

Fizikai Nobel-díj 2024

The Nobel Prize in Physics 2024

Csabai István

az MTA levelező tagja

Eötvös Loránd Tudományegyetem Természettudományi Kar Fizikai Intézet, Budapest
csabai@elte.hu

Összefoglalás

Az idei fizikai Nobel-díj sokakban kisebb meglepetést váltott ki, mivel az avatatlan szemlélők számára a díjazott témakör nem kapcsolódik a fizika hagyományos területeihez. Ebben a rövid cikkben igyekszem elmagyarázni, hogy miként alapozták meg a statisztikus fizikában gyökerező modellek a napjainkban rohamos ütemben fejlődő mesterségesintelligencia-algoritmusokat.

Abstract

This year's Nobel Prize in Physics came as a bit of a surprise to many, as the awarded topic does not seem to relate to the traditional fields of physics for the untrained observer. In this short article, I will try to explain how models rooted in statistical physics have laid the groundwork for the rapidly developing algorithms of artificial intelligence today.

Kulcsszavak: Nobel-díj, fizika, gépi tanulás, mesterséges intelligencia, John J. Hopfield, Geoffrey E. Hinton

Keywords: Nobel Prize, physics, machine learning, artificial intelligence, John J. Hopfield, Geoffrey E. Hinton

A 2024. évi fizikai Nobel-díjat John J. Hopfield (Princeton Egyetem, NJ, USA) és Geoffrey E. Hinton (Torontói Egyetem, Kanada) kapta megosztva, az indoklás¹ szerint: „Az idei év két fizikai Nobel-díjasa a fizika eszközeit használta fel olyan módszerek kifejlesztésére, amelyek a mai nagy teljesítményű gépi tanulás alapját képezik. John Hopfield olyan asszociatív memóriát hozott létre, amely képes képeket és más típusú mintákat tárolni és rekonstruálni. Geoffrey Hinton olyan módszert talált fel, amely önállóan képes az adatokban rejlő tulajdonságokat feltárni, és így olyan feladatokat elvégezni, mint például bizonyos elemek azonosítása képeken. [...] Hinton erre a munkára építve hozzájárult a gépi tanulás jelenlegi robbanásszerű fejlődésének elindításához.” (URL1)

¹ Az idézetek az eredeti angol szövegek automatikus fordításai mesterségesintelligencia-alapú szoftverrel.

Maga a mesterséges intelligencia nagyon szerteágazó témakör, kialakítását több tudományág is magának követelheti. Ahelyett, hogy ennek egy pici szeletét, a Nobel-díj honlapján és a kutatók életrajzában is olvasható tudományos pályákat vázolnám, csupán azt igyekszem megvilágítani, hogy mi benne a fizika, avagy egy, a kezdetek óta a témát követő és művelő fizikusként szerénytelenül fogalmazva: *miért a fizikusok alkották meg a mesterséges intelligencia csíráját?*

Az emberiség ősi vágya, hogy az anyagi világ titkainak feltárása mellett megértse az emberi elme működését is. Sokféle úton meg lehet közelíteni a kérdést, de most nagyon nagy vonalakban vázoljuk csak a természettudományok redukcionista modelljét. Tudjuk, hogy az ember biológiai lény, és a gondolkodásért az agy felelős. Az agy idegsejtekből áll, amelyek egy összetett kapcsolati hálózatban kémiai és elektromos jelekkel kommunikálnak egymással. Egy-egy sejt sokezerféle fehérje, nukleinsav, lipid, ion... komplex, kölcsönható rendszere. Ezek a molekulák a kvantummechanika törvényei szerint működnek. Ez az út azonban nehezen járható, már ha egyáltalán. Egyetlen fehérje térszerkezetének *ab initio* kiszámolását se tudjuk megvalósítani, habár épp az ebben való jelentős előrelépésért adták idén a kémiai Nobel-díjat. Egy másik sejtösszetevő család, a mikro-RNS-ek vizsgálatáért szintén külön Nobel-díj járt az idén. Sok ezer hasonló díjat kell kívárnunk, amire egy sejtet részleteiben megértünk. És akkor az még csak egyetlen sejt. 2024 októberében tették közzé az *acetabularia* mákszemnyi agyának hatalmas erőfeszítéssel és technikai bravúrral (köztük mesterséges intelligencia felhasználásával) elkészített térképét (URL2), amely „csupán” 130 000 sejtet és 50 millió kapcsolatot tartalmaz. A molekulák szerkezetétől kezdve a génhálózatokon át a (valódi) neuronhálózatokig egy ma még megkerülhetetlen *probléma a komplexitás*, a lehetséges állapotok számának kombinatorikus robbanása.

A fizikának, ezen belül is a statisztikus fizikának van egy receptje, amely sikeresen le tudott írni egy addig rejtélyes jelenséget, a fázisátalakulást. Ha a vizet elkezdjük hűteni, akkor egy ideig nem sok változás tapasztalható, de nulla foknál az addigi folyadék hirtelen szilárd anyaggá, jéggé változik. Noha a víz csak H₂O-molekulák sokasága, még ezt se egyszerű nagy skálán leírni, hisz a jelenség megértéséhez kvantummechanikai pontossággal kellene molekulák milliárdjait követni. Próbáljuk tovább egyszerűsíteni a rendszert! Tegyük fel, hogy a részecskék állapotát nem végtelen dimenziós komplex hullámfüggvények írják le, hanem egyetlen szám, ráadásul az is csak két értéket vehet fel: +1 vagy -1 lehet. A vízmolekuláknak ez ugyan nem jó modellje, de a mágnesek leírására alkalmas. A mágneseknél is van fázisátalakulás: az úgynevezett Curie-hőmérséklet fölé melegítve elveszítik a mágnességüket. Itt a két fázis nem a szilárd és a cseppfolyós, hanem a mágneses és a nem mágneses állapot. Egy mágnes spinnel jellemezhető atomok sokaságából áll, amelyek kis elemi mágnesként képzelhetők el. Ha csak egy irányt vizsgálunk, akkor, ha minden spin felfelé áll (+1) vagy mind lefelé (-1), akkor összességében egy erős mágneset kapunk. Ha összevissza állnak fel és le, akkor az összeg nulla lesz, nincs makroszkopikus mágnesség. A szomszédos spinek számára „kedvezőbb”, azaz ala-

csonyabb energiájú, ha ugyanolyan irányba állnak, mint a szomszédjuk. Így alacsony hőmérsékleten a spinek beállnak a legalacsonyabb energiájú állapotba, azaz a szabályos rendbe. A növekvő hőmérséklet szétzilálja a rendszert, és a kritikus pont fölött már nincs kimutatható mágnesség. A makroszkopikus mágnességet sok kis spin kölcsönhatása, együttes viselkedése hozza létre, ez egy úgynevezett *kollektív jelenség*. Egyetlen vízmolekula esetén értelmetlen arról beszélni, hogy szilárd vagy folyékony, de a molekulák „kollektívája” létrehozza ezt az új jelenséget (angolul ez az *emergence*).

Mi lenne, ha az emlékezet egyszerűsített modelljét is kicsi, kétállapotú részecskékből próbálnánk összerakni? Spinek helyett legyenek mesterséges neuronok, amelyek vagy tüzelnek (elektromos impulzusokat küldenek ki a kimeneti nyúlványukon), vagy hallgatnak. A mágneses kölcsönhatás analógiájaként az egyik neuron kimenete hasson a többiek bemenetén (szinapszisain), alkossanak így egy hálózatot. Vegyünk először egy nagyon egyszerű hálózatot: kössük össze minden sejt kimenetét minden másik bemenetével. Ha spinekkel írjuk fel ezt a rendszert, akkor ez az úgynevezett spinüveg, amelynek viselkedését Giorgio Parisi tárta fel, aki ezért 2021-ben kapott fizikai Nobel-díjat. Ha a spineket átnevezzük neuronoknak, a mágneses hatást szinapsziserősségnek, akkor pedig a Hopfield-hálózatot kapjuk. Kisebb módosítással pedig a Boltzmann-gépeket, amelyek megalkotása Hinton nevéhez fűződik.

De hogyan lesz ebből memória? Rendezzük a spineket egy négyzetrácsba, és értelmezzük képként. A +1 legyen fehér, a -1 fekete. Ha az egyszerű mágnest felmelegítjük, összevissza állnak a spinek, véletlenszerű fekete-fehér „QR-kódot” látunk. Lehűtve „emlékezik”, de csak egy nagyon egyszerű, tiszta fekete vagy fehér képre. A Hopfield-hálózat a mágnesnél annyival összetettebb, hogy a sejtek közötti kapcsolat erőssége (szinapsziserősség) változtatható, sőt nemcsak a kollektív beállást erősítő pozitív érték lehet, mint a mágneseknél, hanem negatív is, azaz akkor kisebb a kölcsönhatási energia, ha ellentétesen állnak. Így a legalacsonyabb energiájú állapotok már komplex képeket, például egy számjegyet vagy betűt is kirajzolhatnak. John Joseph Hopfield kitalált egy olyan *tanítási mechanizmust*, amelyik képletesen szólva az összes lehetséges állapotok fölé feszített energiefelszínen ott alakít ki mélyedéseket, ahol a megjegyzendő képek vannak. Mivel az energiefelszín viszonylag sima, ha a képet kicsit eltorzítjuk (kimozdulunk az energiagödör aljáról), vagy csak egy töredékét adjuk a bemenetre, a rendszer akkor is a megfelelőközeli energiaminimumba konvergál, mintegy részleteiből is felismerve a képet.

Az itt vázolt *energiefelszín*-metafora, a végletekig leegyszerűsített „fizikus” modell nagyon termékeny kiindulási alapnak bizonyult. A hálózat topológiájának módosításával, a tanítási módszerek finomításával sok változatot alakítottak ki a kutatók az évek során, amelyek számos, egyre összetettebb feladatot tudtak megoldani. Talán az egyik legsikeresebb, mindenki által ismert eredmény a generatív nyelvi modellek családja, például a ChatGPT, amely az úgynevezett transzformer architektúráján alapul. A transzformereket leíró alapcikknek (Vaswani et al. 2017) tréfás címet adtak:

„Attention Is All You Need”.² Az inkább párkapcsolatokban elhangzó szemrehányást idéző cím az elemzendő szöveg egyes részei közötti korrelációk, asszociációk fontosságára utal. Három évvel később megjelent egy tanulmány „Hopfield Networks Is All You Need”³ címmel (Ramsauer et al. 2020), amelyben kimutatták, hogy a transzformerek algoritmusai ekvivalens Hopfield eredeti asszociatív memóriamodelljével, elegánsan összekötve a már tényleg intelligens nyelvi modellt az azt negyven évvel megelőző, most a legrangosabb tudományos elismeréssel díjazott korai modellel.

Alfred Nobel szándéka szerint a díjat azoknak kell odaítélni, „akik az előző évben a legnagyobb hasznot hozták az emberiségnek” (URL3). A díjazás nem tud ilyen azonnali módon működni, de az eredeti végakarát szellemében olyan fundamentális felfedezéseket díjaznak, amelyek jelentős társadalmi hasznot is hozhatnak. Mára már nem kétséges, hogy a mesterséges intelligencia hatalmas jelentőséggel bír, nagy hatással lesz mind a tudomány, mind pedig általában az emberiség jövőjére. Arra a közös felelősségre pedig, hogy ez a hatás pozitív legyen, éppen Geoffrey Hinton elmúlt egy-két évben végzett tevékenysége hívja fel a figyelmet. A díjat bejelentő első interjúban elhangzott gondolata szerint „reméli, hogy a díj hatására az emberek komolyabban veszik az általa hangoztatott félelmeket” (URL4).

Irodalomjegyzék

Ramsauer, Hubert – Schäfl, Bernhard – Lehner, Johannes et al. (2020). „Hopfield Networks is All You Need”. arXiv:2008.02217v3.

<https://doi.org/10.48550/arXiv.2008.02217>

Vaswani, Ashish – Shazeer, Noam – Parmar, Niki et al. (2017). „Attention Is All You Need”. In: *Advances in Neural Information Processing Systems 30 (NIPS 2017)*.

<https://www.semanticscholar.org/reader/204e3073870fae3d05bcbcf2f6a8e263d9b72e776>

URL1: <https://www.nobelprize.org/prizes/physics/2024/press-release/>

URL2: Cikksorozat a *Nature* folyóiratban az ecetmuslica agyának feltérképezéséről.

<https://www.nature.com/immersive/d42859-024-00053-4/index.html>

URL3: <https://www.nobelprize.org/alfred-nobel/alfred-nobels-will/>

URL4: <https://www.nobelprize.org/prizes/physics/2024/hinton/interview/>

² „Csak egy kis figyelemre van szükség.”

³ „Csak Hopfield-hálózatokra van szükség.”