

Az idő szerepe a technológiai és biológiai információ tárolásban és feldolgozásban

The role of time in storing and processing information in biology and technology

VÉGH János, MTA doktora, egyetemi tanár¹, BERKI Ádám József²

¹Kalimános BT, Debrecen, 4032, Magyarország, Vegh.Janos@gmail.com

²Marosvásárhelyi „George Emil Palade” Orvosi, Gyógyszerészeti, Tudomány és Technológiai Egyetem, 540142, Marosvásárhely, Románia, berki.adam@yahoo.com

Abstract

The time is not the subject of computing science: the classic computing paradigm does not consider data transfer time needed in physical implementations, and the theoretical and empirical execution times tend to differ. The biological computing is aware of the time (its computing system is „spatiotemporal”), but its currently used separable variable method covers that the space and time coordinates are connected through the conduction speed. Because of using that method, biology cannot interpret the method of storing and processing information. Furthermore, the biology-mimicking technical computing systems also discover that the present technology is not suitable for preparing brain-like systems. Introducing the time-aware computing model, these issues are resolved. The model also explains, why learning and machine learning have only the word in their name common; otherwise their methods are orthogonal.

Keywords: information storing; spatiotemporal; biomorph computing; machine learning

Kivonat

Az idő nem tárgya a klasszikus számítástudománynak: a klasszikus paradigma nem veszi figyelembe a fizikai implementációkban az adatok átviteléhez szükséges időt, és egyre inkább eltér a számítógépeken futtatott alkalmazások elméleti és empirikus futási ideje. A biológiai számítás ugyan tud arról, hogy számítási rendszere „spatiotemporális” jellegű, de a használt megközelítésben a szeparált tér és idő koordináták elfedik hogy a koordinátákat a vezetési sebesség összekapcsolja. Emiatt nem tudja értelmezni az információ tárolás és feldolgozás módszerét sem. A technológiai rendszerekkel imitált biológiai rendszerek pedig felfedik, hogy jelenlegi technológia megoldásainkkal nem lehet agy-szerű rendszereket készíteni. Az időt tartalmazó számítási modell bevezetése feloldja ezeket az ellentmondásokat. A modell azt is értelmezi, hogy a tanulás és a gépi tanulás csak nevében hasonlít; a módszereik ortogonálisak.

Kulcsszavak: információ tárolás; spatiotemporális; biomorph szerkezetek; gépi tanulás

1. BEVEZETÉS

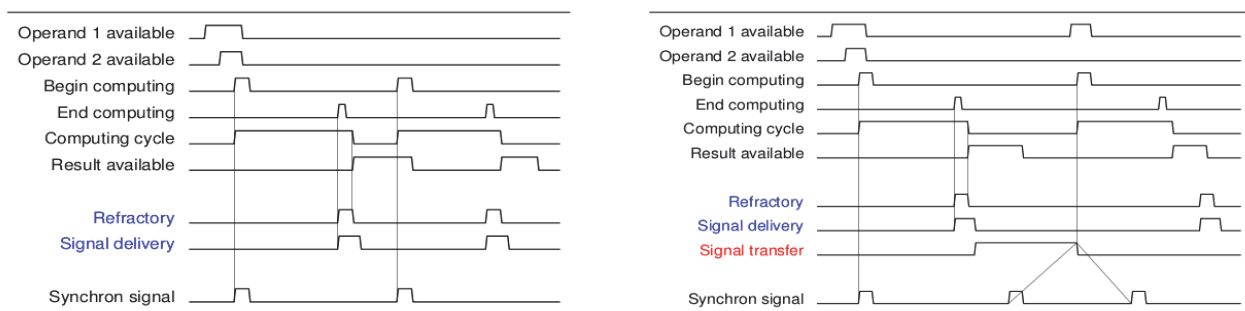
Bár tudjuk, hogy az idő létezik és egyre gyorsabban múlik, tudományos és filozófiai értelemben már a fogalmát is nagyon nehéz meghatározni [1], bár kulturális történelmünk során sokan megpróbálták. Az viszont mindennapi tapasztalatunk, hogy az idő mindent befolyásol. Meg azt is tudjuk, hogy az információ robbanás korát éljük, és a két fogalom összekapcsolódik: társadalmilag és személyesen is egyre nagyobb mennyiségű információ egyre gyorsabb feldolgozására kényszerülünk. Az idő (és az információ) ilyen értelemben is mindinkább főszereplővé válik.

Az információt évmilliók során kifejlesztett biológiai feldolgozó rendszerünkkel is feldolgozzuk, továbbá évtizedek alatt kifejlesztett (elektronikus) technológiai rendszerek használatával igyekszünk megbirkózni az exponenciálisan növekvő mennyiségű információval, a feldolgozásba beleértve a megfelelő szelektálást is. A feladat egyre nehezebb. Mind agyunk, mind technikai rendszereink növekvő nehézséggel boldogulnak a feladattal. Agyunkról a működés tanulmányozásának megkezdése óta tudjuk, hogy (a neurális folyamatok

lassúsága miatt) működése ún. spatiotemporális jellegű, azaz ugyanazt az információt az agyon belül más helyütt más időben (és a környezet által okozott kis torzulásokkal) látjuk. Mivel a jel terjedési sebesség 10 milliószor lassúbb, mint az elektromágneses hullámok esetén, a biológia természetes módon alkalmazza a „spatiotemporális” működést, bár a szó jelenlegi értelmezése nem azt jelenti, hogy a neurális hálózatokban a tér és idő koordinátákat a terjedési (vezetési) sebesség összekapcsolja, hanem hogy egy matematikai értelemben szeparálható, de önkényesen választott függvény írja le a viselkedést.

2. A SZÁMÍTÁSI MODELL

A műszaki technológiában használatos információ közvetítési folyamatok – éppen biológiai működésünk lassúsága miatt – érzékelhetetlenül gyorsan zajlanak. Az intenzívebb használat és a nagyon nagy rendszerek iránti igény azonban felfedte, hogy a „nagyon gyors” azért nem végtelenül gyors, és – éppen azért, mert mindkét információ feldolgozási rendszerben véges a terjedési és a feldolgozási sebesség – a nagy technológiai rendszerek viselkedése éppúgy eltér a mindennapos használatú „játék” rendszerek viselkedésétől, mint ahogyan csupán a neuronok működésének ismeretében nem tudjuk levezetni agyunk működésének tapasztalt viselkedését, pl. hogy van tudatunk. A kétféle megvalósítás tulajdonságai alapjaiban különböznek. Tudjuk, hogy a biológiai rendszerek élethosszig tanulnak, rövid és hosszú távon egyaránt; amit a technológia rendszerek csak nagyon rossz határfokkal és különféle (tanító és normál) üzemmódokkal tudnak elérni.



1. ábra

A számítási és adat átviteli idők viszonya Neumann egyszerűsített (bal oldali) és teljes (jobb oldali) számítási modelljében. Láncolt műveletek esetén nyilván figyelembe kell venni az átviteli időt; ez rövid számítási idők és/vagy nagy fizikai méretű számítási rendszerek esetén alapvető különbséget jelent.

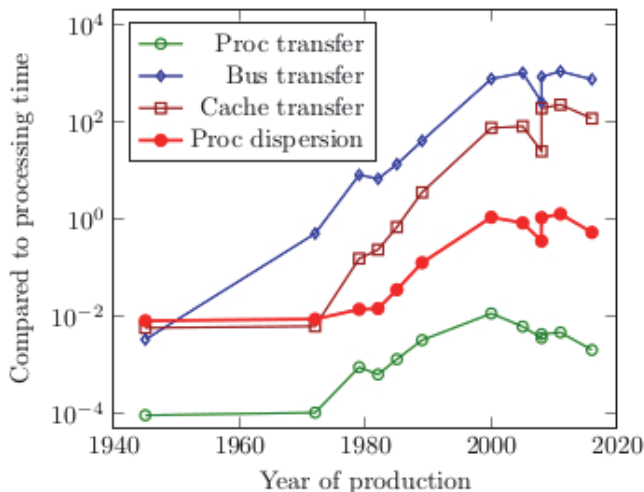
A technológia oldaláról sikerült felfedni, hogy – bár elvileg tudtuk, hogy az elektromágneses hullámok terjedési sebessége véges és a nagyon gyors elektronikus állapot változások követését már az anyagi tulajdonságok korlátozzák [2] – a fizikailag nagy méretű, nagyon sok, és a feladat megoldása során egymással adatokat cserélő számító elemet tartalmazó, nagyon sok adattal dolgozó nagy számítógépek teljesítőképessége már elérte a fizikailag lehetséges határt [3]; ezért bizonyos típusú feladatok megoldására a jelenlegi technológiai kivitel már elvileg sem alkalmas. A biológia oldaláról pedig, bár szinte szó szerint atomjaira szedték a neurobiológusok az agyat, nem sikerült megtalálni, hogy valójában hogyan tárolja a biológiai számítási rendszer az információt [4]; ezért annak pontos részleteit sem, hogy hogyan dolgozza azt fel. Mindennek ellenére ambiciózus tervek születtek és világszerte folynak jól támogatott kutatások jelenleg is az agy működésének számítógéppel történő modellezésére; inkább kevesebb, mint több sikerrel [5]. Háromnegyed évszázada Neumann János a neurális működés ismeretének akkori szintjén alkotta meg azt a számítási modellt, amelyen mai számítógépeink működése alapszik; bár már akkor figyelmeztetett, hogy sokkal gyorsabb processzorok esetén az általa használt elhanyagolások nem érvényesek; továbbá, hogy az általa javasolt „klasszikus paradigma” nem használható neurális működés leírására. Mára a technológia fejlődése következtében számítási szerkezeteink működése szinte nem is hasonlít arra, amit a számítás tudomány alapját képező egyszerűsített paradigma feltételez [6].

Neumann alap gondolata az volt, hogy a számításokat lánc-szerűen végezzük, azaz az egyik elemi számító egység eredményét egy másik számító egység bemenő adatként használja, és hogy a számító egység működése ideje mellett az egységek közötti átviteli időt is figyelembe kell venni. A számítás és az átvitel kölcsönösen akadályozzák egymást: a műveletet nem is tudjuk elkezdni, amíg az operandusok meg nem érkeztek; fordítva pedig addig nem lehet elszállítani az eredményt, amíg a számítás el nem készült. Ez akkor is igaz, ha a technológiai kivitelben fizikailag ugyanaz a processzor végzi az egymás utáni számításokat: az

eredményt az „output section”-ből át kell szállítani az „input section”-be, mielőtt a következő művelet végeznék. Abban a korban azonban a művelet végzési idő néhány millisekundum nagyságú volt, az átviteli idő meg néhány mikrosekundum nagyságú; ezért Neumann teljes joggal mondotta, hogy az akkori technológiát nagyon jó közelítéssel leírja a ma klasszikusnak nevezett, egyszerűsített modell. A lényegi elhanyagolást mutatja az 1. ábra. A teljes modell (jobb oldalt) tartalmazza az átviteli időt is, az egyszerűsített model (bal oldalt) pedig nem (azaz, feltételezi az azonnali kölcsönhatást). A teljes modell mind technológiai, mind biológiai számításra alkalmazható, bár a folyamatok sorba rendezését más módon előállított jelek végzik el.

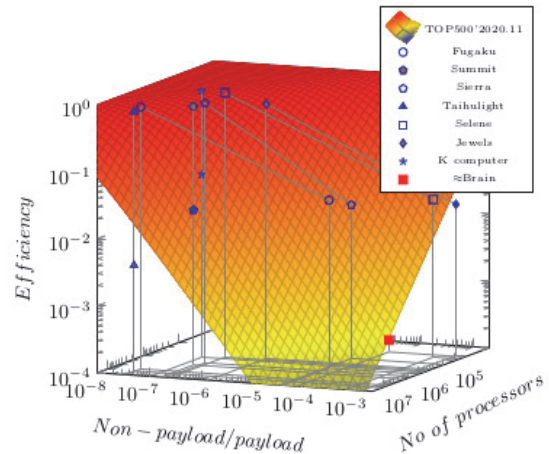
3. SZINKRONIZÁLÁS

Összetett rendszerek esetén külön nehézséget jelent az elemi számító egységek működésének összehangolása, a szinkronizálás: mind a műveleti idő, mind a szállítási idő esetről esetre változik. A technológia és a biológia különböző szinkronizálási módot használ. Neumann az említett időzítési viszonyokra javasolta a központi órajel használatát szinkronizálásra technológiai számító rendszerekben. Ez a megoldás kétségtelenül egyszerűsíti kis rendszerek esetén a kivitelezést, de a szinkronizálás egyúttal valamelyik művelet várakoztatását (a hasznos teljesítőképesség csökkenését) is jelenti. Az alapfeltétel, hogy minden művelet befejeződjön egy órajel alatt, továbbá hogy a gazdaságos működtetéshez a műveletek időbeli hosszának diszperziója nagyon kicsi legyen. A technológia fejlődés során óriási mértékben megnőtt a processzorok (elemi kapuk számában kifejezett) mérete és egyre kevésbé teljesül, hogy kicsi legyen a diszperzió (2. ábra). A diszperzió mérete még aránytalanabban nő, ha a számítógép összetevőinek fizikai méretét is figyelembe vesszük (a Moore-megfigyelés csak a processzoron belüli alkatrész sűrűsége teljesül, úgy, ahogy). A következmény, hogy modern processzoraink az energia nagy részét hőtermelésre fordítják, számolás helyett.



6. ábra

A processzor diszperziójának fejlődése a technológia változásával. Viszonyítási alapul a mindenkori működési sebesség szolgál.



7. ábra

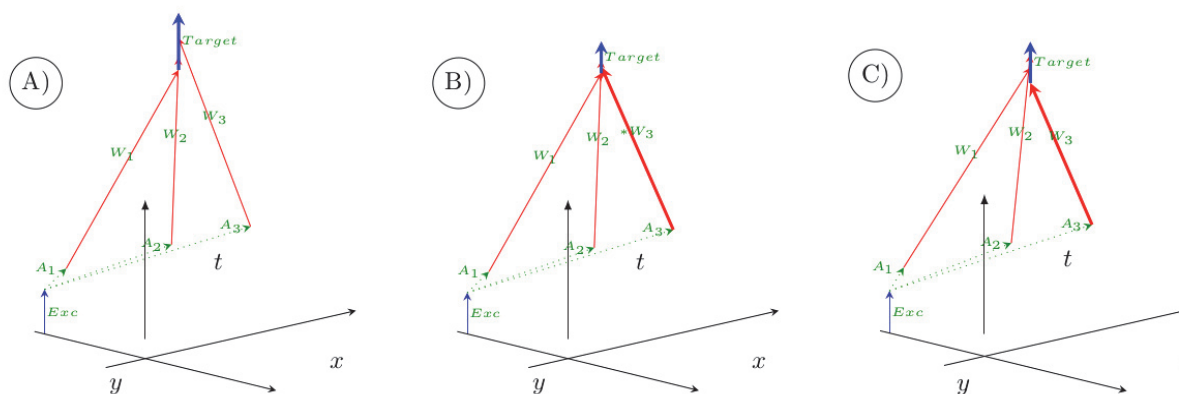
A párhuzamosítással működő (elosztott) rendszerek határfokának változása a processzorok száma és a nem-számítási hasznos teljesítőképesség arány függvényében

További hatások miatt, de lényegében az idő elhanyagolása következtében, nagy rendszereknél óriási méretű határfok csökkenés következik be (3. ábra). Mivel a számítógép processzorok működését egy-szálú végrehajtásra optimalizálják, nagy rendszerek és a neurális működést imitáló alkalmazások esetén már jól látható, hogy a jelenleg használatos módokon nem lehetséges megoldani a feladatot. Nem növekszik a magányos számítógép processzor teljesítőképessége, nincs újabb millió-processzoros nagy szuperszámítógép, korlátokba ütközött a mesterséges neurális hálózatok fejlesztése is [7]. Az újonnan megjelent, főleg GPU-alapú szuperszámítógépek valódi feladatok esetén már processzoraiknak csak töredékét tudják használni.

A biológiai rendszerek lényegesen nagyobbak lehetnek, mint a technológiaiak és természetesen módon a végrehajtó egységek együttműködésére fektetik a hangsúlyt, nem pedig az egyes neuron képességeit fejlesztik az elképzelhető legjobbra. Az együttműködés keretében az egyes neuronok kísérletezik ki a tanuláshoz szükséges időzítést, de azt a neuronhálózat véglegesíti. A számítási hatékonysághoz a biológiának másféle

szinkronizálási módszert (aszinkron ön-szinkronizálás) kell használnia: a neurális impulzusok megérkezése szolgáltatja a „Begin Computing” jelet, amelyek közül az első ilyen impulzus indítja el a „Computing cycle” folyamatot. Hasonló módon, a neuron maga állítja elő az „End computing” jelet is, amikor elegendő töltést gyűjtött össze membránján, és a potenciál elérte a kisüléshez szükséges küszöbertértéket. Azaz, a biológiai számítási rendszerek (technológiai fogalmakkal) a digitális és analóg technikák egy sajátos keverékét használják, ráadásul anatómiai mechanizmusok lehetővé teszik az „architektúra” paramétereinek menet közbeni megváltoztatását is. Túl azon, hogy a neuronok autonóm módon tudják változtatni a neurális impulzus elküldésének időpontját, az anatómiai paraméterek változtatásával a neuronok hálózata az impulzus terjedési (vezetési) sebességét is képes változtatni. Ennek következtében térnek el sajátágaik a technológiai rendszerektől, bármennyire is „biomorphic” a technológiai rendszer.

4. A TANULÁS



8. ábra

A neuron rövid és hosszú távú tanulásának folyamata. a) egy neuron assembly impulzusai érik el a cél-neuront. B) rövid távon, a cél-neuron módosítja szinaptikus súlyait C) hosszú távon, a rendszer megnöveli a kedvezményezett impulzus szállítási sebességét.

A 4. ábra mutatja azt a két lépésből álló módot, ahogyan neuronjaink tanulnak. Az A) ábrán látható az alapállapot, amikor egy neuron assembly három tagja küld információt ugyanazon cél-neuronnak. Az assembly tagjai függetlenek, különböző helyeken vannak, a jel indulási/érkezési ideje és sebessége is eltérhet. Amikor a beérkező jelek integrálva meghaladják a küszöbértéket (ennek folyamata maga a „számítás”), a cél-neuron (egy belső „Signal delivery” idő után) elküldi az eredményt (spike), ami aztán majd a „Signal transfer” idő elteltével jut el céljához. A neurális információ elküldésének időpontja attól függ, mennyi idő alatt éri el a membrán a küszöb feszültséget: a beérkezett töltéseket a membrán egy „szinaptikus erősség” faktorial szorozva integrálja, azaz a nagyobb szorzó faktorial rendelkező jelek hamarabb vezetnek kisüléshez. A B) ábra mutatja, mi történik, ha az A_3 neuron jelét fogadó szinapszis erőssége 50%-kal megnő, mivel a cél-neuron úgy találja, hogy az A_3 neuron jelét érdemesebb nagyobb súllyal kezelni. Ekkor ez az utolsóként beérkező adalék az integrálás korábbi befejezését jelenti, azaz a cél-neuron hamarabb fog tüzelni. Ez a mechanizmus viszonylag gyors, csupán a neurotranszmitterek helyi gradiensének megváltoztatását igényli. A gradiens fenntartása viszont csak töltések pumpálásával lehetséges, ami jelentős energia befektetést igényli; ez a módszer azonban jól használható rövid távú tanulásra. Ha az A_3 neuron jelének hasznossága tartósan bizonyul, a neuron hálózat megpróbálja ugyanezt a hatást más módon elérni. Amint azt a C) ábra mutatja, ha az A_3 tagtól származó jel sebessége megnő 10%-kal, akkor ugyanezt a hatást érhetjük el. A sebesség növelését a biológia úgy tudja megoldani, hogy növeli az axon szigetelő rétegének vastagságát, ami ugyan napokig/hetekig tartó folyamat, viszont hatása csaknem élethosszig tart. Miután azonban megnő a továbbítási sebesség, már nincs szükség a gradiens fenntartására, azaz kisebb energia befektetéssel érhető el ugyanaz a hatás. A biológia előbb megkeresi az optimális időzítést, majd minimalizálja az energia felhasználást. A biológia látja a tanulással növekvő szigetelő vastagságot (lásd [8] hivatkozásait), de a mechanizmust az időt független változóként kezelve, nem sikerült megértenie.

A biológia tehát két olyan módszert is használ, amelyet a jelenleg használatos technológia nem tud utánozni. Egyrészt kevert digitális/analóg számítási módszert használ, másrészt változtatni tudja az információ

hordozó terjedési sebességét. Emellett olyan szinkronizálási módszert alkalmaz, amelyik alkalmazkodik a változó sebességekhez és nagyon kis energia fogyasztást tesz lehetővé. A biológia előnyt kovácsol az lassúságból: a lassú folyamat maga tárolja az információt (érdekes párhuzam a kezdeti számítógépek késleltető művonalas tárolóival), továbbá a számítás során is kihasználja a működés lassúságát. A biológiai számítógépek működését csak az idő figyelembe vételével lehet megérteni.

A működési modelljünkben már a kiindulási ponttól eltérő technológiai és biológiai megvalósítások így csak nagyon durva közelítésként és csak nagyon egyszerű számítási rendszerekben írják le ugyanazt a jelenséget. Kezdeti hasonlóságuk miatt próbálják a technológiai számítási rendszereket nagy méretű és nagy bonyolultságú biológiai számítási rendszerek imitálására is használni. Az eltérő számítási modell azonban egyrészt természetes módon korlátozza a technológiai számítógépek teljesítőképességét, másrészt nagyobb méretű rendszerek és mélyebb szintű tanulmányozás esetén kifejezetten félrevezető eredményeket adhat.

ÖSSZEFOGLALÓ

Az idő fontos szerepet játszik mind a technológiai, mind a biológiai számítási rendszerekben. Az időt megfelelő módon számításba vevő számítási modell sikeresen írja le mindkét implementáció működését. A modell sikeresen megmagyarázta a nagy-skálájú technológiai implementációk tapasztalt teljesítőképesség korlátait, és értelmezni tudta a biológiai információ tárolás és feldolgozás módszerét. A modell alapján sikerült kimutatni, miért csak nevében rokon a tanulás és a gépi tanulás, ebből következően az intelligencia és a mesterséges intelligencia.

KÖSZÖNETNYILVÁNÍTÁS

A kutatás a Nemzeti Kutatási, Fejlesztési és Innovációs Alap K-136496 sz projektjének támogatásával készült.

Irodalmi hivatkozások

- [1] Buzsáki Gy, *The Brain from Inside Out*, 2019, Oxford University Press, 2019, 978-0-19-090538-5
- [2] Markov I. L.: Limits on fundamental limits to computation. *Nature*, 2014, 512(7513), 147-154
- [3] Végh J.: Finally, how many efficiencies the supercomputers have? *The J. Supercomputing*, 2020, 12(76), 9430-9455
- [4] Sterling P., Laughlin S.: *Principles of Neural Design*, 2017, MIT Press, 978-0-262-53468-0
- [5] Végh J.: How Amdahl's Law limits performance of large artificial neural networks, *Brain Informatics*, 2019, 4(6), 1-11.
- [6] Végh J.: von Neumann's missing "Second Draft": what it should contain, 2020, *Proc. 2020 Internat. Conf. on Computational Science and Computational Intelligence* <https://american-cse.org/sites/csci2020proc/pdfs/CSCI2020-6ScvvdzjqC7bKupZxFmCoA/762400b260/762400b260.pdf>
- [7] Hutson M.: Core progress in AI has stalled in some fields, *Science*, 2020, 6494(368), 927, 10.1126/science.368.6494.927
- [8] Végh J., Berki Á.: Storing and processing information in technological and biological computing systems. 2021, *Proc. 2021 Intern. Conf. on Computational Science and Computational Intelligence*, in print. <https://www.researchsquare.com/article/rs-88297/v1>