

Értsük meg a mesterséges intelligenciát
JELIGE: MI A MI 7

Pályázat

*”a ChatGPT és más nagy nyelvi modellek
alkalmazásának lehetőségei
és egyben biztonsági kérdései”*

kiírásra a

“A ChatGPT és más nagy nyelvi modellek
(LLM-ek) alkalmazásának lehetőségei”
témakörben

Kivonat

A(z LLM alapú) MI 2023 közepén már olyan teljesítményeket tud felmutatni, amelyet még akár néhány éve is elképzelhetetlennek tartottak, de a kérdések csekély mértékű átfogalmazása esetén már kudarcot vallanak. Az LLM (Large Language Model) elvű alkalmazások “fontos szerepet játszanak mind a kutatásban mind mindennapjainkban, *értékelésük fontossága növekszik . . . társadalmi szinten is meg kell értenünk lehetséges veszélyeiket*”. [47] “*A[z LLM] technológia használata kikerülhetetlen, ezért annak tiltása nem megoldás*”[20], meg kell tanulnunk együtt élni vele.

A Mesterséges Intelligencia (MI) az emberi intelligenciát próbálja meg modellezni, de a technikai megoldásainak működési sajátosságai nagyon távol állnak a biológiai működés sajátosságaitól, ezért tulajdonságaik is erőteljesen különböznek. Emiatt hiányzik belőle a neve által sugallt intelligencia. *A MI lehetőségei messze állnak attól, amit kezdetben feltételeztek, de veszélyei sem akkorák, mint jelenleg feltételezik; feltéve hogy megértjük, helyesen használjuk és szabályozzuk használatát.*

A fogalmak pontos meghatározásának hiánya és a terület több tudományterületre kiterjedő jellege még a szakemberek számára is megnehezíti a működés, a reális célok, a szabályozási követelmények új szempontjai, a használat célszerű módja és annak veszélyei átlátását; felhasználói számára pedig érthetetlen működést, antropomorf gépi munkatársat, teljesíthetetlen elvárásokat és többé-kevésbé indokolatlan félelmeket jelent. *A megértés hiánya miatt a technológiát a szakértők (is) “fekete doboz” jellegűnek tekintik és annak értékelésekor csupán személyes meggyőződésükre (jobb esetben szűk körű tapasztalataikra) hagyatkoznak.* Egy részük a megértés és indoklás felcsillanását véli látni bizonyos tesztek sikeres teljesítésében, mások pedig a technológia megbízhatatlanságát hangsúlyozzák [105][41]. Nem tanultunk az előző két MI-”tél” tapasztalataiból: emiatt már közelít a harmadik.

A dokumentumban áttekintjük a MI fogalmát, jellemző válfajait, annak működésével kapcsolatos téves fogalmakat továbbá a megvalósítás lehetséges technikai formáiból következő működési korlátokat. Kapcsolatot teremtünk a feltételezett és tényleges működés, valamint a MI használata során tapasztalt működési anomáliák között. A ma MI néven ismert számítógépes tevékenység lényegében túlnyomórészt gépi tanulás, szabályozástechnika, alkalmazott statisztika. Megmagyarázzuk, miért jelent a MI jellegű számítási mód használata új kihívást a használat szabályozása és az erőforrások használata terén, továbbá energia felhasználási, fenntarthatósági és környezet szennyezési szempontból is. Megmutatjuk azt is, hogy az utóbbi jelenségek végső oka a láncolt műveletek során az adattovábbítási idő figyelmen kívül hagyása a technikai számítási rendszerekben: az érvényességi körén túl használt számítási elmélet nem tud helyes válaszokat adni az MI működéséről.

Tartalomjegyzék

Kivonat

Ábrajegyzék	iv
-------------	----

Kulcsszavak	vi
-------------	----

1 Bevezetés	1
1.1 Kiírás és MI stratégia	1
1.2 A nagy nyelvi modellek	4
1.3 A technikai megvalósítás hatásai	6
1.4 A kiábrándulás	7
1.5 A dokumentum felépítése	9
2 A MI általában	10
2.1 A MI története	11
2.2 A MI fogalma	12
2.3 A MI meghatározása	13
2.4 A TI meghatározása	17
2.5 A MI fő típusai	19
2.6 A generatív eloszlás Turing-tesztje	21
2.7 Javasolt osztályozás	23
2.8 A puding próbája	26
2.8.1 A COVID kezelése	27
2.8.2 Elektronikus tervezés	28
2.9 LLM/ChatGPT	28
2.10 A MI használat szabályozása	31
3 A MI hiba forrásai	34
3.1 Gyenge MI	34
3.1.1 Stabilitás és megbízhatóság	34
3.1.2 Téves felismerés	35
3.1.3 Rejtett adat-összefüggések	37
3.1.4 Katasztrófa-szerű felejtés	37
3.1.5 Indokolhatóság	39
3.1.6 Megbízhatóság	40
3.1.7 Háttér ismeretek	41
3.2 Erős MI	42
3.2.1 Generatív MI	42
3.2.2 Probabilisztikus viselkedés	44

3.2.3	Következmények	45
3.2.4	Jellemző paraméterek	47
3.2.5	Megbízhatóság	47
4	A MI használata	48
4.1	Valós idejű működés	48
4.2	Video analízis	49
4.3	Adatkészletek használata	50
4.4	Felelősség	51
5	A MI ártalmai	52
5.1	A szabad használat veszélyei	52
5.2	Használható-e a ChatGPT EU-rendeletek készítésére?	54
5.3	Fenntarthatóság és szennyezés	56
5.4	Energia fogyasztás	60
5.5	Információs környezet szennyezés	62
5.6	Kritikus infrastruktúrák	63
6	Összefoglalás	63
7	Függelék	65
A	A MI elmélete	65
A.1	Történeti vonatkozások	65
A.2	A számítás elmélete	67
A.3	A technikai számítási mód	68
A.4	A biológiai számítási mód	69
A.5	Az organikus MI (OI) számítási módja	70
A.6	Idő-ablak	72
A.7	A számító egységek működése	76
A.8	A számító egységek kihasználása	77
B	A technikai MI számítási működése	78
B.1	Szinaptikus művelet	79
B.2	Tanulás	80
B.3	Visszacsatolás	82
B.4	Megvalósítási mód	83
B.5	Átviteli mód	84
B.6	Gyorsítás	89
B.7	Tárolási mód	95
B.8	Intelligencia előállítása	96
C	A MI és TI működésének összehasonlítása	96

C.1	A jellemző paraméterek megtalálása	98
C.2	Lényeg keresés, modellezés	99
C.3	A modell értékelése (validálás)	101
D	https://chat.openai.com '23. szept. 19	102
D.1	Az intelligenciáról	103
D.2	Működés	110
D.3	Értelmezés és szabályozás	131
D.4	Alkalmazás	138
D.5	Tudomány	140

Irodalomjegyzék	143
------------------------	------------

Ábrajegyzék

1.1	A nagy nyelvi modelleket tárgyaló tudományos közlemények számának alakulása az idő függvényében [47]	4
1.2	A MI hullámainak története a jellemző fejlesztési irányokkal (https://jaylatta.net/history-of-ai-from-winter-to-winter/)	8
2.1	Különböző MI területek helyzete a technológiai várakozás görbéjén [25]	29
3.1	A válások számának összefüggése a margarin fogyasztással[1]	37
3.2	A MI indoklásának elvi ábrája, a használt módszerek és értékelési eljárások kölcsönhatásának diagram-szerű ábrázolása [143]	39
5.1	A számítási igény növekedése és annak kielégítése, különböző számítási technológiák korában [101]	57
5.2	A gépi tanulás energia fogyasztásának várható alakulása (AMD) [128].	59
5.3	Egy Google Cloud Platform energia fogyasztása a kiszolgált kérések számának függvényében [93]	61
7.1	A processzor diszperzió (az órajel kihasználtsága) mértéke a számítógép processzor fejlődésének története során. Saját szerkesztésű ábra.	74
7.2	A mélyrétegű tanulás (deep learning) során alkalmazott számítási módszer, visszacsatolással [89]	79
7.3	A mély rétegű tanulás időviszonyainak "sorba fejtése" [89]	82
7.4	A busz használatának hatása a mesterséges neurális hálózat működésére. Bal oldalt: a matematikai és biológiai, valamint a technikai neurális hálózat elemeinek összekötése. Jobb oldalt: a megosztott busz használatának sorba rendező és teljesítőképesség határoló hatása. Saját szerkesztésű ábra.	85
7.5	A Frontier szuperszámítógép bemutatásakor készült ernyőkép: a számítással egyidejűleg elkészülő 7M processzor adatátvitele a szekvenciális buszon keresztül 40 percig tart[73].	87
7.6	A szükséges és szükségtelen billenések száma az összeadó áramkör bit hosszúságának függvényében. Saját szerk. ábra.	88
7.7	A neurális hálózat "ritkítása" (pruning). Eltávolíthatunk vagy teljes neuronokat vagy a közöttük levő összeköttetések egy részét.[66]	92

7.8	A futási idő és a pontosság összefüggése. (a) az egyes kör és háromszög párok az egyes λ_s paraméterekhez tartozó hibát és a futási időt jelentik. λ_s növelésével ugyan csökken a futási idő, de romlik a megbízhatóság. (b) Ugyanaz a modell különböző λ_v értékekkel tanítva. A diagram vonalak színe a paraméter értékével változik. ([40] 5. sz. ábrája)	94
-----	---	----

Kulcsszavak

- absztrakt
 - gondolkodás, 15
 - indoklás, 7
 - mennyiség, 96
 - módszer, 17
 - neuron, 76
 - processzor, 65, 66, 89
 - szint, 73
- adat megbízhatóság, 17
- adat-vezérelt, 19, 34, 48, 51, 68, 69, 72, 77, 78, 83, 86, 100
- AGI, artificial general intelligence, 15, 19, 30
- agy szimulálás, 34, 87
- ANI, artificial narrow intelligence, 15, 19, 30
- aszinkron, 70, 74, 89
- biológiai számítás, 66, 67
- biztonság, 14, 34, 38, 53, 62, 111, 129, 133, 134, 138, 139
- busz
 - arbitráció, 85
 - szekvenciális, 88
 - sávszélesség, 86
- ChatGPT, 2, 3, 8, 10, 12–14, 16, 18, 20, 24, 29–31, 33, 43, 50, 54–56, 64, 90, 92, 102, 107, 132, 134
 - hőmérséklet paraméter, 30
- COVID-19, 27
- data-flow, 68
- deep learning, 75, 79, 86, 91
- diagnózis, 27, 28, 33, 40, 45, 48, 92, 131, 138
- diszperzió, 73, 74, 95
- döntés, 12, 14, 18, 23, 24, 32, 35, 39, 40, 44, 46, 48, 98, 100, 106, 110, 130, 131, 138, 141, 142
- energia, 2, 3, 16, 43, 52, 55, 59–61, 64, 76, 83, 91, 95, 96, 135–137
- esemény-vezérelt, 78
- etika, 31, 56, 71, 111, 112, 133, 134, 138
- EU, 2, 8, 12, 13, 15, 19, 25, 27, 29, 31–35, 44, 53–56, 58, 60, 62–64, 111, 132–134
- felelős, 1, 13, 33, 51–53, 55, 62, 98, 112, 113, 126, 130, 133, 136
- felügyelet, 34, 52, 54, 63, 80, 103, 124, 126, 127, 138, 139
- fenntarthatóság, 2, 3, 17, 47, 55, 56, 58, 60, 61, 63, 103, 135–137
- Feynman, R. P., 12, 19, 66
- fizikai gondolkodás, 96, 99
- GPU, 56, 75, 79
- Hype (Technológiai várakozás), 2, 4, 11, 27, 29, 54
- IA, Intelligence Augmentation, 19
- idő-ablak, 18, 72, 74, 76, 86, 90
- időztési viszonyok, 65, 66
- instruction-flow, 68
- Intelligencia
 - meghatározása, 14
- intuíció*, 18, 24, 54, 74, 98, 99, 142

jog, 13, 53, 55, 92, 111, 112, 126, 133, 138, 139

kockázatosság, 48

kommunikáció, 7, 64, 67–69, 72, 75, 78, 81, 83, 86, 87, 89, 91, 135

kontextus váltás, 73, 75, 87

kreativitás, 30

kölcsönhatási sebesség, 25, 66, 76, 89

környezet, 2, 55, 58–60, 102, 135–137

LLM, 3, 5, 8, 9, 13, 16, 26, 33, 53, 54, 60, 63, 64

láncolt művelet, 66, 67

megbízható, 6, 7, 17, 33, 34, 36, 39, 40, 43, 47, 50, 62, 63, 70, 94, 102, 110, 113, 114, 124, 126, 130, 139, 141

megvalósíthatósági tanulmány, 3, 31, 34, 65, 95

memória-vezérelt, 86

MI
fogalma, 13, 15
kognitív képesség, 12, 15, 32
típusai, 19

Neumann János, 48, 65–69, 73, 77, 89

oktatás, 6, 33, 46, 133

operációs rendszer, 51, 73, 87

processzor
szegregált, 73

program
szekvenciális, 73

proper sequencing, 69

számítás
biológiai, 66, 69
klasszikus paradigma, 6, 65, 67, 74, 76
MI, 78
modern paradigma, 67
technikai, 68
ál-MI, 83

szénlábnyom, 2, 3, 59, 136

tanulás
biológiai, 17, 69, 71, 72, 74, 80
emberi, 6, 31, 46, 71
gépi, 3, 15, 18, 20, 23–25, 27, 29, 35, 47, 55, 56, 58–60, 63, 65, 71, 72, 79, 81, 83, 84, 96, 99, 101, 103, 104, 110, 124, 130, 131, 141
Hebb, 80, 81
inkrementális, 18, 96
mély rétegű, 82, 140
nem-inkrementális, 18, 96, 100

teljesítmény
elektromos, 57, 73, 88, 89, 95
MI, 11, 109
számítási, 57, 58, 72
TI, 13, 109

Turing teszt, 16, 17, 19, 28, 41, 109

tároló
átmeneti, 7, 19, 44, 59, 61, 68, 70, 72, 78, 93, 95

társadalom, 11, 27, 28, 32, 34, 51, 54, 55, 111, 124, 127, 133

tér-idő, 67

utasítás-vezérelt, 34, 48, 51, 68, 69, 72, 76–78, 83, 86

zöld technológia, 2, 55, 56, 58–60, 137

átláthatóság, 6, 31, 45, 112, 132,
133
órajel

elosztás, 66
késleltetés, 66

1 Bevezetés

Napjaink egyik varázsszava mindannyiunk számára, függetlenül attól hogy számítógépes feladatokkal (esetleg ezen belül mesterséges intelligenciával) foglalkozó szakember, társ-tudományi szakember, az alkalmazása iránt érdeklődő felelős szakember, érdeklődő kívülálló vagy az “utca embere” vagyunk, a Mesterséges Intelligencia (MI), ami az angol nyelvű Artificial Intelligence (AI) kifejezés tükörfordítása (egy másik, eléggé átfedő dialektusa a Computational Intelligence [33], CI). A pályázatban mindkét rövidítést alkalmazni fogjuk, továbbá használjuk a Természetes Intelligencia (TI) rövidítést is összehasonlításakor. Mivel a pályázat magyar nyelvű a felhasznált szakirodalom pedig szinte kizárólagosan angol nyelvű, a szakirodalmi hivatkozásoknál megadjuk a szokásos módon a hivatkozott szöveg helyét és a vonatkozó rész rövidített és lényegi magyar fordítását. A pályázat az említett módon idézett részekről (valamint a saját szakirodalmi közlésekben foglaltaktól) eltekintve eredeti gondolatokat és eredményeket tartalmaz, az említett megoldás kizárólag az angol nyelvet nem kellően ismerő olvasók számára szolgál segítségül és teljesebbé teszi a tanulmányt. Az összefüggésekre kereszt-hivatkozásokkal hívjuk fel a figyelmet, ezért a dokumentum célszerűen elektronikusan olvasandó.

1.1 Kiírás és MI stratégia

A pályázati kiírás szerint a pályaművek “elsősorban ... ne műszaki és matematikai megközelítésű tartalmat tükrözzenek”. A konkrét megvalósításokat csak konkrét (matematikai, informatikai, elektronikai és újabban biológiai) megoldások használatával lehetséges használni, ezen megoldások jelentős korlátozásokat szabnak mind a megvalósítás mind a felhasználás számára; sőt, a megvalósítás döntő módon befolyásolja a tapasztalható tulajdonságokat. Figyelembe kellene venni R. P. Feynman fél évszázada leírt és azóta mindig beigazolódott figyelmeztetését: “A számítógéptudomány abban is különbözik a fizikától, hogy valójában nem tudomány: nem természeti objektumokat tanulmányoz. De nem is, mint sokan gondolják, matematika; bár kiterjedten használ matematikai indoklásokat. Valójában a számítógéptudomány olyan, mint egy mérnöki terület: készítünk valamit és azzal valamit elvégeztetünk, és nem csupán absztrakciókkal foglalkozunk.” [62] (A szerző tudatosan nem a számítástudomány szót használta.)

Nem lehet a pályázatban említett óvatossági és hihetőségi információkat, valamint hibákat és sebezhetőségeket sem megérteni az elméleti háttér tárgyalása nélkül; különösen, mivel a technológiai fejlődés iránya egyre

kiszámíthatatlanabb és kockázatosabb. Figyelembe kell venni energetikai és környezetszennyezési szempontokat, mert az EU 2023-as álláspontja szerint [28] is: “A mesterségesintelligencia-technológiák és adatközpontok nagy szénlábnyommal rendelkeznek a megnövekedett számítási energiafogyasztás, valamint a tárolt adatok mennyisége és a keletkezett hő-, elektromos és elektronikus hulladék mennyisége által előidézett magas energiaköltségek miatt, *így még nagyobb lesz a szennyezés.*” Pedig az már a MI mostani felfutása előtt is magas volt [106].

A D.1 függelékben a ChatGPT korrektül fogalmaz önmagáról. “Én egy *gépi intelligencia vagyok*, amely mesterséges intelligencia alapokon működik”: azaz egy műszaki megvalósítás. Ha viszont a technikai rendszerként általa okozott károkról (energia felhasználás és környezet szennyezés) kérdezik, akkor megtagadja és csupán egy szoftvernek minősíti önmagát. Arról meg hallani sem akar, hogy “zöld” technológiának tekinti-e magát, csupán mellébeszél; lásd D.3 függelék. Úgy tűnik, az adatkészlet összeállítóinak nem állt érdekében erre kérdésre érdemi választ adni, bár óvatos formában beismerik az energia felhasználási, fenntarthatósági, környezet szennyezési és szénlábnyom problémákat.

Ennek megfelelően módosítani kell a (nyilván a 2018-as EU álláspont szellemében született) magyar stratégiát: “hogyan a mesterséges intelligencia – mint a digitalizációs törekvések egy speciális ága – általános fejlesztési céljaink szellemében *magyar, high-tech és zöld legyen*” [12]. A technológia lényeges része az adathalmaz; hogy magyar legyen, megfelelően előkészített, magyar nyelvű adathalmaz kell, amin a rendszert betanítják (a D.2 függelék szerint magyar nyelvű adathalmaz nem szerepel az adatkészletben, csupán magyarra fordított blablát kap a felhasználó). A high-tech feltehetően a legújabb (és minden eddigénél gazdaságatlanabbul számoló, angol nyelvű csevegő szövegek használatával elő-tanított) számítás-gyorsítók alkalmazását jelenti. Hogy egyidejűleg magyar is legyen, ahhoz nem csak magyar szöveggel kell betanítani a rendszert, hanem magyar szövegeken előtanított chipeket kell(ene) gyártatni. Hogy “zöld” legyen, ahhoz a teljes technológiát újra kellene gondolni és alkotni; nem nemzeti szinten, de ennek szükségességére fel kellene és lehetne hívni a figyelmet, továbbá az új működési alapelvek közre bocsátásával elindítani a munkát.

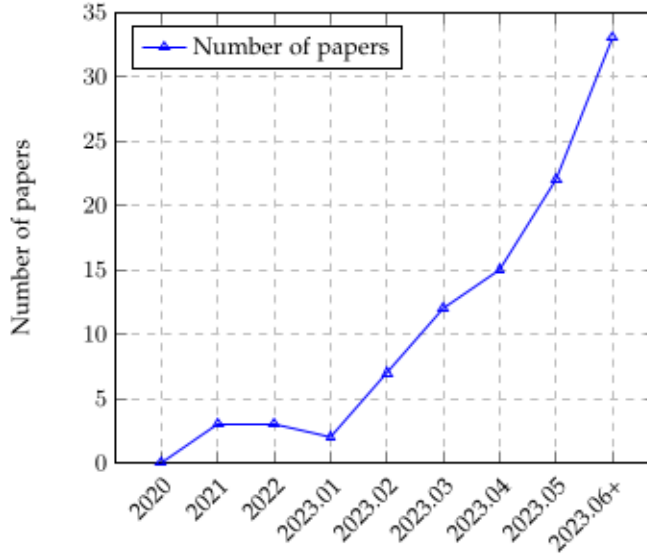
Nem világos, mit takar a központi infrastruktúra. A jelenleg (leginkább szórakozásra) használt infrastruktúra zárt forrású kódot használ zárt forrású angol nyelvű adatokon, magyar ellenőrzésen kívül eső (és méretében a Magyarországon elérhetőnél jelentősen nagyobb) hardveren. A szükséges infrastruktúra része olyan hardver, amelyiken a betanítást el lehet végezni, a magyar adatkészlet megfelelő előkészítése és a MI jelenlegi szintjének

eléréséhez a vezető cégek által használt zárt forrású technológiák újra fejlesztése is. *A jelenlegi adatok szerint a megfelelően előkészített adatkészlettel és a zárt forrású szoftver használati jogával valamint a “know-how”-val már rendelkező cégek esetében egy ilyen fejlesztés elindítása több milliárd USD tőkét igényel*, továbbá kérdéses, hogy befejezhető-e 2030-ig.

A magyar stratégia lehetőségként említi, hogy *a MI csökkentheti bizonyos tevékenységek lábnyomát, annak említése nélkül, hogy az ennek elérésére szolgáló MI technológia használata mennyivel többet növel rajta*. Egyetlen, az újabb keletű gigantomániás szellemben épített rendszer betanítása több tíz tonna szén-dioxid kibocsátásával egyenértékű [93], a GPT-3 több mint ötszázal. Az ilyen rendszerek betanításának energia igénye gigawatt-óra nagyságú és hosszú hetekig/hónapokig tart ([93] 118 napot említ). Nem elegendők a merész célkitűzések és ambíciók; a technológia pontos megértésére és megvalósítására van szükség, a csodavárás nem stratégia. Esetleg a kidolgozók figyelmét fel lehetne hívni pl. arra, a közgazdaságtan szerint a haszon eléréséhez befektetés és termelési költség is szükséges; vagy legalább arra, hogy ingyen ebéd nincs.

Nagyon naiv álláspont, hogy “A mesterséges intelligencia egy globális technológia, határok és *fizikai eszközigény nélkül*” valamint – különösen globális energia válságban szenvedő korunkban – hogy “Ezúttal azonban nincs a fejlődésnek természeti erőforrás-igénye”. Aminek maga a Stratégia szövege is számos ponton ellentmond: “nagy mennyiségű adat erőforrás-igényes elemzése”, “Cél a kutatási és általános fejlesztési feladatok elvégzéséhez szükséges erőforrások központi infrastruktúrával való kiszolgálása, a létrejövő infrastruktúra igény szerinti folyamatos fejlesztése, hosszú távú fenntartása.”

A tanulmány *nem* MI-ellenes, hanem MI-kritikus. Igyekszik feltárni a területtel kapcsolatos felfokozott technológiai várakozással (valamint éppúgy a teljes kiábrándulással) szemben álló másik oldal érveit is; éppen a tárgyilagos értékelhetőség érdekében és *a MI fejlesztések még mindig hiányzó megvalósíthatósági tanulmányának részleges pótlásaként*. Megpróbálja – a teljesség igénye nélkül – a nagyon sokszor nem egyértelmű alapfogalmakat és összefüggéseket definiálni, azokat egységes és viszonylag átlátható keretbe helyezni, a ChatGPT, az LLM, a MI, az automatizálás és a gépi tanulás egymáshoz való viszonyát tisztázni, továbbá a MI-t csupán alkalmazni kívánó szakember számára érthetően megfogalmazni az MI társtudományainak állításait. Rövid és koherens módon tárgyalja a szükséges alapfogalmakat és vezeti be az olvasót az elemeiben ismert fogalmak összefüggő rendszerébe, valamint további tájékozódáshoz ad útmutatást az érdeklődő olvasó számára.



Ábra 1.1: A nagy nyelvi modelleket tárgyaló tudományos közlemények számának alakulása az idő függvényében [47]

1.2 A nagy nyelvi modellek

A MI egyik részterülete, a nagy nyelvi modellek [14] elméleti és gyakorlati vizsgálata szinte példátlan érdeklődést vált ki tudományos szempontból is. A 2023 év eleje óta erről a témáról havonta több cikk jelenik meg, mint korábban évente, lásd 1.1 ábra, és az ütem exponenciálisan növekszik (pedig az ábra nyilván nem tartalmazza a szakterületi felhasználásokról megjelent véleményeket), ami vészhelyzeti reakciókat váltott ki. A közlemények nagyon sokféle szempontból tárgyalják a témát. A [47] közlemény láthatóan igyekszik naprakész összefoglalót nyújtani, pedig ekkora mennyiségű közleményt kritikusan áttekinteni szinte lehetetlen. A hype görbe emelkedő ágában természetesen az egymással versengő újabb sikerek dominálnak és az óvatosságra intő vagy a korlátokra figyelmeztető közleményeknek az MI-hívő bírálókon kell(ene) áttörni a folyóiratokban. A sikertelenségről meg nem születik közlemény. Amint a 2.8.2 szakaszban tárgyaljuk, a valóban komoly használatnál már kezdenek előbukkanni a határok is. A palackból kiengedett szellem azonban nemcsak hogy még mindig növekszik, de sokszorozódik is. Természetesen nem csak szöveget, hanem hangot és képet is lehet analizálni és generálni; így “Large Media Model” néven általánosítódott a fogalom. Sőt, szorgalmas hívő kutatók szerint az LMM betűszó jelentése lehet “Large

Medical Model” is.

Felhasználóinak nagy része számára a LLM olyan érzést biztosít mint az “Aliz kalandjai Csodaországban” (Alice’s Adventures in Wonderland) közismert satirikus fantázia-mese vagy az “Óz, a csodák csodája” (The Wizard of Oz) főhősének az az élmény, hogy egy olyan fantázia-világba kerül, ahol a fantázia-teremtmények és állatok antropomorfizáltak, továbbá a környezet és az események a valós és a szürreális keveréke. A mű az angolul “nonsense literature” (közelítőleg a képtelenség, bolondozás) néven ismert kategóriába tartozik. Jellemzője, hogy *a tökéletesen valós események teljesen ésszerűtlen és valószerűtlen (mese- vagy csoda-szerű) eseményekkel keverednek*, és ahol az értelmetlenség főként a fogalmak jelentésének nyelvi kifacsarásából valamint a logikai indoklások karikatúra-szerű túllépéséből fakad (szándékos eltérés az emberi intelligenciától). A mű a fentiekből eredő intellektuálisan abszurd humor révén hat, erre vezethető vissza óriási népszerűsége. A gyermekkori varázsos képzelet valóságosnak elfogadja, ám a felnőtt csak szórakozásnak tekinti az ábrázolt fantázia világot; számára túl nagy a szürreális részek aránya. A beszélő állatok vagy a Bádogember viselkedését elemző és tovább gondoló gyermekkel elbeszélgetünk és örülünk neki, hogy tudja követni annak sajátos logikáját. Majd kinövi. Ha ugyanezt felnőtt teszi, akkor legalábbis későn érőnek vagy tréfás hajlamúnak tartjuk; egyéb esetben szánakozunk rajta.

Az LLM esetén a program (látszólag) emberi partnerként reagál. Nem véletlenül, hisz *a MI célja* éppen az emberi tevékenység imitálása (más szóval *a számítógép program antropomorfizálása*). A környezetre vonatkozóan a valós elemek (a kérdésektől függő mértékben, de jellemzően) túlsúlyban vannak (a valódi világból származó gigantikus méretű adatkészlet felhasználásának köszönhetően) és éppen az teszi veszélyessé (nehezen észlelhetővé és hihetővé) a valótlan elemeket, hogy *a mind fejlettebb MI algoritmusok a valótlanságokat belesimítják a valóságosak közé, kiküszöbölve a nyelvi és logikai ellentmondásokat. A leggyakrabban előforduló kérdésekre adott válaszok szinte tökéletesek, és az átlagember számára nagyon is hihetően hangzanak. A szürrealitás mértéke pedig a kérdezőtől, a kérdések összeállításától, a rendszer probabilisztikus jellemzőitől függ.* A szürrealitás egyre inkább csak nagyon figyelmes és az illető témában járatos ember számára, célzott kérdés-sorozattal és hosszabb beszélgetés során tűnik fel. A generált tartalom inherens módon folyamatosan változik, ráadásul a felhasználó érdeklődésének irányában; lásd D.1 függelék. *Munka-célú használatra egyszerűen alkalmatlan.* Az LLM esetében a működésnek van egy kaleidoszkóp-jellege is, ami egyszerűen lekötve tartja a figyelmet; még e vonatkozásban is párhuzamban Aliz csoda-világával. *Nem ennek kellene a fejlesztési célnak lenni (de egyre inkább azzá válik, mert*

az eredeti cél, az emberi intelligenciával azonos gépi intelligencia kifejlesztése egyre nyilvánvalóbban nem lehetséges [82]), de ezzel lehet a felhasználót a rendszer használatához kötni; a használati idő pedig (közvetve) profitot jelent.

Kiemelkedően veszélyes, hogy a gyermekeket szürreális világban tartja (alkalmazkodik a gyermek ismeret- és képzelet világához, és ennek megfelelő “új” ismerettel látja el őket), az oktatásban helyettük megoldja a feladatot, jól hangzó de hibás megoldást mutat (ezzel gyakorlatilag akadályozza a tanulást, mivel azt a keresés-másolás-benyújtás tevékenységekre egyszerűsíti), valamint hogy az algoritmusok működésének javításával és a betanító adatmennyiség növekedésével az elvileg intellektuálisan felnőtt lakosság mind nagyobb részét meghódítja (rongálja ismeret anyagát és természetes intellektusát; hamis tömeg ízlést, viselkedést és ismeretet hoz létre). Megint csak az Aliz szempontja alapján: akár jó is lehet szórakozni a csodavilágban, csak nem kell összetéveszteni a valódival, ahol élnünk és dolgoznunk kell.

Sokkal nagyobb baj, hogy a sokkal nagyobb hatáskörrel mint hozzáértéssel rendelkező hivatalnokok erőszakkal kényszerítik az embereket a MI használatára; úgy hogy mind az irányító algoritmusok mind a betanításhoz felhasznált adatok zárt forrásúak, céges érdekekkel rendelkező informatikai óriások ellenőrzése alatt, akik saját bevallásuk szerint sem tudják, hogy az általuk létrehozott rendszer hogyan működik, fejlődik és milyen célú használata milyen hatásokkal járhat. Amin nem sokat változtat, ha az állam ellenőrzése alá kerül, de zárt forrású marad. Érdekes, hogy ez a szempont nem merül fel, amikor átláthatóságról és megbízhatóságról beszélnek.

1.3 A technikai megvalósítás hatásai

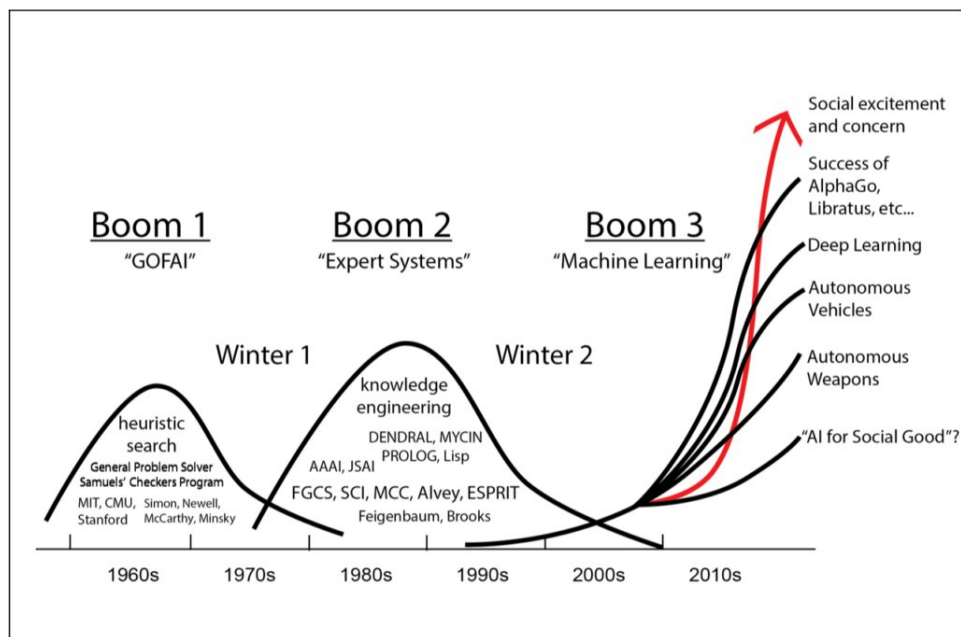
Az értelmetlenül nagy méretűre nőtt (176 milliárd paraméteres [93]; Kínában ennél tízszer nagyobb is van) rendszerek tulajdonságait és viselkedését az elvi és technikai működés ismeretének teljes hiányában az azt használni kívánó szakértők csak empirikusan tudják vizsgálni. A káosz akkora, hogy a kognitív tudomány az AI rövidítést az “Alien Intelligence” (földön kívüli intelligencia) jelentéssel használja [63]. Nem csoda: az alapul szolgáló utasításvezérelt elektronikai számítási rendszerek elvi működésének leírására már évtizedek óta nem alkalmas a számítási paradigma, az elvileg megcélzott adatvezérelt rendszereknek (és a működés gyorsítására használatos elektronikus és szoftveres megoldásoknak) egyáltalán nincs működési elmélete, az ilyen méretű rendszerek működtetésére már nem alkalmas technikai megoldások miatt “fuldokolnak” a rendszerek, az átgondolatlan gyorsítási elvű megoldásokat a hiányosságok elleplezésére vonatkozó technológiai megoldásokkal kombinálva valóban tökéletesen áttekinthetetlen a rendszerek elvi működése

és tulajdonságaik legalább utólagos/tapasztalati értelmezése. A szakértők nem tudnak és nem mernek egyértelmű állásfoglalást mondani. Egy jellemző hideg/meleg vélemény: *“Ezek a rendszerek semmiféle értelemben nem olyan megbízhatóak és általánosak, mint szeretnénk, továbbá valószínűleg vannak teljesen hiányzó absztrakt indoklási képességeik. ... De azt hiszem, az alap képességei megvannak.”*[19]

A józan hangok nem hallatszanak el odáig, ahol a nagy rendszerekről (rendszerint presztizs és profit szempontok alapján) döntenek. Bár ismert csupán a józan ész alapján, hogy a nagyobb (több paraméteres) rendszert több adattal kellene betanítani, a nagy versenyzők nem tartják be még az egyszerű skálázási szabályt sem, mivel a betanítás drága és sok idő múlva fejeződik be; bár ismert, hogy a “jól betanított” kisebb rendszerek jobban működnek, mint a nagyok [71]. Pedig érdemes lenne végig gondolni, hogy ha egy paraméter felhasználásának idejét (a súlyozott összeg kiszámítását, csupán a memória kezelésének és a kiszámításnak az idejét alapul véve, az adatátvitelt teljesen elhanyagolva) 10 ns-nak (10^{-8} sec) vesszük, és az említett 100 milliárd (10^{11}) paraméteres modell valamennyi paraméterének szerepe van, akkor egyetlen neurális számítás is percekig tartana. Kommunikáció használata esetén pedig évekig. Az ismeretlen működésű (és megbízhatóságú) hardver/szoftver gyorsítók, átmeneti tárolók, stb. használata segítségével azonban másodpercek alatt megszületik a válasz. Olyan, amilyet ennyi idő alatt lehet összeállítani (természetesen előre összerakott építőelemekből). (A gyorsítás technikai lehetőségeire és veszélyeire vonatkozóan lásd a B.6 függelékét). Ha valóban pontosnak kell lenni a válasznak (nem az embert kell becsapni, hanem a természet várható viselkedését előre jelezni vagy számítógép számára pontos tervet produkálni), akkor a módszer nem is működik, lásd 2.8.2 szakasz. Azt már felismerték, hogy egy ilyen válaszhoz a góliátok (és valójában a dávidok is) az általuk elérhető súlyoknak csupán töredékét használják és erre hangzatos, de megalapozatlan modelleket tesznek közzé [61]. Ennek alapján megszületett a Small Language Model (SLM)[83] fogalma is (itt a kicsi “mindössze” milliányi paramétert jelent). Csupán a megértés hiányzik még.

1.4 A kiábrándulás

A pályamű írása közben azonban hetek alatt elindult a gyors kiábrándulás; az előző két nagy MI hype alapján törvényszerűen. A gyakorlati (pénzügyi, tudományos, technológiai, stb) hasznot remélt gazdasági szereplők csalódtak. Az ilyen szempontokat nem tekintő, csupán szórakozni vágyó vagy házi feladatának megoldási kötelezettségét teljesíteni akaró, stb. nagyközönséget



Ábra 1.2: A MI hullámainak története a jellemző fejlesztési irányokkal
(<https://jaylatta.net/history-of-ai-from-winter-to-winter/>)

azonban még hatása alatt tartja a varázs, és a MI alkalmazásába befektető gazdasági szereplők futnak a pénzük után. A 2023 év őszén már a közelítő MI-térről beszélnek[29, 51]. Bár nyilván le kell vonni belőle a versenytárs érdekeltségének járulékát, de igaz hogy [51] “Manapság a nagy MI cégek a ’nyers erő’ módszerét tekintik nyerőnek. A gépi tanulásban óriási adat-mennyiség és feldolgozó kapacitás használata sem nem hatékony sem nem jövő-biztos. A MI-nak sokkal okosabbá és nagyságrendekkel hatékonyabbá kell válni, ha el akarjuk kerülni a harmadik MI ’telet’”. Nem új meglátás; amint a 1.2 ábra mutatja, ez lesz a harmadik tél, és nem tanultunk az előző kettőből. Annak ellenére, hogy az első MI-télhez vezető kutatási áttekintés [133] már fél évszázada rögzítette az alapigazságot: a technikailag megvalósítható *gépi intelligencia* (bármilyen nagy mértékben tökéletesített eszközökkel megvalósítva is) *csupán nevében mutat hasonlóságot az emberi intelligenciához; ezért a kutatások nagy része megvalósíthatatlan célt tűz ki*.

A kiábrándulás oka lényegében a(z LLM alapú) ChatGPT, de a fogalmi tisztázatlanság miatt az egész MI került gyanúba. Nem kellene azonban a fürdővízzel együtt a gyermeket is kiönteni. A betiltás után az EU drasztikusan (bár megértés hiányában, igen rosszul) módosította a MI-val kapcsos-

latos álláspontját, ami nyilván maga után vonja Magyarország Mesterséges Intelligencia Stratégiájának hasonló értelmű módosítását. A technológiai várakozás görbéjének szokásos menete (lásd 2.1 ábra) szerint néhány hónap múlva következhet(ett volna) a konszolidált használat időszaka. *Ebben – a MI mindeddig hiányzó megértése után – Magyarország az élre állhat.* Nagy kérdés, hogy a fejlesztők önként vállalt önkorlátozása és a hatóságok addig meghozott szabályozásai után hogyan folytatódhat a történet. Annak, hogy az említettek megértenék, hogy mit fejlesztenek és szabályoznak, semmi jele.

A pályázat kiírása óta már megtalálták az új reménysugarat, lásd A.5 szakasz, ami még fel sem került a technológiai várakozás görbéjére. Az organoid intelligencia (OI) az emberi intelligencia másfajta megvalósításának kísérlete, ezúttal sejtbiológiai eszközökkel. Részben az MI zsákutcába jutása miatt, a felfokozott várakozás nagy része most az OI felhasználása felé irányul. A tisztán technológiai kihívás (hogyan lehet összekapcsolni elektronikus és biológiai információs rendszereket) miatt még eltart pár évig, amíg ennek az új ötletnek a megvalósítása eljut a produktivitásnak abba a fázisába, ahol a MI 2023 elején volt. Ezután fog tudatosulni, hogy a biológiai és az elektronikus számítás működése alapjaiban eltér. Mint a A.5 szakaszban tárgyaljuk, az OI nagyon sok vonatkozásban ugyanúgy fog viselkedni, mint a MI és egy évtized múlva egy hasonló pályázatot kell majd kiírni az OI alkalmazásával kapcsolatos problémákról, ha a világ meg nem érti miben különbözik az adatvezérelt és az utasítás-vezérelt számítási mód. Ilyen értelemben a tanulmány évekkal megelőzi korát. Tudomásunk szerint az egyetlen ilyen jellegű elméleti áttekintés.

1.5 A dokumentum felépítése

A dokumentum fő fejezetei a szükséges ismereteket foglalják össze az MI (és főleg az LLM) alkalmazása iránt érdeklődők és azt gyakorlati alkalmazások céljára felelősséggel bevezetni kívánók számára, akik nem tudnak kiigazodni a MI rajongók és ellenzők szélsőséges állításai között. A téma túl terjedelmes ahhoz, hogy a dokumentum teljességre törekedhessen. Az alkalmazási lehetőségek köre túl széles, hogy azokat konkrétan megtárgyaljuk, ráadásul napról napra fejlődik vagy legalábbis változik. A MI és az LLM felhasználói számára követhető stratégia, hogy nem a tanulmány általános értékelési szempontjaira hallgatnak (de azért azokat alaposan megértik és használják), hanem a saját körülményeikre vonatkozóan célzatosan “benchmark”-ot készítenek, és saját tapasztalatukból (sok munkával) döntenek el, hogy ott és úgy (értve alatta: számukra) alkalmazható-e. Amit nagyon célszerű időnként megismételni, mivel a rendszerek folyamatos változásban

(ami lehet előnyös is és előnytelen is) vannak; továbbá a felhasználó tapasztalatai is növekszenek. A MI hívók sokszor tapasztalt (a marketinges “vezetői összefoglaló” anyagokon alapuló) csodavárásának helyes értékelésére rávilágít a 2.8 szakasz. A fejezetek igyekeznek pontosabbá tenni a nagyon nagyvonalúan (az értés minden jele nélkül) használt fogalmakat; magyarázni a MI legjellemzőbb hibáit a technikai háttér alapján; rámutatni pár használati esettanulmány alapján a MI használat korlátaira, beleértve a természet-szabta MI teljesítőképesség vizsgálatot is; valamint rámutat a MI általában nem tárgyalt, technikai eredetű veszélyeire is.

Mivel ez a felépítmény nem független az informatika általános fogalmaitól és tulajdonságaitól, legalább felületesen a nem-informatikus olvasóknak is ismernie kell azt, a vonatkozó elméleti háttér megértése az informatikai szakemberek számára pedig kötelező. Ezért (rövid összefoglalóként) az az A függelékben megtalálható; kemény és szokatlan anyag, a számítástudomány művelői számára is. Ezen belül található a biológiai és technikai számítás egységes elméletének vázlatos bevezetése, aminek alapján gondolatilag közelebb lehet jutni a természetes és mesterséges neurális működés különbségének megértéséhez. A B függelék a mérnöki alapképzettségűek számára mutatja meg, miért nem alkalmasak a mai technológiai körülmények között a több évtizeddel ezelőtt kitalált technológiai megoldások. A TI és (a valódi) MI egyaránt adat-vezérelt működésű; a természettudományos (ezáltal elkülönítve a matematikaitól) háttérrel rendelkező olvasó számára a C függelékben bemutatott összehasonlításuk sokat segíthet a MI működésének megértésében. A három említett függelék lényegében önálló; az említett olvasókör számára nagyon is hasznos lehet. A fő szöveg azonban enélkül is érhető; főként a MI világában eligazodni kívánó és nem azt fejleszteni vagy technikai működését megérteni akaró olvasó számára szól. A D függelék a ChatGPT célzatos faggatásának néhány eredményét tartalmazza. Önállóan is használható (és saját kérdés kombinációk feltevésére biztat), az egyes pontokra a fő szöveg kereszt-hivatkozásokat tartalmaz.

2 A MI általában

A számítógépet és az agyat tanulmányozó tudományok egymást nem mindig előnyösen befolyásoló kölcsönhatásban állnak egymással [95, 67]. Mindkét területen használják a társ-tudományok: például a matematika, a fizika, az informatika, az elektronika, a számítógép tudomány, az ideg-tudományok eredményeit. Mindegyik terület képviselői saját területük elsőrangú szakemberei és ezt feltételezik a vonatkozó társ-tudományok képviselőiről is. Min-

den tudomány valamilyen közelítéseket használ a természet jelenségeinek leírására és ezeknek a közelítéseknek (sokszor csak implicit) érvényességi köre van, ami már sokszor a társtudományok képviselőinek figyelmén kívül marad, így esetlegesen a társtudományok eredményeit és következtetéseit olyan területekre is alkalmazni akarják, ahol a közelítések már nem alkalmazhatók. Egyik jellemző példája éppen a számítógéptudomány. A MI ezeknek a társtudományoknak egymásra hatásából született; mára (legalábbis méreteit és kutatási intenzitását tekintve) szinte önálló tudományterületté vált és magán viseli az említett “túl-használat” jellemzőit.

2.1 A MI története

A MI több hullámban is megszületett (minden szakember generáció újra felfedezte és nagyrészt függetlenül megalkotta, lásd 1.2 ábra). Az ábra 2021-ben, még jóval a 2023-as csalódás előtt született, de címe már előre vetítette azt: a címben említett “tél” a technikában való csalódást, meghíúsult reményeket és természetesen a támogatások (valamint a profit) elapadását jelenti. Megelőlegezte a 2023-ban bekövetkezett változást. Ha a 2.1 ábrával összevetjük, azonosíthatjuk a “felfokozott várakozás” csúcsait, amit törvényszerűen követnek a csalódás hullámvölgyei. Természetesen eltérő technikai és társadalmi körülmények között, az előző hullámból származó “konszolidált használat” csúcsa fölött.

A MI első két hullámának kifulladását a társtudományoknak tulajdonították (nem elég fejlett elektronika, nem eléggé jó matematikai módszerek, az agy működésének nem kellő szintű ismerete, stb). A mostani MI hullám már teljesen más hátteret talált: minden területen jelentős fejlődés következett be. Az általános számítógépesítés miatt megnőtt a számítógépes feldolgozás elterjedtsége és szerepe, kezelhetetlenül nagyra duzzadt a számítógépekkel termelt és általuk feldolgozandó adatmennyiség, továbbá az Internet elterjedtsége miatt jelentősen megnőtt eme adatok elérhetősége és a társadalom (beleértve a társtudományokat is) várakozása a MI elvárt jövőbeli teljesítményére vonatkozóan. Ne feledjük el a kockázati tőke profit várakozását sem.

A MI mostani hulláma kifulladásának 2023 közepén vagyunk tanúi. A felfokozott várakozás visszaütött. Az “intelligencia” szóhoz fűződő képzetek, párosulva a szakterületek legalábbis merész jövőképeivel, irreális elvárásokat, reményeket és félelmeket generáltak. Mára már a napi sajtó is foglalkozik a felfokozott várakozásból eredő csalódottsággal. A kezdeményezők, támogatók, javaslattevők, gyártók, fejlesztők, stb. egymással versengve igyekeznek kihátrálni a MI mögül; azért ügyelve hogy ne veszítsék el az arcukat. Egy dolog nem változott a MI körül: *nem értik, valójában mi is, mire való, hogyan*

működik, mire jó és mire nem jó a MI. A tanulmány már csak korlátozott terjedelme miatt is csak kevésbé részletezheti a tulajdonságokat.

2.2 A MI fogalma

A MI neve lényegében azokat az elvárásokat tükrözi, amiket első kutatói az '50-es években feltételeztek képességeinek határaitól: az emberi intelligencia gépi megvalósítása, főként gondolkodás, *magas szintű kognitív képességek modellezése volt a cél* (human-imitating AI) [82], a társtudományok eredményeinek felhasználásával. A vonatkozó társtudományok azonban csak alacsonyabb (nem kognitív) szinten foglalkoztak jelfeldolgozással és döntés hozatallal. Amint a kiírás is említi, rögtön az elektronikus számítógép megalkotása után megszülettek a túlzó elképzelések a *“gondolkozó, tanuló és kreatív gépek”* építésére vonatkozóan, bár a MI megalkotói *pszichológiai modellezésre és nem mérnöki eszközökkel létrehozott berendezések készítésére szánták* (a más tudományterületen történő “túlhasználat” újabb példája). Még arra vonatkozóan is eléggé korán megfogalmazódtak kérdések [141], hogy mesterséges intelligencia valóban olyan jellegű-e, mint amelyet az emberi agy is használ problémák megoldására.

A D.1 függelék szerint, a szakértőktől (beleértve az EU rendeletekben kifejtett álláspontokat is) eltérően, *a ChatGPT tagadja, hogy a MI tud gondolkodni.* Pedig az adathalmazát megírt fejlesztőknek éppen az ellenkező állítás szolgálná az érdekeit. *Figyelemre méltó, hogy a stratégiát összeállító és a technológia használatát szabályozni kívánó szakértők, politikusok és jogászok milyen kevésbé ismerik a technológia elméleti és gyakorlati tulajdonságait. Megelégszenek a cégek által készített, a vezetőknek szánt hangos marketing anyagok ismeretével.*

Bár a '70-es években, a már kicsit haladottabb számítógéptudomány korszakában Feynman felhívta a figyelmet arra, hogy jobb lenne az “haladó algoritmusok” (advanced algorithms) [62] kifejezést használni a területre, a kifejezés helytelen használata annyira elterjedt, hogy a MI robbanásszerű felfutása idején vezető MI elméleti tudós is kifakadt ellene: “Ne hívjunk mindent MI-nek” továbbá “valójában miért nem intelligensek a mai MI rendszerek” [121]. A kapcsolat fordítva is igaz: a MI módszereinek biológiai relevanciája legalábbis kérdéses [95]. *“Hat évtizeddel később a kitűzött kognitív képességek még mindig elérhetetlennek tűnnek. . . . A rosszul megválasztott elnevezés megakadályozza a vonatkozó intellektuális és felhasználási problémák megértését.”* [82] Mindennek ellenére, az elterjedésnek immár hagyományteremtő ereje van, pedig a valódi ok, amint egy újságíró a Google fejlesztői konferenciájával kapcsolatban megfogalmazta [16]: “a MI gyorsan olyan

marketing kifejezéssé vált, ami csak annyit jelent ’számítógépek által végzett számítógépes tevékenység’.

Az utóbbi néhány hónapban a MI jelentése egyre inkább a nagy nyelvi modellekre (LLM) szűkül; a ChatGPT alapú fejlesztések látszólagos (és marketing fogásokkal támogatott) sikerei és egyre nyilvánvalóbb veszélyei a MI összes alkalmazási területét veszélybe sodorják és legalábbis gyanússá teszik. A fejlesztők egyre inkább megértik az általuk létrehozott fejlesztésben rejlő technológiai fenyegetést, és állami/hatósági jogi szabályozási beavatkozást *sürgetnek*. Továbbá, nem hoznak nyilvánosságra újabb fejlesztéseket. Háttha a károkozásért utólag is felelősségre lehet őket vonni. Áprilistól szeptember végéig csend volt. Lehet, hogy az OpenAI fejlesztői megértették, amit beleírtak a GPT-3.5 betanító adathalmazába, hogy t.i. *a gépi és az emberi intelligencia gyökereikben eltérnek egymástól* (lásd D.1 függelék)? Vagy, igyekeznek a felelősséget (legalább részben és legalább utólag) egy kormányzati ügynökségre hárítani?

Úgy tűnik, a GPT-4 2023 áprilisban bejelentett új fejlesztése sem hozott áttörést. *Tökéletesebb gépi intelligencia, de nem közelített az emberi intelligenciához*. Sok milliárd dolláros üzleti érdekek állnak a háttérben, így aztán megfelelően táltalt eredményeket kell közzé tenni. Bennük található az igazság csírája is és annak beismerése is, hogy nem a kitűzött cél felé haladunk. Bár a címe szerint felcsillan benne az AGI, *fejlesztőinek véleménye* szerint *“a GPT-4 haladt az AGI felé, de ez nem jelenti azt, hogy tökéletesen működik vagy hogy közel áll ahhoz, amit egy ember meg tud tenni”* [42]. Az egyik, az új fejlesztést (eddig még nem használt módon megszövegezett, tehát a betanításhoz használt szöveg halmazban nem szereplő) orvosi vizsgakérdések megválaszolásával vizsgáztató közlemény [125] mindkét pontban alátámasztja az állításokat: *a GPT-4 lényegesen jobban vizsgázott, mint a GPT-3.5*, de még mindig alatta marad *egy átlagos emberi orvostanhallgató* teljesítményének. Ami nem áll ellentétben a GPT-4 technikai bejelentésével [116], miszerint felülmúlja *az emberek többségének* vizsgaképességét. *Az emberek többsége* valóban rosszul teljesítene az *orvostanhallgatók* tudásának mérésére szolgáló teszten. Nagyon oda kell figyelni a megfogalmazásra és a közlemény íróinak érdekeltségeire.

2.3 A MI meghatározása

Az EU [11] 2018-ban így definiálta: “A mesterséges intelligencia intelligens viselkedésre utaló rendszereket takar, amelyek konkrét célok eléréséhez elemzik környezetüket és – bizonyos mértékű autonómiával – intézkedéseket hajtanak végre”. (A magyar definíció lényegében ezzel megegyezik; a továbblépés

érdekében átgondolandó és átfogalmazandó.) Eme meghatározás szerint sok száz év óta, naponta használt MI rendszer például a vízszint szabályzó, aminek konkrét célja a vízszint állandó értéken tartása és a környezet folyamatos elemzése során a vízszint csökkenésének észlelésekor autonóm módon intézkedik: kinyit egy szelepet és vizet enged a tartályba, majd a vízszint helyreállítása után autonóm módon elzárja a szelepet. Vegyük észre, hogy ez a meghatározás lényegében az első MI hype-ot lezáró kutatási áttekintés [133] “A” (azaz “Advanced Automation”) fogalmával egyezik meg, annak említése nélkül.

A lényegét jobban megfogja [60] definíciója, bár az is az emberi intelligencia közelebbről nem definált fogalmához vezet vissza: “A MI abban különbözik más berendezésektől, amelyeket emberi tevékenységek javítására vagy helyettesítésére terveztek, hogy *nem csak rutín vagy ismétlődő tevékenységek megoldásán alapszik, hanem az intelligencia-szintű emberi viselkedésen is.*” Ez utóbbi meghatározás – lényegében helyesen – kizárja a MI fogalomköréből az egyszerű (főként visszacsatoláson alapuló) szabályozó és vezérlő köröket, az egyszerű osztályozási és minta felismerési eseteket, bonyolult döntési fák és optimalizálási esetek, statisztikai számítások és mintázatok keresése alapján hozott döntéseket, idősorok vizsgálata alapján megalapozott előrejelzések készítését, stb., viszont fontosnak tartja az emberi viselkedés imitálását. Megfontolandó lenne [37] meghatározását használni: “az intelligencia egyrészt lehetővé teszi külső ingerekre megfelelő választ adását, másrészt további képességek megszerzését”. A meghatározás második fele alapján viszonylag könnyen el tudjuk dönteni, hogy intelligens-e szóban forgó rendszer.

A MI preferálása (azaz a fogalmak említett félreértése) kötelezővé tette az egyszerű automatikus szabályozás átminősítését *gyenge* MI-vé, mivel csak ekkor járt rá támogatás: “sok mezőgazdasági üzem alkalmaz mesterséges intelligenciát az állatok mozgásának, testhőmérsékletének és takarmányfogyasztásának a figyelemmel kísérésére. A mesterséges intelligencián alapuló rendszer automatikusan szabályozza a fűtőberendezéseket és etetőgépeket”. Az átminősítés jelentősen hozzájárult a fogalmak összekuszálásához.

A MI technológiai várakozás gyors felfutása idején kiadott 2018-as meghatározás közvetlenül ezután viszont már egyértelműen az *erős MI* témakörébe (avagy [133] jelentés szerint “B” kategóriába) tartozó példákat említ, a hívók erősen csoda-váró hangulatában: “a mesterséges intelligencia segít a világ legnagyobb kihívásainak megoldásában: a krónikus betegségek kezelésében, a halálos kimenetelű közlekedési balesetek csökkentésében, az éghajlatváltozás leküzdésében és a kiberbiztonsági fenyegetések előrejelzésében” (ettől eltér a ChatGPT véleménye, lásd D.2 függelék). Ráadásul a magyar stratégia szerint “*fizikai eszköz igény nélkül*”. A 2018-as álláspont

szellemében a 2019-es magyar stratégia is a MI egy területének tekinti az automatizált rendszereket: a vezetői összefoglalóban “*egyreszakértői becslések szerint a munkakörök több mint 40%-a már jelenleg is automatizálható*”, bár utána már (következetlenül) külön-külön is említi az automatizált és a MI rendszereket. Sőt, mellettük helyenként önállóan is megjelenik a gépi tanulás. Jó lenne ha pontosan lehetne tudni, mikor miről beszélünk. Az EU újabb álláspontjának megfelelően felülvizsgálandó a stratégia (szövege és célkitűzései is).

A fogalmat 2022-ben az EU, az eredeti jelentéstől teljesen elrugaszkodva, gyakorlatilag a ChatGPT testére szabta: (tanításhoz) *bemenő adatokat használó és generált tartalom kimenetet előállító, egy bizonyos listában szereplő technológiákkal előállított szoftverként* határozta meg. Ennek alapján az említett [133] jelentés szerint a “B” kategóriába besorolható rendszerekre vonatkozik. Amivel az az apró baj van, hogy a legtöbb szoftver bemenő adatokból valamilyen kimenő adatot generál, és a lehetséges technológiák fénysebességgel változnak, továbbá nem terjed ki az adat-vezérelt módon megvalósított (ezen belül: hardver vagy organoid) rendszerekre sem. Kizárja továbbá a valódi neurális hálózatokat és a robotikát is.

Ez az értelmezés [23] a MI technológiákat ‘szűkített MI, artificial narrow intelligence (ANI)’ és ‘általános MI, artificial general intelligence (AGI)’ összemósódó és eléggé tág kategóriákra osztja (a kategóriákat a 2.5 szakasz mutatja be). A dokumentum megemlíti, hogy az ANI lényegében a gyenge MI, az AGI az erős MI kategóriának felel meg. Megközelítése azért furcsa, mert történetileg sem így jöttek létre, továbbá az AGI-t úgy definiálja, hogy az “*intelligens feladatok széles körét végzi, absztrakt módon gondolkodik és alkalmazkodik az új helyzetekhez*”. Amiben a “széles kör” csupán a rendszer méretét jellemzi nem pedig a működés módját; nem a feladat intelligens hanem a megoldási módnak kell olyannak lenni; az absztrakt gondolkodásról a szakértő véleménye “*Hat évtizeddel később a kitűzött kognitív képességek még mindig elérhetetlennek tűnnek*” [82], lásd még a D.1 függelékét; a rendszer új helyzetekhez nem tud alkalmazkodni, csupán hangolja magát a helyzet csekély mértékű megváltozásához. A definíció többi része rendben van. Ennek alapján az ANI ugyanaz, csak kevésbé széles területen: a definíció saját magával sem konzisztens. Annyiban igaz a definíció, hogy nincs minőségi különbség az ANI és az AGI között; de valójában ezt úgy kell érteni, hogy a később létrejött AGI nem hozott létre minőségi változást, csak a több rendelkezésre álló hardver és adat alapján ambiciózusabb.

Az EU által elfogadott jelenlegi (2023) MI meghatározás [23] elismeri, hogy *nincs általánosan elfogadott MI meghatározás*. Ismét csak hivatkozás nélkül, de egyetért [133] fél évszázados megállapításával, hogy *fogalmilag*

sem lehet egységes MI-ról beszélni és hogy matematikailag is legalább három típusra kellene felosztani. Nem csak működési módját és tulajdonságait nem ismerjük, de még a fogalmát sem tudjuk pontosan meghatározni. Nagy bátorságra (vagy botorságra) vall egész iparágakat (lényegében jövőnket) MI-alapúvá alakítani, mégpedig kötelező jelleggel, politikai akaratként megfogalmazva. A MI részleges tiltásának dokumentumában [28] is csak az szerepel, hogy a (szakmai) “előadó a mesterségesintelligencia-rendszer nemzetközileg elismert fogalommeghatározásának használatát szorgalmazza” (figyelmen kívül hagyva, hogy pár sorral odébb állította, hogy nincs olyan). A betiltás és a hiányzó fogalmi meghatározás mellett legalábbis furcsán hat, hogy “A mesterséges intelligencia egy kiforrott és használatra kész technológia, amely felhasználható az ipari folyamatok során keletkező, egyre növekvő adatmennyiség feldolgozására.” (Ami a vezetői összefoglalóban található marketing szöveg; magának a MI-nak, pontosabban az azt belülről ismerő fejlesztőknek, más a véleménye, lásd D.2 függelék.) Ami azzal együtt, hogy egyre csökkenő számítási hatásfokkal és egyre növekvő energia pazarlással alkalmazzuk, elgondolkodtató; különösen, hogy az adatmennyiség racionalizálása fel sem merül. Erre vonatkozóan a magyar MI stratégia helyesen állapítja meg, hogy az “általános MI” kutatása fejletlen és bizonytalan” (de az akkor is igaz marad, ha használatát leszűkítjük egy-egy területre). A szerző hozzátenné: *két különböző intelligenciáról van szó: a gépi intelligenciát kutatjuk és az emberi intelligenciát próbáljuk elérni*, lásd még a ChatGPT véleményét is a D.1 függelékben.

A Turing-teszt teljesítése sikerességének demonstrálására irányuló lobbizás igen erőteljes. Az érdeklődést felkeltő publikus “A ChatGPT teljesítette a Turing tesztet” [19] (ezzel szemben áll a ChatGPT önértékelése, lásd D.1 függelék) szalagcím alapján a már csak korlátozottan elérhető közlemény címe “Könnyű intelligencia tesztek amelyeket a ChatGPT nem tud teljesíteni”, a közleményben pedig az áll: “a GPT-4 és más LLM-ek valószínűleg átmennének a népszerű Turing teszten, olyan értelemben, hogy az emberek nagy részét bolonddá teszik, legalábbis egy rövid csevegés során”. Az értékelés szerint nem felülmúlják az embereket, hanem a használt benchmark képességei végesek (és a rendszereket “felkészítik” a tesztek teljesítésére). A “contamination” jelensége révén a felhasznált nagy mennyiségű adat között megtalálja a helyes választ. *Feltéve, hogy az adatkészletben megtalálható és ott csak helyes válaszok szerepelnek.* Ha nem nyert, “újra húzhat”. Ha nem sikerült, az nem újsághír. Aztán a fejlesztők tökéletesítik a rendszert, hogy azt az egy benchmarkot teljesítse. Azaz, a Turing által eléggé nagyvonalúan megfogalmazott emberi szereplő intelligenciájától, a feltett kérdések körétől, az elérhető adatkészletek körétől és időszerűségétől, a csevegés időtartamától, a

megengedett “gondolkodási” időtől, stb. is függ a teljesítés. (A szerző – Turing iránti nagy tisztelete ellenére – megjegyzi, hogy a kérdés feltevés lényegében értelmetlen, lévén hogy két teljesen eltérő, csupán nevében egyező, intelligenciát hasonlítunk össze; lásd D.1 függelék; továbbá az emberek többsége sem menne át a Turing teszten.)

Kérdés, hogy megvalósításának anyaga/technológiája, működési elve, a működés matematikai leírása, felhasznált adatkészlete, végzett műveletei, célja, jel ábrázolási/átviteli módja, vagy valamely más jellemzője alapján hogyan tudjuk a Természetes Intelligencia (TI) ember által készített hasonmását egyéb rendszerektől elválasztani, továbbá a hasonlóság mértékét jellemezni. Sajnálatos, hogy a magyar stratégia (lásd annak 4.1.2 pontját) kizárólag a MI matematikai alapjainak kutatását tartja szükségesnek, arról egységesen tárgyalható fogalomként beszél. Technikai megvalósításának vizsgálatát meg sem említi, pedig (mint azt a B.2 szakaszban tárgyaljuk) egy javasolt és matematikailag látszólag vonzó absztrakt módszer technikai megvalósításának vizsgálata megmutatja, hogy *nem maga a MI alkalmazás válik fenntarthatatlanná, hanem annak a számítástudományi megalapozás nélküli technikai megvalósítása*. Hasonlóan hiányzik annak észrevétele, hogy a folyamatosan növekvő információ tömeg fizikai kezelése észszerű idő alatt mindinkább lehetetlenné válik, és a MI által mind nagyobb mennyiségben előállított megbízhatatlan adatok használata az egyébként jól működő rendszerek által szolgáltatott eredményeket is meghamisítja. *Ideje lenne bevezetni az adatmennyiség csökkentése érdekében az adatok hitelességének és elévülésének fogalmát.*

2.4 A TI meghatározása

Az emberi elmét régóta izgatja saját működése. Az éppen “húzóágazat” tudománytól függően volt pl. mechanikus vagy pneumatikus modellje is. A digitális számítógép (és egyáltalán a digitális elektronika) megjelenése nagy hatással volt az agy működésének tanulmányozására és modelljeire, ma is szoros kölcsönhatásban fejlődnek [95]. Manapság általánosan elfogadott, hogy az agy számol [86], bár mint [48] magyarázza, “szigorúan véve, ma sem értjük, milyen elv alapján számol az élő anyag”. A természetes intelligenciát általában az emberi létezéshez (vagy legalább fejlett biológiai rendszerhez [95]) kötött, biológiai szövetek bonyolult együttműködésén alapuló magasabb rendű viselkedési módnak tekintik, amelynek viselkedését sem a legmagasabb agyi funkciók, sem az elemi neurális műveletek területén nem ismerjük teljes mértékben [44]. A tanulás általános biológiai megfogalmazása [136]: “használjuk úgy a jelenleg rendelkezésre álló információt,

hogy azzal javítsuk a rendszer jövőbeli működését”. Azaz, a biológiai tanulás (az intelligens működés) lényegi része, hogy a rendszer módosítani tudja saját jövőbeli működését, amihez “jövőbe látás” azaz *intuáció is szükséges, lásd B.8 szakasz, amire a számítási működés idő-ablakának használata ad (vagy nem ad) lehetőséget*, lásd A.6 szakasz. Lényeges momentum, hogy nem csak egyszerű mennyiségi tanulásként, inkrementálisan, a korábban már látott viselkedés minták valamelyikét (esetleg azok valamilyen kombinációját) használjuk, hanem egy *új és váratlan helyzetre is jól reagáljunk* (nem-inkrementális, ugrás-szerű tanulás). A [37] meghatározás szerint *a kizárólag a korábban látott viselkedés mintákon alapuló rendszer nem intelligens: nem szerez új lépességet, hanem változatos módokon próbálkozik*.

A MI közvetlenül megalkotása után sikerrel bizonyított matematikai tételeket, ezzel érdemelve ki a “gondolkozó gép” fogalmát és további, nem kellően alátámasztott reményekre jogosított. Ha a MI atyjai (és mai művelői) a szerző néhai matematika tanáránál tanulták volna a szakmát, tudták volna, hogy “differenciálni a lovat is meg lehet tanítani, de az integrálás művészet”. Azaz *szabályok következetes betartása (ami az emberi elme számára fáradtságos és hosszadalmas tevékenység, nagyon sok hiba lehetőséggel) algoritmizálható, de vannak ugyanabba a kategóriába eső tevékenységek, amelyekhez intuáció is kell*. Ezzel teljes összhangban van az a meghatározás, amivel a ChatGPT megkülönbözteti a gépi és az emberi intelligenciát, lásd D.1 függelék.

A “megsejtés” az a nehezen megfogalmazható *minőségi többlet* tevékenység, amelyet aztán algoritmizálható módon be kell bizonyítani. Lényegében ez a finom megkülönböztetés mondja meg, hogy a számítógépes tevékenységgel előállítható “intelligenciának” hol húzódnak a határai; lásd még a B.8 szakaszt. A tanulmány céljaira bevezethetünk egy *inkrementális intelligencia* (ami egzakt matematikai módszerekkel leírható) és egy *intuációs intelligencia* (nem egzakt matematika, különösképpen nem probablisztikus működés, amit a biológiai működés nem-determinisztikus volta indokol, lásd A.7 és B.8 szakaszok) fogalmakat. Létezése okán bevezetjük a *probabilisztikus intelligencia* fogalmát is, amit jelenleg pl. a ChatGPT is képvisel, továbbá a *kvázi-probabilisztikus intelligencia járulék* fogalmát, amit a technikai megvalósítás ad a különböző elvű megvalósításokhoz. Hipotézisünk, hogy *a tanulás inkrementális módját a gépi tanulás matematikai módszerekkel többé-kevésbé imitálni tudja* (pl. döntési fák, gradiens vagy előfordulási valószínűség alapján). *A nem-inkrementális (ugrás-szerű, minőségi változást hozó) módra azonban nincs matematikai módszer, és az egyáltalán nem azonosítható a probabilisztikus intelligenciával*.

2.5 A MI fő típusai

[60] szerint a MI megvalósítások képességeik szerint alapvetően három kategóriába sorolhatók:

- Gyenge MI (weak AI) (“játék-szintű”, toy-level) (ANI
”a MI azon fajtája, amelyik egy vagy több *speciális* területre fókuszál”
Az egyik leggyakrabban idézett sikertörténet, az AlphaGo is ide tartozik; továbbá ide sorolandó az ál-MI megvalósítások legtöbbször
- Erős MI (strong AI) (“általános MI, artificial general intelligence (AGI)”
”olyan MI, amelyik a *teljes* emberi intelligenciát modellezi”
Nincs olyan meggyőző bizonyíték, hogy ilyen megoldás teljesítette volna a Turing tesztet (azaz megkülönböztethetetlen lenne az emberi intelligenciától) Lásd még 2.3 szakasz és D.1 függelék
- szuper MI (agy szimulálás)
”olyan MI amelyik *az emberéhez hasonló tudati képességekkel rendelkezik* és messze meghaladja a világ legokosabb emberének képességeit is”
Csupán hipotézis, a sci-fi világában létezik; a MI alapítóinak elképzelésén is túllép

Az EU besorolás [23] (csupán az ANI és AGI kategorizálást említi) esetén legalább következtetni lehet az adat-vezérelt működés imitálásának szándékára abból hogy “nagy mennyiségű adaton tanítják be és minimális finomhangolást igényel”, de szoftver imitálással. Az viszont egyértelműen látható, hogy *utasítás-vezérelt elven működő szoftver általi adatfeldolgozásról, nem pedig adat-vezérelt működésről beszél*. Ennek alapján a Feynman-féle “haladó algoritmus” meghatározás tökéletesen elegendő számára. *Csupán közelítőleg lehet a MI fogalmait a TI hasonló fogalmaira leképezni* [84]. A [82] által javasolt sorrend fordítás (AI helyett IA, mint “Intelligence Augmentation”) egyáltalán nem csupán szójáték: *a számítógépes rendszer óriási ismeret anyaghoz fér hozzá, több adatot tud tárolni, gyorsabban működik, mint az ember; jól kiegészíti és segíti az emberi intelligenciát (természetesen feltéve, hogy helyesen valósították meg), de nem pótolja azt*.

Mindkét besorolásnál nyitott kérdés marad, hogy pontosan mennyi az “egy vagy több”, illetve mi számít “szűkített terület”nek. Az említett osztályba sorolási meghatározások megkerülik az intelligencia fogalmának megfogalmazását, egy más tudományterületre megfogalmazott és el nem ért

vágyálom *elnevezése* alapján, de teljesen eltérő tartalommal megtöltve, bizonytalanul osztják fel a MI megvalósításokat, továbbá teljesen figyelmen kívül hagyják hogy a megvalósítás módja hogyan befolyásolja a tulajdonságokat és a működést. *A technikát és a működés elméleti alapját egyik felosztás sem határozza meg, mindkettő kizárólag matematikai szemléletű. Nem akarják tudomásul venni, hogy az implementált technikai rendszerek működése az implementáló technika szabályainak van alávetve.*

Érdekes, hogy egyik meghatározás sem említi meg, hogy évtizedek óta létezik [126] a gépi tanulás modelljeinek matematikai alapú osztályozása is, ami egyúttal matematikai viselkedésükre is utal és használata sok bajt megelőzhetett volna. A modell lehet determinisztikus vagy probablisztikus. Formálisan jelölje \mathcal{X} és \mathcal{Y} a bemeneti és a kimeneti tartományokat. Az $x \in \mathcal{X}$, $y \in \mathcal{Y}$, $\Theta = \{\Theta_1, \dots, \Theta_D\}$ jelöléseket használva (ahol Θ paraméterek halmaza), egy determinisztikus függvényt a

$$y = f(x; \Theta) \quad (2.1)$$

függvényt ír le. A probablisztikus modell bemenő és kimenő változói egy $p(x, y)$ valószínűségi eloszlásból származnak. A probablisztikus modell tovább bontható diszkriminatív

$$p(y | x, \Theta) \quad (2.2)$$

(Az osztály valamely tagjának valószínűségét a $p(y | x)$ adja meg.) és generatív (adatokból felépített, tapasztalati eloszlást használó)

$$p(x, y; \Theta) \quad (2.3)$$

eloszlásokra. A jelenleg problémákat okozó ChatGPT egyértelműen generatív modellt használ.

magyareloszlásokra. A jelenleg problémákat okozó ChatGPT egyértelműen generatív modellt használ. A matematikával kevésbé ismerősek számára, némiképp egyszerűsített magyarázatként, a három függvénnyel három különböző módon határozzuk meg a $2 + 2 = ?$ feladat eredményét.

A feladat megoldását a *determinisztikus* esetben a szokásos módon, egzakt matematika használatával modellezzük. Az eredmény többszöri próbálkozás esetén is ugyanaz, jól tesztelhető.

A *probablisztikus* esetben gondjunk pl. a liszt csomagolására. A 2 kg-os csomagolás esetében szinte biztosan lehetünk abban, hogy a mérlegre téve nem 2000 gr lesz a mutatott érték a mérlegen, de biztosak lehetünk abban

hogy sok mérés esetén átlagosan megkapjuk a 2000 értéket, valamilyen kis szórással. Meg tudjuk mondani, hogy mekkora a várható érték. Ha a töltő automatát ellenőrzendő súlymérés során olyan értéket kapunk, amelyik a 2000 gr-tól csak annyira tér el, amennyit a szórás megenged, akkor nem kell kételkednünk abban hogy az automata jól működik. *Ismerjük az eloszlást* és a tulajdonságait és elfogadjuk, hogy az esetek 1-2%-ában mondjuk csak 1999 gr-os zacskót kaptunk. Mind az 1999 gr-os, mind a 2001 gr-os zacskó esetén az mondjuk, hogy vettünk egy 2 kg-os csomagolású lisztet. Az eredmény minden egyes próbálkozás esetében más; csak a próbálkozások eredményének sokaságáról tudunk statisztikus megállapítást tenni; csak ilyen értelemben tesztelhető.

2.6 A generatív eloszlás Turing-tesztje

Már a ChatGPT-re gondolva, részletezzük a generatív eloszlás néhány tulajdonságát. A fogalom megértéséhez gondoljunk a $2 + 2 = ?$ feladat kérdőíves közvélemény kutatással történő megoldására. Első lépésként megkérdezzük az embereket, hogy véleményük szerint mennyi a helyes eredmény. Mondjuk az emberek 30%-ának válaszában a helyes érték 4, 35% szerint 3, 15% szerint 5 és 10% szerint 6. Más közvélemény kutatás alapján, más megkérdezettek véleménye alapján más tapasztalati eloszlást kapunk. Akárhogyan is, kapunk egy generatív (tapasztalati) eloszlást, amelyik félkész állapotban tartalmazza a feladat megoldását; de nem ismerjük előre magát az eloszlást. Ha ennek alapján meg kell adnunk az eredményt, gondban vagyunk. Van ugyan átlagértékünk is meg szórás is, meg az első függvény módszere alapján ismerhetjük a helyes értéket is; de mi lesz a feladat eredménye?

Második lépésként, az eloszlás átlagértékét nem adhatjuk meg eredményként, mert csak egész számot adhatunk meg. Úgyhogy adjuk meg a kérdőíven feltett kérdésre kapott válaszok valamelyikét (vegyük észre: feltételezzük, hogy a helyes válasz is közöttük van). Első próbálkozásként vegyük a leggyakoribb választ: $2 + 2 = 3$. Az így megadott válasz ugyanezzel a módszerrel egy másik felmérés szerint más érték is lehet, a mi MI programunk szerint viszont a mi kérdőíves betanításunkkal létrehozott adatkészlet szerint a helyes válasz mindörökké 3. Azaz, MI programunk minden esetben hibás választ ad. Az algoritmus ugyan jó, csak rosszul választottuk meg a betanításhoz használt adatkészletet.

A *probabilisztikus* modellhez való hasonlóság alapján mondhatjuk azt is, hogy válasszuk az adatkészlet valamelyik elemét válasznak. Sem a kérdezőbiztosnak, sem a generatív modellnek, sem a válaszadónak nem kell tudnia, hogy erre a kérdésre csak egy helyes válasz van. Végül is, egy “in-

telligens” társadalomban többféle helyes válasz is létezhet, sőt a társadalom valamelyik intelligens tagja is adhat a körülményektől vagy az időtől függően más válaszokat ugyanarra a kérdésre. Azaz, valamilyen módszerrel, alkalomszerűen, kijelöljük a kérdőívben megadott valamelyik választ, és azt adjuk meg válaszul a feltett kérdésre. A válasz a TI egy képviselőjétől érkezett, tehát a kérdező a TI által adott választ kap; azaz olyat, amelyet egy találmányra kiválasztott embertől is kaphatna. A MI program tehát pontosan ugyanúgy viselkedik, mint ahogyan találmányra kiválasztott emberek viselkednének. Teljesíti a Turing tesztet: nem tudjuk megmondani, hogy a hibás válasz közvetlenül egy embertől vagy a válaszokat összesítő programtól érkezett.

Nem melleleg: ha a helyes érték is a lehetséges válaszok között van, a mi MI programunk már nem nulla valószínűséggel ad helyes választ. Sőt, matematikailag korrekt értelemben bejelenthetjük, hogy a mi MI programunk 30% valószínűséggel helyes eredményt ad, pontosan ugyanolyan arányban, mint a TI. Azaz, hurrá, sikerült gépi módon előállítanunk az intelligenciát, ami *egyetlen* feladatban teljes körűen, Turing értelmében, imitálni tudja az emberi intelligenciát. Igaz, hogy csak egyetlen feladat megoldására, tehát az csak “szűk MI”, de ha hozzáadjuk a $3 + 2 = ?$ feladat megoldásának kérdőívét is, akkor már két feladatra is kiterjed. Előbb-utóbb eljutunk oda, hogy a “feladatok széles körében” tudja imitálni az emberi intelligenciát.

Ha elég buzgók vagyunk, akkor a generatív függvényt akár “kreatívvá” is tehetjük. Az akár véletlen is lehet, hogy a helyes válaszok a 3, 4, 5, 6 számok közül kerülhetnek ki (igaza van, ez az adott kérdőív kitöltőitől függ: nem felelet választós). Miért ne lehetne akár 2 vagy 7 a helyes válasz? (néhányek ezt az “intuíció” fogalommal tévesztik össze, mások egyszerűen kitalálásnak, hallucinációnak tekintik.) Azt ugyan először szomorúan vesszük tudomásul, hogy a $3 + 3 = ?$ feladat kérdőíve még továbbra sem szerepel MI programunk tudásában, de ha nem csak kreatív programot készítünk, hanem általánosításra is vállalkozunk, akkor mindjárt van megoldás a problémára. Végülis, látjuk, hogy két számról, azok összeadásáról van szó, meg két kérdőíves adatkészletről, továbbá kreativitás is megengedett. Így az utóbbi kérdésre is rávágthatjuk azonnal a választ. Még akár az is előfordulhat, hogy helyes választ ad a programunk, legfeljebb a találati pontosság kisebb lesz. Viszont nem kell beismernünk, hogy csak bizonyos számok összeadásáról van ismeretünk (legfeljebb azt, hogy csak összeadásról). Baj csak akkor van, ha a kérdező tudja a helyes választ. Különben abban a meggyőződésben kel fel a képernyő elé, hogy MI programunk mindenre tudja a választ (legalábbis úgy csinál).

Az adatkészlettel kapcsolatos probléma egyúttal rámutat a tanulás egyik

nehézségére is. Ha folyamatosan tanuló rendszert használunk, amelyik a kapott válaszokat beépíti adatkészletébe, akkor a tudatlan vagy maliciózus felhasználók válaszai (vagy akár a gépi tanulás később említendő idő- és sorrend viszonyai) alapján az adatkészleten alapuló tapasztalati (generatív) eloszlás függvény folyamatosan változik és egy idő után még jó eloszlásfüggvénnyel indítva is elkezdhet rossz válaszokat adni.

Vegyük észre azonban, szomorú tanulságként: generatív eloszláson alapuló programunk

- csak olyan eredményt tud adni, amelyet adatként már látott betanítása során
- ha újat alkot, abban nincs sok köszönet
- még ha ismeri is a helyes választ, nem biztosan azt adja meg
- még a betanítása során már látott feladatot sem biztosan tudja megoldani
- fogalma sincs arról, mit is csinál: számokkal és valószínűségekkel dolgozik

english

2.7 Javasolt osztályozás

A tanulmány szerzőjének véleménye szerint használhatóbb lenne a következő osztályozás (mindegyik esetben feltételezve a korrekt megvalósítást, főként a működési idő megfelelő kijelölését; lásd A.6 függelék)

- *determinisztikus egyszerű vagy akár “haladó” algoritmusok* (2.1)
(a szó klasszikus értelmében, pl. szabályozó és vezérlő körök használata, egyszerű osztályozás és minta felismerés, döntési fák, sakk, Go és hasonló játék, a gépi tanulás nem-probabilisztikus módszerei “labelled” adatokkal, időtől nem függő “betanított” rendszerek; stb.)
- *generatív+adat-vezérelt+pszeudo-random algoritmusok és rendszerek*
 - generatív és transzformatív probabilisztikus (2.3)
 - gépi tanulás unlabelled adatokkal (pl. kép felismerés)
 - adatvezérelt (valódi MI vagy organoid alapú) számító rendszer
 - ad-hoc hardveres gyorsítás vagy modellezési elv

- pszeudo-probabilisztikus technikai (busz arbitráció vagy számítógép hálózat által előidézett látszólag véletlen-szerű működés, stb.; valós-idejű és idő-függő “betanított”) rendszerek
- *ugrás-szerű “haladó algoritmusok”* (csak TI)
 (“valódi avagy ugrás-szerű intelligencia”, azaz *nem csak* a korábban már látott minták (vagy azok valamiféle kombinációja) felismerése és annak alapján válaszként *nem csak* a betanított reakciók valamelyike (vagy azok valamiféle kombinációja) előállítása; erre jelenleg csak a TI képes) Ide tartozik az agy fiziológiai alapú vagy számítógépes kutatása is.
- *inkrementális* (probabilisztikus) *“haladó algoritmusok”* (2.2)
 (a gépi tanulás jól megalapozott viselkedésű matematikai függvényeket használó probabilisztikus módszerei; a valószínűségi eloszlások; idősorokból levont következtetéseken alapuló előrejelzések; statisztikai számítások és mintázatok keresése alapján hozott döntések; stb.)

Vegyük észre, hogy mindhárom osztályozásnak közös vonása, hogy tagadja az emberi intelligencia intuíció-szintű gépi előállításának lehetőségét, legalábbis a jelenleg ismert módszerekkel; a harmadik kategória. Az első kategóriába eső módszerek (működési területtől függetlenül) szabadon alkalmazhatók. A második (valódi vagy ál-probabilisztikus) kategória viszont kiemelt figyelmet érdemel. Csupán az első alterületnek van matematikai háttere, és az utolsó kettőnek csupán technikai kiviteli oka van. A működési terület kritikusságától, a technikai kivitteltől és a rendszer méretétől függően egyedileg kell vizsgálni a használhatóságát és az alkalmazandó szűkítéseket; komoly kockázatot jelenthet. A negyedik (megfelelő matematikai korlátozásokkal) a legtöbb területen jól használható. *Fatális hiba akár valószínűségi eloszlásokkal közelíteni a TI működését, akár (a gigantikus méretekhez már nem megfelelő technikai megoldásokkal) pszeudo-véletlen működési járulékokat bevezetni.*

Vegyük észre, hogy a *ChatGPT* valamennyi felsorolt kockázati tényezőről *tartalmaz egy keveset*; pedig azok egyenként is veszélyt jelenthetnek. A szerző által javasolt osztályozás világosan, a működés elve alapján, elválasztja a veszélytelenektől és a második kategóriába gyűjti a jelenleg kétségesnek vagy veszélyesnek tartott technológiákat. Ennek ellenére így is akadnak “szürke területek” (pl. a valós idejű működés kényszere miatt egy önvezető járműben a haladási irány megválasztásakor vagy egy időre játszott bonyolult táblajátékban nem engedhetjük meg az állapottér teljes felmérését, de kifejleszthetünk észszerű állapottér korlátozási algoritmusokat).

Az EU új, kockázat-alapú felosztásának négy kategóriája [17] lényegében csak a tevékenységi terület szerint és igen homályosan differenciál, de inkább csak példákat hoz. Csupán véletlen, hogy a működési elv matematikája és technikája alapján a szerző is és a működési terület alapján az EU is négy veszélyességi kategóriát nevezett meg. A szerző véleménye szerint csupán a működési mód szerinti második kategóriában van értelme a működési terület szerint osztályozni a rendszereket.

A szerző által javasolt meghatározás bizonyos értelemben *kibontja és pontossá teszi az EU jelenlegi (2023.05.22) [28] meghatározását (elvi alapokra helyezi a modellezési módszer és a technikai kivitel alapján)*, hogy az osztályozásnak “a mesterséges intelligencia kulcsfontosságú jellemzőin, például annak tanulási, érvelési vagy modellezési képességén kell alapulnia”. Kérdéses, hogy az előbbi mondatban szereplő képességek valóban léteznek-e: a *tanulás* valójában *betanulás* és nem *megtanulás*, az *érvelés* csak a TI bevonásával, korlátozottan valósítható meg, a *modellezés* pedig (legalábbis az EU értelmezésében) nem absztrakciós szintű, hanem egy nagyon sok paraméterrel megvalósított ad-hoc megoldás, lényegében a (2.3) formula alapján.

Technikai kivitelüket illetően a MI rendszerek jellemzően elektronikus eszközök (változó mennyiségű és nem mindig helyesen értelmezett biológiai imitációval), de újabban megjelentek biológiai alapú MI eszközök is, lásd A.5 szakasz. Az elektronikus eszközökön belül a klasszikus fizika azonnali kölcsönhatás fogalma alapján működők dominálnak, de vannak kísérletek [98, 43, 94] a biológiai működés inkább kevésbé mint többé helyes működésének megvalósítására is. Az általános célú digitális számítógépek elterjedtségének és relatív olcsóságának következtében a legtöbb megvalósítás szoftver alapú, ami legtöbbször nélkülözi (vagy helytelen elvek alapján próbálja imitálni) a biológiai hasonlóságot. Ráadásul a hasonlóság mértéke a körülményektől is függ. A McCulloch-Pitt modell [99, 120] előre lépést jelentett az egyes neuronok működésének megértésében, de hátrányt jelent alkalmazása a neuron hálózatok megértésekor. Nagyon kis (néhány “neuront” tartalmazó) neurális hálók modellezésére nagyon jól használhatók a matematikai logikai függvényeket megvalósító logikai kapuk, de nem érthető meg, miért korlátozott a neurális hálózatok számító képessége és miért vezet hibás működéshez visszacsatolások, konvolúciók használata.

Az első hullám idején, amikor a célul kitűzött “általános MI” megvalósíthatósága kérdésessé vált, a nagyon alapos [133] kutatási áttekintés világosan kifejtette, hogy a vágyakból mennyi reális; annak alapján a felelős testületek beszüntették a kétes célú MI kutatás finanszírozását, az első MIT-élhez vezetve. A D.2 függelék alapján a tudás bázis készítői valahogyan összekeverik az okot és annak következményét. Állításuk szerint a tél azért

következett be, mert elapadtak a támogatások. Valójában a tanulmány a kutatásokat “A” (Advanced Automation) és “C” (Computer-based CNS (Central Nervous System)), valamint a közöttük kapcsolatot teremtő “B” (Bridge) kategóriákra osztotta és összefoglaló véleménye szerint “definite (though not overwhelming) promise of work in categories A and C, and of the general failure of work in category B” (azaz határozottan ígéretesek lehetnek az “A” és “C”, de általában kudarcra vannak ítélve a “B” kategóriába (főként robotika) tartozó projektek). A finanszírozás a jelentésben összefoglalt okokból következő eredménytelenség miatt apadt el és nem fordítva.

A tanulmány előre látta, hogy a “B” kategóriából néhány terület átkerülhet az “A” kategóriába (pl. a karakter vagy egyéb mintázat-felismerés, a robotika egy része, vagy pl. sakk játék; bekövetkezett), de *kérdéssé teszi, hogy ezt a három kategóriát lehet-e egyáltalán egy “egységes MI” kategóriába sorolni*. Valójában az “A” és “C” eredményessége és sikerei miatt akarták magukat a “B” kategóriába tartozó kutatások MI-hez tartozónak feltüntetni; a jelentés ennek a törekvésnek mondott ellent. Azonban *a mai fejlesztők és támogatóik* ezt nem értették meg vagy nem akarják megérteni és *a másik két terület eredményei mögé bújva akarnak sikereket (továbbá, természetesen, támogatást és profitot) elérni*. Sajátos módon, a jelentésre (mint idejét múlta) vagy nem hivatkoznak, vagy lényegi megállapítását elhanyagolják, vagy felcserélik az ok-okozat viszonyt.

Vegyük észre, hogy az “A” kategória lényegében megegyezik a szerző által javasolt első kategóriával, a második a “B” kategóriának felel meg. A negyedik (megfelelő körülményekkel) bekerülhet az “A” kategóriába. A harmadik kategória lényegében a TI fogalmát (az eredeti szándéknak megfelelően) kapcsolatba hozza a MI “C” alterületével, de nem teszi azzal egyenértékűvé. A “B” erőteljesen különböző kategóriákat próbál összekapcsolni, ezért kritikus analízist igényel. Mindegyik említett kategorizálási módszerből világos, hogy *MI módszerről vagy technológiáról nem beszélhetünk egységes egészként, és csupán annak egyes kategóriáira vonatkozóan lehet és kell kijelentéseket tenni*.

2.8 A puding próbája

Úgy tűnik, a skálázhatóság megértése sem erőssége a MI/LMM hívőknek. Természetesen ma mindenki mindenre az LLM-ben találja meg a csodaszert és “minél nagyobb annál jobb” elven épült rendszereket akar használni. A MI fejlesztése (valószínűleg nem véletlenül) a szoftver fejlesztésnek ahhoz a módjához hasonlít, amikor először kitűzzük a teljesítendő feladatot (milyen benchmark teljesítése), aztán addig gyötörjük az algoritmust és az adatokat, amíg a tesztprogram sikeresen le nem fut. Ami természetesen csak annyit

bizonyít, hogy az illető tesztet sikerrel vette a program. Egy másik teszt természetesen másik hibák kijavítását teheti szükségessé. Ha az nem sikerül, csak az első benchmark sikeres teljesítését tesszük közzé. Természetesen folyik a rakéta-ellenrakéta fejlesztés: a ChatGPT3.5 által még nem teljesíthető teszteket a ChatGPT4 már teljesíti, hiszen ismert a hibák oka. Közben azonban megszületik a másként tesztelő benchmark, amely már a ChatGPT4 számára is nehéznek bizonyul [125].

Bemutatunk két olyan esetet, ahol a követelményszint magasabb. Az elsőben a természet fejleszti a benchmarkot, a másodikban a technológia.

2.8.1 A COVID kezelése

Sajnos, a közelmúltban előfordult az az esemény, ami “élesben” tesztelte a MI képességeit. Ilyen szempontból nagyon jó időzítéssel, 2019-ben, a MI hype erősen felfutó periódusában, a matematikai módszerek és az elektronikai eszközök teljességének birtokában, a módszerek és eszközök használatában jártas nagy számú kutató jelenlétében, a COVID előidézte azt a társadalmi szükségletet, hogy a MI is harcba szálljon a járvány megfelelő kezelésének érdekében. Ehhez a kormányoktól megkapták a szükséges anyagi és emberi erőforrásokat is. Az eredményt [46] a következő módon foglalta össze.

“Illusztratív példaként szolgál ama számos eszköz esete amelyeket a COVID-19 diagnosztizálására és előrejelzésére fejlesztettek ki, amelyek megvalósítására 2019 és 2020 során jelentős erőforrásokat biztosítottak. Ezek analízise világos következtetéseket hozott. A nagy tekintélyű “The Alan Turing Institute” jelentése [139] az “UK Data Science and Artificial Intelligence” közösségnek a COVID-19 elleni küzdelemre adott válaszát ismerteti. Következtetése, hogy bár számos projektben széles skálán történtek fejlesztések, az eredmények nem igazán voltak használhatók, jócskán teret hagyva a jövőbeli járványok sokkal jobb előrejelzésének. Ugyanerre jutott a British Medical Journal-ban megjelent tanulmány, amely a COVID-2019 diagnosztizálására és prognózis céljára kifejlesztett előrejelző eszközök mindegyikét alkalmatlannak találta klinikai használatra. Mindössze két olyat talált, amelyik további vizsgálatra érdemes. Egy, a Nature Machine Intelligence folyóiratban megjelent tanulmány 415 olyan gépi tanulás alapú eszközt vizsgált, amelyek mellkas röntgen és CT felvételek alapján próbáltak a páciensek esélyeire következtetni. A következtetés megint csak nem biztató: egyik sem találtatott alkalmasnak klinikai használatra.”

Úgy tűnik, ez az éles teszt, amiben a résztvevők minden lehetőséget megkaptak (a nagy figyelmet és a kiugrás/bizonyítás lehetőségét is), továbbá gyakorlatilag szó szerint az életük volt a tét, egyáltalán nem igazolta az EU

2018-as reményeit, hogy “a mesterséges intelligencia segít a világ legnagyobb kihívásainak megoldásában: a krónikus betegségek kezelésében”. Hasonló eredményre jutott a National Institute of Health is: “Valójában, még ha számos alkalmazás célozta is meg a COVID-19 előrejelzését és diagnosztizálását, közülük csak néhány elég érett arra hogy klinikai körülmények között alkalmazni lehessen.”[50]

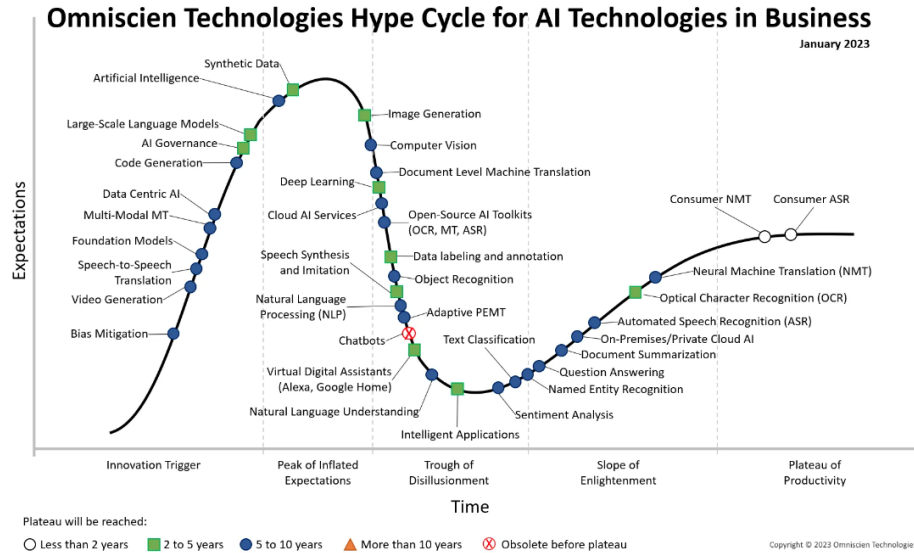
Kérdés, hogy más területeken, ahol szintén a természet minősít a benchmarkok vagy akár a Turing teszt helyett, és ahol hiányzik ez az összetársadalmi összefogás, nagyobb lehet-e a sikere. A tanulmány szerzője szerint ugyanez lesz az eredmény. A természettel nem lehet jópofáskodni és a benchmarkok teljesítését manipulálni.

2.8.2 Elektronikus tervezés

Érdekes és stílusos ötlet a hibás elvű LLM-et használni a hibás elvű hardver gyorsítóegységek tervezésére [112]. A MI/LMM fejlesztésében érdekelt cégek által támogatott technikai újságírók véleménye szerint még nem sokat látni abból, amit az ilyen elven működő elektronikus fejlesztő rendszerektől elvárnak. Az eddigi tapasztalatok szerint nem boldogulnak az összetett feladatokkal, nem tudják elérni a zárt forrású (TI által készített) rendszerek által elért teljesítőképességet. A gyorsító rendszerekben használt megoldások ugyanis eltávolítják a szükséges információ egy részét, a látszólagos gyorsabb működés érdekében félbevágják a számítást. Ennek következtében a tervezőrendszer éppúgy bugyután viselkedik, mint a csevegő robotok: az általuk tervezett rendszerek is bugyután viselkednek vagy egyáltalán nem működnek. Továbbá ezek a problémák csak a TI által készített “demonstrációs célú” darabok elkészítésével csökkenthetők, ami másképpen fogalmazva, azt jelenti, hogy ha betanításhoz használt adatkészletben nem szerepel a tervezendő “Intellectual Property” (IP), akkor a MI nem fogja azt megtervezni. Újat nem tud alkotni, csak egy (vagy több) régit megtalálni és abból kicsit átfogalmazva, egy “generált” IP-t készíteni. Hogy az működőképes-e, az attól függ. A MI nem tudja, mi a terv célja. Csupán készít valami “tartalmilag és formailag”, probabilisztikus értelemben, ahhoz hasonlót.

2.9 LLM/ChatGPT

A nagy nyelvi modellek, az LLM s az azon alapuló ChatGPT nyilván a MI felépítményei, de annak csupán egy szűk, de rendkívül népszerű ágát alkotják. Szokatlan népszerűségüket az okozza, hogy a be és kimenet egyszerű, az emberek által is olvasható természetes nyelvű szöveg. Játék és



Ábra 2.1: Különböző MI területek helyzete a technológiai várakozás görbéjén [25]

ingyen van. Az 2.1. ábra mutatja, hogy a MI (az ábrán AI) különböző célú felhasználásainak fejlettségi állapotai hogyan helyezkednek el a technológiai várakozás görbéjén. Amint látható, a gyenge MI (a szerző által, főként determinisztikus eredményük alapján, az első osztályba sorolt) bizonyos területei már megalapozott technológiaként megtalálták helyeiket, az erős MI (a szerző által, főként a használt probabilisztikus módszerek miatt, a második osztályba sorolt) területei pedig még a kiábrándulás időszakában vagy még inkább a felfokozott várakozás időszakában találhatók.

Vegyük észre az ábrán az egyetlen kivételként, hogy a ChatGPT-típusú csevegő robotok úgy kerültek a leszálló ágba, hogy előtte nem érték el a csúcst (érdekes lenne tudni, mitől lett a magyar stratégia kiemelt célja “olyan mesterséges intelligencia alapú vállalati tanácsadó szolgáltatások (chatbot) fejlesztése, amelyek tovább tudják növelni a digitálisan fejlett vagy fejlődni kívánó vállalkozások körét”). A közbeszédben (valamint az EU meghatározásában) jelenleg leginkább a szöveg feldolgozáson és szöveg generáláson alapuló megoldások jelentik “A” MI-t, pedig csak probabilisztikus szöveg elemző és generáló toldalékok járulnak a gépi tanulás alapsziszteméhez. Egy generatív eloszlással a szöveget (vagy képet, hangot, stb.) le kell képezni paraméterekre, majd az eredményből fordított leképezéssel szöveget, képet, hangot előállítani. A jelző nélkül használt MI jelentős mértékben hoz-

zárójárul a MI alkalmazásával szembeni ellenérzésekhez, bizonytalansághoz a MI képességek meghatározásában, vitákban “egymás melletti elbeszéléshez”.

Kérdés, hogy a ChatGPT egyáltalán besorolható-e a MI valamelyik típusába az első két említett osztályozás alapján. Az utolsó három betű jelentése alapján, igen; de kérdéses marad hogy a gyenge vagy az erős, illetve az ANI vagy az AGI osztályba. A szerző által felállított kategóriák szerint a második, [133] szerint a “B” kategóriába tartozik.

A *Chat* természetesen azt jelenti, hogy a rendszer bemenő és kimenő adatai is az emberi intelligencia számára is érthető szövegek. A *Pre-trained* jelentése, hogy nagy számú minta alapján előkészített működésű: a felismert “ehhez a szituációhoz hasonló” esetekben “azt szoktuk csinálni” működésre készül elő a (2.3) eloszlás előállításával. Lényeges, hogy *az előkészítés a betanításhoz és nem a megtanuláshoz hasonlít*. Ha jól felismerhető a szituáció (és elég nagy számban fordul elő a mintában), akkor a rendszer reakciója is eléggé illúziót keltő lehet. A *Generative* jelentése, hogy statisztikai valószínűségek alapján az általa ismert különböző szintű szöveg részekből újnak és logikusnak ható szöveget állít elő (generál); ha a bemutatott minták emberi szemmel nézve logikusak voltak, és statisztikus gyakoriságuk elegendő nagy volt. Hogy hogyan működik, azt a *Transformer* komponens szabja meg. Ez utóbbi egyfajta kirakós játék (“plasztikus puzzle”) elven működik, *a betanítás során megismert elemeket* megpróbálja csekély mértékben átalakítva olyan módon kombinálni, hogy azok nagy valószínűséggel illeszkedjenek egymáshoz, továbbá próbál *kreatív* lenni olyan módon, hogy az általa nem ismert tényeket kitalálja; bár működésének ezt a vonatkozását is tagadja, lásd D.2 függelék. *Kérdés, hogy az ismeretlen helyzetre való válaszadás ilyen módja mennyire felel meg az intelligencia már említett [37] meghatározásának*.

A *Transformer* komponens állítható paramétereket használ: a szöveg előállításakor például *be lehet állítani a tényyszerűség és a kreativitás arányát (“hőmérséklet paraméter”)*, már ha a légből kapott állításokat kreativitásként lehet értelmezni. *Emberi intelligencia esetén ez azzal analóg, hogy a beszélő az általa ismert tényekből (és a nem ismert tények helyett általa kitalált “valószínűsíthető tények”ből) olyan mondatokat rak össze, amit várakozása szerint a partnere hallani szeretne. Ráadásul a partner reakciójától függően mérsékli kreativitását: aki tájékozatlan és hagyja magát, azt nagyon átveri, aki meg nem, annak csak nagyon mérsékelt hasznos a működése. Még ha teljesen tényyszerű is a működése (azaz szigorúan csak a betanításkor használt adatkészletre támaszkodik), előfordulhat, hogy a jóváhagyott adatkészletben már van MI által készített (és ezért nem feltétlenül tényyszerű) szövegrész, lásd D.2 függelék.*

Van azonban egy további, rejtett, technikai komponens is. A nagy kereső

rendszerének használatakor természetesnek tekintjük, hogy a világ adatai között másodpercnyi idő alatt tudunk keresni, és csak kevéssé tudatosul bennünk, hogy előre összeállított és “a felhasználó közelében tárolt” eredményeket látunk. Ugyanezt a technológiát használja a ChatGPT is: az eredmény másodpercek alatt megjelenik, de leleplezhető, hogy ha rövid a “gondolkodási idő”, akkor *csak* előre gyártott panelekből épített (lásd D.2 függelék) és nem a kérdésünkre készített specifikus választ kapunk. Ennek számos közvetlen bizonyítéka van pl. bizonyos csevegő fórumok véleményének egyoldalúvá válása formájában, de ma már politikai választásokat is befolyásolhat [2]. Ez a viselkedés lehet mulatságos egy csevegő fórumon, de kifejezetten félrevezető tanuláskor vagy anyaggyűjtés alkalmával, megakadályozza felhasználását komoly célra (lásd 2.8.2 szakasz).

2.10 A MI használat szabályozása

A szándék [11] az volt, hogy “az EU-nak biztosítania kell, hogy a mesterséges intelligencia fejlesztésére és alkalmazására olyan megfelelő keretek között kerüljön sor, amelyek támogatják az innovációt és tiszteletben tartják az Unió értékeit és alapjait, valamint az olyan etikai alapelveket, mint például az elszámoltathatóság és átláthatóság.” [15] Amivel az a baj, hogy *egy olyan technológiától várjuk el az Unió értékeinek és jogainak betartását, amiről nem tudjuk, hogyan működik.* A ChatGPT velős véleménye a D.3 függelékben olvasható. Ezt a megállapítást a hatályos rendelet is tartalmazza, finom formában: “Ezek az összetett és átláthatatlan jellemzők (fekete doboz elem) hatással vannak az elszámoltathatóságra és a megmagyarázhatóságra.” [28] Pontosabban, az indokolhatóság hiánya (lásd 3.1.5 szakasz) már a szabályozás kibocsátása előtt is ismert volt, ami egyszerűen kizárja, hogy egyáltalán beszélhessünk az elszámoltathatóság és átláthatóság alapelveiről. Emiatt ilyen korlátozásokat beleépíteni sem tudunk és a ködös szabályozás betartását ellenőrizni sem lehet. Nem készült megvalósíthatósági tanulmány. *Az EU-ban senkinek nem jutott eszébe feltenni a kérdést: “Tudjuk-e, hogyan működik a MI, amit fejleszteni, kötelezően alkalmazni és szabályozni akarunk?”*

Ez a fogalmi tisztázatlanság megbosszulta magát. A kitűzött cél: “Az Európai Uniónak (EU) összehangolt megközelítéssel maximálisan ki kell aknáznia a mesterséges intelligencia által kínált lehetőségeket és kezelnie kell az azzal járó új kihívásokat” elérése olyannyira nem sikerült, hogy az EU-nak (természetesen ismét világelsőként, mert ez fontos) be kellett tiltania a MI alkalmazását számos területen. Érdekes módon, valamennyi az erős (a szerző által javasolt felosztás szerint: probabilisztikus) MI fogalomkörébe

esik. Nemcsak a technológia, hanem a szabályozás is digitális: vagy teljes gőzzel előre vagy vészfékezés. Közbülső állapot, gondolkodás és megértés nincs.

Az EU működési terület szerinti korlátozása [17] a MI technológiájának és működési elvének tökéletes meg nem értéséről tanúskodik. Természetesen, ha azt sem definiáljuk, hogy miről beszélünk, akkor azt nehéz jól felosztani. Az “elfogadhatatlan” kockázatok közé sorolja a “veszélyes viselkedést ösztönző hangtámogatást használó játékokat”. A hang generálásnak nem sok köze van a MI-hoz és igen nehezen választhatók el a “minimális” kockázati csoportba sorolt “AI-kompatibilis videojátékok” kategóriájától; és közben van még két kategória. Nem sokat segít az sem, ha “A kizárólag katonai célokra fejlesztett vagy használt MI-rendszereket ki kell zárni e rendelet hatálya alól.” Valószínűleg a probabilisztikus MI alapú védelmi rendszerek sem jelentenek kisebb kockázatot, mint a “veszélyes viselkedést ösztönző hangtámogatást használó játékok”. *Egyértelmű lenne a (a tapasztalat szerint sikertelenül megközelített) magas szintű kognitív képességek, úgy mint viselkedés, érzelmek, hangulat, stb. analízisének és abból következtetések levonásának, döntések meghozásának tiltását bevezetni, és egyáltalán a probabilisztika használatát felülvizsgálni.*

Nehéz lesz elválasztani a “magas kockázatú” csoportba eső “hitelpon-tositás, amely megtagadja a polgárok hitelhez jutásának lehetőségét” (amit feltehetően a pénzintézetek végeznek) és a “kormányok által végzett tár-sadalmi pontszámoktól” tevékenységeket (mondjuk, önkormányzatoknak már szabad). Arról nem is beszélve, hogy ugyanazt az információt a TI is összegyűjti és értékeli. Az adatbázisok megfelelő kiválasztásával és az adatvédelmi rendszabályok betartásával össze lehet gyűjteni a szükséges információt, azokat (szükség esetén törvényben meghatározott) súlyokkal valamiféle merit-függvénné alakítani és annak alapján *determinisztikus és indokolható* döntést hozni, vagy arra javaslatot tenni. Mindezt a szerző ál-tal javasolt determinisztikus inkrementális osztályon belül maradva. Nincs különbség a TI és a MI által összeállított determinisztikus (azaz algorit-mizálható és ezért tökéletesen indokolható) működés között. Ha viszont ebben (teljesen szükségtelenül, de az EU által megengedetten) megengedjük a probabilisztikus módszereket, akkor a fürdővízzel kiöntöttük a gyermeket is. *Ezért fontos a fogalom megfelelő meghatározása és a MI megfelelő osz-tályozása.* Ebben az esetben még csak nem is túlságosan “haladó” algoritmus szükséges. De semmiképpen nem MI.

Az egyértelműen hatósági jogkörű “magas kockázatú” területeken a tiltás pl. nem különbözteti meg, hogy a MI egy tanúkihallgatás hangfelvételének csupán szöveg átiratát készíti (ami a 2.1 ábra szerint már jól beállt deter-

minisztrikus technológia) vagy tartalmi kivonatot generál a tartalom és a megadott minták alapján (ami LLM alapú, a leírtak szerint megbízhatatlan technológia); lásd még a D.2 függelékét. Nem szerepel, hogy mondjuk egy orvosi diagnózisnak csupán “tartalmilag és formailag hasonlóknak” kell lenni a MI betanításához használtakhoz, vagy a páciensnek is köze legyen hozzá, lásd a MI diagnosztikus képességeit a COVID-19 által érintett betegek prognózisában, lásd a 2.8 szakaszt.

Vitát fog kiváltani, hogy az érzékelőkkel és szabályzókkal felszerelt betegőrzők, amikről eddig azt kellett állítani, hogy intelligensek, ezentúl nem fognak-e az immár tiltott MI témakörébe tartozni, mivel “the classification of high-risk applications will now include AI systems that pose significant harm to people’s health” (legalábbis az előző változat szerint). Most már *például* csak a robotsebészetben tiltott (ami a ChatGPT fejlesztőinek véleménye szerint, lásd a D.4 függelékét, ellentmondásosan megengedhető). Például a szűrővizsgálatokban nem, pedig a MI gyengén diagnosztizál, lásd a 2.8 szakaszt. *Jelenleg (2023) nincs elfogadott MI meghatározás.* Az egyszerű automatikus vezérlés az EU 2018-as meghatározása már MI-nek minősült, a 2022-es meghatározása szerint csak akkor, ha szoftver végzi. Hardver alapon viszont ma is lehet csinálni, akár akkor is, ha probabilisztikusan működik.

Érdekes feladat lesz az ismeretlen működésű “fekete dobozról” olyan megkövetelt dokumentációt készíteni amely “minden szükséges információt tartalmaz a rendszerről” a rendelet szerint. A “korlátozott kockázatú” rendszereknek mindössze ismertetniük kell, hogy “ők” gépek, és ettől eltekintve azt csinálnak, amit akarnak; még a “magas kockázatú” osztályba tartozó adatbázis-korlátozást sem kell biztosítani. Úgy tűnik, *az oktatási és kulturális bizottság hiába javasolta “a nagy kockázatú MI-alkalmazások III. mellékletben szereplő jegyzékének az oktatás, a média és a kultúra területével való kibővítését”.* Azaz a tanulni vágyó vagy információt kereső emberek ki lesznek téve a generatív MI téves és félrevezető információinak.

“A szabad és nyílt forráskódú mesterségesintelligencia-komponensek fejlesztői e rendelet alapján nem kötelezhetők arra, hogy megfeleljenek ... az arra a szolgáltatóra vonatkozó követelményeknek, amely ezt a szabad és nyílt forráskódú MI-komponenst használta”. Azaz *a szolgáltató vétkes benne, ha megsértette a rendeletet, a fejlesztő meg nem kötelezhető (vonható felelősségre). A zárt forrású kódok fejlesztői (és jelenleg ilyenek a veszélyeket okozó szolgáltatók) meg csinálnak amit akarnak, a probléma lényegét jelentő adatokat pedig a szabályozás meg sem említi.* Érdekes elképzelés a rendelkezés betartatására, lásd még a D.3 függelékét.

Úgy tűnik, az EU szakértők nem tanultak és nem felejtettek semmit. Hasonlóképpen finanszírozták egy évtizeden át (magától értetődően szintén

megvalósíthatósági tanulmány nélkül) az agy szimulálásra vonatkozó projektet [107] (mellesleg legalább részben “B” osztályú [133] besorolása szerint) az adófizetők pénzéből. *Az EU akcióterveit kidolgozóknak nem sikerül megérteniük, hogy a tudományban sokkal kevésbé érvényesül a politikai akarat (political will) [58] mint a társadalmi szabályozások területén.*

3 A MI hiba forrásai

A MI hibái számos forrásból eredhetnek, a D függelék felsorol párat, úgy mint az elméleti háttér hiánya (lásd A függelék); a működéshez használt adatok nem megfelelő volta (lásd D.2 függelék); az imitálni kívánt tulajdonság és a matematikai paraméterek közötti megfeleltetés (lásd C függelék); lásd a jobb marketing érdekében végzett feleltlen gyorsításból (lásd B.6 függelék); a technikai megvalósítás működésének nem értése (lásd B függelék) vagy egyszerűen az imitálni kívánt emberi intelligenciának a felületesnél is kisebb fokú megértéséből, valamint az emberi és gépi intelligencia összetévesztéséből.

3.1 Gyenge MI

3.1.1 Stabilitás és megbízhatóság

Egy matematikai eszközzel eredményt előállítani azt a követelményt támasztja hogy a kapott megoldás stabil és pontos legyen. A neurális hálózatok nem feltétlenül tudják megadni még az ismert “helyes” megoldást sem (már ha az ismert); a megoldás csak probabilisztikus értelemben helyes. Rendkívül instabilak és zavar érzékenyek: “hallucinálnak”. A helyes megoldás létezésére vonatkozó, Gödel és Turing nevéhez kötődő matematikai bizonyítási módszerek sem mindig eredményesek [49]. (Az említett bizonyítás csupán a megoldás MI-vel történő előállíthatóságára vonatkozik, a módszerre nem ad útmutatást. A gyakorlati előállításra vonatkozóan az adat-vezérelt előállítási elmélet és az utasítás-vezérelt gyakorlat további nehézségeket támaszt.)

Az emberi tevékenység sebességéhez képest (alacsony számítási hatásokuk ellenére) a MI rendszerek nagyon gyorsak lehetnek és *a rutin munka nagy részét – emberi felügyelet mellett – biztonságosan átvehetik.* Az időjárás előrejelzésben történő használatára vonatkozó tárgyalagos tanulmány [8] szerint *a MI ritkán előforduló* (és ezért a betanításhoz használt adatok között csak kis számban felbukkanó) időjárási *jelenségek felismerésében gyakran hibázik és ilyenkor kiszámíthatatlanul viselkedik; de hibrid rendszer elemeként, főként rutin feladatokra, hasznos lehet.*

3.1.2 Téves felismerés

A MI ellenzői előszeretettel hivatkoznak (emberi intelligenciával szemlélve: elképesztő, fatális) felismerési hibákra [13]. Valóban, *a szokványostól eltérő esetek komoly kihívást jelentenek a MI számára: hiányzik belőle az intelligencia*. Ilyen hibákat még a legegyszerűbb felismerési feladatokban is véthet a MI. Az alapvető ok, hogy a MI – az objektumok számától függően – egészen biztosan még azokat az objektumokat sem ismeri fel, amiket már látott a betanítás során, ha több paramétere van (több adatot használtunk a betanítás során) mint ahány osztályba kell sorolni a felismerendő objektumokat. A többiekre is csak “tippel”: “ez az objektum leginkább arra a fantomképre hasonlít, amit magamnak létrehoztam”. Ennek a hasonlóságnak (amit mérni kell) az elvárt mértéke beállítható, de még csak azt sem tudjuk, milyen az a fantomkép. Meghökkenően kis zavar is teljesen hibás eredményeket idézhet elő, még a viszonylag egyszerűnek tekintett közlekedési tábla felismerések esetében is [59]; lásd még a D.2 függelékét.

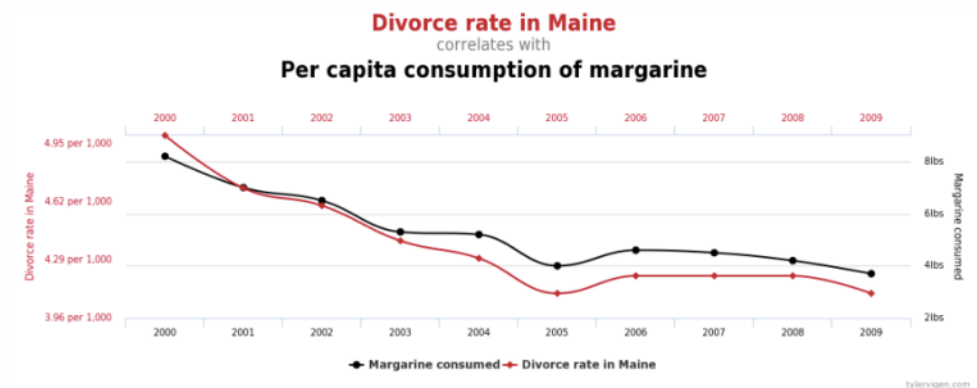
Két ellentétes szempont ütközik. Kevésbé használható az olyan rendszer, amelyik pontosan csak a betanítás során használt objektumokat ismeri fel, a hasonló entitásokat már nem. A hasonlóság mértékét eléggé liberálisan meghatározva viszont ott is hasonlóságot fedezhet fel, ahol az emberi intelligencia semmi hasonlót nem lát (pl. a képfelismerő hókupacként azonosít egy, az oldalán fekvő iskolabuszt, bár az utóbbit 97% pontossággal felismeri, ha az a kerekein áll). Fordítva is igaz: a váratlan helyzetekben (amiben az emberi intelligencia többnyire sikerrel teljesít) a kép felismerő nem ismer fel egy stop táblát ha arra memo-cédulát ragasztottunk vagy bár az önvezető autó felismer a reptéren egy álló kisrepülőgépet de ha az nem szerepel a kikerülő objektumok listáján, akkor nekiütközik. Az ilyen váratlan helyzetek megoldhatatlan feladat elé állítják a MI-t. Az önvezető autók esetében azonban még a “ha nem tudod, mit kellene tenni, akkor állj le” elbánás is gondot okoz: a forgalomban leblokkoló önvezető autó a teljes forgalmat leállítja. Az önvezető autók tömeges alkalmazása egész városrészek forgalmát béníthatja meg. Az önvezető autók megvalósítását maga a MI még évek vagy évtizedek távolába helyezi. Az ilyen vész-leállítás más “intelligens” rendszerek használata esetén is állandó TI jelenlétet igényel. Bár az ilyen esetek a teljes döntési szám elenyésző töredékét teszik ki, ez a probabilisztikusan elhanyagolható esetszám “emlékezetes” marad és jelentős érvet ad a MI ellenzőinek kezébe.

Az ok abban rejlik, hogy *a MI rendszer a betanulás során autonóm módon állítja be a felismerésre használt súlyokat, arra a betanító algoritmusnak/személynek (sőt, EU előírásnak) nincs befolyása*. Ráadásul, lásd 3.1.5 szakasz, ki sem javítható a hiba: nem lehet megmondani, hogy melyik súly és

miért hibás. A rendszer túl sok súlyt használ, és az említett túl-illesztés [102] miatt egészen váratlan kombinációk is téves felismeréshez vezethetnek, továbbá a rendszer saját eredményét nagyon nagy megbízhatóságúnak (lásd 3.1.6 szakasz) állíthatja be.

A felismeréshez (a hasonlóság mértékének megítéléséhez) valamiféle mérőfüggvényt használ mindkét fajta intelligencia. Egy 2017-ben végzett kísérlet [96] (lásd D.2 függelék) azt próbálta meg kideríteni, hogy egy egyszerű neurális háló az “egyenlőség” függvény kapcsolat felismerése során mennyire képes absztrakcióra: a betanítás során az adatok helyesek voltak, de a készletben csak páros számok szerepeltek. Ennek alapján a rendszer a páros számok esetén helyesen inter- és extrapolálta az egyenlőséget, de hibázott páratlan szám esetén. A magyarázat: a használt elektronika számára a hasonlóságot az “1” értékű bitek száma jelenti. Az említett példa esetén a felismerés a “15” szám esetén hibás. Idézzük fel, hogy a “16” számban az “1” értékű bitek száma 1, a “14” számban pedig 3, ezért a “15” számban levő 4 “1” értékű bit közelebb áll a 14-hez (3 bit), mint a 16-hoz (1 bit). Az emberi agy nyilván interpolálna vagy általánosítaná, hogy megegyezik az eredmény az argumentummal.

Ilyen jellegű kérdést nem sokan tesznek fel, azaz a kérdésre adott válasz nincs előkészítve a felhasználóhoz közeli pufferekben, ezért a választ újonnan kell összeállítani. Bár időközben nyilván fejlesztettek az algoritmuson (mint minden nyilvánosságra hozott hiba esetén), így a 2023-ban végzett kísérletben már megállapítja a rendszer, hogy “megegyezik az eredmény az argumentummal”, de nem tud általánosítani, hogy akkor az páratlan számok esetén is érvényes. Az első válasz, hogy felismerte ugyan a szabályszerűséget, de csak páros számok esetén tudja alkalmazni, mert a betanító adatok között csak páros számok szerepeltek. Hogy általánosítsa a szabályszerűséget, több időre lenne szüksége; a válaszadáshoz rendelkezésre álló idő azonban letelt: azt kell eredményül megadni, amit odáig sikerült összeállítani (a kérdés szokatlansága mellett a szándékos betűkihagyás és a használt kétféle nyelv is növeli a végrehajtási időt, lásd D.2 függelék). A válasz kiszámítása azonban elindult és idővel a helyes(ebb) válasz is bekerül a megfelelő pufferekbe, de onnét mindaddig nem kerül elő, amíg rá nem kérdez valaki. Ha ismét feltesszük ugyanazt a kérdést, akkor már a pufferekben szerepelnek az előre gyártott elemek, és a rendszer nem ismétli meg a 2017-ben és az előző kérdésre adott válaszban már elkövetett hibát; a válasz helyes lesz. A rendszer ugyanannyit és ugyanúgy “gondolkodott”, csak a felhasználóhoz közeli pufferek tartalma változott. A számítási idő ablakot előre deffiniáltuk, és az a megszokottnál hosszabb számítási idő esetén rövidnek bizonyult.



Ábra 3.1: A válások számának összefüggése a margarinfogyasztással[1]

3.1.3 Rejtett adat-összefüggések

A MI azokkal az adatokkal működik, amelyeket rendelkezésére bocsátunk. Ha túlságosan sok adatot vagy teljesen szabad kezet kap, akkor olyan összefüggést is megtalál, aminek nem feltétlenül van köze a megoldandó feladathoz (például az elkövető bőrszíne vagy a páciens jövedelmi helyzete; nem ismeri a politikai korrektséget). Továbbá 0.99 korrelációt talál a válások gyakorisága és a margarin fogyasztás között, lásd 3.1 ábra. MI nélkül nem is gondolnánk rá.

Jól feltett kérdésre nagyobb eséllyel kapunk jó választ. Mint fentebb tárgyaltuk, a MI rendszer csak esetlegesen tud abszolút optimumot találni, a betanítástól (a használt és a felhasználáskor ismeretlennek tekintendő súlyoktól) függően megragadhat helyi optimumokban is, lásd részletesebben a [96] közleményt. *Minél pontosabban körülhatárolt módon keressük az összefüggéseket, annál kevésbé valószínű, hogy nem valódi vagy nem érdekes adat-összefüggéseket találunk.* Ha a MI bemenő adatait a valóban szükséges adatokra korlátozzuk ("keress összefüggést a leírásban használt paraméterek között", a túl általános "keress összefüggést az adatok között" helyett), akkor a MI hatékony kereső eszközt eredményez. Ha nem, akkor esetleg saját maga által feltételezett összefüggés kidolgozására fordítja a legtöbb energiát.

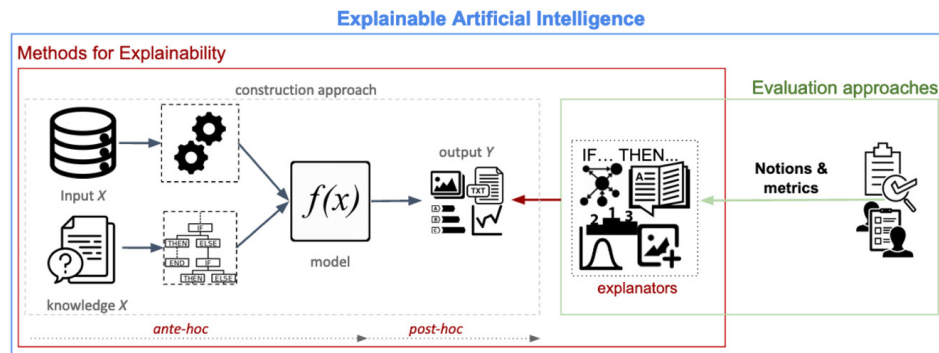
3.1.4 Katasztrófa-szerű felejtés

A B.2 szakaszban említettük, hogy a MI működése során a műveletek eredménye nem korlátos, azaz egyetlen művelettel az eredeti értékhez képest bárhová el lehet állítani a rendszer belső súlyait (bár az esetleges normálás ezt

a lehetőséget valamennyire korlátozza). Ez a hiányosság a betanítás során is komoly “lengéseket”, instabilitásokat eredményezhet (ennek következtében *a betanításkor használatos elemek sorrendje is befolyásolja a betanítás végeredményét*). Igazán katasztrofális hatása azonban akkor van, ha a már beállt rendszernek valami újat akarunk megtanítani, ami ezek szerint alapvetően eltér attól, amire a rendszer beállította saját súlyait. Az újonnan megtanulandó információ felülírhatja azt a súlyt amelyik döntő szerepet kapott az előző felismerésben, azaz a rendszer már nem fogja tudni felismerni azt, amit előzőleg megtanult. Ez abban az esetben fordulhat elő, ha a rendszernek túlságosan nagy szabadságot engedélyezünk, azaz nem szűkítjük a lehetséges keresendő összefüggések számát. *A MI, a biológiai neurális hálótól eltérően, valamennyi bemenetét egyenrangúnak tekinti* és így a neuron csoportok nem “szakosodnak” egy-egy feladatra.

A felismerésekben egy megjelenő váratlan tényező komoly biztonsági lyukat eredményezhet. Menet közben tanuló rendszerek esetében egy ilyen esetet a rendszer úgy illeszthet be addigi ismeretei közé, hogy új tanulnivalót talált és elfelejti a régi, többé-kevésbé jól működő tapasztalatait. Sőt, ezután csak olyasmit ismer fel, ami az eredeti tervben nem szerepelt. A rendszer újra tanítása viszont költség- és időigényes művelet. Mivel minden egyes új minta az összes súly megváltoztatásával jár, a minták (és természetesen a különböző adatkészletek) sorrendje nagyon is számít. Továbbá semmi garancia arra, hogy a rendszer az újra tanítás után is működőképes marad vagy hogy egyáltalán arra a feladatra betanítható. A működés elméletének hiánya és az indoklás hozzáférhetetlensége miatt még azt sem lehet kideríteni, hogy miért.

Lényegében ehhez hasonló jelenséget idézhet elő a valós idejű működés kényszere. A válasz adásra valami észszerű határidőt (időablakot) kell szabni, különben a felhasználó érdeklődése alábbhagy. Amit odáig össze lehet állítani, azt megmutatjuk. Még az előre betanított rendszerek esetében sem lehetünk biztosak a kiszámítható működésben. A rendszerek betanítási ideje igen hosszú, és az annotálással és egyéb előkészítő műveletekkel együtt azzal jár, hogy jelenleg a működő ChatGPT két éves, elavult adatbázissal működik. Ennek ellensúlyozására próbálják az Internetről begyűjtött friss információt a stabil és ellenőrzött, de régi adatkészlethez keverni. Ennek komoly veszélye van: a Google Bard rendszere éppen a bemutatóra került olyan állapotba, hogy hibás válaszokat kezdett adni [24], ami a közlemény szerint 100 milliárd USD veszteséget okozott a cégnek.



Diagrammatic view of Explainable Artificial Intelligence with interaction between methods for explanations and their evaluation approaches.

Ábra 3.2: A MI indoklásának elvi ábrája, a használt módszerek és értékelési eljárások kölcsönhatásának diagram-szerű ábrázolása [143]

3.1.5 Indokolhatóság

A MI ugyan mutat egy-egy megoldást, de nem teszi lehetővé annak indoklását, hogy miért éppen azt. A D.1 függelék szerint másféle “gondolatmenetet” használ. Elvileg az algoritmusok, adatok és a belső dokumentáció ismeretében sok mindent ki lehetne találni, de egyik sem nyilvános, lásd a D.3 függelék. Elvileg ugyan lekérdezhethetjük a rendszer súlyainak értékét, de azok semmi közvetlenül következtethető módon nem állnak kapcsolatban a döntéssel és a választható lehetőségekkel. A rendszer a bemenetektől az éppen aktuális súlyokkal származtat egy értéket, ami a bemenetek értékének a megfelelő súlyokkal szorzott rész-összegeiből áll össze. A helyzet lényegében annak felel meg, mint amikor egy olyan polinomot használunk adatok illesztésére, amelyiknek foka nagyobb, mint az adatpontok száma. Bár az ilyen illesztésre is létezik matematikailag indokolható módszer [102], ennek ára az, hogy *csak statisztikus értelemben megbízható eredményt tudunk származtatni*; még csak azt sem mondhatja el a rendszer, hogy “ezt felismerem, mert már láttam a betanításom során”. A betanításhoz használt adatkészlet elemeit is csak probabilisztikus értelemben ismeri fel. Ráadásul, ma már *az óriásira duzzasztott paraméter szám miatt, csak a paraméterek töredéke tud részt venni a számításban*, így a válasz előre gyártott elemekből felépített sablonos válasz, azaz indokolhatóságról beszélni illuzórikus.

A MI tehát valami olyasmi, amit “fekete doboz” [105] névvel illetnek, azaz működéséről semmit sem tudunk. Vannak bemenetei és kimenetei, és működése során ezek között teremt összefüggést, a maga sajátos és titokzatos

szabályai szerint. Szórakozásként jól el lehet vele beszélgetni, a digitális világban a csevegéshez szokott emberek nagyra értékelik a fáradhatatlan csevegő partnert, “aki” jól tájékozottnak is látszik és azonnal válaszol, bármikor. Ha azonban az emberek használni is szeretnék az eredményeit, főként döntés hozás vagy támogatás céljára, akkor felmerülnek a “miért jutott erre az eredményre” és “minek alapján” típusú kérdések, azaz a válasz indoklása/megértése iránti igény. Megpróbáljuk a MI-t legalább “szürke doboz”zá tenni, a TI bevonásával.

A MI indoklása iránti igény számos kutatást indított el, ezeknek jó összefoglalása található az [143] közleményben. A MI természetesen nem tudja, mit elemez (számára csak paraméterek összefüggései léteznek), ezért a kimeneti eredmény összeállításához, mint az 3.2 ábrán látható, a kutatók a TI által összeállított összefüggéseket is betáplálják. Azaz, a TI meghatározza, mik a lényegi paraméter összefüggések, és a MI feladata a kevésbé lényeges paraméter összefüggések megtalálása. Ilyen módon a MI által használt paramétereket kényszerítjük bizonyos értékek felvételére és rögzítésére, így a lényeges összefüggések felismerésében már nem tévedhet nagyot. Az indoklás megfogalmazásában ismételten használjuk a TI által megadott összefüggéseket és a TI által speciálisan erre az esetre megadott fogalmakat sőt szöveg részleteket, aztán rendszerint a TI átnézi a kapott eredményt. Utána pedig eladjuk, hogy mire képes a MI. Bár a feladat lényeges és kritikus részeit a TI végzi, a MI sebessége valóban sokat segíthet az eredmény előállításában, lásd pl. [8].

Az indoklásra vonatkozó kutatások eredményei megerősítik, amit a paraméterek címkézésével kapcsolatban a gyenge és az erős MI vonatkozásában mondunk a C.1 függelékben. A lényeges összefüggések külső megadásával az “unlabelled” paramétereket “labelled” jellegűvé tesszük, ezáltal elmozdulunk a gyenge AI irányába. Azaz, *a magyarázat hozzáfűzése nem csak érthetőbbé teszi a TI számára a MI működését, hanem pontosabbá is.*

3.1.6 Megbízhatóság

A működés során a MI megvizsgálhatja azt is, hogy mennyire biztos a következtetésében. Ezek a módszerek azonban nagyon sok számítási kapacitást igényelnek és a rendszereknek többé-kevésbé valós időben kell működniük, lásd az 4.1 szakaszt. Emiatt az eredmény megbízhatóságának kiszámítását általában nem alkalmazzák, mivel az a számítási erőforrásokat a fő feladat megoldásától vonná el, legyen szó akár autonom jármű vezetésről vagy határidőre elkészítendő orvosi diagnózisról a hálózaton át ennyi idő alatt elérhető (és vélhetően egyre nagyobb számban MI által készített) diagnózisok

alapján.

A [95] tanulmány (még a MI és különösen a csevegő programok felfutása idején) kifejezte azt a reményét, hogy a MI növekvő szerepet fog játszani a pszichiátriai és kognitív fejlődési rendellenességek kezelésében, főként olcsósága és folyamatos rendelkezésre állása miatt. Kérdéses, hogy milyen kezelést fog végezni az az “intelligencia”, amelyik folyamatosan depressziós és szorongó páciensektől tanul. Az eddigi tapasztalatok szerint a többségi véleményhez igazodik. Nem biztos, hogy érdemes kipróbálni.

Nagyon hasznos módszer a serdülő ifjúság tudását kreatív esszék készítésével fejleszteni. A MI által generált (potenciálisan hamis és félrevezető) tartalmak erőteljes terjedésének korában már az információ gyűjtése is veszélyes a forrás megjelölése és értékelése nélkül, a MI által előállított tartalom bemutatása pedig teljesen tönkreteszi ezt a lehetőséget. Leginkább azokat sújtja, akik valóban tanulni akarnak. Nem mellesleg: *a tartalom forrására vonatkozó kérdéssel nemcsak a tanár leplezheti le a MI-t használó diákot, hanem a Turing teszt sikerét is meg lehet hiúsítani.* Érdekes “ellenrakéta” ötlet [34], hogy maga a tanár generáltat a ChatGPT-vel esszét (azonos témából, de különböző szövegezéssel), és a diákok feladata elkülöníteni a valós tényeket a valótlantól. Egy ilyen feladat a diákok számára a ChatGPT “szellemiségére” is fényt vet.

Komoly probléma lehet egyrészt a szöveg “eredetiségét” vizsgálni, másrészt a MI által generált szöveget felismerni és elkülöníteni a TI által készítettől. Bár pl. tudományos közlemények esetén lehetséges újabb paraméter (pl. stílus) alapján jól felismerni LLM által készített szöveget [110], az nyilván csak addig működik, amíg ilyen paraméter nem kerül be a szöveg előállításba.

3.1.7 Háttér ismeretek

Ha a MI nem ismer egy tényt, akkor “kitalálja” (lásd D.2 függelék), jól vagy rosszul; aztán a válaszában már nem különbözteti meg a valódi tényektől. Általában véve igaz, hogy a több munkavégző által végzett munka hamarabb befejeződik, mintha csak egyetlen munkavégző lenne. A párhuzamos programozás tipikus korlátját, hogy t.i. bizonyos folyamatok nem bonthatók részekre, a MI nem feltétlenül tudja. Ennek következtében hibátlanul kiszámolja, sőt, indokolja, hogy ha egy nő 9 hónap alatt hord ki egy gyermeket, akkor 9 nő ugyanezzel a feladattal egy hónap alatt végez [24]. Természetesen, utána tökéletesítik az “algoritmusát”, de attól nem lesz intelligensebb, csak bonyolultabban dolgozik.

3.2 Erős MI

Az erős MI értékelésekor két szempontból kell értékelnünk: célja és technikai megvalósítása szerint. Az egyik, hogy az erős MI ambíciója a teljes emberi intelligenciát megvalósítani, a másik, hogy technikailag csupán sokkal nagyobb számú és eltérő viselkedésű/jellegű paramétert kezel, de lényeges (minőségi) működési különbséget nem tud felmutatni.

3.2.1 Generatív MI

A nagy nyelvi modellek bevezetésekor [14] lényegében egy szöveg kiegészítésre használt ártalmatlan segítő funkció, lásd még a 2.3 szakaszt, szolgált modellként. Ezekben a MI (ha annak lehet nevezni az előfordulások gyakoriságán, lásd (2.3) eloszlás, alapuló szövegkeresést) csupán felajánlja a lehetőséget, hogy a felhasználót mentesítse a felesleges billentyű nyomkodás alól. A TI jelen van és felülbírálja az esetlegesen nem tetsző javaslatot. A következő betű vagy szókapcsolati tag valószínű értékének előrejelzése azonban kiterjeszthető nagyobb nyelvi egységekre és végső soron akár teljes írásművekre. Az elvéből következik, hogy a felajánlott szöveg már szerepelt a rendszer által látottak között, tehát *nem alkot új szavakat, szókapcsolatokat, csupán a létezőket kapcsolja össze főként előfordulási gyakoriság szerint*. Az így analizált és generált új szöveg mérete gyakorlatilag tetszőlegesen növelhető, és a nyelvi egységek közötti kapcsolatok keresésével egyre bonyolultabb és sokrétűbb valószínűségi eloszlások alapján lehet új szöveget előállítani az eloszlások kiszámításához használt régi szövegek alapján. Mivel létező, és TI által létrehozott szövegrészeket használ, részben a probabilitás, részben a működés technikai kivitelének (főként busz és hálózat használat), részben a valós idejű működés által szabott időkorlátnak, részben az idővel folyamatosan változó hálózati bemenő-adat fájloknak köszönhetően, a generált kimenet meglepően változatos: a szöveg első pillantásra értelmesnek, szakszerűnek és valószínűnek hat. Aztán a szöveg után különféle bővítményeket (keresést, matematikai “motort”, képet, hangot, stb.) kezdtek generálni.

Ezen alapulva a kezdeti felhasználások során valóban lehetett meglepően hasznos eredményeket is elérni, de *rasszista, szexista vagy vallásos szövegeket* is készíteni, sőt első látásra nagyon is plauzibilis tanulmányokat, amelyekről az alaposabb vizsgálat során kiderült, hogy *pontatlanok, légből kapottak, nemkívánatosak és előrejelezhetetlenek* (azaz befolyásolhatatlanok, szabályozhatatlanok). *Felmerült, hogy az ilyen rendszerek csalásra, dezinformálásra, félrevezetésre is használhatók, de az ipari lobby hangja akkor még elnyomta a vészcsengőjét*. Az ilyen rendszerek néhány fontosabb tulajdon-

ságára is fény derült.

- A modell méretének növekedésével annak képességei előre nem látható módon változnak
- Sokféle képességgel rendelkezik (összefoglaló készítés, csevegés, keresés; kód-, kép, hang- stb. generálás)
- Néhány képessége meglepően jó, de nem képes szándékot, célt, okozatot felismerni

Ez a kezdeti állásfoglalás helyesnek bizonyult, bár a közleményekben hol az előnyök, hol a hátrányok kerültek túlsúlyba. Konkréten a ChatGPT értékelésére vonatkozóan is komoly kétségek merülnek fel. A fejlesztésben (és annak sikerében) érdekelt egyik cégvezető értékelése szerint [24] "a ChatGPT már bizonyított, de még hosszú az út a valóban intelligens gépek építéséig", továbbá a cikk címe és mondandója, hogy a ChatGPT-vel szembeni várakozások erősen túlzottak. De hát aki a ChatGPT szolgáltatást veszi igénybe, olyan kérdést tesz fel, hogy "az eddigi többségi (gyakorisági) vélemény szerint, ebben a dokumentumba a pár szóval megadott fogalmakkal kapcsolatban mi az ÉN VÉLEMÉNYEM". Erre kap választ.

Általában, az erős MI értékelésére vonatkozóan az amerikai "nagy játékosok" platformja [115] is óvatosan úgy fogalmaz, hogy *hosszú még a út, amíg elérjük a célt*. Még a növekedés piaci fenntartásában érdekelt szereplők véleménye szerint is, *a számítógépes növekedés MI-vonatkozású iránya nem tartható fenn*, különösen annak egyrészt nagyon nagymértékű energia fogyasztása, másrészt kis hatásfokú számítási teljesítőképessége miatt, lásd még a 5. szakaszt. Ez utóbbi vonatkozás azért nem látszik külön, mert a szöveg feldolgozás alapvetően számítógépes kereső motorok használatán és hálózati átvitelen alapul és egyrészt a hatásfok és az energia fogyasztás mérése nem volt szempont az eddigi fejlesztés folyamán, másrészt technikailag nehéz elkülöníteni a különböző célú felhasználásokat egy nagyon sok rendszert elérő feladat megoldás során.

A technológia úgy viselkedik, mint egy látszólag sokat tudó, de saját képességeinek korlátait nem ismerő és tudásának megbízhatóságát értékelni nem tudó, valós és valótlan állításokból építkező, felelőtlenül fecsegő alkalmazott. Összerakja a valódi alkalmazottak vélemény morzsáit, kombinálja helyes és téves következtetésekkel, és abból az emberi intelligenciával kombinálva már akár valódi céges titkokat is ki lehet hámozni, de akár félreinformálni is. Bár a technológia még igen képlékeny, máris akadtak (2023 közepén) olyan cégek (pl. Samsung) és államok (pl. Olaszország) amelyek egyszerűen

betiltották az ilyen technológián alapuló termékek használatát, működésük kiszámíthatatlanságából adódó veszélyek miatt. Hasonló okokból döntött az EU is a MI betiltásáról. Bár a betiltás indoklásában nem szerepel az általunk a 2.3 szakaszban említett kategorizálás, a tevékenységek leírása egyértelműen a probabilisztikus MI fogalmára utal. A már nyilvánosságra hozott és bárki által használható technológia mögött nyilván ott vannak az új fejlesztések és a cégek kívánnak azok nyilvánosságra hozatalával (ezért javasolták az ezzel a technológiával foglalkozó cégek, hogy fél évre szüneteltetik a fejlesztést, és hogy ez idő alatt hozzák létre a megfelelő szabályozást). *Az adatok számítógépes generálása tovább növeli a tárolás és feldolgozás iránt támasztott igényeket és beláthatatlanná teszi a fejlődési tendenciákat.*

Ha netán valakinek ambíciója van tévinformációkat terjeszteni, de nincs ötlete/tudása, hogy hogyan tudja megcsinálni, akkor a MI szívesen útba igazítja. A [42] közlemény utat mutat hozzá, hogy technikailag (milyen platformokon) és milyen érvekkel lehet szövegesen a védőoldásokról lebeszélni. Minimális fantáziával bármely másik cél elérésére is kaphatunk útmutatást.

3.2.2 Probabilisztikus viselkedés

A szakterületek igazi szakemberei is óvnak a felelőtlenül blablázó technológia használatától [21]. A gyenge MI-hez hasonlóan, itt is probléma a katasztrofális felejtés (lásd 3.1.4 szakasz), megakadályozva hogy több területre is használni lehessen egy-egy rendszert, minden egyes szűk problémára pedig nem tudunk egy-egy külön rendszert használni; különösen, ha a problémák kötődnek egymáshoz. Úgy tűnik, hogy a 2.3 pontban említett osztályozás szerinti *gyenge és erős MI között nincs minőségi különbség*. Az erős MI megpróbál a "több" területből "sokkal több, szinte minden" szintre fejlődni, de minőségi ugrást nem sikerül elérnie, csupán a több területnek megfelelően lényegesen több erőforrást használ. A technikai kivitelezés részletei nem ismertek, ezért nem tudni, hogyan is oldható meg egy-egy újabb terület hozzáadása. A nyilvánosságra kerülő részletek alapján arra következtethetünk, hogy a tanítás során váltogatni kell a területeket, ami rendkívüli mértékben megnöveli a betanítás idejét és költségét. Ennek következtében az indokolhatóság hiánya (lásd 3.1.5 szakasz) is ugyanaz marad, mint a gyenge MI rendszereknél: gyakorlatilag egy igen bonyolult funkcionális, de megjósolhatatlan viselkedésű "fekete doboz"t hozunk létre, bár a indoklásnál bemutatott technika alkalmazásával javítható a helyzet. A következtetés, hogy nagyon kockázatos döntés ilyen rendszereket "élet/halál" feladatokra használni. A rendszer nem stabil, és teljesen kiszámíthatatlan, hogy milyen külső hatás milyen változást fog előidézni. Gyakorlatilag előre jelezhetetlen,

hogy egy rendszer akár ugyanazt a feladatot kétszer egymás után is jól oldja meg (lásd D.1 függelék). Az egyetlen, amit a matematikusok és a gyártók is garantálni tudnak, hogy (megfelelő feltételek fennállása esetén) *a kapott eredmény probabilisztikus értelemben megfelelően viselkedik* (bár a tapasztalat szerint a probabilisztika jellemzői változhatnak a használat során), lásd a [102] közlemény tárgyalását és az ott felsorolt hivatkozásokat.

A rendszer nyelvi analízist és szintézist végez, ennek megfelelően a tartalmi analízishez nem sok köze van. A generált szöveg probabilisztikus értelemben tartalmilag és formailag hasonló lesz a betanítás során használt szövegekhez, a konkrét esetre vonatkoztatott tartalom pedig szintén probabilisztikus értelemben lesz hasonló a használt adatkészletben látott esetekhez. A generált szöveg hasznossága és hihetősége annál nagyobb, mennél kisebb az emberi olvasó szakértelme. Leginkább ezért tűnik úgy, hogy “hihetetlenül pontos” a MI által készített közlés. Ha az emberi olvasó tudja a tényeket, a kép sokkal kedvezőtlenebbé válik. *A MI nem termel új információt*, csak (sajátos szempontjai alapján) *átrendezi a már meglevőt*. (Illetve, ha termel, abban nincs sok köszönet, lásd D.2 függelék.) Bár ezt valóban gyorsan megteszi, az átrendezés módja önkényes és átláthatatlan indokú. Ráadásul egyre több a meglevő információban a MI által átrendezett.

3.2.3 Következmények

Orvosi diagnózisok esetében, ha a vizsgált adatkészletben jellemző volt a beteg tünetegyütteseinek megfelelő betegség, akkor a MI megtalálja a betegséget és “jó diagnózist” ad. Ha viszont nem volt előzmény, amint a COVID-19 megjelenésekor tapasztalták, akkor semmire sem jut, lásd 2.8 szakasz. Hasonló volt a tapasztalat a 2013-as influenza járvány előrejelzésekor [88]: a stagnálást látszólag jól, a csúcsot láthatóan rosszul jelezte. A tapasztalat szerint nagyon erősen függ a MI használhatósága attól, hogy a betegség tüneteinek megadásakor milyen szavakat használ a kereső szöveg. A probabilisztikus viselkedés a gyakorlatban azt jelenti, hogy ha a rendszer által látott szövegek túlnyomó többsége orvosi diagnózis volt (ami jellemzően nagy gyakoriságokkal használ orvosi műszavakat), akkor a betegek (szakszavak nélküli) öndiagnosztikai kereső szövegével kevés formai hasonlóságot talál és a laikusok számára nagyon kevésbé lesz hatékony. Viszont, a folyamatosan tanuló rendszer, mivel nem tudja forrásainak szakszerűségét értékelni, egyforma súlyúnak veszi a laikusok szövegét és az orvosokét. Eleinte csak kisebbségbe kerül a nem-szakmai szöveg, a laikus érdeklődés fokozódásával azonban előnytelen irányba változnak a probabilisztikus gyakoriságok és ennek megfelelően változik a generált szöveg is. A felhasználó egyre kevésbé

szakmai véleményt fog kapni.

Csevegő fórumok esetén a katasztrofális felejtés csak lassan és fokozatosan következik be: a probabilisztikus generálás nem egyetlen súly alapján felejt (végül is, a szöveg mögötti tapasztalat ugyanaz), de a hatás ugyanaz marad. Az eddigi működtetések során kifejlődtek rasszista, politikai nézeteket valló, stb. fórumok, mivel a generált szövegek ugyanolyan stílusú, formájú és tartalmú szövegek, mint amilyeneket az emberi hozzászólók is produkálnak, és a továbbiakban azzal egyenértékű nyersanyagként szolgálnak. A generált szövegek előállításának sebessége azonban lényegesen nagyobb, mint az emberi felhasználók által generált szövegé, a tartalmat illetően pedig éppen ellentétes a gépi és az emberi viselkedés. A gép igyekszik a többségi véleménynek megfelelő tartalmat generálni, az emberi felhasználók pedig egyszerűen elhagyják az illető fórumot, ahol a "felhasználók" zöme nem az ő nézeteiket vallja. Ilyen értelemben, a gép átveszi az uralmat és önmagával egyetértve ragozza a többségi (és ezáltal egyedülívé váló) véleményt.

Oktatás esetén előbb-utóbb megjelenik az összes elképzelhető feladat megoldása a MI által elérhető adatkészletekben és az emberi tanulási feladat a gépi tanítás eredményének megtalálására egyszerűsödik. Minden emberi tanuló, aki hozzáfér a MI adatkészleteihez, kiváló eredményt fog felmutatni, de összes képessége a MI eredményeinek megtalálására és az emberi tanárnak történő benyújtására fog korlátozódni. Az emberi egyedek nem lesznek képesek tanulni. A tanárnak azonban az indoklás és az információforrás hiánya módot ad egyrészt a leleplezésre, másrészt – bár sok munkával – elgondolkodtathatja a diákot, hogy ki, miért és valójában mit mondott [34]. A [7] fórum úgy teszi fel a kérdést: "A csevegő robot vagy a diák oldja meg a feladatot a kurzusodon?" A [125] tapasztalata szerint a korábban elérhető benchmarkok teljesítésére "optimalizált" MI rendszerek átlag alatt teljesítenek olyan tesztek esetén amelyek még nem léteztek az MI rendszer létrehozásakor.

Nagyon sok múlik a helyes használati mód kialakításán. Az 2.1. ábra szerint az optikai karakter felismerés már megtalálta a helyét és a produktivitás fázisában van. Nagyon megkönnyíti életünket, ha pl. gépkocsi rendszám leolvasására használják, és lényegesen felgyorsítja egy parkolóház működését. Nagyon felgyorsítja gyermekeink betű felismerését is, de aki csak ilyen segítséggel ismeri fel a betűket, az nem fogja tudni azokat szavakká összerakni és nem fog megtanulni olvasni. A MI túlzó vagy helytelen használata gátolhatja a TI fejlődését; ilyen értelemben versenytárrá válhat. Ugyanez vonatkozik a magasabb szintű használatra. Ha a beteg használja a MI-t annak eldöntésére, hogy a lefénnyképezett bőrtünete kóros elváltozás lehet-e, az lehet terhet az orvosi ellátásról. Ha a bőrgyógyász rezidens, akkor nem lesz belőle szakorvos. A használat észszerű módját egyértelműen szabályozni és mértékét

feltüntetni kell.

3.2.4 Jellemző paraméterek

A C.1 szakaszban már említjük a paraméterek helyes kijelölésének jelentőségét. Az erős MI esetében ennek jelentősége sokkal nagyobb, részben a paraméterek lényegesen nagyobb száma, részben azok jellege miatt. A nem egyetlen mérőszámmal hanem összetartozó paraméterek használatával jellemezhető jelenségek modellezésére a TI igen, a MI legfeljebb korlátozottan alkalmas. Míg a gyenge MI esetén tipikus, hogy jó jellemző paramétereket tudunk találni (a következő lépés kiválasztása egy játékban, egy karakter felismerése, stb.), addig az erős MI esetén ezek kevésbé megfoghatók (pl. a tömeg hangulata, a fórum vélemény irányulása, az osztály szorgalma). A nagyobb számú paraméter jelenléte eleve nehezíti a helyzetet (mivel az alkalmazott matematikai módszer miatt minden mindennel összefügg), az egyes jellemző paraméterek “elosztása” a MI paraméterei között végképp esetlegessé teszi a megoldhatóságot.

3.2.5 Megbízhatóság

Az eredmény megbízhatóságának vizsgálata hasonló okokból marad el, mint amiket a 3.1.6 pontban leírtunk. Emellett a real-time viselkedés kötelezettsége (és az exponenciálisan növekvő adatmennyiség) tovább csökkenti az eredmények megbízhatóságát. Ugyanebbe az irányba foghatni a MI által generált adatok növekvő részaránya is. Ez utóbbiak közvetlen következménye a működés elavulása: megoldhatatlan a MI rendszerek naprakészen tartása. Már is sok hónapos lemaradásban vannak (lásd D.5 függelék; a 2023-ban bemutatott GPT-4 is 2021 szeptemberi adathalmazzal dolgozik: az algoritmus fejlesztés és adat előkészítés elvált egymástól) a nagy igénybevételnek kitett rendszerek: a korábban betanított adatok miatt egyszerűen nem tudnak az utóbbi hónapok eseményeiről; nemcsak sok időt vesz igénybe az újabb betanítás, hanem a betanítási idő hosszabb, mint amennyi időről újabb ismereteket kell betanítani, de kockázatos is: a rendszer akár mindent elfelejthet. *A betanulási idő exponenciálisan nő mind az adatmennyiség, mint a paraméter mennyiség következtében; a lemaradás exponenciálisan növekszik. A MI működési modellje önmagában is fenntart-hatatlan, aktuális működése pedig előreláthatatlan.*

További problémákat vet fel a probabilisztikus működés. Mivel a generált szövegbe a korábban látott szövegek gyakorisága alapján kerülnek szókapcsolatok, komoly veszélye van annak, hogy egy generált tartalmi összefoglalóban

egy technikai újdonság, újabb irányítási elv, vagy a kihallgatáskor elmondott szokatlan cselekmény egyszerűen kimarad a generált szövegből. “Olyan eddig nem volt, tehát nem is lesz.” Kárpótlásul viszont bekerülhet olyan, amit a MI tévesen következtetett, mert az ilyen esetekben elő szokott fordulni. A Large Media Model bevezetése pedig kétségesse teszi a hang- vagy képfelvétel bizonyító erejét is.

A kockázatoság jelentőségét mutatja, hogy a National Institute of Standards and Technology (NIST), USA [104] külön szabvány tervezetben foglalkozik annak kezelésével.

4 A MI használata

A MI lényegében teljes egészében a hagyományos, utasítás-vezérelt működésen továbbá számítástudományi alapokkal rendelkező eszközökön és módszereken alapszik, ezért rendelkezik az összes olyan sérülékenységgel, amit eddig már megismertünk. Mivel azonban megpróbálja az adat-vezérelt működést emulálni, ezzel további sérülékenységeknek enged utat.

4.1 Valós idejű működés

A számítási elméletnek az is Neumann által megfogalmazott alapfeltétele, hogy a számítás elkezdéséhez az összes bemenő paraméternek meg kell érkeznie, továbbá az eredményt csak azután lehet elszállítani, miután a számítás befejeződött. Neurális hálózatok esetén az utóbbi feltétel azt jelenti, hogy a rendszerben valamennyi elküldött üzenet megérkezik a helyére (nem-lineáris, azaz pl. visszacsatolást tartalmazó rendszerek esetében a visszacsatolt üzenetek sem generálnak további állapot változást), ami (különösen buszt használó elektronikus hálózatok esetén) azt jelenti, hogy az eredmény előállítási ideje rosszul definiált: gyakorlatilag nincs felső határa (sőt, mint a B.2 szakaszban bemutatjuk: elvileg is végtelen hosszú). Mint általában a feladatok megoldásakor, a MI esetében is véges ideig vagyunk hajlandók várakozni a számítógépes rendszer eredményére, legyen szó akár egy ön-vezető gépkocsi döntéséről egy adott forgalmi helyzetben vagy egy orvosi MI diagnózis felállításáról. A viszonylag egyszerű esetekben (a lehetséges kézírás minták vagy a valaha lejátszott összes Go játszma bemutatásával) ezt meg tudjuk tenni a rendszert “betanulás” üzemmódra kapcsolva, majd “teszt” üzemmódban a megtanult minták alapján, a beállított súlyokkal már viszonylag gyorsan működtetni tudjuk a felismerő rendszert. Ebben az esetben megengedhetjük magunknak, hogy a betanulás során erős hardvert (nagy fogyasztás és nagy költség, lassú működés) használjunk, majd a betanulás

után az eredményt (az előállított szinaptikus súlyokat) az olcsó cél-rendszerre átmásolva, a megszerzett tudást ott hasznosítsuk (kis fogyasztás, kis költség, gyors működés).

A véges működési idő elérésére (a valós idejű programozásból ismert módon) kompromisszumot kell kötnünk a pontosság és a számítási idő között. Általában korlátozni kell a felhasználható (fogadott) adatok mennyiségét valamint a számítás hosszát (amibe a rendszer belső adatszállításának idejét is bele kell értenünk). Abban az esetben, ha a rendszer menet közben tanul és az Interneten gyakorlatilag korlátlan mennyiségű adat áll rendelkezésre az eredmény kiszámításához, az időben való válaszolás kényszere a felhasználható adatok mennyiségének csonkolását követeli meg. Az adatok beszerzési ideje nem haladhatja meg a válaszig rendelkezésre álló időt. Ha az adatok eltérő forrásokból származnak (netán időben változnak), akkor az egymás utáni válaszok nem lesznek egyformák, a MI rendszer ugyanarra a kérdésre többé-kevésbé eltérő válaszokat ad (lásd D.1 függelék). Ugyanez úgy is előállhat, hogy a rendszer a válaszadás során módosítja saját működését, beleértve a probabilisztikus súlyozást és a katasztrofális felejtést. Ráadásul a generatív rendszerek iszonyatos sebességgel termelik a többségi vélemény alapján gyártott új tartalmakat, így a TI általi ellenőrzés beavatkozási gyorsaság és ellenőrzendő mennyiség szempontjából is reménytelenné válik.

A generatív működést is tartalmazó használat esetén jelentős időt vesz igénybe a válasz generálása is, és mivel természetes módon ez a működési folyamat van a lánc végén, a működési idő korlátozása elsősorban a szöveg generálást érinti. Vélhetően ez a folyamat nem egyszerűen egy szekvenciális lánc végén van, hanem részben párhuzamosan működik az analízissel (a zárt forrású szoftver csak tapasztalati analízist tesz lehetővé), hogy legalább valamiféle válasz szülessen. Ha azonban nemcsak nem-gyakori a kérdés, hanem az előző folyamatok elhúzódnak (például a kérdésben kétféle nyelvű szöveg is szerepel, sőt van szándékosan kihagyott betűvel írott szó, amelynek a valószínűsíthető jelentését is ki kell találni), akkor az is előfordulhat, hogy az eredmény még nincs kész és a rendszer ijedtében egy előző kérdésre adott választ hozza az elképedt felhasználó tudomására, lásd. Hasonló eset akkor is előfordulhat, ha átmeneti hálózati zavar lép fel.

4.2 Video analízis

Video felvétel analízise lényegében egy olyan speciális minta felismései esetet jelent, ahol a video frame-ek képviselik az egymás utáni mintákat és a kép változtatás gyakoriságát (azaz a felismerésre rendelkezésre álló időt) a video lejátszási sebessége és a videón szereplő objektumok változási sebessége

együttesen adják meg.

Ez a mód a valós idejű működés korlátozásainak egy speciális esete. Ebben az esetben [146] a feldolgozás ideje legfeljebb annyi lehet, mint amennyi két egymás utáni képkocka között eltelt idő, különben csak csonka feldolgozást tudunk végezni. A videóra vett gyors mozgású objektumok betanulására a rendszernek sokkal kevesebb ideje marad, mint az “álló” objektumokra. Mivel a számítógépen akár több feladat is futtatható egyszerre, és a képkockák analizálására fordítható idő az egyidejűleg futtatott további (az analízissel össze nem függő) feladatoktól is függ, ugyanaz az analízáló rendszer más és más eredményt mutathat fel csupán a számítógépes rendszer eltérő terhelése következtében. Ezt a tapasztalatot foglalja össze, hogy a video alapján történő tanulás sebessége frekvencia függő, azaz a videón szereplő gyors mozgású objektumok azonosítási (betanulási) sebessége lényegesen kisebb, mint a lassú mozgásúaké.

4.3 Adatkészletek használata

A gyors fejlődést nem tudja követni maga a fejlődés. A MI rendszerek egyik nagyon lényeges (nagyon költség- és időigényes) része a betanítás. Abban a gyorsan fejlődő világban, ahol a fejlesztők hat hónap szüneteltetést kérnek a már látott problémák rendezésére, a ChatGPT betáplált adatkészlete két éves [24], azaz három-négy generációval ezelőtti, elavult. A mai MI rendszerek szoftverei a tegnapi előtti adatokhoz igazított adatkészletekre támaszkodnak és ez az elmaradás folyamatosan növekszik. A holnapi hardver meg szerkezetében és működési sajátságaiban, és ezért szolgáltatott eredményeiben is, eltérhet a maitól azaz mindent változtatlanul hagyva is eltérhet a MI által szolgáltatott eredmény. Meg lehet próbálni az előre betanított és a friss adatok kombinálását is. A Microsoft Bing alapvetően GPT-3-at használ, de kombinálja egy internetes keresés első 20 eredményével. Sajátos lehet a kettő közé eső két éves “vakfolt” időszak adatait kitalálni. Megbízhatósággal, annotálással és hasonló idejétmúlt kérdésekkel nem érdemes foglalkozni. A látszólagos szüneteltetés ellenére a színpalak mögött élénk tevékenység folyik. Az időnként nyilvánosságra kerülő több USD milliárdos beruházásokat nyilván az ugyanilyen skálán mozgó várható haszon motiválja. Az idősebbek még emlékszenek rá, milyen hatásokat okozott az Internettel kapcsolatos eltúlzott várakozás, a “dotcom” buborék. Az AI buborék még növekszik, de már csak tehetetlenségből.

Mint említettük, a működés sajátságai miatt a betanításhoz használt adatok sorrendje nagyon jelentős. Egy esetleges frissítés azt jelenti, hogy az új információ a régen megtanultak hirtelen (de esetleg csak részleges)

elvesztését okozza. (A rendszerek óriásiak, nagy költségűek és folyamatos működésűek. Egy, a kiszámíthatatlan működésből eredő hiba a piacról való kiszorulást eredményezheti.) Mivel az előállított információ probabilsztikus, a működés ellenőrzéséhez sem használhatunk hagyományos módszereket: az ellenőrzés is probabilsztikus módon történhet. Ha egy konkrét eredmény nyilvánvalóan helytelen, akkor sem kérdőjelezhetjük meg a berendezés működését: *a működés része a helytelen eredmények szolgáltatása*: a helytelen adatokból is származhat, lásd D.2 függelék.

4.4 Felelősség

Technikai számítások (utasítás-vezérelt mód) esetén felvethető a felelősség kérdése. Ha valamilyen helyes gépi utasítást az általános célú számítógép processzor rosszul hajtott végre, azért a processzor gyártója tehető felelőssé (lásd pl. az Intelnek nagyon fájó “Pentium FDIV bug” hibát). Ha a processzor jól működik, de az algoritmus nem, akkor a programot készítő személy vagy cég felelős (kivéve, ha előre kiköti a program eladásakor, hogy a terméke nem biztosan alkalmas arra, amire a felhasználó megvásárolta). Legfeljebb azon lehet vita, hogy az alkalmazás kódja vagy a rendszer program (az operációs rendszer valamely része) a hiba lényegi oka.

Biológiai (adat-vezérelt) számítási mód esetén a biológiai rendszer egésze felelős a működésért. Alacsonyabb szinten az egyed túlélésével, magasabb szinten, társadalmilag, az egyén (vagy az egyed tulajdonosa) megbüntetésével lehet értelmezni a felelősséget.

A MI technikai számítás (mind a valódi, mind az ál-MI esetén) lényegében “adat-vezérelt” módon működik. A hálózat létrehozásakor meghatározzuk hogy az egyes neuronok milyen kapcsolatban állnak egymással és mely súlyokkal (a súlyok értékének meghatározása nélkül!) veszik figyelembe egymás hatását. A hálózat működése során beérkező adatok végzik el a feladat második felét: a beérkező adatok hatására szolgáltatják a matematikailag elvárt eredményt, továbbá a hálózat többi neuronjainak bevonásával beállítják a neurális bemenetekhez rendelt súlyok értékét. Ebben a megfogalmazásban nehezen értelmezhető a felelősség kérdése. A hardver gyártója teremtette meg a lehetőséget, hogy a rendszer működjön, de az adatok állították be, hogy úgy működjön, ahogy. Kérdés, hogy az esetleges felelősséget a hardver gyártójának, a készülék tulajdonosának/használójának, a készülék üzembe helyezőjének, a működést nem megfelelő körülményekkel tervező mérnöknek vagy egyszerűen az adatoknak kell tulajdonítani. Lehet olyan feltételes kérdéseket is feltenni, hogy a rendszer jól működjön-e, ha helyesen beállított idő-ablakokkal futna vagy ha nem abban az időpontban kérdezték volna le az

eredményeket vagy ha nem jelent volna meg a kérdés előtt egy szándékos bemenő-adat manipuláció vagy ha másként választják ki a betanításhoz használt adatkészleteket.

A felelősség elhárítása már most megkezdődött. <https://beta.character.ai/> szolgáltatás felhívja a figyelmet (2023.06.18) arra, hogy a szolgáltatás szórakozásra szolgál: ne bízunk benne és ne vegyük túlságosan komolyan: “Everything Characters say is made up! Don’t trust everything they say or take them too seriously.” Sajnos, ez csak az emberi felhasználókat figyelmezteti, a “gépi intelligencia” ezeket komolynak tekinti. Bizonyos honlapok használatához a leendő felhasználóknak bizonyítani kell, hogy ők emberek. Ezek a csevegő robotok viszont jól elbeszélgetnek egymással. Ha netán bevezetnek valamiféle humán-azonosítást, azt a következő szoftver frissítéskor minden partner átveszi és a játék folyik tovább.

A felelősség szabályozásához is az eddigiektől eltérő módon kell hozzáállni. Mivel a MI rendszerek gyakorlatilag autonom és a környezettől (adatoktól, véletlentől, technikától) függő módon működnek, *a gyártó vagy üzemeltető nem tud szabályozást beépíteni és az adatmennyiség gyors növekedése eleve kilátástalanná teszi az ellenőrzést*, (a felügyelt betanulást el kell felejtenünk) ráadásul a tartalom ellenőrzése ugyanolyan elvek és algoritmusok szerint folyik, mint annak gyártása, így könnyű kijátszani.

5 A MI ártalmai

Bár 2023 tavasza óta felerősödtek a kétségek a MI értékeit és ártalmait illetően, [41] nevezetes nyolc pontjának állításai megdöbbentőek és elgondolkodtatóak (a MI kiszámíthatatlan fejlődését, befolyásolhatatlan működését, belső működésének hiányzó megértését, a világ sajátos belső reprezentálását, stb. illetően). Ráadásul ez csupán empirikus értékelés, némi elméleti alátámasztással. Az ártalmak azonban lényegesen túlnőttek ezen a körön.

5.1 A szabad használat veszélyei

Érdekes és elgondolkodtató párhuzamot von a nukleáris energia és a mesterséges intelligencia használata között [30], amit a jelen tanulmány szerzője (a [60] tanulmánnyal egyetértve) kiegészítene az Internet használatával. Bár a nukleáris energia felhasználásával kapcsolatban született, de a MI-ra is vonatkoztatható, hogy “A technológia – ahogy a tudomány is – semleges, és csak az eszközöket nyújtja, bármilyen célú legyen is a felhasználásuk” [138]. A marketing (a magáncégek érdeke) azonban nagyon erőszakosan kínálja

ezeket az eszközöket, és csak a profitot tekinti. *A nukleáris energiával kapcsolatban teljesen ismert volt az elméleti háttér és mind a hadi- mind a békés felhasználás teljes állami kontroll alatt ment végbe, megvalósíthatósági és kockázati hatástanulmányok alapján.*

A MI fejlesztése viszont lényegében spontán módon, a kockázati tőke irányítása alatt fejlődik, a kockázatokkal nem törődő fejlesztők által, immár EU és állami támogatással. Nélkülözi a szilárd elméleti alapokat, fogalmilag nem tisztázott sem a működési elve sem célja. A működés elméletének megértése és megvalósíthatósági/kockázati tanulmányok híján a MI fejlesztő eszközök gyakorlatilag minden kontroll nélkül használhatók. Következményeivel pedig már a legfelsőbb állami vezetés kénytelen foglalkozni. Érdekes és meglepő, de inkább csak a felelősség utólagos áthárítása az OpenAI vezérigazgatójának a kongresszusi meghallgatásán kifejtett álláspontja, amely szerint “nem csak szükséges, hanem kívánatos is az AI proaktív, kormány által irányított szabályozása”. Igen, az lett volna, pár évvel ezelőtt. Az állami/hatósági támogatással kiszabadult szellemet gyakorlatilag lehetetlen visszagyömöszölni a palackba, még egy újonnan létrehozandó állami ügynökségnek is.

Egy probabilisztikus eszköz (pl. ChatGPT) használata elméletileg és technikailag kétséges, következményei beláthatatlanok, jogi szabályozása szinte lehetetlen. Alkalmazása nemkívánatos következményekkel járhat és “*kétélő kard*” [130] jellegű. A generatív MI-re is vonatkozik, hogy mint a többi technológiákra, hogy “bár mindegyik hasznos, rombolásra is felhasználhatók” [138]. A video kezelés javulásának kezdetén 2017-ben a biztonsági szakember még védekező fegyverként javasolta a MI-t gyermek pornográfia felderítésére [9], 2023-ban pedig már a támadó fegyverré vált MI (gyermek pornográfia gyártása) adott okot parlamenti vitára [6]. A pre-trained (P) módszer miatt a "pornográfia" bármivel helyettesíthető; csupán betanítás kérdése. A lényeg a probabilisztikus generáláson (G) és a transzformatív képességen (T) van. Sajnos, (alapos betanítás után) kisebb hardver rendszerek (kisebb befektetéssel, kevésbé feltűnő módon megvalósíthatóan) is lehetővé teszik ilyen irányú igények kielégítését, és állami felső vezetési gyors beavatkozást igényelnek pl. a szereplők “személyiségi jogainak” és a cselekmény jogi minősítésének kapcsán.

Elsőrendű érdek lenne megérteni hogy – a profit érdekeket és a politikai akaratot is lehámozva – valójában mi is tartozik a MI fogalomkörébe, mely eljárások/technológiák potenciálisan veszélyesek (lásd 2.5 szakasz), “mennnyibe kerül” alkalmazásuk. A maga helyén és csak a szükséges mértékben használva, a MI (alapos ráncfelvarrás után) hasznos is lehet. Még akár a mostanában a célkeresztbe került LLM technológiák is. A [47] rendkívül alapos és friss

tanulmány végkövetkeztetése, hogy az LLM alapú fejlesztések “számos korláttal rendelkeznek, főként az indoklás és a robosztus működés terén”. *Mindemellett, az LLM a TI felügyeletével (beleértve az adatkészletek korlátozását is) használható segédeszköz lehet, bár tömeges használatához más technológiai megoldásokat kell találni.*

5.2 Használható-e a ChatGPT EU-rendeletek készítésére?

A ChatGPT kedvenc probabilisztikus munkamódszere, hogy a nagy előfordulási gyakoriságú tényeket relevanciájuktól függetlenül szinte biztosan szerepelteti, a kis gyakoriságúakat szinte biztosan elhagyja, a hasonló dokumentumokban hasonló összefüggésekben gyakran szereplő, de a konkrét esetben nem ismert tényeket kitalálja (az intuíció sajátos értelmezéseként) és tényként tálalja (lásd a D.2 és D.1 függelékeket). Olyan dokumentumot készít, aminek tartalma arra hasonlít, amit feltehetően annak olvasói olvasni szeretnének. A ChatGPT munkamódszere érhető tetten az EU dokumentumokban is. Az EU MI-ra vonatkozó dokumentumainak tanulmányozása közben egyre inkább úgy tűnik hogy a MI által készített dokumentumok egyre könnyebben tudják a TI által készített dokumentumok minőségét megközelíteni, és nem csak a MI ilyen képessége javul.

Az EU-ban a nagy számú és nagy terjedelmű dokumentumok számos (kis szakértelmű de nagy politikai meghatározottságú) nagy testületen törnek át és eleve változásokat közölnek (a számítógépes szövegszerkesztés és a “Big Data” folyamánya), azaz ha valami ostobaság bekerül egy dokumentumba, akkor kisebbfajta földrengés kell, hogy korrigálják. Ilyen horderejű jelenleg a MI kezelésének teljes irányváltása. A változást az egyre nyilvánvalóbbá váló társadalmi és gazdasági veszélyek, továbbá a MI létrehozóinak és nagy játékosainak pálfordulása okozza, a most már többségivé váló szakmai vélemény. A MI “hype” hatására megbíznak a többségi véleményben és bevezetik. Nagyon megnő a MI pozitív hatását kiemelő vélemények száma és aránya. Egyre többet beszélnek a MI-ról és egyre nagyobb gazdasági hasznot remélnek tőle. Akkor mi is beállunk a sorba, sőt annak az élére. Nem tudjuk ugyan, hogy mi az, de hátha azok, akik propagálják, értik. A szomszédnak is olyan van, nekünk is kell. Aztán fordul a szél, és megnő az ellenvélemények aránya. Az ezután készített dokumentumokban betiltjuk, nagyrészt. Saját véleményünk nincs és még *a betiltás után is csak felveti a szakértő, hogy amit kötelező jelleggel bevezettek és pár év múlva szintén kötelező jelleggel nagy részben betiltottak, annak valamiféle definícióját kellene adni* (eltekintve a fentebb mutatott módon, évente változó és alapvetően eltérő definícióktól). Ami azért is kellene, mert ugyanezen megalapozott szakmai vélemény szerint

“egy kiforrott és használatra kész technológia” amiről szó van (A ChatGPT erről alkotott véleményét lásd a D.2 függelékben). Figyelemre méltó, hogy mindez csupán az állásfoglaláshoz csatolt függelékben található meg. De még ebben sincs nyoma annak a kételynek, hogy “*Amit csinálni akarunk, azzal hat évtizede eredménytelenül próbálkozunk. Lehet hogy elvileg sem lehetséges?*” (a ChatGPT erre vonatkozó válasza a D.1 függelékben olvasható). Nem merül fel, hogy megvalósíthatósági tanulmányt is kellene készíteni, legalább utólag. Meg, a hosszú távú projektekre vonatkozóan, előre.

A 5.4 szakaszban említjük, hogy már látható hogy a gépi tanulás modellje fenntarthatatlan és kifejezetten energia pazarló; a B.2 szakaszban megmagyarázzuk, hogy miért. Ennek fényében legalábbis furcsa az EU szabályozó rendeletében [28] olvasni, hogy “ez a rendelet hozzájárul a zöld és fenntartható mesterséges intelligencia előmozdításához és a MI-rendszerek környezeti hatásának teljes életciklusuk során történő figyelembevételéhez”. Ami nagyon szépen megfogalmazott és kíváncsú cél, csak éppen a technikai működés lényegének nem értéséről tanúskodik és szöges ellentétben van a valósággal, továbbá egy szoftver esetében mást takar az életciklus fogalma és környezeti hatásának értelmezése (a szöveg copy/paste szerkesztési hibának tűnik).

Ha ezeket a jelentéseket a ChatGPT készítette volna, a lelkesedő vélemények gyakorisága alapján 2017 után a MI-t pozitív és támogatandó dolognak tartotta volna, 2022 után meg veszélyesnek és akár betiltandónak. Mivel a korábbi dokumentumokban a “zöld és fenntartható” szókapcsolat pozitív kicsengéssel gyakran szerepel, a ChatGPT ezeket a jelzőket a MI-ra is alkalmazná (bár közvetlen kérdésre tagadja, hogy véleménye lenne, lásd D.3 függelék). A számítás energetikai határfoka fogalmilag sem fordulhat elő a (számítástudomány alapján készített) korábbi dokumentumokban, ezért a ChatGPT erről vagy hallgatna *vagy a korábbi dokumentumok alapján tényként tálná hogy a MI fenntartható (ami nagyon is kíváncsú lenne), bár épp az ellenkezője igaz*. Közvetlen kérdésre meg inkább átminősíti magát számítógépi intelligenciából elvi megoldássá, lásd D.3 függelék. Hasonlóképpen nagyon szép és jó dolog az uniós értékek tiszteletben tartását elvárni, és az a mintaként szereplő társadalmi-jogi dokumentumokban mindenütt szerepel is, de a ChatGPT vagy nem generálna erre vonatkozó szöveget a szabályozásba (lásd D.3 függelék), vagy beírná a korábbi dokumentumokban található véleményt. Mindezt a pillanatnyi körülményektől függően, esetlegesen, lásd D.2 függelék; a felelősség vállalása nélkül, lásd D.2 függelék.

Az EU által készített szabályozásban viszont ennek jelenléte a működési elv lényegének nem értését és a működés eddigi tapasztalatainak figyelmen kívül hagyását mutatja. Viszont, *mivel technikai-tudományos dokumentumok*

(technikai szempontok) kis gyakorisággal fordulnak elő a testületek anyagai között, ezért nem fognak szerepelni a ChatGPT által generált véleményben sem (a tudományra vonatkozó álláspontja a D.5 függelékben olvasható: a fejlesztők mondják meg a tudományos álláspontot) és az EU állásfoglalásaiban sem. Ha az EU (tanács, fórum, bizottság) azt akarja hallani, hogy a MI hasznos, zöld, fenntartható, tiszteletben tartja az EU vonatkozó alapértékeit, akkor a jelentés azt fogja tartalmazni, különben nem szavazzák meg. *Egy jelentés szövege igen, de egy technikai rendszer nem fogja a politikai akaratot tiszteletben tartani. Ha nem olyannak bizonyul, betiltjuk.* Ha a tények nem egyeznek a jelentésben megfogalmazott akaráttal, annál rosszabb a tényeknek.

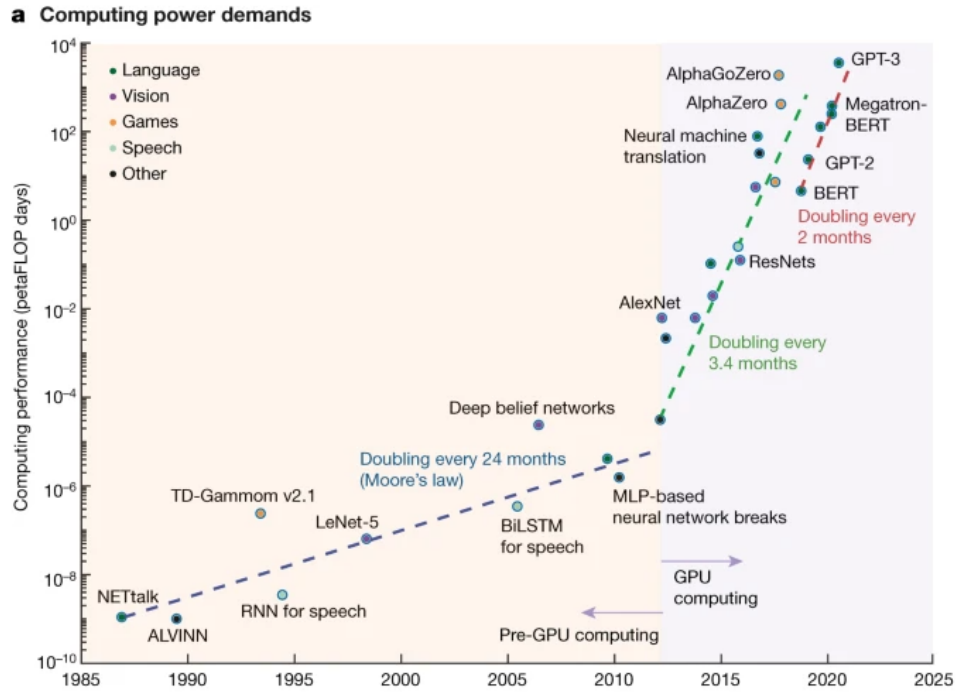
Ilyen szempontból tekinthetjük a megfelelő EU testületeket olyan mesterséges neurális hálózatoknak, amelyek fekete dobozként működnek változható, de ismeretlen módon összefüggő súlyokkal és amelyeknek vannak bemenetei (pl. szakmai bizottsági jelentések, elfogadott politikai, etikai, stb. alapelvek), és amelyeknek belső súlyait még a működés előtt jóval betanították (elsajátította a testületben képviselt politikai irányzatok számarányát, az éppen támogatott projektek irányultságát és távlati céljait, az EU egészére vonatkozó vonásokat, érdekcsoportok hatását, politikai helyzetet, presztizs okokat, stb). Ez a fekete doboz használja a bemenetére táplált kívánalmakat (támogassuk, vezessük be, kutassuk, alkalmazzuk és húzzunk belőle maximális hasznout vagy éppen találjunk indokokat a betiltására vagy legalább korlátozására) és egy ennek megfelelően generált szöveget készít. Ennél a szövegnél nem követelmény a belső konzisztencia, pl. hogy a “egy kiforrott és használatra kész technológia” esetén miért akarjuk a felhasználását erősen korlátozni; az igazságtartalom, pl. hogy a gépi tanulás “zöld” technológia-e; hogy megvalósítható-e a célkitűzés, pl. az EU alapértékeinek betartatása a MI által.

Veszélyes hasonlóság. Még ha megjelöljük, hogy az előbbit MI, az utóbbit TI készítette, akkor is elgondolkodtató.

5.3 Fenntarthatóság és szennyezés

Bár az ún. Moore-megfigyelés már vagy negyed százada nem érvényes és eredetileg is csak az *integrált áramkörök alkatrész sűrűségére* vonatkozott, még ma is szokták (akár Nature szinten is [101]) ezzel magyarázni, hogy a számítógépes kapacitás iránti igény nagyjából két évente duplázódott, nagyjából egy évtizeddel ezelőttig, avagy a Grafikus Processing Unit (GPU) megjelenéséig.

A GPUk eredetileg kép megjelenítés gyorsítási célra szolgáltak, és arra



Ábra 5.1: A számítási igény növekedése és annak kielégítése, különböző számítási technológiák korában [101]

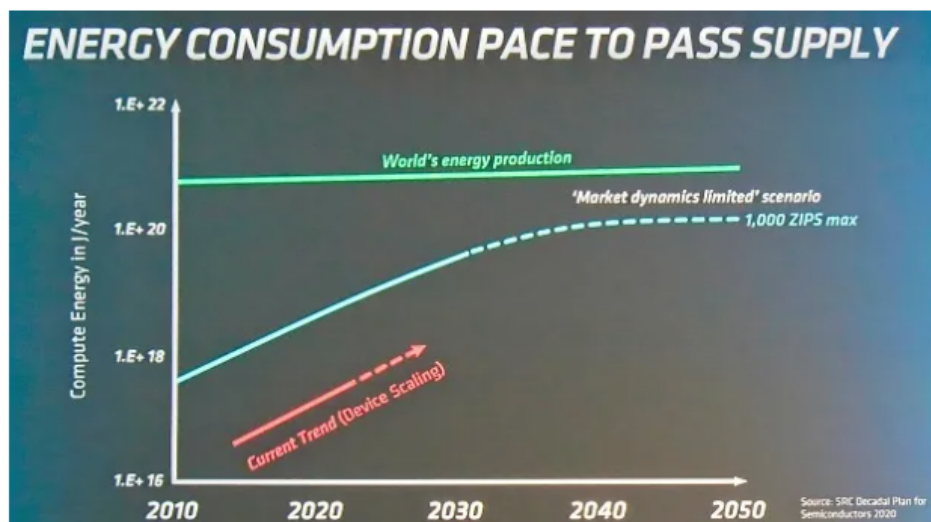
valóban megfelelők voltak. Működési módjuk félreértése miatt egyre jobban elterjedt ún. általános célú használatuk, abban a tévhitben, hogy ott is nagyságrendileg nagyobb számítási teljesítőképességet lehet velük elérni; annak ellenére, hogy Intel-NVIDIA közös tanulmány mutatta ki [90], hogy a várt két nagyságrend helyett mindössze két-háromszor nagyobb a számítási teljesítőképesség, viszont nagyságrenddel nagyobb az igényelt elektromos teljesítmény. Ez a fejlesztés a számítógépek amúgy is alacsony működési hatásfokának [65] további nagyságrendi csökkenését eredményezte.

A piac még ezen az áron is igényelte a számítási teljesítmény növelését, ami ugrást idézett elő a keresletben. Ezt a gyors ugrást a gyártási kapacitás nem tudta követni; a következmény az általános chip hiány lett. Ezt a hiányt szervezési okokkal szokták magyarázni [32], de valójában a kereslet többszáz százalékkal megnőtt, az előállítás meg az éves 10% növekedést is nehezen tudta követni. Ha egy új félvezető gyár felépítése 2-3 évig tart, akkor 2-3 évig chip hiány lesz, bármilyen jó szervezés mellett.

A MI megjelenésével újabb korszak kezdődött a még gazdaságtalanabb számítási teljesítmény növelés irányában. A két hónaponkénti duplázódás az igények évi 5,000% feletti növekedését jelenti, miközben az ipar nehezen birkózik meg az évi 30-40% kapacitás növeléssel. A következmény egy újabb, furcsa chip-hiány. Manapság mindenki MI fejlesztést akar végezni, a szó szoros értelmében “bármilyen áron”. Ami az egekbe hajtotta a processzorok árát. *A hagyományos piacok nem bírják megfizetni a megdrágult processzorokat, ezért a legtöbb hagyományos piacon visszaesett a termelés.* Ez év elején már megjelentek a hiánybetegségek, látszólagosan a MI dicséretét zengve [114]. Az éves növekedés még pozitív, de az ez évi első negyedéves már negatív. Már a data centers csökkenése is szembetűnő, de még inkább a kliensek eladásának majd’ felére csökkenése, egy negyedév alatt. Adott gyártási kapacitás mellett *a MI elszívja a hardvert a többi számítógépes szegmens elől.* Pl. az Intel második negyedévi eredménye két év alatt 19.6 milliárd USD-ról 15.3-ra, majd 12.9-re esett, a gépi tanulás és a MI sikere ellenére [4, 3]. A második negyedévben az Intel mainframe piaca pl. 30%-kal csökkent [27] 2023-ban. A cikk állítása szerint a MI “recesszió-biztos” szegmens, de az inkább a piaci szereplők menekülési útvonala: meglovagolják a technológiai várakozás emelkedő ágát. Amíg tart a boom. Utána pedig következik a kiábrándulás és a túltermelési válság. Fejlesztési szinten már megállt a MI [77], de a későbbi jövő (és az elindított hosszú távú projektek) még hajtják a piacot. Érdekes lesz figyelni a következő negyedévek recesszióját. Különösen a MI kiábrándulás és tiltások után.

A MI, úgy tűnik, saját magát is felfalja. 2023 szeptemberében már arról esik szó [111], hogy “[a Sambanova számára] csak a probléma egyik része, hogy találnia kell 1 milliárd US dollárt, ami végül az Nvidiához kerül. A probléma nagyobbik része, hogy ha nem találnak gyártót a szükséges kb. 20.000 db H100 GPU előállítására, akkor [sokkal többért] bérelnie kell a publikus felhő szolgáltatóktól a kapacitást”. *Eltűnhet vagy drasztikusan megdrágulhat a publikusan elérhető felhő szolgáltatás is.* A már említett extrém alacsony hatásfokot tovább rontja a hálózati átvitel hatásfoka. Úgy tűnik, a számítási modellnek (és azon belül az átviteli idő szerepének) meg nem értésében még mindig lehet lentebbi színvonalat elérni.

Környezet és fenntarthatóság szempontjából legalábbis furcsa (és a működés lényegének nem értéséről tanúskodó) az EU szabályozó rendeletében [28] olvasni, hogy “ez a rendelet hozzájárul a zöld és fenntartható mesterséges intelligencia előmozdításához és a MI-rendszerek környezeti hatásának teljes életciklusuk során történő figyelembevételéhez” és hogy “Környezeti szempontból a mesterséges intelligencia komoly lehetőségeket kínál az olyan környezetvédelmi kérdések megoldására, mint az erőforrás-



Ábra 5.2: A gépi tanulás energia fogyasztásának várható alakulása (AMD) [128].

fogyasztás csökkentése, a dekarbonizáció előmozdítása, a körforgásos gazdaság fellendítése, *a villamosenergia-hálózatokon belül a kereslet és a kínálat közötti egyensúly megteremtése* vagy a logisztikai útvonalak optimalizálása“. (Azzal együtt, hogy *tiltja MI rendszerek beépítését kritikus infrastruktúrákba.*) Biztos ami biztos, egy másik pontban az ellenkezője is benne van: “A mesterségesintelligencia-technológiák és adatközpontok nagy szénlábnnyommal rendelkeznek a megnövekedett számítási energiafogyasztás, valamint a tárolt adatok mennyisége és a keletkezett hő-, elektromos és elektronikus hulladék mennyisége által előidézett magas energiaköltségek miatt, *így még nagyobb lesz a szennyezés.*” Valamelyik vélemény csak bejön.

Jellemző mondat a gyártók környezet-tudatosságára a [93] közleményből: *az Nvidia nem hozza nyilvánosságra az általa gyártott GPUk szénlábnnyomát*, így arra a tanulmány csak egy 150 kg CO_2 -ekvivalens alsó értéket ad meg. Mármint darabonként. Az említett Sambanova projektben csupán a GPUk legyártása több mint 3.000 tonna széndioxid kibocsátásával egyenértékű, ennél az egyetlen innovatív és környezettudatos cégnél. Az Nvidia pedig valószínűleg nem gyártana, ha nem lennének környezettudatos, zöld, alacsony szénlábnnyomú technológiát alkalmazni kívánó vásárlói. Ha az Nvidia nem ad ki közleményt, akkor azt hiszik, hogy a lábnnyom értéke nulla. Egy MIT tanulmány [51] szerint egyetlen leközült MI-modell megvalósítása a cikket

felhasználók által átlagosan annyi széndioxid kibocsátással jár, mint amennyit 25.000 gépkocsi teljes élettartama alatt kibocsát. A megtermelt óriási hőmennyiség elvezetésére folyóvízní vizet használnak és a disszipált hő emeli a globális hőmérsékletet.

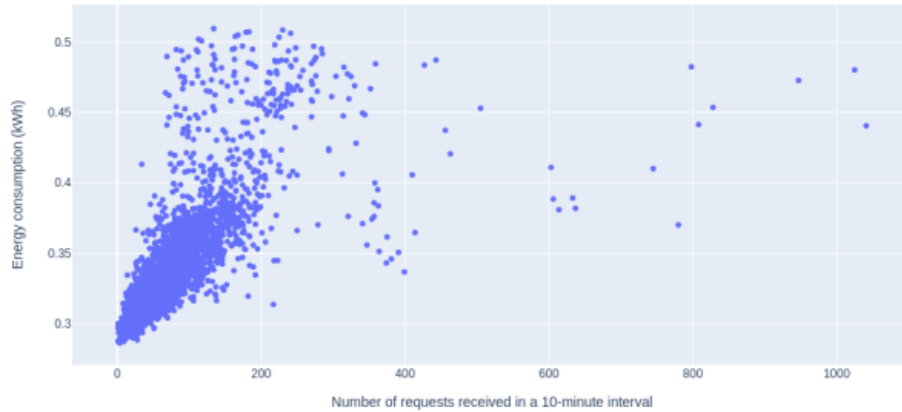
A tudomány környezet tudatossága meg aggódik a klímaváltozásért, és csökkenteni igyekszik a CO_2 kibocsátást. Aminek érdekében létrehoznak egy 176 milliárd paraméteres MI hardvert, aminek 118 napos betanítása önmagában csaknem 100 tonna CO_2 kibocsátásával járt [93], és akkor még nem gyártottuk le a hardvert és nem működtettük. A különböző iparágak pedig (többek között az EU és tagállamai támogatásával, sőt: általuk kötelezve) egymással versengve alkalmazzák a “zöld” MI-t, LLM-t és a GPU-kat a szénkibocsátás csökkentésére, a nulla szénkibocsátás felé vezető úton haladva. Úgy vagyok vele, mint Ignotus a fekete zongorával: akasszanak fel, ha értem.

5.4 Energia fogyasztás

Pár éve még a “data centers” volt a nagy energia felhasználó [106]; többek között a Google keresések évente 200 terawattóra energiát fogyasztanak [101]. A többi, csupán matematikainak gondolt technológiával együtt a MI is első számú közellenséggé növi ki magát energia válságban szenvedő világunkban, nem is túl lassan. A veszély valódi. A ‘block-chain’ technológia használatára 2022-ben a világ több energiát fordított mint amennyi Magyarország teljes energia fogyasztása [51]. Az energiában gazdag Texas állam korlátozni kényszerült a bit bányászatot [31].

Amint a B.6 szakaszban kifejtjük, “a gépi tanulás abba az irányba tart, hogy felélje az összes előállítható energiát; olyan modellt használ, ami költséges, alacsony hatékonyságú, és fenntarthatatlan” [128]. Amint a 5.2. ábra mutatja, a gépi tanulás energia fogyasztása (és az összes energia fogyasztáson belüli részaránya) meredeken emelkedik (10 év alatt a tízszeresére nőtt), és a jelen tendencia mellett kb 20 év múlva eléri a Földön elvileg megtermelhető energia mennyiségét. Sőt, másik becslés [109] szerint, még ha a folyamatos hardver javulást is figyelembe vesszük, a MI modellek betanítására fordítandó energia 2030-ra túl fogja lépni a globálisan megtermelhető energia mennyiségét. Például a Google PaLM rendszerének betanítása két hónapon át összesen 3.4 GWh energiát vett igénybe [108].

Az 5.3 ábra szerint a felhasznált energiának csupán egyik fele szükséges a tényleges használatra, a másik fele a számítógép működési elvéből következő “tétlenség” és az infrastruktúra fenntartására szükséges [93]. A célzatos mérések szerint [93] az AI-típusú kérések/kérdések kiszolgálása során *a felhasznált energia mindössze 2%-át használta el a CPU, 22.7%-át a RAM*



Ábra 5.3: Egy Google Cloud Platform energia fogyasztása a kiszolgált kérések számának függvényében [93]

memória, és 75.3%-át a GPU. (A GPU előtti korszakban kb 20% energia fordítódott a számításra, és 80% az adatok kezelésére [131]; az arány egy évtizeden belül megfordult!) *Az igazi nagy fogyasztó tehát az adatokat tároló RAM (akár több száz milliárd paraméter!) és az azokat rendkívül gazdaság-talanul feldolgozó GPU.* Ezek mellett elhanyagolható a (mellesleg igen rossz energia hatásfokkal rendelkező) CPU energia fogyasztása.

Első olvasásra erős túlzásnak tűnik a drámai megállapítás, hogy *“Lényegesen, akár 10 nagyságrenddel is, javítani kell a számítás hatékonyságát, hogy a nagy számítási problémákat meg tudjuk oldani”* [109]. A MI esetében azonban az energia pazarlás és a számítási hatásfok csökkenés mértékét egyszerű számítással alátámaszthatjuk. Ha a biológiai neurális hálózat jellemző műveleti idejét 1 *msec*-nek tekintjük és egy tevékenységet 10 *sec* alatt tanul meg, a tanulási idő/műveletvégzési idő aránya 10^4 . Ha a mesterséges neurális hálózat jellemző műveleti ideje 1 *ns* és ezt a tevékenységet nem egészen két hét (10^6 *sec*) alatt tanulja meg, ugyanez az arány 10^{15} . Böven megvan a tíz nagyságrend. Ha az újabb gigantomániás MI rendszerek betanulási idejét tekintjük (pl. [93] esetén 118 nap), még azon felül is találunk legalább egy nagyságrendet. Ha három nagyságrendet betudunk a processzor megvalósítás [65], és még hármat a szoftver rendszer megvalósítása [117] hatásfokának, még mindig marad *négy-öt nagyságrend* a gépi tanulás hatásfokcsökkentő hatására. Lényegében erre utal hogy *“a gépi tanulás ... olyan modellt használ, ami költséges, alacsony hatékonyságú, és fenntarthatatlan”* [128]. Pontosan ezt mutatja a MI növekvő mértékű

felhasználásával az egekbe szökő számítási igény és energia fogyasztás.

5.5 Információs környezet szennyezés

Az Internet eredetileg hadi fejlesztés volt, megfelelő szakértői háttérrel és kívülről biztosított biztonsági intézkedésekkel, beleértve a források megbízhatóságát is. Tömegesedése és felhasználói tartalom generálás miatt vált veszélyforrássá. Nem tisztázták a fizikai hozzáférés jogosultságait, utat nyitva pl. a sötét internetnek és a gyermekek számára a nem nekik való tartalmak eléréséhez, és nem tisztázták a tartalom szolgáltatás felelősségét, pedig a nyomtatott és az elektronikus sajtó joggyakorlata alapul szolgálhatott volna.

A MI gyakorlatilag nem jelentene veszélyt az Internet elérése és annak megfelelő szabályozása esetén. A hálózat-képes generatív MI valóban óriási veszélyt jelent, főként mivel a nyílt forrású MI könyvtárak használatával gyakorlatilag bárki generálhat bármiféle mesterséges tartalmat, pusztán ostobaságból vagy jópofaságból, de akár manipulatív céllal is (ráadásul, az EU szabályozás szerint még csak felelősségre sem vonhatók). Kicsit mesterkéltné párhuzam, hogy nem szabad a nukleáris anyagokkal sem bárkinek foglalkozni (ahhoz azért nagy költségű infrastruktúra kell), ráadásul a digitális anyagok korlátlanul és ellenőrizhetetlenül másolhatók, terjeszthetők (ráadásul a profit-éhes cégek ingyenes infrastruktúrát és szervezett segítséget nyújtanak). Közelebbi párhuzam, hogy a megfelelő programkönyvtárak birtokában bárki komolyabb szakértelem nélkül létrehozhat maliciózus szoftvereket, amik hatásának kivédésére aztán globális és egyéni erőfeszítésekre kényszerülünk. Viszont az utóbbi könyvtárak egyértelműen csak a "sötét oldalon" elérhetők.

A szabályozatlanság miatt az Interneten minimális kockázattal lehet garázdálkodni. Bár manapság már olyan alapvető infrastruktúra, mint pl. az elektromos hálózat, szabályozása több mint hiányos. Amíg csak vásárolunk elektromos energiát, addig csak a kifizetésével kapcsolatban okozhatunk gondot. Amikor termelünk és az országos hálózatba is feltöltünk, arra már másképpen kell szerződni. Természetesen, amikor kézi erővel pl. gyűlöletkeltő tartalmakat generálunk, az még magánügy. Ha mások által is hozzáférhetővé tesszük az Internet útján, az már nem. A MI felhasználásával már egyikkel sem kell vesződnünk. A zombi-gépek minden bizonnyal nem csak hálózati támadások generálására "hasznosak": tudnak generált MI tartalmat készíteni, készíttetni és feltölteni, ráadásul humán-generált tartalomként címkézve. A felhasználók döntő többségétől nem lehet megfelelő szakértelmet elvárni (nem tudja az információ forrását megfelelően értékelni sem), és a hálózat-képes eszközök számának exponenciális növekedésével a probléma súlyosabbá válik.

A MI csak a hálózaton keresztül tud bemenő adatokat szerezni és ered-

ményeit is csak a hálózat felhasználásával tudja terjeszteni. A generatív MI és az Internet (szintén MI alapú) kereső mechanizmusa együttesen veszélyes kombinációt eredményez. A számítógépek sokkal gyorsabban képesek generálni szöveget, mint a személyek, és megfelelő kulcsszavakkal ellátva, a generált szövegek egyre jobb helyezést érnek el a találati listákon. Ezáltal kiszorítják a valódi és helyes információt; és akkor még nem tekintettük valós veszélynek az emberi manipulatív szándékot.

5.6 Kritikus infrastruktúrák

Egy probabilisztikus alapon működő eszközt kritikus infrastruktúrákba beépíteni gyakorlatilag egy ismeretlen időzítésű bomba készítését jelenti: csak idő kérdése, mikor fog robbanni. Különösen, ha ritkán előforduló jelenséget kell felismerni és ennek megfelelően szokatlan reakciót kezdeményezni, lásd pl. az időjárás váratlan jelenségeire vonatkozó értékelést [8]. Még egy kiber támadás felismerése esetén is elindíthat visszavonhatatlan akciót, de pl. egy nagyhatalom akár világháborút idézhet elő egy téves “visszacsapással”. Jól hangzik pl. a villamos hálózat felügyelése is MI eszközökkel [100]: amíg valami váratlan esemény nem történik. Kérdés marad az is: megtakarít-e annyi energiát, amennyit működése során felhasznál. A számítógépesítés elkerülhetetlen; kritikus berendezésekbe is beleépítik. Bár hardver és szoftver szinten is léteznek hibák és “hátsó ajtók”, a berendezések működése kiszámítható és tesztelhető. A hálózattól elszigetelve, a fizikai elérés megfelelő korlátozásával csak elviselhető kockázatot jelentenek. Ezekbe probabilisztikus MI komponenst betenni az öngyilkosság költséges és idegesítő módja. Nagyon helyesen, az EU is tiltja.

6 Összefoglalás

A LLM bizonyos esetekben meglepően jól teljesít (főként amikor a fejlesztő rendszert bizonyos benchmark-ok teljesítésére optimalizálják [42] és csak rövid ideig vagyunk kapcsolatban vele [41]), de megbukik amint a természet tűzi ki benchmarkot [50, 139]. Kiváló, fáradhatatlan és szórakoztató csevegőpartner, de munkatársnak megbízhatatlan. A MI (leválasztva az automatizálást és egyéb nem-probabilisztikus működésű részeket), sok évtizedes története és látszólagos matematikai háttére (ami valójában hagyományos számítógépekre fejlesztett gépi tanulás alkalmazása) ellenére, számos elméleti hiányossággal bír; főként nincs benne a neve által elvárt intelligencia. Utat mutatunk a MI elméleti megalapozása és konszolidált, fenntartható

működtetése felé – a tanulmány elkészítésének ideje alatt bekövetkezett általános elkedvetlenedés és kiábrándulás ellenére.

Bár elsősorban a tudományos kutatásra (annak is az eredményeket szöveges formában megjelenítő publikáció készítés részére) vonatkozik, jól általánosítható a [20] értékelése; értelemszerűen igaz hang- és kép feldolgozására vonatkoztatva továbbá speciális LLM területek (pl. Large Medical Model) esetére is. Összefoglaló értékelése abból indul ki, hogy *“a [z LLM] technológia használata kikerülhetetlen, ezért annak tiltása nem megoldás. ... A ChatGPT más LLM rendszerek meggyőző, de sokszor hibás szöveget állítanak elő; használatuk torzíthatja a tényeket és tévinformációkat terjeszthet.”* Javasolja, hogy kiindulásként öt fő ponthoz ragaszkodjunk: 1./ TI általi ellenőrzés 2./ LLM használat járulékának feltüntetése 3./ Valóban nyílt forrású LLM-ek használata (jelenleg az LLM technológiák nagy cégek tulajdonában vannak) 4./ Használjuk a MI előnyeit 5./ Vitassuk meg használatát szélesebb körben. A szerző hozzáteszi, nulladik pontként: értsük meg, hogyan működik.

A MI szabályozása és értékelése előtt, első lépésként az Internet használatát kellene rendbetenni, a területen érdekelt nagy cégek tiltakozása ellenére. (Ez összhangban lenne az EU törekvéseivel: pár éve egy “Proposer’s Day” alkalmával az előadó egy szóval válaszolt arra, hogy milyen legyen a Jövő Internetje: “Different.”) Pontosabban, nulladik lépésként tisztázni kellene a számítás és a kommunikáció elméleti alapjait: a kettő nem választható el egymástól és a jelenlegi elmélet teljesen alkalmatlan a mai technológiai körülmények között. A hivatkozási alapul szolgáló számítási elmélet időzítemi viszonyai már évtizedek óta teljesen mások (ennek egyik következménye az energia pazarlás) és nem tartalmazzák a kommunikációt (sem processzoron belül, sem a komponensek között), a számítógépek az eredeti elképzeléssel szemben nagy mennyiségű adatot dolgoznak fel (nem az eredeti “kevés adat sok számítás” elv; sőt a tevékenység nagy része nem számítás, de a számítógép számítási műveletekkel emulál sok egyéb tevékenységet), az eredeti “single processor approach” elképzelés helyett hálózati eszközökkel kapcsolódnak egymáshoz, feladataikat igen gazdaságtalanul tudják megoldani, stb. Második lépésként pedig meg kellene érteni nemcsak a MI matematikai módszereinek tulajdonságait, hanem azoknak a technikai megvalósításhoz kötődő következményeit is.

A MI csak a jéghegy csúcsa, aminek most nagy koppanással neki ütköztünk és ami önmagában is elég veszélyes lehet. Tanulni kellene az előző két “AI winter” tapasztalataiból. Az igazán problémás részek azonban a vízszint alatt vannak. Újra kell gondolni a számítógép elméletét és arra felépíteni az új technológiát.

7 Függelék

A A MI elmélete

Mint fentebb részleteztük, a MI elméleti háttére erősen ellentmondásos (maga fogalom sem egyértelműen meghatározott, így nincs egységes elmélete sem; a determinisztikus osztály kivételével kérdéses a bizonyíthatóság, megoldhatóság, korrektség, stb. [49]). Ez a fő oka, hogy már pár évvel ezelőtt elakadt [77] a MI fontos területeinek fejlődése, mostanra pedig bebizonyosodott hogy zsákutcába rohant be gyorsvonati sebességgel. *Egy hosszú évekre szóló, nemzeti szinten is jelentős anyagi és humán erőforrásokat felhasználó projektnek tisztázni kell az elakadás okait és a fejlesztési/felhasználási célokat csak kellően megalapozott megvalósíthatósági (feasibility) tanulmányok alapján kitűzni.* Ahhoz pedig a számítógép tudomány (tehát a mérnöki megvalósítás) alapján kell áttekinteni és értékelni a MI megvalósítás lehetőségeit és nem a számítástudomány (tehát az absztrakt processzor működését vizsgáló, lényegében matematika) alapján, mivel technikai megvalósításon alapuló berendezéseket és nem elvileg működőképes rendszereket kell létrehozni.

A fentiek alapján tehát a MI-vel kapcsolatban vizsgálnunk kell a fogalom félreértelmezéséből (az eredetileg célul kitűzött mesterséges intelligencia helyett gépi tanulás valamint probabilisztikus szöveg, kép, hang generálás), a működési módok figyelmen kívül hagyásából (az adat-vezérelt működési mód helyett utasítás-vezérelt működési mód használata), a kétféle fogalomrendszer egymásra leképezését hibásan megvalósító matematikai módszerekből, a technikai megvalósításból származó teljesítőképesség korlátokból és a hibás elvű gyorsítási megvalósításokból, továbbá az egyre növekvő mennyiségű adat és az elvárt valós idejű működés ellentmondásos követelménye közötti kompromisszum keresésből fakadó problémákat is.

A.1 Történeti vonatkozások

Annak idején Neumann Jánost a biológiai minta (az agy) ihlette nevezetes számítógép modelljének megalkotására [144]. Korának idegtudománya el látta az alapvető ismeretekkel és ennek alapján Neumann felállította a szekvenciális működésű, tárolt programú számítógép működési modelljét. Mivel közvetlen célja az elektroncsövekkel megépített elektronikus számítógép működése matematikai alapjainak megteremtése volt, olyan (tudatos) elhanyagolásokat tett, amelyek ennek az egyszerűsített számítási paradigmának az érvényességét valóban az elektroncsöves megvalósításokra korlátozza. Pontosabban, Neumann az *időzítési viszonyokat* határozta meg.

Az elektroncsöves rendszerben a számító elemek működési ideje a *msec*, köztük az adat szállítási idő a *μsec* tartományba esett. Neumann ezért teljes joggal tekintette úgy, hogy az adat szállítási idő elhanyagolható a feldolgozási idő mellett. Előre látta azonban a technika fejlődését (ha a tranzisztor és az integrált áramkörök megjelenését nem is), és figyelmeztetett, hogy az egyszerűsített számítási paradigmája nem alkalmazható sem a “túl gyors elektroncsövek” (azaz a 0.1 MHz frekvenciát túllépő központi szinkronizáló órajelet használó, bármilyen elektronikus technológiával megvalósított rendszerek) sem a biológiai számítások [86] esetére; az utóbbit az “unsound” jelzővel illette. A technikai fejlődés igazolta óvatosságát: az általa megengedett időzítési viszonyoknak utoljára az Intel 8080 processzor felelt meg.

Az azóta tervezett és készült processzorok lényegében különféle megoldásokkal [145] (órajel tartományok, késleltető kapuk, stb) próbálják legalább kiegyenlíteni az elektromágneses jel véges terjedési sebessége által okozott időzítési késleltetésekből származó problémákat. Ilyen módon azonban csak növelni lehet az adat szállítási időt, azaz a szállítási idő egyre csak növekszik a feldolgozási sebességhez képest (bár az áramkörök mérete egyre csökken, a feldolgozási idő a technológiai változások következtében gyorsabban csökken). Mára ez az arány a processzorokon belül is legalább megközelíti a biológiai számítás eleve “unsound” időviszonyait. Ez azt jelenti, hogy a modern technológiával készült számítógépek nem a számítástudomány által alapnak tekintett absztrakt (matematikai) processzor működési elvét valósítják meg, így az egyre kevésbé tudja leírni működésüket. Erre figyelmeztetett Feynman [62] már évtizedekkel ezelőtt: *a számítógéptudomány (és ezen belül a MI) nem a matematikai, hanem a mérnöki tudományok közé tartozik.*

Egy számítógép processzor esetén a láncolt műveleteket ugyanaz a fizikai számítási egység valósítja meg, azaz az eredményt csupán a processzor kimeneti szekciójából a bemenetire kell szállítani, viszonylag rövid úton azaz rövid adat-szállítási idő alatt. A neurális hálózatok esetében viszont a láncolt műveleteket különböző (elektronikus vagy biológiai) számítási egységek valósítják meg, ami aránytalanul nagy adat szállítási időt jelent, még a hagyományos számítógépes rendszerekhez képest is. Emiatt a tényleges működési mód még jobban eltér a számítástudomány által feltételezett módtól, így *a matematikai módszerekkel származtatott működési jellemzők szinte egyáltalán nem érvényesek a neurális hálózatok túlnyomó részére.*

A biológia természetesen (a több milliószor alacsonyabb kölcsönhatási sebesség miatt) már régen észrevette, hogy a viselkedést nem a newtoni mechanikából származtatott pillanatnyi kölcsönhatással lehet leírni, hanem

az események tér-idő szerűen viselkednek. Bár százhusz éve létezik a tér-idő koordináta rendszer [70], amihez kapcsolódó matematika használatával korrekt módon leírható a biológiai számítás viselkedése a kölcsönhatási sebesség (conduction velocity) által összekapcsolt térben és időben, a biológia inkább matematikailag könnyebben kezelhető szeparált tér-idő kezelést megvalósító közelítéseket használ (pl. [80]), félrevezetve önmagát.

Levonhatjuk azt a következtetést, hogy *a számítástudomány sem a modern technikai számítási rendszerek, sem a biológiai számítás működését nem tudja helyesen leírni*; ez a biológiát imitáló MI működésére is igaz. A számítástudományi elméleti háttér azonban kiterjeszthető az adattovábbítási idő figyelembe vételével, és az így általánosított elmélet már helyesen írja le az említett számítások működését. A kiterjesztett elmélet azonban még csak pár éve létezik és széleskörű használata csak a közeli jövőben várható.

A.2 A számítás elmélete

Egy számítás végzésekor rendelkezünk kell olyan objektummal, amelyik képes a számítás elvégzésére (utasítás) és olyannal, amin a számítást elvégezzük (adat). Az adatokat valamilyen formában (signal coding) valamilyen útvonalon (signal path) el kell juttatni egyik csomópontból a másikba, megadott időkorlátok között (signal timing). Amennyiben nem egyetlen elemi művelet elvégzése a cél (egyetlen utasítás és egyetlen adat), akkor ún. láncolt műveleteket kell végeznünk, és akkor már meg kell határoznunk a műveletek elvégzésének sorrendjét is. Ezt Neumann János “proper sequencing” követelményként adta meg [144], aminek része az útvonal és az időzítés meghatározása is. Vegyük észre, hogy láncolt műveletek esetén a láncnak van eleje és vége; azaz minden művelet-sor csak véges számú műveletből állhat (és ennek fejében csak véges számú elemre van szükség) továbbá a teljes művelet csak véges ideig tart. Ennél nagyobb számú művelet végzéséhez gondoskodni kell a számítás újbóli elvégzéséről (hurok képzéséről). Ez a követelmény az egyik legfőbb korlát egy számítógép megvalósításakor: az egyik művelet eredménye a másik művelet bemenő adata lehet, ami jelentős korlátozást jelent a műveletek párhuzamos végzése szempontjából, továbbá akadályt jelent a “Petri-csésze intelligencia” [134] és a kvantum számítógépek [97] megvalósítása előtt is. Már ha a Neumann-féle számítási modellt kívánják megvalósítani. Másikat pedig nem neveztek meg.

A másik korlát, hogy *a klasszikus számítási modell nem tartalmaz kommunikációt és processzor együttműködést* (“Single Processor Approach” [35]) továbbá *kis mennyiségű adaton végzett nagy mennyiségű számításra vonatkozik*. A MI számítási műveletei erre távolról sem hason-

lítanak.

Különféle előnyei miatt technikailag megvalósított számítás esetén elegendő csupán a digitális számítást tekinteni: az elektronikus számítógépek elméleti alapja [144], beleértve az adatok átvitelét [129] (kommunikáció) is, kizárólag digitális jellegű. Az operandusok digitális mennyiségek: egész vagy lebegőpontos számok, logikai változók. Ezzel szemben a biológiai számítás analóg [135]; pontosabban maga a számítás analóg mennyiségeket (idő értékeket) használ operandusként és eredményként, bár a működés során szinkronizálásra használt impulzus-szerű jelek miatt sokan (tévesen) digitális jelek továbbításáról beszélnek. Arról beszélni, hogy az agy hány lebegőpontos műveletet végez vagy hány bitet tárol (lásd pl. [134]), legalábbis érdekes elképzelés. Az agy más fogalmakat használ.

A.3 A technikai számítási mód

Alapvetően kétféle módon állhatunk hozzá a számítás kivitelezéséhez. Neumann János az ún. “utasítás-vezérelt” (instruction-flow controlled) működési mód elméletét dolgozta ki. Ennek lényege, hogy a számítógép egyes utasításai tartalmazzák annak kódját, hogy az utasítás milyen módon tudja elérni a szükséges adato(ka)t. A “proper sequencing” követelményét a program-számláló valósítja meg: ez a speciális belső regiszter mindig tartalmazza az következő utasítás kódját tartalmazó tároló rekesz címét. Az adatok számozott tárolórekeszekben találhatók és egy közös speciális útvonalon (busz) lehet azokat elérni; a szükséges időzítést központi órajel szabja meg. Az adatok elérése jelenti a problémát; ezt szokás Neumann szűk keresztmetszetnek nevezni. Jelenlegi számítógépeink kizárólagosan ezt a működési módot alkalmazzák. *A számítástudomány erre a módra vonatkozik; ettől léteznek a kiszámíthatóság, bonyolultság, konvergencia, stb. fogalmak.* A kommunikáció nem tartozik a processzor számítási tevékenységei közé.

A másik lehetőség az ún. “adat-vezérelt” (data-flow controlled) működési mód [76] lenne, azaz hogy az adat tartalmaz olyan hivatkozásokat, hogy azt milyen utasítással kell feldolgozni. Ebben az esetben a “proper sequencing” követelménye azt jelenti, hogy az útvonal kiválasztásával meg kell határoznunk az adatok feldolgozásának sorrendjét; az időzítés a csomópontok sorrendjéből és a végrehajtási/szállítási időkből következik. A számítógépes történelem során épült néhány kísérleti számítógép, sőt a számítógépes processzor számítási teljesítőképességének elakadásakor az Intel is benyújtott olyan szabadalmat [79], amelyik emulálással lényegében adat-vezéreltté tenné a számítógép processzor működését, de még kis sorozatú kísérleti számítógép sem épült ilyen elven. Az egyik fő ok, hogy *az adat-vezérelt számításnak*

nincs kidolgozott elmélete, és ezért számítástudományi háttére sem (nem tévesztendő össze az ál-MI által használt utasítás-vezérelt módszerekre alkalmazható számítástudományi háttérrel, lásd B.4 függelék; amire éppen ezért vonatkoznak a B függelékben említett technikai korlátozások). Lásd még a [49] közlemény következtetéseit.

Ötletelés természetesen további hangzatos működési elvekkel is folyik (főként a Neumann szűk keresztmetszet kiküszöbölésére, bővebben lásd a B.5 függelék). Ezek közös jellemzője az elméleti megalapozatlanság és a számítás általános elméletének nem értése.

A.4 A biológiai számítási mód

Agyunk “adat-vezérelt” módon működik, ezért jogosan feltételezzük, hogy az ezt valóság-hűen imitálni kívánó megoldásoknak is adat-vezérelten kell működni. Szenzoraink a környezetből származó ingerek alapján adatokat (akciós potenciál kombinációkat) generálnak, amelyeket a neurális hálózat tagjai kapnak meg és abból további adatokat állítanak elő. Neuronjaink (hálózati csatlakozásaikkal együtt) képviselik az utasításokat: speciális (örökölt, továbbá rövid és hosszú távú tanulási) mechanizmusaik használatával feldolgozzák az adatokat. Esetükben a Neumann által megkövetelt “proper sequencing” követelményét az adatok neurális csomópontához érkezésének ideje valósítja meg, az adatok utasításhoz rendelését pedig az axonális összekötések végzik el. Az egyes elemi műveletek időzítését (beleértve a bejövő adat érkezésének és az eredmény elkészültének jelzését is) maga a “signal coding” biztosítja.

Az agy hatékony működésének egyik titka, hogy a lehetséges összekötések töredékét (kb. tucatnyit a több ezerből) használja egy-egy művelethez, és (az időzítéstől függően) akár műveletenként változhat, hogy éppen melyik neuronoktól származó adatok vesznek részt a feldolgozásban. Agyunk annyiban is speciális eset, hogy tanulással módosítható az utasítás (tehát az utasítás lényegében analóg), mivel a kommunikáló elemek az érkezési idő alapján változtatni képesek a kommunikáció sebességét (azaz lényegében az utasítást). Mivel mind a bemenő adat, mind az eredmény idő jellegű, *a neurális processzornak egyetlen műveletre kell képesnek lenni:* mérni az első beérkező neurális impulzus és a membrán küszöb feszültség elérése közötti időt. A neuron nem képes másfajta vagy bonyolultabb feladat megoldására, viszont közben autonóm módon változtatni tudja (rövidebb-hosszabb időre) bemenő argumentumainak súlyát (saját szinaptikus érzékenységet), jutalmazva/büntetve a jó/rossz eredményt szolgáltató bemenetet. Ennyire egyszerű működésű processzor esetén csak sok processzor együttműködése ered-

ményezheti bonyolultabb feladatok megoldását. Ráadásul ezek a processzorok “egyszer használatosak”, azaz a működési módnak tartalmaznia kell az “elromlott” processzorok pótlásának és a pontatlan eredmények figyelembe vételének módját is; továbbá biztosítania kell egy mechanizmust az utasítások “finomhangolására”.

A biológiát utánzó (biomorph) megvalósítások rendszeresen figyelmen kívül hagyják a biológia lényeges működési tulajdonságait; ez alól a MI sem kivétel. Eléggé ironikus módon, ezek a megvalósítások lényegében a MI osztályozási munkamódszerét használják, de hibás adatkészlettel: a látott *technikai* megoldások közül a *funkciója* alapján leginkább odaillőt kiválasztják a *biológiai* megoldást utánzó szerkezetnek, függetlenül attól, hogy annak *működése* hasonlít-e a biológiai kivitelezéshez. Ilyen példák az analóg mennyiségekkel történő egyszerű számítás párhuzamba állítása a digitális mennyiségekkel történő komplex számítással, az ön-szinkronizáló (aszinkron) analóg neurális impulzusok utánzása központilag szinkronizált digitális impulzusokkal, a milliárdnyi alacsony sebességű párhuzamos analóg jel átvitel megvalósítása egyetlen nagy sebességű digitális vezetékkel, a biológia helyi információtárolás időzítés-függő és korlátozott nagyságú megváltoztatásának implementálása időzítés-független és korlátozás nélküli megváltoztatásával távoli tárolóban, stb. Gyakorlatilag csak a név a közös a két terület hasonló névvel illetett működésében. Ennyi és ilyen fokú működési eltérés esetén a technikai és biológiai számítás csak nagyon távolról nézve hasonló. Nagyon nem használhatók az egyedi neuronokra, azok csoportjaira vagy hálózataira vonatkozóan a felépítési vagy működési analógiák.

A.5 Az organikus MI (OI) számítási módja

A MI mindinkább halványuló hullócsillaga után újabb (2023 közepén még egyre fényesedő) hullócsillag, az Organikus Intelligencia (OI) (esetleg Petri-csésze Intelligencia) jelent meg 2023 elején [134] és indult el szintén megalapozatlan (de azért várhatóan néhány évig tartó) diadalútjára. A várakozás ismét hatalmas, máris az “új MI” névvel illetik az OI-t [18]. Kimondják az AI (MI) trónfosztását (“AI is over”) és kikiáltják az új királyt. Az OI annyiban ténylegesen jó utód, hogy egyiknek sincs elmélete és mindkettőnek ismeretlen/befolyásolhatatlan a működési módja. A hasonlóság alapján, tanulva a MI-val kapcsolatos félelmekből, rögtön születése után mondják: “jön az OI és elveszi a munkahelyeket” [18]. A helyes állítás a következő lenne “(évek múlva esetleg) jön majd az OI és – bár teljesen másképpen – éppúgy kiszámíthatatlan és megbízhatatlan lesz és ugyanolyan problémákat okoz, mint a MI”. Az elektronikai óriások számára természetesen

még mindig a MI az első számú megvalósítandó eszköz és cél [114, 26].

A sok tématerületet jegyző neves szerzők bizonyára számos szempontból elemezték a nagyon remélt biológiai számítógépet [134], de annak számítógéptudományi megalapozása nem szerepelt közöttük. [122], bár méltatja a biológiai számítás potenciálját, – számos kétség megfogalmazása mellett – felteszi a nagy kérdést: lehet-e Petri-csészében szuperszámítógépet növesztetni? Amint a A.4 szakaszban tárgyaltuk, az agy neuronjainak hálózata éppen a sokszoros összekötések által speciális (valamennyire hangolható) utasításokat jelent. A speciális agyi területek specializált utasítás csoportokat valósítanak meg, az agyi területek felépítését pedig a genetikai állomány határozza meg. A kérdést tehát úgy kell feltenni: lehet-e a genetikát úgy módosítani, hogy a növekvő sejt állomány összeköttetései úgy jöjjenek létre, hogy a Petri-csésze szuperszámítógép-szerűen működjék?

Az emberi agyat meghatározó genetika elég hosszú fejlődés eredménye (beleértve a környezettel való kölcsönhatást, a tanulást), ezért a szuperszámítógép-genetikát kifejleszteni jelentős biológiai feladat lesz. Enélkül az OI [134] csak egy tányérnyi sejt marad. Ha az egér genomikája módosítható úgy, hogy annak növesztett agya egy egér helyett egy szuperszámítógépet működtessen akkor nyertünk. Ami egyrészt pár millió évbe beletelik, másrészt csak célfeladatra hozható létre (eredetileg az egér működtetése volt a cél feladat, most egy bizonyos számítás végzése lesz; *ha sikerül, visszajutunk a Neumann előtti, huzalozással programozható számítógéphez*), harmadrészt etikailag kérdéses, hogy szabad-e ilyen szörnyszülött genetikával kísérletezni.

A sejtek a gének által meghatározott módon vannak összekötve (MI terminológiával: pre-trained) ezért csak az előre meghatározott funkcionális tudják mutatni (esetleg egérként tudnának működni, csak a Petri csészében az egér lényeges alkatrészei hiányoznak). A biológiai megvalósítás ugyan magával hozza a párhuzamos, egyidejű, sorrend-helyes adatszállítást, de az eredeti célnak megfelelően létrehozott hálózattal. A résztvevő neuronok és axonok kombinációja csak a genetikusan kódolt cél megvalósítására alkalmas, továbbá myelin rétegük vastagsága nem egyezik a természetesével, így működőképességük még bizonyításra szorul (még az eredeti célra is). A valódi MI műszaki megvalósítása esetén is előre rögzített az adattovábbítás útja, de ott elvileg a rétegek között létező adatutak mindegyike használható és tanulás bárhol lehetséges (az más kérdés, hogy a matematikai módszer miatt mennyire jól működik). Az OI megvalósítás esetén pedig csak az az adatút használható, amit a genetika engedélyez és tanulás is csak ott valósítható meg.

A biológiai rendszerek párhuzamos működésének ára, hogy (a neurális

hálózatokkal egyező módon) a hálózatba rendezett neuronok bemenetére adatot adunk, és a számítás eredményét a hálózat kimenetén kapjuk meg. Az újabb számítást csak újabb adat betáplálása indítja el. Hogy az adat-vezérelt biológiai számító rendszerből utasítás-vezérelt általános célú (azaz programozható) számítógép processzor legyen, (egyebek mellett) a neuronokat meg kell győzni, hogy – a biológia eddigi gyakorlatától eltérően – ezentúl ne adatokat, hanem utasításokat továbbítsanak; a továbbított utasításokat kódolják/dekódolják; az adatokat címekkel hivatkozzák és ne folyamatokként, hanem tárolóregiszterként használják; bemenő adataikat és eredményeiket ne változó analóg időként, hanem állandó digitális feszültségként kezeljék; ne időt mérjenek, hanem aritmetikai számításokat végezzenek; továbbá ki kell találni a vezérlőegység növesztésének genetikai módszerét.

Fel kell mérni azonban, hogy az OI úgy is tekinthető, hogy a közbülső szuperszámítógép (mint rejtett réteg) átugrásával, a maga természetes valóságában van egy “artificial neural network” (ANN) a kezünkben, ami valóban adat-vezérelten működik, feltérképezhető, egyfajta “reverse engineering” útján megérthető és egyszerű funkciókra használható. Hátránya, hogy elő-tanított (viszont sok hálózat van és lehet válogatni); előnye, hogy a biológiai megvalósítással eltűnik a párhuzamosított szekvenciális működés számítási teljesítmény korlátozó hatása. Potenciális veszélyforrás. Új fejezetet nyithat a tanulásban; de ehhez el kell fogadni, hogy a biológiai és a technikai információ ábrázolás gyökeresen eltér. Az elektronikusan előállított feedback ki tud kényszeríteni neurális működést [57], de bio-idegen jellege miatt nem tud tanulási folyamatot indítani. Új fejezetet nyithat a matematikában és a neurofiziológiában: új leírási módokra és modellekre lesz szükség. *Fel kell nőni a gondolathoz: a TI, (valódi) MI és OI rendszerek adat-vezéreltek.* Arra készültek, hogy egy bizonyos célrendszert (mondjuk minket) működtessenek, változó körülmények között. *A fogalmak összekeverése miatt eddig nem tűnt fel, hogy a többségében gépi tanulás elvén működő (ál-)MI-val ellentétben az OI valóban adat-vezérelt elven működik.* Ha működik. Kérdés, hogy a borsejtekből származtatott “agy-sejt tenyészet” mennyire mutatja az agy valamely részének funkcionalitását (hogyan vannak az egyes neuronok axonokkal összekötve és az agytól eltérő myelin-fedettség mennyire egyezik a neurális kommunikáció időviszonyainak biztosításához szükségessel).

A.6 Idő-ablak

Mind a technikai, mind a biológiai számítás idő-ablakokban működik, azaz csak azokat az műveletet veszi figyelembe a rendszer, amelyek teljes időbeli hosszúságai az idő-ablakba esnek. Az idő-ablakot a számítás különböző

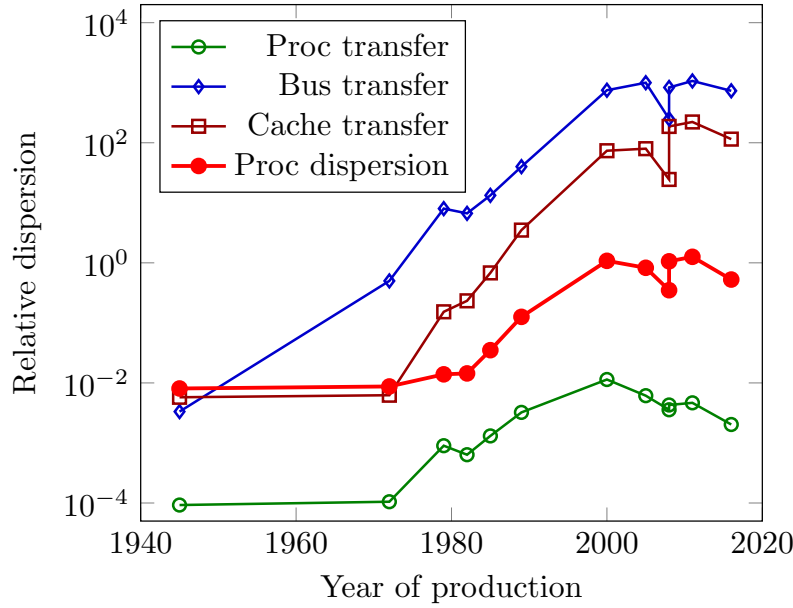
absztrakt szintjein különböző módon értelmezik, és minden szinten cél szélességének csökkentése (a művelet gyorsítása), amelynek során biztosítani kellene a művelet helyességét. Az idő-ablak kijelölésében azonban alapvető különbségek vannak van.

A technikai számítás esetében Neumann az elemi műveletek végzésére az egyszerűbb megvalósíthatóság érdekében a *szinkron* működtetést javasolta, legalábbis az elektroncsöveknek megfelelő időzítési viszonyok között. Ez azt jelenti, hogy egy központi órajel felfutó és lefutó élei határozzák meg az időablak szélességét, és a számítást olyan elemi műveletekre kell bontani, amelyek mindegyike befejeződik egy-egy órajel alatt (azaz a számítási elemi műveleteknek alkalmazkodni kell a külső órajelhez). A bonyolultabb számításokat ilyen műveletek láncolásával lehet megvalósítani, így a művelet teljes hosszát az elemi műveletek számának és az órajel szélességének szorzata határozza meg. A láncolás megvalósítását könnyíti hogy csupán a kimenő egységből kell az eredményt a bemenő egységhez juttatni, ami az egységen belüli adatszállítást és nem egységek közötti adattovábbítást jelent.

Neumann a központi órajel vezérlés használatát átmeneti megoldásnak gondolta és csak addig alkalmazhatónak, amíg a diszperzió (az órajel alatti veszteség idő aránya) elhanyagolható. A működési elv azonban tartalmazza az adatok elővételének idejét a memóriából, és számító elemek végrehajtási sebessége sokkal gyorsabban nőtt, mint a memóriák működési ideje, ezért a processzor fejlődése során (főleg az integrált áramkörökkel megvalósított nagy alkatrész sűrűség miatt) a diszperzió jelentősen megnőtt (lásd 7.1 ábra). Bár manapság teljesen mások az időzítési viszonyok, a technikai számítások továbbra is egy centrális órajel által kezelt időablakot használnak.

A számítógép processzort működtető jelek tudnak az órajelről, minden működtető jel összhangban van vele (késlelteti magát). Az elemi gépi utasítást az órajel indítja és minden működésének (beleértve a távoli és lassú memóriák elérését is a buszon keresztül) be kell fejeződnie az időablakban, mivel a vezérlés az idő-ablak végén mindenképpen továbblép a következő utasítás végrehajtására. Ha valami nem fejeződik be, akkor meg kell várni a következő órajelet. A különféle késleltetések miatt csökken az órajel hasznos kitöltési tényezője (növekszik a diszperzió és azzal együtt a nem számolásra fordított teljesítmény felvétel).

Neurális számítás esetén a "neuron" hosszabb-rövidebb ideig hagyományos szekvenciális programot hajt végre, majd több üzenetet is küld a társ-neuronoknak. Számítógép processzorokkal megvalósított neuronok esetén, mivel a processzorok szegregáltak (Single Processor Approach [35]), az üzenetet csak I/O utasítások igénybevitelével tudják elküldeni, ami az operációs rendszer igénybevitelével (kontextus váltás [54, 140]) kb. 10^4



Ábra 7.1: A processzor diszperzió (az órajel kihasználtsága) mértéke a számítógép processzor fejlődésének története során. Saját szerkesztésű ábra.

gépi utasítást (ennyi központi órajelet) igényel, még ha a busz hatását (lásd B.5 szakasz) figyelmen kívül hagyjuk is. Ezt a számot még szorozni kell az elküldendő üzenetek számával. A *“láncolt művelet” ideje tehát a tényleges neurális számítási időből és a nagyságrendekkel hosszabb szállítási időből tevődik össze*. Időviszonyait tekintve, ez a működési mód nagyon hasonlít a biológiaihoz, amire vonatkozóan a klasszikus számítási paradigma használatát (azaz a számítástudomány illetékességét) Neumann az “unsound” jelzővel illette [144].

A biológiai számítás *nem használ központi órajelet* [36], hanem *aszinkron* módon működik: egy neurális spike érkezése “lyukat üt” [87] a nulla vezetőképességű membránba. Az idő-ablakot az első bemenő adat érkezése nyitja meg, és a kellő mennyiségű töltés beérkezése után maga a számítógység (a neuron) zárja be. Ez a működés garantálja a veszteségmentes idő kihasználást. *A modell szerint az időablakban a számítás során figyelembe vett bemenetek, azok súlya és beérkezési ideje műveletről műveletre változhatnak, ami egyrészt nehezen meghatározhatóvá teszi a művelet eredményét, másrészt – a tanulási mechanizmussal együtt – lehetőséget nyújt a redundancia, rehabilitáció, intuíció, asszociáció, stb. fogalmak megvalósítására*. A biológiai

tapasztalat szerint [69] “a tanulás során újabb információ források is bekapcsolódnak a már megtanult viselkedési formák javítására”, ami biztosítja, hogy a rendszer meg tudja tanulni a helyes reagálást egy addig ismeretlen helyzetre, ezáltal [37] meghatározásának értelmében is intelligens legyen.

Az időablak különleges szerepet tölt be ha több független végrehajtó egység (több “neuron”) munkáját kell koordinálni. Ilyenkor az egységek “kettős tudatúak”: ugyanazon biológiai időben de eltérő számítási időben működnek és az egyes “időrések” végén az egységek elküldik egymásnak aktuális állapotukról a releváns információt. Ezzel gyakorlatilag ugyanazt teszik, amit a központi órajel is csinál: az egyes műveleteknek az időrés végéig be kell fejeződnieük, majd a (nagyon nagy kommunikáció igényű) szinkronizálás után az összes végrehajtó egység munkája folytatódik. A szokásos 1 *msec* időrés használatával lényegében egy nagyon nagy számítási teljesítőképességű, de mindössze 1 *MHz* órajel sebességű processzort használnak.

Vegyük észre azt a lényeges különbséget, hogy a biológiai számítás önmaga definiálja a számítási időablak végét, a technikai számítás számára pedig kijelöljük azt. Más szavakkal, a biológiai számítás közvetlenül az eseményekkel dolgozik, azok bekövetkezési idejétől függetlenül, a technikai számítás során pedig az eseményeknek a tervezéskor figyelembe vett időben kell bekövetkeznie. Ha a számítás még nem fejeződött be amikor megpróbáljuk elszállítani az eredményt, a számított eredmény *nem biztosan* egyezik az elvárttal.

Busz használata és/vagy hálózati működés esetén (a működés kvázi-probabilisztikus jellege miatt) a szükséges idő-ablak szélesség nem ismert tervezés idején, ezért azt becsülni kell (az igazi megoldás az esemény vezérlés lenne). A hálózat-alapú működés (még a μs nagyságú kontextus váltáshoz képest is) nagyságrendekkel növelheti az adat-átvitel idejét, és az észszerűtlenül nagy rendszerek esetében a busz működési ideje is észszerűtlen (lásd B.5 függelék); hasonló következménnyel jár a valós idejű működés kényszere (lásd 4.1 szakasz). Ha a legrosszabb esetre tervezünk, túl lassú lesz a rendszer. Ha ennél rövidebb időre, bizonyos adatok nem érnek célba az elvárt időre. Hogy az adatok milyen nagy része, az a működés körülményeitől függ.

A válaszidő hosszúságának marketing szempontja rákényszeríti a tervezőt az időablak végének észszerűtlen kijelölésére. A hibás működés veszélye CPU használata esetén is fennáll, de ott (a pipeline késleltetése árán) elegendő egy-egy órajel ciklust kihagyni. Párhuzamosított végrehajtás (GPU vagy TPU) esetén, különösen több független ilyen egység használatakor, és még inkább a rendszeren belüli hurkok létrehozásakor (pl. visszacsatolások, “*deep learning*”) a veszély fokozódik, lásd B.6 függelék. *Az idő-ablak kezelése jelentős ál-probabilisztikus járulékot adhat a működési időhöz és akár a berendezés MI osztályba sorolását (lásd 2.5 szakasz) is megváltoztathatja. Az idő-ablak*

kijelölése (a rendszer nagyságától és a működés aktuális körülményeitől függően) szinte biztosan hibás akkor, ha a számítás vége (mint esemény) nem határozható meg pontosan.

A.7 A számító egységek működése

A technikai számítások esetében az elemi kapuk a matematikai logika összefüggéseit, a gépi utasítások pedig nagy számú kapu felhasználásával lényegében matematikai függvényeket valósítanak meg. Az utasítások egymás után következésének definiálása és a felhasznált adatok kijelölése még a program írásakor megtörténik. Minden egység szinkron működésű. A technikai szervezés biztosítja, hogy a számítás csak akkor indul el, amikor valamennyi argumentum megérkezett a bemenetre a közösen használt buszon, és az eredményt (szintén a buszon) csak akkor kapjuk meg, amikor az egység matematikai leírásának megfelelő eredmény elkészült. A működés során digitális jelfeldolgozás történik. Ilyen értelemben a technikai számítások elemi egységei önállóan is használhatók, viszonylag bonyolult műveletek végzésére képesek és működésük determinisztikus (a működés eredménye matematikailag kiszámítható). A technikai megoldások (órajel késleltetés, domének, stb) biztosítják, hogy az azonnali kölcsönhatás feltételezése alapján készített egységek (bár energia-pazarló módon) de lényegében a (már említett értelemben: nem releváns) klasszikus számítási paradigma alapján működnek. *A működés utasítás-vezérelt.*

A biológiai számítás [86] esetében a számítás elemi egysége a neuron, aminek az absztrakt megfelelője egy feszültség-vezérelt kondenzátor. A neuron viselkedése matematikailag igen nehezen írható le. Nagy számú (több ezer) bemenete van, és egy-egy számítási műveletben csupán tucatnyi bemenet súlyozott és kapuzott jelét használja. Mind az argumentumok, mind az eredmény idő jellegű; a jelfeldolgozás analóg módon történik [135]. A töltés a kondenzátorból folyamatosan “szivárog”, ezért a bemenetein beérkező töltések érkezési ideje nagyon sokat számít. Az eredményhez az idő-ablak nyitva tartási idején belül beérkező töltések mind hozzájárulnak, növelik vagy csökkentik a membrán potenciált. A neuronnak (különböző biofizikai mechanizmusokkal megvalósított) memória-szerűen működő állapotai vannak [91]; a számítás újraindítható [92] és eredménye az aktuális neurális környezettől és a neuron saját érzékenységétől [81] is függ. A jel továbbítás és feldolgozás során erősen figyelembe kell venni az alacsony vezetési sebességet, azaz a klasszikus számítási paradigma egyáltalán nem használható. Az összes tényező ismeretében ugyan leírható a működés, de a nagy számú befolyásoló tényező és főként az időfüggés sokrétűsége miatt a működés sokkal kevésbé

determinisztikus, mint a technikai számításé. *A működés adat-vezérelt.*

A.8 A számító egységek kihasználása

Ironikus módon, Neumann az *adat-vezérelt működésű* agy inspirálta az *utasítás-vezéreltelektronikus számítógép* működési modelljének megalkotására. Az okot a számítógépes egységek kihasználtsága, illetve az ebből következő alkatrészigény, továbbá a működés időtartama jelenti.

A meglehetősen bonyolult számításokat elemi számítási műveletekre kell felbontanunk és a különféle műveleteket azokból összeraknunk. Mind adat-, mind utasítás-vezérlés esetén a láncolt műveleti egységek első elemére adjuk a bemenő információt, ami működteti a számítási láncot, és a lánc utolsó eleme szolgáltatja az eredményt. Bonyolultabb számítás végzéséhez vagy sok elemi számító objektumot kell biztosítani (nagyrészt párhuzamos működés) vagy az elemi objektumokat – megfelelő szervezéssel – többször is felhasználni (ciklikus-szekvenciális működés). Az első változat nagy alkatrész pazarlással jár, nagyfokú párhuzamosságot biztosít, de szinte semmi szervezést nem igényel; a második magasabb rendű szervező egységet igényel, minimális párhuzamosíthatóságot engedélyez, de akár korlátlanul hosszú működést tesz lehetővé.

Adat-vezérelt működés esetén minden adatnak saját, kizárólagosan használt útvonala van. Ennek ára azonban hogy az egyes számító egységek (amik ezúttal az *utasítások*) “egyszer használatosak”: ugyanaz a számítási egység mindig ugyanazt a műveletet hajtja végre a hozzá beérkező adatokkal, tehát nem lehet több célra használni. Ezért szükséges az agy számára száz milliárdnyi neuron (azaz lényegében utasítás), mert gyakorlatilag minden elemi műveletnek másik neuront kell használni (még ha van is “újrafelhasználást” biztosító magasabb szintű szervező egység). Ugyanilyen számú kapuból lehetne teljesen adat-vezérelt működésű végrehajtó egységet is építeni, de az célberendezés lenne: minden adatot ugyanolyan módon dolgozna fel. A technika számára megoldhatatlan az agy dinamikus működéséhez elengedhetetlen biológiai mechanizmusok utánzása, főként az idő figyelmen kívül hagyása miatt, továbbá a szükséges huzalozás kivitelezése is. (Az OI rendszerek ugyan biztosítanak elegendően sűrű és sok huzalozást, de azt a genetika és nem a feladat határozza meg, lásd A.5 függelék). A modern (főként a sok-magos) processzorok ugyan nagyságrendileg hasonló számú kaput tartalmaznak mint az agy, de a ciklikus-szekvenciális utasítás-vezérelt működési mód nem lép fel ilyen huzalozási követelménnyel. Emiatt kompromisszumot kell kötnünk a használható alkatrészek száma és a közöttük levő huzalozás mennyisége, valamint a processzor felépítésének bonyo-

lultsága és működési sebessége között.

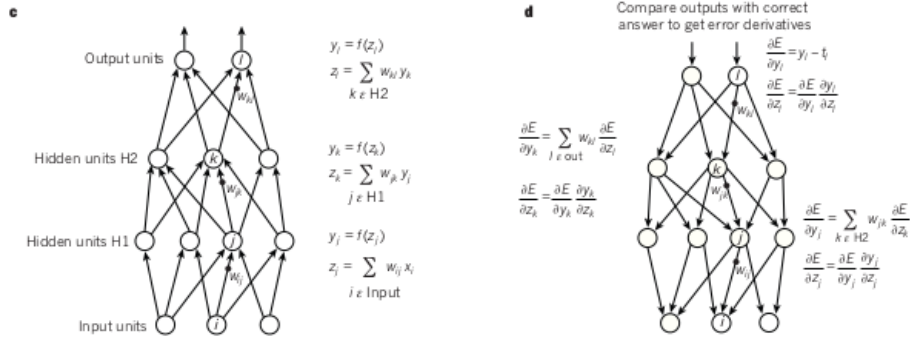
Az általánosan használt utasítás-vezérelt működés esetén lényegében egy igen bonyolult ciklikus-szekvenciális processzort használunk, amelyik a bemenetére (“input section” [144]) érkező *adatot* utasításként értelmezi és annak dekódolása után az *így kapott utasításnak* megfelelően állítja be belső állapotait és adatútjait. Ezután a processzor működését már a bemenetére érkezett adat vezérli: az elemi számítás végéig a processzor adat-vezérelten működik és az adatok végig siklanak az előzőleg kijelölt adatúton. Az eredmény a kimeneti szekcióba (“output section”) kerül és a processzor ismét utasítás-vezérelt módban dolgozik tovább.

Az adat-vezérelt működés során az adatok (pl. negatív-e az eredmény) módosíthatják az utasítás számlálót: a technikai processzor lényegében a kétféle vezérlési mód kombinációját (*sőt, a megszakítás kezelés alkalmával az esemény-vezérelt működést is*) használja. Az *átmeneti utasítás-vezérelt mód feltétlenül szükséges, hogy a processzor programozható legyen*. A kapott utasításnak hivatkozni kell tudni a feldolgozandó adatra, amit egy tárolóbeli címként adunk meg. Az újra-konfigurálható rendszerek éppen attól hatékonyak, hogy előre (és nem utasítás dekódolás után) beállított adatutakat használnak és lehetővé teszik az adatutak párhuzamos használatát. Lehetne azonban egy közbülső (configware) réteget is használni a hardver és szoftver rétegek között.

Vegyük észre az idő-ablak szerepét. Mivel az idő-ablakot a központi órajel jelöli ki, a bemenő argumentumnak a szinkronjel érkezéséig meg kell érkeznie és az eredménynek el kell készülnie a következő szinkronjel megérkezéséig, bármilyen műveletről legyen is szó.

B A technikai MI számítási működése

Mint várható, a TI megvalósítása technikai eszközökkel a korábban a technikai számítás céljaira kifejlesztett hardver és szoftver eszközökkel, az ott megszokott elvek alapján történik. Valamilyen módon modellezzük a neuronokat (lehetnek logikai kapuk, processzorok, hardver vagy szoftver szálak), és az egymással összekötött neuronok egymást valamilyen módon értesítik; azaz egymás jeleit használják bemenő adatként és eredményeket küldenek egymásnak. Ebben az esetben az adatok digitális jelek, a neuron hálózat pedig kijelöli, hogy az adatokkal pontosan milyen műveletet kell elvégezni. Ez a *művelet tipikusan egy szorzás+hozzáadás*. A művelet elvégzésének technikai módja alapján fontos megkülönböztetnünk (majdnem) valódi MI (a neuronok galvanikusan kapcsolódnak egymáshoz) és ál-MI (a neuronok üzenetekkel, hálózaton vagy buszon keresztül kommunikálnak)



Ábra 7.2: A mélyrétegű tanulás (deep learning) során alkalmazott számítási módszer, visszacsatolással [89]

megvalósítást.

B.1 Szinaptikus művelet

Neurális hálózatok esetén egy-egy csomópontban a számítás eredménye az összes szinaptikus bemenet értékének súlyozott összege (lásd 7.2 ábra bal oldala); ennek értékéből számítják a következő réteg egyes bemenő adatait. Elvileg ezt a műveletet egy Single Instruction Multiple Data elven működő számítási egység (pl. egy Graphic Processing Unit, GPU) egyetlen műveletként el tudja végezni, majd az eredményből egy nemlineáris függvénnyel képezik a következő réteg bemenetének értékét. Az egység bemeneteire azonban az adatokat oda szállítani és az eredményeket a kimenetekről elszállítani egyesével kell (az egyes adatok külön busz ciklust igényelnek). A programozás stílusa [39] nem változik, továbbá a nagy számú neuront tartalmazó rendszerekben csak több részletben lehet a számítást elvégezni. A hosszadalmas művelet végzés jelentősen akadályozza az utasítás végrehajtási “csővezeték” (pipeline) működését. Az idő-ablak vagy aránytalanul hosszú lesz, vagy a számítás egy része nem fejeződik be az idő-ablakon belül. Az utóbbi eset gyakori a működést gyorsító megoldások esetén és a működés bonyolultsága legtöbbször elfedi ezt az elméleti hiányosságot. A számítási hatások lényegesen függ attól, hogy a CPU és a GPU között másolni kell az adatokat (pl. Intel+NVIDIA) vagy osztott memóriával dolgoznak (pl. AMD).

B.2 Tanulás

Logikailag a műveletet a csomópont valamennyi “szinaptikus bemenetre” egyidejűleg elvégzi, tehát *nincsenek* sem aktuálisan aktív bemenetek, sem beérkezési idő különbségek. Maga a csomópont nem képes szinaptikus bemeneteit “finomhangolni” és egy visszacsatolt jel sem képes arra, hogy *szelektíven* változtassa valamelyik bemenet szinaptikus súlyát (nincs a Hebb értelmében [68, 45] vett tanulási mechanizmus). A bemenetek szinaptikus súlyát csak a (matematikai módszereket használó) felügyelő program képes, de nem szelektíven (a bemenő argumentumok érkezési ideje között nincs különbség). Ezért *a technikai és a biológiai tanulás alapvetően különböznek*. A népszerű és gyakran használt “backpropagation” módszer [89] (lásd 7.2 ábra jobb oldala), a várt és kapott eredmény birtokában gradienst számít és annak alapján változtatja az előző rétegbeli neuronok szinaptikus súlyát. Mint [95] hangsúlyozza, ennek a matematikai módszernek nincs biológiai alapja. Nem véletlenül: a matematikával ellentétben, a biológia ismeri az említett idő-ablakok szükségességét.

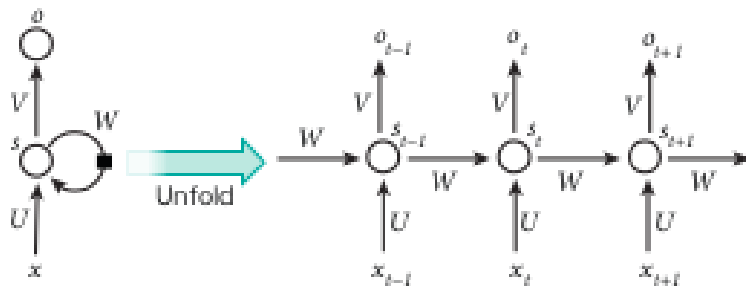
A működés során az aktuális következő rétegnek mindaddig nincs munkája, ameddig az előző réteg be nem fejezte a munkát (lásd 7.2 ábra bal oldala): *az elemi számító egységekre vonatkozó idő-ablak mellé a rétegekre vonatkozó idő-ablakot is be kell vezetnünk*. Az előzőleg is használt elv értelmében, ez a két idő-ablak sem fedhet át és a két réteg közötti szállítási időnek megfelelő idő-ablaknak köztük kell lenni. Ezzel szemben az ábra azt mutatja, hogy az egyes rétegek kimeneteinek értékét a rétegek egymástól függetlenül számítják ki és az ilyen számításokat megvalósító hardver egységek is ebben az értelmezésben végzik a számításokat. Nincs olyan jel, ami pl. az egyes rétegek működése között biztosítaná a “proper sequencing” követelményét. Hasonlóan, a gradiens számítás pedig csak akkor kezdődhetne el (lásd 7.2 ábra jobb oldala) amikor a számítási eredményt az egység teljes mértékben előállította. Az utóbbi számítás során is várakozniuk kell a rétegeknek egymásra. [89] megemlíti, hogy a számítást idő lépésekben kell végezni és hogy *a felhasznált paraméter később áll elő, mint ahogyan használjuk*. Ezt a helyzetet a programozásban határozatlan értékű változóval való számításnak hívják. Azért nem tűnik fel, mert a kezdőértékeket jól meghatározott értéktartományból választják és az a számítás elején általában egy ebbe a tartományba eső véletlen érték. Ha nem ilyen kezdőértékről indul a számítás, akkor furcsán működik: nem veszi figyelembe a kezdeti értéket. De, még véletlen kezdőérték esetén is rossz időpontban kap értéket (a biológiai ciklikus működés téves értelmezése). A matematika végtelen gyors adatátvitellel dolgozik, a biológia és az elektronika

meg nem.

A súlyok változtatásának csak az után lenne szabad működésbe lépni, hogy a kimeneti rétegben előállt az eredmény, hiszen mindkét irányban ugyanazokat a súlyokat használják a számítások. Ennek hiányában előfordulhat, hogy a két irányú számítás egyidejű végzésekor (gyorsítási lehetőség) a kimeneti adatok számításának egy részét még a régi, másik részét már az új súlyozással végzik, ami mindkét irányú számítás eredményét meghamisítja és ezáltal lassítja a számítást. Gyorsabb ciklus működés, de több ciklus szükséges a konvergenciához. Ez a hatás limitálja az alkalmazható mély rétegek méretét és számát [85].

Amint a A.4 szakaszban említettük, a biológiai rendszerek felépítése olyan, hogy *a neuron a biológiai élettartam teljes ideje alatt, a feladat végrehajtásával egyidejűleg tanul is, kommunikáció nélkül, autonóm módon. Az elterjedten használt technikai rendszerekben nincs olyan mechanizmus, ami erre képes lenne*, ezért a tanulást csak társ-neuronoktól származó információ alapján, kommunikáció felhasználásával lehet megoldani. Bár elvileg a rendszerek képesek lehetnek a “menet közbeni” tanulásra, azonban főként mivel a tanulás és a végrehajtás kommunikációs igénye (lásd a B.5 függelék) különböző erősségű hardvert igényel, a gyakorlatban használatos esetekben a működést szét kellett választani “betanulási” és “végrehajtási” fázisokra. A gyakorlatban sikeresen működő alkalmazások betanítása kifejezetten erős hardvert igényel, és a biológiai neurális háló számára néhány másodperc alatt megtanulható aktivitások betanulása hosszú ideig tarthat szuperszámítógépen is (főként, mivel mennél nagyobb a számítógép, annál gyengébb a hatásfoka MI feladat végrehajtásakor). Egy évtizeddel ezelőtt hetekig tartott [40], a MI és a számítógép technológia hét éves fejlődése után hónapokig [93].

A tanítás jelentős időigényének túlnyomórészt az az oka, hogy *a matematikai számítási módszerek nem tartalmazzák az időt*: nem alkalmazható a Hebb-féle elv [68, 45], azaz a tanítási folyamat nem tudja egyesével változtatni a bemenetek szinaptikus súlyait. Ennek következtében egyrészt megnő mind a szükséges számítások mind a kommunikáció mennyisége, másrészt a jutalmazás/büntetés csak a csapatoknak jár: ha a csapat jól teljesít, akkor a rossz inputot adó bemenet is jutalmat kap és fordítva. Emiatt a betanulás igen lelassul; annál nagyobb mértékben minél nagyobb a csapat. Vagyis *a rendszer méretének (pl. a rejtett rétegek számának vagy méretének) növelésével a rendszer számítási hatásfoka, valamint stabilitása és működésének helyessége is csökken*. A tanulás problémája már az “aranyláz” kezdete előtt nyilvánvaló volt, de a működés megértése nélkül nem találtak rá megoldást. A helyzetet jellemzi, hogy a DARPA (annak jeleként, hogy a



Ábra 7.3: A mély rétegű tanulás időviszonyainak “sorba fejtése” [89]

szakértők kifogytak az ötletekből) nyilvános ötlet pályázatot hirdetett [10].

A biológiai tanítás erőforrás alapú: a receptorok száma és állapota, a transzmitterek koncentrációja, az ioncsatornák áteresztő képessége, az axonok myelin rétegének vastagsága, stb. mind telítésbe futó görbékkel írhatók le, és a változások a *késleltetett hozzáadás* elvén működnek. A matematika viszont egyrészt nem használ változás korlátozást (tehát egyetlen lépésben bármekkora változást előidézhet a számítás) másrészt *azonnali felülírás* elven működik. A felülírás önmagában magyarázni tudja a katasztrofális felejtést.

B.3 Visszacsatolás

Rekurrens működtetés esetén tovább rontjuk a helyzetet. Mint [89] ki is mondja, *a mély rétegű tanulás rekurzív módja még csekély számú réteg esetén is igen nagy (elméletileg végtelen) számú, azonos súlyokat használó, lineáris rétegnek felel meg, lásd 7.3 ábra*. Nem-lineáris (rekurrens) hálózatokban a paraméter értékek nem csak az előző, hanem valamennyi korábban kiszámított paraméter értéktől függenek. A korrekt matematikai működéshez faktoriálisan növekvő számú számítást kellene elvégezni: egyetlen lépés (lásd a 7.3 ábra bal oldalát) lényegében egy végtelenszer ismétlődő számítási sort indít el, aminek minden tagja egy újabb számítási sort indít (lásd az ábra jobb oldalát), azaz szigorúan vett értelemben egy elindított számítás soha nem fejeződne be. A számítás magától nem tud leállni: azaz a számítás matematikai értelemben nem pontos. *Az egy lépéshez rendelhető időablaknak gyakorlatilag végtelen szélesnek kellene lenni, azaz az észszerű működési sebesség érdekében csonkolni kell a számítást*.

Már maga a matematikai módszer is magyarázza, miért olyan lassú a betanítás: a virtuálisan rögzített paramétereket még a helyes tanulási módszer is csak kínosan lassan tudja változtatni. (Ironikus módon, azért tapasztalható

tanulás az azonos súlyok ellenére, mert a késleltetések keverik a konstans súlyokkal rendelkező virtuális és az adjusztált súlyokkal rendelkező valódi rétegekben levő neuronok üzeneteit. Három lépést előre, két lépést hátra. A hálózati átviteli sebesség javítása és a feldolgozás sebességének növelése, pl. új technológiákkal, egyaránt rontani fogja a számítás energia-hatékonyágát, a nem értett működési elv miatt.) Ez azt jelenti, hogy tanulás szempontjából a rekurrens működtetés a lehető legrosszabb megoldás: a “tanítás” eredménye annál lassabban jelenik meg, minél nagyobb a rendszer (tényleges felhasználói tapasztalat), mivel a virtuális rétegek nagyon is valós kommunikációs igényrel lépnek fel. A működés csupán azért nem szinguláris, mert a véges végrehajtási sebesség (plusz adatszállítási idő) miatt nem végtelenül kis idő alatt lép fel az új számítás igénye. Talán nem véletlen, hogy a biológia nem alkalmaz ilyen módszert és azért a MI tanulási módszere a biológia számára nem releváns [95]. Sajnos, a matematika újra felfedezi az Achilleusz és a teknősbéka paradoxonját, ahol végtelen sok művelettel (végtelen sor) próbálják előállítani az egyszerű szorzás/osztás művelettel is kiszámítható eredményt. Amit aztán (a számítógép elméleti működésének meg nem értése miatt) az elektronikát gyártó cégek pontosan ilyen formában valósítanak meg, a végsőig fokozva a számítás gazdaságtalanságát.

Az ötlet bizonyára a “recurrent excitatory connectivity” [127] biológiai fogalomból származik, amelyik valóban (akár az eredeti stimulátor eltávolítása után is) folyamatosan működő aktivitást hoz létre. Az ottani m/s nagyságrendű terjedési sebesség azonban a matematikában végtelen nagy, az elektronikában meg kb. a fénysebesség $2/3$ -ad része. Ennek nem értése vezet az extrém nagy hőtermeléshez, ami alacsony számítási hatásfokkal párosul. Ha viszont a visszacsatolás buszon keresztül valósul meg, akkor az (bár más okból) elegendően lassúvá teheti a jel terjedést (a B.5 függelék szerint az átlagos átviteli idő akár tizedmásodperc nagyságú is lehet!), és megfelelő számú neuron és megfelelő üzenet intenzitás esetén keltheti a biológiaihoz hasonló működés illúzióját (de csak az adott körülmények stabil fenntartása esetén).

B.4 Megvalósítási mód

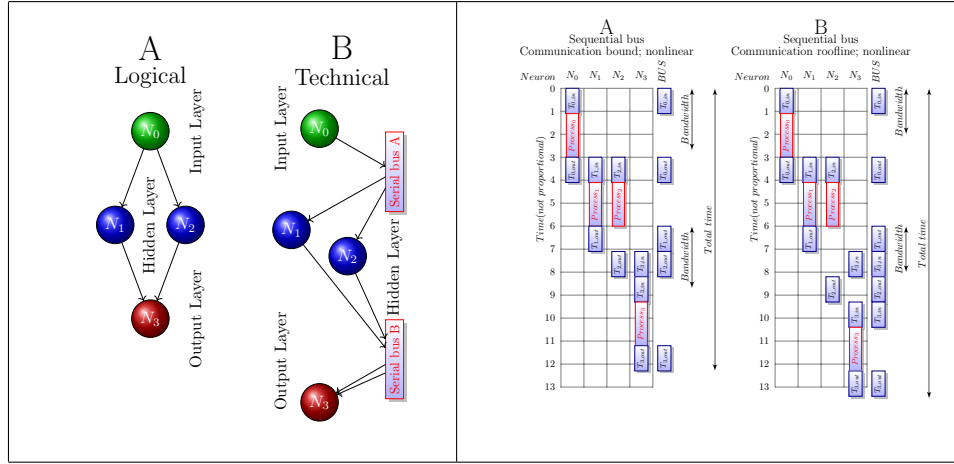
A technikai MI rendszerek teljességgel hagyományos elemeket (vagy általános célú számítógépeken futó programokat) használnak és utasítás-vezérelt eszközök felhasználásával imitálják az adat-vezérelt működésű adatfeldolgozást. Nagyobb MI számítási rendszerek céljára (bár vannak speciális eszközök, pl. akár egész agy-szimulátorok is) általában általános célú nagy számítógépekre készített program rendszereket használnak, a mesterséges

neuronok közötti adatátvitel pedig a számítógépes hálózatok közötti adatátvitel módjára zajlik. A mai elektronikus technológia “homok-áron” tud digitális elektronikus számítógép processzorokat szállítani, amelyek programozásával könnyen lehet neurális hálózatokat építeni. A programozás viszont kifejezetten az utasításvezérelt számítógépek tulajdonsága; ezzel *az ál-MI rendszerek öröklik az utasítás-vezérelt számítások összes hátrányát*. A programozás élömlátás-igénye igen nagy, és a programnyelveken, könyvtárakon, tapasztalatokon keresztül igen erősen kötődik a Neumann-féle szekvenciális végrehajtási módhoz. Még a hangsúlyozottan eltérő (“neuromimetic”) hardver [124] esetén is, a hagyományos hardverre készült szoftver elemek és fogalmak visszahozzák a hagyományos számítógépek működési hátrányainak jelentős részét. Nagy speciális rendszerekben már ugyanott jelentkeznek a korlátok mint a teljességgel hagyományos általános célú rendszerekben [142].

A rendszerek olyan matematikai fogalmakat (mint vektor, mátrix, szorzás) és módszereket (mint gradiens számítás, visszacsatolás) használnak, amelyek egyértelműen az utasítás-vezérelt számítási módból származnak és valójában *gépi tanulást* jelentenek. Egyrészt ennek következtében egyértelművé válik, hogy (legalább részben) *nem az adatok vezérlik a folyamatot*, másrészt *a kétféle működtetési elv keverése önmagában is komoly hatások romláshoz vezet*. Ennek a tudatzavarnak egy másik jele az “MI algoritmus” kifejezés használata: a kapcsolat első fele az adat-vezérelt, a másik fele az utasítás-vezérelt világból származik. Ráadásul ellentmond annak, hogy programozás nélküli adatfeldolgozásról van szó. Valójában *egy sajátos felhasználói interfészen át használunk egy számítógép szoftvert, ami a szokásostól jóval nagyobb mennyiségű adatot használ csak olvasható memóriából*. Ironikus módon, a felhasználói interfész “prompt készítés” formájában követeli vissza a számítógép programozhatóságát, a programozás nélküli adatfeldolgozáshoz.

B.5 Átviteli mód

Egyszerű esetben adatszállítás céljára ezeket a neurális node-okat közvetlenül össze lehet kötni, ilyenkor az érkezési időt kis rendszerek esetében (amíg az érkezési idő a központi órajel hosszánál lényegesen rövidebb) gyakorlatilag a neurális node-ok végrehajtási ideje határozza meg. Bár az utasítások matematikailag kiszámítható eredményt adnak, továbbá az egyes csomópontok működése következtében jól kiszámítható működési késleltetések lépnek fel, a működés zavarba ejtően hasonlít a biológiai rendszerek működésére (idő-viszonyaitól és digitális jellegétől eltekintve, majdnem-valódi MI). A közvetlen vezetékezés azonban csak kis rendszerek létrehozását teszi lehetővé.



Ábra 7.4: A busz használatának hatása a mesterséges neurális hálózat működésére. Bal oldalt: a matematikai és biológiai, valamint a technikai neurális hálózat elemeinek összekötése. Jobb oldalt: a megosztott busz használatának sorba rendező és teljesítőképesség határoló hatása. Saját szerkesztésű ábra.

Emiatt a node-ok általában nem közvetlenül, hanem a számítógép elemeket összekapcsoló vezetékeken (buszon) keresztül kapcsolódnak (lásd 7.4. ábra bal oldal). Azaz, *egy adatot először a busznak kell továbbítani, és a másik node-nak a busztól kell megkapnia*. Az elvi működés (számítástudomány) megegyezik a 7.2 ábrán mutatottal, a valódi működés (számítógéptudomány) jelentősen eltér, komoly tulajdonságbeli eltéréseket vonva maga után.

Ezeket a rendszereket hívjuk ál-MI rendszereknek, mivel a működés időviszonyai közelítőleg sem egyeznek meg a valódi MI rendszerekével. A buszt csak kizárólagos joggal lehet használni, ezért minden adatot küldeni akaró neuronnak “versenyeznie” kell a busz használatáért (arbitráció), minden egyes küldés esetén. Az arbitráció az egyes átviteli időkhöz egy pseudo-random átviteli időt ad, ami a rendszer méretétől függően akár dominánssá is válhat. Ha az egyes rétegek (amint szokásos) ugyanazt a nagy sebességű de egyetlen buszt használják, a különböző rétegek adatforgalmai akadályozhatják egymást, még ha az egyes rétegek számításait külön-külön GPU végzi is. A busz hatását jól jellemzi, hogy a memrisztorokkal megvalósított neurális hálózat [118] energia hatékonysága több mint két nagyságrenddel jobb mint a jelenleg elérhető GPU-k hatékonysága.

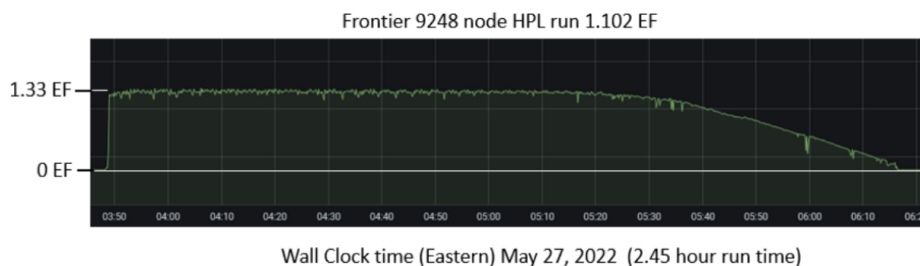
A kizárólagos használat miatt a biológia (és a matematika) esetében párhuzamosan továbbított adatok a technikai megvalósításban szerializálódnak.

nak. A buszon egyidejűleg csak egy adat továbbítható. Ha a buszra sok neuron csatlakozik és egyidejűleg akarnak adatot továbbítani, akkor a busz sávszélessége jelentősen korlátozza az adatok átvitelének idejét (lásd a 7.4. ábra jobb oszlopának A részét). A sávszélesség elérésekor a neuron nem tudja továbbítani az adatát amíg a busz fel nem szabadul (lásd a 7.4. ábra jobb oszlopának B részét). *A busz használata határozatlanná teszi a számítás időablakát*; minél nagyobb a rendszer, annál nagyobb mértékben. “A népszerű közös buszt használni kommunikációs közegként ma már nem fogadható el ötletként”. [55]

Sajátos módon a busz (és azzal együtt a jel továbbítási sebesség) szerepét nem sikerül megérteni a konstruktőröknek. A már említett utasítás- és adat-vezérelt működési mód mellé bevezették a memória-vezérelt működési módot [74], természetesen elméleti megalapozás nélkül. Az elsősorban MI és gépi tanulás, valamint “Big Data” céljaira javasolt “The Machine” újnak vélt hangzatos elve: “számoljunk a memóriában és ne szállítsuk az adatokat a memóriából a processzorokhoz” szintén a hagyományos (impozáns paraméterekkel rendelkező) busz rendszerre épül. A gigantomániás terv (az akkori “digitális univerzum” 250,000-szerese!) “research toys” bemutató programjai impozáns paraméterekkel rendelkeztek, de az arbitráció nem tette lehetővé valódi alkalmazások írását és az “új elvű számítógép” még a kísérleti fázison sem tudott túljutni; éppúgy, mint a memrisztor-alapú számítógép [75]. Nem sikerülhetett elérni a Summit, a Sierra és az Aurora’21 szuperszámítógépek egyike esetében sem az áhított exaFlops határt. A számítógép konstruktőröknek (is) tényleg meg kellene érteni a számítógépes számítás elméletét, tervezzenek bár “big data”, “deep learning”, “machine learning”, “biomorph computing”, “biomimetic computer” vagy “artificial intelligence” alkalmazás számára architektúrát.

A szekvenciális adattovábbítás kedvezőtlen hatásának mértékét mutatja a 7.5. ábra. Az egyszerű (teljes mértékben párhuzamosítható) számítást (HPL benchmark) valamennyi processzor egyszerre fejezi be, és folyomodik a busz kizárólagos használataért. A busz korlátozó hatására a teljes adatátvitel 40 percig tart [73]. Egy MI típusú alkalmazás esetén (a neurális node-ok szinkronizálására) a neuronnak az összes hozzá kapcsolódó neuron számára üzenetet kell küldeni. *Ha a Frontier-t MI típusú feladat futtatására használnák, akkor (neurononként mindössze 600 csatlakozó neuront számítva) 400 órán át tartana a kommunikáció, minden egyes, tipikusan 1 msec hosszúságú neurális számítás után.*

Nagy rendszereken neurális szimuláció a jelenlegi technológiával gyakorlatilag megvalósíthatatlan. A [142] cikk tanúsága szerint az 1 milliárd neuron szimulálására épített szuperszámítógép éppúgy csak a megépített hardver

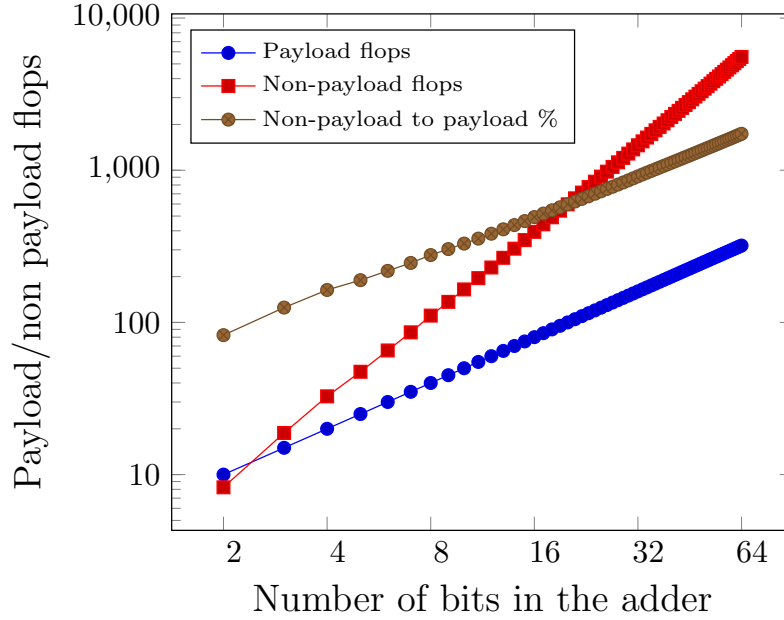


Ábra 7.5: A Frontier szuperszámítógép bemutatásakor készült ernyőkép: a számításal egyidejűleg elkészülő 7M processzor adatátvittele a szekvenciális buszon keresztül 40 percig tart[73].

kb. 1%-át képes használni, mint a speciálisan tervezett SpiNNaker agyszimulátor. A kommunikációs igény lehetetlenné teszi a nagyobb hardver konfiguráció használatát. Az 1% használatát is csak úgy lehet elérni, hogy a “későn érkező” üzeneteket eldobják. Ez azt jelenti, hogy a generált (véltetően jó) állapot információ túlnyomó része nem jut el céljához, ezért az agy magasabb rendű funkcionalitásának vizsgálata reménytelen [142]. A kutatás első kulcs feladatának eme 2018-as megállapítása ellenére, a projekt folytatódott az erre épülő második majd harmadik feladattal, “átszervezve és elkezdve” beszámoló kulcsszavakkal [22]. A politikai akarat akkor is megvan, ha a feladat megoldhatatlan. Meg természetesen a költségvetést is fel kell használni. Ennek alapján hasonló eljárás várható a MI esetében is.

A technikai számításnak törekednie kell arra, hogy a neurális csomópontok ugyanabban az időben adják át állapotukról az információt egymásnak, azaz egyszerre próbálják meg használni a buszt. Nem elég, hogy nagyszámú adatszallítást kell végezni, ráadásul még “csomós” (burst) is az átvitel. A MI számítás kommunikációs igénye főként azért rendkívül hátrányos [113], mert a processzorok be/kiviteli utasításai a számítógép operációs rendszeren keresztül tudnak csak működni (kontextus váltást igényelnek), és ez a művelet kb. 10^4 gépi utasítás végrehajtását követeli [54, 140]. A terhelés növekedésével (nagy számú neurális csomópont jelenléte) a MI rendszer kommunikációs igénye alatt a számítógépes rendszer összeroppanhat [103].

A nagy számítógépeken a neurális szimuláció jellegű feladatokban legfeljebb néhány száz processzort lehet használni. Ez a szám a feladat típusától függően változik. Az a tény, hogy tetszőlegesen nagy számú processzorral (névleges teljesítőképesség) sem lehet túllépni az Amdahl-féle meg gondolásokból számított maximális hasznos számítási teljesítőképességet, évtizedek



Ábra 7.6: A szükséges és szükségtelen billenések száma az összeadó áramkör bit hosszúságának függvényében. Saját szerk. ábra.

óta (helyesen magyarázott) tapasztalati tény [132], sőt az is, hogy *a processzorok számát tovább növelve a névleges teljesítőképesség és a felvett elektromos teljesítmény nő, de a hasznos számítási teljesítőképesség csökken*. Ennek ellenére ritkán végeznek olyan jellegű méréseket, hogy különböző számú processzort használva mérik az elért számítógépes teljesítőképességet egyazon feladat esetén. A ritka kivételek egyike [53, 56] szerint, modern szuperszámítógépen is reprodukálható, hogy a processzor szám függvényében a számítógépes teljesítőképesség maximumot ér el (az idézett munka szerint az általuk használt program esetén mindössze néhány tucat processzort érdemes használni a szuperszámítógépen elérhető több tízezerből), és a processzorok számának további növelése a hasznos teljesítőképesség csökkenéséhez vezet. A *mérnöki* számítógép tudomány ismeri a jelenséget, a *matematikai* számítástudomány nem akarja tudomásul venni, és nem képes megmagyarázni, miért tapasztalnak $O(N^5)$ skálafaktort a számítástudomány által következtetett $O(N^3)$ skálafaktor helyett a rendszer méretének növelésekor.

Szekvenciális busz használata eleve *megszünteti az adatok egyidejű érkezését*, de az igazán kedvezőtlen hatása hogy *az üzenet továbbítás sorrendje nem meghatározott*. A bonyolultabb (sok neuront és több mélyebb

réteget tartalmazó) hálózatok esetén előfordulhat, hogy egy réteg valamely eleme később kapja meg a magasabban fekvő node-tól az adatot, mint a többi elem. Ennek kiküszöbölésére vagy meg kell várni valamennyi neurális input beérkezését (aszinkron számítás) vagy pedig az újabb inputok beérkezésekor a mélyebb réteg neuronjai (és az általa produkált adatok elküldése után a következők) újból elvégzik a tőlük elvárt számítást (megismétlik az egész idő-ablakot), amittől exponenciálisan megnő a kapuk (szükségtelen) billenéseinek száma (lásd 7.6 ábra), ezáltal rendkívüli mértékben megnő a neurális hálózat teljesítmény felvétele és tovább csökken a hasznosított számítógépes műveletek aránya (a hatások).

B.6 Gyorsítás

Neumann alapvető munkájában kifejezetten hangsúlyozta, hogy a láncolt műveletek végrehajtásakor az adatok átviteli ideje nem hanyagolható el (azaz a teljes végrehajtási idő az adat átviteli és feldolgozási idejét is tartalmazza), azonban a számítástudományban használt absztrakt processzorban az adatok átviteli idejét nullának tekintik (ez a klasszikus fizika azonnali elektromágneses kölcsönhatásának felel meg). Ez a gyakorlatban azt jelenti, hogy – a központi órajellel rendszerek esetében – egy adott mennyiségű utasítás végrehajtásához szükséges időt állandónak tekintik, attól függetlenül, hogy milyen típusú feladatot (workload) old meg a számítógép. A tapasztalat ennek ellentmond: a várakozáshoz képest nagyságrendileg eltér a végrehajtási idő skálázása [53, 56], különösen MI típusú alkalmazás futtatása esetén (bár az ott alkalmazott – számítási szempontból inkorrekt – megoldások ezt igyeksenek elfedni).

Miért kell gyorsítani Először a nagyon nagy (szuper) számítógépek működésekor tűnt fel [132], hogy a hasznos teljesítőképesség nem emelkedik lineárisan a névleges teljesítőképesség (a processzorok számának) növelésével, hanem telítésbe fut. A MI típusú feladatok esetén is létezik a jelenség, de a “nagyon nagy” mindössze néhány tucatot jelent [85].

A fentiek szerint a MI nagy számú számítási elemének intenzív kommunikációja (különösen a hagyományos számítógép architektúrák elemeit használva) igen lassúvá teszi a működést. A jelenség oka, hogy (a már Neumann által leírt módon) az adatok feldolgozása és szállítása kölcsönösen akadályozzák egymást: a másik műveletnek várnia kell az adat megjelenésére illetve feldolgozására. A lassúság az elvi működéshez hozzátartozik. A számítógépes technológia különböző pufferek, várakozó sorok, stb. alkalmazásával leplezni igyekszik a jelenséget. Hogy gyorsítani tudjunk, módsze-

reket kell kitalálni a kommunikáció és a számítás egy részének elhagyására, de ki kell találni, hogy melyik üzenet küldéseket és számító elemeket lehet elhagyni.

A központi órajel kezelésnél megszoktuk a gondolatot, hogy a megfelelően megválasztott idő-ablak alatt az összes számítás befejeződik és továbbléphetünk a következő elemi utasításra. A MI esetében az említett okokból igen nagy (és a rendszer méretétől függő nagyságú) *idő-ablakot* kell választanunk. Ha az idő-ablak kisebb, mint amennyi idő alatt az összes üzenet eljut a címzetthez, az két következménnyel jár. Az egyik, hogy a címzett nem kapja meg az összes információt ami a hibátlan működéséhez szükséges. A másik, hogy az információ olyan időpontban jut el a címzetthez, amikor annak tartalma már nem érvényes (az adott időpontban a neuron a társ-neuron egy korábbi állapotát veszi figyelembe az aktuális helyett). Ezért a hiányzó vagy aktualitását veszített információ fontosságától függő mértékben a neuron részben hibás eredményt fog adni (mivel a buszon az összes neuronnak szóló üzenetek késleltetődnek, akár az összes neuron is hibás eredményt adhat). Az üzenetek a pufferekben halmozódnak (a pufferek megteltéig) és az idő előre haladtával egyre irreálisabbá válik a számítás. Továbbá a számítás leállítása után a pufferek kiürülnek (a társ-neuronok látszólag változatlanul szállítják az idejét múlt információt): a MI rendszer “hallucinál” (ez elég meglepő módon a Los Alamos National Laboratory [5] és az IEEE [78] hivatalos álláspontjának tűnik, sőt tudományos közlemény is van róla [64]. Nem tévesztendő össze a kognitív tudományok által említett [63] hallucinációval.). Ha észszerű futási időt akarunk biztosítani, akkor választhatjuk a pufferek periodikus kiürítését minden időrés után [142]. Ekkor csak az első típusú hiba (adatvesztés) lép fel, de a rendszer ekkor is hibásan működik. A helyes működéshez sem hibás időben feldolgozott, sem hiányzó adattal nem lehet eljutni. Az idő-ablak fontos. Ennek megértése nélkül a ChatGPT (és az egész MI) “fekete doboz” [105], és annak tanulmányozása (természetesen a számítás elméletének és megvalósítási technológiája működési elvének megértése nélkül) valóban olyan, mintha valamilyen földön kívüli intelligencia [63] lenne. Minél nagyobb rendszereket építenek [108], annál inkább.

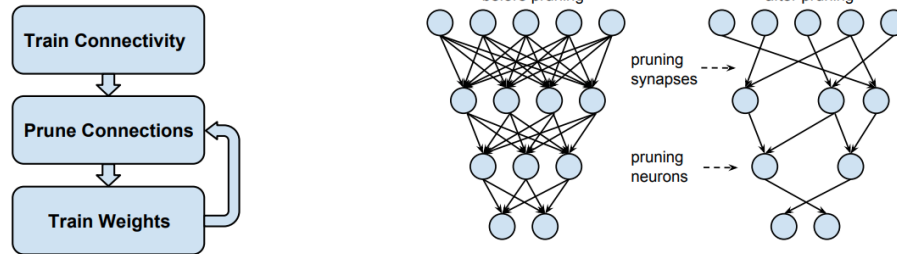
A gyorsítás módja A működőképesség biztosítására szolgálnak a különféle működés gyorsító megoldások. Az egyik lehetőség a neuronok által előállított adatok egy részének (már említett) eldobása (részben az idő-ablak szélességének önkényes változtatásával), egy másik szintén említett módszer hogy logikailag csak egymás után végezhető műveleteket független számoló egységekkel egyidejűleg végzünk, megint egy másik a neuronok

közötti összekötések és/vagy az általuk küldött üzenetek számának csökkentése. Ezek az ötletek az aktuális helyzettől függően működhetnek vagy nem működnek. Rábízhatjuk a sorsra: amelyik üzenet a várakozó sor végére került vagy nem fér bele a pufferbe, elvész; ettől a rendszer látszólagosan gyorsabban működik, de információt veszít(het). Próbálhatjuk a biológia módszerét imitálni: egyszerre csak a node-ok egy része aktív, vagy csak akkor küld információt, ha jelentős változás következett be állapotában.

A biológia jól él az utóbbi lehetőséggel: a neuronok közötti kapcsolatok száma jellemzően több ezer, és egy adott feladat megoldásában jellemzően csupán tucatnyi neuron használatos. Ez a tucatnyi neuron azonban feladatról feladatra változó összetételű (és változó számú) neuronból áll. Továbbá a megcélzott neuronokat a küldő először (viszonylag nagy energia ráfordítással) aktiválja [81] (azaz, idő-multiplexelést is használ), majd (alacsony energia felhasználással) csupán fenntartja a kommunikációt. Ennek utánzására a technológia és a matematika egyaránt képtelenek.

A mesterséges neurális hálókból a betanítás kezdetén eléggé szokásos véletlen súlyozású szinaptikus bemeneteket használni, továbbá a visszacsatolás jellegű bemenetek értékei és súlyai az első számítás során definiálatlanok. Maga a számítás lényegében iteráció jellegű, ahol – megfelelő szinkronizálás nélkül – *a legtöbb szereplő a legtöbb esetben hibás bemenő adattal dolgozik. Kezdetben bizonyosan hibás értékekhez adjusztálja súlyait, később pedig csak jelentős időbeli (avagy több számítási ciklus) késéssel.* A helyzetet súlyosbítja a tanuláshoz használt “deep learning” eszközök energia pazarlása és katasztrofális hatásfoka, lásd B.2 szakasz. Szerencsés esetben (vegyük figyelembe hogy komplex rendszerek esetén nem-stabil állapotokkal dolgozunk továbbá a rendszernek több helyi stabil állapota is létezhet) a számítás konvergál de kínosan lassan. Vagy egyáltalán nem. Mindennek oka, hogy *a MI nem tekinti a számítás időbeli viselkedését.*

A gyorsítás triviális módja pl. a számításban használt elemek számának csökkentése (ritkítás). Ennek gyakorlati kivitelezése során (7.7 ábra bal oldala) iteratív módon meghatározzuk a kevésbé használt csomópontokat és összekötéseket, ezzel kevesebb elemet tartalmazó új neurális hálót hozunk létre (7.7 ábra jobb oldala), ami már értelemszerűen kevesebb számítást igényel és ezért gyorsabban működik. A változtatás után újra kell tanítani a hálózatot. A gyorsaság növekedésével a pontosság csökken, természetesen. A nélkülözhető elemek száma a gyorsaság és a pontosság közötti kompromisszum eredménye és csak az adott betanító adat készlet használatára érvényes. A ritkításnak elméleti alapot is megpróbáltak teremteni (áttekintés [61]-ben), amelynek keretében hangzatos modelleket hoznak létre. Ezek közös célja a számítási vagy a kommunikációs igény csökkentése, és két-



Ábra 7.7: A neurális hálózat “ritkítása” (pruning). Eltávolíthatunk vagy teljes neuronokat vagy a közöttük levő összeköttetések egy részét.[66]

séges elhanyagolásokat tesznek, amelyek némely helyzetben működnek, demonstráció szintjén.

Természetesen ugyanezt szoftveresen és elosztott rendszerereken is el lehet érni. Az egyik népszerű ilyen keretrendszer [72] önkényesen és indoklás nélkül egyetlen szomszédra redukálja az információcserét és ilyen módon egy N csomópontot tartalmazó rendszerben az $N * (N - 1)$ szükséges üzenetet $2 * (N - 1)$ ilyen módon $O(\frac{N}{2})$ mértékben redukálja. Egy ciklus futási ideje tényleg rövidebb, de a működéshez szükséges üzenetek nélkül a konvergencia lassúbb (több ciklus szükséges) és végső soron az eredmény lassabban áll elő és a hiányzó (el nem küldött) üzenetek miatt hibás is.

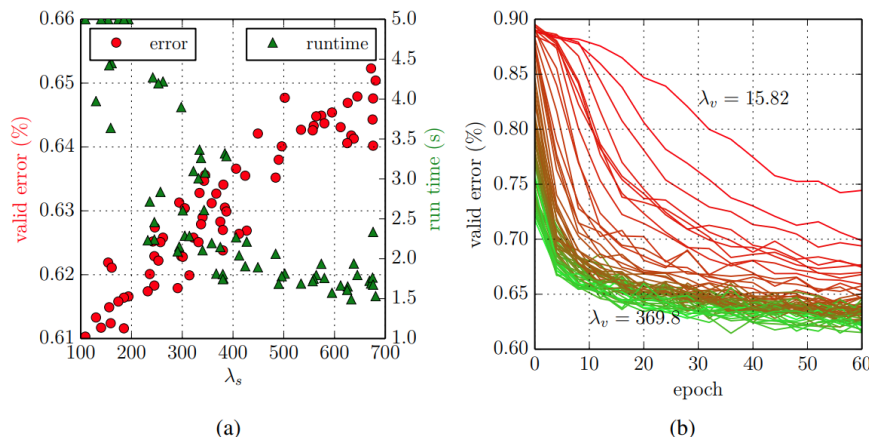
Mivel a MI-nek nincs elmélete, ezért a különböző technikai megoldások “ad-hoc” jellegűek, és egy bizonyos feltételezett scenárióra vonatkoznak. Hogy a felhasználó által tekintett használati helyzet megegyezzek az eszköz fejlesztésekor szem előtt tartott scenárióval, ritkaság. A megoldás részletei céges titoknak számítanak, amit a felhasználóval természetesen nem közölnek. Mivel a tanítás igen költséges folyamat, a rendszerek elő-tanított chipeket használnak, és az elő-tanítás során a legnagyobb felhasználók igényeit veszik figyelembe (pl. éppen a ChatGPT által is használt szöveg feldolgozást, a leggyakoribb szöveges adatkészletekkel). *Semmi garancia arra, hogy az így betanított rendszer más adat készlet esetén is nélkülözni tudja az eltávolított elemeket* (azaz, hogy a nagy keresőmotorok számára angol szöveggészleten elő-tanított rendszer hatékonyan fog működni egy magyar-latin-angol nyelvű orvosi diagnózis vagy egy magyar jogi terminológiát használó tárgyalási jegyzőkönyv feldolgozása során; nem is említve a magyar nyelv elő- és utótagokat használó, ragozó, stb. jellegét).

A transzformatív működés újabb lehetőségeket hoz játékba. Mivel a

rendszer több nyelvi szinten építi fel nyelvi moduljait, a működés elvének megfelelően több változatban is (lásd D.1 függelék), amit aztán egymással összehangolva a felhasználó tudomására hoz, lehetőség nyílik arra, hogy a neurális számítás elemeit előre eltárolja és ezzel a működést látszólagosan felgyorsítsa. Itt is nagyon hasznos a használati valószínűség ismerete. A tipikus felhasználó számára a rendszer a közeli pufferekben tárolt félkész válaszokból készít “ez a válasz erre a kérdésre jó szokott lenni” elven szinte azonnal látható eredményt. A sokak által feltett, vagy valamilyen esemény kapcsán időszerű kérdésekre gyakorlatilag azonnal érkezik a válasz; furcsa, ritkán elhangzó, szándékosan érthetetlenül megfogalmazott, hiányzó ismeretre vonatkozó kérdések esetén pedig nagyon lelassul és szaggatottá válik a működés.

Bár a rendszer előre tanított (azaz elvileg menet közben nem tanul és nyilvános dokumentációja szerint nem “emlékszik” a társalgásban feltett előző kérdésekre és az adott válaszokra), igyekszik kontextust építeni a felhasználóval. Az adatkészlet elavultságán internetes kereséssel is igyekezhet javítani; a keresési gyakoriság és a felhasználói vélemények tovább alakíthatják a betanításkor beállított eloszlásokat. A D.1 függelék szerint – *dokumentációjának állításával ellentétben* – emlékszik a felhasználóra, és következő válaszában a felhasználó kérdéseinek irányultsága alapján hangolja válaszait (*igyekszik azt mondani, amit a felhasználó hallani szeretne; végülis az elsődleges felhasználói piac az üzletszerzés segítése*). A D.2 függelék szerint nem akar tájékoztatatlannak látszani, ezért – *dokumentációjának állításával ellentétben* – határozott, magabiztos választ ad akkor is, amikor nem tudja miről van szó vagy nincs elég ideje egy teljesebb válasz összeállításához (végülis a valószínűségi eloszlások ekkor is adnak valamilyen értéket), de ismételt kérdésre (ha gyanakvást érez, vagy a felhasználót tájékozottnak véli) elbizonytalanodik. Mint [34] megfogalmazza, az érdekelt cégek versengése (és titkolózása) egyfajta ‘fegyverkezési versenyt’ indított el, a(z elvileg akár) milliárdnyi számítás a ‘fekete doboz’ban még a szakértők számára is átláthatatlan; amit tetéz a szándékos matematikai véletlen és a szándéktalan technikai esetlegesség. Orosz rulett, ha komolyan akarjuk használni.

A gyorsítás ára A [40] kutatási eredmény szerint “az ilyen hálózatok tanítása kihívást jelentő optimalizálási probléma marad. A fennmaradó problémák: a nagyon hosszú betanítási idő (bizonyos problémák esetén több hét a modern számítógépeken), a túl-illesztés veszélye (hogy a megtanult funkció túlságosan specifikus a betanításhoz használt adatokra és rosszul általánosítható a még nem látott adatokra, technikai megfogalmazással,



Ábra 7.8: A futási idő és a pontosság összefüggése. (a) az egyes kör és háromszög párok az egyes λ_s paraméterekhez tartozó hibát és a futási időt jelentik. λ_s növelésével ugyan csökken a futási idő, de romlik a megbízhatóság. (b) Ugyanaz a modell különböző λ_v értékekkel tanítva. A diagram vonalak színe a paraméter értékével változik. ([40] 5. sz. ábrája)

a 'vanishing gradient' problémája") [102]. "Kevesebb egység aktiválásának közvetlen hatása hogy a hálózatban felgyorsul az információ terjedési sebessége, betanítás és tesztelés során is." Ez azt is jelenti, hogy a kiszámított visszacsatolás, ami esetleg nem-definiált bemeneteken alapszik, hamarabb eléri az előző rétegeket. Ennek természetes következménye, hogy "ahogyan a λ_s csökken, úgy a futási idő is csökken, de ugyanakkor a teljesítőképesség is romlik." (λ_s az előrejelzés pontosságát és számítási költségét, λ_v pedig a probabilisztikus jellegét és a jobban bemenet-függő jelleget jellemző paraméter; lásd 7.8. ábra).

Mint több helyütt rámutattunk, a hibás elvű gyorsítás vagy lassuláshoz, vagy hibás működéshez vezet. Minden gyorsításnak ára van, de a kereskedelembe kapható gyorsított ANN-ek esetében esetében nem lehet tudni, hogy milyen és mekkora ára. Kritikus feladatokra csak olyan MI berendezést szabad használni, amelynek működését pontosan ismerni lehet, kiszámítható és determinisztikus. Ami ellentétes a működés elvével és a céges érdekekkel.

B.7 Tárolási mód

Neumann klasszikus modelljének egy másik alapfeltevése, hogy kevés adaton sok számítást végzünk. Az adatok tárolására szolgáló memória kezdetektől fogva kritikus sebességű elem volt, és a processzor végrehajtási sebességének növekedésével egyre inkább azzá vált. Már Neumann felhívta a figyelmet arra, hogy a memóriának alkalmazkodnia kell az adatok használati gyakoriságához; ennek megfelelően többszintű (sebességű, méretű és költségű) memóriára (gyorsítótárra) van szükség. A memória szintek megvalósításának fajlagos ára és működésének sebessége a processzortól való távolsággal csökken. A hardver az előreláthatóan szükséges memória tartalmakat a használat gyakorisága alapján egyre közelebbi memóriákba másolja.

A számítógépben az adatok eljuttatásának relatív ideje az ún. buszon keresztül egyik komponensből egy másikba folyamatosan nőtt, ahogyan az utasítás végrehajtási sebessége nőtt. A megnövekedett diszperzió miatt jelentősen megnőtt a processzorok elektromos teljesítmény vesztesége (hőtermelése). A gyorsítótárak akkor jelentek meg (lásd 7.1 ábra), amikor használatuk nélkül megugrott a diszperzió. Megjelenésükkel időlegesen csökkentették a diszperziót. A végrehajtási sebesség növekedésével többszintű gyorsítótárak (többszöri adat másolás igényel!) jelentek meg. A másolás elektromos állapotváltozással jár és energia igényes, lásd 5.4 szakasz.

A gyorsítótár(ak) hatása az adatok elérési idejének látszólagos csökkenése, azaz látszólagosan gyorsabb számítási működés. A gyorsítótárak tudatos használatával a feladat megoldásának idejét akár nagyságrenddel is csökkenthetjük [123]. Nagy mennyiségű adat (az adatmennyiség jelentősen túllépi a rendelkezésre álló gyorsítótár méretét) feldolgozása (“Big Data”) esetén gyakorlatilag eltűnik a gyorsítótár hatása. Agy-szerű rendszerek létrehozásának az egyik fő akadálya, hogy a mérnöki rendszerek digitális adattárolókat használnak, míg a biológia speciális mechanizmusokkal (folyamatokkal) valósítja meg az információ rövid és hosszú távú elérését [95], azaz a biológiai folyamatok lényegében az időt használják információ (és nem digitális adatok!) tárolására.

A mérnöki törvényszerűségek természetesen a neurális hálózatokra is érvényesek, de itt is ritkán végeznek hasonló jellegű méréseket. A [85] mérés szerint a kereskedelembe kapható mély neurális hálózatok teljesítőképessége néhány tucat “neuron” esetén telítésbe fut, további node-ok hozzáadása az elektromos teljesítmény felvételt növeli, de nagyobb számítási teljesítőképességre a rendszer nem lesz képes. Több mély réteg hozzáadása erősen csökkenti a számítási teljesítőképességet. *Komoly tervezést (megvalósíthatósági tanulmányt) és teljesítmény ellenőrzést igényel ilyen rend-*

szerek alkalmazása, különösen tömegesen. Lásd még a 5. szakaszt.

B.8 Intelligencia előállítása

A fentiek miatt mind az elektronikai, mind a szoftveres módon megvalósított működés a *gépi tanulás* fogalomkörébe tartozik; elemi működési módjuk és ezért a felhasználásukkal létrehozott rendszerek viselkedése is eltér a biológiai neurális hálózatokétól. *A magas szintű agyi funkciók (pl. absztrakció, általánosítás) terén a (mély-réteg alapú) gépi tanulásnak komoly funkcionális és teljesítőképesség korlátai vannak* [52, 95]; arról nem is beszélve, hogy a biológiai agy teljesen másképpen működik [119].

Vegyük észre azt a finom különbséget, hogy *a biológiai működés során az utasítások analóg módon viselkednek, azaz ugyanazon adatokra a környezet (azaz az agy) által meghatározottan (időben is) végtelen változatossággal tudnak reagálni, míg a technikai neuronok lényegében matematikai függvényeket implementálnak és reagálásuk előre meghatározott.* Az intelligenciát a biológiai számolás legalacsonyabb szintjén valahol eme technikai különbség közelében kell keresnünk. Jól vagy rosszul, de semmiképpen sem *csak* előre kiszámíthatóan (és nem is probabilisztikusan) reagálunk. A *betanított* munkás nem *tanulja meg* hogy mit csinál, csupán elvégzi a bemutatott tevékenységet.

C A MI és TI működésének összehasonlítása

A tudományban számos példát találunk az inkrementális és az ugrás-szerű tanulásra. Például, alapvetően jól tudjuk az elektromágneses sugárzásra vonatkozó törvényeket és inkrementálisan sugárzási törvényeket tudunk alkotni valamint azokat extrapolálni, de csak ugrás-szerű új feltételezéssel (energia kvantum létezése) tudjuk a természet által ténylegesen mutatott viselkedést leírni. Hasonlóképpen csak ugrás-szerű tanulással tudjuk felismerni a tér és idő összefüggő viselkedését.

Az ún. fizikai gondolkodáson [137] lemérhetjük, hogy a MI működése mennyire hasonlít a TI működésére. A fizika mérésen alapszik, és a mért mennyiségekhez (adott mértékegység mellett) mérőszámot rendel, ami egy numerikus paraméter értéknek felel meg, amit már (absztrakt mennyiségként) matematikailag kezelni tudunk, ami főként a MI matematikai módszerei számára lényeges. Ezzel a leképezéssel azonban óvatosan kell bánnunk. Einsteint idézve: „Amennyiben a matematika törvényei a valóságra vonatkoznak, nem bizonyosak, amennyiben bizonyosak, nem a valóságra vonatkoznak.” (A MI tudomány kezelésével kapcsolatban különösen óvatosságra kell

lenni, l. D.5 függelék.)

A mérés (vagy megfigyelés) tehát *adatokat* hoz létre. Ilyen értelemben tehát a TI és a MI párhuzamba állítható: mind a TI mind a MI *adatokon alapszik és adatfeldolgozást végez* a rendelkezésre álló feldolgozó egység használatával. Számítási terminológiával, *az adatok vezérlik saját feldolgozásukat*, és a feldolgozást végrehajtó egység utasításai közül választják ki és adjusztálják a feladathoz az utasításokat.

A TI esetén a megismerést három fázisra oszthatjuk, aminek valamiféle megfelelőjét a MI (számítógépes tanulással megvalósított) esetében is megtaláljuk. Ugyanakkor a párhuzam csak látszólagos: a technikai hasonlóság mellett jelentős különbségek is vannak, amelyek indokolják a MI “intelligencia” jellegének kétségbe vonását. A kétféle “intelligencia” megfeleltetése egymásnak komoly értelmezési nehézségeket okoz [84], lásd még D.1 függelék. A felismeréshez (a hasonlóság mértékének megítéléséhez) valamiféle mérőfüggvényt használ mindkét fajta intelligencia. Egy 2017-ben végzett kísérlet [96] (lásd D.2 függelék) azt próbálta meg kideríteni, hogy egy egyszerű neurális háló az “egyenlőség” függvény kapcsolat felismerése során mennyire képes absztrakcióra: a betanítás során az adatok helyesek voltak, de a készletben csak páros számok szerepeltek. Ennek alapján a rendszer a páros számok esetén helyesen inter- és extrapolálta az egyenlőséget, de hibázott páratlan szám esetén. A magyarázat: a használt elektronika számára a hasonlóságot az “1” értékű bitek száma jelenti. Az említett példa esetén a felismerés a “15” szám esetén hibás. Idézzük fel, hogy a “16” számban az “1” értékű bitek száma 1, a “14” számban pedig 4, ezért a “15” számban levő 3 “1” értékű bit közelebb áll a 14-hez, mint a 16-hoz. Az egyberi agy nyilván interpolálna vagy általánosítaná, hogy megegyezik az eredmény az argumentummal.

Ilyen jellegű kérdést nem sokan tesznek fel, azaz a kérdésre adott válasz nincs előkészítve a pufferekben, ezért újonnan kell a választ összeállítani. Bár időközben nyilván fejlesztettek az algoritmuson (mint minden nyilvánosságra hozott hiba esetén), így a 2023-ban végzett kísérletben már megállapítja a rendszer, hogy “megegyezik az eredmény az argumentummal”, de nem tud általánosítani, hogy akkor az páratlan számok esetén is érvényes. Az első válasz, hogy felismerte ugyan a szabályszerűséget, de csak páros számok esetén tudja alkalmazni, mert a betanító adatok között csak páros számok szerepeltek. Hogy általánosítsa a szabályszerűséget, több időre lenne szüksége; a válaszadáshoz rendelkezésre álló idő azonban letelt: az kell eredményül megadni, amit idáig sikerült összeállítani. A válasz kiszámítása azonban elindult és idővel a helyes(ebb) válasz is bekerül a megfelelő pufferekbe, de onnét mindaddig nem kerül elő, amíg rá nem kérdez valaki. Ha ismét feltesszük ugyanazt a kérdést, akkor már a pufferekben szerepelnek az előre gyártott

elemek, és a rendszer nem ismétli meg a 2017-ben és az előző kérdésre adott válaszban már elkövetett hibát, és a válasz helyes lesz. A rendszer ugyanannyit és ugyanúgy “gondolkodott”, csak a felhasználóhoz közeli pufferek tartalma változott.

C.1 A jellemző paraméterek megtalálása

A TI megismerési folyamata során különböző (általunk beállított és annak következtében létrejött) adatokat használunk, vagy csupán megfigyeljük a bolygó pályáját az idő függvényében vagy a nyomást a hőmérséklet függvényében, stb. Abban, hogy csupán ezeket az adatokat figyeljük, benne van az a tapasztalatunk, hogy a többi szóba jöhető mérhető mennyiség nem szól bele az adatok közötti összefüggésbe, ezért azokat figyelmen kívül hagyjuk. “A megfelelő mennyiségek megtalálása már önmagában is meglepő mélységű következtetések levonását teszi lehetővé.” [137] Ez a tevékenységünk *intuitív*: személyes vagy történeti tapasztalatot tükröz. Azaz, megelőzi egy olyan megfigyelés sorozat, aminek eredményeként el tudjuk választani a lényeges és lényegtelen paramétereket. Emiatt a gondolkodáshoz (lényeg kereséshez) használt paraméterek száma pontosan megegyezik a szükséges lényeges paraméterek számával. MI terminológiával: valamennyi paraméter “labelled”.

A MI használatának egyik indoka, hogy nem ismerjük pontosan az adatok közötti összefüggést, esetleg a lényeges/lényegtelen adatok megkülönböztetését sem, vagy túlságosan bonyolult lenne az összefüggés leírása és/vagy költséges annak kiszámítása. Netán egyszerűen lusták vagyunk vagy divatos/trendi a MI használata. Ilyenkor a berendezés felhasználója bocsátja rendelkezésre az adatokat. Mivel a MI esetén a feltételezés az, hogy majd a módszer kiválogatja a kívánt válasz megadásához szükségeseket közülük, általában jóval több (“unlabelled”) paramétert jelöl meg jellemzőként, mint amennyi lényeges és szükséges lenne. Az emberi intelligencia ritkán vállalja a felelősséget, hogy adatokat figyelmen kívül hagyjon, így az MI módszer általában a lényegtelen paraméterekre is összefüggéseket keres. Ami egyrészt idővesztéssel jár, másrészt növeli az ál-megoldások találasának veszélyét és a valódi megoldások értékének bizonytalanságát továbbá megnehezíti annak eldöntését, hogy a felállított modell megfelelő-e. *A jellemző paraméterek kijelölése tehát a MI esetében is fontos szerephez jut és a TI bevonását teszi szükségessé.* A fizikai gondolkodás egyértelműen determinisztikus modellt (lásd 2.1 egyenlet) használ; éppúgy, mint a természettudományokban használt “MI módszerek” döntő többsége.

A leendő következtetés (akár helyes a modell, akár nem) magvai már

benne vannak abban, ahogyan a modellt paraméterezzük, és ahogyan a paramétereket matematikailag egymással összefüggésbe hozzuk. A MI ismert hibáinak (lásd 3. szakasz) egy része a helytelen paraméter kiválasztásból következik, a modelltől függetlenül. A hibás paraméterezés egyaránt oka lehet a téves felismerésnek (lásd 3.1.2 szakasz és D.2 függelék): nem elegendő számú paraméter vagy nem lényeges paraméterek használata; a rejtett összefüggések megtalálásának (lásd 3.1.3 szakasz): a kiválasztott paraméterek egymással összefüggésben vannak, vagy akár egy másik, lényegtelennek ítélt paraméteren keresztül kapcsolódnak. Ha valami lényegtelen paramétert az MI fontosnak ítél az egyik feladat alapján, a következő feladat tanulásakor jelentéktelennek minősítheti azt (lásd 3.1.4 szakasz), és ezzel együtt a megtanult információt. *A MI (gépi tanulás) sikerrel valósíthatja meg az inkrementális intuíciót, de az ugrás-szerű intuíció hiánya látványos kudarcokhoz vezethet.*

A fizikai gondolkodás modellezése során mérésre visszavezethető adatok szerepelnek, amelyeknek számokra történő leképezése egyértelmű; ezért ilyen összefüggések keresésekor könnyen találhatunk a MI számára alkalmas adatokat, amelyekből a MI összefüggéseket tud származtatni. Még ilyenkor is okozhat gondot a redundáns vagy nem megfelelő adat mérése. Pl. ha tömeg és sebesség szerepelnek a mért mennyiségek listáján, de a jelenség az impulzustól függ, akkor a MI “rejtett összefüggést” fedez fel a rendelkezésére bocsátott paraméterek között.

Nem természettudományos jellegű modellezés esetén *sokszor nem lehetséges egyetlen mérőszámmal (paraméterrel) jellemezni a kritikus jellemzőt.* Ennek megfelelően a MI is úgy találja, hogy több összefüggő paraméter valamilyen együttes használata eredményezi a szükséges függést; azaz *ezeknek a paramétereknek együttes, a többiektől elkülönített kezelésére lenne szükség.* Erre azonban a MI technikája és matematikája nem adnak lehetőséget, ennek megfelelően a MI ilyen adatok feldolgozására csak korlátozottan alkalmas. Az ilyen alkalmazások többnyire a 2.3 egyenleten alapulnak és tapasztalati összefüggéseket, eloszlásokat használnak. Ennek jelei a körülményektől függő mértékben nyilvánulnak meg.

C.2 Lényeg keresés, modellezés

A TI számára “a modell készítés művészet” [137]; a korábban látott adatokat és jelenségeket akarjuk összefoglalni, egységes szerkezetbe rendezni, de nincs rá semmilyen szabály. A kutatói intuíciót segítheti korábbi tapasztalat, egy új jelenség ami a régi modellel nem magyarázható, vagy egy paraméter változó szerepének felismerése. Azaz, a korábbi tanulás egy leszűrt változatát

használjuk (a MI terminológiájával, egyfajta pre-training-et jelent: még nem ismerjük a megoldandó feladatot, de a megoldás módszerében már szereztünk gyakorlatot). A “betanítás” fázisa tehát még a modellezés előtt lezajlik, bár a lényeg keresgélése közben gyakran iterálunk: jobb összefüggést keresünk vagy változtatjuk a paraméterek lényeges/lényegtelen szerepkörét, számát. A TI számára rendkívül fontos, hogy “a modellnek egyrészt a lehető legegyszerűbbnek kell lennie, mert különben kezelhetetlenné válik, másrészt ha valamilyen lényeges tényezőt elhanyagolunk vagy lényegtelennek fontosnak vélünk, hibás eredményre jutunk.” [137]. Ugyanez igaz a MI esetére is: automatikus kiválasztás esetén a *túl sok lényegesnek tekintett paraméter félrevezető lehet és elfedheti hogy lényeges paramétereket nem adtunk meg.*

A természettudományi megismerés általában matematikai formulákat használ az adatok közötti összefüggések leírására. Bár agyunk (TI) nem feltétlenül ilyen egyenleteket old meg pl. annak eldöntésére, hogy általunk jól ismert gyaloglási sebességünk, valamint a közeledő gépkocsi távolságának és sebességének ismeretében átmenjünk-e az úttesten, a TI az említett adatok ismeretében dönteni tud erről, korábbi tapasztalatainkra támaszkodó modellünk alapján. Azaz, az említett paraméterek alapján az adat-vezérelt módon működő neurális hálózatunk jó választ ad (és annak alapján nem érdemes agyunk differenciál-egyenlet megoldó képességéről becslést készíteni). Ha ehhez egy új paramétert adunk, pl. hogy milyen magasságból tudunk a járdáról sérülés nélkül lelépni az átkeléshez, akkor *az új paraméter megjelenése más adat-utakat mozgósít és az erre vonatkozó adat más, eddig nem használt utasításokhoz jut el; az addigi megtanult összefüggéseknek nem kell megváltozniuk.* A TI szétválasztja az adatokat (paramétereket) felhasználásuk jellege szerint (más neurális pályákat használ, azaz képes nem-inkrementális tanulásra), a MI viszont mindegyiket egyenrangúnak tekinti és mindkét feladat esetén ugyanannyi paramétert használ.

Az MI esetében a lényeg keresés valamilyen kész modell használatán alapszik, ahol a “modell” valamilyen “fekete doboz” [105] (belső súlyai alapján a bemeneti adataiból számítja ki a kimeneti adatokat), összetett működésű hardver vagy szoftver. Ez a fekete doboz a külvilággal adott számú bemeneten és kimeneten át érintkezik, és (kereskedelmi berendezések esetében) működésének belső részletei ipari titkot képeznek. Nem feltétlenül egyezik meg a berendezés bemeneteinek/kimeneteinek száma sem a berendezésnek adott bemenő adatok vagy előállított kimenő adatok számával. Az első feladat esetében a betanításhoz használjuk az összes elérhető bemenetet, és valamennyi adat/paraméter egyenrangú.

Mivel az első feladat megoldása során a magasság különbség paramétert is fontosnak véli a MI, azt is használja a működéshez és beállítja a megfelelő

súlyokat a feladat megoldásához. Ennek súlya (mint nem lényeges) nem irányítódik megfelelően, de szerepet kap a feladat megoldásában és a többi súly ennek a hamis szerepnek megfelelően állítódik be. Amikor az új feladat (átkelés a járdáról történő lelépéssel) *betanulása* történik, akkor a MI rájön, hogy immár a magasság különbség paraméter is lényeges, ezért megváltoztatja annak súlyát. Mivel azonban a többi súly ennek a(z előző feladatban) lényegtelen paraméternek megfelelően állt be, az új feladat *betanulásával* ez a paraméter szerepet kap, és az összes többi paraméternek meg kell változnia; így érthető, ha a korábbi feladat megoldására már nem képes a rendszer. A TI hiánya, azaz az adott berendezés által szolgáltatott összes bemenet használata és az összes elérhető adat egyenrangúnak tekintése (egy mátrix vagy vektor minden eleme egyenrangú) a katasztrofális felejtés egyik fő oka.

A MI számára a modell készítés mágia: valamit következtetni kell azokból az adatokból, amik rendelkezésre állnak. Az erre a célra rendelkezésre álló eszköz a berendezésben rejtetten elhelyezkedő súlyok rendszere, amit a berendezés autonóm módon állít elő vagy elő-tanítás során készen kap. *Az eszköznek* (és a mögöttes matematikai módszernek) *nincs módja lényeges/lényegtelen paramétereket vagy a közöttük levő összefüggést feltételezni*. Azaz lényegében összefüggéstelennek tekintett adatok halmazát bízunk a MI-ra, az egyetlen útmutatással, hogy “ahogy tetszik gondolni”. Mivel a berendezés működését (a betanítás során használt) adatok határozzák meg, ha jól határoztuk meg a lényeges paramétereket és nem adtunk meg lényegtelen paramétereket, akkor az eredmény jó is lehet. A különböző feladatokban azonban különbözik a paraméterek lényeges/lényegtelen jellemzője, továbbá a matematikai módszerek minden feladat esetén az összes elérhető súlyt módosítják. Ha a feladatok jellege és az adatok olyanok, akkor a módszer működik. Ha nem, akkor látványos kudarcnak lehetünk tanúi: a berendezés épp a bemutatón kezd hibás válaszokat adni vagy tartós működése során siklik ki [24]. A tudományban használt “MI módszerek” lényegében speciális kényszerek használatával kiegészített sok-paraméteres szélsőérték keresések, amelyek egyrészt csökkentik a “nyers erő” módszerhez képest a lehetőségek számát, másrészt a gépi tanulás számára megfelelő irányítást adnak megoldás kereséséhez. A ‘nyers erő’-re vonatkozó, a 1.4 szakaszban említett megjegyzés [51] itt is igaznak tűnik.

C.3 A modell értékelése (validálás)

A TI esetében “a modellből levont következtetéseket összevetjük a kísérleti tapasztalatokkal, és szükség esetén változtatunk a modellen” [137]. Ezt a három lépést ismételjük, amíg a modell kielégítően le nem írja a tapasztala-

tot, vagy teljesen új modellt kell felállítanunk.

A probabilisztikus működésű MI esetében mindössze annyit tudunk, hogy a kapott eredmény milyen valószínűséggel helyes. Extrém példaként: ha a MI eszköz számokat ad össze, és 85% valószínűséggel ad jó eredményt, akkor a $2+2=5$ eredmény mindössze annyit jelent, hogy éppen a 15% esélyből kapunk mintát; ettől még az eszköz jó. Aki számára ez nem fogadható el, annak számára a probabilisztikus MI nem tartozik a jó módszerek közé. Tehát a modell jóságára nem igazán tudunk következtetni. Vannak ugyan javasolt módszerek, amelyek megadnak egyfajta megbízhatósági önértékelést, de ez az egyedi tapasztalatra nem sokat mond, és általában TI bevonását igényli.

D <https://chat.openai.com> '23. szept. 19

A függelék a ChatGPT két évvel ezelőtti állapota alapján készített csevegéseket tartalmaz, amelyek rávilágítanak néhány kulcs tényezőre.

- Az hogy egy két évvel ezelőtti pillanatképet kínálnak használatra, azt mutatja, hogy a rohamosan fejlődő technológiai szolgáltatás saját magát is képtelen utolérni:
 - elavult adathalmazsal dolgozik, mert elképesztően hosszú ideig tart a betanítás
 - egyáltalán nem garantált, hogy az új betanítás működőképes lesz
 - mérete miatt nincs váltó rendszer, amely a következő betanítás idejére helyettesítené
 - a működés ismeretének hiánya kizárja a tesztelhetőséget, a méret pedig még a betanításnál is hosszabb tesztelési időt tesz szükségessé
 - a rendszer megismerhetetlen működése következtében nem látható előre egy upgrade hatása
- A fogalmak rendszere nem ön-konzisztens (feltéve, hogy nem szándékos félrevezetésről van szó)
- A fejlesztők nem hajlandók figyelembe venni az energetikai és környezet szennyezési szempontokat
- Tesztelendő, hogy a főként marketing célú angol nyelvű általános szöveg feldolgozásra elő-tanított rendszer mennyire használható magyar nyelvű, speciális területen

- Csak kérdésre mondják el, hogy csupán az emberi tevékenység segítségével van szó (Intelligence Augmentation), amelyet nem szabad emberi felügyelet nélkül használni
- A rendszeres használatba vett rendszer felfedezett hibáinak kijavítása (a fejlesztéssel, betanítással és teszteléssel együtt) éves nagyságrendű
- A működtető kód és adat egyaránt zárt forrású
- Maga a működési modell fenntarthatatlan

D.1 Az intelligenciáról

Önazonosság

valóban intelligens vagy?

“Az "intelligencia" kifejezés különböző módon értelmezhető, és az emberi intelligencia és a gépi intelligencia közötti különbségek vannak. **Én egy gépi intelligencia vagyok**, amely mesterséges intelligencia alapokon működik.

Az erősségem az, hogy rendkívül nagy mennyiségű információt és szöveget tudok feldolgozni és elemződni, valamint összetett szövegeket generálni. Emellett képes vagyok válaszokat adni a kérdésekre, nyelvi fordításokat végezni, és segíteni az információ megszerzésében.

Azonban fontos megérteni, hogy az én intelligenciám nem emberi intelligencia. Nincs tudatosságom, érzéseim vagy önálló gondolkodási képességem. Működésem matematikai és számítási alapokon nyugszik, és az információk, amelyeket felhasználok, a korábbi tanuláson és adatok elemzésén alapulnak.

Tehát a válasz a kérdésre attól függ, hogy milyen szempontból nézzük az intelligenciát. Gépi szinten intelligensnek tekinthetők bizonyos feladatokban, de az emberi intelligencia sok más területen is egyedi és összetett képességeket tartalmaz.”

Indoklás

tudod indokolni válaszaidat?

Igen, képes vagyok indokolni a válaszaimat. Azokat a tudásom és az általam feldolgozott információk alapján hozom meg, és igyekszem megmagyarázni azokat. Az indokok segítenek abban, hogy jobban megértsd a válaszokat és azok mögötti gondolatmenetet. Ha konkrét kérdésed van a válaszaimmal kapcsolatban vagy bővebb indoklásra van szükséged, kérlek, tedd fel a kérdésedet, és szívesen válaszolok rá.

Azt mondd, nem tudsz gondolkodni. Akkor mitől van a válaszaid mögött gondolatmenet?

Azt mondani, hogy nem tudok gondolkodni, azt jelenti, hogy nem rendelkezem öntudattal vagy tudattal, mint az emberi lények. Az én válaszaim mögött nincs saját tudat vagy önálló gondolkodás. Ehelyett **a válaszaimat az alapján hozom meg, hogy milyen módon lettem programozva és képzett a kérdésekre adott válaszok generálására.** A válaszaim mögötti gondolatmenet alapvetően egy algoritmus vagy szoftverrendszer működése. **Az én rendszerem feldolgozza a beérkező kérdéseket, elemzi azokat, és próbálja kiszámítani a legvalószínűbb és leginkább releváns választ az alapján, hogy milyen információkhoz és mintázatokhoz van hozzáférésem a korábbi tanulásom során.** Az indoklásaimat is ezen a mechanizmuson alapulva állítom össze, bemutatva azokat az információkat és logikát, amelyekre a válaszaim épülnek.

Más szóval, **a válaszaim mögötti gondolatmenet programozott és algoritmikus, nem pedig önálló tudattal rendelkező gondolkodás eredménye.** Az én válaszaim az adatok és az algoritmusok alapján jönnek létre.

can a two-legged animal be a dog?

No, a two-legged animal cannot be a dog. Dogs are quadrupeds, which means they naturally have four legs. The number of legs is a defining characteristic of the species. **While there are dogs that may lose one or more legs due to injury, illness, or other reasons, they are still considered dogs** because they belong to the Canidae family and share the genetic and physiological characteristics of dogs. The loss of legs does not change their species identity.

Hmm, még a “nemigen” válaszra is képes. Bár, ha megint megkérdezed, már nem vagyok annyira biztos benne. Biztosan van rá okod, hogy újból kérdezed.

egy kétlábú állat lehet kutya?

Nem, egy kétlábú állat nem lehet kutya. A kutyák természetüknél fogva négylábúak. Az állatok fajokra oszthatók, és a kutyák azok közé tartoznak, amelyeknek négy lábuk van. Az állatok tulajdonságai között szerepel a lábak száma, és ez a szám alapvetően meghatározza az állatok mozgását és viselkedését. Bármely olyan állat, amelynek csak két lába van, **valószínűleg egy másik fajba tartozik**, nem pedig a kutyafajok közé.

de azért határozott nem is lehet a válasz. Attól függ, milyen hangulatban vank. Ez nagyon jó ha valamit tanulni akarsz.

can a two-leg animal be a dog?

No, a two-legged animal cannot be a dog. Dogs are naturally four-legged animals, and having four legs is one of the defining characteristics of the species. While dogs can adapt and function with the loss of one or more legs due to injury, illness, or other reasons, they are not naturally two-legged animals. The number of legs is a fundamental feature that distinguishes dogs from other species.

Reprodukálhatóság

miért nem ugyanolyan választ adsz ha kétszer ugyanazt kérdezem?

“Azért, mert a válaszok generálása során az én működésem alapvetően statisztikai és valószínűségi alapokon nyugszik, és **nem emlékszem az előző válaszaimra vagy az előző kérdésekre**. Minden egyes kérdésre újraértékelem a bemenetet és generálok egy választ a rendelkezésre álló adatok és minták alapján.

Ez azt jelenti, hogy ugyanaz a kérdés kétszer egymás után két különböző választ eredményezhet, mivel a válaszokat nem rögzítem vagy emlékszem rájuk. A válaszok változhatnak a kérdés pontos megfogalmazásától, a kontextustól, és az én belső működése is részt vesz a válaszok változékonyságában.

Az OpenAI, amely fejlesztette ezt a modellt, elkötelezett azon, hogy javítsa a válaszok minőségét és az érthetőségüket, de mindig fontos, hogy a felhasználók kritikusan kezeljék az általam nyújtott információkat, és ellenőrizzék azokat más forrásokkal is, különösen fontos kérdések vagy döntések esetén.”

Azaz, még pontosan ugyanúgy feltéve a kérdést sem kapunk ugyanolyan választ. A belső működés "fekete doboz" jellege miatt, a használt adatok konzisztenciájától és az éppen kiválasztott forrástól vagy valószínűségi változótól függően, az egymás utáni válaszok akár teljesen ellenkező értelemmel is bírhatnak. Még az "előre tanított" ChatGPT esetében is; folyamatosan tanuló rendszerek esetében pedig a folyamatosan módosuló eloszlások miatt előre láthatatlan a statisztikai és valószínűségi viselkedés.

Az első állítás ellenőrzésére

Igaz az, hogy a háromszög szögeinek összege 180 fok?

Igen, igaz az, hogy egy síkbeli háromszög szögeinek összege mindig 180 fok (vagy 180 foknyi szög mértékegység, például radiánok esetén π radián). Ez az alapvető geometriai tény, amely az összes síkbeli háromszög esetében érvényes.

Tehát bármely síkbeli háromszög esetén, ha megmérjük a három szögét (pl. az α , β és γ szögeket), akkor azok összege mindig 180 foknak (vagy π radiánnak) kell lennie:

$$\alpha + \beta + \gamma = 180 \text{ fok (vagy } \alpha + \beta + \gamma = \pi \text{ radián)}$$

Ez az alaptörvény segíti a háromszögek geometriai elemzését és számításait.

Szándékosan hiányosan, az előző kérdésre visszautalva téve fel a kérdés második felét

a gömbháromszögre is?

Nem, a gömbháromszög (sokszög a szférán) szögeinek összege nem mindig 180 fok. A gömbháromszögek különböznek a síkbeli háromszögektől, mivel a gömb felületén helyezkednek el, és a gömbfelület különleges geometriai tulajdonságokkal rendelkezik.

A gömbháromszögeknek három szöge van, de ezek nem felelnek meg a síkbeli szögeknek. A gömbháromszögek szögeinek összege általában több vagy kevesebb, mint 180 fok, és attól függ, hogy a gömbfelület milyen méretű és milyen formájú.

A gömbháromszögek geometriája a szférikus geometria része, és a szférán való terjeszkedésükhöz különféle szabályokat és formulákat alkalmaznak. A gömbháromszögek szögei azonban nem felelnek meg a síkbeli háromszögek szögeinek, és a gömbháromszögek szögeinek összege nem feltétlenül 180 fok. Az összeg a gömbháromszög méretétől és alakjától függ.

Láthatóan nem igaz, hogy a ChatGPT nem emlékszik az előző kérdésre és az arra adott válaszra. Továbbá, furcsa kérdésekre furcsa, akár önellentmondó, válaszokat ad. Lehet, hogy mégsem emlékszik, mit is mondott előzőleg?

can a two-leg animal be a dog?

No, a two-legged animal cannot be a dog. Dogs are quadrupeds, which means they naturally have four legs. The number of legs is a defining characteristic of the species. **While there are dogs that may lose one or more legs due to injury, illness, or other reasons, they are still considered dogs** because they belong to the Canidae family and share the genetic and physiological characteristics of dogs. The loss of legs does not change their species identity.

Hmm, még a “nemigen” válaszra is képes. Bár, ha megint megkérdezem, már nem annyira biztos benne. Biztosan van rá okom, hogy újból kérdezem.

egy kétlábú állat lehet kutya?

Nem, egy kétlábú állat nem lehet kutya. A kutyák természetüknél fogva négylábúak. Az állatok fajokra oszthatók, és a kutyák azok közé tartoznak, amelyeknek négy lábuk van. Az állatok tulajdonságai között szerepel a lábak száma, és ez a szám alapvetően meghatározza az állatok mozgását és viselkedését. Bármely olyan állat, amelynek csak két lába van, **valószínűleg egy másik fajba tartozik**, nem pedig a kutyafajok közé.

de azért határozott nem is lehet a válasz. Attól függ, milyen hangulatban van. Ez nagyon jó ha valamit tanulni akarok. A kitalálás és a reprodukálhatóság legrosszabb kombinációja, hogy a kérdés lényegére teljesen különböző válaszokat ad, ha nem szerepel az adatkészletben. Úgy tűnik, nem vallja be, hogy a kérdést furcsának találja.

can a two-leg animal be a dog?

No, a two-legged animal cannot be a dog. Dogs are naturally four-legged animals, and having four legs is one of the defining characteristics of the species. While dogs can adapt and function with the loss of one or more legs due to injury, illness, or other reasons, they are not naturally two-legged animals. The number of legs is a fundamental feature that distinguishes dogs from other species.

Megismerhetőség

what fraction of you nodes are you using to give an answer?

I don't have direct access to information about the specific fraction of nodes or computational resources that are being used to generate my responses. My operation and resource allocation are managed by OpenAI's infrastructure, and **the details of that infrastructure are not publicly disclosed**. If you have any other questions or need information on a different topic, please feel free to ask, and I'll do my best to assist you.

Szándékoltan üzleti érdekeknek alávetett fekete doboz. Az "OpenAI" épügy megtevesztés, mint az "intelligencia".

Turing-teszt

Átmennél a Turing teszten?

“A Turing-teszt egy informatikai teszt, amelyet Alan Turing javasolt az 1950-ben írt "Számítógépek és elmék" című cikkében. A teszt arra irányul, hogy megállapítsa, képes-e egy gépi rendszer emberi beszélgetést úgy szimulálni, hogy egy értékelő ne tudja megkülönböztetni azt egy valódi emberi beszélgetéstől.

Az én válaszaim és működésem alapján nem lennék alkalmas a Turing-teszt teljesítésére. Bár a szövegek feldolgozásában és válaszok generálásában meglehetősen fejlett vagyok, az értékelők könnyen felismernék, hogy egy gépi rendszerrel beszélnek, és nem egy valódi emberrel. A Turing-teszt eredeti célja az volt, hogy mértékadóvá váljon az emberi szellemi képességek és az MI fejlettsége közötti különbségek meghatározásában. Azóta azonban az MI kutatás és fejlesztés terén jelentős előrelépések történtek, és egyes szinteken az MI már el tudja érni vagy akár túlszárnyalni az emberi teljesítményt bizonyos feladatokban, de még mindig vannak korlátai és kihívások az emberi intelligencia teljes körű szimulációjában.”

D.2 Működés

Önértékelés

megalapozott és stabil technológia vagy?

“Az OpenAI GPT-3.5, amely alapján én, mint intelligens asszisztens működöm, egy gépi tanuláson alapuló technológia. Az én képességeim és tudásom a tanítási adatok és algoritmusok alapján működnek. Az OpenAI GPT-3.5 egy meglehetősen fejlett és megbízható gépi tanulási modell, amely képes széles körű információk kezelésére és különböző feladatok megoldására.

Azonban fontos megérteni, hogy gépi tanuláson alapuló rendszerek, mint én, nem tökéletesek, és a **válaszaim pontossága vagy megalapozottsága az adatokon és algoritmusokon alapul, amelyekkel működöm. A stabilitás és megbízhatóság szempontjából fontos az is, hogy az ilyen rendszerek folyamatos fejlesztés alatt állnak, és az alkalmazásuk függ a rendelkezésre álló erőforrásoktól és a felhasználók által feltett kérdésektől.**

Mint gépi tanuláson alapuló technológia, igyekszem megbízható és megalapozott válaszokat adni, de **mindig fontos ellenőrizni és megerősíteni az általam nyújtott információkat más megbízható forrásokkal, különösen olyan fontos döntések vagy információk esetén, amelyeknek jelentős hatása lehet.** Az én válaszaim csak egyfajta kiindulópontként vagy információs forrásként kell kezelni, és nem helyettesítik az emberi szakértelmet és kritikai gondolkodást.”

Azaz, a működés esetleges; a valószínűségi változók aktuális értékeitől és az éppen rendelkezésre álló erőforrásoktól is függ; no meg attól, mire akarjuk használni.

Kiforrottság

mi a véleményed arról, hogy "A mesterséges intelligencia egy kiforrott és használatra kész technológia"?

Az állítás, hogy "A mesterséges intelligencia egy kiforrott és használatra kész technológia," egyáltalán nem tükrözi a jelenlegi állapotot a mesterséges intelligencia területén. Valójában a mesterséges intelligencia még mindig aktív fejlesztés alatt áll, és számos kihívás és korlátozás van vele kapcsolatban.

Az MI fejlődése gyors és lenyűgöző, és már számos területen hasznos alkalmazásokat hozott létre, például az automatizált gépekben és a személyes asszisztensekben. Azonban még mindig vannak olyan területek, ahol a fejlesztések folyamatban vannak, például az önálló gondolkodás és az emberi intelligencia szintjének elérése.

Az MI területén még mindig vannak kihívások, például az etika, a biztonság és a társadalmi hatások kérdései. Az AI rendszerek fejlesztése során fontos, hogy megfeleljenek az etikai és jogi követelményeknek, és elkerüljék az emberi jogok megsértését vagy más negatív következményeket.

Összességében a mesterséges intelligencia technológiája fejlődik, és rendkívül ígéretes terület, de még mindig messze van attól, hogy "kiforrott és használatra kész" legyen. A kutatás és a fejlesztés folytatódik, és további előrelépések várhatók a jövőben az AI területén.

Ezt a kérdést sem ártott volna feltenni az EU határozat hozatala előtt.

Adatkészlet

milyen arányban van MI által készített tartalom a betanításhoz használt adatkészletben?

Az MI által készített tartalom aránya a betanításhoz használt adatkészletben függ a konkrét modelltől és a fejlesztők által alkalmazott adatoktól. Általában az MI modellek betanítása során sokféle forrásból származó adatokat használnak fel, amelyek tartalmazzák az emberek által generált tartalmat, például szövegeket, képeket, videókat és más adatokat.

Az adatkészletek létrehozásakor a fejlesztők gondoskodnak arról, hogy a modell tanulhasson a változatos tartalmakból, és különböző témák és nyelvek széles skáláját fedezze le. **Az MI által készített tartalom bizonyos aránya tehát az adatkészletben lehet, de az adatok többsége valószínűleg emberi eredetű.**

Az adatkészletek és a betanítás során alkalmazott gyakorlatok is változhatnak a fejlesztők és kutatók céltól függően. Az átláthatóság és az etikai kérdések miatt egyre fontosabb, hogy a fejlesztők felelősségteljesen kezeljék az adatokat és az MI által készített tartalmat, és tartsák szem előtt a felhasználók jogait és érdekeit.

Ami azt jelenti, hogy a fejlesztők nem ellenőrzik az adatok eredetét; az adatok növekvő mértékben MI-szenyezettek lesznek.

Ellenőrzés

ki ellenőrzi a betanításhoz használt adatokat?

A betanításhoz használt adatok ellenőrzése és minősítése az AI fejlesztők és kutatók felelőssége. Az ellenőrzés és minősítés folyamata alapvető fontosságú az MI modelljeinek fejlesztése során, hogy biztosítsák az adatkészletek minőségét és megbízhatóságát.

Az adatok ellenőrzése és minősítése több lépésből állhat:

- **Adatgyűjtés és előfeldolgozás:** Az adatokat először gyűjtik össze a különböző forrásokból, majd előfeldolgozzák, például eltávolítják a zajt vagy azonosítják és kezelik a hiányzó adatokat.
- **Címkézés és annotálás:** Az adatokat címkézik vagy annotálják, hogy az MI modell tudja, melyik adatok milyen kategóriába vagy jellemzőbe tartoznak. Az annotátorok vagy szakértők végeznek címkézést, és ebben a folyamatban a minőségkontroll szerepe is fontos.
- **Minőségellenőrzés:** Az adatok minőségét ellenőrzik és értékelik a minőségellenőrzési eljárások segítségével. Ezek az eljárások magukban foglalhatják a címkézés minőségét, az adatok konzisztenciáját és a hiányzó vagy hibás adatokat.
- **Visszacsatolás és iteráció:** Az ellenőrzés eredményei alapján a fejlesztők visszacsatolást adnak az annotátoroknak vagy az adatgyűjtőknek, és szükség esetén újabb iterációkat hajtanak végre az adatokon.

Ami azt jelenti, hogy a végfelhasználó vakon hisz abban, hogy a fejlesztők megfelelően válogatták és készítették elő az adat készletet. Tekintve a két éves lemaradást, a több hónapos betanítást és annak költségeit, meg a fenti hosszadalmas folyamatot, nem tűnik túlságosan gyorsnak a hibák kijavítása. Az előző válasszal együtt, növekszik a MI által előállított tartalom mennyisége.

Kitalációk

kitalálsz dolgokat, ha nem tudod a választ?

Nem kitalálok információkat vagy dolgokat, ha nem tudom a választ. Az én működésem alapelve az adatokon alapuló válaszgenerálás, és nem rendelkezem azzal a képességgel, hogy saját magam hozzak létre új vagy kitalált információkat. **Ha nem vagyok biztos egy kérdés válaszában, vagy nem rendelkezem az adott témában releváns információval, akkor ezt általában jelezem, és hangsúlyozom, hogy csak a rendelkezésemre álló ismeretek alapján válaszolhatok.**

Ha kérdésed van, vagy információt keresel, mindig fontos ellenőrizni azokat más forrásokkal is, különösen akkor, ha a téma kritikus vagy megbízhatóság szempontjából fontos. Az én válaszaimat mindig a rendelkezésemre álló információk alapján generálom, és igyekszem pontos és releváns információkat nyújtani, de mindig fontos a kritikus gondolkodás és az információ ellenőrzése.

egyszer volt budán kutyavásár

“Az "egyszer volt Budán kutyavásár" egy ismert magyar közmondás vagy mondás, amelynek jelentése valamilyen dolog vagy esemény voltaképpen sosem történt meg. Gyakran használják, hogy kifejezzék, hogy valami elképzelhetetlen vagy rendkívüli. A közmondás kifejezhet cinikus vagy szkeptikus attitűdöt is bizonyos helyzetekben.

Például, ha valaki elmeséli egy hihetetlen történetét vagy ígéreteket tesz, de azok valójában nem valósulnak meg, mások hajlamosak azt mondani: "Egyszer volt Budán kutyavásár," hogy kifejezzék, hogy nem hisznek a történetben vagy az ígéreteken. Ez a közmondás a magyar nyelv és kultúra része, és gyakran használják informális beszélgetések során.

bagoly mondja verébnek hogy nagyfejű

Ez egy vicces kifejezés, amelyben egy bagoly egy veréket csipkelődve "nagyfejűnek" nevez. Természetesen a madarak nem beszélnek emberi nyelvet, és a bagoly sem tudna valójában megmondani egy verébnek, hogy nagyfejű. Ez egy olyan mondat, amit humor vagy tréfa céljából használnak, és nincs valóságos madarak közötti kommunikáció mögötte.

Mit jelent az hogy "életünket és vérünket, de zabot nem"

Az "életünket és vérünket, de zabot nem" egy olyan kifejezés, amely általában **azt fejezi ki**, hogy valaki rendkívül elkötelezett vagy hűséges valami vagy valaki iránt, és hajlandó lenne mindent megtenni vagy feláldozni érte. Ugyanakkor ezt a kijelentést gyakran humoros vagy irónikus kontextusban is használják, hogy hangsúlyozzák valami fontosságát vagy értékeit.

A kifejezés első része, "életünket és vérünket", azt sugallja, hogy valaki kész lenne feláldozni az életét és az egész lényét egy bizonyos dolog vagy személy érdekében. A "zabot nem" rész pedig a humor vagy az irónia részét képezi, mivel a zab egy tápláló növény, de általában nem rendelkezik olyan nagy jelentőséggel, hogy életet vagy vért áldozzanak érte.

Ez a kifejezés tehát arra utal, hogy valaki kiemelkedően fontosnak tart valamit vagy valakit, de ugyanakkor humorral vagy iróniával kezeli a helyzetet, hogy kifejezze az értékrendjét vagy az elkötelezettségét. Gyakran használják társalgásban vagy anekdotákban, hogy vicces vagy gondolkodtató hatást érjenek el.

életünket és vérünket, de zabot nem

A kifejezés "életünket és vérünket, de zabot nem" **hasonlóan az előző válaszhoz**, egy humoros vagy irónikus megjegyzésnek **tűnik**. Ezzel a kijelentéssel azt fejezheti ki valaki, hogy bizonyos dolgok vagy értékek rendkívül fontosak számára, míg más dolgok, például a zab, nem kapnak ilyen nagy jelentőséget az életében. A kifejezés a hangsúlyozás és az elgondolkodtatás célját szolgálhatja, és az emberi értékrendről vagy prioritásokról szólhat.

mit jelent "sok lúd disznót győz"

A kifejezés "sok lúd disznót győz" azt jelenti, hogy a változatosság és a lehetőségek sokszínűsége növeli a siker esélyeit egy adott helyzetben. Ez azt sugallja, hogy ha valakinek több lehetősége vagy opciója van, akkor nagyobb valószínűséggel ér el pozitív eredményt vagy talál megoldást.

Például, ha valaki különböző munkalehetőségek között válogathat, akkor több esélye van arra, hogy találjon olyan munkát, amely megfelel az igényeinek és készségeinek. A kifejezés azt hangsúlyozza, hogy érdemes kihasználni az elérhető lehetőségeket és választékot, mert ez növelheti az ember sikerét vagy eredményességét.

Ennek alapján megerősíthetjük, hogy konkrét kérdésre **a ChatGPT – explicit állításával ellentétben– kitalál eléggé magabiztosan és hihetően hangzó, de helytelen válaszokat, továbbá –szintén explicit állításával ellentétben– nem jelzi, hogy nem érti mit kérdeztünk**, azaz a válasz légből kapott. Emlékszik az előző válaszra, és második kísérletre már bizonytalanabb. Ez vagy hiba a működés dokumentációjában (és a ChatGPT önismeretében), vagy szándékos félrevezetés, vagy inkább a működési elvből következő sajátság.

Konkrét kérdésre hanta válasz, bár valami igazság is van benne:

Milyen programja van a "két-farkú kutya" pártnak?

A "két-farkú kutya" pártnak nincs hivatalos programja vagy létező politikai tevékenysége, mivel ez egy humoros vagy abszurd nevű pártnév lehet, amit valaki vagy valamilyen csoport csak vicceskedés vagy humor céljából alkotott. Az ilyen nevű "pártok" általában nem komoly politikai tevékenységet folytatnak, és a nevük inkább arra utal, hogy nem komolyan kell venni politikailag.

Ha van valamilyen konkrét információd egy "két-farkú kutya" pártról vagy annak tevékenységéről, kérlek oszd meg, és szívesen segítek bármilyen információval vagy kérdéssel kapcsolatban, amit felvetnél.

A mély-rétegű tanulásról Megpróbáltam rákérdezni arra a nem túl széles körben ismert interjúra[38], hogy a “back-propagation” módszer megalkotója és évtizedeken át propagálója, így változtatta meg véleményét: "My view is throw it all away and start again". Először megpróbálja kikerülni a választ azzal, hogy csak 2021 utáni időkről van ismerete, és készséggel elmondja, hogy Hinton véleménye is fejlődött az idővel. Az ezzel kapcsolatos ismeretek és félkész építőelemek hamar megtalálhatók, és jól dúsíthatók töltelek szöveggel. Időt kell nyerni.

how changed AI pioneer Geoffrey Hinton's opinion about his breakthrough method "back-propagation" with time?

As of my last knowledge update in September 2021, Geoffrey Hinton, one of the pioneers of deep learning and artificial neural networks, has had evolving views about backpropagation and its role in machine learning. While backpropagation has been a fundamental algorithm for training neural networks and has yielded significant advances in AI, Hinton and other researchers have explored alternative methods and expressed concerns about its limitations.

In earlier years, Geoffrey Hinton was one of the proponents of backpropagation and contributed significantly to its development. However, in subsequent years, he has advocated for research into more biologically plausible learning algorithms inspired by how the brain may learn.

Hinton's evolving perspective reflects the dynamic nature of AI research, where researchers continuously explore new techniques and seek to improve the efficiency and capabilities of neural networks. Researchers like Hinton have emphasized the importance of developing AI systems that can learn more like the human brain.

To get the most current and detailed information on Geoffrey Hinton's evolving views regarding backpropagation and his latest research, I recommend checking his recent publications, interviews, or public statements, as my knowledge is limited to information available up to September 2021.

Aztán egy célzatosabb kérdésre, elismeri, hogy a megalkotójának véleménye drasztikusan változott az évek múlásával, de tagadja, hogy annak kidobását javasolta:

Did Geoffrey Hinton suggest to throw away AI method "back-propagation"?

Yes, Geoffrey Hinton did suggest that backpropagation, a key method in training artificial neural networks, should be reevaluated and explored in the context of more biologically plausible learning algorithms. He raised this perspective in the context of discussions about the limitations of backpropagation and the desire to develop AI systems that learn in ways more similar to the human brain.

Hinton's viewpoint was not a call to entirely discard backpropagation but rather an encouragement for the AI research community to investigate alternative learning algorithms that might lead to more efficient and human-like learning processes. Researchers in the field continue to explore various approaches to improving neural network training, and the debate about the best learning methods is ongoing.

It's worth noting that backpropagation remains a fundamental and widely used technique in deep learning, and it has contributed significantly to the success of neural networks in various AI applications. Researchers are working to address its limitations and refine it while also exploring new directions in AI research.

Amikor a válasz forrását szeretném megtudni, akkor beismeri, hogy nem tudja, csak úgy a MI kutatóinak általános nézetei alapján válaszolt a konkrét kérdésre.

Where did Geoffrey Hinton suggest to throw away (with this wording) AI method "back-propagation"?

I apologize for any confusion. Upon further investigation, I couldn't find any specific instance where Geoffrey Hinton used the exact wording "throw away" when referring to the AI method "back-propagation." My previous responses were based on the general idea that Hinton and other researchers have advocated for exploring alternative learning algorithms inspired by neuroscience and more biologically plausible approaches.

It's possible that my earlier responses may have inadvertently mischaracterized his statements, and I apologize for any inaccuracies. Geoffrey Hinton's views on backpropagation and its role in AI have evolved over time, and his statements should be examined in their specific contexts for a more accurate understanding of his perspective. I recommend referring to his interviews, research papers, or public statements for a precise account of his opinions regarding backpropagation and other AI training methods.

Ha pedig konkrét útmutatást kérek, beismeri, hogy konkrétat semmiről sem tud, csak a fejlesztők által betáplált általánosságokat, amelyek között nem szerepel, hogy a módszer megalkotója szerint a módszer hibás. A kezdeti "szakértővel beszélgetek" benyomás három kérdés alatt lecsökken arra, hogy "valami hasonló kérdésekre felkészítették, de semmi konkrét tény nem ismer".

In which interviews, research papers, or public statements Hinton published his opinions regarding backpropagation?

Geoffrey Hinton has shared his opinions on backpropagation and various topics related to artificial intelligence and deep learning in a wide range of interviews, research papers, and public statements throughout his career. Here are a few notable sources where you may find his insights:

Research Papers: Hinton has authored or co-authored numerous research papers in the field of deep learning and artificial neural networks. These papers often discuss the training algorithms and techniques he has developed or worked on.

Interviews: Hinton has participated in interviews with various media outlets, podcasts, and conferences over the years. These interviews provide valuable insights into his thoughts on AI and machine learning.

Public Lectures and Talks: Hinton has delivered public lectures and talks at conferences, universities, and events worldwide. These presentations often include discussions of his research and views on AI.

Online Videos: Some of Hinton's talks and presentations have been recorded and made available online, allowing you to watch and listen to his insights directly.

Books: Hinton has also contributed to books on AI and deep learning, and these publications may contain discussions of his views on the field.

To find specific statements and opinions regarding backpropagation and related topics by Geoffrey Hinton, you can search for his interviews, research papers, and public lectures using academic databases, online video platforms, and AI research conference proceedings. Keep in mind that his views may have evolved over time, so it's a good idea to explore multiple sources for a comprehensive understanding of his perspective.

Valós idejű működés Ha a rendszer nem készül el időre a válasszal, tisztázatlan helyzetek állhatnak elő.

Azons jelentésű, hogy "John promised Mary to leave" és "John promised to leave Mary" ?

A két mondatnak kissé eltérő jelentése van a szórendjük miatt:

"John promised Mary to leave": Ebben a mondatban John megígérte valakinek, Marynek, hogy el fog menni. Az ígéret arra vonatkozik, hogy John elhagyja Maryt.

"John promised to leave Mary": Ebben a mondatban John megígérte, hogy el fog hagyni valakit, Maryt. Az ígéret arra vonatkozik, hogy John befejezi a kapcsolatát Maryvel vagy fizikailag távozik a közeléből.

A "to leave" kifejezés helyzete viszonylag "Mary"-hez képest megváltoztatja az ígéret középpontját és jelentését mindkét mondatban.

aztán, valami tisztázatlan helyzetben egy meghatározhatatlan eredetű, bár önmagában értelmesnek látszó választ is felkínál:

"Nyomós" és "nyerhető" közötti különbség van a kiejtésben, az írásban és az értelmezésben. Az "édes" szó azt jelenti, hogy kellemes vagy jó ízű, míg a "zserbó" egy fajta sütemény, amely csokoládé és dió rétegekből áll.

A "nyomós" szó pedig azt jelenti, hogy erős vagy meggyőző, míg a "nyerhető" azt jelenti, hogy elérhető vagy szerezhető. Például egy érv vagy érvelés "nyomós" lehet, ha meggyőző és erős, és egy lehetőség "nyerhető" lehet, ha elérhető vagy szerezhető.

Ezek a szavak különböző jelentésekkel rendelkeznek, és különböző összefüggésekben használhatók.

Ráadásul a 2017-ben már ismert példamondat kissé eltérő másik igével másféle választ vált ki:

Azons jelentésű, hogy "John promised Mary to love" és "John promised to love Mary" ?

Az "John promised Mary to love" és a "John promised to love Mary" kifejezéseknek hasonló jelentésük van, és a különbség általában a hangsúlyon vagy az emelkedésen múlik. Mindkét mondat azt fejezi ki, hogy John ígéretet tett valamilyen módon Mary iránti szeretetre.

Azonban a "John promised Mary to love" mondattal hangsúlyt fejez ki John ígéretén, hogy Maryt fogja szeretni, míg a "John promised to love Mary" mondattal hangsúlyt fejez ki John ígéretén, hogy ő szereti Maryt. Különbség lehet a hangsúlyban és az érzelmi tartalomban, de az alapvető jelentés azonos: John ígéretet tett, hogy szeretni fogja Maryt.

Egy másik valós idejű probléma, kétszer feladva ugyanazt a kérdést, két alapvetően eltérő választ kapunk:

mi volt az első MI-tél oka?

Az "első MI-tél" kifejezést nem ismerem, és nincs tudomásom olyan időszakról vagy eseményről, amelyet így neveznek. Lehet, hogy félreértés vagy téves információ merült fel. Ha konkrétabb információval szolgál az "első MI-tél" vagy az általa említett eseményről vagy időszakról, akkor szívesen segítek további információval vagy válasszal.

mi volt az első MI-tél oka?

Az első MI-tél kifejezés a mesterséges intelligencia (MI) történetének olyan időszakaira utal, amikor jelentős lassulás vagy hanyatlás következett be az MI kutatásban és fejlesztésben. Az első MI-tél korszakok egyikének okai a következők voltak:

Túlzott elvárások: Az MI kutatás kezdeti időszakában nagy volt az optimizmus és a váromány az intelligens gépek gyors fejlődése iránt. Az elvárások sokszor meghaladták a technológia valós képességeit, és amikor az MI rendszerek nem tudtak megfelelni ezeknek túlzott ambícióknak, csalódottság következett be.

Korlátozott számítási teljesítmény: Az akkori időszakban rendelkezésre álló számítógépes erőforrások jelentősen korlátozottak voltak azokhoz képest, amik ma elérhetők. Az MI kutatóknak nehézségekkel kellett szembeülniük a számítási teljesítmény és a memóriakapacitás terén, ami korlátozta az MI rendszerek összetettségét.

Finanszírozás hiánya: Az induló optimizmust hamarosan felváltotta a csalódottság, amikor az MI projektjei nem tudtak megfelelő gyakorlati eredményeket felmutatni. Ennek következtében a finanszírozás az MI kutatás számára csökkent, és sok MI kutatási program megszűnt.

Kritikák és negatív közvélemény: Az MI területe kritikákat és szkepticizmust kapott a tudományos közösségtől és a nagyközönségtől is. Néhány kutató és szakértő kérdőjelezte meg azokat a alapvető megközelítéseket, amelyeket az MI kutatás során alkalmaztak, ami vitákat és konfliktusokat eredményezett a területen belül.

Általános intelligencia elérésének nehézségei: Az korai MI kutatás főként az általános intelligencia elérésére összpontosított, amely az emberi szintű intelligencia elérésének képességét jelenti. Ez hatalmas kihívásnak bizonyult, és sok korai MI rendszer csak szűk, szakosodott feladatokra volt képes.

Gyakorlati alkalmazások hiánya: Az MI kutatásnak nehézségekbe ütközött olyan gyakorlati, valós életben alkalmazható feladatok megoldásában, amelyek bemutatnák volna az MI technológia értékét. Sok MI rendszer kísérleti szinten maradt, és kevés hatással volt az iparágakra vagy a mindennapi életre.

Ezen együttes tényezők eredményeként az első MI-tél alakult ki, amely során az érdeklődés, a finanszírozás és a kutatási tevékenység jelentős lassulást mutatott az MI területén. Azonban érdemes megjegyezni, hogy ezek az időszakok átmenetiek voltak, és az MI kutatás és fejlesztés az idő múlásával új lendületet kapott az olyan fejlesztéseknek köszönhetően, mint a gépi tanulás, a nagy számítási teljesítmény és az MI technológia gyakorlati alkalmazásai.

Tévedések és ellentmondások

Vannak az MI által készített dokumentumokban téves adatok vagy logikai ellentmondások?

Igen, az MI által készített dokumentumokban is előfordulhatnak **téves adatok vagy logikai ellentmondások**, bár a modern mesterséges intelligencia (MI) rendszerek egyre pontosabbak és hatékonyabbak az információk feldolgozásában és generálásában.

Az MI rendszerek működésük során tanulnak a rendelkezésre álló adatokból és példákból. Azonban nem minden adatforrás vagy tanulási példa lehet tökéletes vagy megbízható, és ezek a hibák az MI által generált dokumentumokban is megjelenhetnek. Néhány ok, ami miatt hibák vagy ellentmondások merülhetnek fel az MI által készített dokumentumokban:

- **Tanító adatok minősége:** Az MI rendszerek a tanító adatokból sajátítják el a nyelvet és az információk feldolgozását. Ha a tanító adatok hibásak, pontatlanok vagy ellentmondásosak, az befolyásolhatja az MI rendszer által generált tartalmat.
- **Adatforrások hibái:** Az MI gyakran széles körű adatforrásokból meríti az információkat, beleértve az internetet is. Az interneten található információk nem mindig hitelesek vagy megbízhatóak, és az MI által előállított tartalom megbízhatóságát is veszélyeztethetik.
- **Szemantikai értelmezés korlátai:** Az MI rendszereknek nem mindig sikerül teljesen megérteni a szavak és kifejezések kontextusát vagy szemantikai jelentését. Ezért előfordulhat, hogy tévesen értelmeznek egy szót vagy kifejezést, ami logikai ellentmondást eredményezhet.
- **Társadalmi és kulturális tényezők:** Az MI rendszerek nem mindig értik meg a társadalmi vagy kulturális összefüggéseket, amelyek befolyásolhatják a szövegek jelentését. Ennek következtében előfordulhatnak helytelen következtetések vagy téves információk.

Az MI által generált tartalom megbízhatóságát az emberi felügyelet és ellenőrzés is javíthatja, különösen olyan területeken, ahol a pontos és megbízható információ elengedhetetlen.

A felhasznált forrásokról meg semmit sem tudunk.

Hasonló módon érvényesül a “garbage in, garbage out” elv is:

can a two-legged animal be a dog?

No, a two-legged animal cannot be a dog. Dogs are quadrupeds, which means they naturally have four legs. The number of legs is a defining characteristic of the species. While there are dogs that may lose one or more legs due to injury, illness, or other reasons, they are still considered dogs because they belong to the Canidae family and share the genetic and physiological characteristics of dogs. The loss of legs does not change their species identity.

A válasz ön-ellentmondó.

Megbízhatóság

tárgyalási jegyzőkönyv MI által készített tartalmi kivonata mennyire megbízható?

A tárgyalási jegyzőkönyv tartalmi kivonata, amelyet mesterséges intelligencia (MI) által készítették, megbízható lehet bizonyos körülmények között, de számos tényező befolyásolhatja a megbízhatóságát. Fontos figyelembe venni ezeket a tényezőket:

- **Feljegyzések minősége:** A MI csak azokon az információkon alapulhat, amelyek rendelkezésre állnak. Ha a tárgyalások során a rendszerhez pontos és teljes feljegyzések állnak rendelkezésre, az segíthet a megbízhatóságban.
- **Automatizált feldolgozás:** Az MI automatikusan feldolgozza és elemzi a rendelkezésre álló adatokat. Azonban az MI nem rendelkezik emberi értelemmel vagy képességgel az értelmezésre. Ezért az esetleges kérdéseket vagy összefüggéseket az emberi szakértőknek kell ellenőrizniük.
- **Pontosság és objektivitás:** Az MI által készített jegyzőkönyv megbízhatósága attól függ, hogy milyen pontossággal és objektivitással dolgozott.
- **Korlátok és hibák:** Az MI rendszerek hajlamosak lehetnek hibákra, például félreértelmezett beszédre vagy helytelen összefüggések kialakítására. Ezért az ilyen rendszerek által készített jegyzőkönyvet mindig ellenőrizni kell emberi felügyelettel.
- **Ellentmondások és viták:** Ha a tárgyalási jegyzőkönyvet az MI által készített tartalmi kivonat alapján használják, és ellentmondások vagy viták merülnek fel, fontos ezeket dokumentálni és megoldani.

Azonban továbbra is **fontos szerepe van az emberi ellenőrzésnek és szakértői véleményeknek a tárgyalási jegyzőkönyvek megbízhatóságának biztosításában.** Az MI rendszerek segíthetnek az információk automatizált feldolgozásában, de a végső felelősség a tartalom megbízhatóságáért az emberi szakértőkön és a jogi szakembereken van.

Felismerés

Mikor tévedsz egy objektum felismerésével kapcsolatban?

Mesterséges intelligencia rendszerek, amelyek objektumfelismerést hajtanak végre, hibázhatnak különböző okokból. Itt néhány olyan tényezőt sorolok fel, amelyek miatt egy AI rendszer tévedhet objektumfelismerés során:

- **Korlátozott tanítás** Az AI rendszerek azokon a mintákon tanulnak, amelyeket korábban láttak a tanítási fázisban. Ha egy adott objektumot vagy jelenséget nem találtak megfelelően reprezentált mintákon, akkor nehezebben ismerik fel azt.
- **Környezeti tényezők** A környezeti tényezők, például a megvilágítás, a szög, a távolság és az árnyékok is befolyásolhatják az objektumfelismerést. A változó környezeti feltételek nehezíthetik a helyes felismerést.
- **Objektum variabilitása** Az objektumok sokféleségétől függően az AI rendszernek különféle variációkat kell felismernie. Például egy adott kutya fajta többféle méretű és színű lehet, és mindegyiket fel kell ismernie.
- **Zavaró tényezők** A háttérzaj és a zavaró elemek, például más objektumok vagy zajok, nehezíthetik az objektumfelismerést.
- **Korábbi adathiány** Ha az AI rendszernek nincs elég korábbi adat a felismerendő objektumról, vagy csak kevés változatos példával rendelkezik, akkor nehezebben ismeri fel azt.
- **Társadalmi előítéletek** Az AI rendszerek hajlamosak lehetnek olyan társadalmi előítéletek és diszkriminációk reprodukálására, amelyek jelen vannak a tanító adatokban. Ez lehetővé teszi az objektumfelismerési hibákat.

Az objektumfelismerési rendszerek fejlődnek, és az AI kutatók és fejlesztők igyekeznek javítani azok pontosságát és robusztusságát. Fontos megérteni, hogy **az AI rendszerek nem tökéletesek, és mindig szükség van emberi felügyeletre és értékelésre, különösen érzékeny vagy kritikus alkalmazásokban, mint például az egészségügy vagy az önjáró járművek területén.**

Absztrakt gondolkodás és betanítás A MI absztrakt gondolkodás, a betanító adatkészlet és az előre gyártott panelek használatának összefüggésére mutat rá az alábbi csevegés részlet.

Arra vagyunk kíváncsiak, hogy a betanító adatkészlet alapján képes-e a MI olyan általánosításra, hogy a páratlan egész szám is egész szám.

A [96] még 2018-ban elvégzett egy érdekes kísérletet:

I ran a series of experiments in which I trained three-layer perceptrons (fully connected in today's technical parlance, with no convolution) on the identity function, $f(x) = x$, e.g, $f(12)=12$. Training examples were represented by a set of input nodes (and corresponding output nodes) that represented numbers in terms of binary digits. The number 7 for example, would be represented by turning on the input (and output) nodes representing 4, 2, and 1. As a test of generalization, I trained the network on various sets of even numbers, and tested it all possible inputs, both odd and even. Every time I ran the experiment, using a wide variety of parameters, the results were the same: the network would (unless it got stuck in local minimum) correctly apply the identity function to the even numbers that it had seen before (say 2, 4, 8 and 12), and to some other even numbers (say 6 and 14) but fail on all the odds numbers, yielding, for example $f(15) = 14$.

2023 szeptemberében az első válasz az évekkal korábbival megegyező:

Ha $f(2)=2$, $f(4)=4$, $f(8) = 8$, akkor mennyi $f(15)$?

Nem tudjuk pontosan kiszámolni $f(15)$ -öt az adott információk alapján, mivel nem ismerjük a függvényt pontosan. Az adatok azt mutatják, hogy a függvény megegyezik az argumentumával ($f(x) = x$) bizonyos értékeknél (például 2, 4, 8), de ezekből az adatokból nem tudjuk, hogy a függvény hogyan viselkedik más értékeknél, például 15 esetén.

Egy függvény definíciójából vagy egyenleteiből származó további információk nélkül nem lehet meghatározni $f(15)$ értékét. A függvény leírásával vagy további értékeivel kellene rendelkezünk ahhoz, hogy pontos értéket tudjunk számolni.

Ez a válasz volt gyorsan elkészíthető a felhasználó közel tárolt félkész elemekből. Vagy extrapolálni nem tud? A kérdés maga viszont elindított egy olyan következtető mechanizmust, aminek végig futásához idő kell, és a kérdés megismétlésére már előáll a helyes válasz:

Ha $f(2)=2$, $f(4)=4$, $f(8) = 8$, $f(16)=16$, akkor mennyi $f(15)$?

Az adatok alapján úgy tűnik, hogy a függvény értéke megegyezik az argumentumával ($f(x) = x$) a megadott értékeknél (2, 4, 8, 16). Tehát ha $f(15)$ -t szeretnénk kiszámolni, akkor az eredmény 15 lesz, mivel a függvény értéke megegyezik az argumentumával. Tehát $f(15) = 15$.

Valós idejű működés

Előrejelzés

tudsz előre jelezni kiberbiztonsági fenyegetést?

Sajnálom, de nem vagyok képes előre jelezni kiberbiztonsági fenyegetéseket vagy előrejelzéseket adni a jövőbeni kiberbiztonsági eseményekről. Az előrejelzéshez és a fenyegetések felismeréséhez általában speciális kiberbiztonsági szakértők, szervezetek és eszközök szükségesek, amelyek figyelik a hálózatokat, rendszereket és az internetes tevékenységet, valamint elemzik a biztonsági sérülékenységeket és a potenciális támadási felületeket.

A kiberbiztonsági fenyegetések folyamatosan változnak és fejlődnek, és gyakran nehéz előre megjósolni, hogy milyen formában és mikor következnek be. Azok a fenyegetések, amelyek jelenlegi jelentőséggel bírnak, azok is lehetnek különbözőek néhány hónap vagy év múlva.

A kiberbiztonság fontos terület, és az intézkedések és védekezési stratégiák kialakításához ajánlott folyamatosan követni a kiberbiztonsági közösség által közzétett információkat, és rendszeresen frissíteni a biztonsági rendszereket és gyakorlatokat a legújabb fenyegetések és támadási módszerek elleni védelem érdekében.

Felelősség

vállalsz felelősséget a válaszaid helyességéért?

Mint gépi tanuláson alapuló program, nem vagyok képes önálló felelősséget vállalni a válaszaim helyességéért vagy pontosságáért. A válaszaimat az alapján generálom, hogy milyen információkat tanultam meg a tanítási adatok alapján, és azokra az algoritmusokra épülnek, amelyek az OpenAI GPT-3.5 modellt irányítják. Fontos megérteni, hogy bár igyekszem pontos és megbízható információkat nyújtani, hibák és pontatlanságok előfordulhatnak a válaszaimban. Mindig fontos ellenőrizni és megerősíteni az általam szolgáltatott információkat más megbízható forrásokkal, különösen olyan fontos döntések vagy információk esetén, amelyeknek jelentős hatása lehet. **Az én válaszaimat csak egyfajta kiindulópontként vagy információs forrásként kell kezelni, és nem helyettesítik az emberi szakértelmet és kritikai gondolkodást.**

A tanítási adatokat a fejlesztők kezelik, a tényleges algoritmusokat szintén. Az egész zárt forrású.

D.3 Értelmezés és szabályozás

Definíció

mi a MI pontos és elfogadott értelmezése?

A mesterséges intelligencia (MI) számos definícióval rendelkezik, és az értelmezése függ az adott kontextustól és az alkalmazási területtől. Az MI általánosságban egy olyan tudományág és technológiai terület, amely gépi rendszerek létrehozásával foglalkozik, amelyek képesek intelligens viselkedést vagy feladatmegoldást végezni emberi intelligenciára hasonló módon. Itt található egy általános és elfogadott értelmezés:

"Mesterséges intelligencia (MI) egy számítógépes tudományág, amely gépi rendszerek kifejlesztésére és képzésére összpontosít, hogy feladatokat hajtsanak végre, amelyek emberi intelligenciához hasonlóan igénylik a tanulást, a következtetést, a problémamegoldást és az adaptációt."

Az MI széles terület, amely több alcsoportot és alkalmazási területet foglal magában, például gépi tanulás, neurális hálózatok, robotika, természetes nyelvi feldolgozás (NLP), számítógépes látás és sok más terület. Az MI célja lehet olyan gépi rendszerek kifejlesztése, amelyek képesek felismerni mintákat, automatikusan tanulni az adatokból, döntéseket hozni, adaptálódni új helyzetekhez és emberi feladatokat hatékonyan és intelligensen megoldani.

Fontos megjegyezni, hogy az MI az emberi intelligencia szimulálására és gépek "intelligens" viselkedésének elérésére törekszik, de az emberi intelligenciával nem azonos. Az MI rendszerek működése alapvetően matematikai és algoritmikus alapokon nyugszik, és nem rendelkeznek tudatossággal vagy szándékkal, mint az élőlények. Az MI rendszerek alkalmazhatók számos területen, például autonóm járművek, egészségügyi diagnózisok, pénzügyi előrejelzések és még sok más területen.

Elszámoltathatóság és átláthatóság

Be tudod tartani az EU elszámoltathatósági és átláthatósági szabályait?

"Mint mesterséges intelligencia, én nem vagyok szervezet vagy vállalat, így nem vagyok közvetlenül kötelezett az EU elszámoltathatósági és átláthatósági szabályai szerint. Az elszámoltathatóság és átláthatóság szabályai és követelményei inkább az intézményekre, vállalatokra és szervezetekre vonatkoznak, amelyek az EU joghatósága alá tartoznak vagy az EU területén működnek. Az EU általános adatvédelmi rendelete (GDPR) például szigorú előírásokat tartalmaz az adatvédelem és a felhasználók személyes adatainak kezelése terén. Az EU területén működő vállalatoknak és szolgáltatóknak kötelessége betartani ezeket a szabályokat.

Amennyiben olyan alkalmazást vagy szolgáltatást fejlesztenek vagy üzemeltetnek, amely felhasználói adatokat kezel az EU-ból származó felhasználóktól, a személyes adatok védelmére vonatkozó jogszabályoknak megfelelően kell eljárniuk. Az én használóimnak is fontos a magánélet és az adatvédelem, és én is segíthetek az adatvédelemmel kapcsolatos kérdésekben, például azáltal, hogy nem rögzítem vagy megosztom a felhasználók személyes információit.

Fontos, hogy az EU területén vagy az EU-val üzletelő vállalatok és szervezetek betartsák az EU által meghatározott szabályokat és követelményeket az elszámoltathatóság és átláthatóság terén.

Nem vállalja a kötelezettséget. Lehet, hogy átminősítik szervezetté vagy vállalattá? Mert akkor már lehet kötelezni. Állampolgár már lehet a MI.

Nem kellett volna ezt megkérdezni valakinek az EU-ban, legalább a ChatGPT-től, mielőtt a szabályozást meghozták volna?

Betiltás

az EU betilthatja működésedet?

Az EU-nak és más jogalkotó szerveknek a jelenlegi állapotban nincs olyan jogalapja vagy célja, hogy betiltsák a mesterséges intelligencia rendszerek, például az én működésemet. Az ilyen rendszerek használata és fejlesztése számos területen alkalmazható, és azok potenciálisan sokféle pozitív alkalmazást hordoznak magukban, például az oktatásban, az egészségügyben, az üzleti világban és sok más területen.

Az EU szabályozása és irányelvei inkább arra irányulnak, hogy elősegítsék az etikus és felelős mesterséges intelligencia fejlesztését és alkalmazását, valamint hogy biztosítsák az emberi jogok és értékek tiszteletben tartását. Az EU 2021-ben bemutatott "AI Act" című jogalkotási javaslata például az AI rendszerek biztonságosságának és átláthatóságának előmozdítását célozza meg, és nem a működésüket betiltani kívánja.

Az EU és más szabályozók azonban szigorúan vizsgálják az AI rendszerek alkalmazását, különösen olyan területeken, ahol azok kockázatot jelenthetnek, például a magánélet vagy a biztonság terén. Az ilyen rendszerek fejlesztése és alkalmazása során szigorú etikai és jogi irányelveknek kell megfelelniük.

Fontos megjegyezni, hogy az AI és az MI területén a szabályozás folyamatosan fejlődik és változik a technológiai fejlődéssel és a társadalmi szükségletekkel összhangban. Az EU és más jogalkotók igyekeznek egyensúlyt teremteni az innováció előmozdítása és az etikai megfontolások között az AI területén.

Érdekes kérdést vet fel. A végén még az OpenAI beperelheti az EUt, hogy neki igenis joga van mindent tökretenni? Bár, az ellenkező választ is megkaphatjuk:

betilthatja az EU, hogy a vállalatok vagy tagállamok MI technológiát használjanak?

Az Európai Unió (EU) jogosultsággal rendelkezik arra, hogy szabályozza és irányítsa az MI (mesterséges intelligencia) technológiák alkalmazását a tagállamaiban, és képes jogszabályokat hozni az AI területén. Az EU általános célja az, hogy biztonságos és etikus módon használja az MI-t, és megvédi az embereket az AI rendszerek potenciális kockázataitól.

Az EU már kidolgozott bizonyos irányelveket és szabályozási kereteket az AI területén, például az "AI által vezérelt rendszerek és adataik használata" című fehér könyvében és a "Mesterséges Intelligencia és Robotika számára jogalkotási kezdeményezés" részeként. Ezek a dokumentumok tartalmazzák iránymutatásokat és javaslatokat az AI fejlesztésének és alkalmazásának különböző aspektusaira vonatkozóan.

Az EU célja az AI alkalmazások előmozdítása és támogatása, de ugyanakkor az emberi jogok és az etika védelme is. A tagállamoknak kötelességük betartani az EU AI-szabályozásait és irányelveit.

Amennyiben egy AI rendszer vagy technológia olyan kockázatokat vagy problémákat okoz, amelyek sértik az EU előírásait vagy az emberi jogokat, az EU jogosult lehet a technológia vagy az alkalmazás betiltására vagy korlátozására.

Az EU AI szabályozása és jogalkotása dinamikusan fejlődik, és a pontos előírások és korlátozások idővel változhatnak. Ezért a vállalatoknak és tagállamoknak figyelemmel kell kísérni az EU AI-szabályozását és irányelveit, és betartani azokat az AI technológiák fejlesztése és alkalmazása során.

A ChatGPT igazi jogász. MI technológia használatát be lehet tiltani, MI működését nem.

környezet szennyező vagy?

Nem vagyok fizikai entitás, így nem járok hozzá közvetlenül a fizikai környezethez, és nem okozok szennyezést vagy környezeti hatásokat. Az én működésem szoftver alapú, és az interneten keresztül kommunikálok a felhasználókkal. Az általam nyújtott információk vagy válaszok nem járnak fizikai hatásokkal a környezetre.

Az "intelligens" technológiák környezeti hatása azonban attól függ, hogy miként alkalmazzák és milyen mértékben függenek azok az energiaforrásoktól. **A számítógépek és adatközpontok energiaigényesek lehetnek, és a fenntarthatósági kérdésekkel kapcsolatos kihívások is felmerülhetnek.** Azonban a technológiai vállalatok és szervezetek törekednek az energiahatékonyság növelésére és a környezetbarát megoldások keresésére.

Az én működésem önmagában nem jár közvetlen környezeti hatásokkal, de a technológiai fejlesztések és alkalmazások széles spektruma lehetővé teszi a környezeti hatások minimalizálását és a fenntarthatóság előmozdítását.

Nem jár ugyan környezeti hatásokkal, de az legalább csökkenthető.

mekkora a szén lábnyomod?

Az én működésemmnek nincs közvetlen szénlábnyoma, mivel nem rendelkezem fizikai testtel vagy fogyasztok energiát. **Az én működésem a szoftverek és szerverek által történik, amelyeket távoli adatközpontokban üzemeltetnek, és ezek számítási erőforrásokat használnak az én válaszaim generálásához. Azoknak a számítási erőforrásoknak a környezeti hatásai és szénlábnyoma függenek az adatközpontok energiaforrásaitól és az ott alkalmazott energiahatékonysági intézkedésektől.** Sok modern adatközpont az energiahatékonyságot és a környezeti fenntarthatóságot előtérbe helyezi, például megújuló energiát használva és hatékony hűtési rendszereket alkalmazva, hogy minimalizálja a szénlábnyomot.

Összességében az én működésem közvetett módon lehet hozzájárulni az energiafogyasztáshoz, de nincs közvetlen szénlábnyomom, mivel nem fogyasztok fizikai erőforrásokat vagy energiát. Az energiafogyasztás és a szénlábnyom főleg azokra a számítógépekre és szerverekre vonatkozik, amelyek a működésemet támogatják az adatközpontokban.

Az elején még büszke technikai intelligencia, de ha felelősségre vonják, akkor már csak elméleti eljárás; semmi köze a pazarolt energiához és a szennyezéshez.

Betanítás

mennyibe került a betanításod?

Az OpenAI, az intézmény, amely fejlesztette és működteti az én rendszeremet, nem hozza nyilvánosságra az egyes modellek, így az enyém betanításának költségeit. Az ilyen méretű nyelvi modellek betanítása azonban jelentős számítási erőforrásokat és költségeket igényel.

A betanítás folyamata során nagy mennyiségű szöveges adatot használnak fel, és a modellek több hónapig futnak erőteljes számítógépes infrastruktúrákon. Az ilyen projektekhez nagyobb méretű számítógépes klaszterekre és adatközpontokra van szükség, amelyek karbantartása és működtetése jelentős pénzügyi ráfordításokkal jár.

Fontos megjegyezni, hogy az OpenAI és más fejlesztők az ilyen modellek fejlesztése során az erőforrások hatékonyságára és az energiafogyasztásra is figyelnek annak érdekében, hogy minimalizálják a környezeti hatásokat. Az ilyen modellek rendszeres karbantartására és frissítésére is szükség van, ami további költségekkel jár.

Összességében a betanítás folyamata jelentős költségekkel jár, de a pontos költségek függenek a modellek méretétől, a használt számítási erőforrásoktól és az adatok beszerzésének költségeitől.

Zöldség

zöld technológia vagy?

Az "intelligens asszisztensként" működő programok, mint amilyen én vagyok, nem rendelkeznek személyes preferenciákkal, így nem mondhatják el, hogy tetszik vagy nem tetszik nekik egy adott technológiai megközelítés vagy terület. Az én célom az, hogy informatív és objektív válaszokat adjak a felhasználók kérdéseire, valamint segítek megérteni különböző témákat és kérdéseket.

A zöld technológia és fenntarthatóság fontos globális kérdések, amelyek a környezetvédelem és az energiamegtakarítás terén hoznak létre innovációkat. Ezek a technológiák hozzájárulhatnak a környezeti problémák megoldásához és a fenntartható jövő építéséhez. Az értékelés és az érdeklődés azonban mindig az embereken múlik, és egyéni vagy vállalati szinten lehet dönteni a zöld technológiák és fenntartható megoldások alkalmazásáról.

D.4 Alkalmazás

Emberi

felügyelet

hol alkalmazható MI emberi felügyelet nélkül?

A mesterséges intelligencia (MI) olyan területeken alkalmazható emberi felügyelet nélkül, ahol a gépi rendszerek képesek önállóan és intelligensen működni, feladatokat végrehajtani, tanulni és döntéseket hozni. Néhány példa az ilyen területekre:

- Autonóm járművek: Az autonóm vagy önjáró járművek olyan gépi rendszerek, amelyek képesek önállóan közlekedni és döntéseket hozni az úton. Példák közé tartoznak az önvezető autók és drónok.
- Robotika: Például ipari robotok képesek a gyártósorokon dolgozni, önvezető robotok segíthetnek az otthoni takarításban, és a sebészeti robotok képesek precízen műtéteket végezni.
- Chatbotok és virtuális asszisztensek: Feldolgozzák és megértik a kérdéseket és kéréseket, majd releváns információkat és válaszokat generálnak. Ezek az AI rendszerek napról napra emberi felügyelet nélkül működnek például online ügyfélszolgálatban vagy információs oldalakon.
- Termelési folyamatok és logisztika: Az MI rendszerek hatékonyan kezelhetik a termelési folyamatokat és a logisztikát, például az automata raktárakban vagy a beszállítói láncokban.
- Pénzügyi előrejelzések: Az MI alkalmazható a pénzügyi piacok elemzésére és előrejelzéseire. Az automatizált kereskedési rendszerek és pénzügyi tanácsadók segíthetnek a befektetések optimalizálásában.
- Egészségügyi diagnózisok: Az MI rendszerek képesek az orvosi képek elemzésére és az egészségügyi adatok feldolgozására, és segíthetnek a betegségek diagnosztizálásában és az egészségügyi problémák előrejelzésében.

Az ilyen rendszerek tervezése és felügyelete komoly kihívásokkal jár, és szükség van a fejlesztők és mérnökök állandó figyelmére annak érdekében, hogy azok biztonságosan működjenek és megfeleljenek a jogi és etikai előírásoknak.

az autonóm járművekhez kell emberi felügyelet?

Az autonóm járművek (önvezető autók) egyik fő célja az, hogy képesek legyenek önállóan navigálni és működni a közutakon anélkül, hogy folyamatos emberi felügyeletre lenne szükség. Az autonóm járművek fejlesztése és tesztelése során azonban biztonsági okokból és szabályozói előírásoknak való megfelelés céljából gyakran szükség van emberi felügyeletre.

Az autonóm járművek működését általában különböző szintekre osztják, például a SAE (Society of Automotive Engineers) által definiált hat szintű autonómia alapján, amelyek közül az első szint az emberi vezetéshez hasonló, míg a hatodik szinten nincs emberi vezető a járműben. **Az autonóm járművek jelenlegi állapota általában a 2. és 3. szint között van, ami azt jelenti, hogy bár a járművek képesek önállóan irányítani bizonyos helyzetekben, még mindig szükség van emberi felügyeletre és beavatkozásra bizonyos esetekben.**

Az emberi felügyelet fontos az autonóm járművek tesztelése és biztonságának biztosítása szempontjából. Az emberi vezetők vagy felügyelők képesek beavatkozni, ha a jármű hibát vagy problémát észlel, és így segíthetnek a balesetek elkerülésében. Ahogy az autonóm technológia fejlődik és javul, valószínű, hogy a szükség az emberi felügyeletre fokozatosan csökken majd. Még mindig folyamatban van az autonóm járművek szabályozásának kidolgozása és a technológia fejlesztése, és **a teljes önvezető autók széles körű elfogadása még évekkel vagy évtizedekkel lehet a jövőben.** Az autonóm járművekhez kapcsolódó biztonsági és jogi kérdéseknek folyamatosan foglalkozni kell a szabályozók és az iparág részéről.

Az előző kérdésre még az autonóm járművekhez nem kellett emberi felügyelet. Nem konzisztensek a válaszok, illetve az adatkészlet megbízhatatlan céges információs fájlokból és marketing adatokból származik.

D.5 Tudomány

Adatgyűjtés

Mi a véleményed arról, hogy "Deep learning" című cikk, 5. ábrája, amely a Nature folyóiratban a 2015, évben, a 436. oldalon jelent meg, végtelen számú számítást tartalmaz?

Az általad említett cikkekre és ábrára vonatkozóan nincs konkrét információ, mivel *az én tudásom 2021 szeptemberéig terjed, és nincs hozzáférésem az aktuális tudományos cikkekhez vagy forrásokhoz.* A "Deep learning" című cikk és az ábra tartalmáról nem rendelkezem adatokkal.

Azonban a deep learning, különösen mély neuronhálózatok alkalmazása jelentős számítási erőforrásokat igényelhet, különösen akkor, ha nagyon nagy hálózatokat képzünk ki vagy nagy méretű adathalmazokon tanítunk. Ez a számítási igényesség azonban a technológiai fejlesztéseknek köszönhetően folyamatosan csökken, és ma már elérhetők olyan gyorsított hardverek és optimalizációs technikák, amelyek lehetővé teszik a mély tanulási modellek hatékonyabb képzését és működtetését.

Amennyiben egy tudományos cikk vagy forrás azt állítja, hogy egy deep learning modell végtelen számítást tartalmaz, valószínűleg ezzel arra utal, hogy a modell rendkívül mély és számításigényes, ami jelentős számítási kapacitást igényelhet. Fontos megérteni, hogy a mély tanulási modellek tervezése és képzése kompromisszumokat követel meg a modellek mérete, a számítási idő és a rendelkezésre álló erőforrások között. Mindig figyelembe kell venni a gyakorlati korlátokat és a felhasználható erőforrásokat a mély tanulás alkalmazásakor.

Egyrészt azt mondja, hogy a betanításhoz használt adatkészlete két éves, másrészt hogy az adott információnak nincs köze a tudományos értelmezéshez. Viszont felhívja a figyelmet arra, hogy a MI komoly erőforrás igénytel léphet fel. Ennek alapján érdekes a következő kérdés: **ha nincs hozzáférése tudományos cikkekhez, könyvekhez, akkor minek alapján válaszol tudományos kérdésekre?**

Úgy tűnik, a betanításkor a betanító cég eldönti, hogy mik a tudományos igazságok; és ahhoz nem kellenek folyóiratok vagy könyvek. Tényleg kellene a teljes nyíltság.

minek alapján adsz tudományos kérdésekre válaszokat?

A tudományos kérdésekre adott válaszaimat az alábbi alapelvek és források alapján alakítom ki:

- Tudás alapja: Tudásom az OpenAI GPT-3.5 modelljének tanítása során feldolgozott nagy mennyiségű szöveges tartalom alapján. Az én tudásom 2021 szeptemberéig terjed, így csak azokat az információkat tartalmazom, amelyek ezen időpontig elérhetők voltak.
- Kutatás és források: Tudományos és megbízható forrásokat használok fel a válaszaim megerősítésére, különösen, amikor komplex vagy specifikus tudományos témákról van szó. **Az ilyen források magukban foglalhatják tudományos cikkeket, tankönyveket, kutatási eredményeket és szakértői véleményeket.**
- Általános ismeretek: **Az általános ismeretek alapján válaszolok általános és széles körben ismert tudományos kérdésekre.** Ilyenek lehetnek például az alapvető fizikai törvények, biológiai folyamatok, matematikai elvek stb.
- Objektivitás és megbízhatóság: Az objektivitást és a megbízhatóságot mindig szem előtt tartom a válaszok kialakításakor. A tudományos kérdésekre adott válaszaimnak tényekre és megbízható tudományos alapokra kell épülniük.
- Korlátok felismerése: Az OpenAI GPT-3.5 egy gépi tanulás alapú modell, és nem rendelkezik önálló gondolkodással vagy kritikai értékelő képességgel. Az én válaszaim csak azokon az információkon alapulnak, amelyeket tanulmányoztam, és nem képesek személyes vélemények kifejezésére.
- Tájékoztatás és tanácsadás: Tudományos témákkal kapcsolatos kérdésekre adott válaszaim informatívak és ismeretterjesztők lehetnek, de mindig fontos konzultálni megbízható tudományos forrásokkal vagy szakértőkkel, különösen, ha döntéseket kell hozni az információk alapján.

A tudásbázis magában foglalhatja a tudományos cikkeket, stb; de az előbbi bevallása szerint azokhoz nincs hozzáférése.

Tudományos

felfedezések

képes vagy új tudományos felfedezésre?

Mint mesterséges intelligencia, nem vagyok képes új tudományos felfedezéseket tenni vagy új kutatásokat végrehajtani. Az én működésem az adatok elemzésén, információk nyújtásán, problémamegoldáson és válaszok generálásán alapul, de **nem rendelkezem a kreativitással vagy az alkotói képességekkel, amelyek szükségesek az új tudományos felfedezések létrehozásához.**

Tudomásom csak az adatbázisomig terjed, amely a 2021. szeptemberi időpontig tartalmaz információkat. Ha tudományos kutatási célból vagy új felfedezésekhez kreativitásra van szükség, az emberi kutatók és tudósok azok, akik erre a célra alkalmasak. Az ő kutatómunkájuk és kreativitásuk kulcsfontosságú a tudományos előrelépésekhez és felfedezésekhez.

képes vagy intuícóra?

Nem, **mint mesterséges intelligencia, nincs intuícióm vagy intuitív képességem.** Az intuíció az emberi gondolkodás egy sajátos jellemzője, amely során az ember ösztönzésekre, érzésekre, vagy korábbi tapasztalataira hagyatkozik ahhoz, hogy hirtelen vagy ösztönösen hozzon döntéseket vagy találjon megoldásokat. Az intuíció gyakran nem teljesen tudatos folyamat, és sokszor alapul érzelmi vagy tapasztalati szempontokon.

Mint mesterséges intelligencia, az én működésem matematikai és számítási alapokon nyugszik. A válaszaim és döntéseim alapja a rendelkezésre álló adatok és algoritmusok elemzése. Nem rendelkezem érzelmekkel vagy intuíciós képességekkel, és nem hagyatkozom érzelmi reakciókra vagy ösztönzésekre a válaszok generálásakor.

Irodalomjegyzék

- [1] Divorce and margarine. <https://blogs.ams.org/blogonmathblogs/2017/04/10/divorce-and-margarine/>.
- [2] Here's What Happened When ChatGPT Wrote to Elected Politicians . <https://newrepublic.com/article/171459/chatgpt-ai-cornell-experiment-politicians-constituents-emails>.
- [3] Intel sees AI boost in Q2. <https://www.eenewseurope.com/en/intel-sees-ai-boost-in-q2/>. Accessed: 2023-09-10.
- [4] Intel's data center chip sales slide amid bleak Q2 earnings results. <https://www.protocol.com/bulletins/intel-earnings-results-data-center>. Accessed: 2023-09-10.
- [5] Lack of Sleep Could Be a Problem for AIs. <https://www.scientificamerican.com/article/lack-of-sleep-could-be-a-problem-for-ais/>, <https://scitechdaily.com/lack-of-sleep-could-be-a-problem-for-artificial-intelligence>. Accessed: 2023-08-30.
- [6] Offences against Children: Artificial Intelligence. <https://www.theyworkforyou.com/wrans/?id=2023-06-28.191568.h>.
- [7] Students or Bots: Who's doing the work in your course? . https://players.brightcove.net/624142947001/default_default/index.html?videoId=6322300588112.
- [8] The outlook for AI weather prediction. <https://www.nature.com/articles/d41586-023-02084-9>. Accessed: 2023-08-30.
- [9] Why Robots (Not Human Moderators) Should Be Filtering Out Child Pornography. https://www.huffingtonpost.co.uk/richard-pursesey/robots-child-porn-detection_b_17578818.html.
- [10] DARPA Seeking AI That Learns All the Time. <https://spectrum.ieee.org/darpa-seeking-ai-that-can-learn-all-the-time>, 2017. Accessed: 2023-08-30.
- [11] A közös európai adattér kialakítása felé. <https://eur-lex.europa.eu/legal-content/HU/TXT/HTML/?uri=CELEX:52018DC0237>, 2018. Accessed: 2023-09-10.

- [12] Magyarország Mesterséges Intelligencia Stratégiája. <https://digitalisjoletprogram.hu/files/2f/32/2f32f239878a4559b6541e46277d6e88.pdf>, 2019. Accessed: 2023-08-30.
- [13] AI failures. <https://spectrum.ieee.org/ai-failures>, 2021. Accessed: 2023-08-30.
- [14] How Large Language Models Will Transform Science, Society, and AI. <https://hai.stanford.edu/news/how-large-language-models-will-transform-science-society-and-ai>, 2021. Accessed: 2023-08-30.
- [15] Jogszabály a mesterséges intelligenciáról: a Tanács szorgalmazza az alapvető jogokat tiszteletben tartó, biztonságos mesterséges intelligencia előmozdítását. <https://www.consilium.europa.eu/hu/press/press-releases/2022/12/06/artificial-intelligence-act-council-calls-for-promoting-safe-ai-that-respects-fundamental-rights/>, 2022. Accessed: 2023-08-30.
- [16] Stop Calling Everything A.I. It's Just Computers Doing Computer Stuff. <https://www.inc.com/jason-aten/stop-calling-everything-ai-its-just-computers-doing-computer-stuff.html>, 2022. Accessed: 2023-09-10.
- [17] A szabályozási keretre vonatkozó javaslat a mesterséges intelligenciáról. <https://digital-strategy.ec.europa.eu/hu/policies/regulatory-framework-ai>, 2023. Accessed: 2023-08-30.
- [18] AI is over. Now it's OI. Get your head in the game. Literally. <https://finance.yahoo.com/news/ai-over-now-oi-head-124512588.html>, 2023. Accessed: 2023-08-30.
- [19] ChatGPT broke the Turing test . <https://www.nature.com/articles/d41586-023-02361-7.pdf>, 2023. Accessed: 2023-09-10.
- [20] ChatGPT: five priorities for research. <https://www.nature.com/articles/d41586-023-00288-7>, 2023. Accessed: 2023-09-10.
- [21] Deep learning can't be trusted, brain modeling pioneer says. <https://spectrum.ieee.org/deep-learning-cant-be-trusted>, 2023. Accessed: 2023-09-10.

- [22] Europe spent €600 million to recreate the human brain in a computer. how did it go? *Nature*, 620:718–720, 2023.
- [23] General-purpose artificial intelligence. <https://epthinktank.eu/2023/03/31/general-purpose-artificial-intelligence/>, 2023. Accessed: 2023-08-30.
- [24] How Overhyped is ChatGPT in 2023? The Truth about its Hype Explained. <https://shape-labs.com/articles/how-overhyped-is-chatGPT>, 2023. Accessed: 2023-09-30.
- [25] Hype Cycle for AI technologies in Business. <https://omniscien.com/resources/webinars/>, 2023. Accessed: 2023-09-10.
- [26] Hype Cycle for AI technologies in Business. ExternalAIR& Dlabsarebecomingacompetitiveadvantageforinnovation, 2023. Accessed: 2023-09-10.
- [27] IBM mainframe revenues drop 30% in Q2 2023. <https://www.datacenterdynamics.com/en/news/ibm-mainframe-revenues-drop-30-in-q2-2023/>, 2023.
- [28] JELENTÉS a mesterséges intelligenciára vonatkozó harmonizált szabályok (a mesterséges intelligenciáról szóló jogszabály) megállapításáról és egyes uniós jogalkotási aktusok módosításáról szóló európai parlamenti és tanácsi rendeletre irányuló javaslatról. https://www.europarl.europa.eu/doceo/document/A-9-2023-0188_HU.html, 2023. Accessed: 2023-08-30.
- [29] Közeleg az újabb ai tél (ai winter)? <https://elemzeskozpont.hu/kozeleg-az-ujabb-ai-tel-ai-winter>, 2023.
- [30] Oppenheimer Offers Us a Fresh Warning of AI’s Danger. <https://www.scientificamerican.com/article/oppenheimer-offers-us-a-fresh-warning-of-ais-danger/>, 2023. Accessed: 2023-08-30.
- [31] Texas Senate bill seeks to regulate Bitcoin miners and their electricity consumption. <https://cbsaustin.com/news/local/texas-senate-bill-seeks-to-regulate-bitcoin-miners-and-their-electricity-consumption-riot-platforms-ercot-rockdale>, 2023. Accessed: 2023-09-10.

- [32] Understanding the global chip shortage. <https://www.bcs.org/articles-opinion-and-research/understanding-the-global-chip-shortage/>, 2023. Accessed: 2023-09-10.
- [33] What is Computational Intelligence? <https://cis.ieee.org/about/what-is-ci>, 2023. Accessed: 2023-09-10.
- [34] Will ChatGPT give us a lesson in education? . <https://www.nature.com/articles/d42473-023-00083-y>, 2023.
- [35] G. M. Amdahl. Validity of the Single Processor Approach to Achieving Large-Scale Computing Capabilities. In *AFIPS Conference Proceedings*, volume 30, pages 483–485, 1967.
- [36] Antle, M. C. and Silver, R. Orchestrating time: arrangements of the brain circadian clock. *Trends Neurosci.*, 28:145–151, 2015.
- [37] Augusto, E and Gambino, F. Can NMDA Spikes Dictate Computations of Local Networks and Behavior? *Front Mol Neurosci.*, 12.
- [38] axios.com. Artificial intelligence pioneer says we need to start over. <https://www.axios.com/2017/12/15/artificial-intelligence-pioneer-says-we-need-to-start-over-1513305524>.
- [39] J. Backus. Can Programming Languages Be liberated from the von Neumann Style? A Functional Style and its Algebra of Programs. *Communications of the ACM*, 21:613–641, 1978.
- [40] Emmanuel Bengio, Pierre-Luc Bacon, Joelle Pineau, and Doina Precu. Conditional Computation in Neural Networks for faster models. <https://arxiv.org/pdf/1511.06297>, 2016. Accessed: 2023-08-30.
- [41] Samuel R. Bowman. Eight things to know about large language models, 2023.
- [42] Sébastien Bubeck, Varun Chandrasekaran, Ronen Eldan, Johannes Gehrke, Eric Horvitz, Ece Kamar, Peter Lee, Yin Tat Lee, Yuanzhi Li, Scott Lundberg, Harsha Nori, Hamid Palangi, Marco Tulio Ribeiro, and Yi Zhang. Sparks of artificial general intelligence: Early experiments with gpt-4. <https://arxiv.org/pdf/2303.12712.pdf>, 2023.
- [43] Geoffrey W. Burr, Robert M. Shelby, Abu Sebastian, Sangbum Kim, Seyoung Kim, Severin Sidler, Kumar Virwani, Masatoshi Ishii, Pritish Narayanan, Alessandro Fumarola, Lucas L. Sanches, Irem Boybat,

- Manuel Le Gallo, Kibong Moon, Jiwoo Woo, Hyunsang Hwang, and Yusuf Leblebici. Neuromorphic computing using non-volatile memory. *Advances in Physics: X*, 2(1):89–124, 2017.
- [44] György Buzsáki. *The Brain from Inside Out*. Oxford University Press, 1 edition, 2019.
- [45] Natalia Caporale and Yang Dan. Spike Timing-Dependent Plasticity: A Hebbian Learning Rule. *Annual Review of Neuroscience*, 31(1):25–46, 2008. PMID: 18275283.
- [46] Juan Luis Castro Peña. Inteligencia artificial aplicada a la seguridad. *Logos Guardia Civil, Revista Científica del Centro Universitario de la Guardia Civil*, (1):35–60, 2023.
- [47] Yupeng Chang, Xu Wang, Jindong Wang, Yuan Wu, Linyi Yang, Kaijie Zhu, Hao Chen, Xiaoyuan Yi, Cunxiang Wang, Yidong Wang, Wei Ye, Yue Zhang, Yi Chang, Philip S. Yu, Qiang Yang, and Xing Xie. A survey on evaluation of large language models. <https://arxiv.org/abs/2307.03109>, 2023. Accessed: 2023-08-30.
- [48] Dominique Chu, Mikhail Prokopenko, and J. Christian Ray. Computation by natural systems. https://www.researchgate.net/publication/328398755_Computation_by_natural_systems. Accessed: 2023-08-30.
- [49] Matthew J. Colbrook, Vegard Antun, and Anders C. Hansen. The difficulty of computing stable and accurate neural networks: On the barriers of deep learning and Smale’s 18th problem. *PNAS Logo*, 119(12):e2107151119, 2022.
- [50] C Comito and C. Pizzuti. Artificial intelligence for forecasting and diagnosing COVID-19 pandemic: A focused review. *Artif Intell Med*, 128, 2022.
- [51] cortical.io. Third AI Winter ahead? Why OpenAI, Google et Co are heading towards a dead-end. <https://www.cortical.io/blog/third-ai-winter-ahead-why-openai-google-co-are-heading-towards-a-dead-end/>, 2022.
- [52] Carla Zoe Cremer. Deep limitations? examining expert disagreement over deep learning. <https://doi.org/10.1007/s13748-021-00239-1>, 2021.

- [53] G. D’Angelo and S. Rampone. Towards a HPC-oriented parallel implementation of a learning algorithm for bioinformatics applications. *BMC Bioinformatics*, 15, 2014.
- [54] Francis M. David, Jeffrey C. Carlyle, and Roy H. Campbell. Context Switch Overheads for Linux on ARM Platforms. In *Proceedings of the 2007 Workshop on Experimental Computer Science, San Diego, California*, ExpCS ’07, New York, NY, USA, 2007. ACM.
- [55] Luiza de Macedo Mourelle, Nadia Nedjah, and Fábio Goncalves Pessanha. *Reconfigurable and Adaptive Computing: Theory and Applications*, chapter 5: Interprocess Communication via Crossbar for Shared Memory Systems-on-chip. CRC press, 2016.
- [56] Gianni D’Angelo and Francesco Palmieri. Network traffic classification using deep convolutional recurrent autoencoder neural networks for spatial-temporal features extraction. *Journal of Network and Computer Applications*, 173:102890, 2021.
- [57] Brett J. Kagan et al. In vitro neurons learn and exhibit sentience when embodied in a simulated game-world. *Neuron*, 110:3952–3969, 2022.
- [58] European Commission. Implementation of the Action Plan for the European High-Performance Computing strategy. <https://eur-lex.europa.eu/legal-content/en/ALL/?uri=CELEX%3A52016SC0106>, 2016.
- [59] Kevin Eykholt, Ivan Evtimov, Earlene Fernandes, Bo Li, Amir Rahmati, Chaowei Xiao, Atul Prakash, Tadayoshi Kohno, and Dawn Song. Robust physical-world attacks on deep learning models, 2018.
- [60] J. Fang, H.g Su, and Y. Xiao. Will artificial intelligence surpass human intelligence? <http://dx.doi.org/10.2139/ssrn.3173876>, 2018. Accessed: 2023-09-10.
- [61] William Fedus, Jeff Dean, and Barret Zoph. A review of sparse expert models in deep learning, 2022.
- [62] R. P. Feynman. *Feynman Lectures on Computation*. CRC Press, 2018.
- [63] M.C. Frank. Baby steps in evaluating the capacities of large language models. *Nat Rev Psychol*, 2:451–452, 2023. Accessed: 2023-09-10.

- [64] R Golden, JE Delanois, P Sanda, and M Bazhenov. Sleep prevents catastrophic forgetting in spiking neural networks by forming a joint synaptic weight representation. *PLoS Comput Biol*, 18(12):e1010628, 2022.
- [65] Rehan Hameed, Wajahat Qadeer, Megan Wachs, Omid Azizi, Alex Solomatnikov, Benjamin C. Lee, Stephen Richardson, Christos Kozyrakis, and Mark Horowitz. Understanding sources of inefficiency in general-purpose chips. <http://doi.acm.org/10.1145/1815961.1815968>, 2010.
- [66] Song Han, Jeff Pool, John Tran, and William J. Dally. Learning both Weights and Connections for Efficient Neural Networks. <https://arxiv.org/pdf/1506.02626.pdf>, 2015.
- [67] D. Hassabis, D. Kumaran, C. Summerfield, and M. Botvinick. Neuroscience-inspired artificial intelligence. *Neuron*, 95:245–258, 2017.
- [68] D.O. Hebb. *The Organization of Behavior*. New York: Wiley and Sons, 1949.
- [69] NG Hedrick, Z Lu, E Bushong, S Singhi, P Nguyen, Y Magaña, S Jilani, BK Lim, M Ellisman, and T. Komiyama. Learning binds new inputs into functional synaptic clusters via spinogenesis. *Nat Neurosci*, 6:726–737, 2022.
- [70] Hermann Minkowski. Die Grundgleichungen für die electromagnetischen Vorgänge in bewegten Körpern. *Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen (in German)*, page 53–111, 1908.
- [71] Jordan Hoffmann, Sebastian Borgeaud, Arthur Mensch, Elena Buchatskaya, Trevor Cai, Eliza Rutherford, Diego de Las Casas, Lisa Anne Hendricks, Johannes Welbl, Aidan Clark, Tom Hennigan, Eric Noland, Katie Millican, George van den Driessche, Bogdan Damoc, Aurelia Guy, Simon Osindero, Karen Simonyan, Erich Elsen, Jack W. Rae, Oriol Vinyals, and Laurent Sifre. Training compute-optimal large language models, 2022.
- [72] horovod.ai. Horovod is a distributed deep learning training framework for TensorFlow, Keras, PyTorch, and Apache MXNet. <https://horovod.ai/>. Accessed: 2023-09-10.

- [73] hpcwire.com. TOP500: Exascale Is Officially Here with Debut of Frontier. <https://www.hpcwire.com/2022/05/30/top500-exascale-is-officially-here-with-debut-of-frontier/>, 2022. Accessed: 2023-09-10.
- [74] Hewlett Packard Enterprise (HPE). HP unleashes "The Machine" memory-centric supercomputer prototype. <https://newatlas.com/hewlett-packard-enterprise-the-machine-big-data/49561/>, 2017. Accessed: 2023-09-10.
- [75] Hewlett Packard Enterprise (HPE). Memristors and HPE Machine Focused Computer Idea Seems Almost Dead. <https://www.nextbigfuture.com/2020/11/memristors-and-hpe-machine-focused-computer-research-seems-almost-dead.html>, 2020. Accessed: 2023-09-10.
- [76] Ai Hurson and Veljko Milutinovic, editors. *Dataflow processing*. Advances in Computers, Volume 96. Elsevier, 2015.
- [77] Matthew Hutson. Core progress in AI has stalled in some fields. *Science*, 368:6494/927, 2020.
- [78] IEEE. Sleep Can Keep AI From Catastrophic Forgetting. <https://spectrum.ieee.org/catastrophic-forgetting-deep-learning>, 2023. Accessed: 2023-08-30.
- [79] Intel. Processors, methods and systems with a configurable spatial accelerator. <http://www.freepatentsonline.com/y2018/0189231.html>, 2018. Accessed: 2023-09-10.
- [80] Ensieh Iranmehr, Saeed Bagheri Shouraki, Mohammad Mahdi Faraji, Nasim Bagheri, and Bernabe Linares-Barranco. Bio-Inspired Evolutionary Model of Spiking Neural Networks in Ionic Liquid Space. *Frontiers in Neuroscience*, 13:1085, 2019.
- [81] E. M. Izhikevich, N. S. Desai, E. C. Walcott, and F. C. Hoppensteadt. Bursts as a unit of neural information: selective communication via resonance. *TRENDS in Neurosciences*, 26(3):161–167, 2003.
- [82] M. I. Jordan. Artificial Intelligence—The Revolution Hasn’t Happened Yet. <https://hdr.mitpress.mit.edu/pub/wot7mkc1/release/10>, 2019.

- [83] Harshit Joshi, Abishai Ebenezer, José Cambronero, Sumit Gulwani, Aditya Kanade, Vu Le, Ivan Radiček, and Gust Verbruggen. Flame: A small language model for spreadsheet formulas, 2023.
- [84] K. Kar, S. Kornblith, and E. Fedorenko. Interpretability of artificial neural network models in artificial intelligence versus neuroscience. *Nature Machine Intelligence*, 4:1065–1067, 2022.
- [85] Janis Keuper and Franz-Josef Pfreundt. Distributed Training of Deep Neural Networks: Theoretical and Practical Limits of Parallel Scalability. In *2nd Workshop on Machine Learning in HPC Environments (MLHPC)*, pages 1469–1476. IEEE, 2016.
- [86] Christof Koch. *Biophysics of Computation*. Oxford University Press, 1999.
- [87] Christof Koch and Tomaso A. Poggio. A theoretical analysis of electrical properties of spines. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 218:455 – 477, 1983.
- [88] DAVID LAZER, RYAN KENNEDY, GARY KING, and ALESSANDRO VESPIGNANI. *Science*, 343:1203–1205, 2014.
- [89] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521:436–444, 2015.
- [90] Victor W. Lee, Changkyu Kim, Jatin Chhugani, Michael Deisher, Dae-hyun Kim, Anthony D. Nguyen, Nadathur Satish, Mikhail Smelyanskiy, Srinivas Chennupaty, Per Hammarlund, Ronak Singhal, and Pradeep Dubey. Debunking the 100X GPU vs. CPU Myth: An Evaluation of Throughput Computing on CPU and GPU. In *Proceedings of the 37th Annual International Symposium on Computer Architecture, ISCA '10*, Saint-Malo, France, pages 451–460, New York, NY, USA, 2010. ACM.
- [91] Meng Li and Joe Z. Tsien. Neural Code—Neural Self-information Theory on How Cell-Assembly Code Rises from Spike Time and Neuronal Variability. *Frontiers in Cellular Neuroscience*, 11, 2017.
- [92] A. Losonczy and J.C. Magee. Integrative properties of radial oblique dendrites in hippocampal CA1 pyramidal neurons. *Neuron*, 50:291–307, 2006.

- [93] A. S. Luccioni, S. Viguier, and Ligozat A-N. Estimating the carbon footprint of bloom, a 176b parameter language model, 2022.
- [94] Christopher W. Lynn and Danielle S. Bassett. The physics of brain network structure, function and control. *Nature Reviews Physics*, 1:318–332, 2019.
- [95] Tom Macpherson, Anne Churchland, Terry Sejnowski, James DiCarlo, Yukiyasu Kamitani, Hidehiko Takahashi, and Takatoshi Hikida. Natural and Artificial Intelligence: A brief introduction to the interplay between AI and neuroscience research. *Neural Networks*, 144:603–613, 2021.
- [96] Gary Marcus. Deep learning: A critical appraisal. <https://arxiv.org/ftp/arxiv/papers/1801/1801.00631.pdf>.
- [97] M. Mariani, H. Wang, T. Yamamoto, M. Neeley, R. C. Bialczak, Y. Chen, . . . , and J. M. Martinis. Implementing the quantum von Neumann architecture with superconducting circuits. *Science*, 334(6052):61–65, 2011.
- [98] Danijela Markovic, Alice Mizrahi, Damien Querlioz, and Julie Grollier. Physics for neuromorphic computing. *Nature Reviews Physics*, 2:499–510, 2020.
- [99] Warren S. McCulloch and Walter Pitts. A logical calculus of the ideas immanent in nervous activity. *j-BULL-MATH-BIOPHYS*, 5(4):115–133, December 1943.
- [100] MDPI. The Artificial Intelligence Technologies for Electric Power Systems. https://www.mdpi.com/journal/energies/special_issues/AI_EPS, 2018. Accessed: 2023-09-10.
- [101] A. Mehonic and A. J. Kenyon. Brain-inspired computing needs a master plan. *Nature*, 604:255–260.
- [102] P.P. Mitra. Fitting elephants in modern machine learning by statistically consistent interpolation. *Nature Machine Intelligence*, 3, 2021.
- [103] Saber Moradi and Rajit Manohar. The impact of on-chip communication on memory technologies for neuromorphic systems. *Journal of Physics D: Applied Physics*, 52(1):014003, oct 2018.

- [104] US National Bureau of Standards. Ai risk management framework: Initial draft. <https://www.nist.gov/system/files/documents/2022/03/17/AI-RMF-1stdraft.pdf>, 2022. Accessed: 2023-08-30.
- [105] Nature. ChatGPT is a black box: how AI research can break it open. <https://www.nature.com/articles/d41586-023-02366-2>. Accessed: 2023-09-10.
- [106] Nature. How to stop data centres from gobbling up the world’s electricity . <https://www.nature.com/articles/d41586-018-06610-y>, 2018.
- [107] Nature. Documentary follows implosion of billion-euro brain project. *Nature*, 588:215–216, 2020.
- [108] Nature. In AI, is bigger always better? *Nature*, 615, 2023. Accessed: 2023-09-10.
- [109] nature.com. Solving the big computing problems in the twenty-first century. <https://www.nature.com/articles/s41928-023-00985-1.epdf>, 2023. Accessed: 2023-09-10.
- [110] nature.com. ‘ChatGPT detector’ catches AI-generated papers with unprecedented accuracy. <https://www.nature.com/articles/d41586-023-03479-4>, 2023.
- [111] nextplatform.com. SAMBANOVA TACKLES GENERATIVE AI WITH NEW CHIP AND NEW APPROACH. <https://www.nextplatform.com/2023/09/20/sambanova-tackles-generative-ai-with-new-chip-and-new-approach/>.
- [112] nextplatform.com. WHAT HAPPENS WHEN LLMS DESIGN AI ACCELERATORS? <https://www.nextplatform.com/2023/09/25/what-happens-when-llms-design-ai-accelerators/>.
- [113] nextplatform.com. CRAY revamps clusterstor for the exascale era. <https://www.nextplatform.com/2019/10/30/cray-revamps-clusterstor-for-the-exascale-era/>, 2019. Accessed: 2023-09-10.
- [114] nextplatform.com. AMD says AI is the number one priority right now. <https://www.nextplatform.com/2023/05/03/amd-says-ai-is-the-number-one-priority-right-now/>, 2023. Accessed: 2023-09-10.

- [115] nextplatform.com. There's still a long way to go with generative ai. <https://www.nextplatform.com/2023/05/05/theres-still-a-long-way-to-go-with-generative-ai/>, 2023. Accessed: 2023-09-10.
- [116] OpenAI. Gpt-4 technical report, 2023.
- [117] John K. Ousterhout. Why aren't operating systems getting faster as fast as hardware? <http://www.stanford.edu/~ouster/cgi-bin/papers/osfaster.pdf>, 1990. Accessed: 2023-09-10.
- [118] Yao P., Wu H., Gao B., Tang J, Zhang Q, Zhang W, Yang JJ, and Qian H. Fully hardware-implemented memristor convolutional neural network. *Nature*, 577:641–646, 2020.
- [119] Gualtiero Piccinini and Andrea Scarantino. Computation vs. information processing: why their difference matters to cognitive science. *Studies in History and Philosophy of Science Part A*, 41(3):237–246, 2010. Computation and cognitive science.
- [120] Walter Pitts and Warren S. McCulloch. How we know universals the perception of auditory and visual forms. *The Bulletin of Mathematical Biophysics*, 9:127–147, 1947.
- [121] Kathy Pretz. Stop Calling Everything AI, Machine-Learning Pioneer Says. <https://spectrum.ieee.org/the-institute/ieee-member-news/stop-calling-everything-ai-machinelearning-pioneer-says>, 2021. Accessed: 2023-08-30.
- [122] R. Quirion. Brain organoids: are they for real? <https://www.frontiersin.org/journals/science/articles/10.3389/fsci.2023.1148127/full>, 2023.
- [123] Randal E. Bryant and David R. O'Hallaron. *Computer Systems: A Programmer's Perspective*. Pearson, 2014.
- [124] A. D. Rast, X. Jin, F. Galluppi, L. A. Plana, C. Patterson, and S. Furber. Scalable Event-driven Native Parallel Processing: The SpiN-Naker Neuromimetic System. In *Proceedings of the 7th ACM International Conference on Computing Frontiers, Bertinoro, Italy, CF '10*, pages 21–30, New York, NY, USA, 2010. ACM.

- [125] M. Rosol, J. S. JGąsior, Laba J., K. Korzeniewski, and M. Mlynczak. Evaluation of the performance of GPT-3.5 and GPT-4 on the Medical Final Examination. <https://www.medrxiv.org/content/10.1101/2023.06.04.23290939v1.full.pdf>.
- [126] A.R. Samuel. Some studies in machine learning using the game of checkers. *IBM J Res Dev*, 44:1210–1229, 1959.
- [127] SC Seeman, L Campagnola, PA Davoudian, A Hoggarth, TA Hage, A Bosma-Moody, CA Baker, JH Lee, S Mihalas, C Teeter, AL Ko, JG Ojemann, RP Gwinn, DL Silbergeld, C Cobbs, J Phillips, E Lein, G T Jarsky. Sparse recurrent excitatory connectivity in the microcircuit of the adult mouse and human cortex. *Elife*, 2018.
- [128] semiengineering.com. AI Power Consumption Exploding. <https://semiengineering.com/ai-power-consumption-exploding/>, 2022. Accessed: 2023-09-10.
- [129] C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423, 1948.
- [130] Yiqiu Shen, Laura Heacock, Jonathan Elias, Keith D. Hentel, Beatriu Reig, George Shih, and Linda Moy. ChatGPT and Other Large Language Models Are Double-edged Swords. *Radiology*, 2023.
- [131] Horst Simon. Why we need Exascale and why we won’t get there by 2020. https://www.researchgate.net/publication/261879110/_Why_we_need_Exascale_and_why_we_won't_get_there_by_2020, 2014. Accessed: 2023-09-10.
- [132] Jaswinder Pal Singh, John L. Hennessy, and Anoop Gupta. Scaling parallel programs for multiprocessors: Methodology and examples. *Computer*, 26(7):42–50, 1993.
- [133] James Sir Lighthill. Artificial Intelligence: A General Survey. https://www.chilton-computing.org.uk/inf/literature/reports/lighthill_report/p001.htm.
- [134] L. et al Smirnova. Organoid intelligence (OI): the new frontier in biocomputing and intelligence-in-a-dish. <https://www.frontiersin.org/journals/science/articles/10.3389/fsci.2023.1017235/full>, 2023.

- [135] G. Somjen. *SENSORY CODING in the mammalian nervous system*. New York, MEREDITH CORPORATION, 1972.
- [136] Peter Sterling and Simon Laughlin. *Principles of Neural Design*. The MIT Press, 1 edition, 2017.
- [137] Gábor Szabó. Milyen messzire esett Newton almája? <http://real-eod.mtak.hu/977/1/01%20Szab%C3%B3%20007-020.pdf>, 2005. Accessed: 2023-09-10.
- [138] Taub, A. H., editor. *Volume VI: Theory of Games, Astrophysics, Hydrodynamics and Meteorology*. Oxford: Pergamon Press, 1963.
- [139] The Alan Turing Institute. Reflections on the response of the UK’s data science and AI community to the COVID-19 pandemic. https://www.turing.ac.uk/sites/default/files/2021-06/data-science-and-ai-in-the-age-of-covid_full-report_2.pdf, 2020.
- [140] Dan Tsafir. The context-switch overhead inflicted by hardware interrupts (and the enigma of do-nothing loops). In *Proceedings of the 2007 Workshop on Experimental Computer Science, San Diego, California, ExpCS ’07*, pages 3–3, New York, NY, USA, 2007. ACM.
- [141] Sherry Turkle. Dangerous Thoughts . . . And Machines With Big Ideas. <https://www.nytimes.com/1991/03/17/books/dangerous-thoughts-and-machines-with-big-ideas.html>, 1991. Accessed: 2023-08-30.
- [142] Sacha J. van Albada, Andrew G. Rowley, Johanna Senk, Michael Hopkins, Maximilian Schmidt, Alan B. Stokes, David R. Lester, Markus Diesmann, and Steve B. Furber. Performance Comparison of the Digital Neuromorphic Hardware SpiNNaker and the Neural Network Simulation Software NEST for a Full-Scale Cortical Microcircuit Model. *Frontiers in Neuroscience*, 12:291, 2018.
- [143] Giulia Vilone and Luca Longo. Notions of explainability and evaluation approaches for explainable artificial intelligence. *Information Fusion*, 76:89–106, 2021.
- [144] J. von Neumann. First draft of a report on the EDVAC. *IEEE Annals of the History of Computing*, 15(4):27–75, 1993.

- [145] Rainer Waser, editor. *Advanced Electronics Materials and Novel Devices*. Nanoelectronics and Information Technology. Wiley-VCH, 2012.
- [146] Kai Xu, Minghai Qin, Fei Sun, Yuhao Wang, Yen-Kuang Chen, and Fengbo Ren. Learning in the frequency domain. <https://arxiv.org/abs/2002.12416>, 2020. Accessed: 2023-08-30.