a realizáció-funkcionalizmussal, (ld. ehhez: Tózsér, 2008: 61skk, Crane-Farkas, 2004: 605). Éppígy nem döntök itt arról sem, hogy vajon a tudatosság, legalsóbb szinten a kválé, vajon az információ-feldolgozás egy adott komplexitási fokú és típusú meg-valósulásával azonos-e (ez felelne meg a reduktív fizikalizmusnak), vagy az információ-fel-dolgozás egy speciális módjára szükségszerűen ráépülő, nem-fizikai tulajdonságról van-e szó, (tehát nem-reduktív fizikalizmusról). Ezekben a kérdésekben a kutatásoknak ezen a pontján még nem döntünk. Ami egészen bizonyos, hogy egy természettudományos elkötelezettségű, a *megszorított funkcionalizmust* elfogadó elméletről van szó (megszorított funkcionalizmusról beszélünk, amennyiben feltevésem szerint a megfelelő funkciók megvalósításához szükséges egy minimális komplexitási szint, valamint nem minden anyag képes arra, hogy a tudatossá-got megteremtő vagy lehetővé tevő funkciókat ellássa, biztosítsa). Ez az elmélet továbbá *emer-gentista*, amennyiben azt mondom, hogy egy bizonyos komplexitási szint fölött van tudat, alatta pedig nincs; és azt, hogy az egyik szinten miért nincs, a másikon pedig miért van, a komplexitás fokából tudjuk levezetni. (Látni fogjuk alább, hogy ez az emergentizmus sem zár ki minden átmenetet, amennyiben megengedem a tudattalan tudat, illetőleg a szubliminális vagy szubkritikus intencionalitás fogalmát).

Ezek után összefoglalhatjuk az elmélet legfontosabb sajátosságait és elveit. Először is ki akarunk zárni három elméleti lehetőséget: 1) a pánpszichizmust, (William Seager, 2004), amely szerint mindenhol van tudatosság, ahol csak információ-feldolgozás zajlik. Erre, ezen a ponton, röviden annyit kell mondanunk, hogy a pánpszichizmus *bizonyíthatatlan*; (és hozzáte-hetném, hogy kontra-intuitív, de egy érv kontra-intuitivitása soha sehol nem számíthat komoly érvnek). 2) A termosztát-paradoxont. A termosztát-paradoxon értelmében minden visszacsatolással rendelkező, komplex anyagi rendszer tudatossággal kell, hogy bírjon. (Ld. pl. „erős mesterséges intelligencia”-elmélet. Vö. Roger Penrose, 2011). A termosztát-paradoxonnal vagy az erős mesterséges intelligencia-elmélettel kapcsolatban ugyanazt kell mondanunk, mint a pánpszichizmusra: nem bizonyítható, illetve (feltevésem szerint) hamissága a jelen elmélet

egy kellőképp előrehaladott pontján bizonyítható lesz, amennyiben megmutathatóvá válik, hogy egy alacsonyabb szervezettségű rendszer vagy gépezet által végrehajtott információ-feldolgozás komplexitási foka alatta marad a tudatossághoz szükséges kritikus szintnek. 3) A filozófiai zombi. Meg kell mutatni a filozófiai zombi elvi lehetetlenségét, vagyis ki kell zárni azt a lehetőséget, mely szerint létezhet olyan anyagi rendszer, mely működésében maradéktalanul megegyezik egy tudatossággal bíró anyagi rendszerrel, és mégsem bír tudattal. Azt kell világosan kimutatni, hogy az információ-feldolgozás egy bizonyos komplexitási fokát és szerveződési módját elérve nem képzelhető el az, hogy az adott rendszer funkcionális állapotait ne tudatosan valósítsa meg. A helyzet ugyanaz volna, mintha egy filozófus azt bizonygatná, hogy elvileg elképzelhető, hogy holnap nem kel fel a Nap – annak rögzítésével, hogy a Föld körbefordul tengelye körül, és közben a Nap sem tűnik el valamiféle kozmikus katasztrófa következtében. Az ilyen elképzelést hangoztató filozófusnak azt kellene mondanunk – legalábbis én így hiszem –, hogy vagy a vonatkozó természeti jelenségeket nem értette meg, vagy az általa használt fogalmakat, vagy egyiket sem. Az, hogy „másnap felkel a Nap” pontosan azt jelenti, hogy a Föld körülfordul tengelye körül, a Nap megvilágítja a Föld addig homályban lévő oldalát; és ha a Nap nem tűnt el közben, és nem került égitest a Nap és a Föld közé, mely eltakarhatná a Nap sugarait (pl. napfogyatkozáskor), akkor teljesen szükségszerű, hogy a Nap felkeljen.

Az itt előadott elmélet három alapvető elvet fogalmaz meg. 1) *A minimális komplexitás elve*. Az információ-feldolgozás „fázisátalakulásának” az információ-feldolgozás komplexitása és szerveződési módja egy bizonyos fokához kell, hogy kötődjön. Az alatt nincs tudatos információ-feldolgozás. Ezzel szándékozom kizárni a pánpszichizmus lehetőségét, illetve az erős mesterséges intelligencia-elméletét, vagyis közelebbről a „termosztát-paradoxont”. 2) *A rendszerbe integráltság (vagy rendszerbe foglaltság) elve*. Az a feltevés, hogy egy adott információ-feldolgozó alrendszer csak a teljes rendszeren belül funkcionális, azon kívül nem; még akkor sem, ha a megfelelő apparátust „életben tartják”, vagy a rendszeren kívül működtetik. Ezzel a kiegészítéssel – az elsőhöz hasonlóan – az a célom, hogy kizárjak néhány kontra-intuitív következményt. Ezt a kiegészítést – csak úgy, mint az első – azonban csak a projekt előrehaladásának későbbi fokán fogom tudni bizonyítani szigorúan is. A rendszerbe foglaltság elve arra utal, hogy ha egy rendszerből valamilyen módon eltávolítjuk, „kioperáljuk” az egyik tudatos funkciót megvalósító részrendszert, akkor az önmagában nem lesz funkcionális; még akkor sem, ha biztosítják azokat a körülményeket, hogy a részrendszer ne haljon el. A maga speciális funkcióját egy részrendszer csak a teljes rendszeren belül tudja ellátni. Konkrét példával: ha egy különleges műtét során valakiből kioperálják a szemét, és a hozzá-

tartozó teljes vizuális kéregrendszert, a teljes látókérget, úgy hogy az illető túl is élje a műtétet, akkor nem az lesz az eredmény, hogy lesz két szubjektum, akik közül az egyik immár nem lát, a másik pedig csak lát, minden egyéb mentális képesség nélkül. Az eredmény tisztán az lesz, hogy lesz egy szubjektum, aki sajnálatos módon megvakul; de az eltávolított szem, a hozzá tartozó vizuális alrendszerrel együtt – még akkor is, ha egy olyan speciális oldatba helyezik, ami életben tartja a teljes szervi komplexumot – nem valósít meg vizuális funkciót, még akkor sem, ha sikeresen ingerlik. Egy ilyen ingerelt szervi komplexum esetében nincs egy szubjektum, aki látna, csupán anyagi folyamatok összességéről van szó. Nem lehet szó tehát arról, hogy egy szubjektum úgymond osztódással szaporodásnak indulna, ha egy főrendszerből a tudatos funkciókat megvalósító alrendszereket nekiállnánk sorban kioperálni. Ez azonban egyelőre *sejtés*, melynek bizonyítása (adott esetben: cáfolata) csak később válik lehetségessé. 3) *Diszkontinuitási tézis*. A jelen elmélet első, ideiglenes megfogalmazásánál úgy véltem, hogy ugrás van a tudatos és nem-tudatos rendszerek között. Nem tartottam értelmes elgondolásnak a fizikai és a mentális között félúton lebegő valamiféle hibrid entitás koncepcióját. A diszkontinuitási tézissel kifejezetten vitába kívántam bocsátkozni Dennettel, aki szerint a tudat „fekete-fehér, mindent vagy semmit-típusú jelenség”, (Dennett, 1995). Később, Vámos Tibor javaslatainak hatására, ebben a tézisben némiképp visszakoztam (lásd alább), és harmadikként a *megszorított diszkontinuitás elvéről* szeretnék beszélni. A megszorított diszkontinuitás elve azt mondja ki, hogy bizonyos módon a tudatos és nem-tudatos információfeldolgozás között az átmenet valóban nem-folytonos, ugrásszerű módon valósul meg; azonban ez nem azt jelenti, hogy a tudatnak semmilyen módon nem lenne *előformája*; illetve, hogy a tudatos információfeldolgozás határának közelében már másképp zajlik az információfeldolgozás, mint e határtól távol. Ezt fejezi ki a szubliminális vagy szubkritikus intencionalitás fogalma.

II.3. Szubliminális vagy szubkritikus intencionalitás

Ahogy a fázisátalakulás közelében lévő anyag is meglehetősen átmeneti módon viselkedik (pl. forrásban lévő víz), ugyanígy a tudatosság határában lévő információfeldolgozás sem teljesen ugyanolyan módon megy végbe, mint korábban, illetve e határtól távol.

Az érzékelés élettanában különbséget kell tennünk ingerküszöb és érzetküszöb között. Van, hogy néhány receptorsejtben a beérkező ingerek átlépik a szükséges küszöböt, a receptorsejt gerjesztett állapotba kerül, jelzéseket küld további sejteknek és adatfeldolgozóközpontoknak, de mégsem elegendő mennyiségű jelzést, hogy egy teljes információs, illetve kommunikációs ciklus végbemenjen az érzékszerv és az adatfeldolgozó-központ között, ami szükséges volna ahhoz, hogy az adott élőlény érzetet éljen át. Azonban a tudatosság körébe be

nem került jelzések következtében az organizmus teste már másképp viszonyul a környezethez, ahhoz viszonyítva, mintha a külvilág ingerei nem érték volna el az ingerküszöböt. Konkrét példával: a nőtény selyemlepke (*Bombyx mori*) ivari csalogatóanyagát a hím már 11 kilométer távolságból is képes érzékelni. A „mámorszint” kiváltásához azonban bizonyos mennyiségű illatanyagra van szükség. A selyemlepke receptoros sejtjei viszont már egyetlen molekulára is képesek reagálni. A „mámorszint” a konkrét keresési mechanizmus csak akkor indul be, ha a hím selyemlepke 35000 receptorsejtje közül legalább 200 egyidejűleg ingerelve van. Ez mondhatni az érzetküszöb. Ez alatt, amikor ennél kevesebb számú receptorsejt van ingerelve, azonban az állat magatartása bizonyos módon, érzékelhetően megváltozik. Az élőlény teste sajátos készenléti állapotba kerül, az élőlény mintegy érzékeli, hogy nem minden ugyanúgy van, mint korábban, de még nem „tudja”, hogy pontosan mi változott. (Vö. Vogel-Angermann, 1999: 349, utolsó bekezdés a jobboldalon). Az ingerküszöb fölött és az érzetküszöb alatt az élőlény magatartása némileg módosult, a testi reakciók némileg másképpen folynak le, mint az ingerlés előtti állapotban. *Ezt az állapotot nevezem a szubliminális vagy szubkritikus intencionalitás állapotának.*

A test készenléti állapotba kerül (az adott inger típusának megfelelően), a szubliminális szint alatt maradó jelzések meghatározzák a konkrét érzet kialakításának módját is. A pusztán testi jelzések és folyamatok, a testnek a környezettel kapcsolatban kialakított tudat alatti „mikro-intencionalitásai” meghatározzák a tudatosság konkrét módjait és irányait. A tudattalan, szubliminális testi intencionalitás, valamint a tudatossá vált intencionalitás együttműködése és viszonya megköveteli a „tudattalan tudat”, a tudatosság körén kívül maradó jelzések, illetve (egy magasabb szinten) a tudat fókuszán kívül maradó érzetek problémájának a tisztázását, és teljes körű feltárását. A pusztán fizikai lét és a mentális között a tudattalan teremt kapcsolatot és átmenetet. A tudattalan problémája mélyen az állatvilágba nyúlik vissza, és *nem* a tudattal egyidőben születik meg, hanem egy még mélyebb szinten, ahol már előkészíti a tudatosságra való átmenetet. (Arról a koncepcióról van szó nevezetesen, hogy már azok az állatok is, akiknek idegrendszere nem alkalmas arra, hogy érzetet valósítsanak meg, viszont az idegrendszeri működés, a funkcionalitás, nem is marad sokkal a kritikus szint alatt, már kimutathatóan másképpen viselkednek, mint a kritikus szinttől távol eső felépítésű állatok).

Konklúzió

A fentebb előadott elmélet még nem kész, befejezett teória. Ha van kifejezett ellenfele, amellyel le szeretne számolni, az az eliminativizmus, amellyel azonban osztozik egy sajátos vonásban: az eliminativizmushoz hasonlóan *projekt-filozófiáról* van szó; olyan elméleti vállalkozásról tehát, amely sejtései egy részének tudományos, logikailag kényszerítő erejűnek

szánt bizonyítását csak később, a vállalkozás kellőképp előrehaladott fokán tudja nyújtani, vagy legalábbis azzal az ígérettel lép föl, hogy a kutatás egy bizonyos fokán már képes lesz ígéretei beváltására.

A konkrét hipotézis tehát az, hogy ha kellőképpen kiismertük az anyagi rendszerek által végzett, saját külső és belső környezetükre vonatkozó információ begyűjtésének és feldolgozásának speciális módjait, akkor 1) képesek leszünk elhatárolni egymástól a tudatos és tudatlan rendszereket, továbbá 2) megértjük, hogy az egyik rendszer miért rendelkezik szükség-szerűen és apriori okokból kifolyólag tudattal, és a másik miért nem, és végül 3) megértjük azt is, hogy a szubliminális vagy szubkritikus intencionalitás, tehát a tudatosság küszöbe alatt végzett, ehhez a küszöbhez viszont közel álló információ-feldolgozási folyamatok, hogyan készítik elő a tudatosság megjelenését, illetve a tudat alatti (de az ingerküszöböt elért) jelzések hogyan határozzák meg a tudatosság realizációjának egyes konkrét módjait.

Úgy gondolom, hogy a fentiekben körvonalazott elmélet eleget tesz mind annak a követelménynek, hogy a természettudományos világszemlélettel összhangban lévő magyarázatot nyújtson a tudat keletkezésére és mibenlétére (vagyis tiszteletben tartja a fizikai világ oksági zárttságának elvét), mind a szabad akaratot meg tudja magyarázni (vagyis elismeri a mentális okozás létezését). Az organizmus az általa végzett információ-feldolgozás komplexitási fokának függvényében rendelkezik egyre nagyobb és nagyobb mozgástérrel. Itt tényleges autonómiáról, a szóban forgó individuum tényleges önvezérléséről és önrendelkezéséről kell beszél-nünk, ahol a természeti törvények nem írják elő pontról-pontra, hogy az individuumnak ho-gyan kell cselekednie, vagyis nem magyarázzák cselekvésének minden egyes mozzanatát és részletét, hanem csupán e cselekvés *mozgásterét* jelölik ki. A tudatos információ-feldolgozás olyan egymásra épülő szintjeiről van szó, amelyek az emberben érik el a legmagasabb fokot, ahol már a biológiai alapoktól ténylegesen függetlenedő viselkedésformák kialakításának a lehetőségéről van szó, (vö. Dennett, 1998b: 400-412, Frans de Waal, 2006: 37, 167-174, Ma-rosán György, 2010). A tudatosság, valamint a szabad cselekvés magyarázatához a kulcsot az élő rendszerek által végzett információ-feldolgozás evolúciója adja.

Irodalomjegyzék

- Ambrus Gergely (2008): „Fenomenális tudatosság”, in Ambrus Gergely, Demeter Tamás, Forrai Gábor és Tózsér János (szerk.), *Elmefilozófia*, Budapest, L'Harmattan, 2008: 297-318.
- Armstrong, David M. (1968): *A Materialist Theory of the Mind*, London: Routledge.
- Ashby, Ross (1972): *Bevezetés a kibernetikába*, Budapest, Akadémiai kiadó.
- Brenner, E.; Stahlberg, R.; Mancuso, S.; Baluska, F. és Van Volkenburgh, E. (2007): "Response to Alpi et al.: plant neurobiology: the gain is more than the name". *Trends in Plant Sci-*

- ence 12 (7): 285–286.
- Chalmers, David J. (2004): „Szemközt a tudat problémájával”, in *Vulgo* 2004/2: 14-35.
- Churchland, Paul M. (1981/1991): "Eliminative Materialism and the Propositional Attitudes", in David M. Rosenthal (szerk.) *The Nature of Mind*, New York, Oxford: Oxford University Press, pp. 601-612.
- Churchland, Patricia S. (1986): *Neurophilosophy*, Cambridge, Mass.: MIT Press/Bradford Books.
- Crane, Tim és Farkas, Katalin (2004): "Introduction IX", in uő. *Metaphysics, A Guide and Anthology*, Oxford: Oxford University Press, pp. 601-610.
- De Waal, Frans (2006): *Primates and Philosophers (How morality evolved)*, Stephen Macedo-Josiah Ober (szerk.), Princeton University Press.
- Dennett, Daniel (1988): "Quining Qualia". In Marcel, A. Bisiach, E. (szerk.): *Consciousness in Modern Science*. Oxford, Oxford University Press.
- Dennett, Daniel (1991): *Consciousness Explained*, Indianapolis, Little-Brown.
- Dennett, Daniel (1995): "Animal Consciousness: What Matters and Why", in *Social Research*, Volume 62, Number 3, pp. 691-710.
- Dennett, Daniel (1998a): *Az intencionalitás filozófiája*, Budapest, Osiris-Gond kiadó.
- Dennett, Daniel (1998b): *Darwin veszélyes ideája*, Budapest, Typotex kiadó.
- Dennett, Daniel (2004): „Háttal a tudat problémájának”, in *Vulgo* 2004/2: 36-38.
- Elwood, Robert W. és Appel, Mirjam (2009): "Pain experience in hermit crabs?", in *Animal Behaviour* 77 (5): 1243–1246.
- Elwood, Robert W. és Magee, Barry (2013): "Shock avoidance by discrimination learning in the shore crab (*Carcinus maenas*) is consistent with a key criterion for pain", *Journal of Experimental Biology*. doi:10.1242/jeb.072041.
- Fichte, Johann Gottlieb (1972): „Második bevezetés a tudománytanba”, in *Válogatott filozófiai tanulmányok*, Budapest, Gondolat kiadó.
- Floridi, Luciano (2010): *Information. A Very Short Introduction*, Oxford, Oxford University Press.
- Lewis, David (1966): "An Argument for the Identity Theory", *Journal of Philosophy*, 63, pp. 17-25.
- Lowe, E. Jonathan (2000): *An Introduction to the Philosophy of Mind*, Cambridge: Cambridge University Press.
- Ludeman, Danielle A.; Farrar, Nathan; Riesgo, Ana; Paps, Jordi és Leys, Sally P. (2014): „Evolutionary origins of sensation in metazoans: functional evidence for a new sensory

- organ in sponges”, in *BMC Evolutionary Biology* 2014, 14:3 doi:10.1186/1471-2148-14-3.
- Mancuso, Stefano (2010): „The roots of plant intelligence”.
- Hyperlink: http://www.ted.com/talks/stefano_mancuso_the_roots_of_plant_intelligence
- Marosán György (2010): „Mi az ember? – Egy evolúciós szintézis körvonalai”, in *Magyar Tudomány Ünnepe*, konferencia-kötet, Budapest, Budapesti Gazdasági Főiskola, 2010:1-14.
- Penrose, Roger (2011): *A császár új elméje*, Budapest, Akadémiai kiadó.
- Quinton, Anthony (1973): *The Nature of Things*, London: Routledge.
- Ravenscroft, Ian (2005): *Philosophy of Mind, A Beginner's Guide*, Oxford: Oxford University Press.
- Seager, William (2004): „Tudat, információ és pánpszichizmus”, in *Vulgo* 2004/2: 47-62.
- Shannon, Claude (1948): "A Mathematical Theory of Communication", *Bell System Technical Journal*, 27, pp. 379–423 & 623–656, July & October, 1948.
- Shannon, Claude és Weaver, Warren (1949): *The Mathematical Theory of Communication*, University of Illinois Press.
- Smart, John Jamieson (1959/1981): "Sensations and Brain Processes", in V. C. Chappell (szerk.), *The Philosophy of Mind*, New York: Dover Publications, Inc., pp. 160-172.
- Taylor, Charles (1964): *The Explanation of Behaviour*, London, Routledge and Keagan Paul.
- Tózsér János (2008): „Általános bevezetés: a test-lélek probléma”, in Ambrus Gergely, Demeter Tamás, Forrai Gábor és Tózsér János (szerk.), *Elmefilozófia*, Budapest, L'Harmattan, 2008: 9-85.
- Trewavas, Anthony (2007): "Response to Alpi et al.: Plant neurobiology--all metaphors have value.". *Trends in Plant Science* 12 (6): 231–233
- Vogel, Günter és Angermann, Hartmut (1999): *Biológia*, Budapest, Springer Hungarica kiadó.
- Wiener, Norbert (1974): *Válogatott tanulmányok*, Budapest, Gondolat kiadó.