

**Vita:**

**Fellner Ákos: Ne misztifikáljuk túl a mesterséges intelligenciát!**

*(comments are welcome!)*

A mesterséges intelligencia néha túl van misztifikálva, és van, hogy démonizálják is, avagy éppen az emberiség megmentőjét látják benne. Azonban ha megnézzük a matematikai működési elvét, sok alaptalan félelemet és túlgondolást kiiktathatunk.

A mesterséges intelligencia matematikai értelemben tulajdonképpen egy döntésméleti algoritmus, amely meghatározza, hogy bizonyos határértékek esetén egy függvény vagy több, vagy függvények végtelen halmaza miként csoportosítható, klaszterezhető úgy újra, hogy abból olyan új függvények vagy függvény halmazok jöjjenek létre, melyek információs értéke nagyobb mint a kiinduló halmazoké.

Vagyis a mesterséges intelligencia valóban dönt, de csak egy meghatározott transzformációs elv avagy szélső függvény határértékek esetén, és semmi másról mint hogy hogyan csoportosítson, klaszterezzen újra már meglévő függvény halmazokat.

Az hogy ez az algoritmus alkalmas bizonyos új összefüggések feltárására, nem jelenti, hogy ennél több lenne.

Fontos még, hogy a mesterséges intelligencia képes magukat a transzformációs függvényeket is átalakítani (smart moving algorithms), de ezt is csak a meghatározott hibahatárok és szélsőértékek között tudja tenni.

Kiss Károly okvetetlenkedő megjegyzése:

Kedves Ákos, ha az agyműködésünket akarnád leírni a matematika nyelvvel, akkor mit mondanál? Nem ugyanezt? Agyunk, vagy akár Einstein agya sem a semmiből alkotja/alkotta meg az új összefüggéseket; azok nem a meglévő tudás „újraklaszterezéséből származnak”? Az intuíciók isteni sugallatok? Nem a meglévő tudás, tapasztalat áll mögöttük?

Számomra az a hihetetlen csoda, hogy a gép mindezt „valós időben” csinálja, olyan gyorsasággal, ahogy agyunk sejtjei, molekulái és atomjai végzik a kémiai gradiensek cseréjét... - azaz gondolkodunk.

Fellner Ákos:

*Szerintem fontos az agyműködést elkülöníteni az intelligenciától. Nem csak aggyal rendelkező rendszerek lehetnek intelligensek. De ha úgy érted, hogy a humán intelligencia működését akarjuk matematikailag leírni igen, az alapszerkezet ez lenne: smart (azaz önmagát korigálni képes) döntésméleti algoritmusok sorozata.*

*A Big Data elemzésére fejlesztett MI algoritmusok jóval több adatot képesek feldolgozni mint az emberi agy. A probléma az, hogy a filozófia vagy a teológiatörténet olyan fogalmakkal terheli ezt a modellt, mint a gondolkodás és az intuíció, vagy az emlékezet. Léteznek a matematikában is intuíciók és heurisztikák,*

*ellentétben a filozófiával vagy teológiával, de általában nem mondjuk, hogy valamilyen transzcendens dologra vonatkoznának.*

Kiss Károly: Ezzel elérkeztünk Max Tegmarkhoz (Élet 3.0), aki azt állítja, hogy az intelligencia az atomok olyan elrendeződése, ami képes az információfeldolgozásra (azaz az összefüggések és hasonlóságok felismerésére). Továbbmenve: sem az intelligencia, sem a számítás nincs egy meghatározott anyaghoz kötve, anyagsemlegesek (azaz szervetlen anyagon is létrejöhetnek).

Csikvári F. András: Kedves Ákos,

Nem biztos, hogy jól értettem az írását. Nem tőlünk függ, hogy az MI veszélyes lesz vagy csak egy interaktív lexikon vagy csak egy aranyos társalkodónő?

Ha rá lesz bízva a bejövő katonai információk elemzése **és** meg tudja nyomni a kilövő gombot, akkor nem fogja megnyomni, amikor szükségét látja?

Fellner Ákos: Kedves András,

Igen, itt az a gond, hogy a fejekben néha iszonyat keveredik a real science és a science fiction.

A közgondolkodásra a 80-as években nagy hatást tett a Terminátor c. film, amiben Skynet, ami egy centralizált MI, öntudatra ébred és háborút kezd az emberiség ellen.

Sajnos számos félrevezető, féltudományos narratíva is kering az interneten, ami szintén azt sugallja, hogy a MI valamiféle kvázi-akarattal rendelkező ágens, ami egyszer csak elszabadul. Erre aztán különböző összeesküvés elméletek is rakódtak, és máris kész egy teljesen torz kép a MI kapcsán.

Mint az előzőekben is írtam, a MI nem más mint voltaképpen egy Big Data analízisre kifejlesztett analitikus algoritmus sorozat. Szerepe abban áll, amire az első hozzászólásomban is hivatkoztam.

Hogy az Ön példájára hivatkozzak, a MI sosem fog nukleáris rakétát kilőni, csak mert éppen úgy tartja kedve.

Ilyen helyzet csak akkor lenne elképzelhető, ha maguk az emberek úgy programoznának be valamiféle algoritmust, amely a számos katonai adatból azokat a szignálokat keresné, amelyek potenciális atomtámadást jeleznének. Amennyiben ezeknek a szignáloknak a száma túllépné az elfogadható felső határértéket, akkor lenne elképzelhető amit Ön kérdez. De ehhez az algoritmust a humán ágenseknek kell így megalkotni.

Az, amit James Cameron klasszikus filmje felvet, matematikailag nem lehetséges. Egy algoritmus sorozat – még ha önfejlesztő is (ettől nevezzük mesterséges intelligenciának) – sosem lesz képes olyan komplexitást magától elérni, ami ahhoz kellene hogy kilőjön egy nukleáris rakétát, mert éppen erre van kedve. Ez azért van,

mert a MI algoritmusai automatikusan elvetik a felesleges hurkok gyártását, vagyis az Occam-borotvája elvén működnek.

Természetesen biztosan sok anyagot lehet találni ennek az ellenkezőjéről, de ezek matematikailag nem megalapott nézetek. Innentől kezdve inkább az a kérdés, miben hiszünk, vagy miben akarunk hinni.

Remélem válaszoltam a kérdésére.

Kiss Károly: A MI képes önmagát fejleszteni. Ezt a programozói lépten-nyomon felfedezik: az algoritmus új, egyszerűbb utat talált a megoldáshoz. (Ez különösen a deep learning-típusú, az agy működését utánozó MI-nál fordul elő.) Kis idő kérdése, hogy különféle célú algoritmusokat algoritmusok fejlesztésére hozzanak létre.

A szakirodalom nagyjai (Kurzweil, Bostrom, Tegmark, stb.) evidensnek tekintik, hogy a MI idővel (a század közepére becsülik) saját lábra fog állni, függetlenedik az embertől. (Ezt a típusát már ÁMI-nak, általános mesterséges intelligenciának nevezik.) A kérdés, a félelem tehát, hogy mi lesz ekkor, nem fog-e bennünket maga alá vetni, nem fog-e uralkodni rajtunk, indokolt. De Steven Pinker, a Harvard pszichológus-professzora erről egy érdekes véleményt fejt ki. Szerinte a jellegzetes MI-disztópiák „a szűklátókörű alfa-hím pszichológia kivetődései az intelligencia koncepciójára”. (Azaz az intelligensebb lény hatalmát mások meghódítására használja fel.) Szerinte a rettegők téves platformon állnak: azt hiszik, hogy az intelligenciából következik a dominancia vágya.<sup>1</sup>

Bp, 2023. jan. 30.

---

<sup>1</sup> Lásd erről: a John Brockman által szerkesztett: Possible Minds. 25 Ways of Looking at Artificial Intelligence c. könyv ismertetését a Valóság 2021/11. számában. Ez és egyéb, a MI-ról szóló írásaim az Utódaink Jövője Új Technológiák oldalán olvashatók.