

Síklaki István: Ne féljünk a számítógéptől!

Hivatkozás/reference:

Síklaki István, „Ne féljünk a számítógéptől!”,
Információs Társadalom, XV. évf. (2015) 4. szám,
48-50. old.

<http://dx.doi.org/10.22503/inftars.XV.2015.4.4>

Reakció Z. Karvalics László a "Mesterséges intelligencia – a diskurzusok újratervésének kora" című vitaindító írására.

Do not be afraid of computers!

Comments to the paper by Z. Karvalics László, "Artificial intelligence – why to redesign the discourses?"

Információs Társadalom

Vita a mesterséges intelligencia
fejlesztésében rejlő
lehetőségekről és veszélyekről

2015. XV. évfolyam 4. szám

A folyóiratban közölt művek a *Creative Commons*
Nevezd meg! - Ne add el! - Így add tovább! 4.0
Nemzetközi Licenc feltételeinek megfelelően
használhatók.

Síklaki István

Ne féljünk a számítógéptől!

Hozzászólás Z. Karvalics László vitaindító cikkéhez

Előrebocsátom, hogy nem fogok vitatkozni Z. Karvalics László kiváló és gondolatébresztő cikkével. Teljes mértékben egyetértek azzal a nézetével, hogy tévút az alarmista diskurzus arról, hogy a gépi intelligencia előbb-utóbb eléri azt a fejlettségi szintet, ahol legyőzi az emberi intelligenciát, ami az emberiség számára végzetes következményekkel járhat. Mint az emberi elme működése iránt érdeklődő szociálpszichológus, szeretnék néhány adalékkal hozzájárulni Z. Karvalics érveléséhez.

Először szeretném felvázolni, hogy a magam együgyű módján miként képzelem el az emberi elme működését. Ebből a vázlatos modelltől nyomban adódnak következtetések az alarmista nézetek számára. Az emberi elme egy evolúciósan kialakult, olyan hihetetlenül robusztus hálózat, amely folyamatosan fenntart, működtet egy modellt, a külső és belső világunk emulációját, annak történetével, pillanatnyi állapotával és közelebbi és távolabbi jövőjével együtt. Tehát amikor tudatos gondolataink vannak, akkor valójában ehhez a virtuális emulációs modellhez van hozzáférésünk, a modell által a tudatos folyamatok rendelkezésére bocsátott információt használjuk föl. Erre nyomban hozok egy-két egyszerű példát, de előbb néhány adattal érzékeltetni szeretném ennek a modellnek a biológiai nagyságrendjét, amit érdemes összevetni a szuperszámítógép rendszerek nagyságrendjével.

Az agyban hozzávetőlegesen 100 milliárd neuron található. Ezek a neuronok olyan hálózatot alkotnak, ahol a 100 milliárd neuron mindegyike hozzávetőleg 8-10 000 másik neuronnal áll szinaptikus kapcsolatban. Tehát a lehetséges kapcsolatok száma egy adott pillanatban 100 milliárd a tízezrediken. Ez a hálózat megállás nélkül működik, alapjáraton a neuronok körülbelül másodpercenként negyvenszer sülnek ki, az aktív neuroncsoportok esetében ez elérheti a másodpercenkénti ezer kisülést. És ez csak az alap. Ezt a működést modulálja több tucat neurotranszmitter és más aktív molekula, és még egyéb tényezők, csak amikről már van valamelyes tudomásunk. Ez az emulációs modell, amely a tudatunk számára a virtuális valóságot előállítja, moduláris szerkezetű. Részben az evolúció során kialakult és velünk született, részben tapasztalati tanulóssal kialakított tudattalan autonóm modulok óriási és folyamatosan más és más konfigurációkba rendeződő hálózata alkotja. Ez az elrendezés biztosítja, hogy képesek vagyunk tudattalanul óriási információmennyiséget párhuzamosan feldolgozni. Ehhez képest közismert, hogy a tudatos folyamataink igen lassúak, szekvenciálisak, és korlátozott a kapacitásuk.

Lássunk egy egyszerű példát arra, hogy a tudatunk nem fér hozzá a valóságos világhoz. Az egyik közismert példa, amire az interneten számos helyen nagyszerű illusztrációkat találhatunk, a változás vakság. Bizonyítható, hogy ha például egy vizuális bemeneten az adott szituációban irreleváns változás következik be, akkor azt tudatosan képtelenek vagyunk észlelni. Nem azért, mert a bemenet nem járná végig a természetes útját a retinától a V1 látómezőn át az egész vizuális rendszeren, hanem azért, mert a tudattalan modellünk,

amely irrelevánsnak találta a változást, aktívan gondoskodik róla, hogy ne jusson el a tudatunkba, ne terhelje a szűk kapacitást fölöslegesen. Tehát egy ilyen végtelenül egyszerű észlelési feldolgozás mögött is ott van az evolúció és az egyéni fejlődés eredményeként kialakult sajátos virtuális modell, amely primer módon biztosítja azt, amit a mesterséges intelligencia számára lehetetlenség beprogramozni. Ez teljesen összhangban van Z. Karvalics gondolatmenetével.

Egyelőre tehát ott tartunk, hogy a saját elménk működését nem vagyunk képesek olyan mértékben megismerni, még megközelítőleg sem, hogy algoritmizálható módon felhasználhassuk a mesterséges intelligencia építésére. Ennek illusztrálására egy klasszikus kognitív szociálpszichológiai kísérletet szeretnék felidézni. Timothy Wilson (2010) kezébe vette az amerikai fogyasztóvédelmi magazin, a *Consumer Digest* egyik számát, és megnézte azt a cikket, amelyben 16 ételszakértő tesztelte a piacon lévő negyvenvalahány dzsemet, és felállítottak szokás szerint egy minőségi rangsort. Wilson ebből a spektrumból találmra kiválasztott egy fél tucatot lehetőleg egymástól távol a rangsorban, és laikus egyetemi hallgatókat kért meg a kóstolásra és rangsorolásra. Amikor összehasonlította a szakértők rangsorolását a laikus egyetemistákéval, meglepetéssel tapasztalta, hogy igen magas korrelációt mutatott ($r=0,46$). Ekkor ugyanilyen feltételekkel megismételte a kóstolósos rangsorolósos vizsgálatot azzal a csekélynek tűnő különbséggel, hogy most indoklást is kért a rangsorolásra. Tehát tudatosan meg kellett adniuk, hogy mely tényezőket vették figyelembe a rangsor kialakításánál. Ebben a változatban a szakértőkkel való korreláció drámaian lezuhant ($r=0,11$). Tehát amikor átváltottak a kognitív tudattalan párhuzamos üzemmódjáról a tudatos szekvenciális üzemmódra, a teljesítmény látványosan leromlott. Nyilván áldozatos munkával megoldható, hogy a szakértők tudását valamely jól szervezett szakértői rendszerbe szervezzük, s ez által algoritmizálhatóvá és így gépesíthetővé tegyük. De ha belegondolunk, hogy milyen csillagászati apróságból áll össze folyamatosan az emberi teljesítmény, nyilvánvaló, hogy a mesterséges intelligenciának ez az útja nem járható.

Egy egészen más oldalát is fontolóra érdemes venni ennek a kérdésnek. A cikk vége felé Z. Karvalics ír a mesterséges intelligencia és a felelősség kérdéséről. Voltaképpen ugyanezeket a kérdéseket fel lehet vetni az emberi felelősséggel kapcsolatban is. A tetteinkért vállalt felelősség ugyanis azon a meggyőződésünkön alapul, hogy szabad akaratú rendelkező, szuverén személyek vagyunk. Igen ám, de számos kutatás mutatta ki elég meggyőzően az utóbbi időkben (például Wegner, 2006), hogy ez az élményünk nem azon alapszik, hogy közvetlen hozzáférésünk lenne az agyunkban, azaz, a tudattalanunkban lejátszódó döntési folyamatainkhoz. Ez az élmény csupán illúzió, mert az agyunk aktívan elzárja a tudatunktól mindazokat a folyamatokat, amelyek elvégzése után valamilyen eredményre jut, s csupán az eredményhez enged a tudatunknak hozzáférést. A saját indítékainkra tehát éppúgy következtetés útján jutunk, mint más emberek indítékaira. Ám mivel az agyunk elzárja előlünk azt a miriádnyi számítást, ami egy-egy tettünkhöz vezet, az az illúzió, hogy szuverén, szabad akaratú rendelkező emberek vagyunk, s ezt tételezzük föl embertársainkról is. Így megvan a felelősség alapja. Ha egyszer eljutna oda a tudomány, hogy láthatóvá tegye ennek az irdatlanul összetett virtuális emulációs modellnek, a tudattalannak a működését, akkor pontosan azt a kérdést tehetnénk fel a felelősségről, ami mesterséges intelligencia rendszerek kapcsán felvetődött.

Végül csak utalnék arra, hogy a modern idegtudomány nagyon meggyőzően bizonyította, hogy az érzelmi működések nélkül, amelyek pedig a zsigereinkben gyökereznek, még egyszerű döntéseket sem vagyunk képesek meghozni (Damasio, 1996; Panksepp 1998). Nem nehéz belátni, hogy az érzelmeink nagyon szaftos biológiai alapjait is magában foglaló érzelmi rendszert programozni igen kétes vállalkozás.

Irodalom

- Damasio, Antonio, R. (1996): *Descartes tévedése*. Adu Print Kiadó
Panksepp, J. (1998): *Affective Neuroscience*. Oxford University Press
Wegner, D. M. (2006): *A tudatos akarat illúziója*. Kossuth Kiadó
Wilson, T. D. (2010): *Ismeretlen önmagunk*. Háttér Kiadó