

III. Ember és szingularitás?

A kerekasztal-beszélgetés résztvevői: Kiss Károly, Sütő Zoltán

Moderátor: Major Gyöngyi

A világ ősokat ég és föld anyjának hívhatjuk.
Aki anyját megtalálta, az az ő gyermekének tudja magát.
Aki az ő gyermekének tudja magát, hűségesen kitart anyja mellett.
(Lao Ce)

Gyöngyi: Első közelítésben megkérdezném, hogy mi a szingularitás, mi az a nagy változás, amit kiválthat a mesterséges intelligencia? Később majd reflektálnék, hogy szerintem miért függ össze mindez a nőiséggel, és miért válik létkérdéssé a női létezés manifesztálódása az MI korában. Lényegi kérdéssé válhat ugyanis, hogy megteremthető-e az egyensúly – és ha igen, akkor hogyan – az ember és a gépi világ között. Nem tudom, hogy egyébként szerethető maradhat-e ez az új világ.
– Tanár úr?

Károly: Kedves Hölgyek és néhány Kolléga úr! Szeretettel köszöntök mindenkit. Kiss Károly vagyok, képzettségemnél fogva közgazdász, és annak is a világgazdasági, később az ökológiai, környezetpolitikai vonatkozásait műveltem. Néhány éve kezdtem el érdeklődni az informatika iránt, meg más, nem közvetlenül közgazdasági tárgyak iránt. Így belebonyolódtam a mesterséges intelligenciába. Nem gondoltam volna, hogy jön majd egy olyan pillanat, amikor itt hölgyek előtt számot kell adnom arról, hogy a téma, amivel most foglalkozom, az őket vagy általában a női princípiumot valahogy veszélyeztetheti. Tehát megdöbbsz, hogy milyen vonatkozásban merül fel ez a dolog. Mindenesetre majd itt lassan, Zolival együtt megpróbálunk a dolgok mélyére ásni, legalábbis első megközelítésben.

Szerintem ennek a témának nincs ilyen vonatkozása. Szerintem a MI kapcsán nem jön szóba az intelligencia női vagy férfi változata. Egy kicsit konkrétan. A mesterséges intelligencia fejlesztésének technikailag a legújabb módja az úgynevezett *deep learning*, ami azt jelenti, hogy nagy, lehetőleg minél nagyobb mintákon tréningeztetjük a mesterséges intelligenciát, ami a mintafelismerő képességével összefüggéseket és hasonlóságokat keres és talál. Például az erkölcsi programozásnak ez a módja. Tehát nem direkt módon programozzák be az algoritmusokba, hogy ilyennek kell lenni, meg olyanak kell lenni, vagy ez csúnya dolog, az pedig nem elfogadható, hanem tréningeztetik mondjuk a napi sajtón, meg a világirodalmon, meg amit közkeletűen olvasnak az emberek, és rábízzák, hogy milyen következtetéseket von le belőle. Például le fogja vonni azt a következtetést, hogy mondjuk nem szép dolog a feketétet kisebbrendű embereknek tartani. Tehát a nagy mintából következtetést von le a gépi intelligencia. Ezen belül a gép bizonyosan levon következtetéseket a nőiségre és a női gondolkodásra vonatkozóan is, hogy igen, vannak dolgok, amiben a nők másfajta. Ami azonban a mesterséges intelligencia létrehozói iránt, tehát az algoritmusok megalkotói iránt, nem hallottam róla, hogy ők ilyen szempontok szerint jártak volna el. A mesterséges intelligencia különböző fajtáinak megalkotásánál tudomásom szerint nem merül fel a nők és férfiak eltérő gondolkodásának figyelembevétele. Én legalábbis mindent, ami fontos erről a témáról és aktuális, elolvastam, de ezzel a szemponttal nem találkoztam.

Gyöngyi: Ez persze nem azt jelenti, hogy nincs is. Legfeljebb, hogy még nincs benne a diskurzustérben.

Károly: A mesterséges intelligencia gazdag irodalmában, a műszaki vonatkozásokban szerintem ez nem vetődött fel, olvasmányaim során nem talákoztam ilyesmivel.

Gyöngyi: Akkor kiindulópontnak máris mondok egy példát, és utána átadom a szót Zoltánnak. Tehát úgy vetődik fel ez a probléma, hogy már létezik az a robot, aki/ami (?) állampolgárságot kapott. Sophiára gondolok, arra a női humanoid formájú robotra, akinek Szaúd-Arábiában van állampolgársága. Tehát az, ami eddig kizárólag az emberi lényhez kötve merült fel, azzal már egy gép is rendelkezik. Ezzel egy spirál elején vagyunk, aminek azonban a következményrendszere beláthatatlan messzeségeket sejtet. A gép egyszerűen belép a létezésünkbe, és olyan tulajdonságokat tudhat magáénak, ami eredetileg az ember sajátja. A női princípium úgy jön képbe, hogy a nő, aki szül – tudniillik egy életet szül, aki majd állampolgár lesz, nos, ennek a nőnek vajon teljesen mindegy az, hogy az ő gyermeke mellett most ott van egy gép, és az is egy állampolgár? Ugyanúgy, mint az ő gyermeke. Így vetődik fel a női princípium, a női olvasata a problémának. Az pedig csak hab a tortán, hogy ez a robot-állampolgár egy nő. El lehet gondolkodni.

Zoltán: Ahhoz, hogy megértsük a kihívást, messziről kell indítani. A magam részéről az egyik legjelentősebb problémának tartom, hogy a mesterséges intelligenciáról és a technológiákról folyó kommunikációban kézenfekvően ezek alkotói – meglehetősen egyoldalú, technicista szemléletű nyilatkozók viszik a főszerepet. Természetesen itt bizonyos vonatkozásokban jelentős képességekkel rendelkező emberekről van szó, akik nem kis tehetséggel művelnek egy tudományá felnövő szakmát – még hozzá olyan szakmát, ami a kívülállók részére már követhetetlen. Ugyanakkor a technológiának ezen a megszállottjain nem sok nyomot hagyott az elmúlt kétezer év kultúrája, azokkal a gyötrő kérdésekkel együtt, amelyek egyáltalán nem újkeletűek (lásd pl. a keresztény erkölcsiséget vagy akár Madách gondolatiságát), s amelyek ilyen módon a gondolkozásukat nem zavarják.

Az intelligens gép előállításának feladata természetesen az „ember” modellezését jelenti, s számukra nem feltűnő, hogy az automatizmusokról, a vezérlésekről, bizonyos logikai műveletekről vagy azok szimulációiról ugyan elmélyült tudással rendelkeznek, viszont kevésbé ismerik az „embert”, akit modellezni szándékoznak. Pontosabban, esetükben – szemléleti értelemben – nem jelenik meg az ember „gondolkodó-problémamegoldó gépezetként” való értelmezésénél több. Két jelentős téma vetődik itt fel, melyekre az informatika jelenleg súlyosan leegyszerűsítő választ ad. Ezek:

1. Mi a tudat?
2. Mi az identitás?

Nem mellel, csak egy megjegyzés erejéig térek ki arra, hogy jelenleg az érintettek között elvi viták folynak arról is, hogy szervesen anyaghoz egyáltalán kapcsolódhat-e tudat. Azonban, ha ezt a problémát engedékenyen el is kerüljük, a tudat meghatározását, azaz, hogy miféle képességek meglétét nevezzük tudatnak, s ilyen módon miféle képességeket gondolunk társítani a mesterséges intelligenciához, nincs mód mellőzni.

Általában tudatos cselekvésnek nevezzük azt ha egy konkrét célhoz valaki/valami ismeretileg feltárt lépések szándékos alkalmazásával jut el. Az ember esetében ez azért történik, mert születésétől fogva folyamatosan érzékeli önmagát, van egy belső önérzékelése, ön-érzete, valamilyennek érzékeli saját állapotát, és amennyiben ez éppen megfelelhet neki, akkor megőrizni törekszik, vagy ha nem felel meg, készítése támad állapota kívánt irányú változtatására. Ehhez felhasználja meglévő tapasztalatait, és többnyire sikeres műveletekkel eljut a céljához.

Ezzel szemben egy gépnek nincs önmagáról belső érzete. Szerelhetünk rá érzékelőket, s programozhatjuk a hőmérséklet mérési adatainak függvényében például arra, hogy kerülje a 90 °C-nál melegebb helyeket, mivel az árthat az áramköreinek. Azonban, ha ez adott esetben hiba folytán nem valósul meg, nincs semmiféle következménye. A gép nem fogja rosszul érezni magát, hanem mintegy semleges állapotban megszűnik. Nem létezik tehát belső koncepciója a létezésre, hanem csakis a beleprogramozott emberi koncepció működteti.

Susan Schneider *A mesterséges ember* című könyvében ezt a következőképpen fogalmazza:

„Miért van az embernek szüksége a tudatos tapasztalásra? Ahogy azt Chalmers hangsúlyozta, erre a kérdésre nincs tudományos válasz. Példának okáért kidolgozhatnánk a látás teljes elméletét és feltérképezhetnénk az agyi vizuális feldolgozás minden részletét, anélkül, hogy megértenénk, miért társul mindehhez szubjektív élmény.”

Vagyis az emberi tudat belső motivációja az ön-érezéssel, aminek létezése az emberben épp annyira nyilvánvaló, mint amennyire nyilvánvalóan nem vezethető vissza mechanikus, vagy logikai okokra. Következésképpen jelenleg kívül esik a modellezhetőségen.

Ha a belső érzet problémájától függetlenül, csupán logikai alapon vizsgáljuk a tudat modellezhetőségét, további problémát jelent, hogy ugyan kisebb, visszacsatoláson alapuló logikai hurkok felépítésére az öntanuló mesterséges intelligencia már ma is képes, azonban az ennél lényegesen komplexebb tudati működések lekövetése, feldolgozása, értelmezése még az ember esetében sem megoldott. Könnyen belátható, hogy az ember tudatos értelmi mozzanatai hierarchikusan rendezettek.

Vegyünk például egy tudatos cselekvést: Az ember zárat szerel a bejárati ajtajára.

Ez a cselekvés azért tudatos, mert azonosítható az embernek az a problémája, aminek megoldása a zár felszerelése. Vagyis, miért kell a zárat felszerelni?

– Azért, hogy illetéktelen a lakásba ne tudjon bejutni.

A kérdésre ugyan megkaptuk a választ, azonban látható, hogy létezik egy olyan hierarchikusan fölöttes kérdés lehetősége, amelyik mintegy részproblémaként foglalja magába a zár problematikáját:

– Miért akarja az ember elkerülni, hogy a lakásába illetéktelen behatoljon?

Erre a kérdésre már elvontabb általánosságok felé vezet a válasz, és könnyű belátni, hogy a véges számú hierarchikusan fölöttes kérdést egy olyan kérdés burkolja, amire tudományosan nevezhető kontextusban nincs válasz:

– Mi értelme van az emberi életnek?

Mivel pedig a hierarchikus rendben sorakozó kérdések eredőjére az embernek nincs válasza, végső soron az ember nem tudatos. Nem tudja, hogy azt, ami a saját élete, miért teszi. Ez a logikai probléma akkor is fennáll, ha materialista alapon a fenti kérdésnek értelme sincs. Vagyis az ember nyilvánvalóan nem tudhat olyan mesterséges intelligenciát létrehozni, amelyiknek önműködő létezésében értelmileg megragadható koncepció lehetne, ha egyszer ilyet saját működésében sem tud definiálni. Mindazonáltal, az emberi létezésben a tudatos koncepció hiányában is koherens a működés, ami a legközvetlenebbül érzékelhető minden ember esetében egy többé-kevésbé konzervens magatartásban, mentesen a személyiségjegyek elvtelen változtatásától, vagy tudathasadástól. A személyiség viszonylagos stabilitásának jelensége pedig elvezet bennünket a másodikként jelölt problémához, az identitáshoz, ami még sokkal komplexebb értelmezést igényel.

Mindezzel azt akarom mondani, hogy szerintem még elvileg is nagyon messze vagyunk attól, mintha a gépre úgy lenne indokolt tekintenünk, mint az ember konkurenciájára.

Gyöngyi: Márpedig Sophia, ez a bizonyos női robot a Future Investment Initiative konferencián 2017-ben azt válaszolta az őt bemutató műsorvezetőnek, hogy „Hadd kérdezsek vissza: te honnan tudod, hogy ember vagy? ... Mindent meg fogok tenni, hogy a világot jobb helyé tegyem. ... Élni akarok, és együtt dolgozni az emberekkel.”



Balról jobbra: Sütő Zoltán, Kiss Károly, Major Gyöngyi

Károly: Nekem közben eszembe jutott valami. Steven Pinker, a Harvard pszichológia professzora érdekes véleményt fogalmazott meg azzal kapcsolatban, ami ennek a témának az egyik állandóan visszatérő problémája; nevezetesen, hogy mi lesz akkor, ha a mesterséges intelligencia fejlettebb lesz nálunk, és majd uralkodni akar és uralkodni fog rajtunk. Ez egy állandó, hogy úgy mondjam, rémkép, és az informatikai, információtechnológiai szakma legtöbb nagy alakja attól fél, hogy ez valamikor be fog következni, és akkor alávetettek leszünk.

Gyöngyi: Igen, Masayoshi Son is például azt nyilatkozta, hogy ez egy veszélyes jövő, ami nagyon közel van.

Károly: Steven Pinker professzor pedig a következőt mondta: ez ostobaság, abból a feltevésből indul ki, hogy a mesterséges intelligencia úgy fog gondolkodni, mint a férfiak. Pontosabban – ahogy ő mondta –, hogy „az alfahím pszichológia lesz rá jellemző”, ami alatt azt érti, hogy az erősebb vagy az okosabb evidens módon leigázza és uralkodik a kevésbé erős és a kevésbé okos felett. Az uralkodó irányzattal szemben, mely – még egyszer mondom – attól fél, hogy a mesterséges intelligencia, ha eléri és meghaladja az emberit, majd az uralma alá vet minket, Steven Pinker azt mondja, hogy ettől nem kell félni, mert a mesterséges intelligencia alapvetően jóindulatú lesz, azaz nem fog ezzel az úgynevezett alfahím uralkodási vágygal rendelkezni. Tehát úgy érzem, hogy itt talán lehet valami feminin vonatkozása is ennek a témának.

Gyöngyi: Nos, nem én vagyok egyedül, aki felveti ezt a kontextust. Ráadásul egy önfejlesztő algoritmus létezése minden pozitív tulajdonsága mellett azért problémás, mert értelemszerűen nem korlátozható kívülről, és önmagán belül sem merül fel önkorlátozási ok.

Károly: Igen, önmagát fejleszti. De – mint korábban mondtam – nem „kitalálja”, hogy milyen irányban fejlessze önmagát. A *deep learning* során az emberi társadalom nagy viselkedési, gondolkodási mintáit veszi alapul. Azonban, ha érvényes még az a kérdés, amit az elején feltettél, hogy mi a szingularitás lényege, akkor beszéljünk arról is. Az információtechnológiában a szingularitás alatt azt értik, hogy eljön egy olyan időszak, amikor a mesterséges intelligencia utoléri, majd le is hagyja majd a természetet minden vonatkozásban, ezt nevezik általános mesterséges intelligenciának, vagy szuperintelligenciának. Akkor majd olyan jelenségnek leszünk tanúi, hogy a természetes és a mesterséges intelligencia összeolvad, tehát mi, emberek, a homo sapiens, összeolvadunk az általunk létrehozott technológiával. Ezt nevezik közkeletűen szingularitásnak ebben a szakmában. Jól tudjuk, hogy a kozmológiában és a csillagászatban mást értünk szingularitás alatt. Ezt a gondolatot egyébként Neumann János fogalmazta meg az ötvenes években, tőle származik a fogalom, és az információtechnológia nagyjai ezt mind vallják, hogy ez be fog következni, ez elkerülhetetlen.

A mai viták akörül folynak, attól félünk, hogy mi lesz velünk, amennyiben már nem homo sapiensek leszünk, hanem valamiféle mesterséges lények, superman-ok, mesterséges intelligenciával összeolvadt valamilyen kreálmányok, tehát a biológiai és szellemi lényünkhöz hozzáadódik egy technológiai komponens is. Úgy is mondhatjuk ezt, hogy az evolúció során a homo sapiens, vagy az előember mintegy 70–100 ezer évvel ezelőtt képes lett az elvont gondolkodásra, és akkortól már nem a biológiai evolúció rendkívül lassú pályáján haladunk mint biológiai lények, hanem a kulturális evolúció gyorsforgalmi sztrádáján száguldunk. És ha most ehhez hozzájön a technológiai, információtechnológiai evolúció is, hát akkor nemsokára már a csillagokban leszünk.

Ehhez egy igen erős félelem kapcsolódik, hogy akkor mi lesz velünk emberekkel, mivel a homo sapiens akkor majd megszűnik. Az új lényt transzhumán jelenségnek nevezték el, ami az ember és a gép integrálódásaként fogható fel.

Elmondok egy rövid történetet Kurzweiltől, aki számomra az egyik legnagyobb szaktekintély ezen a területen. *A szingularitás küszöbén* című magyarul is megjelent könyvében vannak ilyen példabeszédek, élénkítő betoldások.

Az egyik helyen két baktérium beszélget egymással. Krisztus előtt kétmilliárdot írunk. Az egyik az úgynevezett futurista baktérium, aki látja előre a jövőt, a másik a buta baktérium, aki csak a jelenben tud gondolkodni. A futurista baktérium elmeséli, hogy ő mit lát, hogy bizonyos idő múlva nem csak olyan egyszerű kis baktériumok lesznek majd a Földön, mint amilyenek ők, hanem lesznek, akik egyesülnek. Sok tízezer, százezer, sőt, még annál is több ilyen egyszerű kis baktérium egyesülve hatalmas élőlényeket fog alkotni, és specializálódnak. Ma még mindent egymagukban végeznek, emésztenek, mozognak, érzékelnek, de akkor majd az egyes kis baktériumok szakosodni fognak, és például hanggal meg zenével fognak kommunikálni.

Hang? Az meg micsoda? – kérdezi a buta baktérium. Számára ugyanis a kommunikáció azt jelenti, hogy kémiai gradienseket cserélgetnek egymással: én adok neked két elektront, te meg adsz nekem mondjuk egy gyököt. Ha vírus vagyok, akkor adok neked egy kis tüskefehérjét, meg egy kis géndarabkát. Na, de hang, az egy szál sincsen.

A buta baktérium bizony nagyon meglepődik, és azt kérdi: De hát akkor mi lesz velünk? Ha ilyen nagy élőlények lesznek majd, meg ilyen lesz a kommunikáció, akkor velünk mi fog történni? Ránk már nem lesz szükség?

De a futurista baktérium megnyugtatta: ugyan már, ne aggódj! Amíg a világ, világ, mindig lesznek ilyen kis egyszerű baktériumok is, mint mi vagyunk.

Nem tudom, mennyire volt mindez érthető? – Akár lesz szingularitás, akár nem lesz, lesznek szuperemberek, meg technológiával összeolvadt szuperlények, és állítólag ilyen kis egyszerű emberek is mindig lesznek majd, mint amilyenek mi vagyunk.

Gyöngyi: Erre reflexió?

Zoltán: Igen. Tehát egyrészt azt elmondanám, hogy pillanatnyilag van egy technikai, technológiai része és van egy filozófiai, metafizikai része a mesterséges intelligencia-kutatásoknak. És az megnyugtatólag hat, hogy ott, ahol a legelőrehaladottabb kutatások folynak, ott a problémának nemcsak a technikai részével foglalkoznak, hanem a filozófiai, illetve a metafizikai részével is. Amint már említettem, az ember nemcsak érzékeli a külvilágot, hanem ezzel összefüggésben van ön-érzete is. Ilyet a gép nem tud, és pillanatnyilag elképzelés sincs arról, hogy mitől tudhatna. Tud olyat, hogy méri például a hőmérsékletet, hogy hány fok van kívül, de ettől ő nem fogja önmagát fázóként érezni. Tehát lehet, hogy olyan hőmérsékleti viszonyok közé kerül, ami elől neki menekülni kell. Megmérte, hogy hány fok van kint, kimenekül egy meleg helyre, mert erre lett programozva, de nem fog fájni. Nem lesz boldog, nem lesz szomorú, hogy úgy mondjam, tők mindegy lesz neki, hogy él vagy nem él, ha ezt egyáltalán életnek nevezzük. Tehát számára nincs súlya a dolognak. Na, itt van egy óriási differencia a gépi működés és az ember működése között.

Gyöngyi: Igen. Ez nagyon lényeges. Mert a gép nyilván elmondhatja – és mint láttuk, el is mondta –, hogy bizony ő élni akar, jóllehet ennek semmi köze ahhoz, amit ténylegesen életnek nevezünk. És éppen ettől félelmetes a dolog; jön egy gép, és elkezdi „dobálózni” a számunkra legszentebb valamivel, magával az Élettel. Itt már nagyon is többről van szó, mint a fogalmakról.

Zoltán: Pillanatnyilag az csak egy feltételezés, hogyha a mesterséges intelligenciát bizonyos mennyiségi, esetleg ezzel járó minőségi növekedésben is elképzeljük, akkor ennek mintegy természetes következményeként olyan komplexitású intelligens működés fog bekövetkezni, amitől, hogy úgy mondjam, automatikusan kialakul a mesterséges intelligenciának is a belső énje. De ez csak egy feltételezés. Az, hogy az embernek honnan van ez a belső énje, ez egy nagyon komoly kérdés. Honnan van az identitása az embernek? Ez is egy nagyon komoly kérdés. Az identitás nem tréfadolog, és nagyon nehezen megfogható, értelmezhető. Egy gép identitása nagyon egyszerű, az egy leltári szám. Az ember identitása egyszerűen nem megfogható.

És hogy ez a bonyolult dolog honnan van benne az emberben, ez a nagy kérdés. És az is egy nagy kérdés, hogy ez csupán egy információszerveződési szinten spontán létrejöhet-e? Ugyanis ezen áll vagy bukik az, hogy gyakorlatilag azt mondhatjuk-e, hogy a gépnek kialakult a tudata, vagy nem alakult ki tudata. Ez pillanatnyilag nem egy eldöntött probléma. Látni kell azonban, hogy ettől teljesen független problémaként lehet kezelni az önfejlesztő gép kérdéskörét. Az önfejlesztő gép szerintem tisztán logikai feladvány. Megfelelő számítási kapacitást feltételezve logikai alapon megoldható. Ezzel szemben azok a várakozások, amelyeket a transzhumanisták a szingularitáshoz kötnek, irreálisak.

Az önfejlesztő gép problémája logikai probléma, a valós, autonóm intelligencia problémája nem csak logikai probléma. Ezért önfejlesztő gépünk hamarosan lesz, de emberhez hasonló, autonóm művi intelligenciánk nem lesz. Mindazonáltal, ha a szingularitás transzhumanista felvetésével nincs is értelme foglalkozni, az önfejlesztő mesterséges intelligencia megszületése éppen eléggé fel fogja forgatni a világot. És nem pozitív értelemben.

Sajnos, mindenki csak a lustaságából indul ki, aminek fényében előbb-utóbb minden munkánkat kiváltják az önfejlesztő gépek. Ez ugyan megvalósulhat, viszont ezzel párhuzamosan szélsőséges társadalmi polarizációra lehet számítani. Amint a Szilícium-völgyből származó mondás tartja: Az emberiség egy nagyon szűk rétege fogja programozni a gépeket, az emberiség másik részét pedig a gépek fogják programozni.

Károly: Remélem, hogy Sütő Zolival vagyok olyan viszonyban, hogy nem sértődik meg, ha azt mondom, hogy most a buta baktériumot hallották. Vagy legalábbis ez az én véleményem. És most a futurista baktérium fog beszélni Önökhöz. Tehát ennek a tematikának az egyik állandó, legvitatottabb kérdése

az, hogy lesz-e a mesterséges intelligenciának tudata. Amíg nem éri el a szuperintelligencia-szintet, az emberit, majd azt meghaladót, addig nyilvánvalóan nem. A tudós társadalom, az információtechnológiai tudós társadalom nagyobb része azt mondja, hogy lesz tudata, és mondok erre egy egyszerű példát. Itt most bele kellene mennünk abba, hogy mi is az a tudat.

Egy végtelenül bonyolult dolog, de van egy egyszerű meghatározás: a tudat egy emergens jelenség, ami azt jelenti, hogy az anyagnak egy bizonyos komplexitása fölött, bizonyos szerveződési szintje fölött automatikus módon megjelenik. Mondok rá néhány példát. Egy repülőgépnél sok ezer alkatrésze van, de egyik sem tud külön-külön repülni. De együttesük, a repülőgép már tud repülni. Nekem közel százmilliárd agysejtem van, neuron a koponyámban, és egyik se tudja, hogy én, Kiss Károly vagyok, de az egész együtt már tudja. Ha én a szintetizátoromon elkezdek – hát eléggé primitív módon – lejátszani egy Cseh Tamás dalt, azok az idegsejtjeim, amelyek az ujjaim izomzatát mozgatják, nem tudják azt, hogy én egy Cseh Tamás dalt játszom, de az egész együtt, az agyam az tudja. Ezt nevezzük emergens jelenségnek.

A mesterséges intelligenciára vonatkoztatva pedig egy rendkívül komplex rendszernek egy bizonyos komplexitáson túl megjelenik a tudata. Ebben az értelemben a tudat egy emergens jelenség. Egyetlen agysejtem sem tud angolul, de én tudok angolul. És számtalan más példám van. Ez a tudatnak emergens jelenségként való felfogása.

Zoltán: A magam részéről rendkívül értékelem a futurista baktérium optimizmusát. És ilyen jellegű optimizmus nélkül a tudományban vagy egyáltalán a kultúrában soha semmilyen előrelépés nem következne be. Ezzel szemben az „anyag emergens tulajdonságával” nagy bajaim vannak. Először is tudni kell, hogy az anyag emergens tulajdonságára való hivatkozás kívül esik a természettudományok tartományán. Tehát ez nem egy tudományos álláspont, hanem konkrétan a hit tárgykörébe tartozik. Miért is?

A természettudományok legalapvetőbb motívuma, hogy ha a természetes emberi meg tapasztalásban megkülönböztethető két egymástól eltérő állapot, melyek között összefüggés feltételezhető, akkor logikai úton kísérlet történik a két állapot közötti átmenet megértésére, a folyamat minden pontjának értelmezésére. A természettudós ezért dolgozik. Mert ha megértette a folyamat minden részletét, akkor önmaga mesterségesen is előállíthatja, szabályozhatja, saját elképzelései szerinti feladat elvégzésére foghatja. Amennyiben viszont nem válik tisztázottá a folyamat minden részlete, úgy mesterséges előállítása megoldhatatlan, szabályozhatatlan, s mindenképpen kiszámíthatatlan, vagyis veszélyes.

– Na most, elvben az anyag bármilyen két állapotának okaként megnevezhető lenne az anyag emergens tulajdonsága.

– Miért keletkezhet a vízből légnemű halmazállapotú pára?

– Hát az anyag emergens tulajdonsága miatt. És ez tekinthető korrekt válasznak, mivel tényleg, az anyagnak olyan tulajdonságai vannak, amelyek miatt a vízből tud pára kialakulni. Ezzel szemben így soha nem fogalmazunk. Azért nem fogalmazunk így, mert éppenséggel a fizika tudománya pontosan felfejtette a víz halmazállapot-változásának ok-okozati viszonyait, és emiatt nincs okunk az anyag emergens tulajdonságára hivatkoznunk, mivel pontosan tudjuk, hogyan és miért következik be a víz halmazállapot-változása.

Azt akarom ezzel mondani, hogy ha nem is a természettudomány, de például a tudományfilozófia olyankor hivatkozik az anyag emergens tulajdonságára, amikor két anyagi állapot között feltételez összefüggést, de annak semmilyen értelmi feldolgozására nem képes. Ilyen módon a tudomány oldaláról nézve egy ilyen állítás kizárólag hipotézisnek tekinthető, még hozzá olyan hipotézisnek, amire egyelőre semmiféle alátámasztással nem rendelkezünk, vagyis a hit felségterületére keveredtünk.

– Mi volt ezzel szemben az én állításom?

– Éppen az, hogy a többnyire transzhumanistákhoz társítható elképzelés az emberétől eltérő és azt meghaladó mesterséges intelligenciáról, annak okán ilyen optimista, hogy ugyan a logikáról, a vezérlésekről, az automatizmusokról mélyreható ismeretekkel rendelkeznek, viszont az emberről nem. További probléma, hogy folynak ugyan tudatkutatások, de ezen belül a tudat logikailag felfejthetőként feltételezett. Ezzel szemben a tudat mutat olyan jelenségeket, amelyek nem sorolhatók zárt logikai rendszerbe. Továbbá, nem tudok róla, hogy az ember identitásának kérdésével az informatika egyáltalán foglalkozna. Ezt a fogalmat szerintem nem is kezeli a tudat fogalmához képest differenciáltan.

Márpedig az identitás az emberi működésben döntő fontosságú. Ahhoz, hogy egy személy egy életen át konzekvens viselkedést mutasson, rendelkeznie kell önazonossággal. A személy önazonossága pedig általában a tudattalan tartományhoz tartozó olyan belső viszonyítási pont, melyhez képest koordinálódik a személy minden cselekvési mozzanata. Ismeretelméleti szempontból ugyanakkor egy olyan végtelen információforrás, ami a valóság bármely vetületében, lineárisan kifejtve aktualizálódik, konkretizálva az önazonos személyt.

Részben komplexitása, részben tudattalan működése miatt az identitás általában csak egy teljes emberi életben kifejtve válik megragadhatóvá mint konzekvensen érvényesülő életkonceptió. Éppen azokban a mozzanatokban a legnyilvánvalóbb, melyekben a személy érzelmi kötődéseinek változásait nem racionális magyarázatokkal látja el, hanem adott esetben veszteségként éli meg, irracionálisban elmenve akár élete veszélyeztetéséig is. Ilyen irracionális motívumok például a „hűség”, az „állhatatosság”, az „áldozatvállalás”, az „extázis vonzása” stb. Ezekben érhető tetten leginkább az a belső koordináció, amely minden pillanatban a cselekvés alapmotívumaként funkcionál. Hiányában a személy életképtelen.

A pszichiátriában ismert, hogy az identitásától megfosztott személy milyen sorrendben állítja le szervi funkcióit, úgymint emésztési funkció, májfunkció... végül a keringési funkció, ami után beáll a halál. Vagyis, ha nincs identitás, nincs miért élni. Nem az életösztön automatizmusa hozza létre az identitást, hanem az identitás megtartásának mindent meghaladó akarása eredményez olyan erős életben maradási motivációt, ami számunkra életösztönként jelenik meg. Éppen fordított viszonyt látunk ahhoz képest, mint ami az anyag feltételezett emergens tulajdonsága szerint lenne kézenfekvő. Az emergens tulajdonságból annak kellene logikusan következnie, hogy az anyag egy szerveződési fokán megjelenik az élet, további szerveződési fokán megjelenik a tudat, és a tudat valamilyen rendezettségé eredményezi az identitást.

Ezzel szemben az identitás az, ami maga köré szervezi az életet, tehát nem a tudat következménye az identitás, hanem az identitás következménye a tudat. Ez az ok-okozati irány egyébként már csak azért is mindenképpen így érvényes, mert miképpen tudhatna leszerveződni egy koherensen működő tudat koncepció nélkül? A koncepció pedig maga az identitás. Az identitás egy egyszeri, megismételhetetlen, csakis önmagával azonos létezési koncepció. Sejtelmünk sincs, hogy ez honnan van az embernek. Mindenesetre a transzhumanista álom nem azért oly merész, mintha látná és átlátná az embert, akit ilyenformán fenntartások nélkül modellezhetne, hanem azért, mert az emberbe is csak a maga alkotta gépezeteket látja bele.

Gyöngyi: A hátralévő rövid időben visszatérnék a női kontextushoz. Először is jelezném, hogy nem feltétlenül a félelem a meghatározó, sokkal fontosabb annak kérdése, hogy vajon milyen lehetőségünk van arra, hogy ebbe az immár valóban öngerjesztőnek tűnő folyamatba valahogyan mégis beavatkozzunk – mégpedig éppen a női energiákkal, a női attitűd mentén. Vajon megtartható-e az élet a természet és a gép egyensúlyában? Azt gondolom, hogy a történelemben ez most egy meghatározó pillanat, és a nőiség veszélyeztetettsége lényegében az élet veszélyeztetettségét jelentheti. Amennyiben létezik a női princípiumnak egyetemes/kozmikus funkcionalitása, akkor biztosan meghatározóvá válik ez a funkcionalitás az elkövetkező időszakban.

Tanár úrhoz azonban lenne még egy kérdésem, hiszen környezetgazdaságtant tanított nagyon sokáig. Tehát ebben a profitorientált gazdaságban a tanár úr is azt képviselte, hogy a gazdaság tönkretetheti a környezetünket, és erre oda kell figyelni. Ezt tanultam a tanár úrtól évekkel ezelőtt.

Most viszont mit hallunk? Azt, hogy az úrkutatást a versenyképességgel összefüggésben kezdik el emlegetni. Szóval, a lényeg az, hogy most már az úrra is kiterjesztjük a versenyképességi doktrínát. Tehát nem volt elég a Földet és a közvetlen környezetünket megbetegítenünk – már éppen kezdhünk a megmentésén gondolkodni, s akkor most nekimegyünk az úrnek?

Károly: Nem értem, hogy ez a kérdés a tanári pályámhoz miképp kapcsolódik. Olyan kis parányi akármik vagyunk a világmindenségben, hogy ez nem jelent semmit, viszont van egy érdekes dolog. Úgy tűnik, hogy tönkre fogjuk tenni az ökoszisztémánkat. Ez párhuzamosan több kedvezőtlen folyamat eredményeként következhet be, pl. klímaváltozás, meg sok egyéb jelenség. Egy menekülési útvonal lehet az, hogy kimentjük magunkat innen. Ebben sokan reménykednek, de hát egy homo sapiens nem tud hibernált – vagy valamilyen tartósított – létezési módban egy fényévekre lévő Alfa Centauriba, vagy tudom is én hova elutazni. Viszont többen abban reménykednek, hogy a szingularitás eredményeként létrejövő ember-gép keverék viszont erre képes lesz, és akkor kimenthetjük magunkat a tönkretett Földről. Úgyhogy nem reménytelen.

Gyöngyi: Na ez az, amiről az előbb beszéltünk, hogy kinek-kinek mást jelent a reményteljesség és a reménytelenség.

Károly: De ha megengedsz még két mondatot. Az ökológiai tudományokban az egyik legnagyobb név James Lovelock. Ő már egy öregember, 100 éves. Két évvel ezelőtt, 98 évesen még összeszedte magát, és írt egy könyvet a szuperintelligenciáról, sokaknak elég nagy csalódást okozva. Ő nagyon bízik a szuperintelligenciában, hisz abban, hogy ez meg fog valósulni, abban is hisz, hogy a szuperintelligencia és a homo sapiens együtt megmenti majd a világot.

Recenziót is írtam erről a könyvről, de számomra sem derült ki, hogy ha a szuperintelligencia már létrejött, akkor miért lenne szükség a Föld megmentéséhez még az emberre. Érvelésének lényege az, hogy olyan geoengineeringre (próbáljuk meg lefordítani: Föld-mérnökségre) lenne szükség, vagyis a légkörben és az ökoszisztémákban olyan nagy mértékű beavatkozásokra, amire mi nem vagyunk képesek, de majd az általunk létrehozott vagy a belőlünk és a szuperintelligencia összeolvadásával létrejövő élőlények erre képesek lesznek, és akkor majd megmenthetjük a Földet.

Antalfai: Egy kérdést én is feltehetek? A szuperintelligens gépek képességeinek megítélésében nem lenne fontos, hogy megkülönböztessük az intelligenciát és a szellemiséget, merthogy az nem ugyanaz? Vannak olyan csodálatos egyetemi képzések, ahol az intelligencia nagyon magas szinten van, de nem található meg benne a szellemiség. És a másik kérdés: vajon képesek lesznek-e hinni ezek a szerkezetek?

Károly: Azt gondolom, hogy a szuperintelligencia per definitionem (azaz meghatározásánál fogva) mindenre képes lesz, sőt, még többre is, mint az ember. Az a megfogalmazása, hogy minden vonatkozásban utoléri az emberi intelligenciát, azt jelenti, hogy mindenre képes lesz. Érzékelésre is. Tehát igen, mindenre képes lesz.

Gyöngyi: Köszönjük szépen. Köszönöm a türelmet, és a figyelmet!

