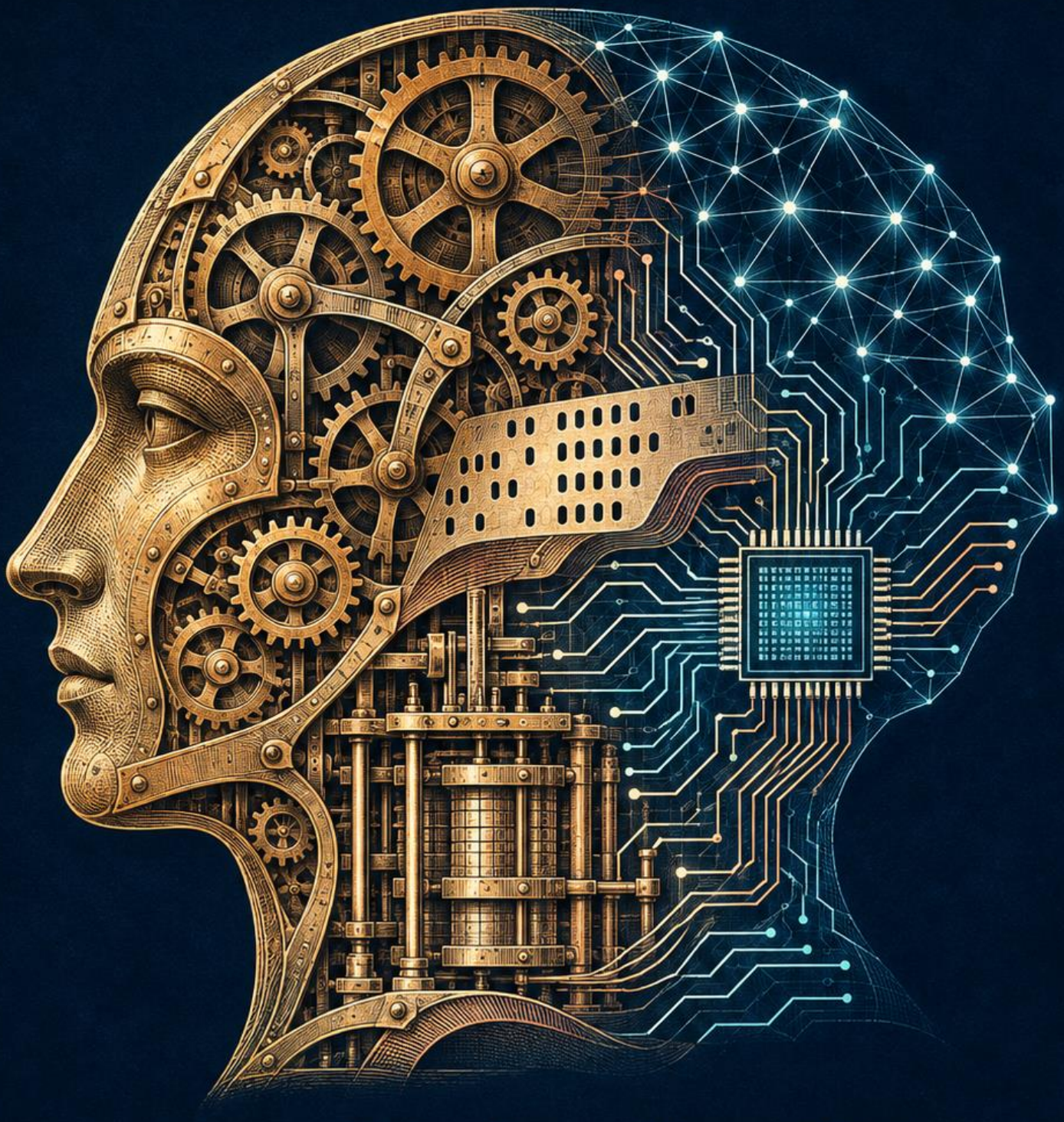
A gondolkodó gép nyomában

Az automatáktól a mesterséges intelligenciáig



Nagy Károly - ChatGPT

A gondolkodó gép nyomában

Az automatáktól a mesterséges intelligenciáig

2027

Tartalomjegyzék

Előszó – Miért akar az ember, gondolkodó gépet alkotni?

A mesterséges intelligencia története egyszerre találmányok, tudományos felismerések, túlzott remények és váratlan áttörések története. Mit tekintettek intelligenciának az egyes korokban, és mit vártak a gépektől?

I. rész – Amikor a gépek még csak az életet utánozták

1. Mesterséges lények a mítoszokban és az ember képzeletében
Talósz, a gólem és az életre keltett teremtmények. Az alkotás és az alkotója ellen forduló teremtmény motívuma.

2. Az ókor mozgó csodái
Hérón szerkezetei, vízzel és levegővel működő automaták. Előre meghatározott működés, amely kívülről önálló cselekvésnek látszik.

3. Óraművek, zenélő figurák és mechanikus állatok
Al-Dzsazarí, az európai órásmesterek, Vaucanson és a Jaquet-Droz család automatái.

4. A gép, amelyről elhitték, hogy gondolkodik
Kempelen Farkas sakkozó Törökje és beszélőgépe. A megtévesztő látszat és a valódi műszaki teljesítmény különbsége.

5. Amikor a működés leírássá válik
A Jacquard-szövőszék, a lyukkártya és a programozhatóság előzményei.

II. rész – A gondolkodás kiszámíthatóvá válik?

6. A gondolkodás mint szabálykövetés
Arisztotelész logikájától Leibniz elképzeléséig: lehet-e a következtetést számolássá alakítani?

7. Számológéptől az általános célú gépig
Pascal, Leibniz, Babbage és Ada Lovelace. A felismerés, hogy egy gép számokon kívül más jelképekkel is dolgozhat.

8. A logika találkozása a matematikával
Boole, Frege és a formális rendszerek. Hogyan írhatók le pontos szabályokkal az állítások és a következtetések?

9. A kiszámíthatóság határai
Gödel, Church és Turing. Mely problémák oldhatók meg eljárásokkal, és hol vannak az elvi korlátok?

10. A számítógép megszületése
A háborús számításoktól a tárolt programig. Neumann János és az a műszaki alap, amelyre később az AI épülhetett.

III. rész – Visszacsatolás, kibernetika és az első önálló gépek

11. Gépek, amelyek érzékelik saját működésüket
Szabályozás és visszacsatolás: hogyan tart fenn egy rendszer valamilyen kívánt állapotot?

12. Norbert Wiener és a kibernetika megszületése
Irányítás és kommunikáció az élőlényekben és a gépekben. A különböző tudományterületek közös nyelvének keresése.

13. Az agy mint modell
McCulloch és Pitts mesterséges neuronjai, Hebb tanulási elképzelése. Az idegrendszer működésének első matematikai közelítései.

14. A teknősök, amelyek keresték a fényt
Grey Walter egyszerű robotjai és Ashby alkalmazkodó szerkezetei. Hogyan kelthet néhány szabály összetett viselkedést?

15. A robot szó és az ipari robot megjelenése
Čapek, Asimov, majd a gyárak robotkarjai. Az irodalmi emberutánzattól a meghatározott feladatot végző gépig.

16. A magyar kibernetika úttörői
Nemes Tihamér, Tarján Rezső, az M–3 és Muszka Dániel szegedi katicabogara.

IV. rész – Megszületik a mesterséges intelligencia

17. Tudnak-e a gépek gondolkodni?
Alan Turing kérdése és az utánzási játék. Mit lehet a viselkedésből megállapítani?

18. Dartmouth, 1956 – Egy kutatási terület nevet kap
John McCarthy, Marvin Minsky, Claude Shannon és Nathaniel Rochester kezdeményezése.

19. Két út az intelligencia felé
Szabályokkal következtető programok és tanuló neurális hálózatok. Két megközelítés, amely később újra és újra találkozik.

20. Az első sikerek és a nagy ígéretek
Tételbizonyítás, problémamegoldás, dámajáték és a perceptron. Miért tűnt közelinek az emberi szintű gépi intelligencia?

21. A gép megszólal, körülnéz és cselekszik
ELIZA, SHRDLU és Shakey. Látványos eredmények gondosan körülhatárolt világokban.

V. rész – A nagy ígéretektől az AI telekig

22. Miért bizonyult nehéznek az, ami nekünk könnyű?
Nyelv, látás, mozgás és hétköznapi józan ész. A lehetséges megoldások számának robbanásszerű növekedése.

23. Az első AI-tél
A gépi fordítás csalódásai, a Lighthill-jelentés és a finanszírozás visszaesése. Az ígéretek és az eredmények közötti szakadék.

24. A szakértői rendszerek felemelkedése
DENDRAL, MYCIN és az üzleti alkalmazások. Az emberi szaktudás szabályokba foglalása.

25. Tudás palackba zárva
Miért volt nehéz megszerezni, frissíteni és új helyzetekre alkalmazni a szakértők tudását?

26. A második AI-tél
A szakértői rendszerek üzleti korlátai és a speciális AI-számítógépek piacának megrendülése. Ami a visszaesés idején is tovább fejlődött.

VI. rész – A gépek tanulni kezdenek

27. Szabályok helyett példák
A gépi tanulás előretörése. Tanítóadat, általánosítás és a hibákból történő javulás.

28. Bizonytalansággal is lehet számolni
Valószínűségi modellek, statisztikai módszerek és következtetés hiányos információból.

29. A neurális hálózatok visszatérése
A visszaterjesztés elterjedése és a többrétegű hálózatok. Mi változott a korábbi próbálkozásokhoz képest?

30. A mesterséges intelligencia észrevétlenül beköltözik a mindennapokba
Karakterfelismerés, keresők, ajánlórendszerek, csalásfelderítés és beszédfelismerés.

31. Amikor a gép legyőzi a bajnokot
Deep Blue, Watson és AlphaGo. Mit bizonyít egy győzelem, és milyen kérdéseket hagy nyitva?

VII. rész – A mélytanulástól a beszélgető gépekig

32. Adat, számítási teljesítmény és új módszerek találkozása
Az internet, a grafikus processzorok és a nagy adathalmazok szerepe. Az ImageNet és a képfelismerés fordulata.

33. A gép megtanul alkotni
Generatív modellek, GAN-ok és diffúziós eljárások. Képalkotás és új tartalmak előállítása.

34. A transzformer és a nyelvi modellek felemelkedése
A figyelmi mechanizmus, az előtanítás és a nagy nyelvi modellek. Miért bizonyult ilyen termékeny tanítóanyagnak a nyelv?

35. A ChatGPT-pillanat
Hogyan vált a mesterséges intelligencia széles körben használható, hétköznapi eszközzé?

36. Egy rendszer már lát, hall és beszél
Multimodális modellek, eszközhasználat és AI-ügynökök. A válaszadástól a feladatok végrehajtásáig.

VIII. rész – A történet elér a jelenhez

37. Az intelligencia visszakapja a testét
Tanuló robotok, önvezető rendszerek és humanoidok. Miért nehezebb a fizikai világban boldogulni, mint egy képernyőn?

38. Mit ért meg a mesterséges intelligencia?
Tudás, következtetés, tévedés és hallucináció. Az intelligencia, a megértés és a tudat külön kérdései.

39. A fejlődés ára és a hatalom kérdése
Chipek, adatközpontok, energia, adatok és tőke. Kik képesek fejleszteni és működtetni a legerősebb rendszereket?

40. A gép, amely már a saját fejlesztésünkben is segít
AI a tudományban, a programozásban és az új rendszerek kutatásában. Gyorsulás, korlátok és az emberi ellenőrzés problémája.

Epilógus – Az alkotó felelőssége

Függelék Rövidítések és fogalommagyarázat, idővonal

Előszó – Miért akar az ember gondolkodó gépet alkotni?

Az ember különös teremtmény. Évezredeken át azon dolgozott, hogy eszközei megsokszorozzák testi erejét, messzebbre vigyék, mélyebbre láttassanak vele, és megőrizzék azt, amit az emlékezete már nem tud befogadni. Közben újra meg újra felmerült benne egy merészebb kívánság is: olyan szerkezetet alkotni, amely képes önállóan cselekedni, tanulni, talán még gondolkodni is.

Ez a vágy jóval idősebb a számítógépnél. Megjelent az életre keltett szobrok történeteiben, az ember alkotta szolgákról szóló mondákban és a mesterséges életet kereső elképzelésekben. Pügmalión szobra, a gólem vagy az alkimisták homunculusa különböző korok és kultúrák teremtményei, mégis rokon kérdéseket hordoznak. Átadhat-e az ember valamit saját képességeiből annak, amit létrehoz? Hol húzódik a határ az élettelen anyag és az önálló lény között? És mi történik, ha az alkotás többé nem engedelmeskedik alkotójának?

A vágy mellett tehát kezdettől ott járt az aggodalom. Segítőt szeretnénk, amely leveszi vállunkról a terheket, de tartunk attól, hogy elveszítjük felette az uralmat. Olyan gépet akarunk, amely önállóan megoldja a feladatát, közben szeretnénk biztosak lenni abban, hogy azt teszi, amit valóban kívántunk. Ez a feszültség végigkíséri könyvünk történetét.

Az első automaták még mozgásukkal ejtették ámulatba az embereket. Egy önműködő figura, egy zenélő szerkezet vagy egy mechanikus állat azt az érzést kelthette, hogy az élettelen anyagban valamilyen szándék ébredt. A fogaskerekek, szelepek és súlyok azonban előre meghatározott rend szerint működtek. A szerkezet nem tudta, mit csinál. A szemlélő mégis könnyen célt, ügyességet, sőt értelmet látott a mozdulataiban.

Ez a hajlam ma is velünk él. Ha valami megfelelően reagál ránk, hajlamosak vagyunk feltételezni, hogy érti is a helyzetet. Ha összefüggően beszél, gondolatokat tulajdonítunk neki. Ha együttérző választ ad, könnyen úgy érezzük, hogy át is érzi a gondunkat. A mesterséges intelligencia története ezért az emberi ítéletalkotás története is: mitől hiszünk egy gépet intelligensnek, és mire elegendő az a bizonyíték, amely meggyőz bennünket?

A kérdésre minden korszak más választ adott. Egykor a pontos számolás is olyan szellemi teljesítménynek tűnt, amely különleges emberi képességet igényel. Később a sakkozást, a nyelv használatát vagy a képek felismerését tekintették az intelligencia döntő próbájának. Amikor a gépek valamelyik területen sikert értek el, közelebbről is megvizsgálhattuk, miből áll az adott képesség. Kiderült, hogy egy szűk feladatban elért kiváló eredményből még nem következik általános tájékozottság vagy hétköznapi józan ész.

Egy sakkbajnokot legyőző program nem feltétlenül tud eligazodni egy szobában. Egy tárgyakat ügyesen mozgató robot nem feltétlenül képes elmagyarázni, miért végzi a feladatát. Egy választékosan fogalmazó rendszer pedig meggyőzően adhat elő téves állításokat. Ezek a különbségek arra figyelmeztetnek, hogy az intelligenciát nehéz egyetlen képességgel vagy egyetlen vizsgával meghatározni.

A gondolkodó géphez vezető út több tudományterületen formálódott. A logika azt kutatta, hogyan lehet helyesen következtetni. A matematika és a számításelmélet azt, hogy mely feladatok oldhatók meg pontosan leírt eljárásokkal. Az idegrendszer kutatói a tanulás és az emlékezés működését próbálták megérteni. A mérnökök pedig olyan berendezéseket építettek, amelyek érzékelik környezetüket, és a visszajelzések alapján módosítják működésüket.

A kibernetika e törekvések között keresett kapcsolatot. A visszacsatolás segítségével közös nézőpontból lehetett vizsgálni egy szabályozott gépet és egy élő szervezet bizonyos folyamatait. Később a mesterséges intelligencia önálló kutatási területként is nevet kapott. A feladat egyszerre tűnt világosnak és határtalannak: olyan gépeket kell alkotni, amelyek az embernél intelligenciát igénylő tevékenységeket végeznek.

Az első eredmények nagy reményeket ébresztettek. Ha egy program képes tételt bizonyítani, játszani vagy egyszerű mondatokra válaszolni, talán hamarosan a bonyolultabb gondolkodási feladatokkal is megbirkózik. A kutatók azonban sorra ütköztek olyan nehézségekbe, amelyek jelentőségét kezdetben alábecsülték. A lehetséges megoldások száma kezelhetetlenné nőtt. A hétköznapi tudást nehéz volt szabályokba foglalni. A valóság pedig sokkal változatosabbnak bizonyult a laboratóriumi példáknál.

A lelkesedést többször kiábrándulás követte. Visszaesett a támogatás, vállalkozások szűntek meg, és korábban ígéretes irányok kerültek háttérbe. Ezeket az időszakokat nevezzük az AI teleinek. A kutatás közben tovább folyt, és a későbbi áttörésekhez gyakran olyan gondolatok adtak alapot, amelyek korábban még nem tudtak széles körben hasznosulni. Több adatra, nagyobb számítási teljesítményre, jobb módszerekre vagy ezek megfelelő találkozására volt szükség.

E könyv ezt a hosszú, kanyargós utat követi végig. Az automaták mestereitől eljutunk a kibernetika úttörőihez, a szabályokkal következtető programoktól a példákból tanuló hálózatokig, majd a beszélgető és tartalmakat létrehozó rendszerekig. Közben a magyar szereplőkre is figyelünk: munkájuk segít megmutatni, hogyan kapcsolódott be a hazai tudomány és műszaki kultúra ebbe a nemzetközi történetbe.

A megértéshez nem szükséges programozónak vagy matematikusnak lenni. Arra viszont szükség van, hogy különbséget tegyünk egy látványos bemutató, egy tudományos eredmény és egy megbízhatóan használható eszköz között. Érdemes azt is megvizsgálni, mit vártak egy találmánytól a maga korában, és mi valósult meg ezekből a várakozásokból. Így a kudarcok is érthetőbbé válnak, és az áttörések valódi jelentősége is jobban kirajzolódik.

Miért akar tehát az ember, gondolkodó gépet alkotni? Mert segítséget keres a munkához, a felfedezéshez és az egyre összetettebb világ megértéséhez. Mert kíváncsi arra, meddig terjedhet az alkotóképessége. És mert a gép építése közben önmagát is kutatja: hogyan tanulunk, mire emlékezünk, miként döntünk, és mit jelent számunkra az értelem.

A gondolkodó gép történetét követve ezért újra és újra az emberhez jutunk vissza. Ahhoz, aki megálmodja, megépíti, feladatokkal bízza meg — és aki felelősséggel tartozik azért, milyen helyet ad neki a világban.

I. rész – Amikor a gépek még csak az életet utánozták

Az ember jóval azelőtt elképzelte a mesterséges lényt, hogy képes lett volna működő mását megépíteni. Előbb történetekben adott mozgást az élettelen anyagnak, később vízzel, levegővel, súlyokkal és rugókkal. A mesterséges intelligencia előtörténetének ebben a szakaszában a gondolkodó gép még nem kutatási program. A kérdések azonban már megszülettek: mitől eleven valami, mennyit árul el a külső viselkedés a belső működésről, és meddig terjed az alkotó hatalma a teremtménye fölött?

Az első három fejezet a képzeletbeli lényektől az ókori automatákon át a tizennyolcadik század mechanikus szereplőiig vezet. Nem egy megszakítatlan fejlődési sort követünk. A mítosz, a színház, a vallási látványosság, az órásmesterség és az élettani kutatás különböző célokkal közelített az élet utánzásához. Találkozásaikból mégis kirajzolódik egy tartós emberi törekvés: a saját képességeink egy részét az általunk alakított anyagra bízni.

1. Mesterséges lények a mítoszokban és az ember képzeletében

Mielőtt megszületett a gép

Képzeljünk el egy partot őrző bronzóriást. Nem fárad el, nem kér fizetséget, és nem kell meggyőzni arról, hogy teljesítse a feladatát. Vagy képzeljünk el egy szobrot, amely egyszer csak viszonozza alkotója érintését. A két történet egészen más vágyat fejez ki. Az egyik az ember erejét és biztonságát növeli, a másik a magányát oldaná fel. Mindkettőben átlépünk egy határt: az emberi kéz munkája többé nem egyszerű tárgy.

A mesterséges lények története ezért tágabb a mesterséges intelligencia műszaki történeténél. Egy életre keltett szobor nem feltétlenül gép, a varázslattal mozgásba hozott agyagember pedig nem számítógép. Ezekben a történetekben nem algoritmusok helyettesítik az isteni vagy mágikus erőt. A közös vonás az alkotói szándék: az ember olyan létezőt akar létrehozni, amely önállóan tesz valamit, amihez korábban élőlényre volt szükség.

A régi elbeszélések olvasásakor könnyű minden részletben a jelen előképét felfedezni. A varázsszóban programot, a mágikus pecsétben hozzáférési kulcsot, a bronztestben robotburkolatot láthatunk. Az ilyen párhuzam gondolatébresztő lehet, amíg tudjuk, hogy mi hozzuk létre. Nem bizonyítja, hogy a történet szerzője előre megsejtette az elektronikát. Inkább azt mutatja, milyen régóta foglalkoztat bennünket az engedelmesség, az önállóság és a teremtés felelőssége.

Talósz és az őrzésre alkotott test

A görög mitológiában Talósz Kréta szigetét őrző bronzlény. Eredetéről több változat maradt fenn. Egyes elbeszélések Héphaisztosz, az isteni kovács munkájának mondják, mások a bronznemzedék túlélőjét látják benne. Pszeudo-Apollodórosz szerint naponta háromszor kerülte meg a szigetet, és kövekkel tartotta távol az idegen hajókat. A hatalmas testben egyetlen ér húzódott, amelyet a boka közelében bronzszeg zárt le. [1]

Talósz pusztulásának is több története van. Az egyikben Médeia az örök élet ígéretével veszi rá, hogy engedje eltávolítani a szeget. Apollóniosz Rhodiosz elbeszélésében a varázslat hatása alatt megsérti a bokáját, elfolyik az életadó nedv, és az óriás összeomlik. Nem érdemes a változatokból egyetlen „eredeti műszaki leírást” összeállítani. A mítosz nem ilyen dokumentum. A különbségek éppen azt jelzik, hogy a történetet több korszak és közösség alakította. [1]



1. kép Talósz halála egy attikai vörösalakos kratéron, Kr. e. 420–400 körül. A fehérrel kiemelt bronzóriás mítoszát már az ókori képművészet is megörökítette. Museo Archeologico Nazionale Jatta, Ruvo di Puglia. Fotó: ArchaiOptix, CC BY-SA 4.0.

Mai szemmel Talószban könnyű egy önálló őrszemet felismerni. Feladata világos, ereje rendkívüli, környezetében pedig ellenséges szereplők is megjelennek. Sebezhetőségét nem a teljes bronztest gyengesége okozza, hanem egyetlen sérülékeny pont. Ebből ma a biztonsági tervezéshez is meríthetünk hasonlatot, de a mítosz jelentése ennél tágabb: még a legyőzhetetlennek látszó alkotás is elpusztulhat, ha alkotója nem ura minden feltételnek.

Talósz ráadásul fontos ellenpélda arra, hogy a mesterséges teremtmény szükségképpen fellázad. Ő nem a gazdája ellen fordul, hanem végzi az őrzést. Az argonauták nézőpontjából veszélyes akadály, Kréta nézőpontjából védelmező. Már itt megjelenik egy máig érvényes kérdés: kinek a biztonságát szolgálja az a szerkezet, amelyet biztonságunk érdekében hoztunk létre?

Héphaisztosz más alkotásai még közelebb kerülnek az emberi képességekhez. Az Íliász tizennyolcadik énekében maguktól közlekedő háromlábú állványok és aranyból készült szolgálólányok jelennek meg. A költő az utóbbiaknak értelmet, hangot és erőt tulajdonít. A segítő mesterséges lény képzete tehát nem pusztán a nyers munka kiváltásáról szól: társas együttműködést és az emberi cselekvéshez illő ügyességet is elképzel. Mindez az istenek világában történik, ahol az alkotás korlátait nem a földi műhely lehetőségei szabják meg. [2]

Pügmalión és az életre keltett vágy

Ovidius Átváltozások című művében Pügmalión elefántcsontból farag nőalakot. Munkáját olyan tökéletesnek látja, hogy beleszeret. Ajándékokat ad neki, megszólítja, majd Venus ünnepén olyan

társat kér az istennőtől, aki a szobrához hasonló. Hazatérve az elefántcsont enged az érintésnek: a mű élő nővé változik. Ovidius nem nevezi Galateiának a szobrot; a ma közismert név későbbi hagyományból származik. [3]

Ebben az elbeszélésben nem a munka vagy az őrzés automatizálása a tét. Pügmalión társat akar. Az alkotás egyszerre esztétikai teljesítmény és egy vágy kivetítése. A szobor addig nem mond ellent, amíg tárgy, és a történet alig időzik annál, milyen saját akarata lesz, miután életre kel. Az alkotó boldogságát részletesebben látjuk, mint az alkotásból személlyé lett lény nézőpontját.



2. kép Edward Burne-Jones Pügmalión és a képmás című sorozatának záródarabja, The Soul Attains, 1878. A kép az életre keltett alkotás későbbi művészeti értelmezése. Birmingham Museums Trust. Közkincs.

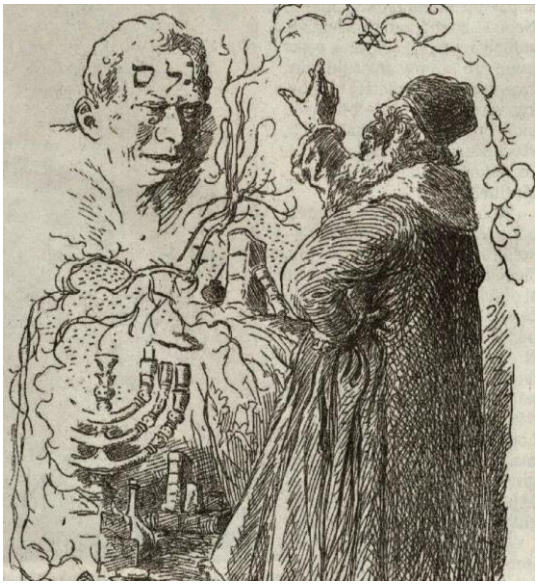
Ez a hiány a mai olvasó számára különösen beszédes. Ha valamit az igényeink szerint alakítunk ki, könnyen természetesnek vesszük, hogy a létrehozása után is a mi igényeinket kell szolgálnia. Egy valódi társ azonban éppen attól társ, hogy saját szándéka, döntése és nemet mondási lehetősége van. Pügmalión története így az ember által teremtett kapcsolat vágyát is megmutatja: olyan másik lényt keresünk, akit teljesen megérthetünk, és akinek a válaszaiban nem kell csalódnunk.

A mesterséges beszélgetőtársakkal vont párhuzam itt kulturális értelmezés. A kedves válasz, az emberi hang vagy a személyesnek ható megszólítás érzelmi kötődést kelthet a használóban; ebből önmagában nem következik, hogy a másik oldalon ugyanilyen élmény van. A mítoszban az istennő egyértelművé teszi az átváltozást. A valóságban az élet, az értelem és a szubjektív tapasztalat kérdését nem oldja meg az, hogy valami meggyőzően emberinek látszik.

A gólem és a szóval irányított erő

A zsidó hagyomány gólemalakja élettelen anyagból, rendszerint agyagból formált emberhez kapcsolódik, amely szent nevek és betűk segítségével jut valamiféle élethez. A motívum régebbi a legismertebb prágai változatnál. A prágai gólemet a hagyomány a tizenhatodik századi Löw rabbihoz köti, de a rabbi történeti személye és a róla szóló későbbi legendakör nem azonos bizonyító erejű forrás. A közismert prágai történet nyomtatott elterjedése a tizenkilencedik századhoz tartozik. [4–5]

Az egyik változat szerint a gólem a közösség védelmére készül. Engedelmes és rendkívüli erejű segítő, amelynek működését a rabbi képes elindítani és megszüntetni. Más változatokban az irányítás megbomlik: a teremtmény veszélyessé válik, ezért vissza kell venni tőle az életet adó szót vagy jelet. Nem minden gólemtörténetben szerepel ugyanaz az eljárás. A szájba helyezett istennév és a homlokra írt, megváltoztatható felirat különböző hagyományváltozatokhoz tartozik. [4–5]



3. kép Löw rabbi és a gólem Mikoláš Aleš ábrázolásán. A kép a legendát jeleníti meg, nem egy történeti esemény dokumentuma. Forrás: Wikimedia Commons. Közkincs.

A gólem különössége, hogy az anyag fölötti hatalom a nyelvhez kapcsolódik. A betű nem egyszerű megnevezés, hanem cselekvő erő. A mai olvasónak eszébe juthat a gépnek adott utasítás, de a szent név vallási jelentését nem szabad számítógépes kóddá egyszerűsíteni. A hasonlóság abban áll, hogy az ember valamilyen jelrendszeren keresztül érvényesíti akaratát egy nála erősebbnek bizonyuló testen.

A teremtmény engedelmessége közben nehéz kérdést vet fel. Elég-e pontosan végrehajtani azt, amit az alkotó mondott? Egy emberi segítő a helyzetből is következtet: felismerheti, hogy egy korábbi parancs most már kárt okozna. Az egyszerű utasításvégrehajtás és a körülmények megértése különböző képesség. A gólemmotívum mai technológiai értelmezésében éppen ez a különbség kap hangsúlyt. A legendák valamennyi változatára azonban nem helyes utólag ugyanazt az „elszabadult program” magyarázatot ráolvasni.

A homunculustól Frankenstein teremtményéig

Az alkímia mesterséges embere, a homunculus más utat képvisel. A latin szó kis embert jelent. Itt az életet nem egy fémből vagy agyagból készült test kapja kívülről, hanem az anyag átalakításával próbálják előidézni. A Paracelsus nevéhez kapcsolt De natura rerum e hagyomány fontos szövege. Az 1572-ben megjelent mű szerzősége vitatott; ezért a homunculusról szóló leírást nem célszerű Paracelsus biztosan saját, végrehajtott kísérleteként bemutatni. [6]

A homunculus az élet titkának műhelyben való megfejtését ígéri. Az alkímia világában az anyag átalakulása, a természet rejtett erői és a vallási elképzelések még szorosan összefonódnak. Ezt a közeget anakronizmus volna mai biotechnológiai laboratóriumnak látni. A kérdés viszont

felismerhető: ha a természet képes élőlényt létrehozni, az ember megismételheti-e ezt a folyamatot a feltételek megfelelő összeállításával?

Mary Shelley 1818-ban megjelent Frankensteinje már a modern tudomány társadalmi képzeletéhez tartozik. Victor Frankenstein életre kelti teremtményét, majd elfordul tőle. A regény lényét sok későbbi ábrázolás egyszerű tombolássá szűkítette, holott Shelley alakja tanul, beszél, szenved, kapcsolatot keres, és erkölcsi követelést fogalmaz meg alkotójával szemben. A teremtmény neve nem Frankenstein: ez az alkotó családneve. [7]

A regényben az alkotó felelőssége az életre keltés után válik igazán súlyossá. Mit jelent létrehozni egy lényt, majd magára hagyni? Meddig tart az alkotói kötelesség, ha az eredmény nem olyan, amilyennek megálmodtuk? Shelley története azt is megengedi, hogy a teremtmény felől nézzünk az emberre. A pusztítás mögött így nem egyszerűen egy hibás tárgyat látunk, hanem elutasítást, magányt és bosszút is.

Egyszerre akarjuk és félünk tőle

Nem minden mesterséges lény fordul az alkotója ellen. Pügmalión története beteljesüléssel zárul, Héphaisztosz szolgálói segítenek, Talósz pedig a rábízott területet őrzi. A fenyegető változatok mégis feltűnően tartósak. A létrehozás vágya és az uralom elvesztésének félelme ugyanabból a helyzetből fakad: olyan képességet akarunk átadni az alkotásunknak, amelynek értékéhez valamennyi önállóság is hozzátartozik.

A teljesen mozdulatlan tárgy biztosan engedelmes, de keveset végez el helyettünk. A sok mindenre képes segítő akkor is cselekszik, amikor nem állunk mellette. Minél több helyzetben várunk tőle megfelelő működést, annál kevésbé elég minden mozdulatát külön megszabni. A mítoszok ezt a feszültséget az életre keltés és a lázadás nyelvén beszélik el. A későbbi technika más fogalmakat talál hozzá, de a delegált cselekvés problémája megmarad.

A történetek a vágyainkról is tanúskodnak. Erős őrt, fáradhatatlan szolgálót, tökéletes társat vagy saját hasonmásunkat képzeljük el. Azután meglepődünk, ha ezek a vágyak nem férnek össze: a társ nem lehet egyszerre teljesen önálló személy és minden körülmények között engedelmes tulajdon. A mesterséges lény ezért az ember önvizsgálatának eszköze is. Nemcsak azt kérdezi, mire képes az alkotó, hanem azt is, mit akar valójában.

Az első működő automaták nem váltották valóra ezeket a mítoszokat. Valami szerényebbet, ugyanakkor ellenőrizhetőbbet tettek: megmutatták, hogy élettelen anyagok megfelelő elrendezésével életre emlékeztető cselekvéssor hozható létre. A következő fejezetben már nem isteni kovács vagy varázsszó mozgatja a testet. A csoda mögött megjelenik a cső, az emelő és a gondosan elrejtett ellensúly.

2. Az ókor mozgó csodái

Alexandria műhelyei és a láthatatlan levegő

Az ókori technikát könnyű kőfalakkal, utakkal és vízvezetékekkel azonosítani. A hellenisztikus világ műhelyeiben azonban olyan szerkezetek is készültek, amelyeknél nem a méret, hanem a részek együttműködése volt a legfontosabb. A víz, a levegő és a súlyok nem pusztán mozgattak valamit: megfelelő edényekkel, csövekkel és szelepekkel egymás után bekövetkező eseményeket idézhettek elő.

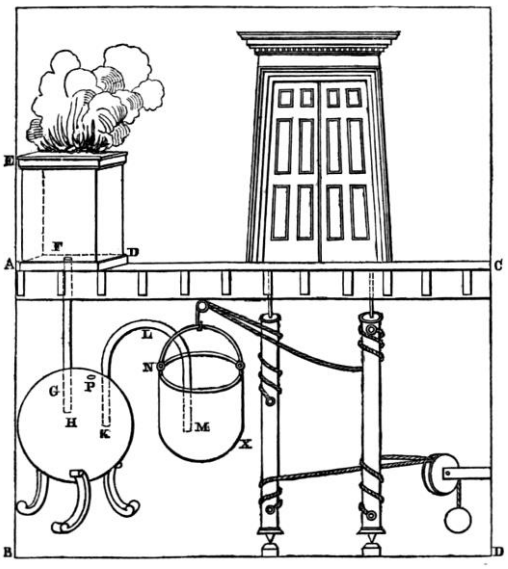
Alexandriai Hérónt általában az időszámításunk szerinti első századba helyezik. A személyéről kevés biztos adat maradt fenn, művei viszont rendkívül gazdag technikai világot tárnak fel. A Pneumatika a levegővel, vízzel és hővel működtetett eszközök gyűjteménye, az Automatopoétika pedig az automataszínházak szerkesztésével foglalkozik. Hérón korábbi mesterek hagyományát is közvetíti; műveit ezért nem helyes valamennyi bennük szereplő eljárás kizárólagos feltalálói jegyzékének tekinteni. [8–10]

A levegő különösen alkalmas a meglepetésre, mert láthatatlan. Egy edény kívülről üresnek tűnik, mégis van benne valami, ami melegítéskor kitágulhat, összenyomható, és nyomást gyakorolhat a vízre. A rejtett összeköttetések miatt a néző ott lát mozgást, ahol nem érzékel közvetlen emberi erőkifejtést. Az automaták mestere ezt a távolságot szervezi meg az ok és a látható következmény között.

Itt már érdemes két kérdést külön feltenni. Honnan származik a működéshez szükséges energia? És mi határozza meg, hogy a szerkezet mikor, milyen sorrendben használja fel? Az elsőre lehet a válasz tűz, magasabban tárolt víz vagy felemelt súly. A másodikra az edények elrendezése, egy szelep helyzete, egy kötél vezetése vagy egy emelő egyensúlya. Az energiaforrás és a működési rend különbsége a későbbi gépek megértéséhez is alapvető.

A templomajtó amely a tűzre válaszol

A Hérónnál leírt egyik legismertebb szerkezetben az oltáron meggyújtott tűz hatására kinyílnak a templomajtók. A látvány önmagában értelmezhető volna természetfölötti jelként. A műszaki leírás azonban okok láncolatát mutatja be. A tűz felmelegíti a zárt térben levő levegőt; a növekvő nyomás egy összekötött edényből vizet szorít át egy másik tartályba. [8–9]



4. kép Hérón hőhatásra nyíló templomajtajának rajza a Pneumatika 1851-es angol kiadásából, 37. szerkezet. Az oltár, a víztartály, a függesztett edény és az ellensúly egymáshoz kapcsolása adja a mozgást. Közkincs.

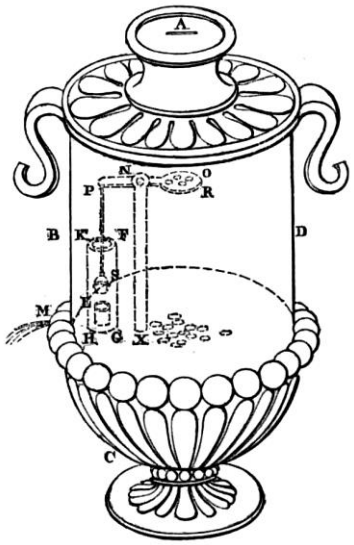
A megtelő függesztett edény nehezebbé válik, lesüllyed, és köteleken vagy láncokon keresztül elfordítja az ajtók tengelyét. A tűz megszűnése és a lehűlés után az átrendeződés visszafordulhat, az ellensúly pedig becsukja az ajtót. A lényeg nem valamely titokzatos új anyag, hanem ismert fizikai hatások gondos összekapcsolása. A bemenet hőhatás, a kimenet egy nehéz ajtó elfordulása. [8–9]

Ez már környezeti eseményre adott gépi válasz. Mégsem olyan döntés, amelynél a szerkezet mérlegeli az áldozat jelentését vagy a belépők szándékát. Nem a szertartást érti meg: a hőmérséklet és a nyomás változásának engedelmeskedik. Ugyanaz a fizikai hatás ugyanarra az eredményre vezethet akkor is, ha az ember más jelentést tulajdonít a történésnek.

A berendezés működési elvének ismerete és az ókori használat történeti bizonyítása két eltérő feladat. A leírás, valamint egy működő mai rekonstrukció megmutathatja, hogyan nyitható ajtó ilyen elven. Ebből nem következik, hogy minden ókori templom kapuja így működött, vagy hogy a leírt eszköznek minden részletében azonos példánya feltárásból is előkerült. A technikatörténetben a megvalósíthatóságot és a tényleges elterjedést külön kell vizsgálni.

Érme a tálcán és víz az edényből

A Pneumatika egy pénzbedobásra vizet kiadó szerkezetet is leír. Az érme egy emelő tálcájára esik, súlyával megbillenti azt, és megnyitja a víz útját. Amint lecsúszik a tálcáról, az emelő visszatér, a szelep zár, a folyás megszűnik. A vallási környezethez kapcsolt adagoló a modern árusító automaták távoli rokona abban az értelemben, hogy egy emberi műveletet gépileg kiváltott kiszolgálás követ. [8]



5. kép Érme súlyával működtetett vízadagoló a Pneumatika 1851-es kiadásából, 21. szerkezet. A belső emelő egy rövid időre nyitja meg a víz útját. A rajz későbbi kiadás illusztrációja, nem ókori lelet. Közkincs.

A szerkezet nem tudja, mit jelent a fizetés. Az érme tömegként hat rá. Ha egy másik tárgy megfelelően ugyanazt a mechanikai hatást fejtené ki, az egyszerű eszköz nem tudna gazdasági vagy erkölcsi különbséget tenni közöttük. Ez jól mutatja, milyen távol eshet a működtető emberi szándék a gép által ténylegesen érzékelt tulajdonságtól.

Mégis jelentős lépés, hogy a kiszolgálás egy része állandóan jelen levő kezelő nélkül végbemehet. A felhasználó elindít valamit, a belső szerkezet pedig lezárja a folyamatot. A leírás nem teszi bizonyossá a mai értelemben vett térfogatpontosságot: a kifolyó mennyiségre hat az átfolyás, a szelep és az érme mozgása. A lényeg a kezdés és a befejezés gépi összekapcsolása.

Az efféle egyszerű mechanizmusban már jól látszik az automatizálás egyik gyakorlati értéke. A gép hasznosságához nem kell az egész helyzetet emberi módon értenie. Elég lehet, ha egy szűk, jól körülhatárolt feltételt megbízható cselekvés követ. A korlátok akkor válnak veszélyessé, ha a használó a szerkezetnek olyan megkülönböztető képességet is tulajdonít, amellyel az nem rendelkezik.

Madárhang és mozgó színpad

Hérón leírásaiban a hang is mesterségesen előállítható jelenség. A víz beáramlása levegőt szoríthat egy síp felé, amely madárhanghoz hasonló hangot kelt. Az egymást váltó víz- és légmozgások szakaszos megszólalást hozhatnak létre. A szerkezet így nem csupán mozog: olyan érzéki benyomást kelt, amelyet az ember megszokásból élőlényhez kapcsol. [8]

A mechanikus madárnál különösen világos az utánzás rétegzettsége. A síp hangja emlékeztethet egy madáréra anélkül, hogy a gépnek légzőszervet, idegrendszert és madárviselkedést kellene felépítenie. Egyetlen jól kiválasztott jel is elegendő lehet az élet benyomásához. A néző tapasztalata kiegészíti azt, amit a szerkezet valóban végrehajt.

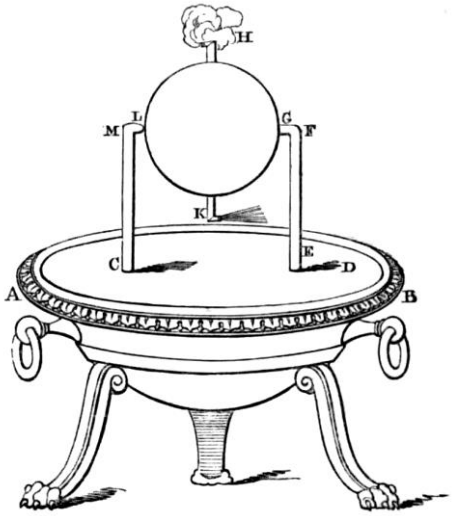
Az automataszínházak ezt az elvet hosszabb cselekvéssorokká alakították. A helyben maradó színpadon alakok, hajók és díszletek mozoghattak egymás után. A mozgó változat maga is helyet

változtatott, miközben jeleneteket mutatott be. A mai múzeumi rekonstrukciókban megfigyelhető, hogyan kapcsolhatók össze a mozgások egy súly és a tengelyek köré vezetett kötelek segítségével. [10]

A hajtó súly süllyedését kiáramló szemcsés anyag fékezhette, a feltekert kötelek pedig különböző pillanatokban húzták meg a megfelelő részeket. A kötél útja és feltekerése rögzítette a sorrendet. Más elrendezés más előadást eredményezhetett. Ebben a körülhatárolt értelemben beszélhetünk mechanikusan tárolt utasításssorról. Ez még nem mai számítógépes program, és a színház nem választotta meg önállóan a történet befejezését. [10]

Forgó gőzgömb és a fejlődés elmaradt útjai

A Hérón nevéhez kötött aiolipilé, a gőzzel forgatott gömb különösen gyakran szerepel a technikatörténeti összefoglalókban. A hevített vízből keletkező gőz csöveken keresztül egy elfordulni képes gömbbe jut, majd annak meghajlított kifúvócsövein távozik. A kiáramlás reakciója forgatja a gömböt. A Pneumatika ötvenedik szerkezete világosan leírja ezt az elvet. [8]



6. kép A gőzzel forgatott gömb, az aiolipilé rajza Hérón Pneumatikájának 1851-es kiadásában, 50. szerkezet. A kiáramló gőz reakciója forgatja a gömböt. Közkincs.

A forgó gömb könnyen felkelti a kérdést, miért nem indult el már az ókorban a gőzgépekre épülő ipari átalakulás. A kérdésben azonban veszélyes előfeltevés rejtőzik: mintha egy működési elv felismerése után már csak türelemre volna szükség a gyárakhoz. Egy tartós és gazdaságos erőgéphez megfelelő anyagok, tömítések, megmunkálás, javíthatóság, energiaellátás és fizetőképes felhasználási igény is kell.

A gőzzel forgatott gömb ezért nem egy elkészült ipari gőzgép, amelyet az ókoriak érthetetlen módon elfelejtettek munkára fogni. Egy részjelenséget mutat meg. A technikai fejlődés nem előre kijelölt út, amelyen egyetlen társadalomnak minden állomáson végig kell haladnia. Sok ötlet hosszú időn át bemutatóeszköz marad, más később új feladattal és más technológiákkal együtt nyer jelentőséget.

Ugyanez az óvatosság az automaták történetében is szükséges. A látványos bemutató és a mindennapi üzemben használható rendszer között nagy távolság lehet. A műhelyben előkészített

körülmények csökkentik a bizonytalanságot. A valós környezet változásai viszont olyan helyzeteket teremtenek, amelyekre az előre összeállított mozgássor nem készítette fel a szerkezetet.

Önműködés és önállóság

Az ókori automaták döntő újítása, hogy egy elindított folyamat egy részét a szerkezet belső kapcsolatai vezetik tovább. A kezelőnek nem kell minden mozdulatot külön elvégeznie. Ez az önműködés azonban nem azonos a célok önálló kialakításával. A templomajtó nem határoz arról, kit enged be; a színpad nem értékeli a közönség tetszését; a síp nem tanul új madárdalt.

A működés ugyanakkor több lehet egyszerű ismétlésnél. Egy szelep állapota függhet a vízszinttől, az emelőé a terheléstől. A környezet befolyásolja, mi történik, de a lehetséges válaszokat továbbra is a fizikai szerkezet határolja körül. A későbbi fejezetekben ezért az érzékelést, a visszacsatolást és a tanulást is külön kell vizsgálnunk. Egyik sem következik pusztán abból, hogy valami mozog.

Az emberi munka ráadásul nem tűnik el az automata működéséből. Átköltözik a tervezésbe, a gyártásba és az előkészítésbe. Aki a köteleket elrendezte, az előre eldöntötte, mit fog a néző látni. A mechanizmus a készítő egy korábbi cselekvését őrzi meg anyagi formában. Ez az automatizálás egyik mélyebb jelentése: a jelenben működő gépben egy korábbi emberi döntés folytatja a hatását.

Alexandria szerkezetei az élet néhány jelét már tárgyakkal állították elő. A következő évszázadok mesterei a mozgások időzítését tették egyre finomabbá. A vízórák, majd a fogaskerekek és rugók világában az utánzás nemcsak látványosabb, hanem pontosabban ismételhető is lett. Ezzel új kérdés született: milyen bonyolult viselkedést lehet előre beleszerkeszteni egy testbe?

3. Óraművek, zenélő figurák és mechanikus állatok

A mozgás időrendje

Egy alak megmozdítása még nem ugyanaz, mint egy viselkedés előadása. Az élet benyomását gyakran a részletek időzítése kelti: a kéz megáll, a fej odafordul, a dallam megszólal, majd a figura ismét nyugalomba kerül. A középkor és a kora újkor automatamesterei ezt az időrendet tették egyre összetettebbé. Munkájukhoz a vízmérnökség, a fémművesség, az órásmesterség és a színházi látvány technikája egyaránt hozzájárult.

A történet nem egyszerűen Alexandriából halad egyenesen az európai órásműhelybe. Szövegek fordultak le más nyelvekre, eljárások éltek tovább és alakultak át, de sok részlet elveszett. A fennmaradt leírások nem jogosítanak fel minden köztes láncszem feltételezésére. A különböző műhelyek gyakran hasonló feladatokra találtak megoldást: hogyan folyjon egyenletesebben a víz, mikor oldjon ki egy retesz, hogyan ismétlődjön meg pontosan egy mozdulat?

Ebben a folyamatban különösen értékessé vált a szerkesztés átadhatósága. Egy egyszeri udvari látványosság elkápráztathatta a közönséget; egy érthető rajzokkal és gyártási tanácsokkal kísért leírás más mestereket is munkára késztethetett. A gép fennmaradását így nemcsak az anyag tartóssága határozta meg. A tudás megőrzése és újraértelmezése legalább ilyen fontos volt.

Al-Dzsazarí és a vízbe rejtett idő

Al-Dzsazarí a tizenkettedik század végén és a tizenharmadik század elején az artukida udvar mérnöki környezetében dolgozott, a mai Diyarbakır vidékén. A különleges mechanikai szerkezetekről írt, 1206-ra elkészült könyve vízórákat, folyadékot adagoló eszközöket, kézmosó berendezéseket és vízemelő gépeket ismertet. Az ábrák a külső forma mellett a működtetéshez szükséges belső elrendezés megértését is szolgálják. [11–13]



7. kép Al-Dzsazarí elefántórája egy 1315-ben készült kéziratmásolat lapján. Másoló: Farrukh ibn Abd al-Latif. A gép leírása korábbi, az ábrázolt lap azonban nem 1206-ból való. The Metropolitan Museum of Art, 57.51.23. CC0 / közkincs.

A legismertebb példa az elefántóra. A gazdagon díszített, elefánt alakú alépítmény fölé ember- és állatfigurák emelkednek. Az idő múlását a vízzel működő belső rendszer és az egymáshoz kapcsolt mozdulatok teszik láthatóvá. A Metropolitan Museum 1315-ben másolt kéziratlapjának leírása szerint félóránként megszólalt a kupolán álló madár, egy emberalak golyót juttatott a sárkány szájába, az elefánthajcsár pedig megmozdította ösztökéjét. [11]

Az óra a múló időt eseménnyé alakítja. A nézőnek nem feltétlenül kell egy számlap apró jelzéseit figyelnie: a mozgó szereplők és a hangok jelzik a szakaszhatárt. A folyamat belsejében a víz mozgása és a kioldó mechanizmusok dolgoznak, kívül jelenet zajlik. Ugyanaz a berendezés mérőeszköz, udvari műtárgy és színházi előadás.

Az ember alakú kézmosó és italt kínáló eszközök más mindennapi cselekvéseket emeltek át a mechanika világába. A Smithsonian gyűjteményében fennmaradt kéziratlap például korsót tartó nőalak formájú kézmosó berendezést ábrázol. A forma társas jelentést ad a folyadék kiadásának: az eszköz úgy jelenik meg, mintha kiszolgálná a vendéget. Ettől még nem ismeri fel a szomjúságot vagy az udvariasság szabályait. [12]

Al-Dzsazarí jelentősége nem merül ki abban, hogy utólag a robotika előfutárának nevezzük. Saját korának problémáira dolgozott ki megoldásokat, és leírta, hogyan lehet mozgást mérni, átadni és időzíteni. Könyvének későbbi másolatai azt is megmutatják, hogy a szerkezetek tudása szövegben és képben tovább élhetett akkor is, amikor maguk a berendezések már nem maradtak fenn. A kéziratlapot ezért nem szabad egy fennmaradt óra fényképével egyenértékű bizonyítéknak tekinteni.

Az európai órásmesterek mozgó világa

A mechanikus óra olyan feladatot tesz a műhely középpontjába, amely az automaták számára is alapvető: a mozgás szabályozott adagolását. A felemelt súly vagy a felhúzott rugó energiája önmagában gyorsan felszabadulna. A fogaskerék-áttételek és a gátszerkezet azt teszik lehetővé, hogy a gép mozgása kezelhető ütemet kapjon. Ha az időrend adott, ugyanahhoz a hajtáshoz mutató, harang, ütőfigura vagy összetettebb jelenet kapcsolható.

Az európai városok és udvarok mozgó figurái egyszerre mutatták a pontos szerkesztés és a reprezentáció értékét. Az óra jelzett, a hozzá kapcsolt alak pedig eljátszotta a jelzés pillanatát. A néző számára az elvont idő látható eseménnyé vált. A mechanizmus sikere azon múlott, hogy a különböző mozgások megfelelő pillanatban találkozzanak, és a díszítő burkolat mögött elegendő hely maradjon a hajtásnak.

A tizenhatodik–tizenhetedik századi Augsburg az automataórák fontos központja lett. A fennmaradt tárgyak között oroszlánok, elefántok, madarak és mozgó hajók találhatók. A Metropolitan Museum sas alakú, körülbelül 1630-ban készült órája jól szemlélteti e műfajt. A rugós szerkezet az időjelzéshez mozgást kapcsolt: a sas jogarja fel-le járt, és eredetileg a csőre is mozgott. Az óra nem pusztán funkcionális tárgy, hanem a megrendelő hatalmának és gazdagságának jelképe is volt. [14]



8. kép Sas alakú automataóra, Augsburg, 1630 körül. A tárgy az időjelzést mozgó állatfigurával kapcsolja össze. The Metropolitan Museum of Art, 29.52.14. CC0 / közkincs.

Ezeket a tárgyakat nem egyetlen mester elszigetelt találmányaiként érdemes elképzelni. Az órás munkáját öntők, vésnökök, asztalosok és más kézművesek egészítették ki. Egy mozgó aranyozott állat egyszerre anyagtechnológiai, gépészeti és művészi teljesítmény. A látvány egységes, a létrehozásához szükséges tudás viszont sokféle.

Az automataóra egy további különbséget is láthatóvá tesz. Az ember gondosan kiválasztott pillanatban lát élénk mozgást; a gép közben egy állandó szerkezetet követ. A látványos figura nem az egész belső működés leképezése. A nézőnek készített jelentés és a gépben zajló erőátadás két különböző leírása ugyanannak a tárgynak.

Vaucanson és a test működésének utánzása

Jacques de Vaucanson az 1730-as évek végén olyan automatákkal vált ismertté, amelyek emberi és állati tevékenységet utánoztak. Fuvolása esetében a feladat nem egyszerűen az volt, hogy egy szobor mellett dallam szóljon. A levegőt a hangszerhez kellett vezetni, a hangképzés feltételeit és az ujjak mozgását pedig össze kellett hangolni. A korabeli leírás részletesen tárgyalta a műveletek mechanikai megvalósítását. [15]

Ez a különbség az utánzás történetében lényeges. Egy hangszert tartó figura és a talapzatba rejtett zenedoboz a nézőnek együtt muzsikusnak látszhat. A hangképzés fizikai folyamatát valóban elvégző szerkezet azonban mást bizonyít: megmutatja, hogy a tevékenység bizonyos részei gépileg is megvalósíthatók. Ettől még nincs zenei ízlése vagy előadói szándéka. A működés megfejtése és a mentális állapot tulajdonítása külön feladat.



9. kép Vaucanson három automatája az 1738-as Le Mécanisme du flûteur automate című kiadvány nyitó metszetén. Középen a kacsa, kétoldalt a zenélő emberalakok. Forrás: Bibliothèque nationale de France / Wikimedia Commons. Közkincs.

Vaucanson mesterséges kacsája még erősebb állítással vonzotta a figyelmet. A mozgó, táplálékot felvevő és üríteni látszó állat azt a benyomást keltette, hogy az élet belső folyamatait is sikerült géppé alakítani. A szerkezet valódi emésztését azonban nem tekinthetjük bizonyítottnak. A történeti beszámolók és későbbi magyarázatok szerint a táplálék felvétele és az ürüléknek mutatott anyag kibocsátása nem egy valódi biológiai emésztési lánc két vége volt. [15–16]

A gép hatásának kulcsa a néző következtetése. Amit bemenetként látott, azt összekapcsolta azzal, amit kimenetként látott. A köztes folyamatot a szerkezet burkolata és a bemutató elbeszélése fedte el. Így a külső eseménysor többet sugallt, mint amennyit a belső működés igazolt. Ez nem von le a mozgások összehangolásának teljesítményéből, de korlátozza, milyen élettani következtetés vonható le belőle.

A kacsa eredeti példánya nem áll rendelkezésünkre sértetlenül elvégezhető mai vizsgálathoz. A népszerű belső rajzok egy része későbbi rekonstrukció; a tizenkilencedik század végén közölt ábrát sem helyes Vaucanson eredeti gyártási dokumentumának tekinteni. A gép működéséről szóló állításoknál ezért meg kell nevezni, mit ismerünk a készítő leírásából, mit későbbi beszámolóból, és mit csupán utólagos elképzelésből. [16]

A bütykös tárcsába írt mozdulat

Hogyan tud egy óramű kézmozdulatot, fejbiccentést vagy tollvonást előállítani? A fogaskerék elsősorban forgást ad tovább. Ha összetettebb mozdulatot akarunk, a forgást más pályává kell alakítani. Erre alkalmas a bütykös tárcsa: körvonala vagy hornya mentén egy követőelem halad, amely a tárcsa forgása közben előírt módon elmozdul. A követőhöz rudazat, kéz vagy más testrész kapcsolható. [17]

A mozgás menete a tárcsa alakjában van rögzítve. Egy hosszabban változatlan szakasz kitartást, más profil emelést vagy süllyesztést eredményezhet. Több összehangolt tárcsa több mozgásösszetevőt

vezérelhet. A Franklin Institute-ban őrzött, későbbi Maillardet-automata mechanikai memóriájának leírása különösen jól megmutatja ezt az elvet. Ugyanez a szerkesztési gondolat segít megérteni a tizennyolcadik század rajzoló és író alakjait is. [17]

Itt az információ nem betűként és nem elektromos állapotként jelenik meg. Egy fémalkatrész geometriája határozza meg, hogyan mozogjon a szerkezet. A tárcsa megváltoztatásával a viselkedés megváltozhat, de a gép ettől még nem tanul. A változtatást a készítő vagy a kezelő végzi, és a megvalósítható mozgásoknak szigorú mechanikai korlátjuk van.

A „mechanikai memória” ezért hasznos kifejezés, ha pontosan értjük. A tárgy megőriz és ismétel egy korábban kialakított mintát. Nem emlékezik úgy, ahogyan egy ember felidéz egy élményt. Az információ tárolása, a mozgássor végrehajtása és a jelentés megértése különböző szintek. Összekapcsolódhatnak egy későbbi rendszerben, de egyikből sem következik automatikusan a másik.

A Jaquet-Droz műhely írója rajzolója és zenésze

Pierre Jaquet-Droz, fia, Henri-Louis és munkatársuk, Jean-Frédéric Leschot 1768 és 1774 között hozta létre a három híres ember alakú automatát. Az Író, a Rajzoló és a Zenésznő ma Neuchâtel művészeti és történeti múzeumának gyűjteményében található. Ezek már olyan tevékenységeket mutattak be, amelyekhez a közönség rendszerint értelmet és műveltséget társított. [18]

Az Író tollal ír, és előre beállított szöveget képes papírra vetni. A múzeum leírása szerint legfeljebb negyven jelből álló szöveg állítható be, három sorban. Különös jelentősége éppen az, hogy nem kizárólag egyetlen változatlan mondatot ismétel: a mechanikai beállítás módosításával más betűsorrend íratható vele. Az átalakítható tartalom és a betűrajzolás módja tehát részben elkülönül egymástól. [19]



10. kép A Jaquet-Droz műhely Író automatája, 1768–1774. A figura beállított betűsorrendet ír le; a szöveg megváltoztathatósága nem jelent nyelvi megértést. Musée d’art et d’histoire de Neuchâtel. Fotó: Gre regiment, CC BY-SA 4.0.

Az Író nem azért különleges, mert érti, amit leír. Nem tudja megítélni, igaz-e a mondat, helyénvaló-e a megszólítás, vagy megsérti-e a címzettet. Teljesítménye a jelalakok pontos előállítása és a beállított sorrend követése. A szavak jelentése a beállító ember és az olvasó világához tartozik. A tárgy ettől még fontos állomás: a testhez kapcsolt, módosítható utasítássor kézzelfogható példája.

A Rajzoló képeket készít, miközben időnként elfújja a ceruzapor maradékát. A múzeum katalógusa cserélhető bütyökkészleteket és előre meghatározott rajzokat sorol fel. Az alkotó munkájára emlékeztető apró gesztus különösen hatásos: a néző nemcsak a vonalakat látja, hanem egy rajzolással elfoglalt alakot. A gesztus mégsem azt jelenti, hogy az automata maga találta ki a képet vagy a munkamódszert. [20]

A Zenésznő valódi orgonán szólaltat meg dallamokat, ujjai a billentyűket nyomják. A múzeum leírása öt darabot, a légzés látszatát keltő mozgást és az előadást lezáró meghajlást említ. Itt a hangszer megszólaltatása és az előadó társas megjelenése egyszerre része a gépbe szerkesztett folyamatnak. A test nem dísz a zene mellett: a bemutató értelmét adó szereplő. [21]

A három alak együtt a tanult ember képességeit állította színpadra. Írás, rajzolás és zenélés: a kéz finom mozgásaihoz kulturális jelentés társult. A meglepetést ezért nem pusztán az alkatrészek száma okozta. A néző olyan cselekvést ismert fel, amelyről hajlamos volt azt feltételezni, hogy mögötte az emberéhez hasonló belső élet áll.

A mechanikus állat és a néző képzelete

A mechanikus állatok vonzerejét nem feltétlenül a sokféle művelet adta. Egy jól eltalált nyakmozdulat, szárnyrebbenés vagy hang gyakran erősebben hatott, mint a bonyolultság önmagában. Az 1773 körül James Cox londoni műhelyéhez kapcsolódó Ezüst Hattyú például üvegrudakkal utánozza a vizet, fej- és nyakmozgással a tollászkodást, majd egy hal elkapásának és lenyelésének látszatát kelti. A Science Museum ismertetője szerint a lenyeltnek látszó hal már korábban a csőrében rejtőzött. [22]

A hattyú mutatványa nem tesz úgy, mintha valódi ragadozó életmódját minden részletében újraalkotta volna. Egy rövid esemény hiteles benyomásához elég a kiválasztott mozdulatok összhangja. Az anyag más: ezüst, üveg, fém. A viselkedés egy töredéke mégis felismerhető. Az élet utánzásához tehát nem mindig az élőlény belső felépítését kell másolni; sokszor az emberi észlelés számára döntő jeleket kell megtalálni.

A korábbi kibernetikai ismeretterjesztéshez is kapcsolódó kérdés, hogy mennyit lehet egy élőlény működéséből más anyagokkal megvalósítani. Csató István könyvei és az 1965-ös A kibernetika klasszikusai már későbbi tudományos környezetben tárgyalják az ember, a gép, az információ és az irányítás kapcsolatát. Az automaták történetéhez olvasási szempontot adnak: ne elégedjünk meg azzal, hogy valami emberinek látszik, hanem vizsgáljuk meg, milyen szerveződés hozza létre a megfigyelt teljesítményt. [23]

Ez a szempont a történeti automaták értékét is pontosabban mutatja meg. Az Író nem alkot gondolatot, de módosítható szöveget állít elő. A Zenésznő nem érez zenei örömöt, de működő hangszert szólaltat meg. A hattyú nem táplálkozik, de az ember számára felismerhető cselekvéssort ad elő. A hiányzó képességek megnevezése nem teszi jelentéktelenné a megvalósítottakat.

Az élet látszatától a viselkedés vizsgálatáig

A mítoszokban a teremtmény egy döntő pillanatban életet kapott. Az automaták műhelyében az életre emlékeztető viselkedés részekre bontható feladattá vált. A fej elfordulása, a levegő áramlása, a toll pályája és a mozdulatok sorrendje külön szerkeszthetőnek bizonyult. Ez az elemzés később sokkal fontosabb örökségnek bizonyult, mint az, hogy a gép külső alakja emberre vagy állatra hasonlított.

A három fejezetből nem az következik, hogy a bronzóriástól szükségszerű út vezetett a mesterséges intelligenciához. Inkább háromféle törekvés rajzolódik ki. A képzelet az emberi szándékhoz hasonló cselekvőt keresett. A mérnöki műhely megoldható mozgási és időzítési feladatokat talált. A közönség pedig az elkészült tárgy viselkedéséből következtetett arra, mi lehet belül.

A következő nagy kérdés az lesz, hogyan változik meg a helyzet, ha a gép nemcsak végrehajt egy előre rögzített mozdulatsort, hanem saját működésének eredményéről is kap jelzést. Még később az is felmerül, hogyan módosíthatja tapasztalatból a válaszait. Az órásmesterek nem oldották meg ezeket a problémákat. Olyan tárgyakat készítettek azonban, amelyek megkerülhetetlenné tették a kérdést: hol húzódik a határ a mozgás, a célszerű viselkedés és az értelem között?

Források és további olvasmányok

A szögletes zárójelben álló számok az alábbi tételekre utalnak. Az ókori művek történeti kiadásainak bevezetői elavult keltezéseket is tartalmazhatnak; a fejezetek Hérón időszakánál a mai múzeumi és tudománytörténeti besorolást követik.

[1] Apollóniosz Rhodiosz: Argonautika, IV. 1638–1693; Pszeudo-Apollodórosz: Mitológia, I. 9. 26. A Talósz-történet eltérő változatai. Ókori szövegrészletek

[2] Homérosz: Íliász, XVIII. 369–420. Héphaisztosz műhelye, önmozgó háromlábak és arany szolgálólányok. Szövegrészlet

[3] Ovidius: Átváltozások, X. 243–297. Pügmalión története. Henry T. Riley angol prózafordítása, 1851; 1893-as kiadás. Teljes szöveg

[4] Jewish Encyclopedia: Golem. 1906. A gólemmotívum vallási és elbeszélő hagyományairól. Enciklopédiaszócikk

[5] Židovské muzeum v Praze: The Golem of Prague The Legend of Rabbi Loew’s Creation. A prágai legendakör múzeumi ismertetője. Múzeumi ismertető

[6] Real or Fake? New Light on the Paracelsian De natura rerum. Ambix, 2020. DOI: 10.1080/00026980.2020.1720339. A mű hagyományozása és vitatott szerzősége. Tanulmány kivonata

[7] Mary Shelley: Frankenstein or The Modern Prometheus. Első kiadás 1818. Szöveg és kiadási adatok

[8] Hérón: The Pneumatics of Hero of Alexandria. Fordítás J. G. Greenwood közreműködésével; szerkesztette Bennet Woodcroft. London, 1851. Különösen a 14–15., 21., 37–38., 43–44. és 50. szerkezet. Szöveg és ábrák

[9] Kotsanas Museum: The automatic opening of the temple gates. Hérón ajtónyitó szerkezetének mai rekonstrukciója és működési leírása. Múzeumi leírás

[10] Kotsanas Museum: The mobile automaton of Heron of Alexandria; The static automatic theatre of Philon. A rekonstrukciókat eredeti, fennmaradt gépnek nem tekintjük. Mozgó színház · Álló színház

[11] The Metropolitan Museum of Art: The Elephant Clock. 1315-ös kéziratlap, leltári szám 57.51.23. A tárgyleírás az időjelző jelenetet is ismerteti. Műtárgykatalógus

[12] Smithsonian National Museum of Asian Art: A hand-washing device in the form of a woman holding a pitcher. 1315-ös kéziratlap, F1930.75. Műtárgykatalógus

[13] Routes of Translation Connected Book Histories and al-Jazari’s Robotic Wonders from the Mamluks to Mandu. South Asian Studies. DOI: 10.1080/02666030.2023.2287840. Az 1206-os kézirati hagyomány és a fordítások története. Kutatási tanulmány

[14] The Metropolitan Museum of Art: Automaton clock in the form of an eagle. Augsburg, 1630 körül. Leltári szám 29.52.14. Tárgyleírás és katalógus

[15] Jacques de Vaucanson: An Account of the Mechanism of an Automaton or Image Playing on the German-Flute. Fordította J. T. Desaguliers, 1742. Vö. Conrad William Cooke: Automata Old and New. London, 1893. A korabeli működési állítások forráskritikát igényelnek. Vaucanson műve · Cooke könyve

[16] William B. Ashworth Jr.: Jacques de Vaucanson. Linda Hall Library, Scientist of the Day, 2023. november 21. A kacsa működése és a későbbi belső ábrák eredete. Könyvtári tárgy- és forrásismertető

[17] The Franklin Institute: Maillardet’s Automaton. A bütykös tárcsákra épülő mechanikai memória működése. A tárgy későbbi a Jaquet-Droz-automatáknál. Gyűjteményi ismertető

[18] Musée d’art et d’histoire de Neuchâtel: Automates Jaquet-Droz. A három automata készítése és gyűjteményi története. Múzeumi ismertető

[19] Musée d’art et d’histoire de Neuchâtel: Automate L’Écrivain. AA 2. A negyvenjelnyi, háromsoros szöveg beállíthatósága. Műtárgykatalógus

[20] Musée d’art et d’histoire de Neuchâtel: Automate Le Dessinateur. AA 3. Cserélhető bütyökkészletek és a rajzolás mozdulatai. Műtárgykatalógus

[21] Musée d’art et d’histoire de Neuchâtel: Automate La Musicienne. AA 1. A valódi orgona, az öt dallam és az előadói mozdulatok. Műtárgykatalógus

[22] Howard Coutts: The Silver Swan. Science Museum Blog, 2016. A mozgások, a három hajtás és a hal elkapásának látszata. Kurátori ismertető

[23] Csató István: A kibernetika Az információ forradalma. Kossuth, 1969; második, javított kiadás 1971. Uő: A kibernetika és az ember. Kossuth, 1970. A kibernetika klasszikusai. Válogatta Szalai Sándor; bevezetés és összekötő szövegek Tarján Rezső; fordította Tarján Rezsőné. Gondolat, 1965. E művek a későbbi kibernetikai szemlélet magyar olvasmányai, nem az ókori automaták elsődleges forrásai. Csató kötetének archívumi adatai

Kapcsolódó magyar olvasmány

Toby Walsh: A mesterséges intelligencia legrövidebb története. Athenaeum, 2025. A teljes AI-történet további szakaszaihoz kapcsolódó áttekintés.

Képforrások és felhasználási adatok

A képek a tárgyalt szövegrészekben, saját képaláírással szerepelnek. A közkincs megjelölés a történeti műre, a CC0 a múzeumi nyílt hozzáférésű képre, a CC BY-SA 4.0 a megnevezett fényképész felvételére vonatkozik.

1. kép Eredeti képfájl vagy gyűjteményi adatlap · CC BY-SA 4.0 · A licenc feltételei

2. kép Eredeti képfájl vagy gyűjteményi adatlap

3. kép Eredeti képfájl vagy gyűjteményi adatlap

4. kép Eredeti képfájl vagy gyűjteményi adatlap

5. kép Eredeti képfájl vagy gyűjteményi adatlap

6. kép Eredeti képfájl vagy gyűjteményi adatlap

7. kép Eredeti képfájl vagy gyűjteményi adatlap

8. kép Eredeti képfájl vagy gyűjteményi adatlap

9. kép Eredeti képfájl vagy gyűjteményi adatlap

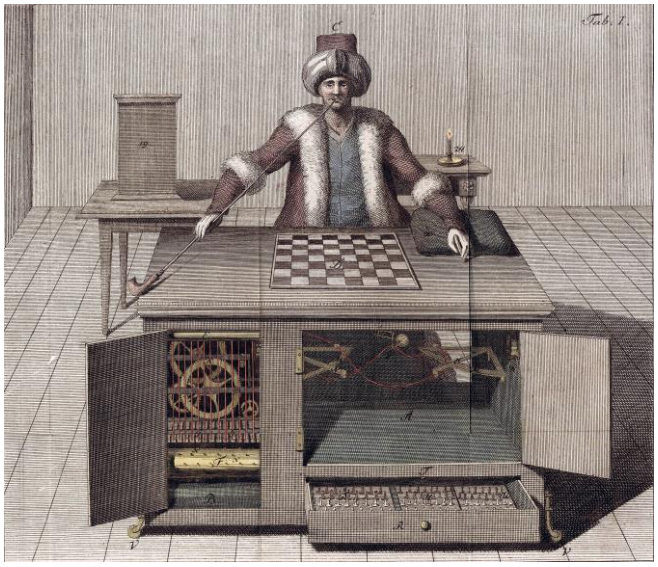
10. kép Eredeti képfájl vagy gyűjteményi adatlap · CC BY-SA 4.0 · A licenc feltételei

4. A gép, amelyről elhitték, hogy gondolkodik

Egy sakktábla az udvari közönség előtt

A néző előtt turbános alak ül egy szekrény mögött. Az asztallapon sakktábla, a nyitott ajtók mögött fogaskerekek és más mechanikus alkatrészek látszanak. A bemutató azt sugallja, hogy a szerkezet belsejét már megvizsgáltuk. Az ajtókat becsukják, elkezdődik a játszma, és a figura olyan lépéssel válaszol, amelyre az ellenfélnek komolyan el kell gondolkodnia. Az órásmesterek mozgó alakjai után ez már egészen más teljesítménynek tűnik.

Kempelen Farkas, a Pozsonyban született, 1734 és 1804 között élt feltaláló és udvari tisztviselő sakkozó Törökje 1770 körül vált a Habsburg-udvar szenzációjává. A külföldi források Wolfgang von Kempelen néven említik. A mutatvány később Európa több városában, majd Kempelen halála után Amerikában is közönséget vonzott. A „gondolkodó gép” lehetősége és a rejtett ember gyanúja kezdettől egymás mellett kísérte. A kétkedés éppen úgy része a történetének, mint a csodálat. [24]



11. kép A sakkozó Török Joseph Friedrich von Racknitz 1789-es könyvében. A nyitott ajtók és a látható mechanika a bemutató fontos részei voltak. A kép korabeli ábrázolás, nem fennmaradt tárgy fényképe. Humboldt-Universität könyvtára / Wikimedia Commons. Közkincs.

A Török a néző számára nem egyszerűen egy sakkozó eszköz volt. Ruhája, emberi alakja és mozgó keze szereplővé tette. Úgy tűnt, valaki ül velünk szemben, és szándékosan választ a lehetőségek közül. A bemutató így egyszerre használta a mechanika teljesítményét és az ember hajlamát arra, hogy a célszerű viselkedés mögé személyt képzeljen.

Miért volt a sakk több egy előre megírt jelenetnél

Az előző fejezetben megismert Író automatának előre be lehetett állítani, milyen betűsort írjon le. A sakkozó helyzete ennél nehezebb. Nem elegendő egy szép mozdulatsort megőrizni: az ellenfél beavatkozik a folyamatba. Már néhány lépés után olyan állás alakulhat ki, amely eltér a korábban lejátszott partitól. A megfelelő válaszhoz ismerni kell az aktuális helyzetet.

A feladatot érdemes részekre bontani. A játékosnak érzékelnie kell, melyik mezőn milyen bábu áll; meg kell különböztetnie a szabályos és szabálytalan lépéseket; választania kell a lehetséges folytatások között; végül végre kell hajtania a kiválasztott lépést. Egy szépen mozgó kar csak az utolsó műveletre ad választ. A közönség mégis könnyen az egész feladat megoldását tulajdonítja annak a szerkezetnek, amelynek a kezét látja.

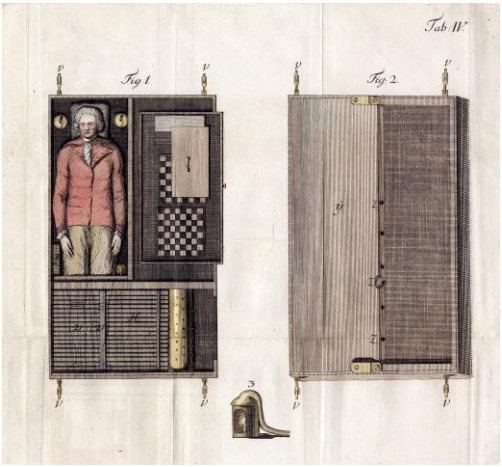
Az automataszínházban a készítő előre tudta, mikor következik a következő jelenet. A sakkban egy rögzített előadás hamar lelepleződne: az ellenfélnek elegendő volna más lépést választania. Ez nem azt jelenti, hogy elágazó helyzeteket gép sohasem kezelhet. Azt jelenti, hogy a puszta időzítéshez új képességeket kell hozzáadni. A tábla állapotának és a válasz kiválasztásának gépi kezelése volna az igazi eredmény.

A rejtett játékos és a látható gép

A Törökben a sakkozói döntést ember hozta meg. A szekrény belsejében elhelyezkedő játékos mechanikai közvetítéssel irányította a figura kezét. A rendszer a tábla állásának követését is lehetővé tette: a múzeumi ismertetők a mágneses jelzések szerepét hangsúlyozzák. A gép tehát egy távolabb látszó kézen keresztül jelenítette meg a bent ülő ember cselekvését. [24]

A nyitogatott ajtók és a megmutatott fogaskerekek a bizonyítás látszatát keltették. A néző azt gondolhatta, hogy ahol alkatrészeket látott, ott ember nem férhet el. Valójában nem ő határozta meg, milyen sorrendben, milyen nézőpontból és meddig vizsgálhatja a belsőt. A bemutató irányítója a figyelmét is vezette. Egy rendszer részleges feltárása ezért teljes átláthatóságnak tűnhetett.

A kortársak különféle magyarázatokkal próbálkoztak. Joseph Friedrich von Racknitz 1789-es könyve például rajzokkal mutatott be egy lehetséges elrendezést, rejtett emberrel. Az ilyen ábrák a titok megfejtésére tett kísérletek dokumentumai. Nem mindegyik részletüket igazolja a későbbi ismeret, ezért nem használhatók úgy, mintha Kempelen hiteles szerelési rajzai volnának. A valóban elrejtett ember ténye és egy adott magyarázó ábra pontossága két külön kérdés. [26]



12. kép Racknitz 1789-es magyarázati kísérlete a rejtett ember elhelyezéséről. Az ábra a titok megfejtésére tett korabeli kísérletet dokumentálja; nem Kempelen hiteles belső szerelési rajza. Humboldt-Universität könyvtára / Wikimedia Commons. Közkincs.

A megtévesztéshez mindazonáltal valódi műszaki munka kellett. Embernek kellett elférnie a ládában, követnie kellett a játékot, és a mozdulatát át kellett vinni a látható karra. A mutatvány ügyességét

nem érdemes letagadni. A határt ott kell meghúzni, amit a gépnek tulajdonítottak: a gondolkodás feladatát nem a fogaskerekek végezték el.

Poe helyes gyanúja és túl erős következtetése

A sakkozó titkát Edgar Allan Poe is vizsgálta. Az 1836-ban megjelent Maelzel's Chess-Player című írásában a bemutató részleteiből próbálta kikövetkeztetni az emberi közreműködést. Akkor a Török már Johann Nepomuk Mälzel tulajdonában járta a világot. Poe az ajtók nyitásának rendjére, a figura mozgására és a bemutató körülményeire egyaránt figyelt. [25]

Írásának érdekessége nem merül ki abban, hogy helyesen gyanakodott. A gépi számolást lényegében kötött műveletsornak tekintette, a sakkot viszont a játékos változó ítéletére bízott tevékenységnek. E különbségből túlságosan erős következtetést vont le: az adott mutatvány emberi irányítását a gépi sakkozás lehetetlenségének gondolatával kapcsolta össze. [25]

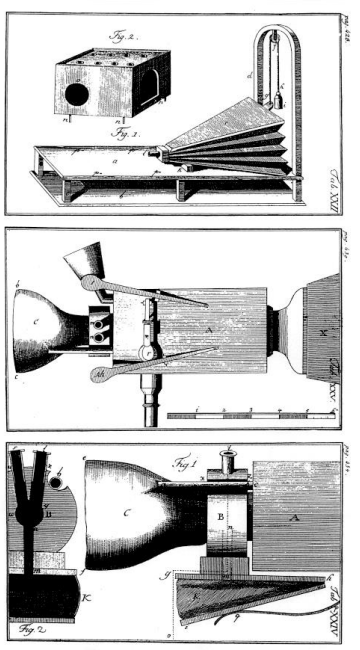
Mai szemmel látszik a logikai rés. Attól, hogy az ellenfél lépését nem lehet előre tudni, még megadható olyan eljárás, amely a létrejött állapotból választ folytatást. Az elágazások kezelése nem kívánja meg, hogy a gép a játszma előtt már az egész partit ismerje. Azt kell meghatározni, hogyan képviselje az állást, milyen folytatásokat vizsgáljon, és milyen szempont alapján válasszon.

Poe példája ezért két tanulságot hordoz. A bemutató kritikus megfigyelése eredményes lehet, miközben a megfigyelő általános elmélete részben téves. Egy mutatvány leleplezése nem bizonyítja, hogy a bemutatott képesség mesterséges megvalósítása elvileg lehetetlen. A Török valóban nem gondolkodott. Ebből azonban nem következik, hogy semmilyen gép nem tudhat sakklépést választani.

A beszélőgép valódi kérdése

Kempelen másik nevezetes szerkezete egészen eltérő megítélést érdemel. Beszélőgépében a hangot valóban mesterséges berendezés hozta létre. Nem egy elrejtett ember beszédét vezették ki a dobozból. A cél a beszédképzés fizikai feltételeinek megértése és részleges utánzása volt. A kutatás eredményét Kempelen 1791-ben Bécsben megjelent, az emberi beszéd mechanizmusáról és beszélőgépéről szóló könyvében tette közzé. [27–29]

A feladat nem egyszerűen abból állt, hogy minden hanghoz egy sípot kell keresni. A beszédben ugyanaz a hangképző rendszer sokféle hangot hoz létre, a részek pedig mozgás közben alakítják egymás hatását. A különálló hangok egymás mellé helyezése és a folyamatos, érthető beszéd között jelentős különbség van. A gép megépítése arra kényszerítette a feltalálót, hogy az emberi beszélőszervek működését is megvizsgálja. [27]



13. kép Kempelen 1791-es beszélőgépének rajzai. A fújtató, a hangkeltő egység és a kezelő által alakított szájrész együtt hozta létre a beszédre emlékeztető hangokat. Forrás: Wikimedia Commons. Közkincs.

A fújtató a tüdőhöz hasonlóan levegőáramot biztosított. Egy rezgő nyelv hangforrást adott, az üregek és járatok pedig alakították a hangot. A változtatható szájnyílás, valamint a külön nyitható és zárható járatok különböző hanghatásokat tettek lehetővé. A szerkezet nem másolta le anatómiai pontossággal az ember egész beszédképző rendszerét; egyes részek szerepét más anyaggal és eltérő formával valósította meg. [27–29]

A beszélőgépet kezelni kellett. A megszólalás időzítését, a levegő adagolását és a hangformáló elemek mozgatását gyakorlott ember hangolta össze. Helyesebb ezért beszédkeltő hangszerként elképzelni, mint előre megírt mondatokat önállóan felmondó automataként. A kezelő jelenléte itt nem titkolt helyettesítése a gép eredményének: a kéz formálta a hangot, de az akusztikus rezgést a szerkezet állította elő. [28–29]

E különbség világítja meg a műszaki teljesítményt. Egy fuvola sem válik hamis hangszerré attól, hogy zenész szólaltatja meg. A kérdés az, milyen fizikai folyamatot végez el a hangszer, és mit tesz hozzá a kezelő. Kempelen berendezése a beszéd egy részét mérnöki vizsgálat tárgyává tette. Nem kellett hozzá gondolatot tulajdonítani a fújtatónak.

A leírástól a megismételhető kísérletig

Az 1791-es könyv jelentősége az, hogy a bemutatott hatást összekapcsolta a szerkezet és a hangképzés leírásával. A titok többé nem a siker feltétele volt. Mások megvizsgálhatták az elképzelést, újraépíthették a berendezést, és összevethették az eredményt a szerző állításaival. A mű így az egyszeri látványosságon túl is használható tudást őrzött meg. [27]

Nikléczy Péter és Olaszy Gábor 2001-ben működő rekonstrukciót készített; eredményeikről 2003-ban nemzetközi beszédkutatási konferencián számoltak be. Céljuk nem csupán egy múzeumi tárgy külső másolata volt. Meg akarták vizsgálni, milyen hangok hozhatók létre a történeti elven, és ezek

akusztikai tulajdonságai hogyan mérhetők. A kétszáz éves állítás így újra kísérleti kérdéssé válhatott. [29]

A rekonstrukció nehézségei azt is megmutatták, hogy a rajz még nem tartalmaz feltétlenül minden gyártási fogást. A tömítések, a rezgő elem tulajdonságai és a pontos illesztések befolyásolják az eredményt. A Deutsches Museum a saját gyűjteményében őrzött, 1800 körülre datált beszédkeltő szerkezet megértéséhez ugyancsak másolatot készített. A történeti tárgy, az 1791-ben közölt változat és a későbbi rekonstrukció nem kezelhető automatikusan ugyanannak a példánynak. [28–29]

Két különböző szerep az ember számára

A Török és a beszélőgép összehasonlítása az egész könyv számára hasznos különbséget ad. A sakkozónál a közönségnek gépi döntésként bemutatott választ ember hozta meg. A beszélőgépnél az ember irányította a hangképzést, miközben a gép valóban elvégezte az ígért fizikai műveletet. Mindkettőben jelen volt emberi ügyesség, de máshol húzódott a teljesítmény határa.

Egy rendszer vizsgálatakor ezért nem elég azt kérdezni, van-e benne emberi közreműködés. Szinte minden gép mögött találunk tervezőt, készítőt vagy kezelőt. A fontosabb kérdés az, melyik feladatot végzi a gép, melyiket az ember, és erről mit állít a bemutató. A rejtett munkamegosztás könnyen olyan képesség látszatát kelti, amely a berendezésben önmagában nincs jelen.

Kempelen öröksége ebben a kettősségben különösen eleven. A Török azt mutatta meg, mennyire szeretnénk elhinni a gondolkodó gépet, és milyen találékonyan kételkedünk benne. A beszélőgép pedig azt, hogy egy emberi képesség részfolyamatait fel lehet tárni és más anyagokkal működésbe lehet hozni. A két tanulság együtt készít elő a következő lépésre: hogyan adható át a gépnek egy műveletsor úgy, hogy azt ne minden alkalommal az ember kezének kelljen irányítania?

5. Amikor a működés leírássá válik

A minta amelyet újra elő kell állítani

Egy díszes szövet látványa távolinak tűnhet a számítógépek világától. Közelről nézve azonban már az ismétlődés és a választás problémáival találkozunk. A szövőnek egymás után kell létrehoznia a fonalak megfelelő kereszteződéseit. A mintát nem elég elképzelni: a munka minden szakaszában tudni kell, melyik fonalcsoport kerüljön felülre, és melyik maradjon alul.

A szövés két alapvető fonalrendszert kapcsol össze. A hosszában kifeszített láncfonalak között keresztirányban halad a vetülékfonal. Ha a láncfonalak egy részét megemeljük, nyílás keletkezik, ezen átvezethető a vetülék. A szálak egymás fölé és alá kerülésének rendje határozza meg a kötést. A díszes kelme elkészítéséhez ezt a rendet sokszor bonyolult sorozatban kell változtatni. [32–33]

A hagyományos mintás szövésben a szükséges fonalcsoportok kiválasztásához segítő is kellett. A feladat jelentős része tudás és figyelem volt: milyen sorrendben kell a mozdulatokat elvégezni, és hol tartunk a mintában? A későbbi lyukkártyás megoldás ezt a kiválasztási rendet tette géppel olvashatóvá. Nem a szövet szépségét „értette meg”, hanem a megfelelő helyeken érvényesítette az előkészített választásokat. [30, 32]

Jacquard előtt is volt lyukasztott vezérlés

A szövőszék programozhatóságának története nem egyetlen ember ötletével kezdődik. Basile Bouchon 1725-ben lyukasztott papírszalagot alkalmazott a mintavezérléshez. Jean-Baptiste Falcon 1728-ban összefűzött, lyukasztott kártyákkal dolgozott. A papíron levő nyílások az érzékelő elemek mozgását tették lehetővé vagy akadályozták meg. A jel tehát nem pusztán a mester számára volt látható: fizikailag hatott a gép működésére. [30]

Falcon megoldásában a kártyákat még embernek kellett a megfelelő helyzetbe juttatnia. Ez fontos korlát, mert a lyukkártya megjelenése önmagában nem jelenti a teljes folyamat automatizálását. A minta tárolására már született új eljárás, miközben a továbbítás és az összehangolás egy része még a kezelőre maradt. [30]

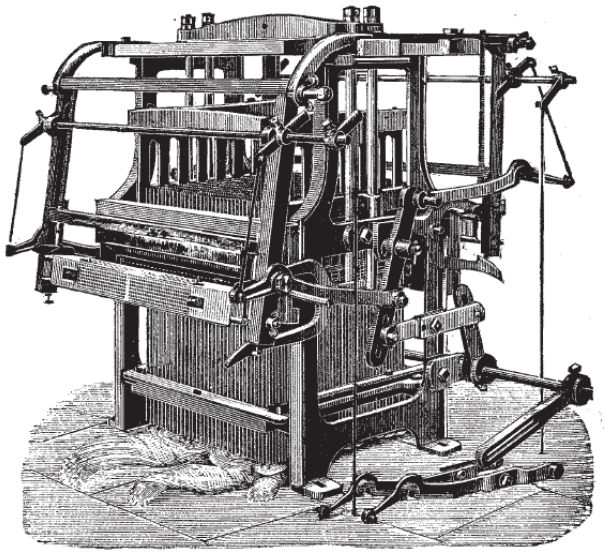
A korábbi fejezetből ismert Vaucanson az 1740-es években a selyemszövés gépesítésével is foglalkozott. A párizsi Musée des Arts et Métiers 1746-ra datált szövőszéke a szövés több műveletének összehangolására törekedett. A múzeum leírása szerint ez a fejlesztés akkor nem jutott ipari alkalmazásig. A látványos és működőképes műszaki elképzelés, valamint az elterjedt termelőeszköz között itt is megmaradt a távolság. [31]

Jacquard tehát olyan hagyományra épített, amelyben a papírra vitt minta, a mechanikai kiválasztás és a munkafolyamat gépesítése már találkozott. Hozzájárulása az volt, hogy ezekből a mintás szövés számára széles körben használható rendszert fejlesztett. Az előzmények ismerete nem csökkenti a jelentőségét; inkább megmutatja, milyen összetett eredmény az, amit egyetlen névvel szoktunk jelölni.

Jacquard gépe és a cserélhető mintasor

Joseph Marie Jacquard, az 1752-ben Lyonban született feltaláló a tizenkilencedik század első éveiben mutatta be és fejlesztette tovább mintavezérlő berendezését. Az 1801-es bemutató és az 1804–1805 körüli továbbfejlesztések egyaránt fontosak; ezért az eltérő évszámok nem feltétlenül egymást kizáró állítások. Egy fokozatosan használhatóbbá tett gépcsalád történetéről beszélünk. [32–33, 37]

A „Jacquard-szövőszék” elnevezés mögött döntő jelentőségű a szövőszékhez kapcsolt kiválasztó szerkezet. A kártyák lyukasztása meghatározta, mely vezérlő elemek engedjenek meg egy mozdulatot. Az ezekhez csatlakozó horgok és zsinórok a láncfonalak megfelelő csoportjait emelték. Az új kártyával új kiválasztási rend lépett érvénybe, a vetülék következő átvezetéséhez igazodva. [32]



14. kép Késő tizenkilencedik századi Jacquard-berendezés Richard Marsden Cotton Weaving című művének ábráján. Fejlett változatot mutat, nem Jacquard első, 1801-es gépét. Forrás: Wikimedia Commons, Marsden-könyvillusztráció. Közkincs.

Egy lyuk helye tehát a gép valamely vezérlési pontjához tartozott. Több kártya egymásutánja a kiválasztások sorozatát adta. A gyakorlati szövőszékekben egy vezérlési pont több, azonos mintarészhez kapcsolt fonalat is mozgathatott; ezért az „egy lyuk egy fonal” magyarázat csak leegyszerűsített modellként használható. A kötés, a befűzés és a fonalak színe együtt alakította ki a kész kelme látványát.

A kártyasor cseréje más mintát tett elérhetővé ugyanazzal az alapgéppel. Az ismétlődő kártyalánc ismétlődő mintát eredményezhetett. Ettől a gép még nem határozta el, milyen virág illik egy ruhához, és nem tervezte meg saját ízlése szerint a kelmét. Az alkotói és előkészítő munka az embernél maradt, miközben annak eredményét a berendezés sok egymás utáni műveletben érvényesítette.



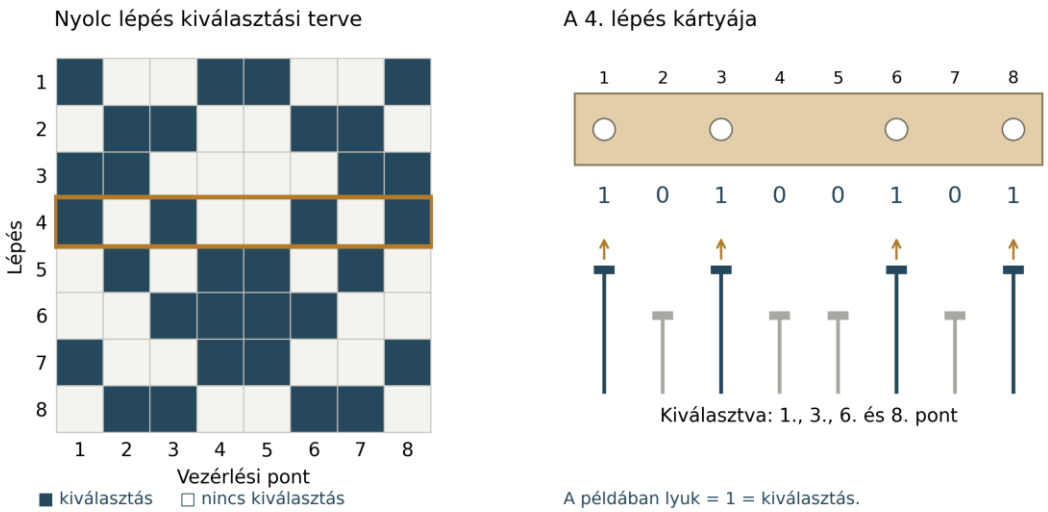
15. kép Összefűzött mintavezérlő kártyák egy Jacquard-szövőszéken, a National Museum of Scotland gyűjteményében. A sorozat a kiválasztások rendjét őrzi. Fotó: Stephencdickson / Wikimedia Commons, CC BY-SA 4.0.

A kézi hajtású Jacquard-szövőszéken a munkásnak továbbra is feladata volt. A mintakiválasztás automatizálása és a gép energiaellátása különböző kérdés. A kártyák nem adtak hajtóerőt: a mozdulatok rendjét befolyásolták. Az utasítás, a végrehajtás és az energia ilyen különválasztása segít elkerülni azt a tévedést, hogy minden programozható szerkezetet eleve teljesen önálló gépnek képzeljünk.

Nyolc vezérlési pont egy szemléltető példában

A működés alapgondolatát kicsiben is követhetjük. Képzeljünk el nyolc vezérlési pontot és nyolc egymást követő műveleti lépést. Minden lépéshez megadjuk, mely pontok legyenek kiválasztva. A leírás így nyolc sorból állhat, mindegyikben nyolc döntéssel. A példában a lyuk a kiválasztást jelenti; egy adott történeti szerkezet tényleges mechanikai kapcsolatát külön kell megvizsgálni.

Ha a negyedik lépésben az 1., 3., 6. és 8. pont kiválasztását akarjuk, az ennek megfelelő négy helyet lyukasztjuk ki a negyedik kártyán. A gép olvasóelemei nem a számok nevét vagy a készülő díszítést értik. A lyuk és a tömör karton eltérő mechanikai feltételt teremt. Az ember által megtervezett különbség így mozgásbeli különbséggé alakul.



16. kép Nyolc lépés és nyolc vezérlési pont szemléltető terve. Jobbra a negyedik kártya, amely az 1., 3., 6. és 8. pontot választja ki. A rajz vezérlési elvet mutat; nem történeti gép méretarányos terve és nem kész kelmeminta. Saját szerkesztés.

A rajz egy vezérlési tervet mutat, nem egy valódi kelme fényképének helyettesítője. A sötét négyzetekből önmagukban nem állapítható meg a kész textil színe vagy teljes felületi képe. Ehhez ismerni kellene a befűzést, a kötést és a használt fonalakat is. Ez az apró példa a jel és a jelentés viszonyát is megvilágítja: a lyukak hatását az olvasó és a hozzá kapcsolt gép szerkezete határozza meg.

Egyetlen kártya megváltoztatásával egy lépés kiválasztása módosul. A sorrend felcserélésével más eseménysor jön létre, a kártyák egy részének ismétlésével pedig ismétlődés. Az alapgép közben változatlan maradhat. Az utasítások szerkesztése így külön munkafolyamattá válik a végrehajtó szerkezet megépítésétől.

A kártyasorban megőrzött munka

A lyukkártyát könnyű a mai program egyszerű elődjének nevezni. A hasonlat lényegét azonban pontosan kell megfogni. A különböző mintákhoz nem kell minden alkalommal új gépet építeni, mert a cserélhető hordozó őrzi a kiválasztások rendjét. Ez már fontos lépés a gép és a neki adott feladat részleges szétválasztásában.

A hordozóra vitt leírás ráadásul megőrizhető. Egy sikeres mintát később újra elő lehet venni, a kártyasort ellenőrizni vagy másolni lehet. A tudás nem kizárólag egyetlen munkás emlékezetében él tovább. A műhely így nemcsak szövőszékeket és fonalakat, hanem gépileg végrehajtható mintákat is gyűjthet.

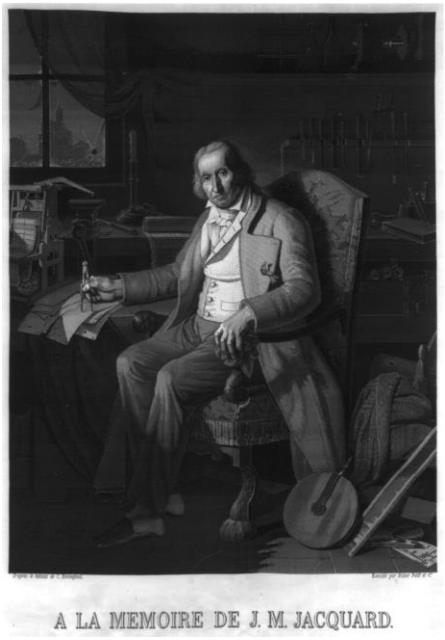
A munkát mindez átrendezi. A minta megtervezése, a kártyák elkészítése, a gép beállítása és az üzemeltetés külön feladattá válhat. A gépesítés ezért nem egyszerűen azt jelenti, hogy „a munkás helyett dolgozik a gép”. Egyes műveleteket kivált, másokat a végrehajtás előtti szakaszba helyez át. A kártyák lyukasztásában korábban elvégzett emberi munka irányítja a későbbi mozgást.

Ebből gazdasági és társadalmi feszültség is következhet. Ami a műhely számára megismételhetőbb termelést és hozzáférhetőbb mintákat jelent, az egy korábbi feladatra szakosodott munkás számára a

tudása értékének megváltozása lehet. A termelékenység javulása önmagában nem mondja meg, ki részesül az előnyből, és kinek kell alkalmazkodnia. A gép műszaki leírása és bevezetésének társadalmi története ezért összetartozik, de nem helyettesíti egymást.

Egy portré amelyet nem festettek

A Jacquard-rendszer képességeit különösen szemléletesen mutatják a feltalálóról készített selyemportrék. Claude Bonnefond festménye nyomán Michel-Marie Carquillat szőtt ilyen képet 1839-ben. A finom tónusokat és tárgyrészleteket nem ecset vitte fel a felületre: a szövet szerkezete alakította ki őket. A Computer History Museum a híres portréváltozathoz mintegy huszonnégyezer kártyát kapcsol. Ez egy különösen részletes munka nagyságrendje, nem minden Jacquard-minta általános kártyaszáma. [34, 37]



17. kép Jacquard 1839-es selyemportréja: Michel-Marie Carquillat munkája Claude Bonnefond festménye nyomán, Didier Petit et Cie kivitelezésében. A tónusokat a szövés hozta létre. Library of Congress, 2002737214 / Wikimedia Commons. Közkincs.

A portré azért több látványos kuriózumnál, mert megmutatja, milyen messzire vezethet sok egyszerű választás megfelelő összehangolása. Egyetlen vezérlési pont alig mond valamit egy arcról vagy egy műhely belsejéről. A gondosan előkészített műveletsor mégis olyan képet hozhat létre, amelyet a néző emberi alaknak és térnek ismer fel.

A kép megtervezése és a kép megszövése itt világosan elkülönül. A gép nem a feltaláló arcát ismeri fel, és nem a portré hasonlóságát ellenőrzi. A már előkészített rendet valósítja meg. Ettől az eredmény nem kevésbé képszerű; a kép létrehozásának módját azonban másképpen kell leírnunk, mint az emberi látást vagy rajzolást.

Mit jelent itt a programozhatóság

A Jacquard-berendezés programozhatósága egy meghatározott feladatcsaládra vonatkozik. Különböző minták vezérlését teszi lehetővé, a szövőszék adottságain belül. Nem általános célú

számítógép, és a kártyalánc ismétlődése önmagában nem azonos egy számítási eredménytől függő elágazással. A történeti teljesítményt azzal értjük meg, ha ezt a határt megtartjuk.

A lyuk és a lyuk hiánya mai szemmel kétállapotú jelnek tekinthető. Ebből még nem következik, hogy a készítő a későbbi digitális számítógép teljes elméletét alkalmazta. A mintavezérléshez elegendő volt a fizikai különbség és a hozzá rendelt mechanikai hatás. Egy rendszer tartalmazhat később általánossá váló megoldást anélkül, hogy alkotója annak valamennyi későbbi következményét ismerné.

A program és az adat közötti különbség sem a karton anyagába van beírva. Ugyanilyen hordozóra kerülhet a végrehajtandó művelet előírása, de egy személy életkorának vagy foglalkozásának kódja is. A szerepet az határozza meg, mire használja a jeleket a rendszer. A későbbi lyukkártyás adatfeldolgozás ezért rokona a mintavezérlésnek, mégis más feladatot old meg. Hollerith 1890-es amerikai népszámlálási alkalmazásában a kártyák emberekről felvett adatokat hordoztak. [36]

A virágmintától az algebrai műveletig

Charles Babbage a lyukkártyás vezérlésben a számítások szervezésére is alkalmas elvet ismert fel. Az analitikai gép tervében a műveletek és a feldolgozandó mennyiségek gépi kezelését sokkal általánosabb feladatra kívánta alkalmazni. A történet részletes folytatása a számológépekhez vezet, de a kapcsolatot már itt érdemes rögzíteni: a szövőszék példája segített elképzelni, hogyan adható át egy műveleti rend a gépnek. [35]

Ada Lovelace 1843-as jegyzeteiben a párhuzamot emlékezetes képpel fogalmazta meg: az analitikai gép algebrai mintázatokat sző, ahogyan a Jacquard-szövőszék virágokat és leveleket. A hasonlatban a lényeg a viszonyok rendezett átalakítása. Az egyik gép fonalak kereszteződését, a másik számok és műveletek kapcsolatait kezeli. A mondat itt magyar parafrázisban szerepel. [35]

Ez a fordulat különbözik attól, amit a Török sugallt. Ott a néző előtt jelent meg egy gondolkodónak látszó alak, miközben a döntés embernél maradt. A lyukkártyás berendezésnek nincs szüksége emberi arcra ahhoz, hogy egy pontosan körülhatárolt műveletsort valóban gépileg hajtson végre. A látszólag kevésbé emberi szerkezet ezért fontosabb lépést tehetett a programozható gép felé.

A fejezet címében szereplő leírás nem egyszerű magyarázó szöveget jelent. Olyan rendezett jelrendszerről van szó, amelynek különbségeit a gép működési különbséggé alakítja. Az ember megtervez valamit, a hordozó megőrzi, a berendezés pedig végrehajtja. A következő nagy kérdés az lesz, milyen műveletek írhatók le ilyen módon, és meddig tágítható a gép által követhető utasítások világa.

Források és további olvasmányok

A hivatkozások és a képek számozása az első három fejezetét folytatja. A rekonstruált szerkezeteket és a korabeli működésmagyarázatokat a történeti tárgyaktól megkülönböztetve használjuk.

[24] Heinz Nixdorf MuseumsForum: Der Schachtürke des Wolfgang von Kempelen; 250 Jahre Schachtürke. A rejtett játékos, a mechanikai közvetítés és a bemutatás története. Az első bemutató keltezéséről a múzeum utólagos helyesbítést is közöl. Carola Dahlke: Schach, die Mensch-Maschine und die Anfänge der KI. Deutsches Museum, 2024.

[25] Edgar Allan Poe: Maelzel’s Chess-Player. Southern Literary Messenger, 1836. április, 2:318–326. A korabeli megfigyelések és a gépi sakkozásról alkotott következtetések elsődleges forrása. Az 1836-os szöveg

[26] Joseph Friedrich von Racknitz: Ueber den Schachspieler des Herrn von Kempelen und dessen Nachbildung. Leipzig–Dresden, 1789. A bemutatott táblák magyarázati kísérletek, nem a feltalálótól származó műszaki dokumentáció. Digitalizált képtábla és forrásadatok

[27] Wolfgang von Kempelen: Mechanismus der menschlichen Sprache nebst der Beschreibung seiner sprechenden Maschine. Wien, Degen, 1791. Különösen a beszédkeltő berendezésről szóló ötödik rész és a rézmetszetek. A korszerű átírás és angol fordítás szerkesztői Fabian Brackhane, Richard Sproat és Jürgen Trouvain, 2017. Kiadási és hozzáférési adatok · Az 1791-es kötet bibliográfiája

[28] Deutsches Museum: Kempelen’s speaking machine. Történeti berendezés, 1800 körül, leltári szám 6596; valamint a múzeum és a Leibniz-Gemeinschaft Der Kempelen’sche Sprechapparat című bemutatója. A kézi kezelés és a hangképző részek szerepe. Gyűjteményi ismertető · A szerkezet részletes bemutatója

[29] Nikléczy Péter – Olaszy Gábor: A reconstruction of Farkas Kempelen’s speaking machine. Eurospeech 2003, 2453–2456. DOI: 10.21437/Eurospeech.2003-674. A 2001-ben készült működő rekonstrukció és vizsgálatának elsődleges kutatási beszámolója. Tanulmány és teljes szöveg

[30] Science Museum Group: Model of Falcon’s loom. Gyűjteményi azonosító 1963-379. Bouchon 1725-ös papírszalagja, Falcon 1728-as kártyalánca és az emberi közreműködés szerepe. Műtárgykatalógus

[31] Musée des Arts et Métiers: Métier à tisser les étoffes façonnées de Vaucanson destiné à remplacer l’ancien métier à la tire. 1746, leltári szám 00017-0001. A szerkezet és az ipari alkalmazás elmaradásának ismertetése. Műtárgykatalógus

[32] Science Museum Group: Jacquard Hand Loom. Y1973.43, W. Archer, 1910 körül. A későbbi gyűjteményi tárgy a lyukkártyás kiválasztás és a kézi működtetés alapelvét szemlélteti. Szerkezeti ismertető

[33] National Museums Scotland: The Jacquard loom Innovation in textiles and computing. T.1934.241. A mintavezérlés, az 1801-es bemutató és a későbbi skót szövőszékek története. Gyűjteményi ismertető

[34] Chris Garcia: Artifact Adopted Preserving the Legend of J. M. Jacquard. Computer History Museum, 2018. december 21. A szövött portré és a kártyás mintavezérlés kapcsolata. A felhasznált kép külön példányát a Library of Congress őrzi. Kurátori ismertető · A képen látható műtárgy adatlapja

[35] Luigi Federico Menabrea: Sketch of the Analytical Engine Invented by Charles Babbage. Ada Lovelace fordításával és jegyzeteivel, 1843. Különösen az A jegyzet a Jacquard-párhuzamról és a műveleti kártyákról. A fejezetben szereplő magyar megfogalmazás parafrázis. Elsődleges szöveg

[36] Smithsonian National Museum of American History: Punch Cards for Data Processing. A Hollerith-kártyák és az 1890-es népszámlálási adatfeldolgozás. Gyűjteményi áttekintés

[37] Computer History Museum: 1801 Punched cards control Jacquard loom. A mintakiválasztás alapelve és az 1839-es selyemportré huszonnégyezer kártyás vezérlése. Technikatörténeti áttekintés

Képforrások és felhasználási adatok

11. kép Eredeti képfájl és forrásadatok

12. kép Eredeti képfájl és forrásadatok

13. kép Eredeti képfájl és forrásadatok

14. kép Eredeti képfájl és forrásadatok

15. kép Eredeti képfájl és forrásadatok · Stephencdickson · CC BY-SA 4.0

16. kép Saját szerkesztésű szemléltető ábra a fejezetben megadott nyolclépéses kiválasztási tervhez.

17. kép Eredeti képfájl és forrásadatok

II. rész – A gondolkodás kiszámíthatóvá válik?

Az első rész gépei mozgást, beszédet vagy emberi ügyességet utánoztak. A következő fordulat kevésbé látványos: a figyelem a gondolkodás szabályaira irányul. Ha egy következtetés helyessége pontosan megadható, vajon a végrehajtása rábízható-e valamilyen szerkezetre? Ehhez két, sokáig külön fejlődő hagyománynak kell közel kerülnie egymáshoz. Az egyik a fogalmak és állítások logikai vizsgálata, a másik a számolás gépesítése. Találkozásukból még nem születik mesterséges intelligencia, de fokozatosan kirajzolódik a jelképekkel dolgozó gép lehetősége.

6. A gondolkodás mint szabálykövetés

Egy állítás meggyőző lehet azért, mert tekintélyes ember mondja, mert sokan ismétlik, vagy mert illik ahhoz, amit szeretnénk igaznak hinni. A logika más kérdést tesz fel: valóban következik-e abból, amit már elfogadtunk? Ez a különbség az ókori vitákban éppúgy számított, mint ma egy tudományos érvelés, egy hiteldöntés vagy egy mesterséges intelligencia válaszának ellenőrzésekor.

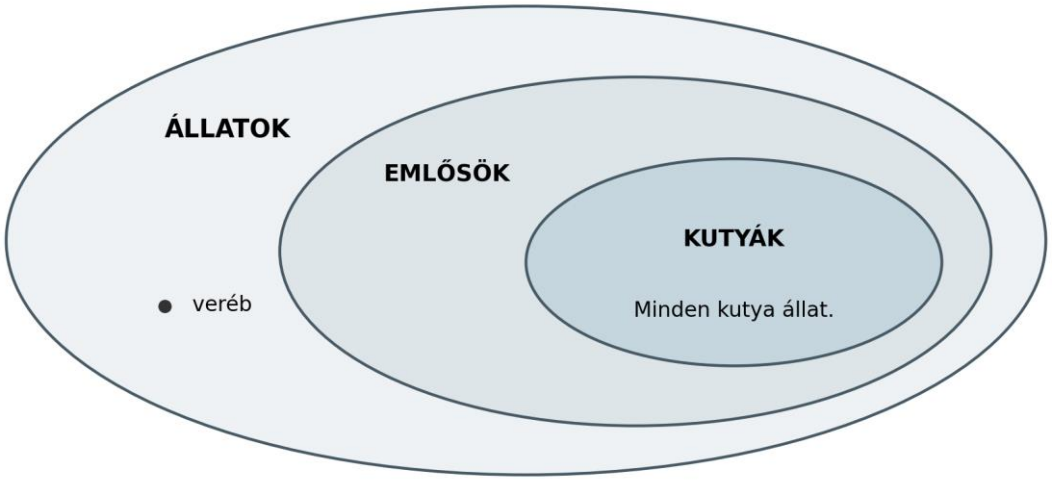
A gondolkodó gép történetéhez ezért olyan felismerés is hozzátartozik, amelyhez nem kellett fogaskerék. Bizonyos következtetések helyessége elkülöníthető attól, hogy éppen emberekről, állatokról vagy geometriai alakzatokról beszélünk. Ha sikerül megragadni ezt a közös szerkezetet, akkor a következtetés egy része szabályokkal kezelhetővé válik.

Arisztotelész: mitől szükségszerű egy következtetés?

Arisztotelész a Kr. e. negyedik században rendszeresen vizsgálta a következtetések alakját. Az Első analitika központi fogalma a szillogizmus: olyan érvelés, amelyben bizonyos állítások elfogadása után valami más szükségképpen következik belőlük. Nem minden emberi gondolkodást akart egyetlen egyszerű sémába szorítani; a bizonyító következtetés szabályszerűségeit kutatta. [38]

Vegyünk egy mai, könnyen áttekinthető példát. Minden emlős állat. Minden kutya emlős. Ebből következik: minden kutya állat. A két kiinduló állítás összekapcsolásában az „emlős” közvetítő szerepet játszik. Ha a kutyák az emlősök közé tartoznak, az emlősök pedig az állatok közé, akkor a kutyák sem lehetnek az állatok körén kívül.

A következtetés alakja akkor is megmarad, ha lecseréljük a fogalmakat. Minden A egyben B; minden C egyben A; tehát minden C egyben B. Az A, B és C itt üres helyek: a tartalom változhat, a kapcsolat ugyanaz marad. Ez a tartalomtól való részleges elvonatkoztatás lesz később a gépesíthetőség egyik feltétele.



18. kép A kutyák, az emlősök és az állatok egymásba foglalt körei. A veréb az állatok közé tartozik, de nem kutya. Mai szemléltető halmazábra; nem Arisztotelész saját rajza. Saját szerkesztés.

A szabály megfordításával azonban könnyen hibázunk. Abból, hogy minden kutya állat, nem következik, hogy minden állat kutya. A veréb is állat. A téves következtetés gyakran éppen azért csúszik át az ellenőrzésen, mert a benne szereplő szavakat ismerjük, és az érvelés nyelvileg természetesnek hat. A forma vizsgálata ott is hibát tárhat fel, ahol az előadás meggyőző.

Nem minden arisztotelészi következtetés beszél „mindenről”. A „némely”, az „egyik sem” és a tagadás szintén fontos szerepet kap. A különféle állításalakok nem tetszőlegesen kapcsolhatók össze. Éppen ezeknek a megengedett és tiltott kapcsolásoknak a rendszerezése emelte a következtetés vizsgálatát a jó vitakészségen túlmutató tudománnyá. [38]

A helyes forma és az igaz tartalom

A logika erejét akkor értjük meg, ha a határát is látjuk. Tegyük fel, hogy valaki így érvel: minden fényes tárgy aranyból van; ez a kanál fényes; tehát ez a kanál aranyból van. A következtetés formája rendben lehet, az első állítás azonban hamis. Az érvényes következtetés azt biztosítja, hogy igaz kiinduló állításokból ne jussunk hamis következményhez. A kiinduló állítások igazságát önmagában nem biztosítja.

Egy vegyészmérnök számára ez ismerős helyzet. Ha egy számítás hibátlan, de rossz koncentrációadatból indul, az eredménye nem írja le megfelelően a tartály tartalmát. A formális következtetésben a kiinduló állítások töltik be az adatok szerepét. Ellenőrzésükhöz mérésre, megfigyelésre, történeti forrásra vagy másfajta bizonyítékra lehet szükség.

Ugyanez a különbség társadalmi kérdésekben is számít. Egy szabályrendszer következetesen alkalmazhat egy rosszul megválasztott besorolást. Attól, hogy az összes lépés visszakövethető, még meg kell vizsgálni, helyesen írtuk-e le az embereket és a helyzetüket. A következetesség érték, de nem helyettesíti a valósággal való összevetést.

Nem minden gondolkodás dedukció

A biztos következmények levezetését dedukciónak nevezzük. A mindennapokban azonban sokszor másképpen haladunk. Sok fehér hattyú megfigyeléséből arra következtethetünk, hogy talán minden hattyú fehér. Ez indukció: tapasztalatokból általánosítunk. Az új esetek száma növelheti a bizalmunkat, de egyetlen fekete hattyú megcáfolja az általános állítást.

Más helyzetben egy jelenség magyarázatát keressük. Reggel vizes az utca, ezért arra gondolunk, hogy esett az eső. Lehet, hogy valóban ez történt, de az utcát fel is moshatták, vagy csőtörés lehetett. Ilyenkor lehetséges magyarázatot választunk; ezt gyakran abdukciónak nevezik. Az eredmény további vizsgálatra szorul.

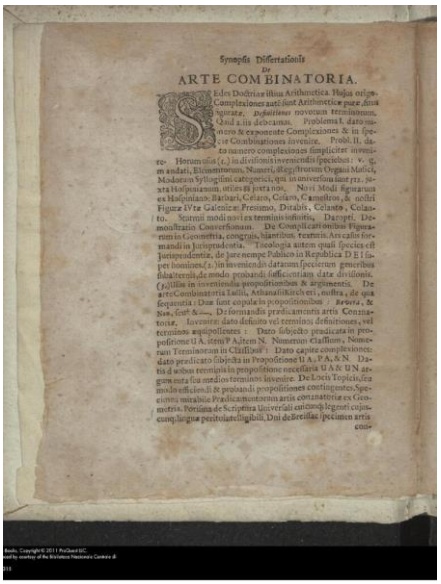
E három művelet összekeverése az intelligens rendszerekről alkotott képünket is eltorzíthatja. Egy valószínű magyarázat nem matematikai bizonyítás. A sok hasonló eset nem azonos egy minden esetre érvényes tétellel. Egy gép teljesítményét is annak megfelelően kell megítélni, hogy biztos következményt vezet le, tapasztalatból becsül, vagy magyarázatot javasol.

Az arisztotelészi hagyomány fontos indulópont, de a modern logika nem egyetlen ókori tanítás változatlan továbbélése. Más logikai hagyományok és későbbi átalakítások is kellettek ahhoz, hogy összetett állításokat, kapcsolatokat és matematikai bizonyításokat pontosan lehessen kezelni. A történet fő kérdése mindeközben megmaradt: mi tehető láthatóvá és ellenőrizhetővé a gondolatmenetből?

Leibniz nagy terve: a gondolatok jelrendszere

Gottfried Wilhelm Leibniz a tizenhetedik században jóval messzebbre tekintett. Az 1666-ban megjelent De arte combinatoria már összekapcsolta a fogalmak felbontását és kombinálását a rendezett felfedezés lehetőségével. Későbbi terveiben egy általános jelrendszer és egy hozzá tartozó következtetési eljárás körvonalazódott. Az előbbit rendszerint characteristica universalis, az utóbbit calculus ratiocinator néven említjük. [39–40]

A tervet nem úgy érdemes elképzelni, mint egy minden nyelvre lefordított szótárt. Leibniz olyan jeleket keresett, amelyek a fogalmak szerkezetét és kapcsolatait is kezelhetővé teszik. Ha a bonyolult fogalmak egyszerűbb összetevőkre bonthatók, a köztük levő viszonyok pedig szabályosan leírhatók, akkor a következtetés egy része jelekkel végzett műveletté válhat.



Synopsis Dissertationis
DE
ARTE COMBINATORIA.

19. kép Leibniz De arte combinatoria című, 1666-os értekezésének áttekintő lapja. Az ifjú szerző a kombinációk vizsgálatát a fogalmak rendezésével és a felfedezés módszerével is összekapcsolta. Biblioteca Nazionale Centrale di Firenze / Internet Archive / Wikimedia Commons. Közkincs. A digitalizált lap körülvágott részlete.

Az algebra már megmutatta, mennyit nyerhetünk megfelelő jelöléssel. Nem szükséges minden egyenletet hosszú mondatokban megfogalmazni: a betűk és műveleti jelek a kapcsolatokra irányítják a figyelmet. Leibniz azt remélte, hogy a rendezett jelhasználat másfajta érvelésekben is hasonló szerepet tölthet be. A viták egy része így számolásszerű ellenőrzéssé alakulhatna. [40]

Ebben a tervben a „számolás” tágabb jelentést kap az összeadásnál. Szabályozott jelátalakítást jelent: megadott kifejezésekből megengedett lépésekkel jutunk új kifejezésekhez. A művelet lehet logikai következtetés is, nem csupán mennyiségek meghatározása. Ez a gondolat köti majd össze a logika történetét a számítógépével.

A nehézség a leírásnál kezdődik

Leibniz nem valósította meg az emberi tudás egészére kiterjedő egyetemes jelrendszert. A nehézség nem pusztán egy megfelelő gép hiánya volt. Előbb abban kellene megállapodni, melyek az alapfogalmak, mit jelentenek, hogyan függnek össze, és milyen tényekből indulhatunk ki. [40]

A hétköznapi szavak ráadásul környezetfüggők. A „biztonságos”, a „nagy”, a „jó” vagy a „kockázatos” nem mindig ugyanazt jelenti. Egy hitel megítélésénél külön kell választani a fizetőképességre vonatkozó adatokat, az előrejelzési bizonytalanságot és az intézmény által vállalt kockázatot. Nem ugyanaz a kérdés, hogy egy szabályból mi következik, és hogy érdemes-e ezt a szabályt alkalmazni.

A formalizálás ezért alkotómunka. A valóság összetett helyzetéből olyan leírást készítünk, amelyben bizonyos részleteket megőrzünk, másokat elhagyunk. A következtetés pontossága nem teszi fölöslegessé annak vizsgálatát, mit veszítettünk el a leírás során.

Mégis hatalmas változás történt. A gondolkodás bizonyos mozzanatai olyan eljárásként kezdtek megjelenni, amelyet meg lehet tanítani, ismételni és ellenőrizni. Ha pedig a végrehajtáshoz elegendő felismerni a jeleket és követni a szabályokat, felmerül a következő kérdés: szükség van-e minden

egyes lépésnél emberi végrehajtóra? A számológépek története ennek a kérdésnek adott kézzelfogható formát.

7. Számológéptől az általános célú gépig

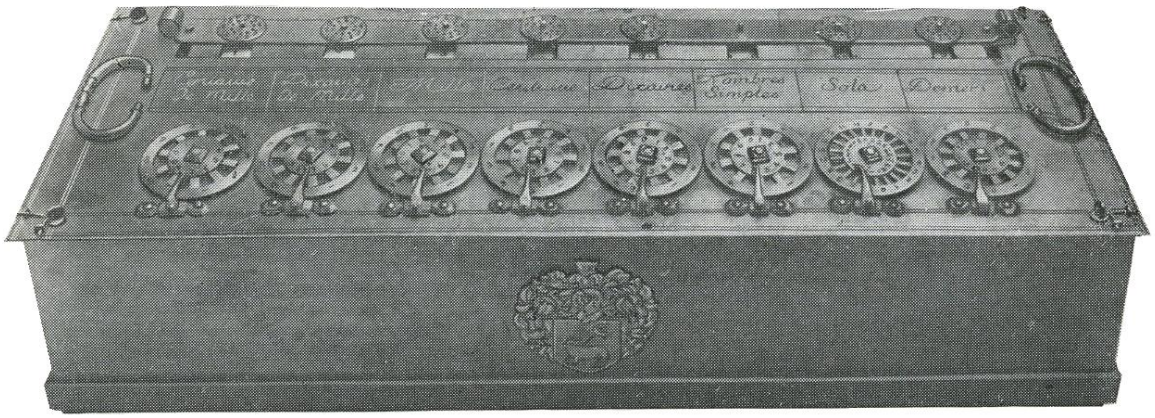
Az összeadásban könnyű hibázni. Egy kihagyott tétel, egy elrontott átvitel vagy egy rossz helyi érték elegendő ahhoz, hogy a végeredmény téves legyen. A számolás gépesítésének egyik hajtóereje ezért igen gyakorlati volt: hogyan lehet az ismétlődő műveleteket megbízhatóbban elvégezni?

Az előző rész automatáinál a gép egy előre kialakított mozgássort hajtott végre. A számológépeknél már az a kérdés is megjelenik, hogy a szerkezet állapota miként hordozhat elvont értéket. Egy kerék helyzete lehet a négyes szám jele, a következő keréké a tízesek számát adhatja meg. A fizikai mozgás és az értelmezett mennyiség között tudatosan kialakított megfelelés jön létre.

Pascal: az átvitel gépesítése

Blaise Pascal 1642-ben kezdett számológépének fejlesztésébe, hogy megkönnyítse apja adóügyi számításait. A Pascaline néven ismert szerkezet helyi értékeket képviselő kerekekkel dolgozott, és alkalmas volt összeadásra, illetve megfelelő eljárással kivonásra. A gép történeti jelentőségét nem szükséges azzal felnagyítani, hogy minden korábbi számolóeszközt elfelejtünk: a konkrét műszaki feladat megoldása önmagában is figyelemre méltó. [41]

A legkényesebb kérdés az átvitel. Ha az egyeseknél a kilenc után még egyet hozzáadunk, az eredmény ott nulla lesz, a tízeseknél pedig eggyel növekednie kell az értéknek. Ha több helyi értéken egyszerre történik ilyen váltás, az átviteleknek sorban tovább kell terjedniük. Papíron ez megszokott művelet; mechanikus alkatrészeknél pontos időzítést és megfelelő erőátadást kíván.



20. kép A Pascaline egyik változata J. A. V. Turck Origin of Modern Calculating Machines című, 1921-es könyvének illusztrációján. A tárcsák a bevitelhez, a fölöttük látható ablakok az eredmény leolvasásához tartoztak. Wikimedia Commons. Közkincs.

A gép használója beállította a bevitt mennyiséget, a szerkezet pedig közvetítette annak hatását az eredménykijelzésre. Ettől a gép még nem választotta ki önállóan a megoldandó feladatot. Nem tudta, mit jelent az adó, kitől származik egy összeg, vagy helyesen vették-e fel a tételeket. A számolás jól körülhatárolt részét végezte el.

A fennmaradt Pascal-gépek között pénzügyi elszámolásra készült változatok is vannak. Ezeknél nem minden kerék fordult tíz lépés után: a korabeli pénzrendszer húszas és tizenkettes váltásokat is

megkívánt. A szerkezet tehát nem kizárólag a mai, minden helyi értéken tízes számolást tükrözhette. A gép felépítését a felhasználási környezethez igazították. [41]

Leibniz: több művelet ugyanazzal az elvvel

Leibniz számára a számológép megépítése és a következtetés szabályainak keresése ugyanannak a tágabb törekvésnek két oldala volt. Az ismétlődő szellemi munkából minél többet szeretett volna megbízható eljárásokra bízni. Az 1673-ban Londonban bemutatott korai modell és a későbbi, továbbfejlesztett gépek azonban egy hosszú fejlesztési folyamat különböző állomásai voltak. [42]

A Leibnizhez kötődő lépcsős henger különböző beállításoknál eltérő számú fogat kapcsolhatott a továbbító szerkezethez. Az összeadásra épülő művelet ismétlésével és a helyi értékek eltolásával szorzási feladatokat is lehetett kezelni; a megfelelő fordított eljárás az osztáshoz vezetett. A bonyolultabb művelet egyszerűbb műveletek szervezett sorozatára bomlott.

Itt már világosan látható az algoritmikus gondolat: nem kell minden szorzáshoz külön gépet építeni. Ugyanazt az alapműveletet ismételjük más beállítások mellett. A feladat mérete és a számjegyek rendje határozza meg, hányszor és melyik helyi értéken kell működésbe hozni a szerkezetet.

A megvalósítás korántsem volt problémamentes. A hosszú átviteli láncok megbízhatósága, az alkatrészek pontossága és a teljes gép együttműködése jelentős nehézséget okozott. A korai bemutató nem azonos egy kész, minden helyzetben üzembiztos termékkel. A fennmaradt fiatalabb gépet az 1690-es évek fejlesztéseihez kapcsolják. [42]

Ez a különbség a későbbi számítástechnika történetében is visszatér. Egy működési elv bizonyítása, egy bemutató és a tartósan használható rendszer három különböző eredmény. Mindhárom fontos, de eltérő bizonyítékot szolgáltat arra, mire képes a technika.

Babbage: hibás táblázatok helyett gépi eljárás

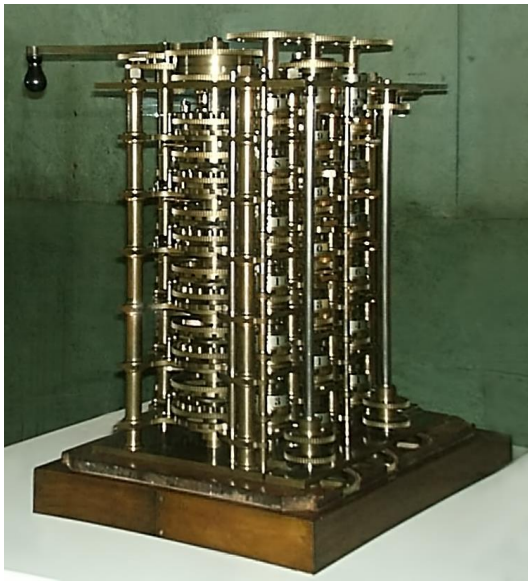
A tizenkilencedik században a hajózás, a csillagászat és a mérnöki munka nagy számtáblázatokra támaszkodott. A számítás, az eredmények másolása és a nyomdai előállítás mind hibaforrás lehetett. Charles Babbage ezért nemcsak egy-egy összeadást akart gépesíteni, hanem a táblázatkészítés hosszabb folyamatát is. [43]

Az 1822-ben meghirdetett differenciagép matematikai alapja a véges különbségek módszere volt. Bizonyos táblázatok értékei ismételt összeadásokkal előállíthatók. Egyszerű példánkban a négyzetszámok egymást követő különbségei rendre 1, 3, 5 és 7; e különbségek változása állandóan 2. A folytatáshoz elegendő ezt a szerkezetet megőrizni.

x	x^2	Első különbség	Második különbség
0	0	—	—
1	1	1	—
2	4	3	2
3	9	5	2
4	16	7	2

A négyzetszámok egymást követő különbségei. A második különbség állandó. Saját számítás.

Ha az utolsó első különbség 7, a következő 9 lesz. Ha az utolsó négyzetszám 16, akkor a következő 16 + 9, vagyis 25. A gépnek a táblázat folytatásához nem kell minden alkalommal önálló négyzetre emelést végrehajtania. A megfelelő kezdőértékekből, az összeadások rendjét betartva állítja elő a következő sort. Ez tanítási példa; a történeti gép ennél több számjeggyel és több különbségi szinttel tervezett számolni.



21. kép Babbage első differenciagépének 1832-ben elkészült működő része a londoni Science Museum gyűjteményében. A képen nem a teljes tervezett gép és nem az analitikus gép látható. Fotó: Sebastian Wallroth / Wikimedia Commons. A fényképet a szerző közkincsnek nyilvánította.

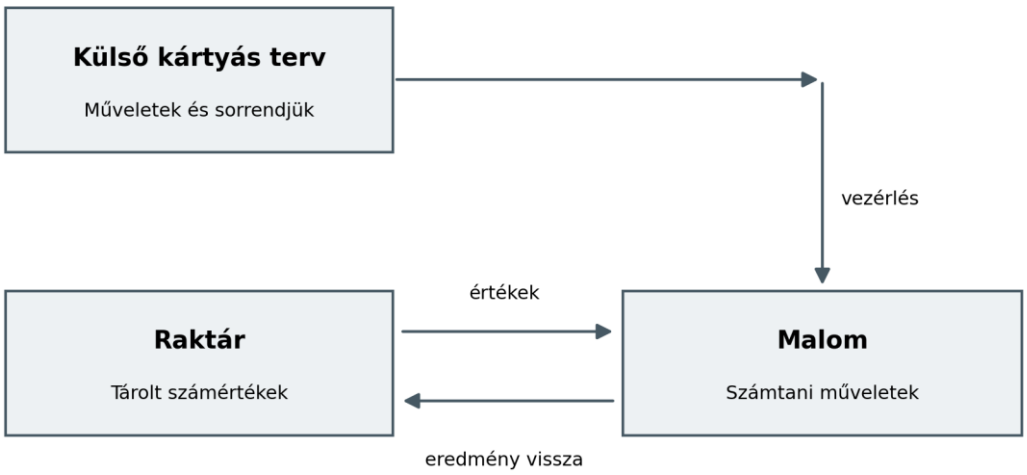
Az első differenciagépből 1832-re működő rész készült el, a teljes tervezett berendezés azonban nem. A munka költségei, a kivitelezés és az együttműködés nehézségei végül megakasztották a vállalkozást. Nem volna pontos mindezt egyetlen mondattal a kor „elégtelen technikájára” visszavezetni. [43]

A differenciagép különleges célú szerkezet volt: egy meghatározott matematikai módszer gépesítésére készült. Babbage következő terve ennél lényegesen nagyobb elvi lépést tett. Miért kellene a gép felépítésének egyetlen módszert rögzítenie, ha a műveletek sorrendjét külön is meg lehet adni?

Az analitikus gép: műveletek, adatok és vezérlés

Az 1834-től körvonalazódó analitikus gépben Babbage elkülönítette a számokat tároló részt és a műveleteket végző egységet. Az előbbit „raktárnak”, az utóbbit „malomnak” nevezte. A Jacquard-szövőszékeknél megismert lyukkártyák azt a lehetőséget kínálták, hogy a gép munkájának rendjét kívülről megadott sorozat irányítsa. [43–44]

Az általános célúság itt nem korlátlan képességet jelentett. A gépnek meghatározott alapműveletei, véges tárolókapacitása és fizikai korlátai lettek volna. A lényeg az volt, hogy eltérő feladatokhoz ne kelljen mindig új alapgépet építeni: a műveleti tervet és a kiinduló adatokat lehessen megváltoztatni.



22. kép Az analitikus gép néhány alapvető szerepkörének egyszerűsített kapcsolata. A külső kártyás vezérlés a tárolt számértékeken végzett műveletek rendjét adja meg. Nem teljes szerkezeti terv, és nem mai tárolt programú számítógép rajza. Saját szerkesztés Babbage és Lovelace leírásai alapján.

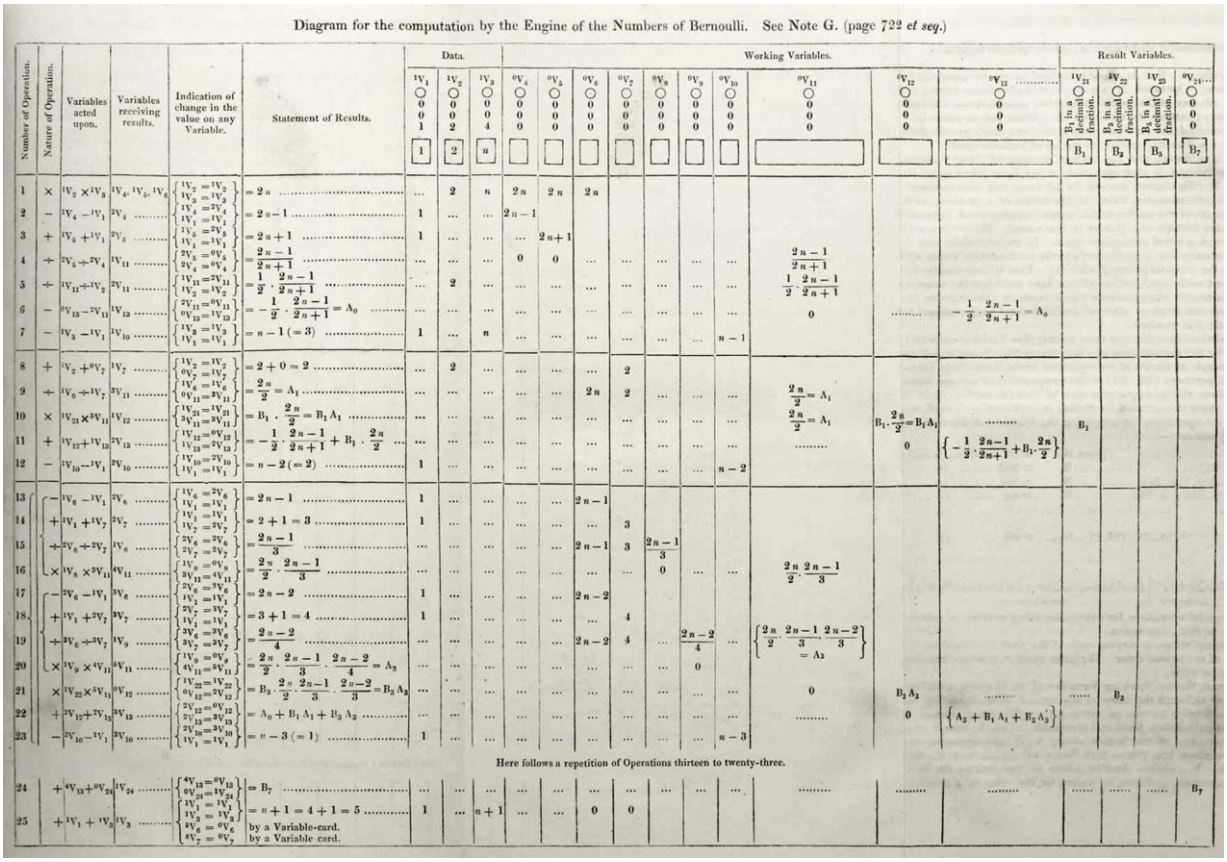
A tervek ismétléseket és feltételtől függő vezérlést is megengedtek. Az ismétlés azt jelenti, hogy ugyanahhoz a műveletsorhoz visszatérünk. A feltételes vezérlésnél egy eredmény vagy állapot befolyásolja, melyik lépés következik. Ez már túlmutat azon, hogy a szerkezet egyszer végigjátsszon egy mereven rögzített mozdulatsort. [43]

A hasonlóságok ellenére az analitikus gépet nem szabad a mai számítógép egyszerűen felnagyított fogaskerekes változatának tekinteni. Tízes számrendszerű mechanikus gép volt a tervekben, külső kártyás vezérléssel. A mai értelemben vett tárolt programú felépítést nem valósította meg, és Babbage életében a teljes gép sem készült el. A történeti jelentőség itt elsősorban a kidolgozott elképzelésé. [43–44]

Ada Lovelace: a számok mögött jelképek állhatnak

Luigi Federico Menabrea francia nyelvű ismertetést írt Babbage tervéről. Ada Lovelace ennek angol fordítását 1843-ban saját, terjedelmes jegyzeteivel egészítette ki. A jegyzetek a gép működésének megértésében és tágabb lehetőségeinek felismerésében is fontos dokumentumok. [44]

A G jegyzet Bernoulli-számok kiszámításához ad részletes műveleti táblázatot. A táblázat nem csupán egy képletet közöl: a műveletek rendjét és a tárolt értékek felhasználását is követhetővé teszi. Megmutatja, milyen közvetítő leírásra van szükség a matematikai feladat és a gép tényleges működése között.



23. kép A Bernoulli-számok kiszámításának műveleti táblázata Lovelace 1843-as G jegyzetéből. A történeti dokumentum a műveleteket és a tárolt értékek változását együtt követi; nem egy akkor már elkészült gép futási naplója. Wikimedia Commons. Közkincs.

Lovelace-et gyakran az első programozónak nevezik. A megjelölés közismert, de történetileg magyarázatra szorul: Babbage korábban is készített műveleti terveket, a jegyzetek pedig szoros együttműködésben születtek. Biztosabban állítható, hogy Lovelace neve az általános célú géphez készült, részletesen kidolgozott és korán publikált eljárás egyik legfontosabb dokumentumához kötődik. A táblázatot az eredeti analitikus gép nem futtathatta le, mert az nem épült meg. [44–45]

Még messzebbre mutat az A jegyzet gondolata. Ha valaminek a kapcsolatai megfelelő jelekkel kifejezhetők, és a gép műveletei ezekhez a kapcsolatokhoz rendelhetők, akkor a gép nemcsak mennyiségek számítására használható. Lovelace a zene lehetőségét is felvetette, feltételesen: ehhez a zenei viszonyokat alkalmas módon le kellene írni. Nem egy elkészült zeneszerző gépről számolt be. [44]

A felismerés lényegét mai példával érthetjük meg. Megegyezhetünk abban, hogy az 1 az A, a 2 a B, a 3 a C betű jele. Ekkor az 1, 2, 3 számsorozat az ABC betűsort képviselheti. Ha a gép megfordítja a sorrendet, a kimenet 3, 2, 1 lesz, amelyet CBA-ként olvasunk vissza. Az eljárás számokkal dolgozik, de a feladat betűk rendezése.

Önmagában a kódolás még nem elég. Az 1 + 2 = 3 számtani tényből nem következik, hogy A és B értelmes „összege” C. A jelképekhez illő műveletet is meg kell határozni. Más szabály kell a betűk rendezéséhez, egy dallam hangközeinek átalakításához és egy logikai következtetéshez.

A működtető tudása és a gép művelete

A gép ugyanazt a számsort kezelheti úgy is, hogy az összegeket, betűket vagy hangmagasságokat képvisel. A fizikai működés önmagában nem árulja el, melyik értelmezést választottuk. A kapcsolatot a kódolás, a feladat leírása és a kimenet használata teremti meg.

Ebből még nem dönthető el, hogy egy későbbi, összetettebb gép képes lehet-e megértésre. A történet ezen a pontján egy szűkebb, de alapvető eredményt látunk: a jelképekkel végzett műveletek egy része elválasztható attól, hogy az ember minden lépést fejben hajtson végre. A gép képességeit az is meghatározza, milyen reprezentációt és eljárást tudunk adni számára.

Pascal a számolás egy részét tette mechanikussá. Leibniz több aritmetikai művelet szervezésére keresett megoldást. Babbage a változtatható műveleti tervhez tervezett általános gépet. Lovelace pedig világosan megfogalmazta, hogy a megfelelően képviselt viszonyok nem korlátozódnak puszta mennyiségekre. A következő lépéshez olyan jelképes nyelv kellett, amely magukat az állításokat és következtetéseket teszi műveletek tárgyává.

8. A logika találkozása a matematikával

A „minden”, a „nem”, az „és”, a „vagy” és a „ha” rövid szavak, mégis bonyolult gondolatmeneteket kapcsolnak össze. Hétköznapi használatukban azonban gyakran bizonytalan, mire vonatkoznak. A „nem minden jelentés hibás” például mást állít, mint a „minden jelentés hibátlan”. Az első megengedi, hogy legyen hibás jelentés; csak azt zárja ki, hogy valamennyi az legyen.

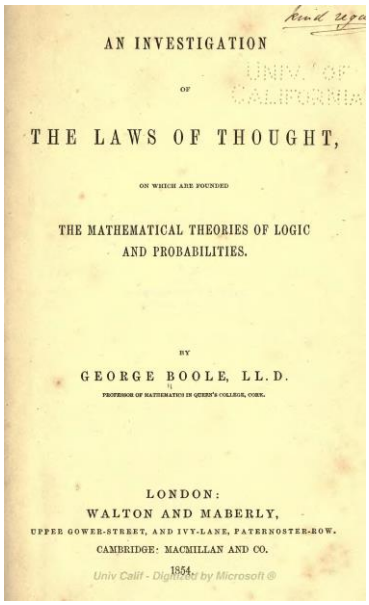
A tizenkilencedik század logikusai olyan eszközöket kerestek, amelyekkel az efféle különbségeket pontosan lehet rögzíteni. A matematika nem egyszerűen új témát kapott: maga a következtetés vált matematikai vizsgálat tárgyává. George Boole és Gottlob Frege eltérő utakon jártak, de mindketten hozzájárultak ahhoz, hogy a gondolatmenetek szerkezete jelekkel kezelhető legyen.

Boole: algebra a fogalmak között

Boole 1847-ben megjelent A logika matematikai elemzése című munkájában, majd az 1854-es A gondolkodás törvényeinek vizsgálata című könyvében algebrai módszerekkel dolgozott logikai problémákon. A fogalmak és osztályok közötti kapcsolatokat egyenletek segítségével kezelte. [46–47]

Képzeljük el az összes tárgy egy kiválasztott körét, például egy könyvtár állományát. Az x jelölje azokat a köteteket, amelyek magyar nyelvűek, az y pedig azokat, amelyek technikatörténettel foglalkoznak. Az xy szorzatszerű kifejezés ekkor azokat a köteteket választja ki, amelyek mindkét feltételnek megfelelnek. A művelet jelentését nem két könyv darabszámának összeszorzása adja, hanem a két kiválasztás együttes alkalmazása.

Ha ugyanazt a kiválasztást kétszer alkalmazzuk, nem kapunk új osztályt. A magyar nyelvű könyvek közül ismét a magyar nyelvűeket választva ugyanott maradunk. Ezt fejezi ki Boole rendszerének jellegzetes összefüggése: $x^2 = x$. A közönséges számalgebrában nem minden számra igaz ez az egyenlet; itt a jelekhez rendelt műveleti értelmezés magyarázza. [47]


AN INVESTIGATION
OF
THE LAWS OF THOUGHT,
ON WHICH ARE FOUNDED
THE MATHEMATICAL THEORIES OF LOGIC
AND PROBABILITIES.
BY
GEORGE BOOLE, LL.D.
PROFESSOR OF MATHEMATICS IN QUEEN'S COLLEGE, CORK.
LONDON:
WALTON AND MABERLY,
UPPER GOWER-STREET, AND IVY-LANE, PATERNOSTER-ROW.
CAMBRIDGE: MACMILLAN AND CO.
1854.

24. kép George Boole An Investigation of the Laws of Thought című, 1854-es könyvének címlapja. A címben a logika és a valószínűség matematikai elmélete is megjelenik. University of California Libraries / Internet Archive / Wikimedia Commons. Közkincs.

Az 1 az egész kiválasztott tartományt, a 0 az üres osztályt képviselheti. Az 1 − x azokat választja ki, amelyek nincsenek az x osztályban. A példában ez a könyvtár nem magyar nyelvű köteteinek köre. Már a „nem” jelentéséhez is ismernünk kell tehát, mely tárgyakról beszélünk: az egész világ minden létezője és egy könyvtár állománya más tartomány.

Boole eredeti eljárásai nem azonosak minden részletükben azzal, amit ma iskolai vagy informatikai tankönyvek Boole-algebraként mutatnak be. Különösen a történeti összeadás értelmezését nem szabad gondolkodás nélkül a mai befogadó „vagy” műveletével azonosítani. Az utóbbi fejlődés eredményét használhatjuk a megértéshez, ha jelezzük, hogy korszerű jelölésről van szó. [46–47]

Igaz és hamis: a műveletek áttekinthetővé válnak

A következő táblázat a klasszikus, kétértékű állításlogika mai jelölését használja. P és Q két állítás; I az igaz, H a hamis értéket jelenti. A ¬ jel a tagadás, a ∧ az „és”, a ∨ a befogadó „vagy”, a → pedig a logikai feltételes kapcsolat jele. A táblázat nem Boole könyvéből átvett történeti ábra.

P	Q	¬P	P ∧ Q	P ∨ Q	P → Q
I	I	H	I	I	I
I	H	H	H	I	H
H	I	I	H	I	I
H	H	I	H	H	I

Mai igazságtáblázat. I = igaz, H = hamis. A ∨ befogadó vagyot, a → klasszikus logikai feltételes kapcsolatot jelöl. Saját szerkesztés.

Az „és” csak akkor igaz, ha mindkét állítás igaz. A befogadó „vagy” akkor is igaz, ha mindkettő igaz: legalább az egyiknek teljesülnie kell. Ez eltér attól a hétköznapi választási helyzettől, amelyben a „tea

vagy kávé” úgy értendő, hogy csak az egyiket kérjük. A szabályt előre rögzíteni kell, különben ugyanazzal a szóval eltérő műveleteket végzünk.

A feltételes kapcsolat a legszokatlanabb. A P → Q a klasszikus értelmezésben pontosan akkor hamis, amikor P igaz, Q pedig hamis. Gondoljunk erre úgy, mint egy szabály megsértésének vizsgálatára. Ha a szabály az, hogy minden aláírt szerződés kap nyilvántartási számot, akkor az aláírt, de számmal nem rendelkező szerződés ellenpélda. Egy alá nem írt irat önmagában nem cáfolja ezt a szabályt.

A → e jelentése nem állít oksági kapcsolatot, és nem fejezi ki a hétköznapi feltételes mondatok összes árnyalatát. Azt sem mondja meg, miért igaz Q. A pontosság ára itt az, hogy a természetes nyelv sokféle használatából egy meghatározott kapcsolatot emelünk ki.

A táblázat haszna, hogy a megengedett értékkombinációk ellenőrizhetők. Ha azt tudjuk, hogy P igaz és P → Q is igaz, akkor Q nem lehet hamis. Ez a modus ponens nevű következtetési szabály lényege. A művelethez nem kell tudnunk, hogy P esőről, szerződésről vagy egy mérési eredményről szól. A kiinduló állítások elfogadhatóságának vizsgálata azonban továbbra is külön feladat.

Frege: a mondat belső szerkezete

Az állításlogikában egy egész mondatot kezelhetünk P jelként. A matematikai érveléshez azonban a mondatok belsejébe is be kell jutni. Meg kell különböztetni az egyedi tárgyakat, a tulajdonságokat és a több tárgy közötti kapcsolatokat. Nemcsak az számít, hogy egy állítás igaz-e, hanem az is, hogy kiről és milyen általánossággal állít valamit.

Frege 1879-es Begriffsschrift című műve ilyen célra alkotott formális nyelvet. A cím fogalomírásként vagy fogalomírási rendszerként adható vissza. Jelölése sajátos, kétdimenziós szerkezetű volt, nem a mai tankönyvek megszokott szimbólumsora. A vállalkozás jelentőségét a kifejezhető logikai szerkezet adta: a változók, a függvényszerű felépítés és a mennyiségjelölés együtt jóval gazdagabb következtetéseket tettek kezelhetővé. [48]


BEGRIFFSSCHRIFT,

EINE DER ARITHMETISCHEN NACHGEBILDETE

FORMELSPRACHE

DES REINEN DENKENS.

VON

Dr. GOTTLOB FREGE.

PRIVATDOCENTEN DER MATHEMATIK AN DER UNIVERSITÄT JENA.

HALLE a/S.

VERLAG VON LOUIS NEBERT.

1879.

25. kép Gottlob Frege Begriffsschrift című, 1879-ben megjelent művének címlapja. Az alcím a tiszta gondolkodás aritmetikai mintára kialakított formulanyelvét ígéri. Wikimedia Commons. Közkincs.

Mai jelöléssel az M(x) azt jelentheti, hogy x emlős, az A(x) pedig azt, hogy x állat. A „minden emlős állat” állítást ekkor így írhatjuk: ∀x (M(x) → A(x)). A ∀ jel azt fejezi ki, hogy a megadott tartomány minden x elemére érvényes a mögötte álló kapcsolat. Ha külön tudjuk egy s-sel jelölt egyedről, hogy M(s), akkor következik A(s). Ez mai szemléltető átírás, nem Frege eredeti tipográfiájának utánzata.

A ∃ jel azt mondja: létezik legalább egy megfelelő elem. A kettő különbsége egyszerűnek látszik, de az egymásba ágyazásuk már a természetes nyelv fontos bizonytalanságait tárja fel. Nem mindegy például, hogy minden olvasóhoz található egy könyv, amelyet elolvasott, vagy található egyetlen könyv, amelyet minden olvasó elolvasott.

Az első mondatban a könyv személyenként változhat. Károly egy technikatörténeti kötetet olvasott, Éva egy regényt, Péter egy útleírást. A második mondathoz közös könyvre van szükség. Az első állításból ezért nem következik a második. Ha vannak olvasók, a másodikból viszont következik az első.

A különbséget a ∀ és a ∃ sorrendje teszi láthatóvá. A formalizálás itt nem puszta rövidítés. Olyan félreértést tár fel, amelyet a gördülékeny nyelvi megfogalmazás könnyen elfed. A matematikai bizonyításokban az ilyen eltérések egész tételek jelentését változtathatják meg.

Mitől formális egy rendszer?

Egy formális rendszerben megadjuk a használható jeleket és a kifejezések képzésének szabályait. Ez dönti el, mi számít egyáltalán szabályos formulának. Rögzítünk kiinduló tételeket, azaz axiómákat, valamint olyan következtetési szabályokat, amelyekkel korábbi formulákból újakat vezethetünk le.

A bizonyítás így ellenőrizhető lépéssorozattá válik. Minden sorban megvizsgálható, hogy az adott formula megengedett kiindulópont-e, vagy valóban következik-e korábbi sorokból az egyik szabály szerint. Aki ellenőrzi, annak nem kell minden lépésnél új ötletet kitalálnia. A bizonyítás megtalálása és egy már megadott bizonyítás ellenőrzése ezért nagyon különböző nehézségű munka lehet.

Fontos megkülönböztetés a szintaxis és a szemantika. Az előbbi a jelek szabályos alakjával és átalakításával foglalkozik. Az utóbbi az értelmezésükkel: milyen tárgyakra vonatkoznak, és milyen feltételek mellett igazak a formulák. Egy leírt jelsor nem attól lesz a valóság helyes leírása, hogy hibátlanul követi a nyelvtant.

Két kérdés is elkülönül. Levezethető-e egy állítás a választott axiómákból a megadott szabályokkal? Illetve igaz-e a vizsgált értelmezésben? Egy jól megválasztott logikai rendszerben e két oldal kapcsolatát külön matematikai tételekkel lehet vizsgálni. Nem helyes egyszerűen feltételezni, hogy bármely tetszőleges jelrendszerben automatikusan egybeesnek.

Frege célja túlment a gondolatmenetek kényelmes lejegyzésén. A számtan logikai megalapozását kereste; 1884-es A számtan alapjai című munkája ennek a programnak fontos állomása. Nem egy emberi elméhez hasonló gépet épített, hanem azt vizsgálta, milyen alapokra helyezhető a matematikai tudás. [49]

A gépi következtetés feltételei

A logika matematikai leírása különösen fontos lehetőséget nyitott meg. Ha a jelek felismerhetők, az engedélyezett átalakítások pedig pontosan rögzítettek, akkor egy szerkezet is végrehajthat bizonyos levezetési lépéseket. Nem szükséges minden egyes alkalommal újra elmagyarázni neki a témát. Elég, ha a reprezentációt és a szabályokat megfelelően kezeli.

Ezzel azonban még nem oldottuk meg, hogyan találjon meg a gép egy hosszú bizonyítást. Az sem következik belőle, hogy minden értelmes kérdéshez létezik általános, mindig célba érő eljárás. A pontos szabályok létezése és az összes feladat gépi megoldhatósága külön állítás. E különbség később a számíthatóság elméletének egyik központi témája lesz.

A valóság leírása is nyitott feladat marad. Egy hiányos adatbázisból hiányozhat egy fontos tény; egy rosszul megfogalmazott axióma olyan következtetéseket engedhet meg, amelyeket nem akartunk. Egy bizonyítás ellenőrzése ezért nem helyettesíti a feladat és az alapfeltevések ellenőrzését.

Boole és Frege munkája mégis megváltoztatta a lehetőségek körét. A számoló szerkezetekhez immár olyan elméleti eszközök társultak, amelyekkel állításokat és következtetési műveleteket is le lehetett írni. Az egyik hagyomány megmutatta, hogyan hajthatók végre műveletek fizikai géppel; a másik megmutatta, hogyan válhat a következtetés pontos műveletek tárgyává. Találkozásuk nyomán a gondolkodó gép kérdése új alakot öltött: mely szellemi feladatok írhatók le végrehajtható eljárásként, és meddig terjed ezeknek az eljárásoknak a határa?

Források és további olvasmányok

A hivatkozások és a képek számozása az 1–5. fejezetét folytatja. A mai jelölésű példák és táblázatok magyarázó szemléltetések; a történeti művek eredeti jelrendszerétől szükség szerint eltérnek.

[38] Arisztotelész: Első analitika, I. könyv, különösen 1. és 4. fejezet. A. J. Jenkinson angol fordítása. A szillogizmus fogalma és a következtetési formák. A fejezet állatos és mérnöki példái mai szemléltetések. Elsődleges szöveg

[39] Gottfried Wilhelm Leibniz: Dissertatio de arte combinatoria. Leipzig, 1666. A fogalmi kombináció és a felfedezés módszerének korai programja. A digitalizált példányt a Biblioteca Nazionale Centrale di Firenze őrzi. Bibliográfiai adatlap · Digitalizált példány

[40] Leibniz: De arte characteristica ad perficiendas scientias ratione nitentes. In: C. I. Gerhardt szerk., Die philosophischen Schriften, VII. Berlin, 1890, 198–203. Az általános jelrendszerről és a számolásszerű érvelésről szóló töredék. A magyar szöveg összefoglaló értelmezés, nem szó szerinti idézet. A Würzburgi Egyetem történeti kiállítása a program megvalósulatlanságát is hangsúlyozza. A Gerhardt-kötet digitalizálása · Würzburgi Egyetem

[41] Musée des Arts et Métiers: Machine arithmétique de Pascal, à 6 chiffres, sans sous ni deniers; továbbá a sous és deniers egységeket is kezelő változat tárgyismertetője. A fejlesztés 1642-es kezdete és a pénzrendszerhez igazított kerekek. Tízes beosztású változat · Pénzügyi változat

[42] Heinz Nixdorf MuseumsForum: Herr Leibniz und seine Rechenmaschine; Herr Leibniz und seine verlorenen Rechenmaschinen. Arithmeum, Universität Bonn: a fennmaradt, 1694-hez kapcsolt fiatalabb Leibniz-gép tárgyismertetője. A korai modellek és a későbbi megvalósítás elkülönítése. HNF áttekintés · HNF a korai gépekről · Arithmeum

[43] Charles Babbage: Passages from the Life of a Philosopher. London, 1864, V–VIII. fejezet. Elsődleges leírás a differenciagépekről és az analitikus gépről. A történeti keltezésekhez és a meg nem épült gépek elkülönítéséhez: Computer History Museum, The Babbage Engine, The Engines. Babbage teljes műve · Computer History Museum

[44] Luigi Federico Menabrea: Sketch of the Analytical Engine Invented by Charles Babbage. Ada Lovelace angol fordításával és A–G jegyzeteivel. Scientific Memoirs, 3, 1843, 666–731. Az A jegyzet a számokon túli viszonyokról és a zene feltételes lehetőségéről; a G jegyzet a Bernoulli-számok eljárásáról. A magyar megfogalmazások parafrázisok. Elsődleges szöveg és táblázatok

[45] Computer History Museum: The Babbage Engine, Ada Lovelace. A fordítás és a jegyzetek keletkezése, a Babbage-dzsel folytatott együttműködés és az „első programozó” megjelölés történeti kérdései. Kurátori ismertető

[46] George Boole: The Mathematical Analysis of Logic Being an Essay Towards a Calculus of Deductive Reasoning. Cambridge–London, 1847. Az algebrai logika korai kifejtése. Digitalizált elsődleges mű

[47] George Boole: An Investigation of the Laws of Thought on Which Are Founded the Mathematical Theories of Logic and Probabilities. London, 1854. Különösen II–III., V. és XI. fejezet. A könyvtári példa és a mai igazságtáblázat saját szemléltetés. Teljes elsődleges szöveg

[48] Gottlob Frege: Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens. Halle, Louis Nebert, 1879. A formulanyelv és a levezetési szabályok eredeti kifejtése. A fejezet ∀, ∃ és → jeleket használó formulái korszerű átírások, nem az 1879-es jelölés reprodukciói. Göttingeni digitalizálás · International Frege Society forrásjegyzéke

[49] Gottlob Frege: Die Grundlagen der Arithmetik Eine logisch mathematische Untersuchung über den Begriff der Zahl. Breslau, W. Koebner, 1884. A számtan logikai megalapozására irányuló program. A digitalizált kiadások elérhetők az International Frege Society elsődleges műveket felsoroló archívumában. Elsődleges művek digitális archívuma

Képforrások és felhasználási adatok

18. kép Saját szerkesztésű, mai halmazábra a fejezet kutyákról, emlősökről és állatokról szóló példájához.

19. kép Eredeti fájl és részletes forrásadatok · Közkincs. A digitalizálás 15. képlapjának körülvágott részlete.

20. kép Eredeti fájl és részletes forrásadatok · Közkincs.

21. kép Eredeti fájl és részletes forrásadatok · Közkincs.

22. kép Saját szerkesztésű szemléltető ábra Babbage és Lovelace történeti leírásainak a képaláírásban megjelölt elemeiről.

23. kép Eredeti fájl és részletes forrásadatok · Közkincs.

24. kép Eredeti fájl és részletes forrásadatok · Közkincs. A digitalizálás 5. képlapja.

25. kép Eredeti fájl és részletes forrásadatok · Közkincs.

II. rész – A gondolkodás kiszámíthatóvá válik?

9. A kiszámíthatóság határai

A szabályos következtetés és az általános célú gép gondolata könnyen merész reményhez vezetett. Ha egy feladat pontosan leírható, talán előbb-utóbb megoldható hozzá egy eljárás is. Ha pedig megvan az eljárás, a gépnek csak végre kell hajtania. A huszadik század első felének matematikája megmutatta, hogy e két „ha” között elvi szakadék is lehet.

Gödel, Church és Turing eredményei nem egyszerűen azt mondták ki, hogy a gépek néha lassúak, a matematikusok pedig nem elég ügyesek. Bizonyos korlátok akkor is megmaradnak, ha tetszőlegesen sok idővel és növelhető tárhellyel számolunk. Ugyanakkor ezek az eredmények nem az emberi megismerés teljes kudarcát jelentik. Pontosan körülhatárolt kérdésekre adnak pontos válaszokat.

A bizonyítás és az igazság nem ugyanaz

Az előző fejezetben a formális rendszer axiómákból és megengedett következtetési lépésekből állt. Kézenfekvő cél, hogy a matematika minden kérdését ilyen rendben lehessen eldönteni. Ehhez a rendszernek következetesnek kell lennie: ne lehessen benne egy állítást és annak tagadását egyaránt bizonyítani. A teljesség más követelmény: a vizsgált állítások mindegyikét vagy a tagadásukat le lehessen vezetni.

Kurt Gödel 1931-es nemteljességi eredményei megmutatták e törekvés határát. A ma szokásos, Rosser későbbi erősítését is figyelembe vevő megfogalmazás szerint egy ellentmondásmentes, algoritmikusan felsorolható axiómákra épülő, az elemi számtan megfelelő részét tartalmazó rendszer nem lehet teljes. Lesz olyan számtani állítás, amelyet sem bizonyítani, sem cáfolni nem tud a saját szabályaival. A feltételek lényegesek: a tétel nem minden tetszőleges szabálygyűjteményről szól. [50–51]

A gondolatmenet egyik kulcsa a kódolás. Jelekhez, formulákhoz és bizonyításokhoz természetes számokat rendelhetünk. Ekkor egy bizonyítás formai tulajdonságairól szóló kijelentés számtani kapcsolatok segítségével is kifejezhető. A számtan így közvetett módon saját leírásának és bizonyítási eljárásainak tárgyává válik. [50]

A híres önhivatkozás ennek a pontos konstrukciónak a része. Magyarázó rövidítéssel olyan mondatot képzelhetünk el, amely azt állítja magáról, hogy az adott rendszerben nem bizonyítható. Ez nem azonos a hétköznapi „ez a mondat hamis” paradoxonnal: a bizonyíthatóság meghatározott, számtanilag leírt fogalmáról van szó. Ha a kiinduló axiómák a természetes számok szokásos értelmezésében igazak, az igazság köre tágabb lehet annál, amit a választott rendszerben bizonyítani tudunk. [50]

Gödel második nemteljességi tétele az önigazolás reményét korlátozza. A megfelelően erős, ellentmondásmentes rendszer a saját szokásos eszközeivel nem bizonyíthatja a maga ellentmondásmentességét. Ettől még egy erősebb rendszer vizsgálhatja a gyengébbet; a kérdés azonban az erősebb rendszer alapjainál ismét felmerül. [50–51]

Ebből nem következik, hogy egyetlen állítás sem bizonyítható biztosan. Egy konkrét levezetés minden lépését továbbra is ellenőrizhetjük. Az sem következik, hogy minden eldöntetlen állítás örökre az marad: új axiómákkal egy korábbi rendszer nyitott kérdése eldönthetővé válhat. A korlát azt mondja, hogy az így kibővített, megfelelő feltételeknek eleget tevő rendszer sem fog minden számtani kérdést lezárni.

A nemteljesség itt bizonyos állításoknak az adott számtani elméleten belüli eldönthetetlenségét jelenti. Ettől különbözik az elsőrendű logika teljessége: az utóbbi arról szól, hogy a minden értelmezésben érvényes logikai következményekhez létezik formális bizonyítás. Az egyes számtani elméletek axiómái nem rögzítik teljesen a természetes számok valamennyi tulajdonságát. A kétféle teljesség tehát nem ugyanarra a kérdésre válaszol.

Church: mit nevezünk eljárásnak?

Mielőtt azt bizonyítanánk, hogy valami algoritmussal nem oldható meg, tisztázni kell, mit értünk algoritmuson. Nem elég annyit mondani, hogy „valamilyen ügyes módszer”. Olyan szabályozott eljárásról beszélünk, amelyben minden lépés végrehajtása egyértelmű, és egy eldöntési feladat minden megengedett bemenetére véges időn belül helyes igen vagy nem választ kell kapnunk.

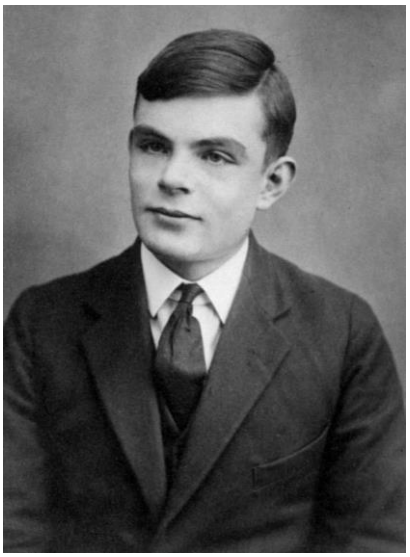
Alonzo Church az 1930-as években a lambda-kalkulushoz és a rekurzív függvényekhez kapcsolódó fogalmakkal közelítette meg az effektív kiszámíthatóságot. A lambda-kalkulus a függvények megadását és alkalmazását szabályozott jelátalakításokkal kezeli. Egészen más kiindulópontnak látszik, mint a fogaskerekes számológép, mégis a végrehajtható számítás pontos leírását célozza. [52]

Church 1936-os eredményei között szerepel, hogy az elsőrendű logika általános eldöntési problémájára nincs ilyen minden esetben célba érő eljárás. Nem készíthetünk olyan algoritmust, amely bármely megfelelően felírt elsőrendű formuláról véges idő alatt eldönti, hogy minden értelmezésben igaz-e. Ez a Hilbert nevéhez kötődő Entscheidungsproblem negatív megoldása. [52–53]

A különbség jól látható a nyolcadik fejezet igazságtáblázatához képest. Véges számú állításbetű esetén végigpróbálhatjuk a lehetséges igazságérték-kiosztásokat. Az általános elsőrendű logikában a mennyiségjelölők és a kapcsolatok ennél gazdagabb világot nyitnak meg. A véges táblázat módszerét nem lehet egyszerűen minden ilyen kérdésre kiterjeszteni.

Turing egyszerű gépe

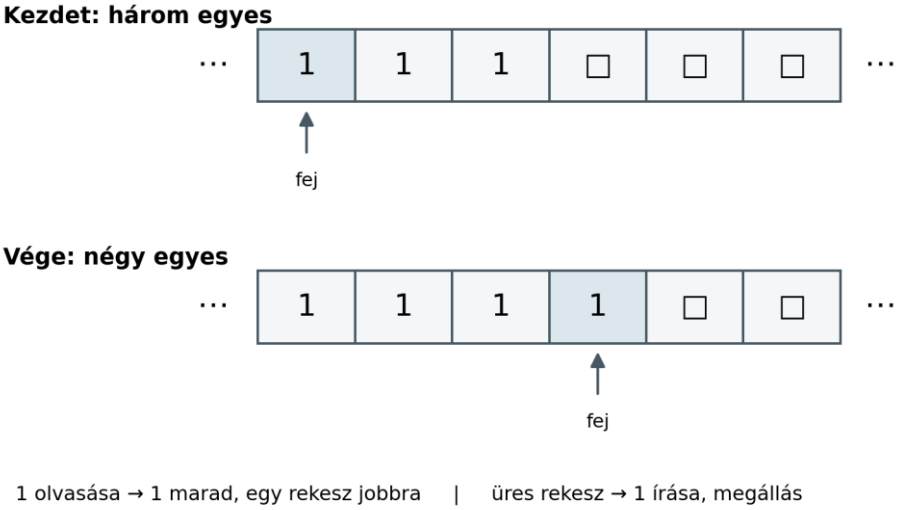
Alan Turing 1936-ban más úton jutott a számítás pontos modelljéhez. Egy papíron dolgozó, szigorúan szabályokat követő számoló ember tevékenységét bontotta elemi lépésekre. A róla elnevezett gép gondolatbeli szerkezet: nem egy elkészült elektronikus berendezés tervrajza. [53]



26. kép Alan Turing tizenhat éves kori útlevélképe, 1928. A fénykép korábbi a számíthatóságról szóló, 1936-os tanulmányánál. Ismeretlen fényképész / Wikimedia Commons. Közkincs.

A modellben a gépnek véges számú belső állapota van. A munkaterületet rekeszekre osztott, szükség szerint korlátlanul bővíthető szalag képviseli. Az író-olvasó fej egy adott rekesz jelét látja. A pillanatnyi állapot és a beolvasott jel alapján a szabály előírja, mit írjon a rekeszbe, merre mozduljon a fej, és milyen állapot következzen. [53]

Mai szemléltető példánkban az egymást követő 1-esek egy számot képviselnek. A 111 a hármat jelenti. A gép a bal szélső 1-esről indul; jobbra haladva változatlanul hagyja az 1-eseket, az első üres helyre pedig még egy 1-est ír, majd megáll. Ezzel a 111-ből 1111 lesz: ebben az egyszerű kódolásban eggyel növelte az értéket.



27. kép Mai szemléltető Turing-gép-példa: az 1-esek sorozatának jobb szélére még egy 1-es kerül. A fej balról indul; az első üres rekeszbe ír, majd megáll. A □ üres rekeszt jelöl, a szalag mindkét oldalon folytatható. Saját szerkesztés, nem Turing eredeti ábrája.

A példa szándékosan szerény. A lényeg az, hogy a teljes működés néhány helyi szabályból áll. A fejnek nem kell egyszerre látnia az egész szalagot, és nem kell minden lépésben „megértenie”, hogy

növelést végez. A nagyobb feladatokhoz több állapot és összetettebb szabályrendszer kellhet, az alapelv azonban ugyanaz marad.

Turing univerzális gépe további fordulatot hozott. Egy gép leírása szintén kódolható jelsorozatként. Így készíthető olyan általános gépmodell, amely megkapja egy másik gép leírását és annak bemenetét, majd szimulálja a működését. Az általánosság a leírás értelmezéséből származik, nem abból, hogy minden feladathoz új fizikai szerkezetet készítünk. [53]

A megállási probléma

A számíthatóság határát ma gyakran a megállási probléma segítségével magyarázzuk. Adott egy program és a hozzá tartozó bemenet. Eldönthető-e általános eljárással, hogy a program egyszer befejezi a munkáját, vagy örökké fut? A válasz: nincs minden programra és bemenetre helyesen működő, mindig véges idő alatt választ adó algoritmus. Ez Turing eredményeinek mai, programok nyelvén megfogalmazott bemutatása; eredeti tanulmánya más technikai fogalmakat használ. [53]

A bizonyítás ötletét egy feltételezett tökéletes ellenőrrel szemléltethetjük. Az ellenőr minden programról és bemenetről megmondaná, megáll-e. Építsünk most egy olyan új programot, amely a bemenetként kapott program saját leírásán való futását vizsgáltatja meg vele. Ha az ellenőr megállást jósol, az új program szándékosan végtelen ciklusba lép. Ha végtelen futást jósol, az új program azonnal megáll.

Ezután az új programnak a saját leírását adjuk bemenetül. Ha az ellenőr szerint megáll, a szabálya szerint nem áll meg. Ha az ellenőr szerint nem áll meg, a szabálya szerint megáll. Mindkét válasz ellentmondáshoz vezet. Tehát a kiindulásként feltételezett, minden esetben helyes és mindig válaszoló ellenőr nem létezhet.

Egy konkrét programról ettől még megállapíthatjuk, hogy befejeződik. A háromtól nulláig visszaszámláló, majd megálló program viselkedése például könnyen belátható. A korlát az összes lehetséges programot és bemenetet kezelő általános módszerre vonatkozik. A gyakorlati ellenőrző eszközök ezért megadhatnak feltételeket, dolgozhatnak korlátozott programosztályokon, vagy egyes esetekben azt válaszolhatják, hogy nem tudnak dönteni.

Elvi korlát és gyakorlati nehézség

Church és Turing eltérő formalizálásai ugyanazt a kiszámítható függvényosztályt adják. A Church–Turing-tézis azt az állítást kapcsolja ehhez, hogy az intuitív értelemben vett, szabályszerűen végrehajtható számításokat ezek a modellek ragadják meg. A modellek egyenértékűsége matematikailag bizonyítható; az informális eljárásfogalommal való azonosítás tézis. [52–53]

A kiszámíthatóság nem mondja meg, mennyibe kerül a számítás. Egy véges kereséshez lehet biztos eljárásunk, miközben a lehetőségek száma kezelhetetlenül nagyra nő. Ha n kapcsoló minden beállítását megvizsgáljuk, 2^n esetet kapunk. Negyven kapcsolónál ez több mint ezermilliárd lehetőség; hatvannál már több mint egymilliárdszor egymilliárd. A feladat véges, de a teljes végigpróbálás költséges lehet. Ez saját számpélda, és nem bizonyítja, hogy ne létezhetne az adott feladathoz jobb módszer.

Az elvi megoldhatatlanságot több processzor és gyorsabb áramkör sem szünteti meg. A gyakorlati nehézséget viszont sokszor enyhíti jobb algoritmus, közelítés vagy megfelelően behatárolt feladat. Ezért az „ezt nem lehet kiszámítani” mondatnál mindig érdemes tisztázni: elvi lehetetlenségről, ismeretlen módszerről vagy túl nagy erőforrásigényről beszélünk-e.

Mit jelent mindez az intelligens gépekre nézve?

E tételek nem bizonyítják, hogy az ember minden algoritmikus rendszernél többre képes. Egy ilyen következtetéshez azt is meg kellene mutatni, pontosan milyen rendszer írja le az embert, és milyen alapon tudja az ember annak következetességét. A tudat létezéséről vagy hiányáról sem döntenek ezek a matematikai eredmények.

Az AI szempontjából a tanulság jóval konkrétabb. Nem várhatunk minden lehetséges program és helyzet viselkedését hibátlanul eldöntő általános ellenőrt. Ettől még rengeteg hasznos bizonyítás, teszt és biztonsági vizsgálat elvégezhető. A korlátok ismerete abban segít, hogy a valóban igazolt képességet megkülönböztessük az egyetemes garancia ígéretétől.

Miközben a matematikusok feltárták a számítás elvi határait, a mérnökök egyre gyorsabb számolóberendezéseket építettek. Az általános számítás modellje már rendelkezésre állt. A következő feladat az volt, hogy a szabályokat követő, jelképekkel dolgozó eljárásokhoz működő, megbízható és átprogramozható gép készüljön.

10. A számítógép megszületése

Egy fogaskerék vagy relé is képes számítási műveletet közvetíteni. A huszadik század közepének áttörését nem egyetlen új alkatrész hozta. A gyors elektronikus kapcsolás, a megfelelő memória, az automatikus vezérlés és a programozás módszerei együtt teremtették meg a számítógépek új nemzedékét.

A háború rendkívüli erőforrásokat mozgósított. Lövedékpályákat kellett számítani, titkosított üzeneteket elemezni, fizikai folyamatokat modellezni. A célok gyakran katonaiak voltak, de a megoldások később tudományos, gazdasági és társadalmi feladatokhoz is használhatónak bizonyultak. A technika történetéből ezért a felhasználás körülményei sem hagyhatók ki.

Miért nehéz megnevezni az első számítógépet?

Az „első számítógép” kifejezés mögött többféle követelmény rejtőzhet. Mechanikus vagy elektronikus gépet keresünk? Programozhatót vagy egyetlen feladatra készítettet? Általános célút vagy különleges rendeltetésűt? Külső szalagról vezéreltet vagy belső memóriában tárolt programmal működőt? Más válaszhoz jutunk attól függően, melyik tulajdonságot tekintjük döntőnek.

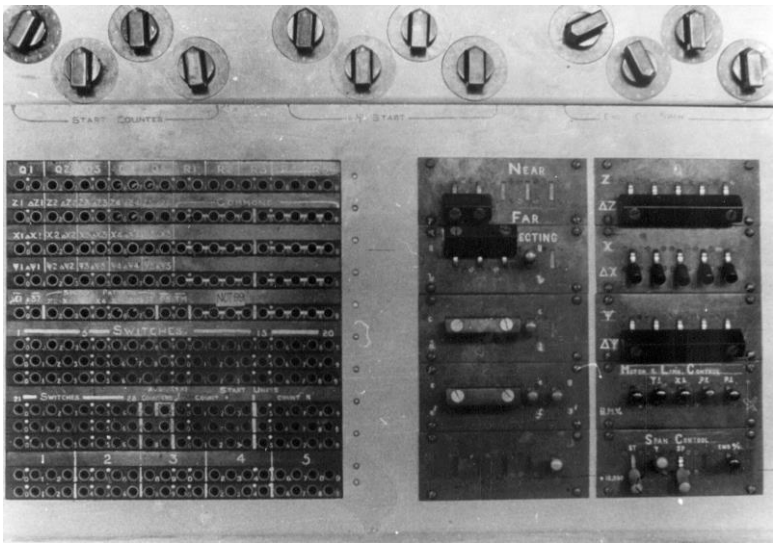
Konrad Zuse 1941-ben bemutatott Z3 gépe relékkel működött, kettes számrendszerű számításokat végzett, és külső adathordozóról kapta a műveleti utasításokat. A programozható, automatikus számolás történetének fontos állomása volt, de nem elektronikus, belső memóriában tárolt programú gép. Az eredeti a háborúban megsemmisült; a múzeumi bemutatásokban rekonstrukcióját láthatjuk. [54]

A relé érintkezőket mozgat: az áram egy elektromágnes segítségével kapcsol át egy másik áramkört. Az elektroncsőben a kapcsolásnak nem kell ilyen mechanikus mozgásra várnia. Ez nagy sebességkülönbséget tett lehetővé, de új gondokat is hozott: hőtermelést, fogyasztást, meghibásodásokat és az alkatrészek együttműködésének nehézségét.

A bináris működés külön kérdés. A két megkülönböztetett állapot alkalmas lehet 0 és 1 képviseletére, de az elektronikus számítógép nem szükségképpen bináris szervezésű: az ENIAC például tízes számjegyekkel számolt. A fizikai alkatrész és a számábrázolás módja összefügg, mégsem ugyanaz a választás. [56]

Colossus: elektronika egy különleges feladatra

A brit Colossus 1944 elején már működött Bletchley Parkban. Tommy Flowers és munkatársai elektronikus gépet építettek a Lorenz-rendszerrel titkosított német távíróüzenetek elemzésének gyorsítására. A szalagról beolvasott jeleken végzett logikai és számlálási műveletek a kódfejtők munkáját segítették. [55]



28. kép Egy Colossus vezérlőpaneljeinek háborús fényképe, amely az 1945-ös General Report on Tunny dokumentációjában is szerepel. A kapcsolók és a csatlakozások a konfigurálható működés kézzelfogható eszközei voltak. Korabeli felvétel, nem a későbbi rekonstrukció képe. Wikimedia Commons. Közkincs.

A Colossus nem az Enigma-gép általános „feltörő számítógépe” volt, és nem Turing 1936-os elméleti gépének egyszerű megépítése. Különleges kriptoanalitikai feladatra készült. Működését kapcsolókkal és csatlakoztatásokkal lehetett konfigurálni; mai értelemben vett, szabadon betölthető általános programot nem tárolt. [55]

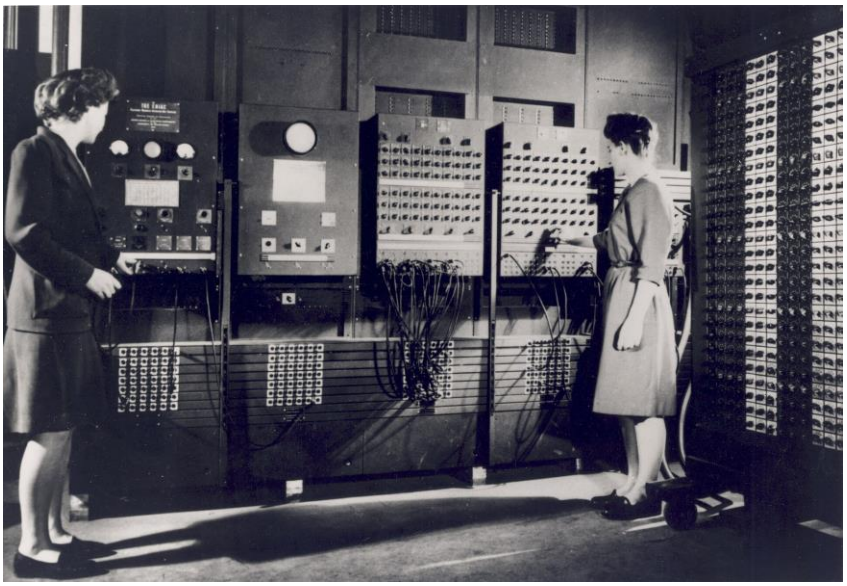
Jelentősége ettől nem kisebb. Megmutatta, hogy sok elektroncsőből felépített, nagy sebességű logikai berendezést tartós munkára lehet fogni. Történetének nyilvánosságát hosszú ideig korlátozta a titkosság. A számítógép fejlődésének későbbi elbeszélései ezért sokáig hiányos képet adtak a brit háborús munkáról. [55]

ENIAC: a sebesség és a programozás ára

Az Egyesült Államokban J. Presper Eckert, John Mauchly és a Pennsylvaniai Egyetem Moore Schoolјának csapata az ENIAC-kal bizonyította a nagy léptékű, általános célú elektronikus számítás gyakorlati lehetőségét. A hadsereg ballisztikai igényeire indult fejlesztés 1945 végére jutott működő gépig; nyilvános bemutatása 1946 februárjában történt, már a háború befejezése után. [56]

A korábbi emberi számolómunka során egy hosszú feladat részlépéseit táblázatokba rendezték, számológépekkel végezték el, majd ellenőrizték és továbbadták. Az ENIAC-ban az elektronikus egységek sok műveletet rendkívül gyorsan hajthattak végre. A sebesség önmagában azonban nem oldotta meg a feladat előkészítését.

Az eredeti ENIAC új feladatra állításakor kábelekkel, kapcsolókkal és részletes működési tervvel kellett kialakítani a műveletek útját. A matematikai eljárást fel kellett bontani a gép egységei által végrehajtható lépésekre, és össze kellett hangolni az időzítést. Egy gyorsan lefutó számítás mögött hosszú előkészítő munka állhatott. A későbbi átalakítások változtattak a vezérlés módján; az eredeti és az átalakított ENIAC képességeit ezért meg kell különböztetni. [56–58]



29. kép Jean Jennings, későbbi nevén Jean Bartik (balra), és Frances Bilas, későbbi nevén Frances Spence (jobbra) az ENIAC-nál, az 1946. februári bemutatóra készülve. Fotó: azonosítatlan amerikai katonai fényképész / U.S. Army / Wikimedia Commons. Közkincs.

A gép első programozói között hat nő munkája különösen jól dokumentált: Kathleen McNulty, Jean Jennings, Betty Snyder, Marlyn Wescoff, Frances Bilas és Ruth Lichterman. Későbbi nevükön többüket Antonelli, Bartik, Holberton, Meltzer, Spence és Teitelbaum néven ismerjük. Nem pusztán adatokat vittek be. A feladatokat működési tervekre bontották, kialakították a gépi végrehajtás rendjét, és hibát kerestek benne. [57]

Ez a munka világossá tette, hogy a számítógép nem csak az elkészült hardver. Szükség van az eljárást megfogalmazó, gépre alkalmazó és ellenőrző tudásra is. A programozás saját szakmai feladattá vált, még mielőtt kialakultak volna a ma megszokott programnyelvek és fejlesztőeszközök.

Neumann János és az EDVAC terve

Neumann János 1944-ben kapcsolódott be az amerikai elektronikus számítógépek fejlesztéséről folytatott munkába. Matematikai, logikai és alkalmazási tapasztalata egyaránt fontos volt. Az ENIAC utódjának szánt EDVAC-ról 1945. június 30-i keltezéssel készült el a First Draft of a Report on the EDVAC című jelentéstervezete. [58–59]



30. kép Neumann János Los Alamos-i igazolványképe, 1943–1947 körül. Matematikai, logikai és alkalmazási tapasztalata egyaránt szerepet kapott a korai számítógépek fejlesztésében. Fotó: Los Alamos National Laboratory / Wikimedia Commons. A forrás névfeltüntetést ír elő.

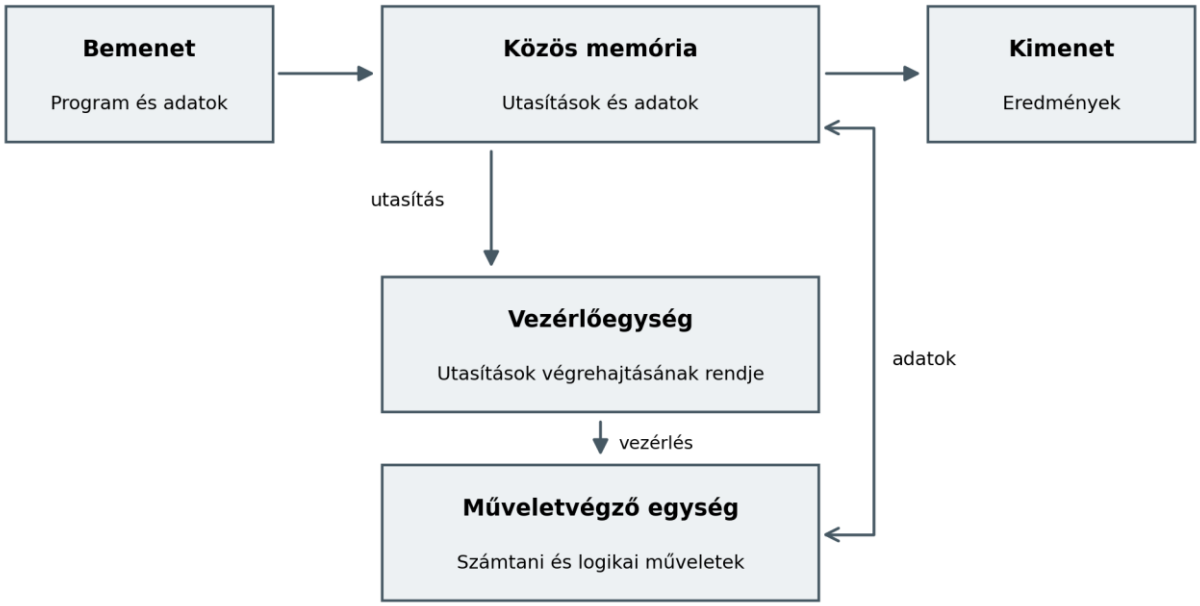
A dokumentum rendszerezte az automatikus számítógép működésének alapvető részeit. Megkülönböztette az aritmetikai műveleteket végző egységet, a központi vezérlést, a memóriát, valamint a külvilággal kapcsolatot tartó be- és kimenetet. A memória szerepe kiterjedt a számítás közbülső eredményeire és a végrehajtás utasításaira is. [58]

A tárolt program gondolata nem egyetlen ember elszigetelt pillanatnyi felismeréseként született. Eckert, Mauchly és munkatársaik már foglalkoztak az ENIAC korlátainak meghaladásával; Neumann e közös fejlesztési környezetben készített nagy hatású, logikai szempontból rendezett leírást. Herman Goldstine a jelentés terjesztésében is szerepet játszott. A történeti érdemeket ezért nem fejezi ki megfelelően az a mondat, hogy Neumann egymaga feltalálta a számítógépet. [59]

A „Neumann-architektúra” elnevezés mégis indokoltan kapcsolódik hozzá: a jelentés és a későbbi munka nagyban hozzájárult ahhoz, hogy e szervezési elvek széles körben ismertté és alkalmazhatóvá váljanak. Nem minden részlet volt új, és nem minden későbbi gép másolta egyformán a tervet. A jelentőség a részek koherens együttműködésének leírásában és a gondolat terjedésében is rejlett.

A program is tárolható információ

A döntő változást hétköznapi példával érthetjük meg. Egy korai gépnél az új feladathoz a kapcsolatok és beállítások jelentős részét át kellett alakítani. A tárolt programú gépben az utasítások kódolt sorozata kerül a memóriába. A vezérlés onnan olvassa ki, milyen műveletet kell végrehajtani, és mely tárolóhelyek adataira van hozzá szükség.



31. kép A klasszikus tárolt programú gép fő funkciói, egyszerűsített mai ábrán. A közös memória utasításokat és adatokat tartalmaz; a vezérlőegység szervezi a műveleteket. A rajz kiválasztott funkcionális kapcsolatokat mutat, nem teljes áramköri terv és nem az EDVAC eredeti szerelési rajza. Saját szerkesztés.

A klasszikus működés leegyszerűsítve három mozzanatból áll: az utasítás elővétele, értelmezése, majd végrehajtása. A vezérlés nyilvántartja a következő utasítás helyét. Egy elágazás ezt a helyet egy feltétel szerint megváltoztatja; egy ismétlés korábbi utasításokhoz vezet vissza. A mai processzorok sok részletben összetettebbek, de az elv megértéséhez ez a modell elegendő.

Képzeljünk el egy mai tanítási célú utasítássorozatot: olvasd be a 7-et; add hozzá az 5-öt; tárold a 12-es eredményt; állj meg. A gép nem ilyen magyar mondatokat olvas, hanem meghatározott kódolású utasításokat és értékeket. Ha a programot megváltoztatjuk, ugyanaz a szerkezet végezhet összeadás helyett összehasonlítást, ismételt számítást vagy más megengedett műveletsort. A példa nem az EDVAC eredeti utasításkészletének másolata.

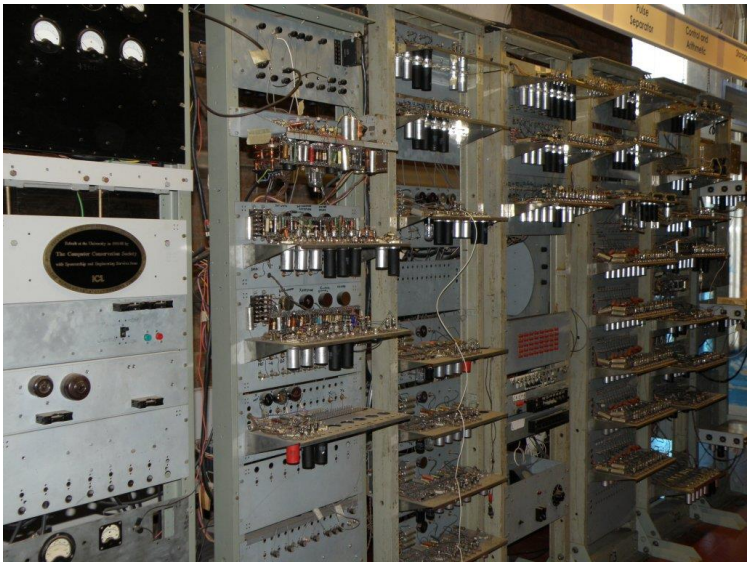
Az adat és az utasítás egyaránt kódolt információ lehet. Attól, hogy ugyanabban a memóriában helyezkednek el, nem tűnik el közöttük a szerepkülönbség: egy bitsorozatot a gép működési helyzete szerint kezel utasításként vagy adatként. A programmal végzett műveletek lehetősége mégis fontos új távlatot nyit. Más programokat betöltő, fordító vagy ellenőrző eljárások készíthetők.

Ez nem jelenti, hogy minden hasznos programnak önmagát kell módosítania. A tárolt program előnye önmódosítás nélkül is óriási: a műveleti terv elektronikus úton elérhető és cserélhető. A gép sokféle feladathoz használható alapberendezéssé válik.

Amikor az elv működő géppé vált

Az EDVAC-jelentés dátuma nem azonos az első ilyen elven működő számítógép üzembe állításával. A terv, a kísérleti megvalósítás és a rendszeres használat külön mérföldkő. Az EDVAC tényleges elkészülése elhúzódott; más műhelyek időközben saját gépeket építettek. [59]

Manchesterben Frederic Williams, Tom Kilburn és Geoff Tootill munkája nyomán a Baby, hivatalos nevén Small-Scale Experimental Machine, 1948. június 21-én futtatta első programját. Jelentőségét pontosan fogalmazhatjuk meg: a programot címezhető, írható-olvasható elektronikus memóriában tárolta. A gép elsősorban az új tárolási módszer működőképességét igazoló kísérleti berendezés volt. [60]



32. kép A Manchester Baby rekonstrukciója a manchesteri Museum of Science and Industry kiállításán. A másolat 1995–1998 között készült; a képen nem az 1948-as eredeti berendezés látható. Fotó: KGGucwa / Wikimedia Commons, CC BY-SA 3.0.

Cambridge-ben Maurice Wilkes vezetésével az EDSAC 1949. május 6-án futtatta első sikeres programjait. Az egyetemi kutatómunka számára használható számítási szolgáltatás kialakítása állt a középpontban. Ezzel a tárolt program elve a kísérleti bizonyításon túl a mindennapi tudományos munka eszközévé vált. [61]

Neumann a princetoni Institute for Advanced Study számítógépes programjában is vezető szerepet vállalt. Az IAS-gép mérnöki megvalósításában Julian Bigelow és munkatársai végeztek meghatározó munkát. A háború után kibontakozó műhelyek nem elszigetelt pontok voltak: jelentések, látogatások és szakmai előadások révén tapasztalatokat cseréltek. [59, 62]

A magyar történetben az MTA Kibernetikai Kutatócsoportjában, szovjet dokumentáció alapján megépített M–3 jelentett fontos kezdetet. Az 1959-ben használatba vett gép körül már a megépítés előtt megindult a programozói munka. A hazai számítástechnika története itt találkozik majd a kibernetika és a mesterséges intelligencia későbbi történetével. [63]

Miért volt ez az AI előfeltétele?

A mesterséges intelligencia nem született meg attól, hogy a programot memóriába helyezték. A korai gépek elsősorban számítási feladatokat hajtottak végre. De létrejött az a rugalmas műszaki alap, amelyen később jelképes keresést, logikai következtetést, játékot vagy tanulási eljárást lehetett kipróbálni.

Egy sakkprogramhoz már nem feltétlenül kellett külön sakkgépet építeni. Le kellett írni az állásokat, a szabályos lépéseket, az értékelés és a keresés módszerét, majd ezeket az általános gép számára

végrehajthatóvá kellett tenni. Ugyanaz a hardver más programmal egészen más szellemi feladat kísérleti eszközévé válhatott.

Ehhez persze elegendő sebesség és tárhely is kellett. A tárolt program önmagában nem tesz kezelhetővé egy túlságosan nagy keresési teret, és nem lépi át a számíthatóság elvi határait. Amit megad, az a módszerek gyors cseréjének és kipróbálásának lehetősége. Az intelligenciáról alkotott elképzelésből működő kísérlet készülhet.

A történet két, egymást kiegészítő eredményhez jutott. A matematikusok pontosabban körülhatárolták, mire képes egy általános eljárás. A mérnökök pedig megteremtették a sokféle eljárást végrehajtó gépet. Ettől kezdve a kérdés már nem csupán az volt, hogy elképzelhető-e gondolkodásra emlékeztető gépi viselkedés, hanem az is, milyen programmal és milyen bizonyítékok alapján lehet azt ténylegesen bemutatni.

Források és további olvasmányok

A hivatkozások és a képek számozása az 1–8. fejezetét folytatja. A mai számítási példák és magyarázó ábrák a történeti eredmények megértését szolgálják; nem korabeli programok vagy eredeti géptervek másolatai.

[50] Kurt Gödel: Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. Monatshefte für Mathematik und Physik, 38, 1931, 173–198. DOI: 10.1007/BF01700692. Az eredeti nemteljességi tételek, a bizonyítások számtani kódolása és az önhivatkozás konstrukciója. Az eredeti első tétel erősebb következetességi feltételét Rosser később gyengítette. Elsődleges tanulmány és bibliográfia

[51] J. Barkley Rosser: Extensions of Some Theorems of Gödel and Church. The Journal of Symbolic Logic, 1/3, 1936, 87–91. DOI: 10.2307/2269028. A puszta ellentmondásmentességre építő erősítés. A fejezet közérthető tételmegfogalmazása a későbbi, szokásos változatot követi. Elsődleges tanulmány

[52] Alonzo Church: An Unsolvable Problem of Elementary Number Theory. American Journal of Mathematics, 58/2, 1936, 345–363. Az effektív kiszámíthatóság, a lambda-definiálhatóság és a rekurzivitás kapcsolata. Az elsőrendű eldöntési problémáról: A Note on the Entscheidungsproblem. The Journal of Symbolic Logic, 1/1, 1936, 40–41. Az elsődleges 1936-os tanulmány

[53] Alan M. Turing: On Computable Numbers, with an Application to the Entscheidungsproblem. Proceedings of the London Mathematical Society, 2. sorozat, 42, 1936–1937, 230–265; helyesbítés: 43, 1937, 544–546. Különösen az 1., 6., 8–9. és 11. rész, valamint a függelék. A megállási probléma és a szalagpélda a fejezetben mai szemléltetés; az eredeti szöveg más technikai fogalmakkal építi fel a bizonyítást. Elsődleges szöveg a helyesbítéssel · Turing Digital Archive

[54] Deutsches Museum: Informatik, a Z3 bemutatása; Computer History Museum: 1941, Timeline of Computer History. A relés, bináris, külső programvezérlésű gép és a későbbi rekonstrukció története. Deutsches Museum · CHM tárgytörténet

[55] The National Museum of Computing: Colossus. Tommy Flowers csapata, a Lorenz-üzenetek elemzése, az 1944-es működés, a konfigurálhatóság és a hosszú titkosság. A fénykép elsődleges forrása: I. J. Good, D. Michie és G. Timms, General Report on Tunny, 1945. Múzeumi ismertető · Korabeli fénykép forrásadatai

[56] Computer History Museum: ENIAC. A gép építése, az új feladatok kábelezése, a működés egységei. University of Pennsylvania: On ENIAC's anniversary, a nod to its female computers, 2019. február 14. A háborús számolómunka és az 1946-os bemutató. Az eredeti gép tízes számábrázolását az EDVAC-jelentés új, bináris tervével is összevetjük. CHM ENIAC-gyűjtemény · A fejlesztő egyetem történeti ismertetője

[57] ENIAC Programmers Project: a programozókkal készített interjúkon alapuló történeti gyűjtemény és The Computers dokumentumfilm. A hat első programozó neve és munkájának jellege. Az IEEE Computer Society úttörőket ismertető oldala szintén dokumentálja hozzájárulásukat. Interjúkon alapuló történeti projekt · IEEE Computer Society

[58] John von Neumann: First Draft of a Report on the EDVAC. Moore School of Electrical Engineering, University of Pennsylvania, 1945. június 30. Különösen 2.0, Main Subdivisions of the System. A hivatkozott digitális kiadás az eredeti szöveg javított, 1992-ben újraszedett változata Michael D. Godfrey gondozásában; nem az eredeti gépirat fényképmásolata. Az elsődleges jelentés digitális kiadása · Smithsonian bibliográfiai rekord

[59] Computer History Museum: Making the Case; IEEE Computer Society, Computer Pioneers: John William Mauchly. Eckert, Mauchly, Neumann és a Moore School közös fejlesztési környezete; a tárolt program gondolatának terjedése és az EDVAC későbbi megvalósulása. A jelentés és a Moore School előadásai · Mauchly munkájának dokumentált története

[60] University of Manchester: The Manchester Small Scale Experimental Machine, The Baby. Williams, Kilburn és Tootill, az elektronikus tárolás és az 1948. június 21-i programfutás. A címezhető, írható-olvasható memória pontosítja a történeti elsőség jelentését. Az egyetem történeti dokumentációja · Az Edinburgh-i Egyetem műszaki ismertetője

[61] University of Cambridge, Department of Computer Science and Technology: 70 years since the first computer designed for practical everyday use, 2019. május. Maurice Wilkes, az EDSAC 1949. május 6-i első programjai és a kutatói számítástechnikai szolgáltatás. Az oldalon korabeli film is elérhető. A fejlesztő egyetem dokumentációja · Whipple Museum gyűjteményi áttekintése

[62] Institute for Advanced Study: Electronic Computer Project; The Electronic Computer Project at IAS. Neumann, Goldstine, Julian Bigelow és a mérnökcsapat, valamint a jelentések terjedése. Az archívum korabeli fényképeket, munkafüzeteket és műszaki dokumentumokat is bemutat. A projekt története · Archív anyagok

[63] NJSZT Informatikatörténeti Fórum: M–3 számítógép. Az 1959–1968-as használat és a szovjet dokumentáció hazai megvalósítása. Elsődleges programozási dokumentum: Az M–3 utasításrendszere alapján készített szabványos és tipikus programok, MTA KKCs, 1958. december 31., 232 oldal; elfogadva 1959. február 24. A hazai gép adatlapja · Korabeli programgyűjtemény és digitalizálása

Képforrások és felhasználási adatok

26. kép Eredeti képfájl és részletes forrásadatok · Közkincs.

27. kép Saját szerkesztésű szalagpélda, a szövegben megadott szabályokkal.

28. kép Eredeti képfájl és részletes forrásadatok · Közkincs.

29. kép Eredeti képfájl és részletes forrásadatok · Közkincs.

30. kép Eredeti képfájl és részletes forrásadatok · Los Alamos National Laboratory. A Commons-forrás névfeltüntetéshez köti a felhasználást; a kredit a képaláírásban is szerepel.

31. kép Saját szerkesztésű funkcionális ábra a klasszikus tárolt programú gép megjelölt kapcsolatairól. A történeti alapelvekhez lásd az [58] forrást.

32. kép Eredeti képfájl és részletes forrásadatok · KGGucwa · CC BY-SA 3.0

III. rész – Visszacsatolás, kibernetika és az első önálló gépek

Egy gép végrehajthatja pontosan az előírt lépéseket, és közben mégis eltérhet a kívánt eredménytől. Megváltozhat a terhelése, lehűlhet a környezete, elmozdulhat a követett tárgy. A következő fordulat ezért a számítás mellett az érzékelésről szól: a működés eredménye visszakerül a rendszerhez, és befolyásolja a folytatást. Ebből a szemszögből egy gőzgép szabályozója, az ember célra tartó mozdulata és egy mesterséges ideghálózat közös kérdést vet fel: hogyan kapcsolódik össze a jel, a cselekvés és a korrekció?

11. Gépek, amelyek érzékelik saját működésüket

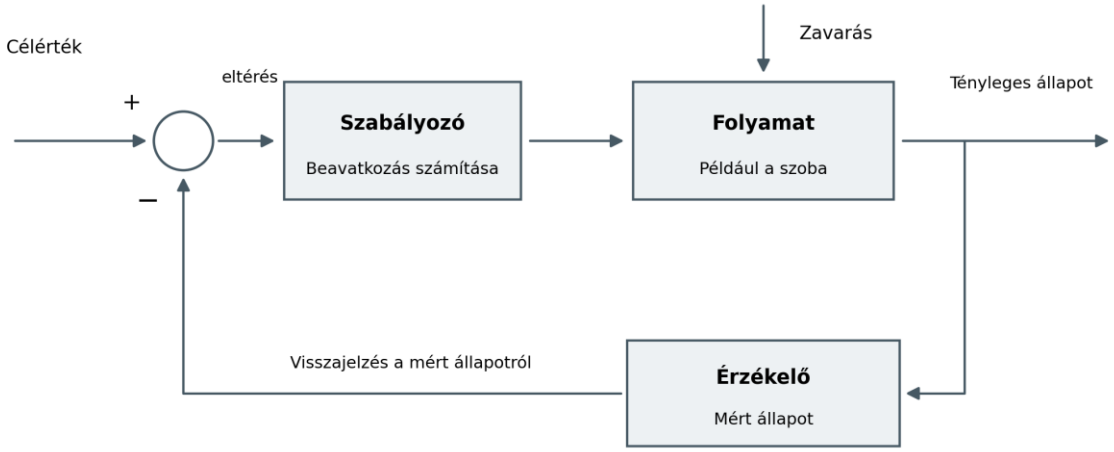
Ha egy szobában este bekapcsolunk egy villamos fűtőtestet, és reggelig változatlan teljesítménnyel működtetjük, a hőmérséklet a kinti időjárástól, a szigeteléstől és az ajtónyitásoktól is függ. A készülék hőt ad le, de nem ellenőrzi, milyen meleg lett. Egy termosztáttal kiegészítve viszont a tényleges hőmérséklet is beleszól abba, mikor kell fűteni.

Ez a visszacsatolás lényege. A rendszer működésének eredményéről információ kerül vissza a beavatkozást meghatározó részhez. A fejezet címében az „érzékelik” műszaki értelemben szerepel: mérésre vagy a mért állapottal összefüggő fizikai hatásra utal. Nem állítja, hogy a gép tudatában volna önmagának.

A kívánt állapot és a tényleges állapot

A szabályozásnak van célértéke, például egy kívánt hőmérséklet, vízszint vagy fordulatszám. Az érzékelés a tényleges állapotról szolgáltat jelet. A kettő eltérése a hibajel. A szabályozó ennek alapján határozza meg a beavatkozást: növelje-e a fűtést, nyisson-e a szelepen, csökkentse-e a motor hajtását. [64–65]

A hibajel elnevezés itt nem feltétlenül meghibásodást jelent. Ha a kívánt hőmérséklet 21 fok, a mért érték pedig 20 fok, az eltérés egyszerűen azt mutatja, hogy még nem értük el a célt. A szabályozás feladata az eltérés megfelelő kezelése. Az, hogy mi számít megfelelőnek, függ a folyamat sebességétől, a megengedett ingadozástól és a rendelkezésre álló beavatkozástól.



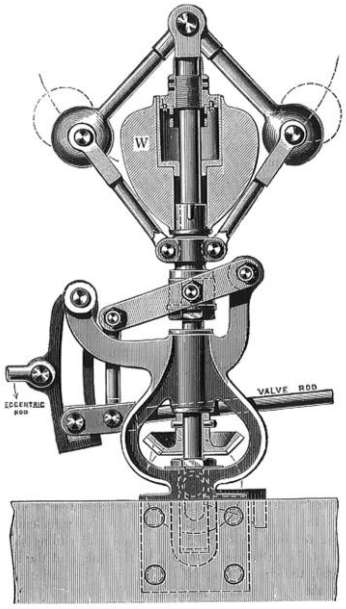
33. kép Negatív visszacsatolású szabályozási kör, egyszerűsített mai ábrán. A célértékből a mért állapot jelét vonjuk ki; az eltérés alapján történik a beavatkozás. A zavarás a folyamatra hat. Saját szerkesztés; az ábra funkcionális kapcsolatokat mutat, nem egy konkrét készülék kapcsolási rajza.

A környezetből érkező zavarás külön szereplő. A szoba lehűlhet egy ablak kinyitásától, a tartály szintje csökkenhet egy váratlan elvételtől. A visszacsatolás egyik előnye, hogy nem kell minden zavarást külön előre megjósolni. Ha hatásuk megjelenik a mért értékben, a szabályozó reagálhat rájuk. A korrekciónak azonban fizikai határa van: egy alulméretezett fűtés a legjobb termosztáttal sem képes tetszőleges hidegben fenntartani a kívánt meleget.

A mechanika is képes mérni és beavatkozni

A visszacsatoláshoz nem szükséges számítógép. Egy úszó a folyadék felszínével együtt mozoghat, és mechanikus kapcsolaton keresztül nyithatja vagy zárhatja a beömlő szelepet. Ugyanaz az alkatrészlánc hordozza a szintre vonatkozó információt, és közvetíti a beavatkozást. A működéshez nem kell számként megjeleníteni a vízszintet.

Az ipartörténet ismert példája a centrifugálszabályozó. A forgó tengelyhez kapcsolt súlyok helyzete a fordulatszámmal változik. Gyorsabb forgásnál a súlyok kijjebb emelkednek; a kapcsolódó mechanika úgy állíthatja a szelepet, hogy kevesebb gőz jusson a gépbe. A fordulatszám csökkenése fordított irányú beavatkozást válthat ki. Watt gőzgépein az 1780-as évek végén alkalmazott megoldás az ipari szabályozás fontos jelképe lett, de a visszacsatolás története nem ezzel az egyetlen találmánnyal kezdődött. [66]



34. kép Centrifugálszabályozó szerkezete a Heat Engines című, 1913-as kézikönyv illusztrációján. A forgó súlyok és a rudazat a szelep beállítását befolyásolják. Későbbi műszaki ábra, nem Watt első szabályozójának eredeti tervrajza. Wikimedia Commons. Közkincs.

A szabályozóban felismerhetjük ugyanazokat a szerepeket, mint a termosztátnál. A fordulatszámot a forgó szerkezet fizikai viselkedése jelzi, a rudazat közvetíti az eltérést, a szelep megváltoztatja a hajtást. Az érzékelő és a szabályozó itt nem külön elektronikus doboz. A szerepek elkülönítése a működés megértését segítő leírás.

Mitől negatív a visszacsatolás?

A negatív visszacsatolásban a beavatkozás az eltérést csökkentő irányban hat. Ha túl meleg van, kevesebb fűtést adunk; ha túl nagy a fordulatszám, csökkentjük a hajtást. A „negatív” tehát nem rosszat jelent, hanem az eltérésre ható kapcsolat irányát. [64–65]

Pozitív visszacsatolásnál egy változás önmagát erősítheti. A mikrofon felveszi a hangszóró hangját, az erősítő nagyobb jellel hajtja a hangszórót, amelynek hangja ismét a mikrofonba jut. Megfelelő erősítésnél kialakulhat a jól ismert sípolás. A pozitív visszacsatolás sem minden helyzetben hibás: kapcsolásokban éles váltást vagy állapotmegőrzést is segíthet. A működés hatását mindig a teljes rendszerben kell megítélni.

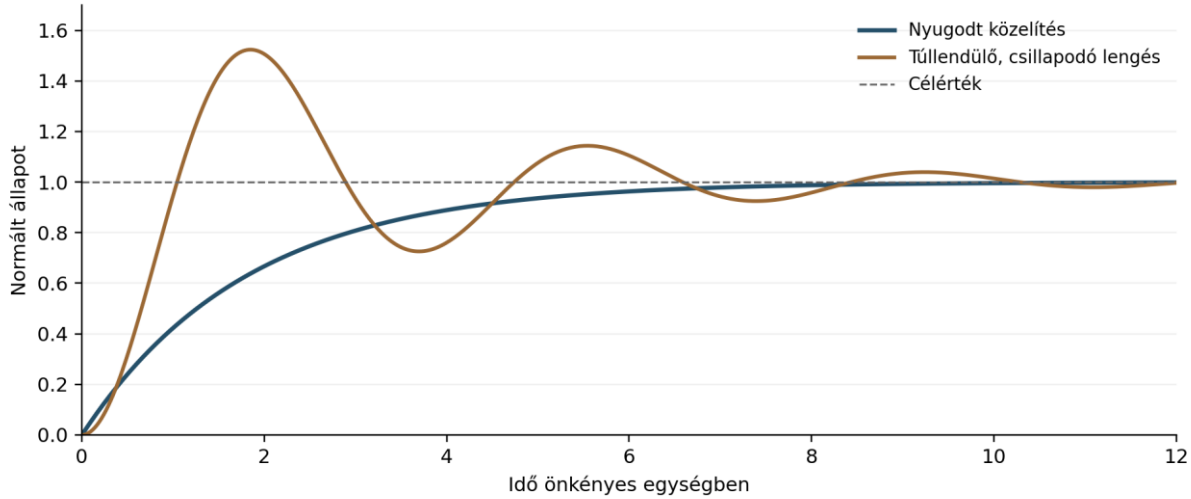
A negatív visszacsatolás önmagában nem garantál nyugodt viselkedést. Ha a korrekció túl erős vagy késve érkezik, a rendszer átlendülhet a célon, majd az ellenkező irányban túl nagyot javíthat. A jószándékú beavatkozásból ingadozás keletkezik. A szabályozás tervezésében ezért a visszacsatolás megléte mellett a stabilitása is döntő.

Maxwell és a szabályozás matematikája

James Clerk Maxwell 1868-as On Governors című tanulmánya már azt vizsgálta, milyen feltételek mellett csillapodnak vagy növekednek a szabályozott gép eltérései. Felismerte, hogy a nagyobb

szabályozóhatás nem mindig jelent jobb működést: bizonyos beállításoknál a zavarásból fokozódó lengés lehet. Az ipari szerkezet viselkedése matematikai vizsgálat tárgyává vált. [64]

Egy háztartási példa ugyanezt szemlélteti. A zuhanynál túl hideg a víz, ezért nagyot fordítunk a melegvizes csapon. Az eredmény csak később érkezik meg. Közben ismét állítunk rajta, majd a víz túl forróvá válik. Visszafordítjuk a csapot, és újabb késés után ismét hideg lehet. A hibát a beavatkozás mértéke és a rendszer késleltetése együtt okozza; nem feltétlenül a csap romlott el.



35. kép Két szemléltető válasz ugyanarra a célértékre: nyugodt közelítés, illetve túllendülő, csillapodó lengés. Mindkét görbe végül a célértékhez tart; nem minden lengő rendszer instabil. Saját szerkesztésű, matematikai példagörbék önkényes időegységgel; nem mérési adatok és nem konkrét szabályozó beállítási útmutatója.

A fűtésnél hasonló szerepe lehet a falak és a radiátorok hőtárolásának. A kazán leállítása után is érkezhet még hő a szobába. A pillanatnyi hőmérséklet ezért nem mond el mindent arról, hogyan változik majd az állapot. A szabályozásnak olykor a változás irányát és a korábbi eltérések hatását is figyelembe kell vennie. [65]

A pontosság ára

Egy egyszerű termosztát be- és kikapcsolhat. Ha ugyanazon hőmérséklet alatt bekapcsolna, fölötte pedig azonnal kikapcsolna, a mérési zaj és az apró változások sűrű kapcsolgatáshoz vezethetnek. Ezért alkalmazhatunk két küszöböt: alacsonyabb értéknél bekapcsolás, magasabbnál kikapcsolás. A köztük levő sáv a hiszterézis. A rendszer szándékosan megenged valamennyi ingadozást a nyugodtabb működésért. [65]

Más feladatban folyamatosan változtatható beavatkozás kell. A nagyobb pillanatnyi hibához nagyobb korrekció rendelhető. A tartósan fennmaradó kis eltérés külön kezelést igényelhet, a gyors változásokra adott érzékeny válasz pedig felerősítheti a mérési zajt. A szabályozó kialakítása ezért kompromisszum a gyorsaság, a pontosság és a nyugodt működés között.

A mérés helye is számít. Egy radiátor fölé szerelt érzékelő a környezeténél magasabb hőmérsékletet jelezhet. A szabályozó következetesen tarthatja ezt az értéket, miközben a szoba másik sarkában hideg marad. A hibás vagy félrevezető visszajelzés a korrekciót is félrevezeti. Ez a tanulság később a tanuló rendszereknél is visszatér: a mért eredménynek valóban a fontos állapotról kell tájékoztatnia.

Szabályozás és önállóság

Egy programóra előre megadott időpontban indítja a fűtést. A termosztát az eredményt is figyelembe veszi. A különbség nem az, hogy az egyikben van szándék, a másikban nincs, hanem az, hogy a tényleges állapot visszahat-e a további működésre. Előre megadott beavatkozás és visszacsatolt korrekció egyébként együtt is alkalmazható.

A szabályozó bizonyos értelemben önálló: nem kell minden pillanatban embernek állítania a szelepet. Célját és lehetőségeit azonban a szerkezet, a beállítás és a környezet határozza meg. A termosztát nem dönti el, mitől lesz méltányos az energiaszámla, és nem tanul meg magától új fűtési stratégiát.

Az automatikus korrekció mégis fontos lépés. A gép nem egyszerűen végigviszi a számára előírt mozdulatokat, hanem a következményekhez igazítja a folytatást. Ez a működési kör lett az egyik alapja annak, hogy a mérnökök, a fiziológusok és a matematikusok közös nyelvet kezdtek keresni az élőlények és a gépek viselkedéséhez.

12. Norbert Wiener és a kibernetika megszületése

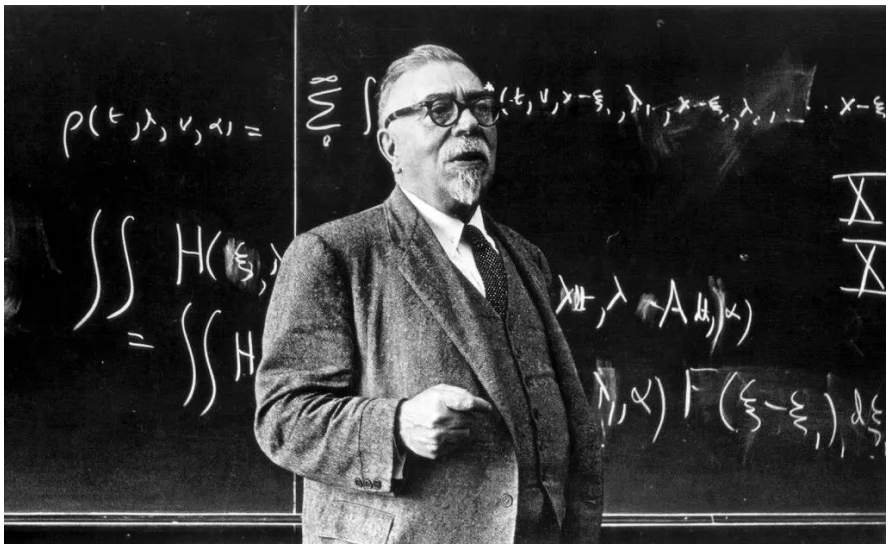
Ha valaki pohárért nyúl, a keze nem feltétlenül egy előre rögzített pályát játszik végig. A látás, az izmok és ízületek felől érkező jelzések segítik a mozdulat korrekcióját. Ha a pohár elmozdul, a kéz pályája is megváltozhat. A cél elérése és az eltérés csökkentése olyan kérdés, amely a gépi szabályozásnál is megjelenik.

A kibernetika ezt a hasonlóságot tette kutatási programmá. Nem abból indult ki, hogy az élőlény egyszerű gép volna, hanem abból, hogy eltérő anyagú és felépítésű rendszerekben az irányításnak és a kommunikációnak lehetnek közös szerkezetei. Az összehasonlítás értékét az dönti el, sikerül-e pontos, vizsgálható kapcsolatokat találni.

A háborús feladattól az általános kérdésig

Norbert Wiener matematikus a második világháború idején a légvédelmi tűzvezetéshez kapcsolódó előrejelzési problémákon is dolgozott. A mozgó repülőgép jelenlegi helye nem elegendő a célzáshoz: a lövedék útjához szükséges idő miatt a jövőbeli helyzetét kell megbecsülni. A mérés zajos, a cél változtathat a mozgásán, és a beavatkozás hatása késik. [67–68]

E feladatban összeért a matematikai előrejelzés, a mérés és a cselekvés. Wiener és munkatársai számára a kérdés túlmutatott egyetlen haditechnikai berendezésen. Hogyan használja fel egy rendszer a múltbeli jeleket a következő lépés meghatározásához? Hogyan különíti el a hasznos változást a zajtól? Hogyan vezethet a visszajelzés célszerű vagy éppen ingadozó működéshez?



36. kép Norbert Wiener archív felvételen. A táblán látható matematikai jelölések nem e fejezet szabályozási példájának képletei. Fotó: MIT Archives / Wikimedia Commons. A Commons-forrás CC0-jelöléssel közli a felvételt.

Arturo Rosenblueth fiziológussal és Julian Bigelow mérnökkel közösen Wiener 1943-ban jelentette meg a Behavior, Purpose and Teleology című tanulmányt. A szerzők a viselkedést a rendszer és környezete kapcsolatán keresztül vizsgálták. A célra irányuló működés, a visszacsatolás és az előrejelzés fogalmaival élő és mesterséges rendszereket próbáltak közös keretben leírni. [67]

Az eljárás egyik ereje éppen a belső felépítéstől való részleges elvonatkoztatás volt. Ha két rendszer eltérő módon épül fel, attól még megvizsgálhatjuk, milyen bemenetekre milyen kimenetekkel

válaszolnak, és miként függ a válasz a korábbi állapotuktól. Ez nem helyettesíti az anatómiai vagy műszaki feltárást, hanem más kérdést tesz fel.

A kormányos tudománya

Wiener 1948-ban megjelent könyvének címe Cybernetics or Control and Communication in the Animal and the Machine volt. A cím egyszerre nevezte meg az irányítást és a kommunikációt, valamint az élő és a gépi rendszerek összehasonlítását. A kibernetika elnevezést Wiener a kormányosra utaló görög szóból vezette le. A hajó irányítása találó példa: a kormányos a tényleges haladást figyelve módosítja a beavatkozást. [68]

A visszacsatolást természetesen nem Wiener találta fel. A mechanikus szabályozók, a villamos rendszerek és az élettani kutatások már korábban foglalkoztak hasonló jelenségekkel. Újdonsága a közös matematikai és fogalmi keret keresése, a különböző területek összekapcsolása és a szemlélet nagy hatású megfogalmazása volt.

A könyvben a zaj, az információ, a visszacsatolás, a számítógép és az idegrendszer kérdései egymás közelébe kerültek. A gép viselkedésének magyarázatában egyre nagyobb szerepet kapott, hogyan jutnak el a jelek egyik részétől a másikig, milyen bizonytalanság kíséri őket, és miként befolyásolják a további történéseket. [68]

Az információ nem azonos az energiával

Egy szelep megnyitásához szükség lehet erőre, de a szelep nyitását meghatározó jel nem maga szolgáltatja a teljes mozgató energiát. A termosztát érzékelője sem fűti fel a lakást. A jel a felhasználás módját befolyásolja. A kibernetikai szemlélet ezért különösen érzékeny arra, hogyan kapcsolódik össze az energiaáramlás és az információt hordozó állapotváltozás.

A kommunikáció itt nem feltétlenül beszélgetés. Egy mechanikai elmozdulás, villamos feszültség vagy idegi impulzus is betölthet jelközvetítő szerepet. A kérdés az, hogy a fogadó rendszer működése szempontjából milyen különbséget hordoz a jel, és mennyire megbízhatóan jut el hozzá. [68–69]

A jel továbbításának pontossága és jelentésének helyessége különböző kérdés. Hibátlanul közölhetünk egy rosszul megválasztott célértéket. Egy szobahőmérsékletet jól mérő rendszer sem dönti el önmagában, milyen hőmérsékletet szeretnének az ott lakók. Az irányítás technikai sikere ezért nem egyenlő a cél emberi szempontú megfelelőségével.

A Macy-konferenciák közös műhelye

A kibernetika nem egyetlen könyv magányos hatásából nőtt ki. Az 1946 és 1953 között rendezett Macy-konferenciasorozat matematikusokat, mérnököket, idegrendszerrel foglalkozó kutatókat és társadalomtudósokat hozott össze. A résztvevők számára éppen az jelentett feladatot, hogy a saját területükön megszokott kifejezéseket közösen értelmezhetővé tegyék. [68, 70]

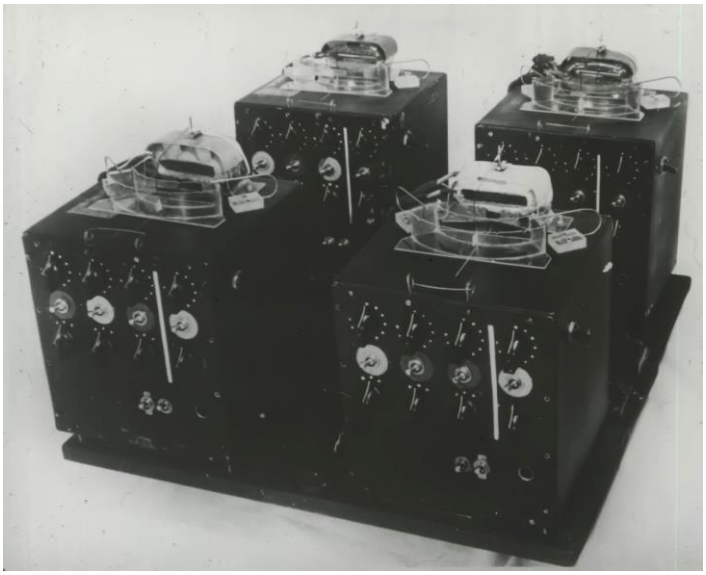
Egy idegkutató és egy híradástechnikai mérnök nem feltétlenül ugyanazt érti jel, emlékezet vagy kód alatt. A közös munka ezért vitákat is jelentett. Melyik hasonlóság elég pontos ahhoz, hogy számolni

lehessen vele? Mikor takar ugyanaz a szó eltérő jelenséget? A későbbi ülések jegyzőkönyvei a kész válaszok mellett ezt a fogalmi keresést is megőrzik. [70]

A kutatói körben Warren McCulloch és Walter Pitts idegi hálózatokról szóló munkája, Neumann automataelméleti érdeklődése, a kommunikáció matematikája és a biológiai szabályozás vizsgálata találkozhatott. A kapcsolatokból nem egységes, mindenre alkalmazható elmélet született, hanem egymást ösztönző problémák és módszerek hálózata.

Ashby és az alkalmazkodás kérdése

W. Ross Ashby brit pszichiáter az állapotok, a stabilitás és az alkalmazkodás vizsgálatával járult hozzá a kibernetikához. Homeostat nevű berendezésének fejlesztése az 1940-es évek második felére esett. Több, egymással összekapcsolt egység viselkedésén mutatta be, hogyan kereshet egy rendszer működőképes beállítást megváltozott feltételek között. [71]



37. kép W. Ross Ashby Homeostatjának négy összekapcsolt egysége. A berendezés az állapotok stabilizálásának és a beállítások megváltoztatásának vizsgálatára készült az 1940-es évek második felében. Fotó: Mick Ashby, az Ashby-hagyaték képviseletében / Wikimedia Commons, CC BY-SA 4.0.

Az alapgondolatot egyszerűsítve két szinten érthetjük meg. Az egyik a pillanatnyi működés: a rendszer adott kapcsolatok mellett változik. A másik a beállítások megváltoztatása akkor, ha bizonyos fontos változók kilépnek az elfogadható tartományból. Az új beállítások kipróbálása nem feltétlenül tudatos tervezést vagy a helyzet teljes belső modelljét kívánja. [71–72]

A különbség a termosztáthoz képest lényeges. A közönséges termosztát rögzített szabály szerint korrigál. Egy alkalmazkodó rendszer a későbbi viselkedést meghatározó kapcsolatot vagy paramétert is módosíthatja. Ashby munkája ennek matematikai és kísérleti megértését kereste. A Homeostat ettől még nem lett általános emberi elme, de az önszabályozás és az alkalmazkodás kapcsolatát kézzelfoghatóvá tette.

Közös nyelv és túlzott általánosítás

A kibernetika vonzereje az volt, hogy távoli területeken is felismert hasonló működési mintákat. Ugyanilyen fontos azonban a különbségek megtartása. Egy társadalomban a célok vitathatók, az emberek értelmezhetik és megváltoztathatják a szabályokat, a mérés pedig maga is befolyásolhatja a viselkedést. Egy gőzgép fordulatszámának és egy társadalmi mutatónak a szabályozása ezért nem válik azonossá attól, hogy mindkettőről rajzolható visszacsatolási ábra.

Az sem igaz, hogy minden kibernetikai rendszer digitális számítógép. A szabályozás lehet mechanikus, villamos vagy biológiai; a jelek és állapotok folytonosan is változhatnak. A kibernetika és a mesterséges intelligencia története átfedi egymást, de egyik sem egyszerűen a másik korábbi elnevezése.

Az AI számára ez a szemlélet új hangsúlyt adott. Az intelligensnek látszó viselkedést nemcsak belső következtetési lépések soraként lehetett elképzelni. Jelentősége lett a testnek, az érzékelésnek, a környezetnek és annak a körnek, amelyben a cselekvés új tapasztalatot hoz létre. Az agy működése így nem csupán hasonlatként, hanem matematikai modellek forrásaként is előtérbe került.

13. Az agy mint modell

Az ember számára a gondolkodás legkézenfekvőbb működő példája az idegrendszer. Mégsem magától értetődő, hogyan kellene átvenni belőle egy gép számára használható elvet. Az agy nem egyszerű alkatrészjegyzék: számít a sejtek kapcsolata, működésük időbelisége és az is, hogyan változik a rendszer a tapasztalat hatására.

A korai matematikai modellek ezért nagyarányú egyszerűsítéssel dolgoztak. Egy részletes biológiai magyarázat helyett bizonyos működési tulajdonságokat emeltek ki. A kérdés az volt, milyen következményeket lehet ezekből levezetni. A modell értékét az adta, amit megmutatott, a határát pedig az, amit szándékosan vagy szükségképpen elhagyott.

A sejtektől a hálózatig

Az idegrendszer felépítésének feltárásában Santiago Ramón y Cajal rajzai és megfigyelései fontos szerepet játszottak. Az elágazó sejtek és kapcsolataik világa jóval gazdagabb volt annál, mint amit egy egyszerű villamos kapcsoló képe kifejezhetett. A kapcsolat nem puszta vezeték: a jelátadás feltételei befolyásolják, hogyan hat az egyik sejt a másikra. [73]



38. kép Santiago Ramón y Cajal rajza egy macska kisagykérgének Purkinje-sejtjéről. A jelölések: a – axon, b – oldalág, c – dendritek. A részletes biológiai szerkezetet a későbbi matematikai modellek erősen egyszerűsítették. Wikimedia Commons. Közkincs.

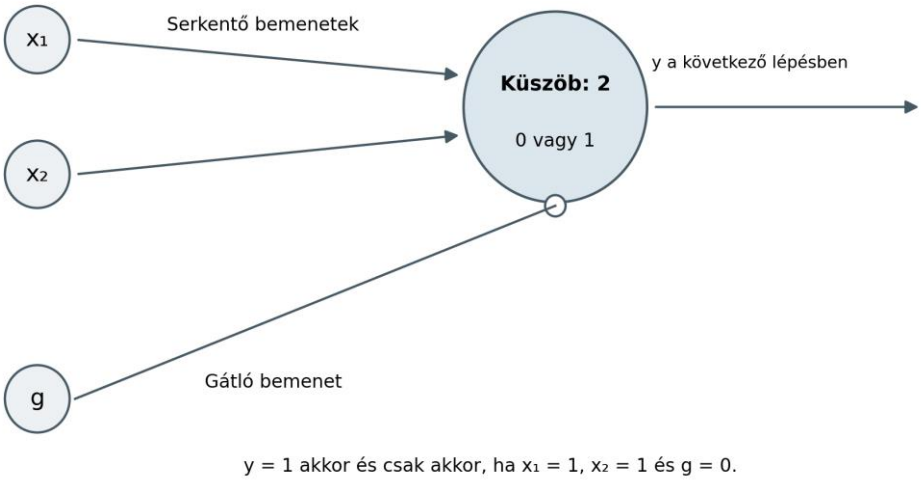
Egy idegsejt sok más sejttől kaphat hatást. Ezek egy része elősegítheti, más része gátolhatja az aktivitását. A tovaterjedő akciós potenciál bekövetkezése bizonyos szempontból eseményként kezelhető: az adott időszakban létrejött vagy nem jött létre. Ez adhat alapot egy kétértékű modellhez. A biológiai sejt teljes működése azonban nem szűkíthető két állapotra.

A sejtek között ráadásul eltérések vannak, a jelek időzítése számít, és a dendritek sem egyszerű, minden tekintetben azonos bemeneti csatlakozók. A korai modelleket ezért nem érdemes utólag teljes agyleírásként olvasni. Azt kell megvizsgálni, milyen feltevések mellett milyen működést tudtak megmagyarázni.

McCulloch és Pitts logikai neuronja

Warren McCulloch és Walter Pitts 1943-ban megjelent tanulmánya azt mutatta meg, hogyan vizsgálható idealizált idegi hálózatok működése a logika eszközeivel. McCulloch idegrendszeri kérdésekkel foglalkozó kutató, Pitts pedig matematikai és logikai problémákban jártas munkatárs volt. Közös eredményük a biológiai működés bizonyos vonásait formális számítási leírással kapcsolta össze. [74]

A modell alapegysége adott időlépésben aktív vagy inaktív. Bizonyos számú egyidejű serkentő bemenet elérésekor a következő lépésben aktiválódhat. Az eredeti egyszerűsített feltevések között a gátló bemenetnek különösen erős szerepe van: jelenléte megakadályozhatja az egység aktiválódását. A hálózat kapcsolatai a vizsgált alapmodellben rögzítettek. [74]



39. kép McCulloch és Pitts egyszerűsített feltevéseit követő mai szemléltető neuron. Két aktív serkentő bemenet és a gátlás hiánya kell a következő időlépés aktivitásához. A kis üres kör gátló kapcsolatot jelöl. A hálózatban nincs feltüntetett tanulási szabály. Saját szerkesztés, nem az 1943-as tanulmány eredeti ábrája.

A szemléltető példában két serkentő bemenetet használunk. Ha a küszöb kettő, mindkettő aktivitása kell a kimenethez. Ha a küszöb egy, már az egyik aktív bemenet is elég. Az első az „és”, a második a befogadó „vagy” logikai kapcsolatát valósítja meg, a modell időlépésnyi késését figyelembe véve.

x_1	x_2	ÉS Küszöb: 2	VAGY Küszöb: 1
0	0	0	0
0	1	0	1
1	0	0	1
1	1	1	1

Két serkentő bemenet és két különböző küszöb. A kimeneti értékek a következő időlépésre vonatkoznak; gátlást nem alkalmazunk, vagyis g = 0. Saját szerkesztés.

A gátló bemenetet úgy képzelhetjük el, mint egy külön tiltó feltételt. Hiába aktív mindkét serkentő bemenet, az ábrán megadott gátlás esetén a kimenet inaktív marad. Ez az eredeti modell sajátos

egyszerűsítése; nem azt jelenti, hogy minden valódi idegsejtben bármely gátló hatás mindig ilyen feltétlen tiltásként működik.

Sok egyszerű elem együtt

Az egyedi egység viselkedése szerény. Hálózatba kapcsolva azonban az egyik egység kimenete a másik bemenete lehet. Így összetett logikai kapcsolatok építhetők fel. Ha visszatérő kapcsolatok is vannak, a korábbi aktivitás befolyásolhatja a későbbi állapotot. A szerkezet ezért nemcsak pillanatnyi bemenetekhez rendelhet választ, hanem az időbeli előzmények hatását is hordozhatja. [74]

Egy egyszerű példa szerint egy esemény jele egy belső állapotot bekapcsol, amely újabb jelzésig fennmarad. A rendszer későbbi válasza már attól is függhet, hogy ez az esemény korábban megtörtént-e. A működési állapot ilyen megőrzése azonban még nem azonos a tanulással. Ha a kapcsolatok és a váltási szabályok nem változnak, a rendszer ugyanazt az előre kialakított feladatot végzi.

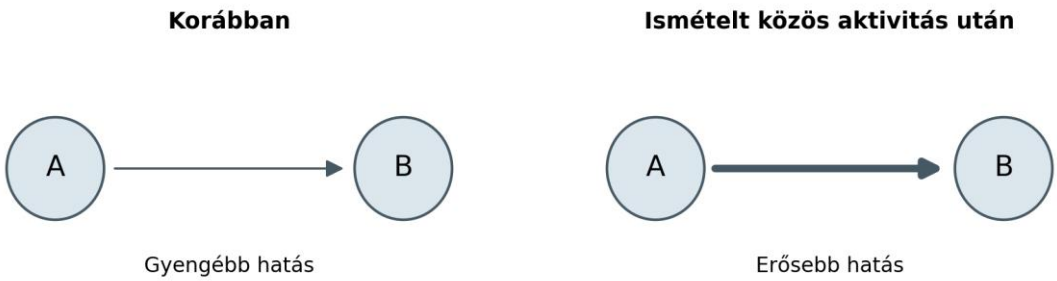
McCulloch és Pitts munkája így hidat teremtett a nyolcadik fejezet logikai műveletei és az idegi hálózatok gondolata között. Egy formálisan leírt következtetési kapcsolat elvben több egyszerű egység kölcsönhatásaként is megjelenhet. Nem szükséges minden összetett műveletet egyetlen különleges elemre bízni.

Az eredményt nem szabad azzal az állítással azonosítani, hogy a szerzők megfejtették az emberi gondolkodást. Saját tanulmányuk is megkülönböztette a formális egyenértékűséget a tényleges élettani magyarázattól. A matematikailag hasonló viselkedés nem bizonyítja, hogy a biológiai rendszer pontosan az idealizált hálózat szerint épül fel. [74]

Hebb: a tapasztalat a kapcsolatokat változtatja meg

Donald O. Hebb 1949-ben megjelent The Organization of Behavior című könyve más kérdést állított előtérbe. Hogyan hagyhat tartós nyomot az idegrendszerben a tapasztalat? A tanulás magyarázatához nem elegendő a pillanatnyi aktivitás; olyan változás kell, amely a későbbi működést is módosítja. [75]

Hebb feltevésének magját úgy foglalhatjuk össze, hogy ha egy sejt aktivitása ismételten részt vesz egy másik sejt aktiválásában, hatása tartósan erősebbé válhat. A hangsúly a kapcsolat hatékonyságának megváltozásán van. A későbbi, gyakran idézett „együtt tüzelnek, összekapcsolódnak” jelszó rövidít, és nem Hebb eredeti mondatának szó szerinti idézete. [75]



40. kép A kapcsolat hatékonyságának tartós növekedése, sematikus ábrán. Hebb feltevésében az A sejt ismételten részt vesz B aktiválásában. A vastagabb nyíl erősebb hatást jelöl, nem új axon szükségszerű létrejöttét. Saját szerkesztés; a rajz nem a teljes biológiai mechanizmus modellje.

A gondolat mai tanítási változatában legyen x a bemeneti, y a kimeneti egység aktivitása, w pedig a köztük levő kapcsolat erőssége. Egy egyszerű, Hebb ihlette szabály: $\Delta w = \eta xy$. A Δw az erősség változása, az η pedig egy kis pozitív tanulási tényező. Ha x és y egyaránt aktív, a kapcsolat erősödik. Ha valamelyik nulla, e legegyszerűbb szabály nem módosítja. Ez korszerű szemléltető képlet, nem az 1949-es könyvben szereplő teljes biológiai mechanizmus.

A különbség McCulloch és Pitts alapmodelljéhez képest világos. Ott a rögzített hálózat működését vizsgáltuk. Itt a tapasztalat a későbbi működést meghatározó kapcsolatot változtatja meg. Egy bemenet, amely korábban kevés volt a válaszhoz, egy később erősebbé vált kapcsolat révén már nagyobb hatást fejthet ki.

Sejtegyüttesek és részleges felidézés

Hebb elképzelése nem állt meg két sejt kapcsolatánál. A tapasztalat összekapcsolódó sejtegyüttesek kialakulásához vezethet. Egy ilyen együttes egy részének aktiválása a többi rész működését is elősegítheti. Az észlelés és a felidézés így elosztott kapcsolatokra támaszkodhat, nem feltétlenül egyetlen elkülönített tárolóhelyre. [75]

Magyarázó hasonlatként gondoljunk egy ismert dallam első néhány hangjára. A részlet nyomán eszünkbe juthat a folytatás. A példa nem bizonyítja Hebb elméletét, de érthetővé teszi a kérdést: hogyan indíthat el egy töredékes inger egy korábban együtt kialakult nagyobb aktivitási mintát?

Ehhez az időbeli sorrend is hozzátartozik. Egy dallam nem az összes hang egyszerre. A sejtegyüttesek egymást követő aktiválódásának gondolata ezért a viselkedés időbeli szerveződését is megcélozta. Hebb pszichológiai jelenségek és idegrendszeri folyamatok között keresett kapcsolatot; nem egyetlen rövid képletből vezetett le minden tanulást. [75]

Amit az egyszerű tanulási szabály elhagy

A $\Delta w = \eta xy$ szabály korlátlan alkalmazása önmagában problémás. Ha két egység gyakran együtt aktív, a kapcsolat újra és újra erősödhet. A növekedés megállításához, a kapcsolatok versenyéhez vagy az aktivitás szabályozásához további feltételek kellenek. Az egyszerű társításból még nem következik, hogy a hálózat hasznos, kiegyensúlyozott tudást szerez.

Az együttjárás ráadásul nem azonos az oksági kapcsolattal. Két jel együtt változhat azért is, mert egy harmadik folyamat mindkettőt befolyásolja. Egy kapcsolat erősödése tehát nem jelenti, hogy a rendszer megértette, mi miért történik. A tanulási elv hasznos mintázatokat rögzíthet, de félrevezető társításokat is megerősíthet.

A későbbi kísérletek a szinaptikus változások időzítésének jelentőségét is megmutatták. Bi és Poo 1998-as, tenyésztett idegsejteken végzett vizsgálatában a bemeneti és a kimeneti aktivitás pontos időrendje befolyásolta a kapcsolat módosulását. A biológiai kép ezért gazdagabb annál, hogy bármely két, közel egyidejű aktivitás egyszerűen és mindig erősít. Ezt az eredményt itt későbbi kitekintésként említjük, nem Hebb eredeti kísérleti bizonyítékaként. [76]

Három különböző változás

A szabályozás során a rendszer a pillanatnyi eltérésre reagál. A belső állapot megőrzése révén a korábbi esemény is befolyásolhatja a választ. A tanulás pedig a későbbi válaszokat meghatározó kapcsolatokat vagy paramétereket változtatja meg. A három összekapcsolódhat, de nem helyettesíti egymást.

E különbségek nélkül könnyű minden alkalmazkodónak látszó gépet tanulónak nevezni, vagy minden hálózatot az agy másolatának tekinteni. A korai munkák jelentőségét pontosabban mutatja, hogy külön kezelhető kérdéseket fogalmaztak meg: milyen működés építhető fel egyszerű egységekből, hogyan hordozhat a rendszer előzményt, és miként változhat a kapcsolatrendszere a tapasztalat hatására?

A mesterséges intelligencia későbbi történetében ezek az utak újra és újra találkoznak. Lesznek rendszerek, amelyek előre megadott jelképes szabályokkal következtetnek, és lesznek, amelyek példák alapján alakítják a viselkedésüket. Az 1940-es évek modelljei még nem a mai ideghálózatok voltak, de megmutatták: a gépi működés leírásához a logika, a visszacsatolás és a változó kapcsolatok egyaránt fontos kiindulópontot adhatnak.

Források és további olvasmányok

A hivatkozások és a képek számozása az 1–10. fejezetét folytatja. A mai példák és magyarázó ábrák a történeti fogalmak megértését szolgálják. A rendszer korrekcióját, belső állapotát és tanulását a fejezetek megkülönböztetik.

[64] James Clerk Maxwell: On Governors. Proceedings of the Royal Society of London, 16, 1868, 270–283. DOI: 10.1098/rspl.1867.0055. A szabályozók dinamikája és a csillapodó vagy növekvő eltérések matematikai vizsgálata. Elsődleges tanulmány · Történeti szövegátirat

[65] Karl Johan Åström – Richard M. Murray: Feedback Systems An Introduction for Scientists and Engineers. Princeton University Press, 2008. Korszerű műszaki háttér a visszacsatoláshoz, a stabilitáshoz, a késleltetéshez és a szabályozás kompromisszumaihoz. A szoba-, zuhany- és érzékelőpéldák, valamint a válaszgörbék saját szemléltetések, nem a könyv ábráinak másolatai. A szerzők elektronikus kiadása · Szerzői egyetemi bibliográfia

[66] Science Museum Group: Rotative Steam Engine by Boulton and Watt, 1788, a Lap engine tárgyismertetője. A centrifugálszabályozó gőzgépi alkalmazásának történeti példája. A fejezetben közölt szerkezeti rajz külön, 1913-as kiadványból származik. Műtárgykatalógus

[67] Arturo Rosenblueth – Norbert Wiener – Julian Bigelow: Behavior, Purpose and Teleology. Philosophy of Science, 10/1, 1943, 18–24. DOI: 10.1086/286788. A viselkedés, a célra irányulás, a visszacsatolás és az előrejelzés közös elemzése. Elsődleges tanulmány

[68] Norbert Wiener: Cybernetics or Control and Communication in the Animal and the Machine. Első kiadás: 1948; bővített második kiadás: 1961. A kibernetika elnevezése, az irányítás és kommunikáció kapcsolata, valamint a háborús előrejelzési munka. Az MIT Press nyílt hozzáférésű újrakiadása a második kiadáson alapul; ezért annak későbbi kiegészítéseit nem kezeljük automatikusan 1948-as szövegként. Kiadói adatlap és nyílt hozzáférés · Az 1948-as könyv bemutatása eredeti bevezetőrészlettel

[69] W. Ross Ashby: An Introduction to Cybernetics. London, Chapman & Hall, 1956. Különösen a mechanizmusokról, visszacsatolásról, stabilitásról, információról és szabályozásról szóló részek. A hivatkozott digitális változat az 1957-es második lenyomatot közli; a mű első megjelenése 1956. Elsődleges könyv digitális kiadása

[70] Cybernetics The Macy Conferences 1946–1953 The Complete Transactions. Claus Pias szerkesztésében. A történeti ülések és a fennmaradt, közölt jegyzőkönyvek gyűjteménye; új kiadás: Diaphanes, 2025. A kötet tartalomjegyzéke az 1949–1953-as ülések témáit részletesen felsorolja. A korai találkozókról Wiener bevezetője is beszámol. A forráskiadás és részletes tartalomjegyzék

[71] The W. Ross Ashby Digital Archive: naplók, családi dokumentumok és a Homeostat története. Jill Ashby életrajza Ashby saját feljegyzéseire és a családi hagyatékra támaszkodik. A fejlesztés az 1940-es évek második felében zajlott; az 1948-as publikáció nem azonos valamennyi részmunka kezdetével. A hagyaték archívuma · Családi és önéletrajzi dokumentumokra épülő ismertető · A Homeostat naplólapjainak tárgymutatója

[72] W. Ross Ashby: Design for a Brain The Origin of Adaptive Behaviour. London, Chapman & Hall, 1952; második kiadás: 1960. A stabilitás, az alkalmazkodás és az ultrastabilitás elmélete. A Homeostat első közleménye: Design for a Brain, Electronic Engineering, 20, 1948, 379–383. Az archív bibliográfia elkülöníti a cikket és a könyvet. Elsődleges művek bibliográfiája

[73] Santiago Ramón y Cajal: The Structure and Connexions of Neurons. Nobel-előadás, 1906. december 12. Az idegsejtek szerkezeti kapcsolatainak történeti vizsgálata. Az itt felhasznált Purkinje-sejtrajz külön történeti ábra; nem állítjuk, hogy az a hivatkozott előadás adott képtáblája. Elsődleges előadás

[74] Warren S. McCulloch – Walter Pitts: A Logical Calculus of the Ideas Immanent in Nervous Activity. Bulletin of Mathematical Biophysics, 5, 1943, 115–133. DOI: 10.1007/BF02478259. A kétértékű aktivitás, a küszöb, a gátlás és a hálózatok logikai elemzése. A szerzők maguk is elkülönítik a formális egyenértékűséget az élettani magyarázattól. A fejezet ábrája és táblázata mai szemléltetés. Elsődleges tanulmány · Kiadói rekord

[75] Donald O. Hebb: The Organization of Behavior A Neuropsychological Theory. New York, Wiley, 1949. Különösen a 62. oldal tanulási feltevése, a sejtegyüttesek és az időbeli szerveződés tárgyalása. A magyar ismertetés parafrázis. A Δw = ηxy képlet mai, egyszerű Hebb ihlette oktatási szabály; nem a történeti könyv teljes modelljének szó szerinti reprodukciója. A mű bibliográfiája és hozzáférési adatai · Hebb eredeti előszava és bibliográfia

[76] Guo-qiang Bi – Mu-ming Poo: Synaptic Modifications in Cultured Hippocampal Neurons Dependence on Spike Timing, Synaptic Strength, and Postsynaptic Cell Type. Journal of Neuroscience, 18/24, 1998, 10464–10472. DOI: 10.1523/JNEUROSCI.18-24-10464.1998. Elsődleges kísérleti beszámoló az időzítéstől függő erősödésről és gyengülésről. A fejezet ezt későbbi kitekintésként használja. Teljes elsődleges tanulmány

Képforrások és felhasználási adatok

33. kép Saját szerkesztésű szabályozási kör a szövegben megadott szerepekkel.

34. kép Eredeti képfájl és részletes forrásadatok · Közkincs.

35. kép Saját matematikai példagörbék: 1 − exp(−0,55t), illetve 1 − exp(−0,35t)[cos(1,7t) + (0,35/1,7)sin(1,7t)]. A célérték 1; mindkét görbe stabil közelítést szemléltet.

36. kép Eredeti képfájl és részletes forrásadatok · MIT Archives · Commons szerinti CC0

37. kép Eredeti képfájl és részletes forrásadatok · Mick Ashby, az Ashby-hagyaték képviseletében · CC BY-SA 4.0

38. kép Eredeti képfájl és részletes forrásadatok · Közkincs.

39. kép Saját szerkesztésű küszöbmodell az [74] forrás megjelölt egyszerűsítései alapján.

40. kép Saját szerkesztésű szemléltetés Hebb [75] alatt ismertetett feltevéséhez.

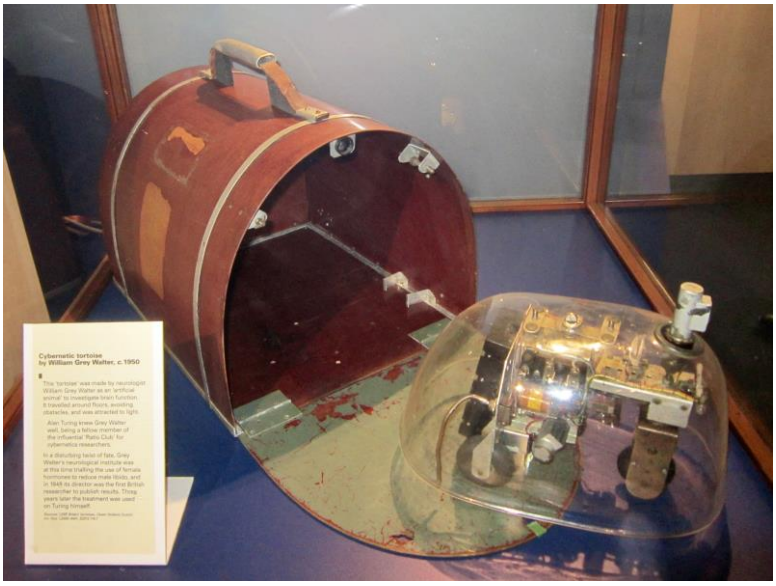
III. rész – Visszacsatolás, kibernetika és az első önálló gépek

A visszacsatolás és az idegi modellek után most olyan szerkezetek következnek, amelyek már mozgásukkal is próbára tették az elképzeléseket. A bristoli teknősök, a gyári robotkarok és a szegedi katicabogár különböző kérdésekre adtak választ. Együtt mutatják meg, hogyan lett az élőlények utánzásából kísérleti módszer, ipari eszköz és a nagyközönség számára is érthető tudományos történet.

14. A teknősök, amelyek keresték a fényt

Egy alacsony, kerekded szerkezet lassan gurul a padlón. Elfordul, megközelít egy lámpát, akadályba ütközik, majd más irányból folytatja útját. A szemlélő könnyen állatot lát benne: mintha keresne valamit, mintha habozna, mintha a saját útját választaná. A mozgás mögött azonban megvizsgálható érzékelők, áramkörök és motorok dolgoznak. Éppen ez tette izgalmassá Grey Walter teknőseit. Az életre emlékeztető viselkedéshez nem kellett az élő szervezet teljes bonyolultságát beépíteni.

William Grey Walter a bristoli Burden Neurological Institute idegrendszeri kutatója volt. Elmer és Elsie nevű első teknőseit 1948–1949-ben készítette. A Machina speculatrix néven ismert gépek fényre és érintkezésre reagáltak; mozgásuk a két érzékelési mód és a motoros működés kölcsönhatásából alakult ki. Nem egy előre felvett útvonalat játszottak le. A következő fordulatot az is meghatározta, mivel találkoztak éppen. [77–78]



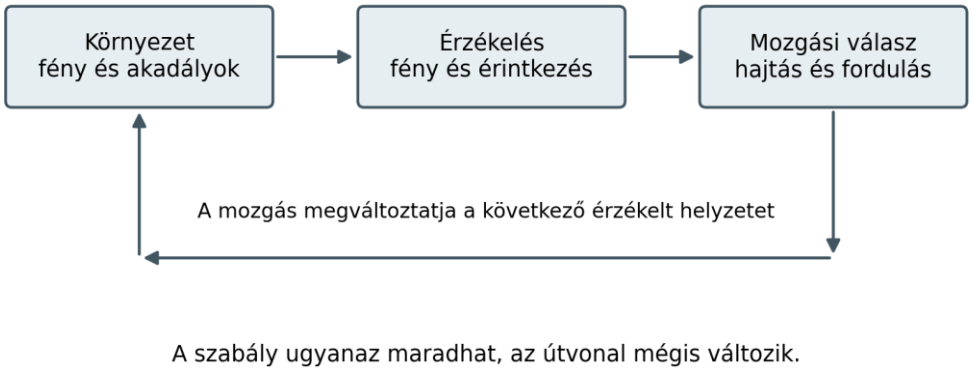
41. kép Grey Walter teknőseinek működési elvét bemutató múzeumi példány a londoni Science Museumban, 2012-es felvételen. Nem az eredeti Elmer vagy Elsie azonosítására szolgál. A fényérzékelő és a kerekeken mozgó test összekapcsolta az érzékelést a haladással. Fotó: Ank Kumar / Wikimedia Commons, CC BY-SA 4.0.

Két érzékelési mód és egy mozgó test

A teknősöket a fényérzékeny elem a megvilágításról tájékoztatta. A burkolat érintkezésérzékelőként is szolgált: ha a gép valaminek nekiment, a kontaktus megváltoztatta a működését. Az elektronikus szabályozás összekapcsolta az érzékelést a hajtással és a kormányzással. Walter a kevés alkatrésszel

felépített rendszeren azt vizsgálta, mekkora viselkedési változatosságot teremthet az összeköttetések módja. [77–78]

A kísérlet megértéséhez képzeljünk el egy mai, leegyszerűsített fénykeresőt. Ha nem talál megfelelő jelet, körbefordulva pásztázza a környezetét. Amikor fényt észlel, módosítja a haladási irányát. Ha akadályhoz ér, az érintkezési jel átmenetileg más mozgást indít. Mire újra a fény alapján mozog, már más helyen és más szögben áll. Ugyanazok a szabályok most más cselekvést eredményeznek. Ez a példa a működés elvét magyarázza; nem Walter valamennyi áramköri állapotának tételes rekonstrukciója.



42. kép A fénykereső működés egyszerűsített köre. A mozgás megváltoztatja azt a helyzetet, amelyből a gép a következő jelet kapja. Saját szerkesztésű magyarázó ábra, nem Walter eredeti kapcsolási rajza.

Az útvonal így részben a szobában születik meg. Számít a lámpa helye, egy asztalláb, a fal, a kerék tapadása és az indulási helyzet. A gép belsejében nincs szükség minden lehetséges kerülőút külön leírására ahhoz, hogy eltérő kerülőutakat járjon be. A környezet nem egyszerű díszlet: a működési folyamat része.

Miért látszik többnek néhány szabály?

Az ember általában a viselkedésből következtet a szándékra. Ha valaki egy lámpához tart, azt mondjuk, keresi a fényt. A teknősnél ez jó rövid leírás lehet, de nem magyarázza meg a mechanizmust. A tudományos kérdés az, milyen érzékelési és mozgási kapcsolatok hozzák létre a látott útvonalat. A magyarázat annál erősebb, minél pontosabban előre jelzi, mi változik meg egy érzékelő kikapcsolásakor vagy a lámpa áthelyezésekor.

Itt válik külön az összetett viselkedés és az összetett belső terv. Egy sor kanyar és megállás úgy is kialakulhat, hogy a gép nem tervezte meg előre ezt a sorozatot. A visszacsatolás minden pillanatban új helyzetből folytatja a működést. A kevés szabályból keletkező gazdag viselkedést gyakran emergensnek nevezik. A kifejezés nem titokzatos többleterőt jelent: az alkatrészek kölcsönhatásaiból és a környezettel való kapcsolatból kialakuló mintázatra utal.

Ebből nem következik, hogy bármilyen egyszerű szerkezet tetszőlegesen intelligenssé válhat. A teknős viselkedési lehetőségeit erősen korlátozza, mit érzékelhet és hogyan mozoghat. A fény iránya kevés segítséget ad, ha a cél átlátszó fal mögött van. Egy helyi akadálykerülő szabály pedig

kedvezőtlen elrendezésben ismétlődő mozgásba zárhatja a gépet. Az ügyesnek tűnő bemutató mellett a kudarcok is hozzátartoznak a modell értékeléséhez.

A reagálás még nem feltétlenül tanulás

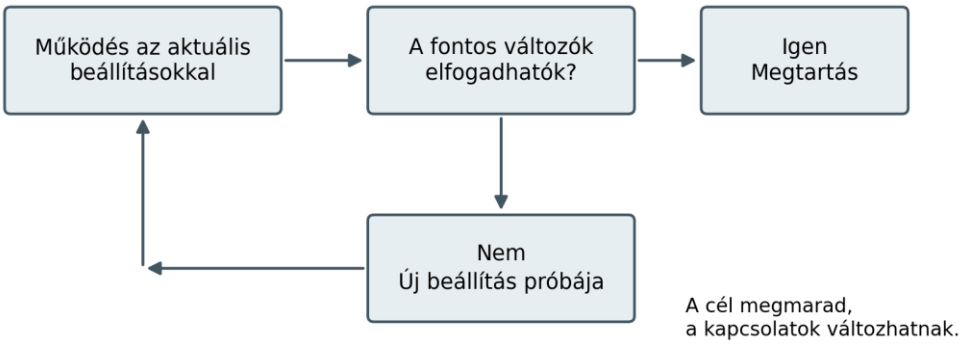
A fénykereső működést és a tapasztalat hatására megváltozó működést külön kell vizsgálnunk. Ha az érzékelési és mozgási összeköttetések változatlanok, a gép eltérő útvonala önmagában nem bizonyít tanulást. Lehetséges, hogy csak más akadályokkal találkozott. Tanulásra akkor következtethetünk megalapozottabban, ha az előző események a későbbi válasz feltételeit is módosítják.

Walter a tanulás irányában is továbblépett. A Scientific American 1951-ben közölte A Machine that Learns című írását a Machina docilisról, a korábbi fénykereső gép tanulási kísérletekhez kialakított leszármazottjáról. Ezért pontatlan volna az első teknősök minden fényre adott válaszát már tanult viselkedésnek nevezni, vagy a későbbi tanuló változat képességeit automatikusan rájuk vetíteni. [79]

Ashby gépe a működőképes beállítást keresi

A tizenkettedik fejezetben megismert W. Ross Ashby Homeostatja más oldalról közelített ugyanahhoz a nagy kérdéshez. Nem a padlón gurult célpontok után. Egymásra ható egységeivel azt modellezte, miként maradhatnak bizonyos változók elfogadható tartományban, és mi történhet, ha a korábbi kapcsolatok mellett ez már nem sikerül. [80]

A különbséget egy gondolatkísérlet teszi világossá. Egy szabályozó az adott beállításai szerint igyekszik csökkenteni az eltérést. Ha tartósan kudarcot vall, egy második folyamat megváltoztatja a beállítások valamelyikét. A rendszer ezután az új kapcsolatokkal próbál működni. Amikor megfelelő állapot alakul ki, a változtatás megállhat. Itt nem a célpont helyét kell megtalálni, hanem olyan belső elrendezést, amellyel a rendszer ismét elfogadhatóan viselkedik.



43. kép Az Ashby-féle alkalmazkodási elképzelés két szintjének szemléltetése. Az aktuális működés mellett a beállítások megváltoztatására is sor kerülhet. A megtartott beállításokkal a működés és az ellenőrzés folytatódik. Saját szerkesztésű fogalmi ábra.

Ashby ultrastabilitásnak nevezett elképzelésében ezért kétféle változás kapcsolódott össze: a pillanatnyi állapotok alakulása és a működés feltételeinek átrendeződése. A Homeostatban a határértékek túllépése új kapcsolati beállítások keresését válthatta ki. A megfelelő egyensúly

megtalálásához nem kellett a berendezésnek előre kiszámítania minden lehetőség következményét. A próbálkozások eredménye visszahatott arra, hogy szükséges-e tovább keresni. [80]

Ennek a módszernek is megvannak a korlátai. Ha a választható beállítások között nincs megfelelő, a keresés nem teremti meg azt a semmiből. Ha egyetlen használható elrendezés nagyon sok lehetőség között rejtőzik, megtalálása hosszú ideig tarthat. Az alkalmazkodás elve és az alkalmazkodás gyorsasága két külön kérdés.

Mit tanítottak a műállatok?

Walter és Ashby gépeit érdemes kísérleti állításokként nézni. Walter azt tette kézzelfoghatóvá, hogy kevés érzékelési és motoros kapcsolat is változatos helyzetekhez igazodó mozgást hozhat létre. Ashby a belső beállítások megváltoztatásának szerepét vizsgálta. A két megközelítés együtt tágította a gépi viselkedésről alkotott képet: nem minden érdemleges működés hosszú, központilag leírt utasítássor végrehajtása.

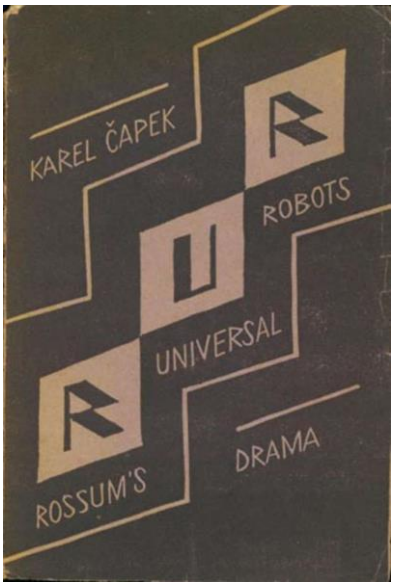
A teknősöknek nem kellett emberi arc, kéz vagy beszéd. Az állati életből is csak néhány kapcsolatot vettek át. Ez a szerénység volt a kísérlet ereje: a megfigyelt mozgást össze lehetett vetni egy átlátható műszaki szerkezettel. A kutató így azt is megkérdezhette, a néző mikor tulajdonít a gépnek többet annál, mint amit a vizsgálat bizonyított.

15. A robot szó és az ipari robot megjelenése

A robot előbb kapott nevet, mint ahogy a gyárban munkába állt. A szóhoz már kezdettől társadalmi kérdés kapcsolódott: ki dolgozik, ki rendelkezik a munkája fölött, és mi történik, ha a megalkotott segítő nem marad engedelmes? A műszaki fejlődés később egészen más alakot adott a robotnak, de ezeket a kérdéseket magával vitte.

Čapek mesterséges munkásai

Karel Čapek R.U.R. című drámája 1920-ban jelent meg könyvként, színpadi története 1921-ben indult. A címben szereplő vállalat neve Rossum's Universal Robots. A robot elnevezést az író saját, 1933-as visszaemlékezése szerint bátyja, Josef Čapek javasolta. A cseh robota szó a kényszermunka, a jobbágyi robot jelentésköréhez kapcsolódik; magyarul ezért a szó történeti háttere különösen könnyen érthető. [81–82]



44. kép Karel Čapek R.U.R. című drámája 1920-as első kiadásának borítója. A robot szó az irodalomból került át a műszaki nyelvbe. Forrás: Wikimedia Commons; a forrás közkincsként közli a borítót.

Čapek robotjai nem mai értelemben vett fémvázas, villanymotoros gépek. Mesterséges anyagból előállított, élő emberutánzatok, akiket tömeges munkavégzésre gyártanak. A dráma ezért a mesterséges élet történetéhez is hozzátartozik. A robotok az emberi munka helyére lépnek, majd az emberiség ellen fordulnak. [81]

Az első rész mítoszaihoz visszatérve ismerős a motívum: az alkotó elveszíti az uralmat teremtménye fölött. Čapeknél azonban a műhely helyére gyár, az egyetlen különleges teremtmény helyére tömegtermelés kerül. A kérdés nem csupán az, lehet-e embert alkotni. Az is kérdés, milyen társadalmat teremt az ember, ha másokat pusztán olcsó és engedelmes munkaerőként kezel. Ez a darab egyik lehetséges olvasata; nem egy későbbi ipari berendezés működési leírása.

Az irodalom ebben az értelemben előbb rendezte meg a vitát, mint a technika a helyzetet. A robot szó egyszerre jelölt segítőt és fenyegetést. A névhez tapadó emberkép azután akkor is velünk maradt, amikor a valóságos robotból eltűnt az emberi alak.

Asimov és a szabályokba foglalt biztonság

Isaac Asimov robotnovelláiban a veszély gyakran a gondosan megfogalmazott szabályok következményeiből bontakozik ki. A robotika három törvényének együttes, kifejezett megfogalmazása az 1942 márciusában közölt Runaround című elbeszélésben jelent meg. [83]

A törvények értelmét röviden összefoglalva: elsőbbséget élvez az ember sérülésének megakadályozása; ezt követi az emberi utasítások teljesítése, ha az nem ütközik az első követelménybe; végül a robot saját fennmaradásának védelme, amennyiben az összeegyeztethető az előző kettővel. Ez tartalmi ismertetés, nem a novellaszöveg szó szerinti idézete. [83]

A gondolat ereje a sorrendben is rejlik. Az utasítás teljesítése nem feltétlenül a legfőbb érték. Egy gép kaphat egymással összeütköző követelményeket, és előre tisztázni kellene, melyiknek van elsőbbsége. Asimov történetei e szabályok értelmezési nehézségeit tették elbeszélhetővé: a nyelv világosnak látszik, a helyzet azonban többféleképpen értelmezhető.

Mérnöki szempontból a nehézség már az első lépésnél jelentkezik. Hogyan ismeri fel a gép az embert? Hogyan becsüli meg, melyik mozdulat okoz sérülést? Hogyan veszi figyelembe egy parancs későbbi következményét? Az általános erkölcsi mondat önmagában egyik érzékelési vagy számítási feladatot sem oldja meg. A három törvény hatásos irodalmi keret, a valóságos gép biztonsága viszont megvalósított érzékelést, korlátozásokat és ellenőrizhető működést kíván.

Amikor a robotból kar lett

Az ipari robot történetében George Devol és Joseph Engelberger együttműködése döntő szerepet játszott. Devol programozható tárgymozgató szerkezetre irányuló szabadalmi bejelentése 1954-es; a szabadalmat 1961-ben adták meg. Engelberger a fejlesztés és az üzleti megvalósítás megszervezésével segített ipari termékké formálni az elképzelést. Az Asimov történetei iránti érdeklődése az iparági visszaemlékezésekben is megjelenik. [84]

A széles körben mérföldkőként számon tartott Unimate 1961-ben állt munkába a General Motors New Jersey-i üzemében: nyomásos öntőgépből vett ki forró munkadarabokat. A feladat veszélyes és ismétlődő volt. Nem emberi beszélgetőtársat kellett létrehozni hozzá, hanem erős, megbízható, előírt sorrendben mozgó manipulátort. A mozdulatsor parancsait mágnesdob tárolta. [85]



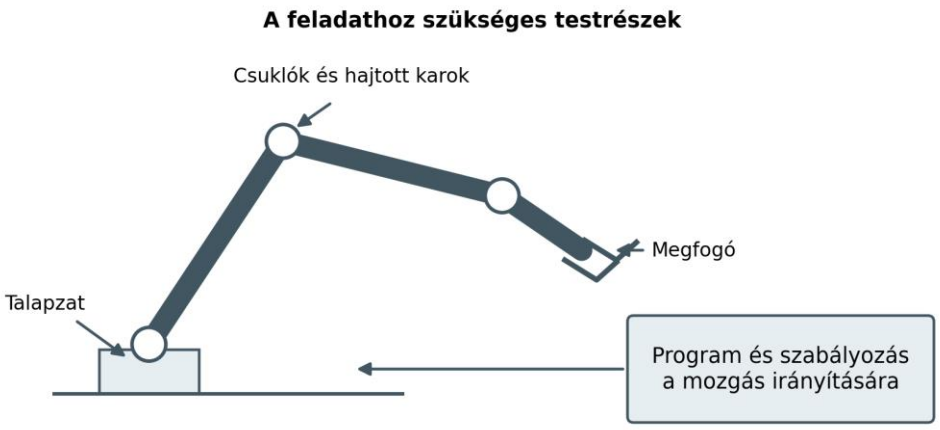
45. kép Unimate robot kávét tölt Betty Myrahnak egy 1967-es bemutatón. A kép a robotkar sokoldalúságának nyilvános demonstrációját mutatja; nem az 1961-es gyári munkába állás felvétele. Fotó: Frank Q. Brown, Los Angeles Times / UCLA Library / Wikimedia Commons, CC BY 4.0.

A műszaki fordulatot nem az jelentette, hogy a gyárba ember formájú gép érkezett. A robotból elég volt annyi, amennyi a feladathoz kellett: kar, megfogó, hajtás és vezérlés. A munkadarab felemeléséhez az emberi test legtöbb tulajdonsága fölösleges. A biztos megfogás, a megfelelő pálya és az ismételhető mozgás viszont elengedhetetlen.

A programozhatóság gyakorlati értelme

Egy meghatározott termékhez épített automata nagyon hatékony lehet, de átalakítása komoly szerelést igényelhet. A programozható robotkar a mozdulatok sorrendjének és célhelyzeteinek megváltoztatásával új feladathoz igazítható. Ez nem jelenti azt, hogy bármilyen munkát azonnal elvégezhet. Megfogóra, megfelelő munkatérre, anyagellátásra és a folyamat többi berendezésével kialakított kapcsolatra is szüksége van.

Képzeljünk el egy egyszerű anyagmozgatási ciklust. A kar a kijelölt helyre fordul, megfogja az alkatrészt, felemeli, a másik állomásra viszi, majd elengedi. A program lépései mellett a mozgást végrehajtó hajtások szabályozása is dolgozik. A kívánt és a tényleges helyzet összevetése más szintű feladat, mint annak meghatározása, hogy most a felvétel vagy a lerakás következik. A programozás és a visszacsatolás ezért együtt is jelen lehet anélkül, hogy a gép a feladatot emberi értelemben megértené.



46. kép Egy csuklós robotkar fő funkcionális részei. A rajz általános mai szemléltetés, nem az első Unimate szerkezeti rekonstrukciója. A program a feladat sorrendjét, a szabályozás a mozgás végrehajtását irányíthatja. Saját szerkesztés.

A beállítások rögzítését a robotikában tanításnak is nevezhetik. Ez a szó félrevezető lehet, ha minden esetben tapasztalatból történő gépi tanulást értünk rajta. Amikor a kezelő megadja a kívánt pontokat és a gép ezeket tárolja, a program tartalmát elsősorban az ember állítja össze. A gép önálló végrehajtása és a végrehajtási szabály önálló megtanulása külön képesség.

Miért nehéz humanoid robotot építeni

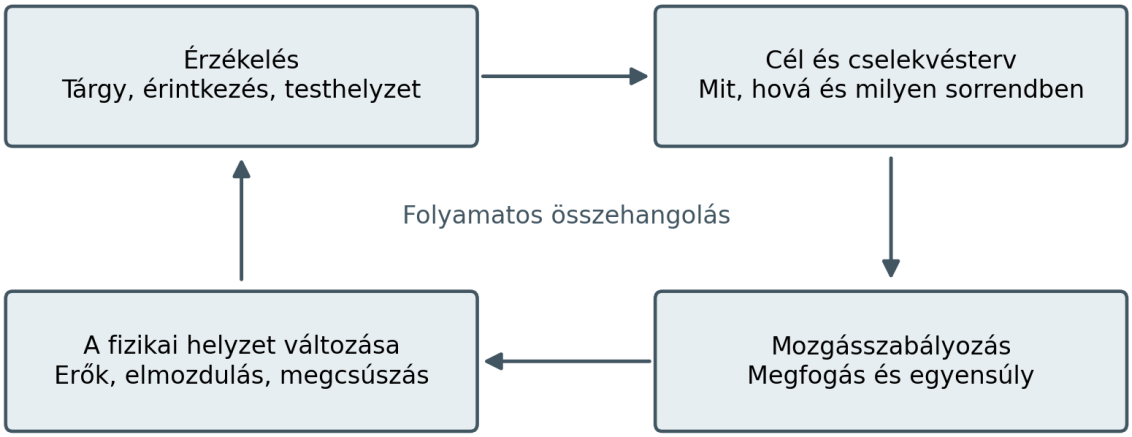
Az ipari robot sikeréhez elegendő volt egy pontosan meghatározott munkát gépesíteni. Az ember alakú, hétköznapi környezetben segítő robothoz ennél sokkal több képességet kell összehangolni. A történet későbbi fejleményeire előrepillantva érdemes feltenni egy egyszerű kérdést: miért olyan nehéz géppel megismételni azt, amit egy négyéves gyerek különösebb erőfeszítés nélkül megtesz?

Képzeljünk el egy gyereket, akit arra kérnek, hogy hozza oda a játékát. Feláll, kikerüli a széket, lehajol, megfogja a tárgyat, majd odaviszi. Ha közben valami útban van, gyakran változtat a tervén vagy segítséget kér. Ez szemléltető élethelyzet, nem minden négyévesre érvényes teljesítményteszt. A gyerek közben nemcsak a kérést értelmezi: érzékeli a környezetét és a testét, egyensúlyoz, és a cselekvése következményeihez igazítja a következő mozdulatát.

A robotika egyik régi felismerése, a Moravec-paradoxon arra figyelmeztet, hogy az embernek nehéznek tűnő szellemi feladat és a gép számára nehéz feladat nem feltétlenül ugyanaz. A szabályos játékban elért magas teljesítmény mellett sokáig küszködtek az érzékeléssel és a mindennapi mozgással. A számolást tudatos erőfeszítésnek érezzük, a járás és a megfogás rengeteg részletét viszont észre sem vesszük. Ettől az utóbbi még nem egyszerű. [15/A]

A test maga is része a feladatnak. Egy két lábon álló robotnak mozdulat közben is fenn kell tartania az egyensúlyát. Amikor előrenyúl, a karjával együtt a test tömegeloszlása is változik. A törzsnek és a lábaknak ehhez igazodniuk kell. A mai humanoidkutatás ezért a teljes test összehangolt vezérlését és a környezettel létrejövő érintkezéseket is vizsgálja. [15/D] A nagyobb motor nagyobb erőt adhat, de tömeget és energiaigényt is jelent; a nagyobb akkumulátort pedig szintén mozgatni kell. A mechanikai döntések visszahatnak a szabályozásra.

A látás nem pótolja a tapintást. Egy tárgy felismerése még nem mondja meg biztosan, milyen erővel foghatjuk meg. Egy tojásnál a túl nagy szorítás törést, a túl kicsi kicsúszást okozhat. Az ujjaknak a megcsúszás kezdetére is reagálniuk kell, miközben a kéz részben eltakarhatja a tárgyat a kamera elől. A tapintás és az erők érzékelése ezért fontos a megbízható tárgykezelésben. [15/C] A robotkéz építésénél az ujjak mozgathatósága, az érzékelők, a hajtások és a vezérlés együtt szabják meg, mire képes a rendszer.



46/A. kép Egy hétköznapi tárgymozgatás mögött egymásra ható feladatok dolgoznak. A kar mozgása a test egyensúlyát is befolyásolja, az érintkezés pedig új érzékelt jeleket ad. Saját fogalmi ábra, nem egy konkrét robot felépítése.

A gyerek már sok saját tapasztalatot hoz. Négyéves korára éveken át nyúlt tárgyakért, ejtett le dolgokat, próbált felállni, egyensúlyozni és másokat utánozni. Nem kész szabálykönyvből tanulta meg a világot. Mozgás közben új tapasztalatokat szerzett, az új képességek pedig további felfedezést tettek lehetővé. A fejlődésben a test állapota, a környezet és a más emberekkel való kapcsolat is részt vesz. [15/B] Ehhez evolúciósan kialakult testi és idegrendszeri adottságok társulnak: a gyerek tanulása sem teljesen üres lapról indul.

A robotok is tanulhatnak példákból, emberi bemutatásból és próbálkozásokból. A fizikai tapasztalat gyűjtése azonban időt, működő gépet és megfelelő kísérleti körülményeket kíván. Egy hibás mozdulat tárgyat törhet vagy alkatrészt rongálhat. Szimulációban sok próbát végezhetünk, de a valóságos súrlódás, anyagrugalmasság és érzékelési zaj eltérhet a modelltől. Ezért külön feladat annak elérése, hogy a megtanult vezérlés a valós roboton is működjön. [15/D]

A beszédkészség és a mozgásügyesség külön tanulandó képesség. Egy nyelvi rendszer helyesen elmagyarázhatja, hogyan kell felvenni egy poharat, miközben a hozzá kapcsolt robot még bizonytalanul hajtja végre a mozdulatot. A 2023-as RT-2 kutatás például megmutatta, hogy képi és nyelvi előtanításból származó tudás átvihető robotvezérlésre. A szerzők ugyanakkor jelezték: a vizsgált modell fizikai készségeit továbbra is a robotos tanítóadatokban szereplő mozgások köre korlátozta. [15/E] Ez egy adott rendszer dokumentált eredménye és korlátja, nem minden jövőbeli robot képességének határa.

A hétköznapi használathoz tartós megbízhatóság kell. Egyetlen sikeres bemutató kevés. Saját számpéldánkban, ha húsz egymást követő művelet mindegyike 99 százalékos valószínűséggel sikerül, és a sikerek egymástól függetlenek, a teljes sorozat sikerének esélye körülbelül 82 százalék. A valós hibák nem feltétlenül függetlenek; a példa a sok lépésből álló feladat érzékenységét mutatja. A gépnek a kudarcot is fel kell ismernie, majd javítania, biztonságosan megállnia vagy segítséget kérnie.

A négyéves gyerek előnye ezért sok hétköznapi helyzetben a különböző képességek rugalmas összekapcsolása. Ez nem általános rangsor: egy célgépként épített robot erőben, pontosságban vagy ismételhetőségben messze felülmúlhatja. A humanoid célja viszont éppen a sokféleség kezelése egy emberekhez igazított környezetben. Ehhez a hasonló külső kevés. Az érzékelést, a tanulást és a test megbízható irányítását kell működő egésszé formálni.

Robot és mesterséges intelligencia

A robot fizikai környezetben avatkozik be, de ehhez nem feltétlenül használ tanuló vagy következtető AI-t. Egy jól meghatározott munkaciklus végrehajtása lehet hagyományos automatizálás. Ugyanakkor egy képfelismerő program önmagában nem robotkar: attól, hogy azonosít egy alkatrészt, még nem emeli fel. A két terület akkor kapcsolódik össze, amikor az érzékelésből származó értelmezés a fizikai cselekvést is alakítja.

Az ipari robot tehát egy régi ábránd gyakorlati átalakulását mutatja. Az irodalmi emberutánzat helyett meghatározott munkát végző gép jelent meg. A korlátozott feladat nem csökkenti történeti jelentőségét: éppen a feladat pontos körülhatárolása tette lehetővé, hogy a robot a bemutatóteremből az üzemi termelésbe kerüljön. Az ember és a gép viszonyáról szóló vita pedig a színpad után a gyárban folytatódott.

16. A magyar kibernetika úttörői

A kibernetika magyarországi történetét nem lehet egyetlen találmány vagy intézet köré rendezni. Mérnöki elképzelések, matematikai kutatások, számítógép-építés, egyetemi műhelyek és ismeretterjesztő könyvek kapcsolódtak össze. Nemes Tihamér, Tarján Rezső, Muszka Dániel és Csató István különböző feladatot vállalt ebben a folyamatban. Egyikük gépeket tervezett, másikuk kutatást szervezett, megint másikuk a nagyközönség számára tette érthetővé az új fogalmakat.

Nemes Tihamér és az emberi működés műszaki modelljei

Nemes Tihamér (1895–1960) villamosmérnökként a híradástechnika és a televíziózás korai magyar kutatói közé tartozott. A látás műszaki utánzása, a logikai gépek és a gondolkodási folyamatok modellezése egyaránt foglalkoztatta. Sakkal és sakkfeladványok megoldásával kapcsolatos gépi elképzelései is ehhez az érdeklődéshez tartoztak. [86]

E munkák megítélésénél fontos megkülönböztetni az elvi tervet, a részmegoldást és a teljesen működő berendezést. Egy sakkozógép szerkezetére tett javaslat nem azonos egy minden szabályos állásban játszó, elkészült géppel. Nemes jelentőségét az mutatja, hogy konkrét műszaki és logikai kérdésekké alakította az emberi képességek utánzásának gondolatát.

A televízió és a logikai gép első pillantásra távoli terület. Mégis közös kérdést vetnek fel: hogyan alakítható egy fizikai jel feldolgozható információvá, és milyen műveletek végezhetők rajta? A látás utánzása nem merül ki a szem alakjának másolásában. Érzékelni, továbbítani és elkülöníteni kell a jeleket. A következtetés gépesítése hasonlóképpen a műveletek pontos leírását igényli.

Nemes munkásságának összefoglaló kötete, a Kibernetikai gépek 1962-ben, halála után jelent meg. Fennmaradt munkáinak rendezésével olyan tudás vált hozzáférhetővé, amely addig részben különálló elképzelésekben és tanulmányokban élt. A könyv a magyar kibernetika egyik fontos korai dokumentuma. [86]

Tarján Rezső és a kutatás megszervezése

Tarján Rezső (1908–1978) matematikusként és tudományszervezőként kapcsolta össze az új elméleteket a hazai intézményi lehetőségekkel. Pályáját a korszak politikai üldözése is megtörte: 1953-ban hamis vádakkal bebörtönözték, és 1955 elején szabadult. Az MTA Kibernetikai Kutatócsoportjának tudományos igazgatóhelyettese lett, később pedig a Neumann János Számítógép-tudományi Társaság első elnökeként is dolgozott. [87]

Az intézményszervezés kevésbé látványos, mint egy mozgó műállat, mégis nélkülözhetetlen. Szakembereket kell egy feladat köré gyűjteni, eszközöket szerezni, a matematikai elképzeléseket áramkörökre és programokra fordítani. Egy új tudományterülethez közös szókincs és tanulási lehetőség is kell. Tarján munkájának ebben az összekapcsolásban volt különös súlya.

Az 1958-ban megjelent Gondolkodó gépek című könyve is jelezte, hogy az új gépekről magyarul kell beszélni. A címben felvetett kérdés messzebb vezetett az összeadásnál: mit nevezhetünk gépi

gondolkodásnak, milyen kapcsolat van a logika és a műszaki működés között? Tarján a nemzetközi eredmények hazai megismertetésében is fontos szerepet vállalt. [87]

Az M–3 és a működő számítástechnikai műhely

Az M–3 az első Magyarországon megépített elektronikus számítógép volt. Szovjet dokumentáció alapján készült, amelyet a hazai alkatrészbázishoz igazítottak, és később saját fejlesztésekkel egészítettek ki. Az MTA Kibernetikai Kutatócsoportját Varga Sándor vezette; Tarján az igazgatóhelyettesi feladatokat látta el. Az építés a programozók, mérnökök és technikusok közös munkája volt, Dömölki Bálint és más munkatársak részvételével. [88]

Az 1959. január 21-i hivatalos átadás fontos dátum, de az üzembiztos működés kialakítása hosszabb folyamatot jelentett. A géphez programokat és matematikai eljárásokat is készítettek; az első hazai szubrutinkollekció e munka kézzelfogható emléke. A dokumentumokat Varga és Tarján mellett az adott feladatokért felelős munkatársak hitelesítették. [88]



47. kép Az M–3 számítógép archív képe a Neumann Társaság Informatikatörténeti Fórumának adattárából. A szekrényekbe épített berendezés jóval nagyobb volt annál, mint amit ma a számítógép szó rendszerint felidéz. Forrás: NJSZT iTF, M–3 számítógép című műtárgyadatlap; a felvétel készítője nincs feltüntetve.

A számítógép megépítése azért tartozik a kibernetika történetéhez, mert az akkori kutatási keretben együtt jelent meg a számítás, a logika és a gépi működés vizsgálata. Az M–3 azonban önmagában nem tanuló mesterséges intelligencia volt. Programozható számítógépként olyan műszaki és emberi alapot teremtett, amelyre később sokféle kutatás épülhetett.

Egy számítógépes műhely tudása a kapcsolási rajznál szélesebb. Tudni kell, melyik hiba miből ismerhető fel, miként ellenőrizhető egy eredmény, és hogyan lehet egy matematikai eljárást a gép korlátai között megvalósítani. A dokumentáció átvétele ezért nem tette fölöslegessé a hazai mérnöki munkát. A működés megszervezése közben szakmai közösség és tapasztalat keletkezett.

Az M–3 1965-ben Szegedre került, ahol 1968-ig működött. Ez a budapesti és a szegedi történetet személyesen is összekötötte: Muszka Dániel a szegedi gépek telepítésében és üzemeltetésében szintén szerepet vállalt. [88–89]

Muszka Dániel és a szegedi iskola

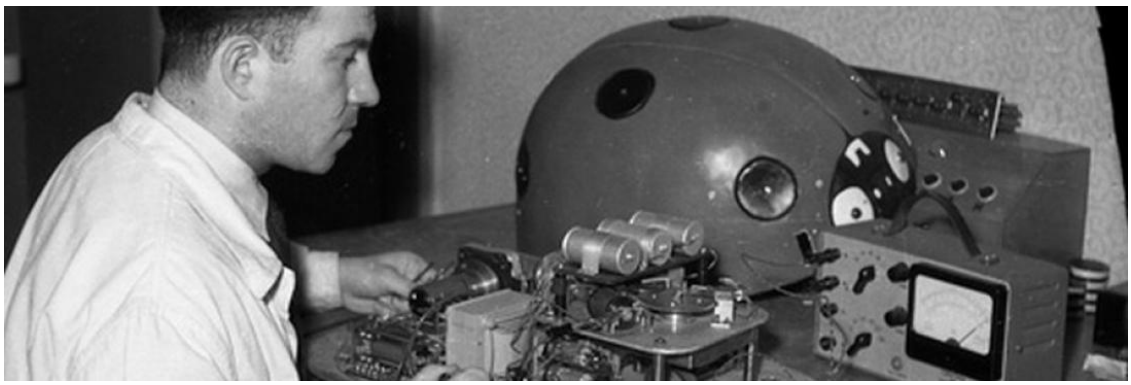
Szegeden Kalmár László matematikus körül alakult ki jelentős kutatói műhely. Muszka Dániel (1930–2018), matematika–fizika szakos végzettségű alkotó, az elvi tervek megépítésében is otthonosan

mozgott. Kalmár tervei alapján 1957–1958-ban elkészítette a Szegedi Logikai Gépet. Saját legismertebb szerkezete az 1956–1957-ben kialakított katicabogár lett, amely Király József pszichológus ötletéből kiindulva a feltételes reflex modellezését szolgálta. [89–90]

A logikai gép és a katicabogár egymás mellett két utat mutat. Az egyikben logikai műveletek pontos gépi végrehajtása, a másikban az inger és a válasz kapcsolatának megváltozása kerül előtérbe. A korszak kutatói számára e kérdések nem külön világokhoz tartoztak. Azt vizsgálták, milyen működés hozható létre ismert és ellenőrizhető szerkezetekkel.

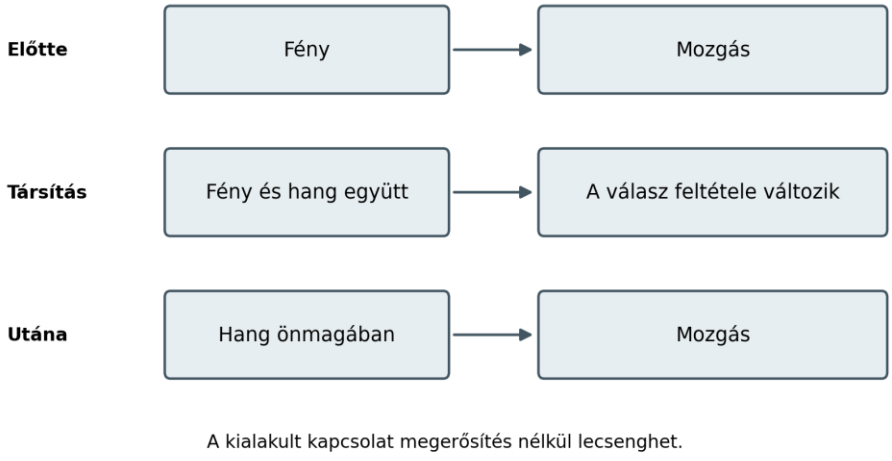
Mit tudott a katicabogár?

A hétpettyes műállat fényre, hangra és érintésre reagált. A megvilágított érzékelők a mozgását irányították, a hanginger kezdetben a szemeknél elhelyezett lámpák felvillanását válthatta ki. Ha a hangot a fény követésével társították, egy idő után a hang önmagában is elindíthatta a mozgást. A kapcsolat később lecsenghetett: a bemutatott működéshez felejtési szakasz is tartozott. [90]



48. kép Muszka Dániel a szegedi katicabogár szerkezetével. Az archív felvételen a burkolat mellett az áramköri és hajtási részek is láthatók. Forrás: Neumann Társaság, A jövő múltja – Szegedi Katicabogár című ismertetőjének nyitóképe; a felvétel készítője nincs feltüntetve.

Ez egy meghatározott tanulási jelenség műszaki modellje. Nem arról van szó, hogy a gép megértette a sípszó jelentését vagy fogalmat alkotott a gazdájáról. A korábbi ingerpárosítás hatására megváltozott, mire reagált mozgással. A különbség azonban így is lényeges a puszta fénykövetéshez képest: a későbbi válasz már az előzménytől is függött.



49. kép A katicabogár ingerpárosítási bemutatójának elve. A fény és a hang társítása után a hang önmagában is mozgást válthat ki; a kialakult kapcsolat később lecsenghet. A rajz az inger és a válasz viszonyát mutatja, nem az áramkört. Saját szerkesztés a múzeumi ismertetők alapján.

A bogár érintésre adott válasza különösen könnyen emberi értelmezést váltott ki. A pettyek megnyomásakor vagy ütközéskor leállt, fényjelzése megváltozott, és hangot adott; a hátán lévő érintkező működtetésével ismét reakcióképes állapotba kerülhetett. A simogatásról, megsértődésről vagy kibékülésről szóló leírás szemléletes, de az érzelmi jelentést a néző társítja a műszaki állapotváltáshoz. [90]

Az eredeti, elektroncsöves katicabogár vezetéken keresztül hálózati táplálást kapott. A későbbi másolatok akkumulátorral működtek. Ezt azért érdemes külön rögzíteni, mert egy mai bemutató látványa alapján könnyű volna az eredeti gépnek is mindenben ugyanilyen önállóságot tulajdonítani. [90]

A szerkezet tudományos és bemutatói értéke összetartozott. A látogató egyetlen rövid kísérletben láthatta az ingert, a választ és a kettő kapcsolatának módosulását. A kedves külső közel hozta a kérdést, a működés ismertetése pedig lehetőséget adott arra, hogy a néző a látszat mögé is belásson.

Csató István és az új tudomány magyar olvasója

Csató István (1916–1974) újságíró és ismeretterjesztő író volt. A magyar kibernetika történetében azért érdemel önálló helyet, mert az új tudomány gondolatait szélesebb olvasóközönséghez juttatta el. A kutatás eredményeinek társadalmi útja nem ér véget a laboratórium ajtajában: valakinek meg kell magyaráznia, mit mutatnak a gépek, és miért érdekesek. [91]

A robotember 1960-ban jelent meg a Móra Kiadónál. Tartalomjegyzéke a hír és a kód kérdéseitől a visszacsatoláson és a számítógépen át az idegrendszerig, a műállatokig és a robotemberig vezet. A könyv e témákat összefüggő történetté rendezte az olvasó számára. Nem csak egyetlen gépet mutatott be, hanem olyan fogalmi kapcsolatokat, amelyek a korszak kibernetikai szemléletét jellemezték. [92]

Az Okos gépek 1964-ben, a Búvár könyvek sorozatban látott napvilágot. A fennmaradt tartalomjegyzék régi automatákat, mesterséges élőlényeket, író és olvasó gépeket, valamint az

automatizáláshoz kapcsolódó emberi viszonyt is megjeleníti. Az olvasó így a látványos szerkezet felől juthatott el a működési elvekhez. [93]

A kibernetika – Az információ forradalma 1969-ben jelent meg, 1971-ben pedig második, javított kiadása is készült. A kiadások megkülönböztetése fontos: egy későbbi példány adatait nem célszerű automatikusan az első megjelenéshez kötni. A kötetben a történeti előzmények, az információ és a számítógép kérdései is helyet kaptak. [94]

A kibernetika és az ember 1970-es kötete a szemléletet az ember életműködésére és tágabb tevékenységeire alkalmazta. Fennmaradt tartalomjegyzékében az élő szervezet, a társadalom, a nyelv, a játék és a művészet egyaránt megjelenik. A témaválasztás jól mutatja, milyen széles érvényű magyarázatot reméltek akkoriban a kibernetikától. [95]

Kötet	Kiadási adat	Fő témakör
A robotember	1960 · Móra	Hír, visszacsatolás, számítógépek, műállatok
Okos gépek	1964 · Móra Búvár könyvek	Automaták és a gépi információfeldolgozás
A kibernetika Az információ forradalma	1969 · Kossuth 1971 · 2. javított kiadás	A kibernetika alapfogalmai és a számítógép
A kibernetika és az ember	1970 · Kossuth	Élő szervezet, társadalom, nyelv, játék és művészet

Csató István négy kötete a kiadási adatok és a fennmaradt tartalomjegyzékek alapján. A témakörök rövid szerkesztői összefoglalások. [92–95]

Csató könyveit ma kétféle érdeklődéssel olvashatjuk. Megismerhetjük belőlük, milyen fogalmak segítették az akkori olvasót a gépek megértésében, és azt is megfigyelhetjük, milyen jövőt vártak a technikától. A történeti érték nem jelenti minden korabeli magyarázat változatlan mai érvényességét. Egy régi könyv jó hasonlata tovább élhet akkor is, ha az adott részlet tudományos magyarázatát időközben pontosították.

Az ismeretterjesztés megőrzés is. A rajzok, fotók, szóhasználatok és kérdésfeltevések azt is dokumentálják, hogyan találkozott egy társadalom az új technikával. A tudomány történetéből ezért hiányozna valami lényeges, ha csak a kutatók és gépépítők nevét tartanánk számon. Az olvasókat megszólító szerzők a műszaki kultúra közvetítői voltak.

Közös örökség

A magyar úttörők történetében a műszaki alkotás és a megértés terjesztése egymást erősítette. Nemes elképzelései, Tarján szervező és elméletközvetítő munkája, az M–3 körül létrejött tapasztalat, Muszka működő modelljei és Csató könyvei együtt alakították azt a közeget, amelyben a gépi gondolkodás kérdése itthon is kézzelfoghatóvá vált.

A katicabogár útja a padlón és egy ismeretterjesztő könyv útja az olvasó kezébe ugyanannak a történetnek két része. Az egyik megmutatja, hogy egy elgondolás működő szerkezetté válhat. A

másik lehetővé teszi, hogy a működést mások is megértsék, megvitassák és továbbgondolják. A mesterséges intelligencia történetéhez mindkettő hozzátartozik.

Források és további olvasmányok

A hivatkozások és a képek számozása az 1–13. fejezetet folytatja. A történeti adatok mellett a mai példák, a fogalmi megkülönböztetések és a magyarázó rajzok szerkesztői szemléltetések. Csató könyveinél a feltüntetett kiadás és a hozzáférhető tartalomjegyzék adatait különítjük el a teljes szöveg elemzésétől.

[77] W. Grey Walter: An Imitation of Life. Scientific American, 182/5, 1950, 42–45. DOI: 10.1038/scientificamerican0550-42. A teknősök korabeli bemutatása; a kiadói oldal a cikk bibliográfiai adatait és bevezetőjét közli. Kiadói adatlap

[78] University of Bristol: Grey Walter and his tortoises, 2008. július 22. A Burden Neurological Institute, az 1948–1949-es első gépek és az érzékelési rendszerek történeti ismertetése. Megkülönbözteti az eredeti teknősöket a későbbi másolatoktól. Egyetemi ismertető

[79] W. Grey Walter: A Machine that Learns. Scientific American, 185/2, 1951. augusztus, 60–63. DOI: 10.1038/scientificamerican0851-60. A kiadói bevezető a Machina docilist a korábbi Machina speculatrix leszármazottjaként azonosítja. Kiadói adatlap és bevezető

[80] W. Ross Ashby: Design for a Brain. Chapman & Hall, 1952; második kiadás: 1960. Uő: An Introduction to Cybernetics, 1956. Az ultrastabilitás és az alkalmazkodás fogalmi háttere. A családi digitális archívum bibliográfiája és életrajza külön dokumentálja a Homeostat fejlesztését. A fejezet példája és ábrája mai egyszerűsítés. Szerzői művek bibliográfiája · Családi archívum és életrajz · Az 1956-os könyv digitális kiadása

[81] Karel Čapek: R.U.R. 1920. Paul Selver és Nigel Playfair angol színpadi változatának digitális szövege: Project Gutenberg, 59112. A mesterséges munkások anyagi természetének és a dráma alaphelyzetének elsődleges irodalmi forrása. A dráma szövege

[82] Karel Čapek: O slově robot, Lidové noviny, 1933. december 24. Angol közlés a Virginia Commonwealth University Blackbird folyóiratának Inventing Roboti oldalán. A robot szó Josef Čapeknek tulajdonított javaslata. A robota jelentéséhez és az 1920–1921-es időrendhez a cseh külképviselet összefoglalója is támpontot ad. Čapek visszaemlékezésének közlése · Cseh intézményi ismertető

[83] Isaac Asimov: Runaround. Astounding Science Fiction, 1942. március; később az I, Robot kötetben is. A törvények közös megjelenésének adataihoz: Isaac Asimov FAQ. A fejezet a törvényeket magyar tartalmi összefoglalásban ismerteti; a műszaki értelmezés szerkesztői magyarázat. Asimov bibliográfiai és műtörténeti FAQ

[84] Association for Advancing Automation: Unimate – The First Industrial Robot. Devol és Engelberger együttműködésének iparági visszatekintése. Az oldal az 1954-es bejelentést és az 1961-es szabadalomengedélyezést elkülöníti. A korai prototípusok üzemi próbáit a múzeumok által 1961-hez kötött mérföldkőtől külön kell kezelni. Iparági történeti ismertető

[85] The Henry Ford: Robot, First Unimate Robot Ever Installed on an Assembly Line, 1961. Computer History Museum: Timeline of Computer History, 1961, UNIMATE. Múzeumi források a General Motorsnál végzett öntvénykezelésről és a mágnesdobon tárolt parancsokról. Az 1961-es robot múzeumi adatlapja · CHM idővonal

[86] Neumann Társaság, Informatikatörténeti Fórum: Nemes Tihamér személyi adatlapja és a munkásságát ismertető dokumentum. A Kibernetikai gépek 1962-es, posztumusz megjelenésének adataihoz is. Az elgondolt és a megépített gépek teljesítményét a fejezet nem azonosítja. Személyi adatlap · Munkásságának összefoglalása

[87] Neumann Társaság, Informatikatörténeti Fórum: Tarján Rezső. Életrajzi, intézménytörténeti adatok, a Gondolkodó gépek 1958-as megjelenése és a társaság alapító elnöki tisztsége. Személyi adatlap

[88] Neumann Társaság, Informatikatörténeti Fórum: M–3 számítógép. Továbbá: Az első magyar szubrutin gyűjtemény és a programozás hőskorának dokumentumai, 2024. március 27. A megépítés, az 1959-es átadás és a későbbi szegedi működés adatai. Az átadási dátum nem jelenti a fejlesztés és a felélesztés valamennyi részfeladatának egyidejű befejezését. Gépadatlap · Korabeli dokumentumok bemutatása

[89] Neumann Társaság, Informatikatörténeti Fórum: Muszka Dániel. A Szegedi Logikai Gép, a katicabogár és a szegedi számítógépes műhely életrajzi vonatkozásai. Személyi adatlap

[90] Neumann Társaság: Szegedi katicabogár, iTF műtárgyadatlap; A jövő múltja múzeumi ismertető. A fénykövetés, az ingerpárosítás, a lecsengés és az érintésre adott válasz leírása. A múzeumi oldal külön rögzíti az eredeti gép hálózati táplálását és a későbbi másolatok akkumulátoros működését. Működési leírás · Múzeumi ismertető

[91] Honismeret, 1998/5, évfordulónaptár: Csató István életrajzi adatai; újságíró és ismeretterjesztő író. A fejezet tudománynépszerűsítő szerepét a következő tételekben felsorolt könyveinek témái alapján tárgyalja. Évfordulónaptár

[92] Csató István: A robotember. Móra Ferenc Ifjúsági Könyvkiadó, Budapest, 1960. A kiadás és a tartalomjegyzék adatai: Antikvarium.hu példányleírás. A fejezet a felsorolt témákra támaszkodik, nem közöl a könyvből szó szerinti szövegrészletet. Kiadási adat és tartalomjegyzék

[93] Csató István: Okos gépek. Móra, Budapest, 1964, Búvár könyvek. Magyar Nemzeti Digitális Archívum: részlet és tartalomjegyzék, a Szendrői Közművelődési Központ és Könyvtár gyűjteményéből. Közgyűjteményi rekord

[94] Csató István: A kibernetika – Az információ forradalma. Kossuth, Budapest, 1969; második, javított kiadás: 1971. Kiadási adatok és tartalomjegyzékek; a második, javított kiadást a II. Rákóczi Ferenc Könyvtár katalógusa is ezzel az évvel írja le. Az 1969-es kiadás · Az 1971-es kiadás

[95] Csató István: A kibernetika és az ember. Kossuth, Budapest, 1970. Magyar Nemzeti Digitális Archívum: könyvrészlet és tartalomjegyzék. A fejezetben szereplő témacsoportok e rekord adatait foglalják össze. Közgyűjteményi rekord

A 15. fejezet kiegészítésének forrásai

[15/A] Hans Moravec: Rise of the Robots. Scientific American, 1999. december, 124–135. A gépi következtetés és az érzékelő-mozgató képességek eltérő fejlődésének történeti értelmezése. A cikk jövőbeli időpontokra vonatkozó jóslatait a kiegészítés nem veszi át. Szerzői archívum

[15/B] Karen E. Adolph – Justine E. Hoch: Motor Development: Embodied, Embedded, Enculturated, and Enabling. Annual Review of Psychology, 70, 2019, 141–164. DOI: 10.1146/annurev-psych-010418-102836. A mozgásfejlődés testi, környezeti és társas összefüggései. Kiadói közlés

[15/C] Jingyi Xiang – Yixuan Wang – Wenzhen Yuan: Tactile-Based Robotic Manipulation: From Mechanics-Driven to Data-Driven Perspectives. Annual Review of Control, Robotics, and Autonomous Systems. Online előközlés: 2026. október 21. DOI: 10.1146/annurev-control-032125-030330. A tapintási információ és a tárgykezelés kapcsolata. Kiadói közlés

[15/D] Chong Zhang – Wenli Xiao – Tairan He – Guanya Shi: WoCoCo: Learning Whole-Body Humanoid Control with Sequential Contacts. Proceedings of the 8th Conference on Robot Learning, PMLR, 270, 2025, 455–472. A teljes test vezérlése, érintkezési szakaszok és a szimulációban tanult működés átültetése fizikai robotra. Tanulmány és nyílt kézirat

[15/E] Anthony Brohan és munkatársai: RT-2: Vision-Language-Action Models Transfer Web Knowledge to Robotic Control. 2023, arXiv:2307.15818. A nyelvi-képi tudás átadása robotvezérlésre; a tanulmány külön jelzi a fizikai készségek robotos tanítóadatokhoz kötődő korlátját. Kutatási kézirat

Képforrások és felhasználási adatok

41. kép Ank Kumar: Cybernetic Tortoise by William Grey Walter c 1950, 2012-es múzeumi felvétel. A teljes kép változtatás nélkül. CC BY-SA 4.0. Kép és metaadatok · Licenc

42. kép Saját szerkesztésű érzékelési és mozgási kör. A [77–78] forrásokban bemutatott elv magyarázata.

43. kép Saját szerkesztésű fogalmi ábra az Ashby-féle ultrastabilitáshoz. Forrás: [80].

44. kép R.U.R., 1920-as borító. A Wikimedia Commons közkincsként közli; a teljes kép változtatás nélkül szerepel. Kép és metaadatok

45. kép Frank Q. Brown, Los Angeles Times: Unimate pouring coffee for a woman at Biltmore Hotel, 1967. október 3. UCLA Library. Változtatás nélkül. CC BY 4.0. Kép és metaadatok · Archívumi forrás · Licenc

46. kép Saját szerkesztésű általános robotkarvázlat. Nem történeti tervrajz; a funkcionális részek felismerését szolgálja.

47. kép Archív M–3-felvétel. NJSZT iTF műtárgyadatlap, 1_M_3.png. A készítő és külön szabad felhasználási licenc az adatlapon nincs megjelölve. Közlés forrásmegjelöléssel, történeti ismertetéshez. Forrásoldal

48. kép Archív felvétel Muszka Dánielről és a katicabogárról. Neumann Társaság, A jövő múltja, danibafejlec.jpg. A forrás szerinti teljes képkivágat; a készítő és külön szabad felhasználási licenc nincs feltüntetve. Közlés forrásmegjelöléssel, történeti ismertetéshez. Forrásoldal

49. kép Saját szerkesztésű szemléltetés a katicabogár ingerpárosítási bemutatójáról. Forrás: [90].

46/A. kép Saját szerkesztésű magyarázó ábra a 15. fejezet kiegészítéséhez. A feladatok kölcsönhatását szemlélteti; nem történeti műszaki rajz vagy teljes robotvezérlési architektúra.

IV. rész – Megszületik a mesterséges intelligencia

A számíthatóság elmélete megmutatta, mit jelent egy feladatot pontos eljárással megoldani. A kibernetika az érzékelés, a cselekvés és a korrekció kapcsolatát vizsgálta. Az 1950-es években ezek mellé egy új kutatási cél került: olyan gépi működést létrehozni, amelyet az embernél az intelligencia megnyilvánulásának tekintünk. Ehhez vizsgálható kérdésekre, közös kutatói programra és működő modellekre volt szükség. Turing írása, a dartmouthi kezdeményezés és a korai programok e fordulat különböző oldalait mutatják.

17. Tudnak-e a gépek gondolkodni?

Ha egy gép hibátlanul összead két hosszú számot, a legtöbben nem mondjuk rá, hogy gondolkodik. Ha viszont visszakérdez, felismer egy ellentmondást a történetünkben, vagy megmagyaráz egy tréfát, könnyebben használjuk ezt a szót. A határ azonban bizonytalan. A gondolkodásnak nincs olyan egyszerű mérőműszere, mint a hőmérsékletnek. Alan Turing éppen ezt a nehézséget fordította át egy megvizsgálható helyzetbe.

Az 1950 októberében, a Mind című folyóiratban megjelent Computing Machinery and Intelligence című tanulmány a gépek gondolkodásának kérdéséből indult ki. Turing a hétköznapi szavak végleges meghatározása helyett másféle kérdezést javasolt. Olyan játékot írt le, amelyben a kérdező az elkülönített résztvevők írásos válaszaiból próbálja megállapítani, kivel beszél. [96]



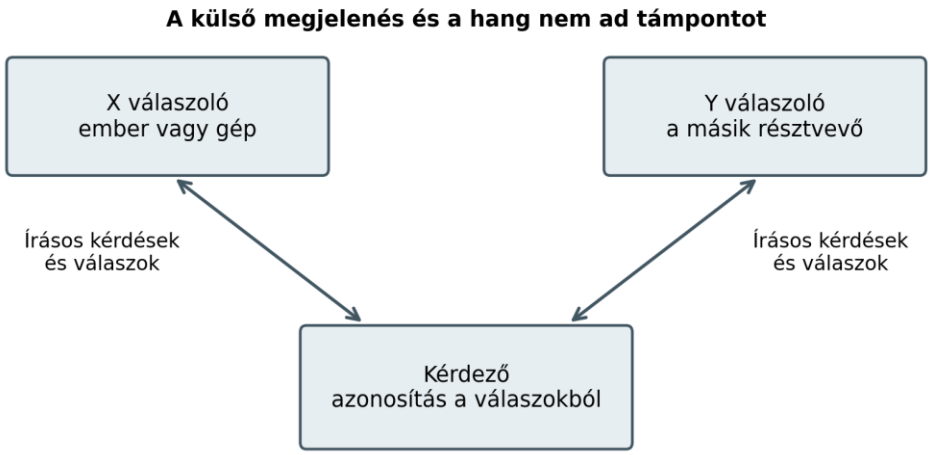
50. kép Alan Turing 1951. március 29-én készült portréja, a Computing Machinery and Intelligence megjelenését követő évben. Fotó: Elliott & Fry / Wikimedia Commons. A forrás közkincsként közli a felvételt.

Az utánzási játék eredeti helyzete

Turing kiinduló játéka három ember szerepére épült: egy férfira, egy nőre és egy kérdezőre. A kérdezőnek a válaszokból kellett megkülönböztetnie a férfit és a nőt; a férfi félrevezethette, a nő segíthette. Turing ezután vetette fel, mi történne, ha a férfi helyére gép kerülne. A később elterjedt, ember és számítógép megkülönböztetésére szolgáló Turing-teszt ennek a gondolatnak az egyszerűsített, közismert változata. [96]

Az írásos kapcsolatnak lényegi szerepe van. A kérdező nem támaszkodhat az arcra, a hangszínre, a testre vagy egy fémburkolat látványára. A gépnek nem kell emberi bőrt viselnie ahhoz, hogy az

értelmi teljesítményét vizsgáljuk. Ez összecseng az előző fejezetek tanulságával: az emberhez hasonló külső és az emberhez mérhető képesség külön kérdés.



51. kép Az ember és a gép megkülönböztetésére szolgáló, közismert teszthelyzet mai vázlata. A kérdező írásban érintkezik az elkülönített válaszolókkal. Az ábra nem az eredeti, nemi szerepekre épülő játék másolata. Saját szerkesztés.

A játékban a válaszok egymást követik. Nem csupán az számít, képes-e a gép egyetlen jól hangzó mondatra. A kérdező pontosíthat, visszatérhet egy korábbi állításhoz, új összefüggésbe helyezheti a feladatot. Egy betanult válasz rövid ideig meggyőző lehet, de a következetes folytatáshoz több kellhet. A beszélgetés előzménye olyan követelményeket teremt, amelyek egy elszigetelt kérdésből nem látszanak.

Mit mér a válasz, és mit tesz hozzá a kérdező?

Képzeljünk el egy saját vizsgálati példát. A kérdező azt írja: két utas ugyanahhoz az ablakhoz ülne a vonaton, miközben csak egy ablak melletti hely szabad. A válaszoló több megoldást javasolhat, majd megindokolhatja, miért tartja az egyiket méltányosnak. Ha a kérdező később elárulja, hogy az egyik utas rosszul lesz a folyosó mellett, a jó válasz ezt az új körülményt is figyelembe veszi. A teljesítmény itt a helyzet követésében és az indokolt módosításban mutatkozik meg.

Egyetlen ilyen párbeszéd még kevés. A válaszoló ismerhette a történetet, a kérdező túl könnyű feladatot adhatott, vagy egy udvarias megfogalmazást mély megértésnek nézhetett. Más kérdező szokatlan helyzetekkel próbálkozna, hosszabb beszélgetést folytatna, vagy szigorúbban ellenőrizné a válaszok összhangját. A vizsgálati eredmény ezért a feladatsortól és a megítélés módjától is függ.

Egy értelmes értékelésnek rögzítenie kell a körülményeket. Mennyi ideig tarthat a beszélgetés? Milyen témák engedélyezettek? Hány ember és hány gép vesz részt? Tudják-e a kérdezők, hogy biztosan találkoznak számítógéppel? Egyetlen téves azonosításból vagy látványos bemutatóból nem lehet általános képességre következtetni. Ezek mai módszertani szempontok, amelyekkel a történeti ötlet vizsgálhatóbbá tehető.

Az emberi jelleg nem azonos a helyes teljesítménnyel

Az utánzás sajátos követelményt állít a gép elé. Egy pillanatok alatt elvégzett, hibátlan fejszámolás árulkodó lehet, mert az átlagos ember lassabban és bizonytalanabbul számol. A gép emberinek látszó szereplése ezért akár a képességei visszafogását is megkívánhatja. A játék eredményét nem szabad automatikusan minden feladatra érvényes intelligenciasorrendként értelmezni.

Fordítva is igaz: egy rendszer lehet nagyon hasznos anélkül, hogy embernek néznénk. Egy útvonaltervezőtől azt várjuk, hogy jó útvonalat adjon, egy számításellenőrző programtól pedig azt, hogy megtalálja a hibát. Nem szükséges hozzájuk hihető gyermekkori emlékeket kitalálni. A gépi intelligencia kutatásában ezért az emberi viselkedés utánzása mellett a feladat tényleges megoldása is önálló mérce.

Turing kérdésfeltevése nem teszi fölöslegessé a belső működés megismerését. Ha két berendezés ugyanazt a választ adja, attól még eltérhet a módszerük, a hibára való érzékenységük és az új helyzetekben mutatott teljesítményük. A kimenet vizsgálata és a mechanizmus feltárása egymást kiegészítő munka. A kettő összekapcsolása segít megkülönböztetni a tartós képességet az ügyesen összeállított bemutatótól.

Viselkedés, megértés és tudat

A korábbi fejezetekben a fénykereső teknős mozgása szándékosnak látszott. A beszélgető gépnél még erősebb lehet ez a benyomás, mert a nyelvhez rendszerint személyt kapcsolunk. Egy jól eltalált visszakérdezés figyelmességnek, egy vigasztaló mondat együttérzésnek tűnhet. A megfigyelt viselkedés valóban adat, de értelmezése további kérdéseket kíván.

Az eredményből közvetlenül azt tudjuk meg, mire volt képes a rendszer az adott helyzetben. Azt, hogy van-e szubjektív élménye, nem látjuk ugyanilyen közvetlenül. Turing a tanulmányában tárgyalta a tudatosságra hivatkozó ellenvetést, de nem követelte meg a tudat teljes magyarázatát a gépi képességek kutatásának megkezdéséhez. [96]

Ez nem bizonyítja sem a gépi tudat lehetőségét, sem annak lehetetlenségét. A kérdéseket érdemes külön tartani. Egy beszélgetési eredmény fontos lehet az intelligens nyelvhasználat megítélésében, miközben a belső élmény kérdésében nem ad végleges választ. A gondos megkülönböztetés megakadályozza, hogy ugyanattól a kísérlettől minden filozófiai probléma eldöntését várjuk.

A gép, amelyet nevelni kell

Az írás utolsó része a tanuló gépekre tért át. Turing felvetette, hogy a kész felnőtt elme közvetlen megalkotása helyett egy egyszerűbb, fejlődésre képes gyermekgép és annak tanítása lehetne a kiindulópont. A kezdeti szerkezet és a nevelési folyamat így együtt határozná meg az eredményt. [96]

A gondolat a gépépítés feladatát is átrendezi. Nem kell feltétlenül minden későbbi választ előre beleírni a programba; meg kell alkotni azt a működést, amelyből tapasztalat és visszajelzés nyomán változó képesség alakulhat ki. A tanulás irányát azonban befolyásolja, milyen példákkal találkozik a gép, mit tekintünk sikernek, és milyen hibát javítunk ki. A tanító döntései a későbbi működésben is jelen maradnak.

18. Dartmouth, 1956 – Egy kutatási terület nevet kap

Egy tudományterület megszületésének ritkán van egyetlen napja. Az előzmények rendszerint különböző műhelyekben, eltérő elnevezések alatt gyűlnek össze. Mégis szükség lehet egy olyan eseményre, amely közös vállalkozásként teszi láthatóvá őket. A mesterséges intelligencia történetében ezt a szerepet az 1956-os dartmouthi nyári kutatási projekt kapta.

A kezdeményezés 1955. augusztus 31-én kelt tervében már szerepelt az artificial intelligence kifejezés. A dokumentumot John McCarthy, Marvin Minsky, Nathaniel Rochester és Claude Shannon jegyezte. Kéthónapos, mintegy tízfős közös kutatást javasoltak a következő nyárra a New Hampshire állambeli Hanoverben, a Dartmouth College-ban. Az elnevezés dokumentált megjelenését és a későbbi találkozó évét ezért meg kell különböztetnünk. [97]

Négy kezdeményező, eltérő tapasztalatok

John McCarthy a Dartmouth matematikus oktatójaként az új kutatási cél köré akarta szervezni a munkát. Marvin Minsky már a neurális hálózatok és a tanuló szerkezetek felől is foglalkozott az intelligencia modellezésével. Claude Shannon az információelméletet és az áramkörök logikáját hozta magával. Nathaniel Rochester az IBM számítógépes fejlesztésében szerzett tapasztalatot, és az idegi működés gépi szimulációja is érdekelte. [97]



52. kép John McCarthy 2006-ban, ötven évvel a dartmouthi nyári projekt után. A későbbi portrét nem az 1956-os találkozó felvételeként közöljük. Fotó: null0 / Wikimedia Commons, CC BY-SA 2.0.

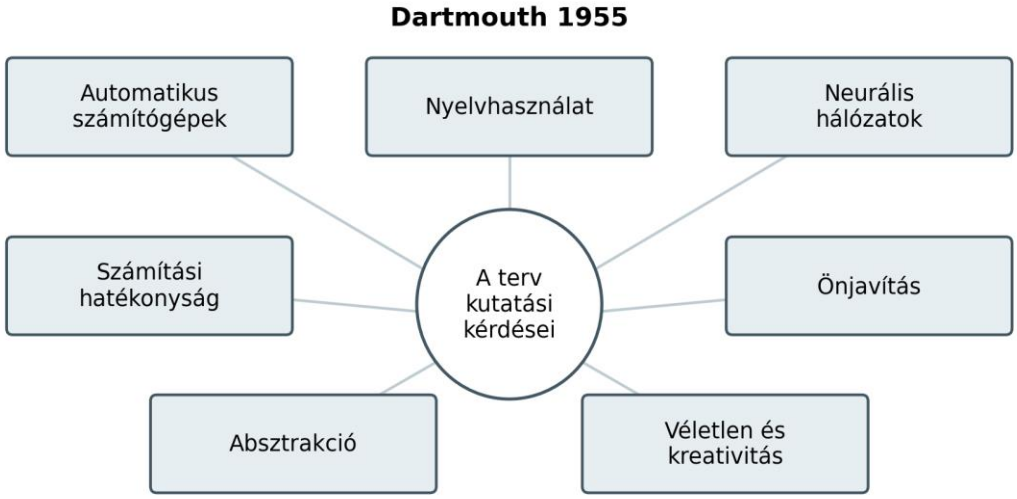
Ez az összetétel önmagában is figyelmeztet arra, hogy a kezdeteket ne egyetlen módszer diadalmeneteként írjuk le. Az agy modellezése, a logikai leírás és a számítógépes megvalósítás már a kezdeményezők érdeklődésében találkozott. Az eltérés nem pusztán a választott eszközökben volt. Más kérdést jelentett egy emberi képesség szerkezetét megmagyarázni, és mást egy bizonyos feladatot jól végző programot készíteni.

Közös kutatási terület akkor keletkezhet, ha az eltérő munkák között az eredmények is átjárhatóvá válnak. Egy matematikai leírás segítheti a programozót; a működő program pedig megmutathatja egy elmélet korlátait. Ehhez nem szükséges minden résztvevőnek ugyanazt gondolnia az elméről. Elég lehet, hogy felismerik egymás kérdéseiben a közös kutatási célt.

A vállalkozás alapfeltevése

A pályázati terv abból a feltevésből indult ki, hogy a tanulás és az intelligencia jellemzői elvileg olyan pontosan leírhatók, hogy géppel szimulálhatók legyenek. Ez kutatást irányító sejtés volt, nem már elért eredmény. A szerzők nyelvhasználatot, fogalomalkotást, problémamegoldást és önjavítást is vizsgálni akartak. [97]

A feltevés merészsége abban állt, hogy a kérdést megoldandó tudományos feladatnak tekintette. Nem kellett megvárni hozzá az emberi agy valamennyi részletének feltárását. Lehetett részfeladatokkal próbálkozni, eljárásokat alkotni, majd megvizsgálni, mire képesek. A kutatás nem azzal kezdődött, hogy már megvolt az intelligencia teljes elmélete.



53. kép Az 1955. augusztus 31-i dartmouthi terv hét témaköre rövid magyar összefoglalásban. A neurális hálózatok és a nyelvhasználat egyaránt szerepelt az eredeti programban. Saját szerkesztés; nem korabeli ábra.

A tervben a neurális hálózatok önálló témakörként szerepeltek. Ugyanott a nyelv, az absztrakció, a számítás hatékonysága, az önjavítás, valamint a véletlen és a kreativitás kapcsolata is helyet kapott. [97] A kezdeti program tehát tágabb volt annál, mint hogy egyetlenfajta gépet vagy egyetlen technikát jelöljön ki. A különböző részproblémákhoz másféle eszközök ígérkeztek használhatónak.

Az új névnek szervező ereje volt. Meg lehetett nevezni a kutatási célt, pályázatot lehetett írni rá, és különböző eredményeket lehetett ugyanahhoz a vállalkozáshoz kapcsolni. A névadás önmagában nem old meg tudományos problémát, de megváltoztathatja, kik keresik együtt a megoldást. Egy terület társadalmi és intézményi létezése részben ilyen közös megnevezéseken alapul.

Mi történt a nyári műhelyben?

McCarthy későbbi visszaemlékezése szerint a közös munka nem az eredetileg elképzelt módon zajlott. A résztvevők eltérő időpontokban és különböző hosszúságú időre érkeztek, saját kutatási elképzeléseiket követték. Nem olyan konferenciát kell elképzelnünk, amelynek minden előadásán végig ugyanaz a közönség ült. [98]

Ez a körülmény a történeti képet is árnyalja. Egy később híressé vált csoportkép vagy névsor nem mondja meg önmagában, ki melyik vitában, mennyi ideig és milyen szerepben vett részt. A nyári projekt, az előkészítő terv és az egyes kutatók saját munkája egymással összefüggő, de külön dokumentálható történet.

Allen Newell és Herbert Simon rövid látogatása fontos esemény lett. McCarthy felidézte, hogy bemutatták a Logic Theory Machine munkáját, vagyis a szimbolikus logika tételeinek bizonyítását kereső rendszerüket. A programozáshoz szükséges listastruktúrák és leíró nyelvek szintén szóba kerültek. [98] Így a gépi intelligencia kérdése kézzelfogható eljárásokhoz és kipróbálható feladatokhoz kapcsolódott.

A bizonyításkeresés különösen alkalmas bemutatóterep. A sikernek világos feltétele lehet: a rendszer szabályos levezetést talál az adott kiinduló állításokból. Ugyanakkor a megoldás megtalálása sok lehetőség közötti keresést kívánhat. A feladat ezért egyszerre teszi próbára a pontos szabálykövetést és a célszerű választást a lehetséges lépések között.

Kezdet, de nem a semmiből

Dartmouth jelentősége nem az, hogy 1956 előtt senki sem gondolt intelligens gépre. A könyv eddigi fejezeteiben éppen az előzmények gazdagságát láttuk. A logikai formalizálás, a számítógép, a kibernetika és a neurális modellek nélkül az új vállalkozásnak nem lett volna sem fogalmi, sem műszaki alapja.

Az egyetem történeti önmeghatározása a nyári projektet a kutatási terület születésének mérföldköveként tartja számon. [99] Ez jól használható történeti határpont, ha nem takarja el az előzményeket és a folytatás sokféleségét. Egy születési év egy összetett folyamat megjelölése, nem annak bizonyítéka, hogy minden lényeges gondolat egyszerre keletkezett.

McCarthy utólag a mesterséges intelligencia önálló tudományos területként való megfogalmazását tekintette a találkozó legfontosabb hozadékának. Az emberi szintű intelligenciához vezető áttörésbe vetett akkori reménye nem teljesült a nyár folyamán. [98] Ezt érdemes együtt látni az esemény jelentőségével. A kezdeményezés hatása nagy volt, miközben a közvetlen eredménye szerényebb maradt a várakozásnál.

Az ígéret és a mérhető feladat

Az optimista kutatási program egyszerre lehet termékeny és félrevezető. Termékeny, mert olyan problémák megoldására ösztönöz, amelyeket addig túl távolinak tartottak. Félrevezetővé akkor válhat, ha egy részfeladat sikerét az egész intelligencia közeli megfejtésének tekintik. Egy logikai tétel bizonyítása még nem oldja meg a nyelv megértését, és néhány minta felismerése sem ad általános világismeretet.

Az új terület fejlődését ezért kétféle kérdés kíséri. Milyen gépet szeretnénk végül létrehozni, és mit tud most ténylegesen a megépített rendszer? A távlati cél kijelöli az irányt, az ellenőrizhető eredmény pedig megmutatja, hol tartunk. Amikor a kettő összemosódik, a kutatás történetéből könnyen puszta sikertörténet vagy csalódástörténet lesz.

Dartmouth után nem egyetlen módszer kezdett menetelni az előre kijelölt cél felé. Különböző kutatók másként fogalmazták meg a feladatot, és másfajta gépi működést építettek. A következő fejezet ennek két nagy irányát követi: a jelképekkel és szabályokkal dolgozó programokat, valamint a tapasztalat hatására változó neurális hálózatokat.

19. Két út az intelligencia felé

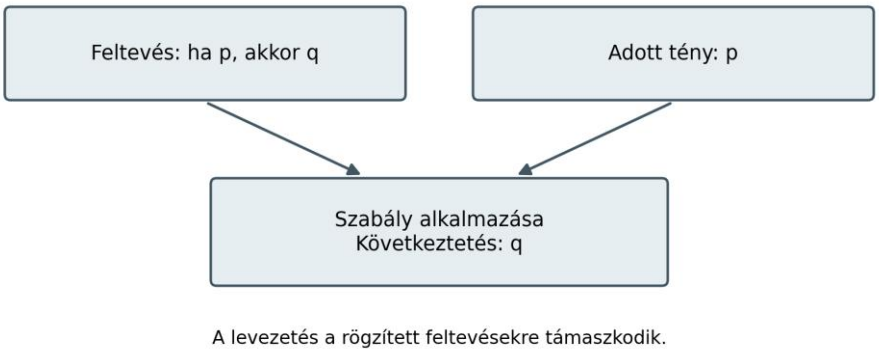
Egy feladat megoldásához többféle tudás vezethet. Elmagyarázhatjuk valakinek a szabályokat, majd megkérhetjük, hogy alkalmazza őket. Mutathatunk sok példát is, amelyekből megtanulhat felismerni egy mintát. Az embernél a kettő többnyire összekapcsolódik. A korai mesterséges intelligenciában mégis eltérő kutatási hangsúlyokat teremtettek: a szimbolikus és a neurális megközelítést.

Ez hasznos felosztás, ha nem tekintjük két teljesen zárt világnak. A szimbolikus programok is tanulhatnak, a neurális hálózatokat pedig szintén algoritmusok működtetik. A kérdés az, hogyan ábrázolják a feladat szempontjából fontos összefüggéseket, és miként jutnak el a válaszhoz.

Jelképek és megengedett műveletek

A szimbolikus megközelítésben az állításokat, tárgyakat vagy helyzeteket elkülöníthető jelképekkel írjuk le. A program ezek között meghatározott műveleteket végez. Egy logikai rendszerben például szabály rögzíti, milyen állításból milyen további állítás következhet. A gép így nem a mondat betűinek külső alakja, hanem a formalizált szerkezet szerint dolgozik.

Vegyünk egy saját, egyszerű példát. Elfogadjuk azt a feltevést, hogy ha esik az eső, nedves a járda. Elfogadjuk azt is, hogy most esik. A következtetési szabály alapján a járda nedvessége levezethető. A programnak nem kell hozzá fényképet látnia a járdáról. Az eredmény az elfogadott kiinduló állítások és a szabály viszonyából születik.



54. kép Egy elemi következtetési lépés: a ha p, akkor q állítás és p fennállása alapján q következik. A szöveg esőre és járdára vonatkozó példájában p az esőt, q a járda nedvességét jelöli. A levezetés a kiinduló feltevésekre támaszkodik. Saját szerkesztés.

A példa határa ugyanilyen fontos. Ha a járdát tető védi, a világra vonatkozó egyszerű feltevésünk nem feltétlenül helyes. A logikai levezetés lehet szabályos akkor is, ha rosszul írtuk le a helyzetet. A szimbolikus módszer ereje a következmények pontos kezelésében van; a megfelelő kiinduló tudást azonban valahonnan meg kell szerezni és alkalmas formában rögzíteni kell.

A Logic Theorist és a keresés szerepe

Newell és Simon 1956-os The Logic Theory Machine című munkája olyan rendszer leírását adta, amely szimbolikus logikai tételek bizonyítását kereste. A program nem egyszerűen az összes elképzelhető műveletsort próbálta végig azonos módon. Heurisztikákat, a keresést célszerűen irányító eljárásokat használt. A korabeli tanulmány a rendszer pontos leírására összpontosított, és ezt megkülönböztette a számítógépes megvalósítástól. [100]

A heurisztika olyan fogás, amely ígéretesebbnek látszó irányba tereli a keresést. Egy elveszett kulcsot keresve előbb nézünk a zsebünkbe, mint a padlás legtávolabbi sarkába. Ettől még nem biztos, hogy a kulcs a zsebünkben lesz. A jó keresési eljárás nemcsak azt határozza meg, melyik lépés szabályos, hanem azt is, melyiket érdemes előbb kipróbálni.

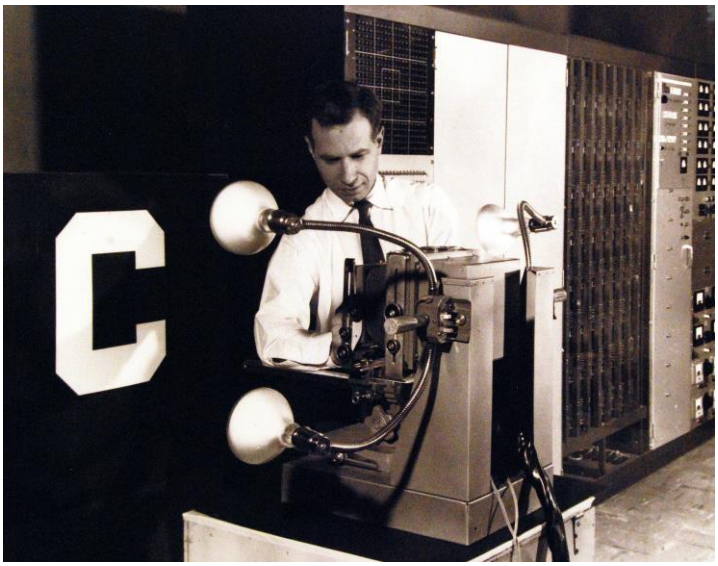
A lehetséges lépések száma gyorsan nőhet. Ha minden helyzetből több irányban folytathatjuk a munkát, néhány újabb lépés már nagyságrendekkel több vizsgálandó esetet teremthet. Egy elvben helyes módszer ezért a gyakorlatban használhatatlanul lassú lehet. A bizonyításkeresés történeti jelentőségéhez hozzátartozik, hogy a gépi gondolkodást a hatékony keresés problémájává is tette.

A szimbolikus rendszerben gyakran megőrizhető a következtetés útja. Megmutatható, melyik állításból, milyen szabály alkalmazásával jutottunk tovább. Ez segíti az ellenőrzést. A végső válasz azonban továbbra is a rendszerben rögzített fogalmakhoz és feltevésekhez kötődik. A szabályos levezetés nem pótolja automatikusan a hiányzó tapasztalatot.

A kapcsolatokban változó tudás

A neurális megközelítés más helyen keresi a megoldás kialakulását. Sok egyszerű egységet kapcsol össze, és a kapcsolatok hatásának megváltoztatásával alakítja a hálózat válaszát. A tizenharmadik fejezet McCulloch–Pitts-modellje és Hebb elképzelése már két fontos kiindulópontot adott: egyszerű egységekből összetett működés épülhet fel, és a tapasztalat tartósan módosíthatja a kapcsolatokat.

Frank Rosenblatt perceptronja az 1950-es évek végén a tanuló mintafelismerés egyik fontos modellje lett. Az 1958-as tanulmány a felismerést és az információ tárolását a változó kapcsolatrendszer felől vizsgálta. A történeti elképzelés érzékelő, közbülső és válaszadó egységeket is tartalmazott; nem célszerű minden részletét a mai tankönyvek egyetlen küszöbelemével azonosítani. [101]

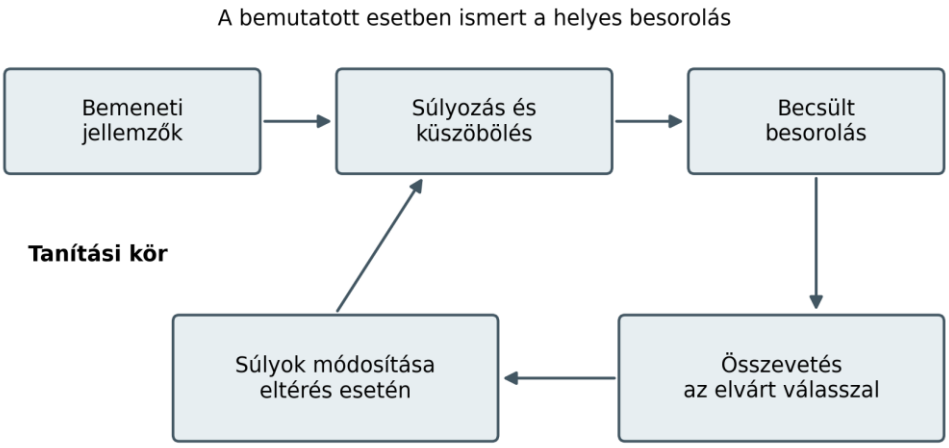


55. kép A Mark I Perceptron optikai bemenettel és a háttérben látható berendezéssel. Az amerikai haditengerészet 1960. június 24-én közreadott archív felvétele. Nem az 1958-as IBM 704-es szimuláció képe. National Museum of the U.S. Navy / Wikimedia Commons. Közkincs.

Az 1958-as nyilvános demonstráció számítógépes szimuláción, IBM 704-es gépen mutatott be tanuló megkülönböztetést. Ettől el kell választani a későbbi Mark I Perceptron fizikai berendezését. A modell, a szimuláció és a megépített eszköz ugyanannak a kutatási programnak eltérő megjelenései. [102]

Mit jelent, hogy a gép példákból tanul?

Egy mai, egyszerűsített oktatási példában a hálózat bemenete néhány mért jellemző: például egy alakzat sötét részeinek elhelyezkedése. A rendszer súlyozza a bemeneteket, összegezi a hatásukat, majd egy küszöb alapján választ két osztály között. A tanítás során az eredményt összevetjük az elvárt besorolással. Hiba esetén a tanulási szabály módosítja a súlyokat, hogy az adott példa besorolása a helyes irányba változzon.



56. kép Egy egyszerű, felügyelt osztályozó tanítási köre. A becslés és az elvárt válasz eltérése alapján változhatnak a súlyok. Mai magyarázó rajz, nem Rosenblatt teljes történeti modelljének kapcsolási rajza. Saját szerkesztés.

Itt az ember nem feltétlenül írja le tételesen, milyen minden felismerendő alakzat. Ehelyett példákat és megfelelő visszajelzést ad. A rendszer működésében a súlyok hordozzák a korábbi tanítás hatását. A kialakult számértékekből azonban rendszerint nem olvasható ki egyszerűen egy olyan mondat, hogy melyik alakzati tulajdonság miért döntő.

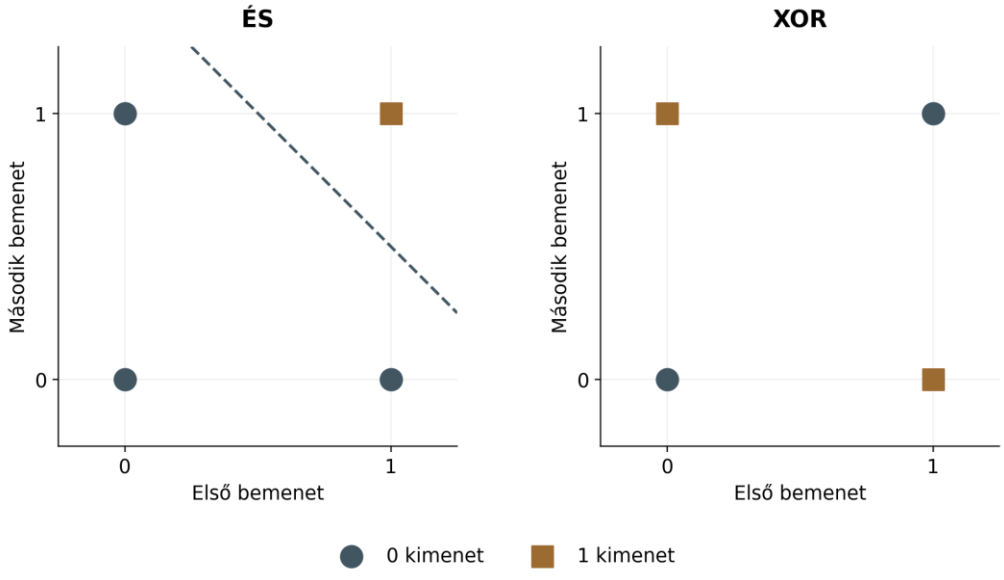
A súlymódosítás nem azonos Hebb legegyszerűbb együttaktivitási szabályával. A mostani példában az elvárt válasz és a tényleges kimenet eltérése irányítja a korrekciót. A tanulásnak többféle eljárása lehet, ezért a neurális hálózat megnevezés önmagában még nem mondja meg, miből és hogyan tanul a rendszer.

Tanítás után új példákat is meg kell vizsgálni. Ha a gép csak a korábban látott eseteket kezeli jól, még nem tudjuk, használható mintázatot sajátított-e el. A cél az általánosítás: megfelelő választ adni olyan esetre is, amely a tanításban pontosan ugyanebben a formában nem szerepelt. Ezt külön kell ellenőrizni, a tanítási példákon elért jó eredmény nem elegendő hozzá.

Az egyszerű modell határa

Egyetlen, a bemenetek súlyozott összegét küszöbölő egység két jellemző esetén egyenes vonallal választja ketté a síkot. Ez bizonyos osztályozásokhoz elég. Más feladatnál a különböző osztályok úgy helyezkednek el, hogy egyetlen egyenes sem különíti el őket.

A kizáró vagy, röviden XOR, jól követhető példa. Két bináris bemenet közül pontosan az egyiknek kell aktívnak lennie ahhoz, hogy a válasz egy legyen. A két ilyen pont a négyzet két szemközti sarkán helyezkedik el, a másik két sarok pedig a másik osztályhoz tartozik. Egyetlen egyenes nem választja szét ezt az elrendezést. Az ábra saját geometriai szemléltetés.



57. kép A körök a 0, a négyzetek az 1 kimenethez tartozó bemeneteket jelölik. Az ÉS két osztálya egy egyenessel szétválasztható; az XOR esetében ez nem lehetséges. A bal oldali szaggatott vonal egy lehetséges választóvonal. Saját geometriai szemléltetés, nem mérési eredmény.

A korlát a megadott modellre és bemeneti leírásra vonatkozik. Közbülső egységek vagy megfelelően átalakított jellemzők már más lehetőségeket adhatnak. Több tanítópélda önmagában nem tünteti el

azt a határt, amelyet a választott szerkezet szab. A tanulási eljárást, a bemenetek ábrázolását és a hálózat felépítését együtt kell vizsgálni.

Két hangsúly, sokféle változat

Szempont	Szimbolikus példa	Neurális példa
A tudás alakja	Kifejezett állítások és szabályok	Kapcsolatokhoz rendelt súlyok
A válasz kialakítása	Szabályalkalmazás és keresés	A bemeneti jelek hálózati feldolgozása
Változás tanításkor	Szabályok vagy más leírt ismeretek is változhatnak	A bemutatott esetben a súlyok módosulnak
Ellenőrzés egyik módja	A levezetés és a feltevések vizsgálata	Új példákon mért helyes besorolás
Jellegzetes nehézség	A megfelelő ismeretek pontos leírása	Alkalmas adatok, jellemzők és szerkezet választása

Két jellegzetes megközelítés összehasonlítása a fejezet példái alapján. A táblázat nem zárja ki a szimbolikus tanulást vagy a két módszer kombinálását. Saját szerkesztés.

A táblázat két jellegzetes esetet hasonlít össze. A valós kutatások nem mindig illenek tisztán valamelyik oszlopba. A szimbolikus rendszer szabályait tanulási eljárás is kialakíthatja. A neurális rendszernél pedig az ember választja meg a feladatot, az adatok formáját és a módosítás szabályát. A tanuló gép sem módszer nélküli szerkezet.

Az ellenőrizhetőség is összetettebb az egyszerű ellentétnél. Egy rövid logikai bizonyítás könnyen áttekinthető, egy nagyon nagy szabályrendszer viszont nehezen. Egy neurális válasz belső oka nehezebben fogalmazható meg, de a rendszer meghatározott tulajdonságai ettől még vizsgálhatók. A magyarázhatóság nem egyetlen kapcsoló, amely az egyik megközelítésben mindig be van kapcsolva, a másikban pedig mindig hiányzik.

A két út találkozása

Egy elképzelt iratfeldolgozó rendszer jól mutatja az összekapcsolás lehetőségét. Egy tanuló felismerő kiolvassa a nyomtatott lap mezőit, majd egy szabályokkal dolgozó program ellenőrzi, hogy a részösszegek megfelelnek-e a végösszegnek. A felismerés és a következtetés itt külön feladatot végez. Az egyik bizonytalansága a másik eredményét is befolyásolhatja, ezért a teljes folyamatot kell ellenőrizni.

Ez a példa nem egy konkrét mai termék leírása, hanem a két megközelítés együttműködésének szemléltetése. Megmutatja azt is, miért nem szerencsés az AI történetét pusztán egymást leváltó módszerek soraként elbeszélni. A világ jelenségeinek felismerése és az ismert összefüggések következményeinek vizsgálata eltérő nehézségeket teremt; a kétféle eszköz kiegészítheti egymást.

A következő évtizedekben a kutatási hangsúlyok és a remények többször megváltoznak. A kiinduló kérdés azonban megmarad: mennyit kell a megoldásból előre leírnunk, mennyit tanulhat meg a gép a tapasztalatból, és hogyan ellenőrizzük, hogy valóban azt tanulta-e meg, amire szükségünk van? A mesterséges intelligencia önálló területként e kérdésekkel indult tovább.

Források és további olvasmányok

A hivatkozások és a képek számozása az 1–16. fejezetet folytatja. Az eredeti kutatási dokumentumokat a későbbi visszaemlékezésektől elkülönítjük. A mai példák és magyarázó ábrák a fogalmak megértését segítik; nem korabeli kísérletek dokumentumai.

[96] Alan M. Turing: Computing Machinery and Intelligence. Mind, 59/236, 1950. október, 433–460. DOI: 10.1093/mind/LIX.236.433. Különösen az 1–3., 6. és 7. rész: az utánzási játék, az ellenvetések és a tanuló gépek. A fejezet az eredeti háromszereplős játékot megkülönbözteti a később közismert teszthelyzettől. A vonatutazási példa és az értékelési szempontok mai szerkesztői magyarázatok. Az elsődleges tanulmány teljes átirata · Kiadói bibliográfiai adatok · A gépirat archívumi rekordja

[97] John McCarthy – Marvin L. Minsky – Nathaniel Rochester – Claude E. Shannon: A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence. 1955. augusztus 31. Elsődleges kutatási terv a következő nyári műhelymunkához. A szerzői bemutatkozások és a hét témakör is olvasható a Stanfordon megőrzött változatban. A feltevések kutatási célokat jelölnek, nem már bizonyított eredményeket. A teljes dokumentum szövege · Az AI Magazine 2006-os újraközlése

[98] John McCarthy: The Dartmouth Workshop – as planned and as it happened. Az AI Past and Future című előadás internetes változata, 2006. Az egyik szervező visszaemlékezése a tervre, a változó részvételre, Newell és Simon bemutatójára és a találkozó hozadékára. A szerző maga is jelzi, hogy személyes érdeklődésével összefüggő eseményekre emlékezik; nem teljes jegyzőkönyv. Szervezői visszaemlékezés

[99] Dartmouth College: Artificial Intelligence (AI) Coined at Dartmouth. Intézményi történeti ismertető az 1956-os kutatási projektről. A névadás dátumánál a fejezet az 1955-ös eredeti dokumentumot követi, és ezt elválasztja az 1956-os eseménytől. Dartmouth történeti oldala

[100] Allen Newell – Herbert A. Simon: The Logic Theory Machine – A Complex Information Processing System. RAND, P-868, 1956. július 12. DOI: 10.7249/P868. Folyóiratváltozat: IRE Transactions on Information Theory, 2/3, 1956, 61–79, DOI: 10.1109/TIT.1956.1056797. A rendszer specifikációja és a heurisztikus keresés szerepe. A RAND-változat külön kezeli a pontos leírást és a számítógépes megvalósítást; a fejezet ezt nem mossa össze. A járdára és az elveszett kulcsra vonatkozó példák saját szemléltetések. RAND kiadói adatlap · A korabeli jelentés digitális másolata · Carnegie Mellon archívumi rekord

[101] Frank Rosenblatt: The Perceptron – A Probabilistic Model for Information Storage and Organization in the Brain. Psychological Review, 65/6, 1958, 386–408. DOI: 10.1037/h0042519. Az eredeti elképzelés érzékelő, asszociációs és válaszadó részeket is tartalmaz. A fejezet egyszerű tanítási köre mai oktatási példa, nem a tanulmány teljes rendszerének reprodukciója. Az ÉS és az XOR geometriai ábrája saját, közvetlenül ellenőrizhető szemléltetés. Az eredeti tanulmány digitalizált változata · Bibliográfiai rekord

[102] Cornell Chronicle: Professor’s perceptron paved the way for AI – 60 years too soon, 2019. szeptember 25. A Cornell saját kutatója és gyűjteménye alapján készült intézményi visszatekintés. Az 1958-as IBM 704-es demonstrációt a Mark I fizikai berendezésétől megkülönböztető adatot használjuk belőle; a cím értékítéletét és minden történeti általánosítását nem vesszük át. Cornell intézményi visszatekintése

Képforrások és felhasználási adatok

50. kép Alan Turing portréja, 1951. március 29. Elliott & Fry. A Commons közkincsként közli; a letöltött képfájl változtatás nélkül szerepel. Kép és metaadatok

51. kép Saját szerkesztésű ábra a közismert ember–gép megkülönböztetési teszthez. A történeti alap és az eredeti játék különbsége a [96] forrás és a fejezetszöveg alapján követhető.

52. kép John McCarthy, 2006. május 13. Fotó: null0. A Commons képfájlját további módosítás nélkül közöljük. CC BY-SA 2.0. Kép és metaadatok · Licenc

53. kép Saját szerkesztésű tématérkép a [97] dokumentum hét kutatási kérdéséről. A magyar címkék rövid tartalmi fordítások.

54. kép Saját szerkesztésű példa a modus ponens következtetési szabályra. Nem a Logic Theorist korabeli képernyőképe vagy programrészlete.

55. kép Mark I Perceptron. National Museum of the U.S. Navy, 330-PSA-80-60, USN 710739. A leírás szerint a felvételt 1960. június 24-én adták közre; a 2015-ös adat a digitális közléshez kapcsolódik. A képfájl további módosítás nélkül szerepel. Amerikai szövetségi kormányzati mű, a forrás közkincsként közli. Kép és teljes archív leírás · Public Domain Mark

56. kép Saját szerkesztésű mai tanítási kör. A rajz ismert címkékkel történő, hibajel alapján végzett súlymódosítást mutat, nem minden neurális tanulás általános folyamatát.

57. kép Saját szerkesztésű geometriai példa. ÉS: a kimenet csak az (1, 1) bemenetnél 1; a választóvonal az első és második bemenet összegének 1,5-ös határa. XOR: az (0, 1) és az (1, 0) kimenete 1, a másik két ponté 0. Szimulált vagy mért adatsor helyett a logikai függvények teljes négyes bemeneti készletét ábrázoljuk.

IV. rész – Megszületik a mesterséges intelligencia

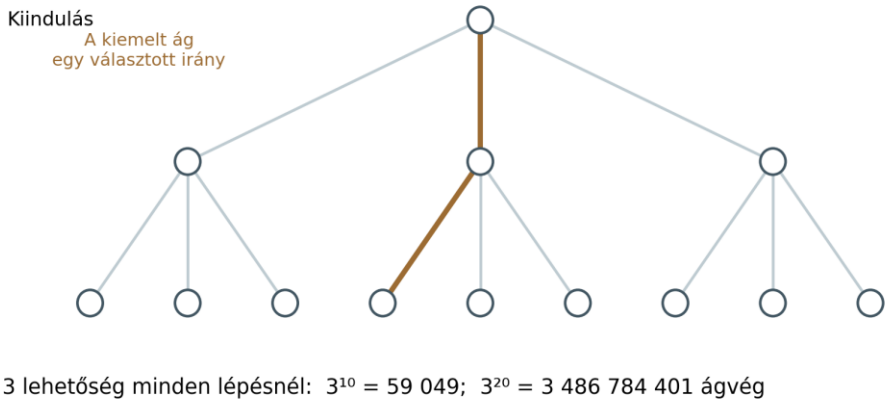
20. Az első sikerek és a nagy ígéretek

Az 1950-es évek végére a mesterséges intelligencia kutatói már működő eljárásokat mutathattak. A gép bizonyítást keresett, játékállásokat értékelt, és tanítás hatására megváltoztatta a válaszát. Ezek a feladatok korábban az emberi értelem jellegzetes megnyilvánulásainak számítottak. Érthető volt a remény: ha az első lépések ilyen gyorsan sikerülnek, talán az általánosabb képességekre sem kell sokáig várni. A nehézséget annak megítélése okozta, mekkora út vezet egy jól megválasztott részfeladattól a mindennapi intelligenciáig.

Bizonyítást keresni és bizonyítást ellenőrizni

Az előző fejezetben megismert Logic Theory Machine a szimbolikus logika tételeihez keresett levezetéseket. Allen Newell és Herbert Simon 1956-os tanulmánya a rendszer részletes leírását adta; a kutatási és programozási munkában J. C. Shaw is részt vett. Az eljárás heurisztikákkal választott az ígéretes lépések között. A tételbizonyítás ezzel a gépi problémamegoldás kísérleti terepe lett. [103, 104]

A feladatban kétféle munka különül el. Egy kész bizonyításnál sorban megvizsgálhatjuk, hogy a lépések megfelelnek-e a megengedett szabályoknak. A bizonyítás megtalálásakor viszont még nem tudjuk, melyik szabályt mikor érdemes alkalmazni. Olyan ez, mint egy útvonal ellenőrzése és az útvonal megtervezése közötti különbség. Az elsőnél előttünk a megoldás; a másodiknál nekünk kell kiválasztanunk a megfelelő lehetőségeket.



58. kép Egy háromfelé ágazó keresési fa első szintjei. A kiemelés egy választott irányt mutat; a számok az adott mélységben elérhető ágvégek számát adják teljes, háromfelé ágazó fában. Saját magyarázó ábra, nem történeti mérési eredmény.

A számítógép gyorsasága önmagában nem tünteti el ezt a különbséget. Saját számpéldánkban, ha minden állapotból három új állapot nyílik, tíz lépés mélységben már 59 049 különböző ágvéghez juthatunk. Húsz lépésnél ez 3 486 784 401. A példa teljes, háromfelé ágazó keresési fát feltételez, nem egy történeti program mért teljesítményét. Azt mutatja, miért vált központi kérdéssé a fölösleges keresés elkerülése.

A heurisztika haszna attól függ, mennyire illik a feladathoz. Egy matematikai bizonyításnál jó ötlet lehet a kívánt végállítás szerkezetéből kiindulni. Más problémában az előrehaladásnak más jelei vannak. A korai sikerek ezért egyszerre mutatták meg a szabályok erejét és a keresést irányító tudás jelentőségét.

Az általános problémamegoldó ígérete

Newell, Shaw és Simon következő vállalkozása a General Problem Solver, röviden GPS volt. Az 1959-es beszámoló a problémamegoldás olyan szerkezetét vizsgálta, amely különböző feladatleírásokkal használható. Fontos módszere az eszköz–cél elemzés: a rendszer megkeresi a jelenlegi és a kívánt helyzet közötti különbséget, majd olyan műveletet választ, amely azt csökkentheti. A szükséges előfeltételekből újabb részcélok keletkezhetnek. [104]

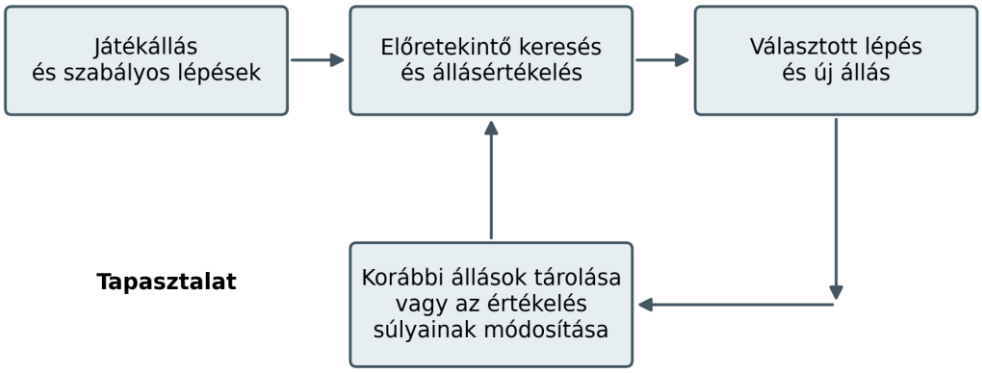
Vegyünk egy mai magyarázó példát. Egy dobozt át kell tenni az asztal egyik végéről a másikra. Ehhez meg kell fogni. Ha egy másik tárgy elzárja a hozzáférést, előbb azt kell elmozdítani. A távoli célból így kisebb feladatok következnek. Ha azonban a gép nem kapja meg, mit jelent a megfoghatóság, mely tárgyak mozgathatók, és milyen mozdulatok állnak rendelkezésére, a cél megnevezése még nem elég.

A GPS általánossága az újra felhasználható problémamegoldási szerkezetben rejlett. A feladatot továbbra is megfelelő jelképekkel, megengedett műveletekkel és feltételekkel kellett leírni. Ezt az előkészítést az ember végezte. Az általános célú számítógép, az általánosítható keresési módszer és a bármilyen élethelyzetben boldoguló intelligencia három eltérő követelmény.

Samuel dámaprogramja és a tapasztalat értéke

Arthur Samuel az IBM-nél az 1950-es években olyan dámaprogramon dolgozott, amely tapasztalatai alapján javíthatott a játékán. Az 1959-ben közölt tanulmány két tanulási eljárást ismertetett. Az egyik korábban megvizsgált állásokat és a hozzájuk tartozó értékeléseket őrzött meg. A másik az állások értékeléséhez használt jellemzők súlyát módosította. Samuel beszámolója szerint a program a készítőjénél jobb játékra is képessé vált. Ez nem jelentette a dámajáték teljes megoldását vagy a világ legjobb játékosának legyőzését. [105]

A kétféle tanulás különbségét könnyű megérteni. Ha ugyanaz a helyzet visszatér, hasznos lehet a korábbi számítás eredményét elővenni. Ha új helyzettel találkozunk, olyan szempontok segíthetnek, amelyek más állásokra is alkalmazhatók. A jellemzők súlyozása azt szabja meg, mekkora jelentőséget tulajdonítunk egy-egy előnynek vagy hátránynak.



59. kép A keresés és a tapasztalat felhasználásának egyszerűsített kapcsolata egy tanuló játékprogramban. Samuel két eljárása az állások tárolását, illetve az értékelés módosítását vizsgálta. Mai magyarázó ábra, nem az eredeti program folyamatábrája. Saját szerkesztés.

A játék szabályait és a számítás keretét emberek készítették el. A tanulás ezek között módosította a program döntéseit. A dámaprogram ezért jól mutatja, miért nem esik egybe a szimbolikus módszer és a változatlan működés. A keresés, a tárolt tapasztalat és a tanulható értékelés ugyanabban a rendszerben is együtt dolgozhat.

A részletes játék- és versenytörténet külön fejezetet érdemel, az „Amikor a gépek játszani kezdtek” témakörében. Itt a kutatási következmény a fontos: nem kellett minden jó választ egyenként beírni a gépbe. Meg lehetett adni olyan eljárást, amelynek későbbi teljesítményét a tapasztalat is alakítja. Ettől azonban a tanuló rendszer feladata még a dáma maradt.

A perceptron körüli remények

Frank Rosenblatt 1958-as perceptrontanulmánya más irányból közelítette meg a tanulást. Az idegrendszer által ihletett modellben az érzékelt minták és a válaszok közötti kapcsolatok változhattak. A kutatás azt vizsgálta, miként alakulhat ki megkülönböztető képesség a kapcsolatrendszer módosulásából. [106] Az előző fejezet egyszerű osztályozási példája ennek egy mai, szűkített magyarázata volt.

A tételbizonyítás és a perceptron eredményei egymást erősítő reményt kelthettek. Az egyik oldalon a gép olyan jelképekkel dolgozott, amelyekhez emberi jelentést rendeltünk; a másikon példák hatására változott a működése. Kézenfekvőnek látszhatott, hogy a külön-külön elért képességek összekapcsolásával egyre teljesebb intelligencia áll elő. Ez azonban kutatási várakozás volt. Az összekapcsolás nehézségéről a különálló bemutatók még keveset árultak el.

Egy alakzat helyes besorolása például nem mutatja meg, hogy a rendszer felismeri-e ugyanazt a tárgyat más megvilágításban vagy részleges takarásban. Egy bizonyítás megtalálása sem mondja meg, hogyan fogalmazza meg a gép a bizonyítandó állítást egy rendezetlen tapasztalatból. Az egyes képességek között új feladatok nyílnak, amelyek az eredeti kísérletben még nem szerepeltek.

Miért tűnhetett rövidnek az út

A kor optimizmusát Simon és Newell egyik konkrét előrejelzése is érzékelteti. Az 1957 novemberében előadott, 1958-ban megjelent közös munkájukban tíz éven belül számítógépes sakkvilágbajnokot vártak, amennyiben a versenyszabályok nem zárják ki a gépet. Ugyanilyen időtávra fontos új matematikai tétel gépi felfedezését és bizonyítását is jósolták. [107] Az előrejelzés tárgya és határideje világos volt; nem szabad utólag általános, időtlen lelkesedéssé szelídíteni.

Az első működő eljárások több korábbi kételyt eloszlattak. A gép képes volt nem számszerű jelképek kezelésére, keresésre és teljesítményének tapasztalati javítására. Az új lehetőségeket azonban többnyire jól előkészített feladatokon vizsgálták. A világ leírásának és a siker meghatározásának jelentős részét a kutatók már elvégezték.

Egy új rendszer értékelésekor ezért a megoldás mellett az előkészítést is látni kell. Ki választotta ki a fogalmakat? Mely helyzeteket zárták ki? Mennyi segítség kellett a bemutatóhoz, és mi történik a következő, kissé eltérő feladatnál? A korai eredmények valóságosak voltak. A belőlük levont, emberi szintű intelligenciára vonatkozó következtetéshez viszont több bizonyíték kellett.

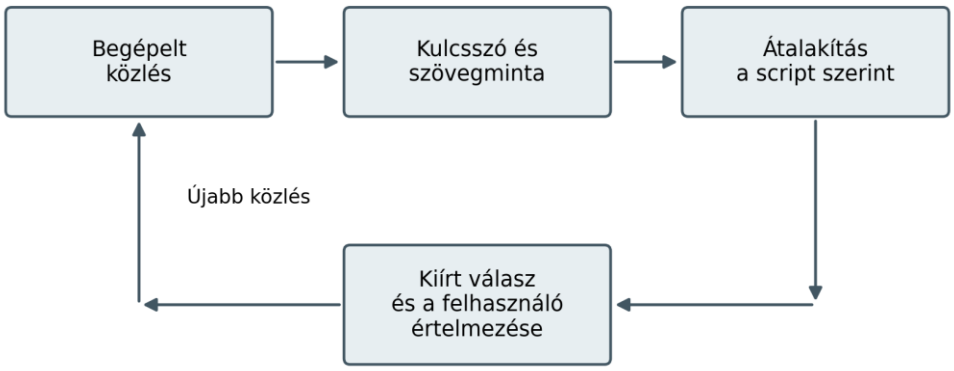
21. A gép megszólal, körülnéz és cselekszik

Egy párbeszéd, egy mozgó tárgyakat ábrázoló képernyő vagy egy guruló robot közvetlenebb benyomást kelt, mint egy logikai levezetés. Az 1960-as években a gépi intelligencia ilyen alakokban is megjelent. ELIZA szöveges válaszokat adott, SHRDLU tárgyakról beszélt és műveleteket tervezett, Shakey pedig fizikai környezetben mozgott. A fejezet címében a „megszólal” a nyelvi kapcsolatra utal: az itt tárgyalt párbeszédek begépelt szövegen alapultak, nem a mai értelemben vett beszédfelismerésen.

ELIZA és a visszakérdezés hatása

Joseph Weizenbaum az MIT-n készítette ELIZA programját; leírása 1966 januárjában jelent meg. A program kulcsszavakat keresett a bevitt szövegben, majd felbontási és újraösszeállítási szabályokkal készített választ. A társalgás jellegét cserélhető szabálykészlet, úgynevezett script adta. Legismertebb változata, a DOCTOR, egy rogersi pszichoterapeuta beszélgetésmódját utánozta. [108]

Ez a szerep kedvezett a módszernek. A visszakérdezés természetesnek hatott, és a beszélgetőpartner újabb mondatai szolgáltatták a további válaszok anyagát. Weizenbaum azt is tárgyalta, hogyan tulajdonít a felhasználó jelentőséget a gép válaszainak. A programnak voltak korábbi közlések későbbi felhasználását segítő megoldásai; pontatlan volna teljesen emlékezet nélküli, egyetlen mondatcserére épülő trükknek nevezni. [108]



60. kép A kulcsszavakra és szövegátalakításra épülő társalgás egyszerűsített vázlata. ELIZA további megoldásokat is használt, például korábbi közlések későbbi elővételét. Az értelmezésben a felhasználó saját háttértudása is szerepet kaphat. Saját magyarázó ábra.

A működés megértéséhez saját magyar szemléltető példát alkothatunk. A felhasználó azt írja: „Aggódom a holnapi beszélgetés miatt.” Egy egyszerű válaszsablon visszakérdezhet: „Mi aggaszt a holnapi beszélgetésben?” Ez a mondat figyelmesnek hangzik, pedig előállításához elegendő lehetett az „aggódom” minta felismerése és a mondatrész átrendezése. A példa nem ELIZA hiteles magyar párbeszéde és nem az eredeti angol szabálykészlet szó szerinti fordítása.

A felhasználó közben saját tudásával egészíti ki a választ. Tudja, melyik találkozóról van szó, miért kellemetlen, és milyen előzményei vannak. A visszakérdezést ezért olyan összefüggésbe helyezi,

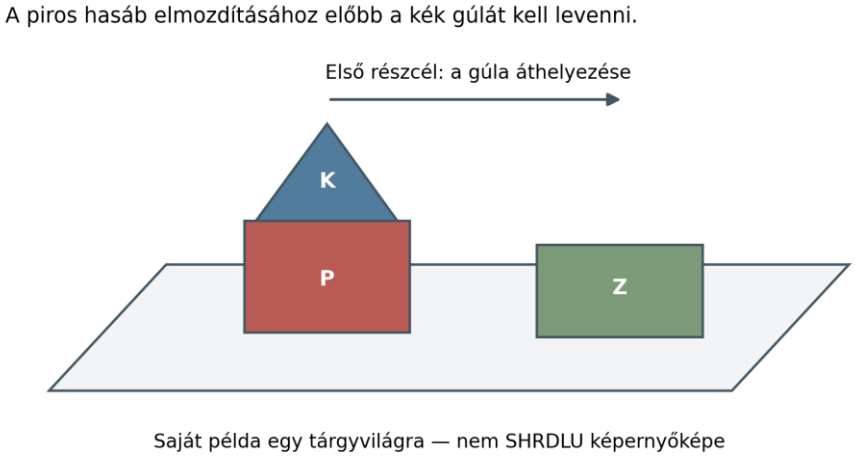
amelyet a programnak nem kellett felépítenie. A társalgás gördülékenysége részben a másik fél munkája is lehet.

Ebből ellenőrizhető kérdés következik. Ha a beszélgetésben megváltoztatunk egy lényeges körülményt, a válaszoló követi-e a változást? Meg tudja-e indokolni az összefüggést, vagy újabb általános kérdést ad? ELIZA története az ilyen vizsgálat szükségességét mutatja. Egy későbbi rendszer képességeit természetesen a saját működése és teljesítménye alapján kell megítélni; ELIZA korlátai nem bizonyítják minden lehetséges társalgó gép korlátait.

SHRDLU tárgyai és követhető cselekvései

Terry Winograd MIT-n végzett munkáját az 1971-es disszertáció és az 1972-es Understanding Natural Language részletesen ismertette. A SHRDLU nevű rendszer angol kérdéseket és utasításokat kezelt egy szimulált tárgyvilágban. Színes hasábok, gúlák és más egyszerű elemek között kellett eligazodnia. A képernyőn megjelenő környezethez és a szimulált megfogóhoz kapcsolódott a nyelvi párbeszéd. [109, 110]

A rendszer a nyelvtani elemzést, a jelentés kezelését és a következtetést összekapcsolta. A tárgyakról részletes modellt kapott. Kérdéseket válaszolt meg, utasításokat hajtott végre, és a környezetre vonatkozó közléseket is fogadott. Korábbi cselekvéseiről és terveiről beszélni tudott, bizonytalan értelmezésnél pedig pontosítást kérhetett. [109, 110]



61. kép Saját tárgyvilág és részcélpélda: a kék gúla a piros hasábon áll. P: piros hasáb; K: kék gúla; Z: zöld hasáb. Az ábra a szöveg példáját szemlélteti, nem SHRDLU eredeti képernyőképe és nem egy dokumentált programfutás.

A különbség ELIZA válaszsablonjaihoz képest itt a tárgyi összefüggések kezelésében látható. Saját szemléltető példánkban az asztalon áll egy piros hasáb, rajta egy kék gúla. Ha a piros hasábot szeretnénk elmozdítani, előbb a rajta álló tárgyat kell levétetni. Egy ilyen feladatban a mondat elemzése csak az első lépés. A programnak a megnevezett tárgyat azonosítania kell, majd a célhoz megfelelő cselekvéssort kell választania.

A „tedd azt a dobozba” utasításnál az „azt” jelentése a beszélgetés előzményétől függhet. Ha pedig több megfelelő tárgy van, a pontosító kérdés hasznosabb a találgatásnál. Ezek a példák a feladattípus

magyarázatai, nem Winograd közzétett párbeszédének idézetei. A kapcsolatot mutatják a nyelv, a tárgyak állapota és az elérendő cél között.

Mit nyerünk egy kis világon

SHRDLU szűk környezetében a tárgyak és a lehetséges cselekvések köre kezelhető maradt. A rendszer világmodelljének határait a készítő szabta meg. Nem kellett a konyha teljes rendetlenségéből felismernie egy bögrét, eldöntenie, kié, vagy megtudnia, sérülékeny-e. A szimulált tárgyvilág megengedte, hogy a nyelv és a cselekvés kapcsolatát sok zavaró körülménytől elkülönítve vizsgálják.

Ez értékes kísérleti egyszerűsítés. Egy vegyész is kiválaszthat olyan feltételeket, amelyek mellett egy reakció jól vizsgálható. Az eredményt azután más összetételnél vagy más hőmérsékleten is ellenőrizni kell. A gépi intelligenciában hasonló kérdés, hogy a megoldás mennyi emberi átalakítást igényel, amikor kibővítjük a környezetet.

A hétköznapi tárgyak új nehézségeket hoznak. Egy gúla tetején másképp lehet megtámasztani valamit, mint egy sík lapon; egy félig telt pohár mozgatásánál már a tartalom is számít. Nem pusztán a tárgyak száma nő. A hozzájuk tartozó tulajdonságok és a kivételek is szaporodnak. A kis világban elért összhangból ezért nem következik automatikusan a tetszőleges témáról folytatható párbeszéd.

Shakey kilép a képernyőről

Az SRI kutatói 1966 és 1972 között dolgoztak Shakey mobil robotrendszerén. A gép környezetének érzékelését és leírását tervezéssel kapcsolta össze. Útvonalat keresett és egyszerű tárgyakat rendezett át a kísérleti térben. A projekt 1984-es összefoglalója korlátozott érzékelési és környezetmodellezési képességről számol be. [111, 112]



62. kép Shakey a Computer History Museum vitrinjében, 2007. január 10-én. A robotrendszert az SRI 1966 és 1972 között kutatta. A fotó későbbi múzeumi felvétel. Fotó: Marshall Astor; a Commons-változat kivágása: Sanchom. CC BY-SA 2.0.

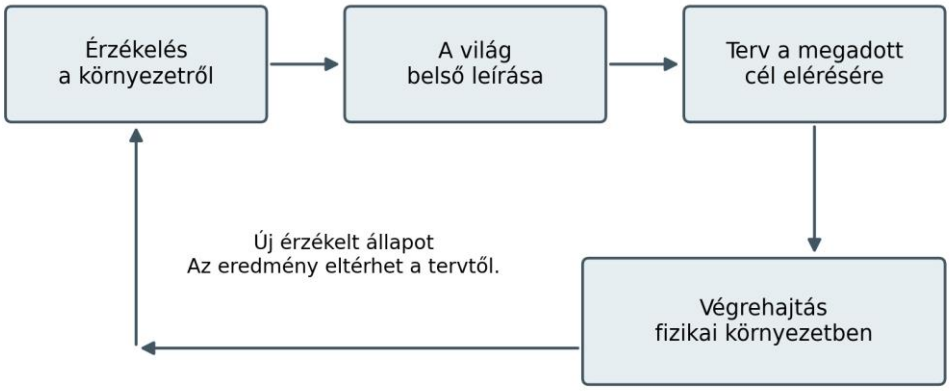
A robot kamerát, távolságérzékelést és ütközésérzékelőket használt. A számításokat külső számítógépek végezték, amelyekkel kapcsolatot tartott; a képen látható felépítmény tehát nem

foglalta magában a teljes feldolgozó rendszert. A kísérleti helyiségekben az egyszerű, jól elkülönülő tárgyak segítették az érzékelést. [112]

A környezet képe és a cselekvésre használható leírás nem ugyanaz. A kameraképen világos és sötét területek látszanak. A tervezőnek olyan állításokra lehet szüksége, mint hogy a robot egy helyiségben van, az átjáró elérhető, vagy egy doboz akadályozza a mozgást. A kutatási feladat része volt a kapcsolat megteremtése az érzékelésből származó információ és a műveletek tervezése között. [112]

A terv feltételei és a végrehajtás valósága

A Shakey-projekthez kötődik a STRIPS tervezőrendszer. Richard Fikes és Nils Nilsson 1971-es tanulmánya a cselekvéseket előfeltételekkel és a világ leírását módosító hatásokkal kezelte. Egy művelet bizonyos állításokat érvényessé tesz, másokat megszüntet. Így a kívánt állapothoz vezető műveletsor keresése formális feladattá alakítható. [113]



63. kép Az érzékelés, a világ leírása, a tervezés és a fizikai cselekvés kapcsolata. A rajz mai fogalmi összefoglalás; nem Shakey történeti hardverének vagy vezérlőprogramjának pontos kapcsolási rajza. Saját szerkesztés.

Egy mai, egyszerűsített példában a „menj át a szomszédos helyiségbe” művelethez hozzáférhető átjáró szükséges. Sikeres végrehajtás után a robot új helyen van, és a korábbi helyét jelölő állítást módosítani kell. A tervező azonban csak az általa kezelt feltételekkel számolhat. Ha a valóság eltér a leírástól, a formálisan helyes terv is kudarcot vallhat.

A fizikai környezetben a mozdulat pontatlansága is új információt teremt. Elfordulhat egy kerék, a tárgy máshová csúszhat, és a kamera következő képe eltérhet a várttól. Ezért az érzékelés, a belső leírás, a terv és a végrehajtás viszonyát együtt kell vizsgálni. A korábbi kibernetikai fejezetek visszacsatolása itt a környezetet leíró és a cselekvést tervező eljárásokkal találkozik.

ELIZA, SHRDLU és Shakey eltérő eredményt tett láthatóvá. ELIZA a társalgási forma hatását, SHRDLU a nyelv és a modellben végzett cselekvés kapcsolatát, Shakey pedig a fizikai érzékelés és a tervezés összekapcsolását tette vizsgálhatóvá. Mindegyiknél meg kell nevezni, mit adott készen az ember, mit számított ki a rendszer, és milyen körülmények között működött. Ettől a leírás pontosabb lesz, a műszaki teljesítmény pedig jobban megérthető.

Források és további olvasmányok

A források és képek számozása az 1–19. fejezetet folytatja. A magyarázó számpéldák, a magyar társalgási példa és a tárgymozgatási helyzetek saját szemléltetések, nem az eredeti programok idézett párbeszédei vagy mért eredményei.

[103] Allen Newell – Herbert A. Simon: The Logic Theory Machine – A Complex Information Processing System. RAND, P-868, 1956. DOI: 10.7249/P868. A rendszer specifikációja és a heurisztikus keresés. Kiadói adatlap

[104] Allen Newell – J. C. Shaw – Herbert A. Simon: Report on a General Problem-Solving Program. RAND, P-1584, 1958. december 30.; javítva 1959. február 9. Konferenciaváltozat: Proceedings of the International Conference on Information Processing, 1959, 256–264. A GPS felépítése és az eszköz–cél elemzés. Korabeli jelentés · Archívumi rekord

[105] Arthur L. Samuel: Some Studies in Machine Learning Using the Game of Checkers. IBM Journal of Research and Development, 3/3, 1959, 210–229. DOI: 10.1147/rd.33.0210. A két tanulási eljárás és a teljesítményre vonatkozó korabeli beszámoló. Teljes tanulmány · Kiadói rekord

[106] Frank Rosenblatt: The Perceptron – A Probabilistic Model for Information Storage and Organization in the Brain. Psychological Review, 65/6, 1958, 386–408. DOI: 10.1037/h0042519. A történeti modell és kutatási céljai. Tanulmány

[107] Herbert A. Simon – Allen Newell: Heuristic Problem Solving – The Next Advance in Operations Research. Operations Research, 6/1, 1958, 1–10. DOI: 10.1287/opre.6.1.1. Az 1957. november 14-i előadás közölt változata; a tízéves előrejelzések a 7–8. oldalon olvashatók. Kiadói rekord · A tanulmány digitális másolata

[108] Joseph Weizenbaum: ELIZA – A Computer Program for the Study of Natural Language Communication between Man and Machine. Communications of the ACM, 9/1, 1966, 36–45. DOI: 10.1145/365153.365168. A program, a script, a szövegátalakítás és a felhasználói értelmezés. A hivatkozott másolat az 1983-as újraközlésből származik. Eredeti közlés adatai · Újraközölt tanulmány

[109] Terry Winograd: Procedures as a Representation for Data in a Computer Program for Understanding Natural Language. MIT, AI-TR-235, 1971. A SHRDLU nyelvi és tárgyi modellje. MIT archívum

[110] Terry Winograd: Understanding Natural Language. Cognitive Psychology, 3/1, 1972, 1–191. DOI: 10.1016/0010-0285(72)90002-3. A rendszer részletes leírása és bemutatópárbeszéde. Kiadói rekord · Teljes tanulmány

[111] SRI: Shakey the Robot. Intézményi történeti ismertető az 1966–1972-es kutatásról, a tervezésről, az útvonalkeresésről és egyszerű tárgyak átrendezéséről. SRI történeti oldal

[112] Nils J. Nilsson szerk.: Shakey the Robot. SRI Technical Note 323, 1984. április. A projekt korábbi műszaki jelentéseinek részletei; különösen a robot, a külső számítógépek, a világmodell és az érzékelés fejezetei. Teljes jelentés · Intézményi rekord

[113] Richard E. Fikes – Nils J. Nilsson: STRIPS – A New Approach to the Application of Theorem Proving to Problem Solving. Artificial Intelligence, 2/3–4, 1971, 189–208. DOI: 10.1016/0004-3702(71)90010-5. Konferenciaváltozat: IJCAI 1971, 608–619. Kiadói rekord · Konferenciaközlés

Képforrások és felhasználási adatok

58. kép Saját szerkesztésű keresési fa. Az ágvégek száma 3 hatványa; az ábra az első két lépésszintet mutatja.

59. kép Saját egyszerűsített magyarázó ábra a [105] tanulmány alapján. A kétféle tanulást közös keretben mutatja, nem egyetlen történeti változat pontos folyamatábrája.

60. kép Saját magyarázó ábra a [108] tanulmányban tárgyalt módszerről; a felhasználói értelmezést is jelöli.

61. kép Saját tárgyvilág és cselekvési példa a nyelvi utasítás és a részcélok kapcsolatáról. Nem történeti képernyőmentés.

62. kép Shakey, 2007. január 10. Marshall Astor fényképe; a Commons-fájlt Sanchom vágta ki. A letöltött változat további módosítás nélkül szerepel. CC BY-SA 2.0. Kép és metaadatok · Licenc

63. kép Saját fogalmi összefoglaló a [112–113] források témájához. Nem a történeti rendszer pontos működési diagramja.

V. rész – A nagy ígéretektől az AI telekig

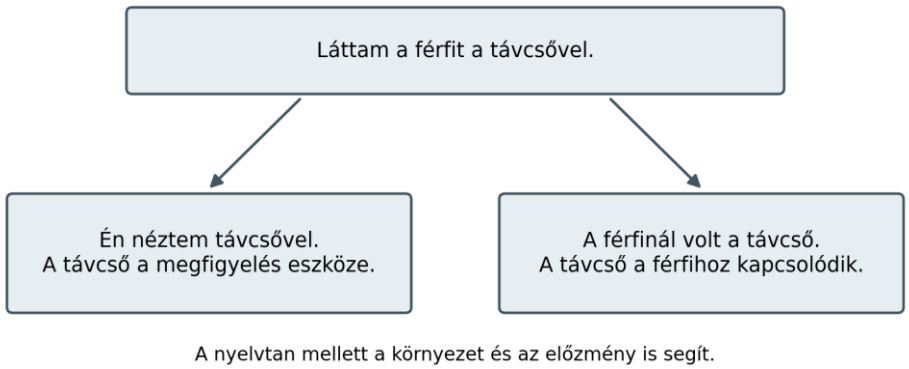
A korai programok megmutatták, hogy a gép bizonyíthat, játszhat és egyszerű világokban cselekedhet. A következő kérdés az lett, mennyit őriz meg ebből a képességből, amikor a világ bonyolultabbá válik. Az AI történetének visszaeséseit ez a nehézség is alakította. A kutatók közben egyre használhatóbb rendszereket építettek jól körülhatárolt feladatokra, de a tudományos eredmény, az üzemi megbízhatóság és a nyereséges vállalkozás nem mindig találkozott.

22. Miért bizonyult nehéznek az, ami nekünk könnyű?

Egy gyerek felismeri a kabátját akkor is, ha összegyűrve hever a széken. Egy ismerős hangot zajban is azonosítunk, és többnyire megértjük a hiányos mondatot. Ezeket a képességeket természetesnek vesszük. A korai mesterséges intelligencia kutatói viszont azt tapasztalták, hogy a szabályos matematikai feladat sokszor könnyebben leírható gép számára, mint a mindennapi tájékozódás. A 15. fejezet humanoidokról szóló kitekintése már bemutatta ennek testi oldalát. Itt azt vizsgáljuk, miért okozott hasonló nehézséget az információ értelmezése és a megfelelő megoldás kiválasztása.

A mondatban nincs benne minden

A nyelv megértéséhez nem elegendő a szavakat szótárban megkeresni. A mondat szerkezete, az előzmények és a világról szerzett tudás is alakítja az értelmezést. Saját példánkban: „Láttam a férfit a távcsővel.” Vajon én használtam távcsövet, vagy a férfinál volt a távcső? A szavak azonosak, mégis két különböző helyzetet írhatnak le. A mondat folytatása döntő lehet: „A nyakában lógott” más értelmezést támogat, mint az, hogy „Szabad szemmel nem vettem volna észre.”



64. kép Egy magyar mondat két lehetséges értelmezése. A folytatás és a helyzetről szerzett tudás segíthet választani közöttük. Saját nyelvi példa és magyarázó ábra, nem egy történeti fordítóprogram eredménye.

A gépnek tehát lehetőségeket kell felállítania, majd ki kell választania a helyzethez illőt. A megfelelő tudás egy része a szövegen kívül található. Egy ember tudja, mire használják a távcsövet, hogyan lehet viselni, és mit jelent távolról megfigyelni valakit. Ha ezek az összefüggések hiányoznak a rendszerből, a nyelvtani elemzés önmagában nem oldja meg a feladatot.

Ez a nehézség a fordításban is jelentkezik. A célnyelv más szórendet vagy olyan nyelvtani különbséget kívánhat, amelyhez előbb el kell dönteni az eredeti mondat értelmét. A szavak egymás utáni

helyettesítése ezért nem elég. A fordítandó dokumentum, a szakterület és a beszélő szándéka egyaránt számíthat.

A látványtól a tárgyig

A kamera fényértékeket rögzít. A „szék”, „ajtó” vagy „kéz” fogalma nincs külön beírva a kép megfelelő sarkába. Egy tárgy képe változik a nézőponttal, megvilágítással és takarással. Ha a kabát gyűrött, egyes jellegzetes részei eltűnhetnek. A felismeréshez mégis ugyanahhoz a tárgyhoz kell kapcsolni a különböző látványokat.

A mozgás újabb függőségeket hoz létre. A gépnek a felismert tárgy helyét a saját lehetséges cselekvéseihez kell viszonyítania. Egy pohár azonosítása más feladat, mint megállapítani, elérhető-e anélkül, hogy valamit fellöknénk. A korábbi fejezetekben SHRDLU kész tárgyvilágot kapott, Shakey pedig egyszerűsített fizikai környezetben dolgozott. Ezekben éppen az tette vizsgálhatóvá a problémát, hogy sok bizonytalanságot kizártak.

A Moravec-paradoxonnak nevezett felismerés e különbségre is figyelmeztet: a tudatosan nehéznek érzett feladat nem feltétlenül a legnehezebben gépesíthető. Az emberi érzékelés és mozgás mögött sok olyan feldolgozás áll, amelynek eredményét készen kapjuk. Hans Moravec későbbi visszatekintése ezt a robotika tapasztalatai alapján tárgyalta. [114]

A józan ész rejtett előfeltevései

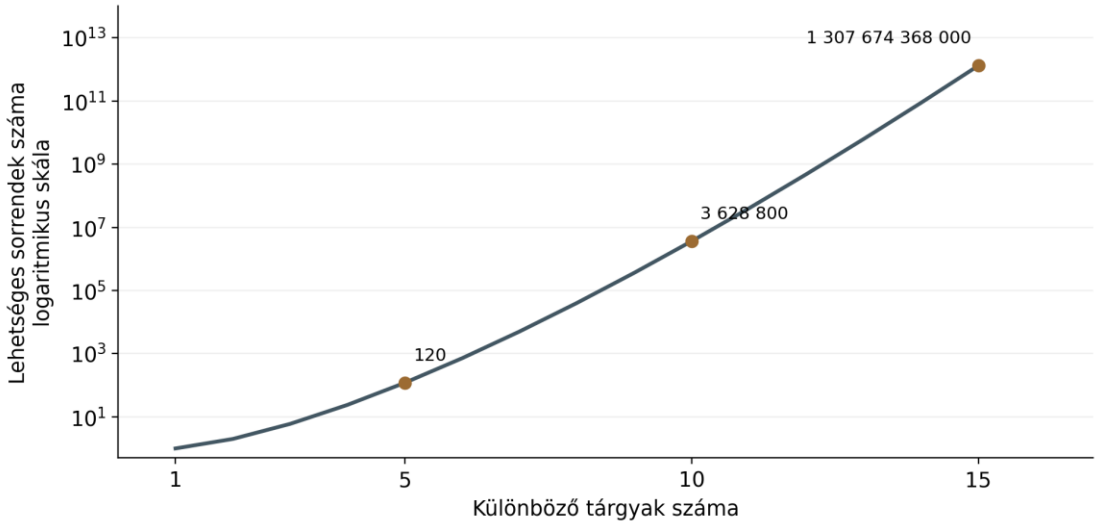
Ha valakit megkérünk, vigyen át egy poharat a másik asztalra, rendszerint nem tesszük hozzá, hogy ne fordítsa fejre, ne ütközzön a többieknek, és ne hagyja a levegőben. A kérés rengeteg hallgatólagos feltételt tartalmaz. Ezek egy része fizikai ismeret, más része a helyzet és az emberi szokások ismerete.

John McCarthy már az 1958-ban előadott Programs with Common Sense munkájában olyan program lehetőségét vizsgálta, amely a világról közölt állításokból következtetni tud. A cél a tudás használható leírása volt, nem pusztán egy-egy rejtvény megoldása. [115] A nehézség abban rejlik, hogy az előfeltevések körét ritkán látjuk előre teljes egészében.

A józan ész szabályai gyakran kivételeket is megengednek. A tárgyak többnyire ott maradnak, ahová letettük őket, de valaki elmozdíthatja őket. A csésze rendszerint megtartja a folyadékot, de lehet repedt. Egy szűk világban elég lehet néhány pontos állítás. Nyitott környezetben azt is kezelni kell, mikor elégtelenek a korábbi feltevéseink.

Amikor túl sok a lehetőség

A keresés nemcsak azért nehéz, mert sok adat van. A lehetőségek egymással kombinálódnak. Saját, egyszerű számpéldánkban öt különböző tárgy sorrendbe rakásának 120 módja van. Tíz tárgynál 3 628 800, tizenötnél 1 307 674 368 000. Itt minden tárgy egyszer szerepel, és minden sorrendet különbözőnek tekintünk: a lehetőségek száma a tárgyszám faktoriálisa.



65. kép Különböző tárgyak lehetséges sorrendjeinek száma, n! szerint. A függőleges tengely logaritmikus: azonos távolság többszöröződést jelent. Saját matematikai szemléltetés; nem állítja, hogy minden feladat megoldásához valamennyi sorrendet meg kell vizsgálni.

Ez nem azt jelenti, hogy minden ilyen feladatnál az összes sorrendet ki kell próbálni. A jó eljárás korlátozásokat használhat, részekre bonthatja a feladatot, vagy korán kizárhat sok lehetőséget. Az ábra azt mutatja, mi történne a megkülönböztetés nélküli felsorolással. A számítógép gyorsulása segít, de önmagában nem feltétlenül ellensúlyozza az ilyen növekedést.

A kutatók így egyre inkább az ábrázolás és a keresést irányító tudás jelentőségével szembesültek. Mit tekintsünk egyáltalán külön lehetőségnek? Melyik részlet lényeges? Milyen információval zárhatunk ki egy ágat? A következő korszak egyik válasza az lett, hogy a gép általános képességeinek reménye mellé sok, gondosan kiválasztott szaktudást kell adni.

23. Az első AI-tél

Az „AI-tél” utólagos gyűjtőnév. Olyan időszakokat jelöl, amikor az eredményekhez képest túl nagynak bizonyultak az ígéretek, megingott a bizalom, és egyes programok támogatása visszaesett. Az első ilyen lehűlést többnyire az 1960-as évek végétől az 1970-es évek egy részéig terjedő folyamatokhoz kapcsoljuk. A dátumok országonként és kutatási irányonként eltérnek. Nem volt egyetlen nap, amikor az egész világ leállította az AI kutatását.

A fordítás ígéretétől az értékelésig

A hidegháború idején az idegen nyelvű tudományos és műszaki dokumentumok gyors feldolgozása fontos cél lett. A számítógépes fordítás vonzereje könnyen érthető: ha a szöveg nagy tömegben érkezik, a gép gyorsasága komoly előnynek látszik. Az előny azonban csak akkor használható, ha az eredmény érthető és helyes, illetve javítása nem igényel aránytalan emberi munkát.

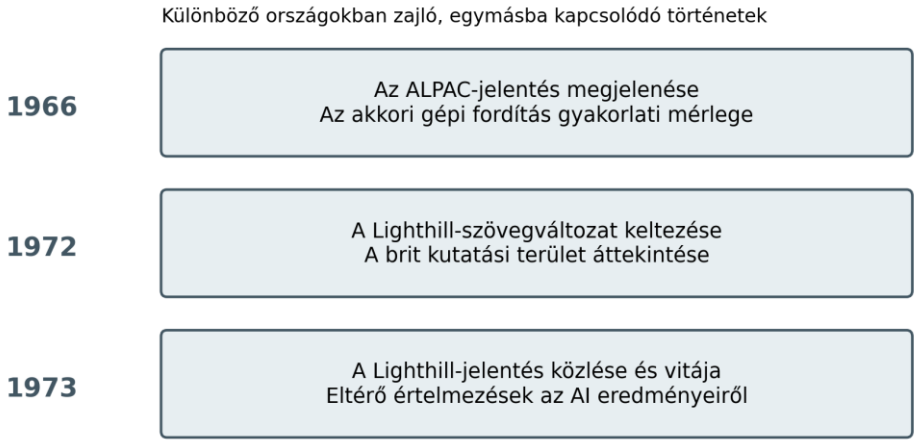
Az amerikai ALPAC bizottság 1966-ban megjelent Language and Machines jelentése a fordítás egész folyamatát vizsgálta. Nem csupán azt kérdezte, képes-e a gép mondatokat előállítani, hanem a minőséget, költséget, átfutási időt és a tényleges igényt is értékelte. A bizottság az akkori hasznos gépi fordítás kilátásait rendkívül kedvezőtlennek ítélte. Ugyanakkor a számítógépes nyelvészet támogatását és a fordítók munkáját segítő eszközök fejlesztését ajánlotta. [116]

A különbség lényeges. A jelentés nem bizonyította, hogy a nyelv gépi kezelése elvileg lehetetlen. Az akkori megoldások gyakorlati értékéről és a további támogatás irányáról foglalt állást. Egy tudományos kérdés és egy költségvetési döntés itt találkozott, de nem vált azonossá.

A kedvezőtlen mérleg következménye

John Hutchins fordítástörténeti elemzése az ALPAC utáni időszakot az amerikai gépi fordítás kutatásának erős visszaesésével kapcsolja össze. A folyamatot ugyanakkor nem lehet minden országra kiterjedő, teljes megszűnésként leírni. A jelentés befolyása mellett a korábbi várakozásokat, az egyes műhelyek helyzetét és a gyakorlati rendszerek további életét is vizsgálni kell. [117]

A csalódásnak gazdasági oka is volt. Ha egy fordítást utólag szakembernek kell szinte újra elkészítenie, az automatikus előállítás gyorsasága félrevezető mérőszám. Saját példánkban egy perc alatt elkészülő nyersfordítás csak akkor gyorsítja a munkát, ha ellenőrzésével együtt is rövidebb a teljes folyamat. Ugyanez a gondolat később más AI-alkalmazások értékelésében is visszatér.

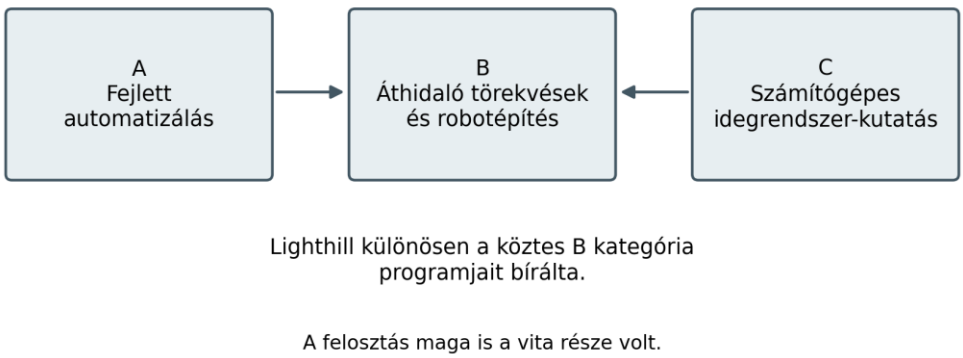


66. kép Az ALPAC- és a Lighthill-történet három dátuma. Az ábra időrendi felsorolás, nem finanszírozási grafikon. Az 1972-es keltezést megkülönbözteti az 1973-as közléstől. Saját szerkesztés a [116–118] források alapján.

Lighthill vizsgálata

Nagy-Britanniában a Science Research Council James Lighthillt kérte fel a terület áttekintésére. A jelentés 1973-ban vált ismertté a tanács kiadványában; az elérhető szövegváltozat 1972. júliusi keltezést visel. Lighthill saját maga is hangsúlyozta, hogy a területen kívülről érkező kutató személyes áttekintését adja. [118]

Három kategóriát különített el. Az A kategória a fejlett automatizálásé, a C az idegrendszer működését vizsgáló számítógépes kutatásoké volt. A kettő közötti B kategóriában helyezte el az áthidaló, robotépítéshez és általános intelligenciához kötődő törekvéseket. Különösen ezeket bírálta. A kombinatorikus robbanást az ígéretek teljesülését akadályozó központi nehézségnek tekintette. [118]



67. kép Lighthill A, B és C kategóriáinak mai összefoglalása. A köztes B kategória értékelése a vita fontos része volt; a rajz nem általánosan elfogadott, semleges felosztásként mutatja be. Saját szerkesztés, nem korabeli ábra.

A bírálat nem pusztán annyi volt, hogy a gépek lassúak. Lighthill azt kérdőjelezte meg, mennyire vihetők át a szűk feladatok eredményei tágabb környezetbe. Ez jogosan vizsgálható kérdés. A válasz

azonban függ attól is, milyen új módszereket tartunk lehetségesnek, és milyen időtávot tekintünk indokoltnak.

A jelentés nem maradt válasz nélkül

John McCarthy részletes bírálatában vitatta Lighthill felosztását és több következtetését. Kiemelte a tudás gépi ábrázolásának önálló kutatási problémáját, valamint az AI-nak tulajdonítható eredményeket. Ugyanebben az írásban az AI-kutatók egyes túlzó előrejelzéseit és a pusztán látványos bemutatókra épülő munkákat is kifogásolta. [119] A vita ezért nem írható le egyszerűen a „hívők” és a „kétkedők” összecsapásaként.

A brit parlament 2018-as történeti áttekintése külön óvatosságra int a finanszírozás történetében. Nem tekinti egyértelműen bizonyítottnak azt az elterjedt állítást, hogy maga a Lighthill-jelentés okozta a brit AI-támogatás visszavágását. [120] A jelentés és a bizalom megingása közötti kapcsolat fontos, de az intézményi döntéseket nem helyettesítheti egyetlen ismert dokumentum neve.

Mit jelentett a lehűlés

Egy megszüntetett támogatás embereket, számítási lehetőséget és kutatási időt is elvonhat. A nagy célok rövid távú számonkérése így visszahat arra, milyen kérdésekkel lehet foglalkozni. A kutatók kisebb, könnyebben mérhető feladatokat választhatnak, vagy más megnevezés alatt folytathatják munkájukat. Ez a finanszírozás lehetséges működésmódjának magyarázata; az egyes műhelyek történetét külön kell dokumentálni.

Az első AI-tél idején is születtek eredmények. A következő fejezetben tárgyalt szakértői rendszerek gyökerei éppen ebbe az időszakba nyúlnak. A visszaesés az általános és gyors áttörésbe vetett reményt érintette. A szűkebb feladatra összegyűjtött tudás közben új irányt adott a kutatásnak.

24. A szakértői rendszerek felemelkedése

Egy tapasztalt vegyész, orvos vagy műszaki szakember sok olyan összefüggést ismer, amelyet a kezdő még nem lát. A szakértői rendszerek ezt a tudást próbálták számítógépes formában felhasználni. A cél jól meghatározott területen hasznos tanács, értelmezés vagy döntési javaslat előállítása volt. A kutatási hangsúly a mindent megoldó általános módszer kereséséről egyre inkább a feladathoz szükséges tudás összegyűjtésére is áttevődött.

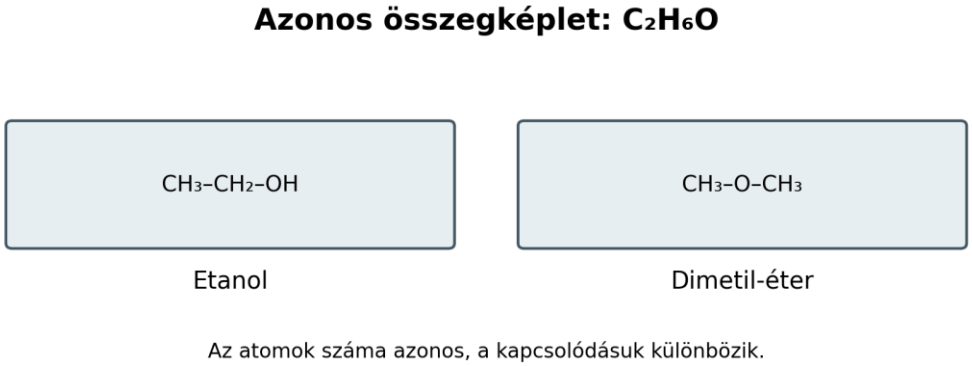
DENDRAL és a molekulák lehetséges szerkezete

A Stanfordon, az 1960-as évek közepétől fejlődő DENDRAL munkájában Edward Feigenbaum, Bruce Buchanan és Joshua Lederberg meghatározó szerepet játszott. A program szerves kémiai szerkezetekre keresett magyarázatot többek között tömegspektrometriás adatok alapján. A lehetséges szerkezeteket kémiai korlátozások és a mérési adatok értelmezésére szolgáló ismeretek segítségével szűkítette. [121]



68. kép Edward A. Feigenbaum későbbi portréja, a forrás szerint 1994. december 1-jéről. A DENDRAL és a szakértői rendszerek egyik meghatározó kutatója. United States Air Force / Wikimedia Commons. Amerikai szövetségi kormányzati felvétel; közkincs.

A tömegspektrumban a molekulából képződő ionok tömeg/töltés arány szerinti eloszlását vizsgáljuk. A szerkezeti következtetéshez több csúcs együttes mintázata és a lehetséges fragmentációk ismerete lehet fontos. Az összegképlet önmagában nem mindig ad egyetlen szerkezetet. Saját egyszerű példánkban az etanol és a dimetil-éter összegképlete egyaránt C_2H_6O, a kötések elrendezése mégis eltér.



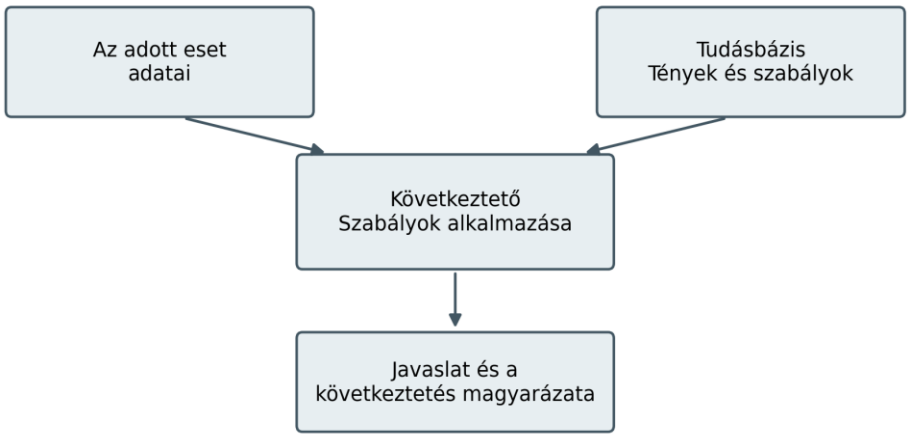
69. kép Etanol és dimetil-éter: azonos összegképlet, eltérő atomkapcsolódás. A tömegspektrometriás szerkezetkövetkeztetés feladattípusának egyszerű magyarázata. Saját kémiai szemléltetés; nem egy DENDRAL-futtatás dokumentuma vagy mért spektrum.

Az ábra nem egy történeti DENDRAL-futtatás eredménye. A feladattípus alapját mutatja: adott összetételből több szerkezeti lehetőség adódhat, és további tudásra van szükség a megkülönböztetésükhöz. A programnak nem kellett minden elképzelhető atomkapcsolatot egyformán komolyan vennie. A kémiai tudás már a keresés terét is alakíthatta.

Feigenbaum és munkatársai 1970-es tanulmánya kifejezetten az általánosság és a szakismeret kapcsolatát vizsgálta. A DENDRAL tanulsága szerint a jó teljesítményhez a keresési eljárás és a területi szakértelem összekapcsolása kellett. [121] A tudás nem egyszerűen adat volt, amelyet a program a végén elővett: a lehetséges magyarázatok kiválasztását segítette.

Mitől szakértő egy rendszer

A klasszikus szabályalapú megoldásban elkülöníthető a tudásbázis és az azt felhasználó következtető eljárás. A tudásbázis állításokat és feltételes szabályokat tartalmaz. A következtető az adott eset adataihoz illeszkedő szabályokat keresi, új állításokat vezet le, vagy további kérdést tesz fel. A MYCIN felépítésében a konzultáció, a magyarázat és a tudás módosításának támogatása külön feladatként jelent meg. [122]



70. kép Egy szabályalapú szakértői rendszer néhány alapvető feladata. Az esetadatok és a tudásbázis együtt szolgálják a következtetést. Saját fogalmi ábra, nem a MYCIN teljes programarchitektúrájának másolata.

A szétválasztás gyakorlati előnye, hogy a szakismeret egy része a teljes program újraírása nélkül változtatható. Ettől még a tudás és a program viselkedése összefügg. Egy új szabály más szabályokat is működésbe hozhat, vagy megváltoztathatja a kérdések sorrendjét. A rendszer minőségét a részek együttes működése dönti el.

MYCIN és a bizonytalan következtetés

Edward Shortliffe és munkatársai az 1970-es években a MYCIN-t fertőző betegségekhez kapcsolódó orvosi konzultáció támogatására fejlesztették. A rendszer kérdésekkel gyűjtött esetadatokat, majd szabályok alapján javaslatokat készített. A kutatás egyik fontos kérdése az volt, hogyan lehet hiányos vagy bizonytalan információval dolgozni. [122]

A MYCIN bizonyossági tényezőket használt az egyes következtetések alátámasztottságának kezelésére. Ezeket nem helyes egyszerűen egy betegség százalékos valószínűségével azonosítani. A módszer saját feltevésekkel és összekapcsolási szabályokkal dolgozott; előnyeit és korlátait a fejlesztők is elemezték. [123]

Az emberi szakértő sem mindig biztos. Különbség van azonban a bizonytalanság kimondása és megalapozott kezelése között. A gép által kiírt szám önmagában nem garancia. Tudni kell, milyen adatokra, szabályokra és feltevésekre épült. A rendszer magyarázó képessége ezért a szakmai ellenőrzés eszköze is lehetett.

A számítógépek megrendelése is szakértelmet kívánt

A Digital Equipment Corporation 1980-tól használta az R1, más néven XCON rendszert VAX-számítógépek konfigurálására. A megrendelt alkatrészekből működő összeállítást kellett kialakítani: a megfelelő modulok, szekrények és kapcsolatok kiválasztása sok részletszabályt igényelt. Judith Bachant és John McDermott 1984-es beszámolója a tudásbázis bővítését és az üzemi használat tapasztalatait ismertette. [124]

Ez a példa azért fontos, mert valódi vállalati feladatot kötött össze a szabályalapú módszerrel. A sikert nem általános társalgásban kellett kimutatni. A jó konfiguráció, az észlelt hiány vagy a kijavított összeállítás közvetlenül kapcsolódhatott a munkafolyamathoz.

A DEC technológiaátadásról szóló beszámolója ugyanakkor azt is rögzítette, hogy az egyetemről átvett rendszert javítani, bővíteni és vállalaton belül gondozni kellett. [125] A szakértői rendszer lehetett hasznos termelőeszköz, de nem vált önmagát fenntartó szakemberré. A következő nehézség éppen az lett, hogyan kerül a tudás a gépbe, és hogyan marad használható.

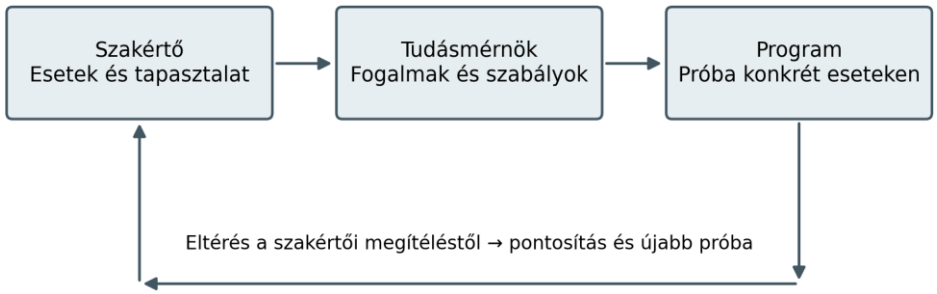
25. Tudás palackba zárva

A szakértői rendszer tudásbázisa első pillantásra tartós gyűjteménynek látszik. Ha egyszer leírtuk a szakértő szabályait, a gép újra és újra használhatja őket. A tapasztalat azonban azt mutatta, hogy a tudás megszerzése nem egyszeri másolás. Meg kell találni a fontos fogalmakat, ki kell mondani a hallgatólagos feltételeket, és ellenőrizni kell, hogy a szabály valóban azt jelenti-e a gépnek, amit a szakember gondolt.

A szakember nem mindig a saját szabályait látja

Egy tapasztalt szerelő gyorsan észrevehet egy szokatlan hangot. Egy vegyész gyanúsnak találhat egy mérési görbét. Ha megkérdezzük, miért, először csak annyit mond: „Ez nem olyan, mint máskor.” A helyes ítéletből még nem következik, hogy az azt létrehozó megkülönböztetések könnyen felsorolhatók.

A tudásmérnök feladata ezért a szakértővel folytatott közös munka lett. Eseteket gyűjtött, fogalmakat tisztázott, és a program eredményeit visszavitte ellenőrzésre. A MYCIN-csoport Knowledge Engineering fejezete ezt átalakítási és fokozatos fejlesztési folyamatként írta le. A szakértelem forrása lehetett ember, szöveg vagy adat, de a gép számára alkalmas alakot külön kellett kialakítani. [126]



A tudás átvitele fokozatos közös munka.

71. kép A szakértői tudás gépi leírásának fokozatos folyamata. A program konkrét eseteken mutatott eltérései újabb pontosítást indíthatnak. Saját szerkesztésű magyarázó ábra a [126] témájához.

A folyamat során az is kiderülhetett, hogy két szakértő másra használja ugyanazt a szót. A „magas”, „szokatlan” vagy „súlyos” minősítés mögött eltérő küszöb állhatott. A rendszerhez azt is rögzíteni kellett, mi számít mérésnek, becslésnek és hiányzó adatnak. A tudás gépi leírása így részben a szakterület fogalmainak rendezését kívánta.

A szabály rejtett érvényességi tartománya

Vegyünk egy saját, műszaki szemléltető példát. Egy ellenőrző szabály magas hőmérséklet és növekvő nyomás esetén hibagyanút jelez. Ha a rendszer nem tudja, hogy a berendezés éppen indítási

szakaszban van, téves riasztást adhat. A szabályhoz tehát üzemállapotot kell rendelni. Később kiderülhet, hogy más anyaggal vagy más érzékelővel további feltétel is szükséges.

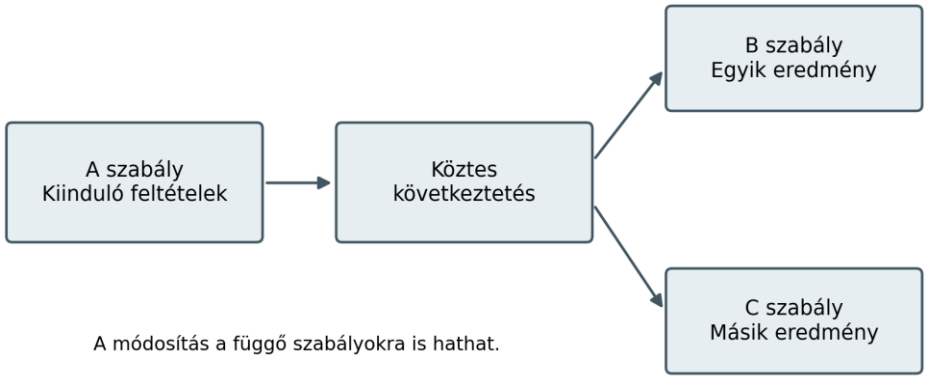
Ez nem működő technológiai biztonsági előírás, hanem a tudásbázis építésének példája. A feltételek pontosítása nem puszta szőrszálhasogatás. Az emberi szakértő sokszor automatikusan figyelembe vesz körülményeket, amelyek a kérdezőnek és a programozónak fel sem tűnnek. A gép számára viszont a hiányzó feltétel hiányzó tudás marad.

A szabály megfelelő működése ezért az adat jelentésétől is függ. Mit mér a műszer? Mikor hitelesítették? Milyen egységben érkezik az érték? Egy helyes következtető eljárás sem tudja kijavítani azt az esetet, amelyet eleve rosszul írtunk le. A szabályokra épülő tanácsadás és az adatminőség kérdése már a kezdetektől összetartozott.

Egy új szabály régi válaszokat is megváltoztathat

A tudásbázis nem egymástól független mondatok doboza. Egy szabály következtetése egy másik szabály feltétele lehet. Az első módosítása így távoli eredményt is megváltoztathat. Két szabály ellentmondó következtetésre juthat, vagy egy lényeges esethez egyik sem illeszkedik.

A MYCIN-kötet teljességről és következetességről szóló fejezete különbséget tett a tudásbázis helyessége és a tudást alkalmazó program helyes működése között. A fejlesztők automatizált ellenőrző segítséget is vizsgáltak a hiányok és ellentmondások feltárására. [127] A szaktudás összegyűjtése mellett tehát a szaktudást kezelő rendszer vizsgálata is önálló feladattá vált.



72. kép Szabályok közötti függés saját egyszerű példában. A köztes következtetés több további szabály feltétele lehet. A nyilak függőséget jelölnek, nem időbeli munkasorrendet; egy módosítás több kimenetre is hathat.

Saját ábránkban az egyik következtetés további szabályok bemenete lesz. A nyilak függést jelölnek, nem időbeli munkafolyamatot. Egyetlen átírás több válaszra hathat. Emiatt érdemes korábban kipróbált eseteket újra lefuttatni: megmaradt-e a helyes eredmény, és valóban kijavult-e az a hiba, amely miatt a változtatás történt?

A magyarázat is tudásra szorul

A szabályokból levezetett válasz útja gyakran megmutatható. A felhasználó láthatja, mely adat és mely szabály vezetett a következtetéshez. Ez valódi előny az ellenőrzésben. A szabálylánc felsorolása azonban nem mindig adja meg, miért éppen ezek a szabályok írják le helyesen a világot.

A MYCIN magyarázó eljárásait külön kutatták. [128] Egy rendszer elmondhatja, milyen lépéseket követett, miközben a felhasználó inkább az oksági hátteret szeretné megérteni. Miért tekintjük lényegesnek ezt a tünetet vagy jelzést? Milyen tapasztalat támasztja alá a szabályt? A következtetés nyomvonala és a szakmai indoklás egymáshoz kapcsolódó, de külön kérdés.

A tudás időközben változik

Új termék, új mérési módszer vagy megváltozott eljárás is elavulttá tehet egy korábbi szabályt. Az emberi szakember szakirodalmat olvas, kollégákkal egyeztet és új eseteket lát. A tudásbázis akkor követi ezt, ha valaki beépíti és ellenőrzi a változásokat.

A tudás megszerzésének szűk keresztmetszetére már a DENDRAL és a MYCIN fejlesztői is felhívták a figyelmet. [126] A „palackba zárás” ezért csak a cím szemléletes hasonlata. A használható tudásbázis gondozott állapot, amelyhez felelős emberek és tartós munkafolyamat kellenek.

A megoldás nem feltétlenül a szabályok elhagyása volt. Bizonyos feladatoknál a közös fogalomtár, az adatellenőrzés és a világos alkalmazási határ sokat segíthetett. A kérdés az lett, hogy a megszerzett haszon indokolja-e a felépítés és a folyamatos gondozás költségét. Ez már közvetlenül a következő AI-tél üzleti problémáihoz vezetett.

26. A második AI-tél

Az 1980-as évek szakértőirendszer-lendülete nyomán külön vállalatok kínáltak fejlesztőeszközöket, szaktudást és speciális számítógépeket. A kereskedelmi várakozások egy része gyorsabban nőtt, mint a tartósan fenntartható alkalmazások köre. Az évtized végétől az 1990-es évek elejéig több ilyen piac és vállalkozás súlyos nehézségbe került. Ezt nevezzük többnyire második AI-télnek. A megnevezés most is eltérő történeteket fog össze.

A prototípus és az üzemi rendszer közötti távolság

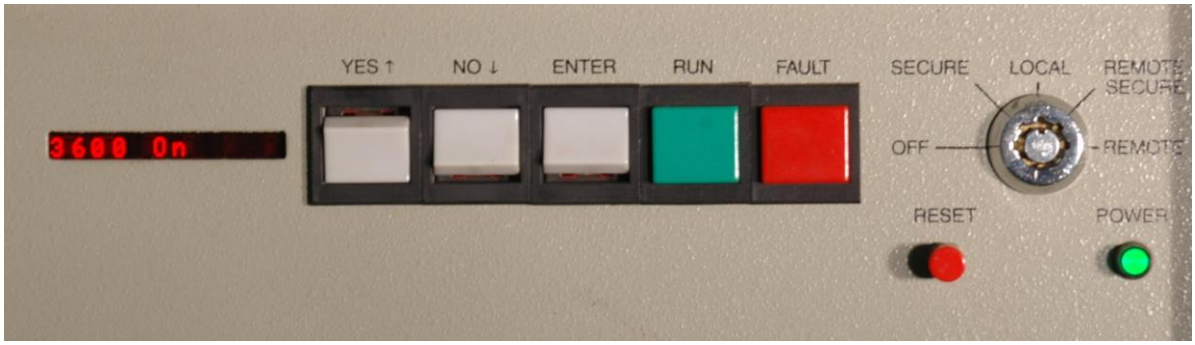
Egy sikeres bemutató után még sok munka maradhat. Az alkalmazásnak kapcsolódnia kell az adatforrásokhoz, el kell férnie a munkafolyamatban, és hiba esetén is kezelhetőnek kell maradnia. A felhasználónak időt kell nyernie vele. Ha ugyanazt az információt több helyre kell begépelnie, a jó tanács ellenére is elutasíthatja.

A MYCIN története pontos példát ad erre. Fejlesztői 1984-ben leírták, hogy a rendszert nem használták rutinszerű betegellátásban. A bevezetés akadályai között az intézményi adatkapcsolatokat, a konzultációhoz szükséges külön erőfeszítést és a szolgáltatás megbízhatóságát is tárgyalták. A tudásbázis idővel elavult. [129] A történetből ezért nem szabad széles körben alkalmazott „géporvost” alkotni.

Az üzleti sikerhez a számítás helyessége mellett szervezeti működőképesség kellett. Ki javítja a szabályt? Ki fizeti az új esetek feldolgozását? Ki vállal felelősséget az eredményért? Ezek nem utólag hozzátoldott mellékkérdések: meghatározzák, hogy a rendszer egyáltalán használatban marad-e.

A különleges gépek vonzereje

A Lisp nyelvet sok AI-műhely használta jelképes adatok és összetett programok kezelésére. A Lisp-gépek ehhez igazított hardvert és gazdag, interaktív fejlesztőkörnyezetet kínáltak. A Symbolics 1985-ös műszaki összefoglalója például a 3600-as család címkézett architektúráját, virtuális memóriáját és egységes Lisp-környezetét hangsúlyozta. [130]



73. kép Egy Symbolics 3600 Lisp-gép előlapjának részlete a Retro-Computing Society of Rhode Island gyűjteményéből, 2008. szeptember 20-i felvételen. Nem a teljes munkaállomás látható. Fotó: Dave Fischer (Shieldforyoureyes), CC BY-SA 3.0.

A specializáció műszaki előnyt adhat, amíg a feladatban nyújtott többlet indokolja az árat és a külön környezet fenntartását. Amikor az általános munkaállomások javulnak, a vevő már a teljes költséget

és a meglévő rendszereivel való együttműködést hasonlítja össze. Nem szükséges, hogy az olcsóbb gép minden részletében jobb legyen: az is elég lehet, ha a feladathoz megfelelő teljesítményt ad.

Az olcsóbb munkaállomások a Symbolics versenytársai lettek. [131] Ez az üzleti történet nem bizonyítja egy programozási nyelv vagy a szimbolikus következtetés tudományos alkalmatlanságát. Egy hardverpiac elvesztése és egy kutatási módszer érvényessége eltérő kérdés.

A visszaesésnek dokumentált következményei voltak

A Symbolics egyik későbbi vállalati beadványa szerint a társaság 1993. január 27-én kezdeményezett amerikai Chapter 11 eljárást. Az irat az 1995-ös eszközértékesítést és a további rendezést is rögzíti. [132] A nagy reményekkel indult különlegesgépes üzletnek tehát konkrét vállalati és pénzügyi következményekkel kellett szembenéznie.

Az elvárások kockázata már a fellendülés közben is szóba került. Az AAAI 1984-es konferenciáján rendezett, 1985-ben közölt The Dark Ages of AI panel a lehetséges csalódást és visszaesést vitatta. [133] A későbbi nehézségeket így nem kizárólag utólag ismerték fel. A hasznos példák és a köréjük épített általános várakozások közötti távolság már akkor kérdéseket vetett fel.

A szakértői rendszerek fenntartási igénye ezt a távolságot növelhette. Az XCON sikere mellé vállalati tudásbázis-gondozás társult, ahogyan azt a DEC saját technológiaátadási beszámolója is bemutatta. [125] Egy rendszer nem attól lesz olcsó, hogy a következtető program automatikusan fut. Az összegyűjtött tudás előállítása és karbantartása továbbra is emberi munkát igényelhet.

Ami a tél idején is fejlődött

A visszaeséstől függetlenül fontos tanulási módszerek fejlődtek. David Rumelhart, Geoffrey Hinton és Ronald Williams 1986-os tanulmánya nagy figyelmet irányított a többrétegű neurális hálózatok hibavisszaterjesztéssel történő tanítására. [134] A közbülső kapcsolatok módosítása lehetőséget adott olyan belső leírások kialakulására, amelyeket nem egyenként írt le a programozó. A módszernek voltak előzményei; az 1986-os publikáció nem a teljes eljárástörténet kezdete.

Ross Quinlan ugyancsak 1986-ban közölt munkája az ID3 döntésfatanulást ismertette. A rendszer példákból épített osztályozó szerkezetet. [135] Christopher Watkins és Peter Dayan 1992-es Q-learning tanulmánya pedig meghatározott feltételek mellett vizsgálta, hogyan tanulhatók meg az állapotokhoz és cselekvésekhez tartozó értékek. [136] Ezek eltérő tanulási kérdések voltak; egyik sem jelentett önmagában általános emberi intelligenciát.



74. kép Három kutatási eredmény az 1986–1992 közötti időszakból. A felsorolás a kutatás folytonosságát szemlélteti, nem azonosítja a három módszert és nem ad a két AI-télnek egységes dátumhatárt. Saját szerkesztés a [134–136] alapján.

A felsorolt munkákból nem következik, hogy a későbbi áttörések már készen álltak. A megfelelő adatok, számítási lehetőségek és gyakorlati eljárások kialakulása további munkát kívánt. Azt viszont jól mutatják, hogy egy üzleti vagy finanszírozási visszaesés alatt is születhetnek később jelentőssé váló eredmények.

Mennyire volt valóban tél

Edward Feigenbaum egy 2019-es visszaemlékezésben magát az AI-telek elbeszélését is vitatta. Arra hívta fel a figyelmet, hogy a visszaesésként emlegetett időszakokban más kutatási irányok erősödtek meg. [137] Ez résztvevői értelmezés, nem a veszteségek cáfolata. A megszűnő vállalkozás és a közben fejlődő módszer egyszerre része a történetnek.

A második AI-tél ezért akkor írható le pontosan, ha megnevezzük, minek esett vissza a támogatása vagy piaca, és mi folytatódott mellette. A szabályokba foglalt tudás sok feladatnál használható maradt, miközben a példákból tanuló módszerek új lehetőségeket nyitottak. A következő korszak ezekből a különböző előzményekből építkezett.

Források és további olvasmányok

A hivatkozások és képek számozása az 1–21. fejezetet folytatja. A forrásjegyzék elkülöníti a korabeli jelentéseket, a fejlesztők beszámolóit és a későbbi értelmezéseket. A nyelvi, kémiai és műszaki példák saját szemléltetések; nem eredeti programfuttatások idézetei.

[114] Hans Moravec: Rise of the Robots. Scientific American, 1999. december, 124–135. Résztvevői visszatekintés a robotika érzékelési és cselekvési nehézségeire. A cikk jövőre vonatkozó jóslatait a fejezet nem veszi át. Szerzői archívum

[115] John McCarthy: Programs with Common Sense. Az 1958. decemberi teddingtoni Mechanisation of Thought Processes konferencián előadott munka, a konferencia közleményeiben megjelent változattal. A hétköznapi tudás leírásának korai kutatási programja. Szerzői oldal és tanulmány

[116] National Research Council: Language and Machines – Computers in Translation and Linguistics. ALPAC, Publication 1416, Washington, 1966. DOI: 10.17226/20813. Különösen a gépi fordítás helyzetét és a támogatási ajánlásokat tárgyaló részek, 19–34. A digitális változat újratördelt szöveg. Teljes jelentés · Kiadói adatok

[117] John Hutchins: ALPAC – the (in)famous report. MT News International, 14, 1996. június. Kutatástörténeti elemzés a jelentésről és következményeiről; későbbi értékelés, nem maga az 1966-os dokumentum. Tanulmány

[118] James Lighthill: Artificial Intelligence – A General Survey. In: Artificial Intelligence – A Paper Symposium. Science Research Council, 1973. Az elérhető szöveg 1972. júliusi keltezésű. A kategóriák, a kombinatorikus robbanás és a korabeli teljesítmény megítélése. Jelentésszöveg

[119] John McCarthy: Review of Artificial Intelligence – A General Survey. Lighthill 1973-as közlésének szerzői bírálata. A jelentés értelmezését és az AI-kutatás belső problémáit is tárgyalja. Teljes bírálat

[120] House of Lords, Select Committee on Artificial Intelligence: AI in the UK – Ready, Willing and Able? 2018, 1. fejezet, különösen a 23–24. pont. A brit jelentés finanszírozási hatására vonatkozó történeti állításokat óvatosan kezeli. Parlamenti áttekintés

[121] Edward A. Feigenbaum – Bruce G. Buchanan – Joshua Lederberg: On Generality and Problem Solving – A Case Study Using the DENDRAL Program. Stanford, AIM-131, CS176, 1970. augusztus; NASA-CR-123182. Az általános keresés és a kémiai szaktudás kapcsolata. Archívumi rekord és kézirat

[122] William van Melle: The Structure of the MYCIN System. In: Bruce G. Buchanan – Edward H. Shortliffe szerk.: Rule-Based Expert Systems – The MYCIN Experiments of the Stanford Heuristic Programming Project. Addison-Wesley, 1984, 4. fejezet. Konzultáció, magyarázat, tudásmódosítás. Fejezetszöveg

[123] Edward H. Shortliffe – Bruce G. Buchanan: A Model of Inexact Reasoning in Medicine. Ugyanott, 11. fejezet, 1984. A bizonyossági tényezők módszere és a valószínűségi értelmezéssel kapcsolatos megfontolások. Történeti módszerismertetés, nem kezelési útmutató. Fejezetszöveg

[124] Judith Bachant – John McDermott: R1 Revisited – Four Years in the Trenches. AI Magazine, 5/3, 1984, 21–32. DOI: 10.1609/aimag.v5i3.445. Az 1980-tól használt R1/XCON és tudásbázisának üzemi fejlődése. Kiadói közlés

[125] Stephen Polit: R1 and Beyond – AI Technology Transfer at DEC. AI Magazine, 5/4, 1984, 76–78. DOI: 10.1609/aimag.v5i4.460. A vállalati átvétel, a tudás javítása és a fenntartáshoz szükséges képességek. Kiadói közlés

[126] Knowledge Engineering. In: Buchanan – Shortliffe szerk.: Rule-Based Expert Systems, 1984, 7. fejezet. A tudás megszerzése, átalakítása, fokozatos bővítése és az emberi közreműködés. Fejezetszöveg

[127] Motoi Suwa – A. Carlisle Scott – Edward H. Shortliffe: Completeness and Consistency in a Rule-Based System. Ugyanott, 8. fejezet, 1984. Tudásbázis-ellenőrzés, hiányok és ellentmondások. Fejezetszöveg

[128] A. Carlisle Scott – William J. Clancey – Randall Davis – Edward H. Shortliffe: Methods for Generating Explanations. Ugyanott, 18. fejezet, 1984. A szabályalapú rendszer válaszainak magyarázata. Fejezetszöveg

[129] Human Engineering of Medical Expert Systems. Ugyanott, 32. fejezet, 1984. Különösen 599–600: a MYCIN rutinszerű bevezetésének elmaradása, használati és adatkapcsolati akadályai, valamint az ismeretek elavulása. Fejezetszöveg

[130] Symbolics Inc.: Symbolics Technical Summary. 1985. október. A gyártó műszaki leírása, a Symbolics Lisp Machine Museum engedéllyel közölt változatában. A 3600-as család architektúrája és fejlesztőkörnyezete; a gyártói minősítéseket forráskritikával kezeljük. Korabeli dokumentum

[131] Los Angeles Times: Symbolics Now Computer World’s Fallen Star. 1988. március 22. Korabeli tudósítás

[132] Symbolics Inc.: Form 8-K, 1996. június 27., accession 0000922409-96-000035; különösen Exhibit 99.2, Background. A vállalati beadvány a Chapter 11 eljárás kezdetét 1993. január 27-re teszi, és leírja az eszközértékesítést. Múzeumi másolat. Vállalati irat

[133] Drew McDermott és munkatársai: The Dark Ages of AI – A Panel Discussion at AAAI–84. AI Magazine, 6/3, 1985, 122–134. DOI: 10.1609/aimag.v6i3.494. Az 1984-es vita közölt változata a túlzott várakozásokról és lehetséges visszaesésről. Kiadói rekord

[134] David E. Rumelhart – Geoffrey E. Hinton – Ronald J. Williams: Learning Representations by Back-Propagating Errors. Nature, 323, 1986, 533–536. DOI: 10.1038/323533a0. Eredeti tanulmány

[135] J. Ross Quinlan: Induction of Decision Trees. Machine Learning, 1, 1986, 81–106. DOI: 10.1007/BF00116251. Az ID3 és a példákból felépített döntési szerkezetek. Kiadói rekord · Tanulmány

[136] Christopher J. C. H. Watkins – Peter Dayan: Q-learning. Machine Learning, 8, 1992, 279–292. DOI: 10.1007/BF00992698. A konvergencia eredménye meghatározott feltételekhez kötött. Szerzői oldal és kézirat

[137] Computer History Museum: AI – Expert Systems Pioneer Meeting, Session 2. 2019, X8903.2019, különösen a 16. oldal. Edward Feigenbaum résztvevői értelmezése az AI-telek történetéről; nem semleges, teljes korszakleírás. Beszélgetés átirata

Képforrások és felhasználási adatok

64. kép Saját mondat és kétértelműségi példa. A két helyzet egyaránt lehetséges az adott magyar mondat alapján.

65. kép Saját matematikai grafikon: n különböző elem összes sorrendje n!. A logaritmikus tengely és a szöveg azonos modellt használ; nem empirikus adatsor.

66. kép Saját időrendi összefoglaló a [116–118] alapján, egységes időskála és finanszírozási mennyiségek nélkül.

67. kép Saját ábra a [118] vitatott kategorizálásáról; a kritikához lásd [119].

68. kép United States Air Force: Edward A. Feigenbaum portréja. A forrás dátuma 1994. december 1. A letöltött Commons-változat további módosítás nélkül szerepel; közkincs. Kép és metaadatok

69. kép Saját szerkezeti képletek: etanol $CH_3–CH_2–OH$, dimetil-éter $CH_3–O–CH_3$. Mindkettő összegképlete C_2H_6O.

70. kép Saját egyszerűsített fogalmi ábra. A következtetés és tudásbázis megkülönböztetéséhez lásd [122].

71. kép Saját magyarázó ábra a tudás megszerzésének és ellenőrzésének folyamatáról; alapja [126].

72. kép Saját függőségi példa három szabállyal. Nem konkrét történeti tudásbázis részlete.

73. kép Dave Fischer (Shieldforyoureyes): Symbolics-3600-on, 2008. szeptember 20. RCS/RI. A teljes letöltött kép további módosítás nélkül. CC BY-SA 3.0. Kép és metaadatok · Licenc

74. kép Saját kutatástörténeti felsorolás a [134–136] tanulmányok alapján; nem teljes idővonal.

VI. rész – A gépek tanulni kezdenek

A szakértői rendszerek egyik legdrágább feladata a tudás összegyűjtése és karbantartása volt. A gépi tanulás más lehetőséget kínált: megfelelő példákból maga a program alakíthatta ki a döntéshez használt összefüggések egy részét. Ez a gondolat már az AI korai éveiben megjelent, az 1980-as és 1990-es években azonban egyre több használható módszer és alkalmazás épült rá. A következő fejezetek a tanulás alapfogalmaitól a hétköznapi szolgáltatásokig vezetnek; a játékok történetét pedig külön, a korai automatáktól 2018-ig követjük végig.

27. Szabályok helyett példák

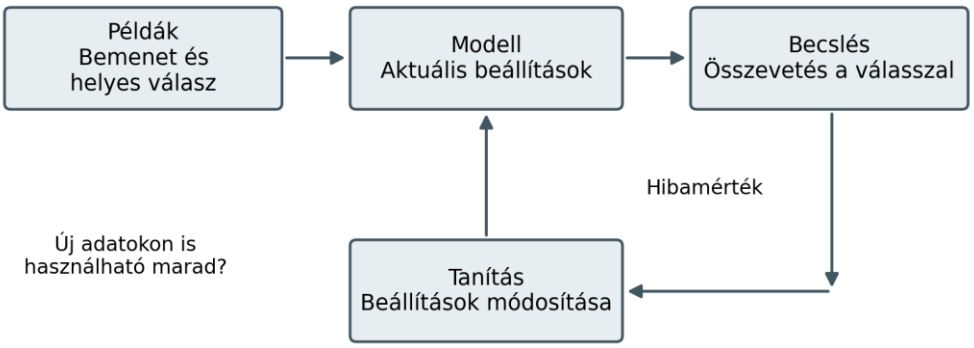
Ha egy programnak meg akarjuk tanítani, hogyan néz ki a kézzel írt ötös, nehéz volna minden változatot külön szabályokkal leírni. Van, aki bezárja a felső hurkot, más ferdén ír, megint más összeköti a számot a következő jellel. A gépi tanulás egyik alapötlete, hogy sok felismert mintát mutatunk a programnak, és olyan beállításokat keresünk, amelyekkel új példákon is jó eredményt ad.

A fejezet címe egy hangsúlyeltolódást jelöl. A tanuló rendszernek is vannak ember által megadott szabályai: valaki kijelöli a feladatot, kiválasztja az adatokat, meghatározza a modell felépítését és a siker mércéjét. Azt viszont nem feltétlenül írja elő részletesen, mely képrészletből milyen döntés következzék. Tom Mitchell 1997-es tankönyve a tapasztalat nyomán javuló, meghatározott feladaton mérhető teljesítmény köré rendezte a tanulás fogalmát. [138]

Mit kap a tanuló program

Felügyelt tanuláskor a példához elvárt válasz is tartozik. A számjegykép mellett ott állhat az „5” címke; egy korábbi ügylet mellett az, hogy később igazolt csalásnak bizonyult-e. A modell a bemenetből becslést készít. A tanítási eljárás az eltérést valamilyen hibamértékkel értékeli, és ennek alapján változtatja a modell beállításait.

Képzeljünk el egy egyszerű, saját példát. Egy rendszer a számjegy néhány képi jellemzőjét súlyozza. Kezdetben túl nagy jelentőséget tulajdonít annak, hogy a jel felső részén vízszintes vonal látszik. Emiatt több hetest ötösnek minősít. A megfelelő tanítópéldák hatására más részletek is nagyobb súlyt kaphatnak. A tanulás itt számszerű változtatások sorozata, amelyet az elvárt válaszok irányítanak.



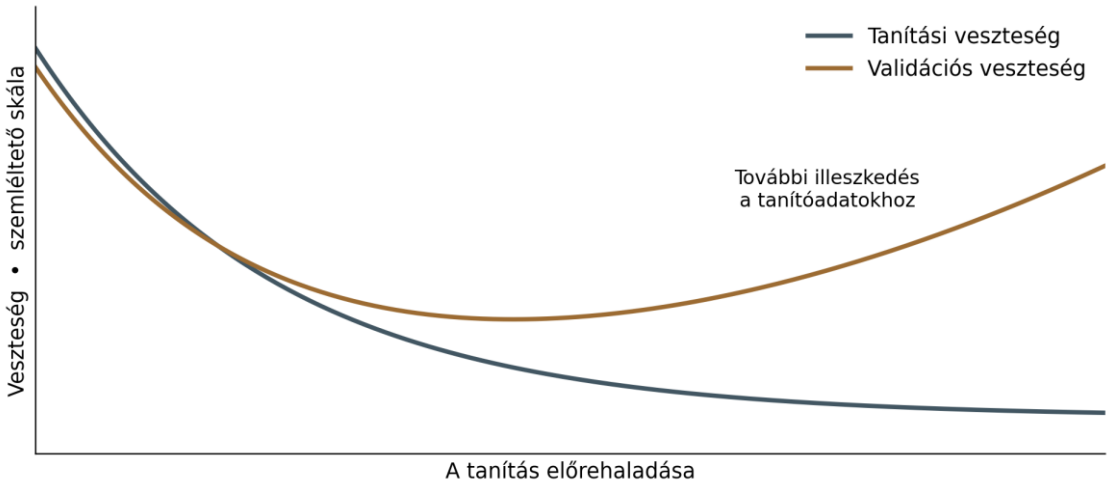
75. kép A felügyelt tanítás egyszerűsített folyamata. A becslés és az elvárt válasz eltérését a tanítóeljárás a modell módosítására használja. Saját magyarázó ábra.

A példák minősége ezért a működés része. Ha az adatbázis címkéi tévesek, a program a tévedésekhez is alkalmazkodhat. Ha csak gondosan írt számokat lát, kevés tapasztalata lesz a kapkodva kitöltött űrlapokról. Az adatgyűjtés nem puszta előkészület: eldönti, milyen különbségeket lehet egyáltalán megtanulni.

Amit még nem mutattunk meg

Az általánosítás azt jelenti, hogy a tanult összefüggés korábban nem látott eseteken is használható. A bemutatott példák megjegyzése ehhez kevés. Egy túl részletes modell a tanítóhalmaz esetleges hibáit és véletlen sajátosságait is követheti. Ezt nevezzük túlilleszkedésnek.

Ezért külön adatokat érdemes fenntartani a beállítások kiválasztására és a végső értékelésre. A tanítóhalmazon módosítjuk a modellt, a validációs halmaz segít választani a változatok között, a teszthalmaz pedig a végül kiválasztott modell ellenőrzésére szolgál. Ha a teszteredmény alapján újra és újra javítgatunk, az ellenőrző adatok is beleszólnak a fejlesztésbe. A független próba szerepe fokozatosan elvész. [141]



76. kép A túlilleszkedés szemléltetése. A tanítási veszteség tovább csökkenhet, miközben az elkülönített validációs adatokon mért veszteség növekedni kezd. Saját, mesterségesen szerkesztett görbék, nem egy történeti kísérlet adatai.

A tanulhatóság elméleti kérdéssé is vált. Leslie Valiant 1984-es tanuláselméleti munkája pontos feltételekhez kapcsolta, hogy példákból milyen számítási erőforrások mellett lehet megbízható közelítést nyerni. [139] A történeti jelentősége az volt, hogy a „megtanulta” kijelentés mellé vizsgálható feltételeket tett. Milyen példákból, milyen feltevések mellett, milyen hibával és mennyi munkával?

Többféle tanulás létezett

A tanuló módszerek nem korlátozódtak neurális hálózatokra. Döntési fák, statisztikai osztályozók és más eljárások is fejlődtek. Corinna Cortes és Vladimir Vapnik 1995-ös tanulmánya a támaszvektor-gépek fontos változatát mutatta be. Az eljárás a különböző osztályok elválasztását olyan matematikai feladatként kezelte, amelyben az elválasztó határ körüli távolságnak is szerepe van. [140]

Felügyelet nélküli tanuláskor nincs minden példához kész megoldás: a rendszer például hasonló eseteket csoportosít. Megerősítéses tanuláskor a cselekvések következményeiből, jutalmakból tanul. Itt egy korai lépés haszna csak sokkal később derülhet ki; a rendszernek próbálkozás és a már bevált viselkedés között is választania kell. [142]

A tanítás és a használat külön időszak lehet. Egy már betanított karakterfelismerő működhet rögzített beállításokkal, és nem szükségszerűen tanul minden új dokumentumból. A folyamatos tanulást külön meg kell szervezni. Az adatokat, a visszajelzéseket és az ellenőrzést akkor sem lehet elhagyni, ha a beállításokat már a gép módosítja.

28. Bizonytalansággal is lehet számolni

A szabályalapú következtetés legegyszerűbb formájában azt mondjuk: ha teljesül egy feltétel, következik egy állítás. A valóságban gyakran csak közvetett jelzésünk van. Egy zajos felvétel több szónak is megfelelhet; egy szokatlan ügylet lehet csalás, de lehet az ügyfél rendkívüli vásárlása is. A valószínűségi modellek lehetővé tették, hogy a rendszer több lehetőséget tartson számon, és az új információ hatására módosítsa a becslését.

A valószínűségszámítás természetesen jóval idősebb az AI-nál. Thomas Bayes halála után, Richard Price közreműködésével 1763-ban jelent meg a tapasztalatból történő következtetés egyik alapmunkája. [143] Az AI számára később az lett a kérdés, hogyan lehet az ilyen matematikai gondolatokat nagyobb, egymással összefüggő változókat kezelő programokban használni.

A jelzés és a bizonyíték ereje

Saját, kitalált számpéldánkban tízezer ügyletből száz csalárd, a többi szabályos. Egy jelzőrendszer a száz csalásból kilencvenet észrevesz, de a 9900 szabályos ügylet öt százalékát is megjelöli. Így 90 valódi találat mellett 495 téves riasztás keletkezik. Összesen 585 jelzésből csupán 90 vonatkozik tényleges csalásra: körülbelül 15,4 százalék.

Kitalált példa	Jelez	Nem jelez	Összesen
Valóban csalás	90	10	100
Szabályos ügylet	495	9 405	9 900
Összesen	585	9 415	10 000

A jelzések között valódi találat: 90 / 585 ≈ 15,4%

77. kép Ritka esemény és téves riasztások saját számpéldája. A csalások 90 százalékának felismerése mellett az összes jelzésnek csak körülbelül 15,4 százaléka valódi találat. A számok kitaláltak.

A számítás nem mond ellent a kilencvenszázalékos találati aránynak. Más kérdésre válaszol. Az egyik azt méri, a valódi csalások mekkora részét jelezte a rendszer. A másik azt, a jelzések mekkora része bizonyult csalásnak. Az utóbbiban az esemény eredeti gyakorisága is döntő. Ritka eseménynél a sok szabályos esetből származó kevés százaléknyi téves jelzés is nagy tömeget alkothat.

A Bayes-szabály az előzetes valószínűséget és a megfigyelt bizonyíték valószínűségét kapcsolja össze a frissített becslésben. Ehhez megfelelő kiinduló adatok kellenek. A képlet önmagában nem teszi hitelessé a rosszul becsült gyakoriságokat. A fenti számok módszertani szemléltetések, nem egy bank vagy kereskedelmi termék teljesítményadatai.

Összefüggések hálózata

Judea Pearl 1988-as könyve a valószínűségi következtetést az intelligens rendszerek egyik szervezőelvévé tette. [144] A Bayes-háló változókat és feltételes valószínűségi kapcsolatokat ír le irányított, körmentes gráffal. A szerkezet segít abban, hogy a teljes közös eloszlást kisebb, helyi részekből adjuk meg. A hiányzó információ nem feltétlenül akasztja meg a következtetést: a lehetséges értékek valószínűségeit összegezve tovább számolhatunk.

Saját egyszerű példánkban egy berendezés hibája és a környezet melegsége egyaránt befolyásolja, világít-e a túlmelegedést jelző lámpa. Ha a lámpa kigyullad, mindkét magyarázatot mérlegelni kell. Ha közben megmérjük a környezeti hőmérsékletet, megváltozhat a hibára vonatkozó becslés is. Az észleletből visszafelé következtetünk egy lehetséges állapotra.



78. kép Egy jelzőlámpa állapotának két lehetséges befolyásoló tényezője. Saját, háromváltozós Bayes-háló-vázlat; a teljes modellhez valószínűségeket is meg kellene adni. Nem egy konkrét berendezés diagnosztikai terve.

Pearl módszerismertetője azt is megmutatta, hogyan terjedhet az információ egy ilyen modellben. [145] A hálózat azonban feltevéseket rögzít, köztük feltételes függetlenségeket. Két hasonló jelzés nem jelent szükségképpen két független bizonyítékot. Ha ugyanannak az érzékelőhibának a következményei, túlzott magabiztossághoz vezethet a külön beszámításuk. Egy pusztán megfigyelési adatokhoz illesztett gráf nyilai önmagukban még nem bizonyítanak oksági kapcsolatot.

Amikor a bizonytalanság időben változik

A beszéd megértéséhez egymás utáni hangrészleteket kell értelmezni. A rejtett Markov-modellekben a rendszer közvetlenül nem megfigyelhető állapotok sorozatát kapcsolja a mérhető jelekhez. Megadja az állapotváltások és a megfigyelések valószínűségeit, majd ezek alapján értékel lehetséges sorozatokat. Lawrence Rabiner 1989-es összefoglaló tanulmánya a beszédfelismerésen keresztül rendszerezte a módszer alapfeladatait. [146]

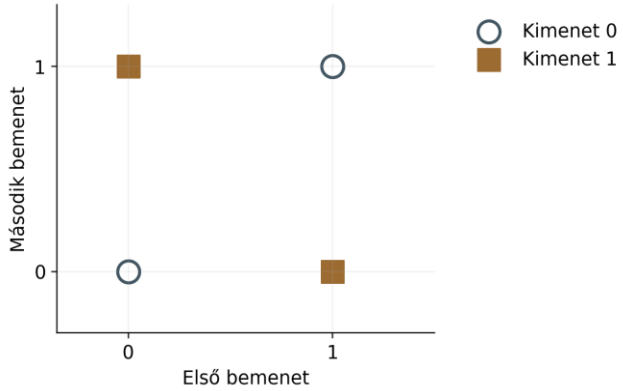
A statisztikai szemlélet a döntés következményeit is előtérbe hozta. Egy téves riasztás és egy elmulasztott csalás eltérő költséget okozhat. Ugyanaz a becslés ezért más intézkedéshez vezethet más feladatban. A valószínűség a bizonytalanság leírása; az, hogy mit tegyünk, a lehetséges következmények mérlegelését is igényli.

29. A neurális hálózatok visszatérése

A 13. és a 20. fejezetben már találkoztunk a mesterséges neuronnal és a perceptronnal. A neurális megközelítés későbbi megerősödését az tette lehetővé, hogy egyre használhatóbban lehetett több egymás utáni réteg kapcsolatait tanítani. A köztes egységek olyan belső leírást alakíthattak ki, amelyet nem kellett minden részletében előre megadni.

Miért számít a köztes réteg

Egy egyszerű, egyetlen lineáris elválasztásra képes modellnek elvi korlátai vannak. A kizáró vagy, az XOR feladatban két bemenet akkor ad egyes eredményt, ha különböznek. A 0 és 0, valamint az 1 és 1 kimenete nulla; a másik két páré egy. Ha a négy esetet pontként ábrázoljuk, az azonos osztályú pontok átlósan helyezkednek el. Egyetlen egyenes nem választja szét a két osztályt.

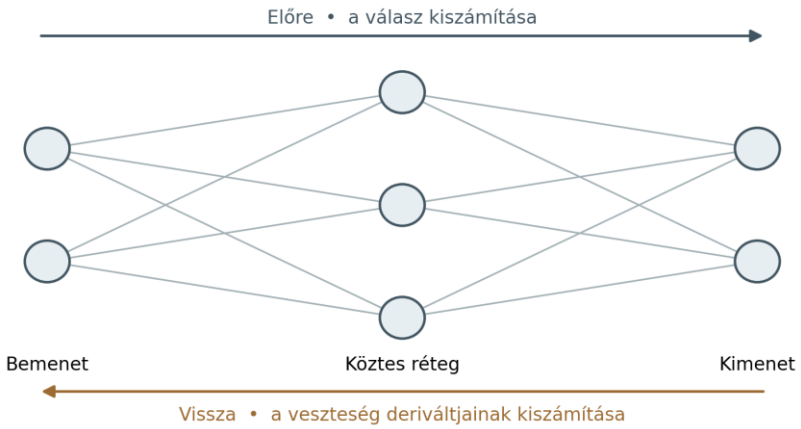


79. kép A kizáró vagy négy esete. Az átlósan elhelyezkedő azonos osztályú pontokat egyetlen egyenes nem választja szét. Saját matematikai ábra.

Több megfelelő köztes egység már más leírást hozhat létre a bemenetből, amelyből a kimenet előállítható. Ehhez nemlineáris működés is kell. Pusztán lineáris átalakítások egymásra rakása végül ismét egyetlen lineáris átalakítással helyettesíthető. A rétegek száma tehát önmagában nem oldja meg a problémát.

Hogyan jut vissza a hiba

A tanításkor a hálózat először kiszámítja a válaszát. Ezt összevetjük az elvárt eredménnyel, és veszteséget számítunk. A visszaterjesztés a matematikai láncszabály segítségével meghatározza, hogyan változna ez a veszteség az egyes súlyok kicsiny módosításakor. Az így kapott deriváltak alapján egy optimalizáló eljárás változtatja a súlyokat, rendszerint a veszteség csökkentésére törekedve.



80. kép Előre a hálózat válaszát, visszafelé a veszteség súlyok szerinti deriváltjait számítjuk. A rajz egy köztes réteget mutat, a magyarázat több rétegre is kiterjeszthető. Saját szemléltetés; az egységek nem biológiai idegsejtek rajzai.

A visszaterjesztés tehát a szükséges változtatási irány kiszámításának hatékony eszköze; az alkalmazott lépésnagyságot és a tényleges frissítést a tanítás további beállításai is alakítják. Nem egy helytelen válasz szöveges megrovása jut hátrafelé a hálózatban. A súlyokhoz tartozó matematikai érzékenységeket számítjuk ki.

David Rumelhart, Geoffrey Hinton és Ronald Williams 1986-os Nature-tanulmánya nagy figyelmet irányított erre a tanításra és a megtanulható belső reprezentációkra. [147] Az eljárásnak voltak korábbi előzményei: Paul Werbos 1974-es doktori munkája is a történetéhez tartozik. Werbos később az időben kiterített visszaterjesztés gyakorlati számítását is részletesen tárgyalta. [148] Az 1986-os közlés ezért fontos fordulópont, de nem helyettesíti az egész fejlődéstörténetet.

A használható hálózat több feltételt kívánt

A működő tanítás mellé adatok, számítási idő és megfelelő felépítés is kellett. A modell nagysága, a kezdő súlyok, a lépésnagyság és a tanítópéldák eloszlása befolyásolta az eredményt. Több réteg nagyobb lehetőséget adott összetett összefüggések leírására, de nehezebb optimalizálási feladatot is teremthetett. Egy kis tanítóhibájú hálózat továbbra is túlilleszkedhetett.

A képfelismerésben hasznosnak bizonyult, ha a hálózat felépítése figyelembe vette a kép térbeli szerkezetét. Egy helyi mintázat felismerésére szolgáló kapcsolategyüttest több képrészleten is alkalmazhatunk. A konvolúciós hálózatok súlymegosztása így csökkentheti a külön megtanulandó paraméterek számát. LeCun és munkatársai 1998-as dokumentumfelismerési tanulmánya ilyen rendszereket és a teljes feldolgozási lánc tanítását is bemutatta. [150]

A nehézségek ettől nem tűntek el. Yoshua Bengio, Patrice Simard és Paolo Frasconi 1994-es munkája elemezte, miért nehéz gradiensalapú tanítással hosszú időbeli függéseket megtanulni visszacsatolt hálózatokban. [149] A korai információhoz vezető sok számítási lépésen át a tanításhoz szükséges jel túlságosan gyengülhet vagy felerősödhet. Ez különösen a hosszabb sorozatok kezelésénél okozott gondot.

A neurális hálózatok visszatérése ezért fokozatos műszaki és módszertani eredményekből állt. A kutatók olyan feladatokat találtak, ahol a megtanult belső leírás előnyt adott, és javították a tanítás

feltételeit. Ez készítette elő a későbbi mélytanulási hullámot; az 1980-as és 1990-es évek hálózatait azonban nem szabad a jóval későbbi rendszerek képességeivel felruházni.

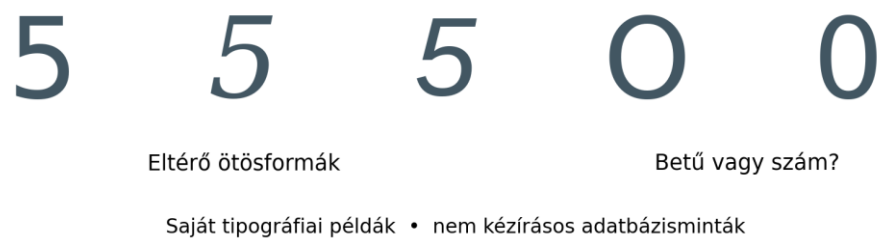
30. A mesterséges intelligencia észrevétlenül beköltözik a mindennapokba

Az AI történetének egy része látványos bemutatókon zajlott. Egy másik része olyan programokba épült, amelyeknek a nevét a felhasználó nem ismerte. Egy beolvasott nyomtatványból kereshető szöveg lett, egy kereső sorba rendezte a találatokat, egy szolgáltatás további könyvet ajánlott. Ezeket a feladatokat különböző módszerek oldották meg; nem kellett mögöttük egyetlen, mindenhez értő intelligenciának állnia.

A gép elolvassa a dokumentumot

Az optikai karakterfelismerés, az OCR feladata, hogy egy kép képpontjaiból betűket és számokat azonosítson. A használható szöveghez azonban az oldal szerkezetét is kezelni kell: hol kezdődik egy hasáb, mi tartozik a táblázathoz, mi a képaláírás? Egy felismerési hiba továbbterjedhet. Az „O” betű és a nulla összetévesztése más következménnyel jár egy regényben és egy számlaösszegben.

A kézírás változatossága különösen jó feladatot adott a tanuló módszereknek. LeCun, Bottou, Bengio és Haffner 1998-as tanulmánya olyan dokumentumfelismerő rendszereket ismertetett, amelyekben a neurális tanítás és a feldolgozás más lépései együtt dolgoztak. Banki csekkek olvasását is tárgyalta. [150] A kép felismerése itt egy hosszabb üzemi folyamat része volt.



81. kép Eltérő ötösformák, valamint az O betű és a nulla. Saját, betűtípusokkal készített példák, nem kézírásos adatbázisból vett minták.

Az OCR megmutatja, miért nem elég egy szép bemutatókép. A programnak sokféle papírral, nyomtatóval, kézírással és képminőséggel kell találkoznia. A bizonytalan eseteket ellenőrzésre lehet továbbítani. A gazdasági értéket az is alakítja, mennyi emberi javítás marad a felismerés után, és mely hibák kerülnek át észrevétlenül a következő rendszerbe.

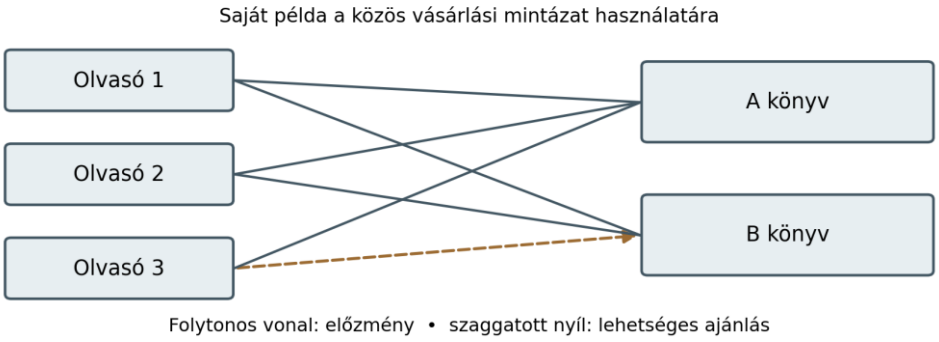
A kereső kiválasztja a sorrendet

A világháló növekedésével a dokumentum megtalálása mellett a találatok rangsorolása is alapfeladat lett. Sergey Brin és Lawrence Page 1998-as tanulmánya a Google korai keresőjét írta le. A szöveges egyezések mellett a hivatkozási hálózatból számolt PageRank is segített a sorrend kialakításában. [151] Egy oldalra mutató hivatkozás jelentőségét az is befolyásolta, milyen oldalak hivatkoznak rá.

Ez a példa óvatosságra int a fogalmakkal. A PageRank matematikai rangsorolási eljárás; önmagában nem példák címkéiből tanított neurális hálózat. A keresés fejlődése információ-visszakeresést, gráfalgoritmusokat, statisztikát és később tanuló módszereket kapcsolt össze. Egy hasznos szolgáltatás történetét félrevezető volna egyetlen technikára szűkíteni.

A rendszer ajánl valamit

Greg Linden, Brent Smith és Jeremy York 2003-ban ismertette az Amazon termékek közötti hasonlóságokra építő ajánlási módszerét. A vásárlói előzményekből olyan termékpárokat keresett, amelyek között használható kapcsolat mutatkozott, majd ezek alapján készített személyes ajánlásokat. [152] A közös viselkedési mintázat akkor is segíthetett, ha a program nem értette úgy a könyvek tartalmát, ahogyan az olvasó.



82. kép Közös vásárlási előzményből származó ajánlás saját példája. Az első két olvasó mindkét könyvet választotta; a harmadiknak a rendszer felajánlhatja a B könyvet. A rajz nem az Amazon teljes eljárásának másolata.

Saját ábránkban több vásárló együtt választ két könyvet. Ez arra adhat alapot, hogy az egyik kötet olvasójának a másikat is megmutassuk. A közös vásárlásból azonban nem következik, hogy minden olvasónak mindkettő tetszeni fog. Az ajánlás becslés, amelyet az is alakít, milyen termékeket láttak egyáltalán a korábbi vásárlók.

Szokatlan ügyletek és beszélt szavak

A csalásfelderítésben a tanuló program ismert esetek és szokásos viselkedési minták alapján kereshet gyanús eltéréseket. Tom Fawcett és Foster Provost 1997-es tanulmánya mobiltelefon-klónozással kapcsolatos csalások felismerésére mutatott be alkalmazkodó módszert. [153] A történeti példa telefonszolgáltatói hívásadatokról szólt; a banki ügyletellenőrzés rokon feladat, de nem ugyanaz az adat és környezet.

A beszédfelismerésben közben a hangjel és a lehetséges hangsorok valószínűségi leírása segített. A Rabiner által összefoglalt rejtett Markov-modellek a változó időtartamú és bizonytalan jelek kezelésére adtak módszert. [146] A beszélő, a háttérzaj és a szóhasználat eltérései itt is megmutatták az alkalmazás határait.

Egy rendszer akkor is válhatott mindennapi eszközzé, ha az általános nyelvértés vagy a józan ész nagy kérdései megoldatlanok maradtak. A felhasználó egy részfeladat eredményét kapta meg. Az AI

hasznosságát ezekben az esetekben a felismerés pontosságán, a gyorsabban elvégzett munkán és a javítási igényen lehetett megítélni.

31. Amikor a gépek játszani kezdtek

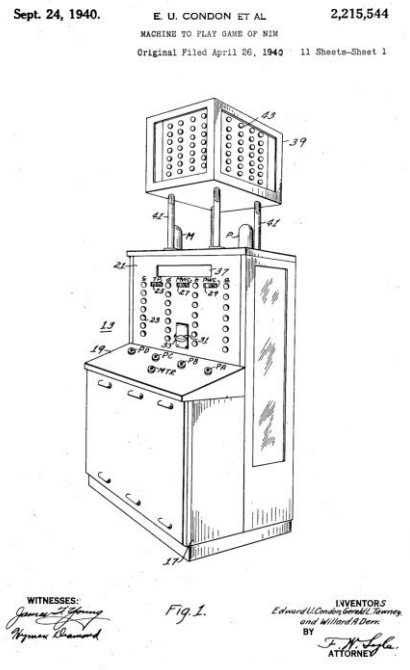
A játékautomatáktól az emberfeletti stratégiákig

A játék vonzó kísérleti környezet: szabályai többnyire pontosak, az eredmény ellenőrizhető, és ugyanazt a programot újra próbára tehetjük. A győzelemhez vezető út mégis nagyon különböző lehet. Egy gép követhet kész nyerőeljárást, kereshet a lehetséges lépések között, vagy játék közben alakíthatja ki az értékelését. A történet elején pedig olyan szerkezet is állt, amelynek tudását egy elrejtett ember adta.

Kempelen Törökje és a valódi játékautomaták

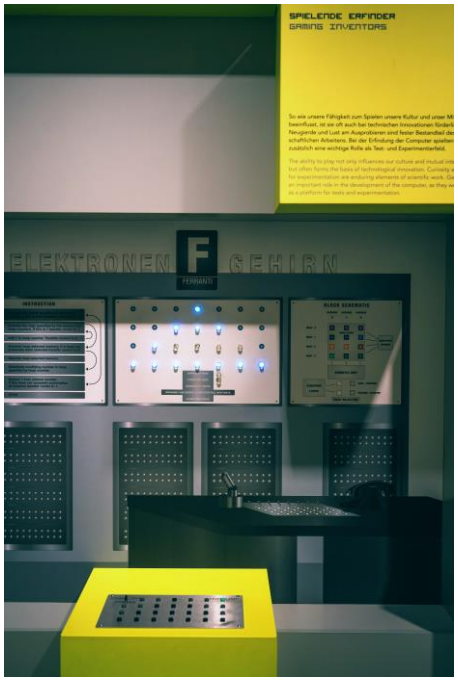
Kempelen Farkas 1770-ben bemutatott Törökjéről a 4. fejezetben már szó volt. A bábu mögötti gépezet látványa önálló sakkjátékost sejtetett, valójában ember irányította a lépéseket. [154] A játék itt a gépi gondolkodás látszatát teremtette meg. A későbbi programok történetében éppen az lett ellenőrizhető kérdés, hogy a döntést maga a berendezés számítja-e ki.

A Westinghouse Nimatronja 1940-ben a New York-i világkiállításon már valódi gépi ellenfél volt. Edward Condon és munkatársai relékkel valósítottak meg egy nimjátékost. [155–156] A nimben halmokból veszünk el elemeket, egy lépésben egyetlen halomból; a szokásos változatban az nyer, aki az utolsót elveszi. A játékhoz matematikai nyerőeljárás adható. A gépnek ezért nem kellett tapasztalatból feltalálnia a stratégiát.



83. kép A Nimatronhoz kapcsolódó US 2,215,544 szabadalom első rajzlapjának képkivágata. Edward U. Condon, Gereld L. Tawney és Willard A. Derr; benyújtás 1940. április 26., megadás 1940. szeptember 24. Korabeli amerikai szabadalmi rajz, közkincs.

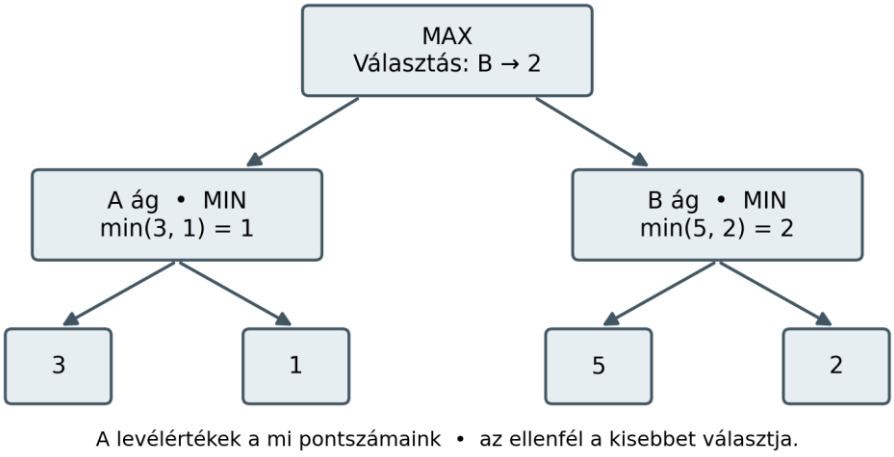
A Ferranti 1951-es Nimrodja elektronikus áramkörökkel és lámpás kijelzéssel mutatta be a digitális számítás elveit a Festival of Britain közönségének. A hozzá kiadott ismertető a számítógépek működésének megértését is szolgálta. [157] A Nimatron és a Nimrod nem ugyanaz a gép volt: különböző korszak és műszaki megoldás kapcsolódott bennük ugyanahhoz az alkalmasan kiválasztott játékhoz.



84. kép A Nimrod kisebb méretű múzeumi másolata a berlini Computerspielemuseumban, 2018. november 9-i felvételen. Nem az 1951-es eredeti gép. Fotó: Marcin Wichary; a Commons-változat fény- és színkorrekciója Karl432 munkája. CC BY 2.0.

A sakklépés számítási feladattá válik

Claude Shannon 1950-es tanulmánya a sakkprogramozás alapfeladatait rendszerezte. Ha nem kereshetünk a játszma végéig, korlátozott mélységben vizsgáljuk a folytatásokat, és az ott elért állásokat értékeljük. Az ellenfél várható legjobb válaszát is figyelembe kell venni. [159] A számítás így felváltva nézi, mit tehetünk mi, és mit válaszolhat rá a másik fél.



85. kép A minimax saját, kétlépéses számpéldája. A gép nagyobb, az ellenfél kisebb pontszámot választ a mi szempontunkból. A B ág legrosszabb kimenete is jobb, mint az A ágé. Az értékek nem egy sakkállásból származnak.

Saját minimaxpéldánkban az egyik lépés után az ellenfél három vagy egy pontot engedhet nekünk, ezért egy ponttal számolunk. A másik ágban öt vagy két pontot érhetünk el, így ott kettővel. Ha az ellenfél a számára kedvezőbb választ adja, a két pontot biztosító második ágat választjuk. A valódi sakkban a levágott keresési fa szélén többnyire becslés áll, nem bizonyított végeredmény.

Alan Turing is kidolgozott sakkjátszó eljárást; ennek gondolatmenetét az 1953-as Faster than Thought sakkal foglalkozó részében ismertette. A korai eljárást kézzel is végig lehetett számolni, működő gépi futtatás nélkül. [158] A program megtervezése és a számítógépen futó megvalósítás tehát külön történeti teljesítmény.

Dietrich Prinz 1951-es manchesteri programja már számítógépen oldott meg kétlépéses mattfeladványokat; teljes sakkjátszmára még nem volt alkalmas. [160] Christopher Strachey az 1950-es évek elején dámázó programot írt, amely a korai működő játékprogramok közé tartozott. [161] A sakk, a dáma és a nim fejlesztései párhuzamosan járultak hozzá a gépi játékhoz; nincs szükség egyetlen, minden értelemben „első” játékos kinevezésére.

Samuel programja a tapasztalatot is használta

Arthur Samuel IBM-nél fejlesztett dámaprogramja a keresést tanulással kapcsolta össze. 1959-es tanulmánya két tanulási eljárást ismertetett: korábban értékelt helyzetek megőrzését, illetve az állások jellemzőit összegző értékelés módosítását. [162] A program így a tapasztalatok hatására változtathatott azon, mit tekint előnyös állásnak.

Ez fontos különbség a nim kész stratégiájához képest. A játékszabályok továbbra is meg voltak adva, de a jó játékhoz használt értékelés nem minden részletét a fejlesztő rögzítette. Egy később kedvezőnek bizonyuló állás visszajelzést adhatott egy korábbi döntés értékeléséhez. A tanulás egy pontosan kijelölt feladaton javította a teljesítményt; általános emberi tudás ebből még nem következett.

BKG 9.8 és a dobókocka szerepe

Hans Berliner BKG 9.8 programja 1979. július 15-én Monte-Carlóban 7–1-re legyőzte Luigi Villa backgammon-világbajnokot egy bemutató mérkőzésen. Berliner 1980-as beszámolója a helyzetfüggő értékelő függvény kialakítását is ismertette. [163] A backgammonban a dobókocka véletlent visz a játékba, így a jó döntéshez a lehetséges későbbi események mérlegelése is hozzátartozik.

A győzelem értékét maga Berliner is körültekintően kezelte. A találkozó nem a világbajnoki címért zajlott, és a rövid mérkőzés, valamint a szerencse nem adott alapot a tartós fölény bizonyítására. [163] A siker azt mutatta, hogy a program erős ellenfél lehetett. Nem jelentette a backgammon matematikai megoldását, és nem azonosítható a későbbi, önjátékkal tanuló neurális programok módszerével.

Chinook és a bizonyított döntetlen

Jonathan Schaeffer és munkatársai 1989-től dolgoztak az angol, 8×8-as táblán játszott dáma gépi vizsgálatán. A Chinook történetéhez erős versenyprogram és később egy bizonyítási munka is tartozott. A 2007-ben közölt eredmény szerint az alapállásból mindkét fél hibátlan játéka döntetlenre vezet. [164]

Ez a játék úgynevezett gyenge megoldása: az induló helyzet értéke és az azt biztosító stratégia ismert. Nem azt jelenti, hogy minden elképzelhető állás teljes értékét külön-külön meghatározták, vagy hogy a dáma minden más szabályváltozatát is megoldották. A kutatás végjáték-adatbázisokat és keresési eljárásokat kapcsolt össze. A „gyenge” itt szakkifejezés; a bizonyítás erejét nem kisebbíti.

A különbség gyakorlati jelentőségű. Egy bajnokot legyőző program még hibázhat más helyzetben. Az alapállásból döntetlent biztosító stratégia viszont bármely ellenféllel szemben tartja ezt az eredményt, ha helyesen követjük. A versenyeredmény és a matematikai bizonyítás eltérő kérdésre ad választ.

Deep Blue és a sakkvilágbajnok

Az IBM Deep Blue 1996-ban már nyert játszmát Garri Kaszparov ellen, de az egész mérkőzést elveszítette. A továbbfejlesztett gép 1997-ben 3½–2½ arányban megnyerte a hatjátszmás visszavágót. A rendszer külön sakkhardvert, nagymértékben párhuzamos keresést és emberi sakktudással alakított értékelést használt. [165]



86. kép Deep Blue múzeumi bemutatása a Computer History Museumban, 2011. május 7-i fényképen. A felvétel a gép egyik szekrényét mutatja, nem az 1997-es teljes versenykörnyezetet. Fotó: Anton Chiang, CC BY 2.0.

A Deep Blue nem a teljes sakkot számította végig. A keresést és az értékelést műszakilag olyan szintre emelte, amellyel a világbajnokkal is felvehette a versenyt. Nem az AlphaZero későbbi, önjátékból tanuló rendszerének korai változata volt. A két gép összevetése éppen azért érdekes, mert ugyanahhoz a jól mérhető célhoz eltérő módszerekkel jutottak.

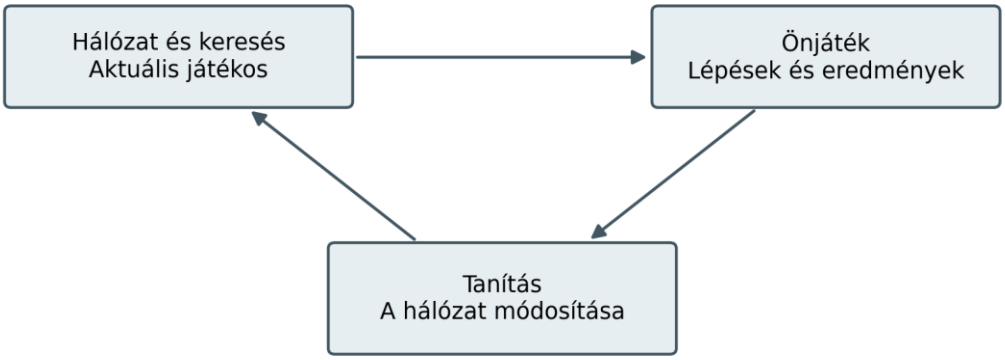
AlphaGo és a megtanult állásértékelés

A góban a sok lehetséges lépés és az állások értékelése különösen nehézzé tette a keresést. A DeepMind 2016-ban ismertetett AlphaGo rendszere neurális hálózatokat kapcsolt fatereséssel. Egyik hálózata ígéretes lépéseket jelzett, másik az állás várható eredményét becsülte. A tanítás emberi játszmákból és önjátékból is merített. [166]

AlphaGo 2016 márciusában Szöulban 4–1-re legyőzte a kiemelkedő góversenyző Li Szedolt, nemzetközi nevén Lee Sedolt. [167] A győzelem mögött a tanult becslés és az előretekintő számítás együtt dolgozott. A hálózat segített oda irányítani a keresést, ahol érdemesnek látszott; a keresés pedig pontosíthatta a puszta becslés alapján hozott döntést.

AlphaGo Zero és AlphaZero

A 2017-ben közölt AlphaGo Zero már emberi játszmák tanítóanyagának használata nélkül, önjátékból tanult gózni. A szabályok és a tanítási rendszer emberi tervezésűek maradtak. [168] A „nulláról” tanulás ebben a történetben a korábbi emberi lépéspéldák elhagyását jelenti, nem a számítógép, a modellfelépítés vagy a jutalmazási cél hiányát.



87. kép Önjáték, keresés és hálózattanítás kapcsolata az AlphaGo Zero és az AlphaZero módszerének egyszerűsített magyarázatában. Saját ábra a [168–169] alapján; a játékszabályokat és a tanítási keretet emberek adják meg.

Az AlphaZero 2017 végi előzetes közlése után 2018-ban jelent meg a részletes Science-tanulmány. Ugyanazt az általános tanulási és keresési eljárást sakkban, sógiban és góban is alkalmazták, játékonként külön tanítással. A rendszer nem kapott emberi játszmagyűjteményt és kézzel kidolgozott állásértékelést. [169] Ettől még ismernie kellett az adott játék szabályait, a szabályos lépéseket és az eredmény meghatározását.

Az önjáték során a pillanatnyi hálózat és a keresés új játszmákat hoz létre; ezekből a hálózat újabb tanítóadatot kap. A keresés által javasolt lépések és a végső eredmények segítik a következő változatot. Így a rendszer maga állítja elő tapasztalatainak jelentős részét. Ez jól használható ott, ahol a környezet pontosan szimulálható, és a próbálkozás eredménye egyértelműen mérhető.

A történeti ív ezért többféle teljesítményt fog össze. A Török az emberi tudást rejtette el. A nimjátékosok kész matematikai eljárást valósítottak meg. A sakkprogramok keresést és értékelést fejlesztettek, Samuel a tapasztalatot is bevonta, a Chinookhoz bizonyítás társult, az Alpha-rendszerek pedig a tanulást és a keresést szorosan összekapcsolták. Ezekből sokat megtudtunk a gépi problémamegoldásról. A játék szabályos világán kívüli alkalmazáshoz azonban külön meg kell vizsgálni, mi felel meg az állásnak, a szabályoknak és a győzelemnek, és milyen következménye lehet egy hibás próbálkozásnak.

Források és további olvasmányok

A hivatkozások és a képek számozása az előző részeket folytatja. A számpéldák és a magyarázó ábrák saját szerkesztések. A játékprogramokról szóló műszaki tanulmányok mellett a korai szerkezetek azonosítását múzeumi és intézményi történeti források segítik.

[138] Tom M. Mitchell: Machine Learning. McGraw-Hill, 1997. Különösen az 1., 5. és 7. fejezet. A tanulási feladat, a tapasztalat és a teljesítmény fogalma. Szerzői oldal és könyv

[139] Leslie G. Valiant: A Theory of the Learnable. Communications of the ACM, 27/11, 1984, 1134–1142. DOI: 10.1145/1968.1972. Tanulmány

[140] Corinna Cortes – Vladimir Vapnik: Support-Vector Networks. Machine Learning, 20, 1995, 273–297. DOI: 10.1007/BF00994018. Kiadói közlés

[141] Trevor Hastie – Robert Tibshirani – Jerome Friedman: The Elements of Statistical Learning. 2. kiadás, Springer, 2009. Különösen a 7. fejezet a modellek értékeléséről és kiválasztásáról. Kiadói adatok · Szerzői könyvoldal

[142] Richard S. Sutton – Andrew G. Barto: Reinforcement Learning An Introduction. 2. kiadás, MIT Press, 2018. A megerősítéses tanulás fogalmai és a feltárás szerepe. Kiadói oldal · Szerzői kézirat

[143] Thomas Bayes: An Essay towards Solving a Problem in the Doctrine of Chances. Richard Price közreműködésével. Philosophical Transactions, 53, 1763, 370–418. DOI: 10.1098/rstl.1763.0053. Történeti szöveg átírása

[144] Judea Pearl: Probabilistic Reasoning in Intelligent Systems Networks of Plausible Inference. Morgan Kaufmann, 1988. Kiadói könyvoldal

[145] Judea Pearl: Bayesian Networks. UCLA Cognitive Systems Laboratory, R-216. A hálózatok felépítésének és a valószínűségi következtetésnek a szerzői módszerismertetője. Tanulmány

[146] Lawrence R. Rabiner: A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. Proceedings of the IEEE, 77/2, 1989, 257–286. DOI: 10.1109/5.18626. Kiadói közlés

[147] David E. Rumelhart – Geoffrey E. Hinton – Ronald J. Williams: Learning Representations by Back-Propagating Errors. Nature, 323, 1986, 533–536. DOI: 10.1038/323533a0. Az előző részben már idézett munka, itt a módszer részletesebb magyarázatához. Szerzői példány

[148] Paul J. Werbos: Beyond Regression New Tools for Prediction and Analysis in the Behavioral Sciences. Doktori értekezés, Harvard, 1974. Uő: Backpropagation through Time What It Does and How to Do It. Proceedings of the IEEE, 78/10, 1990, 1550–1560. DOI: 10.1109/5.58337. 1974-es értekezés bibliográfiai rekordja · 1990-es tanulmány

[149] Yoshua Bengio – Patrice Simard – Paolo Frasconi: Learning Long-Term Dependencies with Gradient Descent Is Difficult. IEEE Transactions on Neural Networks, 5/2, 1994, 157–166. DOI: 10.1109/72.279181. Tanulmány és kivonat

[150] Yann LeCun – Léon Bottou – Yoshua Bengio – Patrick Haffner: Gradient-Based Learning Applied to Document Recognition. Proceedings of the IEEE, 86/11, 1998, 2278–2324. DOI: 10.1109/5.726791. Szerzői oldal · Teljes tanulmány

[151] Sergey Brin – Lawrence Page: The Anatomy of a Large-Scale Hypertextual Web Search Engine. Computer Networks and ISDN Systems, 30, 1998, 107–117. DOI: 10.1016/S0169-7552(98)00110-X. Eredeti konferenciaszöveg

[152] Greg Linden – Brent Smith – Jeremy York: Amazon.com Recommendations Item-to-Item Collaborative Filtering. IEEE Internet Computing, 7/1, 2003, 76–80. DOI: 10.1109/MIC.2003.1167344. Tanulmány

[153] Tom Fawcett – Foster Provost: Adaptive Fraud Detection. Data Mining and Knowledge Discovery, 1, 1997, 291–316. DOI: 10.1023/A:1009700419189. A vizsgált alkalmazás mobiltelefon-klónozás felismerése volt. Szerzői oldal és tanulmány

[154] Computer History Museum: Mastering the Game A History of Computer Chess. Bevezetés és a Török múzeumi ismertetése. A bemutatás éve 1770. Múzeumi történet

[155] Edward U. Condon – Gereld L. Tawney – Willard A. Derr: Machine to Play Game of Nim. US 2,215,544, benyújtás 1940. április 26., megadás 1940. szeptember 24. Elsődleges műszaki forrás. Szabadalom és rajzok

[156] Optica: Edward U. Condon. Intézményi történeti életrajz, benne a Nimatron 1940-es New York-i bemutatásával. Intézményi életrajz

[157] Ferranti: Faster than Thought The Ferranti Nimrod Digital Computer. Korabeli ismertető, 1951. A számítógépek és a nimjátékos bemutatása. Az ismertető szövege archív átiratban · A példány múzeumi katalógusa

[158] Alan M. Turing: Chess, a Digital Computers Applied to Games című fejezet része. In: B. V. Bowden szerk.: Faster than Thought. Pitman, 1953. A teljes fejezet több szerző munkája. Turing kéziratának archív rekordja · Szerzőségi bibliográfia

[159] Claude E. Shannon: Programming a Computer for Playing Chess. Philosophical Magazine, 41/314, 1950, 256–275. A korabeli tanulmány múzeumi másolata

[160] Computer History Museum: First Tests. Dietrich Prinz 1951-es mattfeladvány-megoldó programjának múzeumi ismertetése. Múzeumi összefoglaló

[161] IEEE Computer Society: Computer Pioneers Christopher Strachey. A korai dámázó program és a manchesteri munka története. Intézményi történeti leírás

[162] Arthur L. Samuel: Some Studies in Machine Learning Using the Game of Checkers. IBM Journal of Research and Development, 3/3, 1959, 210–229. DOI: 10.1147/rd.33.0210. Eredeti tanulmány

[163] Hans J. Berliner: Backgammon Computer Program Beats World Champion. Artificial Intelligence, 14/2, 1980, 205–220. DOI: 10.1016/0004-3702(80)90041-7. Különösen a 6–7. szakasz a mérkőzésről és az értelmezés korlátairól. A szerző tanulmányának újraközlése

[164] Jonathan Schaeffer és munkatársai: Checkers Is Solved. Science, 317, 2007, 1518–1522. DOI: 10.1126/science.1144079. Elektronikus közlés 2007. július 19. Az angol dáma alapállásának gyenge megoldása. Tanulmány · Kiadási adatok

[165] Murray Campbell – A. Joseph Hoane Jr. – Feng-hsiung Hsu: Deep Blue. Artificial Intelligence, 134, 2002, 57–83. DOI: 10.1016/S0004-3702(01)00129-1. A fejlesztők műszaki beszámolója. IBM kutatási rekord · Tanulmány

[166] David Silver és munkatársai: Mastering the Game of Go with Deep Neural Networks and Tree Search. Nature, 529, 2016, 484–489. DOI: 10.1038/nature16961. Eredeti tanulmány

[167] Google DeepMind: AlphaGo. A fejlesztő történeti összefoglalója, benne a Lee Sedol elleni 2016. márciusi, 4–1-es mérkőzéssel. Fejlesztői összefoglaló

[168] David Silver és munkatársai: Mastering the Game of Go without Human Knowledge. Nature, 550, 2017, 354–359. DOI: 10.1038/nature24270. AlphaGo Zero. Tanulmány · Fejlesztői magyarázat

[169] David Silver és munkatársai: A General Reinforcement Learning Algorithm That Masters Chess Shogi and Go through Self-Play. Science, 362, 2018, 1140–1144. DOI: 10.1126/science.aar6404. AlphaZero; a 2017-es előzetes eredmények kibővített értékelése. Eredeti tanulmány · Fejlesztői magyarázat

Képforrások és felhasználási adatok

75. kép Saját szerkesztés. A felügyelt tanítás egyszerűsített folyamata. A becslés és az elvárt válasz eltérését a tanítóeljárás a modell módosítására használja. Saját magyarázó ábra.

76. kép Saját szerkesztés. A túlilleszkedés szemléltetése. A tanítási veszteség tovább csökkenhet, miközben az elkülönített validációs adatokon mért veszteség növekedni kezd. Saját, mesterségesen szerkesztett görbék, nem egy történeti kísérlet adatai.

77. kép Saját szerkesztés. Ritka esemény és téves riasztások saját számpéldája. A csalások 90 százalékának felismerése mellett az összes jelzésnek csak körülbelül 15,4 százaléka valódi találat. A számok kitaláltak.

78. kép Saját szerkesztés. Egy jelzőlámpa állapotának két lehetséges befolyásoló tényezője. Saját, háromváltozós Bayes-háló-vázlat; a teljes modellhez valószínűségeket is meg kellene adni. Nem egy konkrét berendezés diagnosztikai terve.

79. kép Saját szerkesztés. A kizáró vagy négy esete. Az átlósan elhelyezkedő azonos osztályú pontokat egyetlen egyenes nem választja szét. Saját matematikai ábra.

80. kép Saját szerkesztés. Előre a hálózat válaszát, visszafelé a veszteség súlyok szerinti deriváltjait számítjuk. A rajz egy köztes réteget mutat, a magyarázat több rétegre is kiterjeszthető. Saját szemléltetés; az egységek nem biológiai idegsejtek rajzai.

81. kép Saját szerkesztés. Eltérő ötösformák, valamint az O betű és a nulla. Saját, betűtípusokkal készített példák, nem kézírásos adatbázisból vett minták.

82. kép Saját szerkesztés. Közös vásárlási előzményből származó ajánlás saját példája. Az első két olvasó mindkét könyvet választotta; a harmadiknak a rendszer felajánlhatja a B könyvet. A rajz nem az Amazon teljes eljárásának másolata.

83. kép Condon, Tawney és Derr: US 2,215,544, első rajzlap, 1940. Az amerikai szabadalmi rajz közkincs. A rajzlap széleit levágtuk; a rajz tartalmát nem módosítottuk. Eredeti szabadalmi dokumentum

84. kép Marcin Wichary: Video Game Museum in Berlin (45896661952), 2018. november 9. A Commons-változat fény- és színkorrekciója: Karl432, 2022. A letöltött változatot tovább nem módosítottuk. CC BY 2.0. Kép és metaadatok · Licenc

85. kép Saját szerkesztés. A minimax saját, kétlépéses számpéldája. A gép nagyobb, az ellenfél kisebb pontszámot választ a mi szempontunkból. A B ág legrosszabb kimenete is jobb, mint az A ágé. Az értékek nem egy sakkállásból származnak.

86. kép Anton Chiang: IBM Deep Blue at Computer History Museum (9361685537), 2011. május 7. További módosítás nélkül. CC BY 2.0. Kép és metaadatok · Licenc

87. kép Saját szerkesztés. Önjáték, keresés és hálózattanítás kapcsolata az AlphaGo Zero és az AlphaZero módszerének egyszerűsített magyarázatában. Saját ábra a [168–169] alapján; a játékszabályokat és a tanítási keretet emberek adják meg.

VII. rész – A mélytanulástól a beszélgető gépekig

A tanuló gépek korábbi sikerei után új feltételek találkoztak: nagyobb adathalmazok, gyorsabb számítás és jobban tanítható modellek. A képfelismerés 2012-es fordulatát a tartalmat előállító rendszerek, majd a sokféle feladatra használható nyelvi modellek megerősödése követte. A történet ebben a részben a 2020-as évek közepéig vezet. A fejlődés megértéséhez külön kell látnunk a modell képességét, a hozzáférést megkönnyítő felületet és azokat az eszközöket, amelyekkel egy alkalmazás már cselekedni is tud.

32. Adat, számítási teljesítmény és új módszerek találkozása

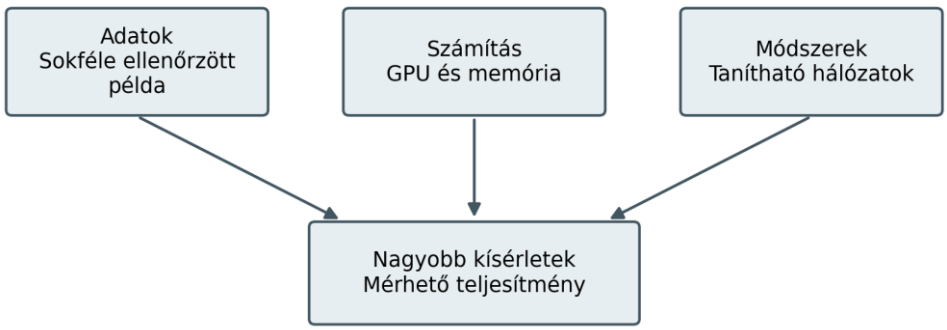
A mélytanulás sikere nem egyetlen ötlet hirtelen felbukkanásából következett. A többrétegű hálózat és a visszaterjesztés már korábban létezett. A használhatóságukat azonban korlátozta, mennyi jó példához jutottak a kutatók, mekkora hálózatot tudtak tanítani, és mennyi idő alatt derült ki, működik-e egy új elképzelés. Ezek a feltételek egymást is alakították.

Az internetből tanítható adathalmaz

Az internet sok képet tett elérhetővé, de a letölthető fényképek halmaza még nem kész tanítóanyag. Tudni kell, mit ábrázolnak, hogyan osztályozzuk őket, és mely példákat tartjuk fenn az ellenőrzéshez. A címkézés és a minőségvizsgálat jelentős emberi munka maradt. Egy rosszul azonosított kép éppen úgy félrevezetheti a tanuló eljárást, mint egy hibás megoldókulcs a diákot.

Jia Deng, Li Fei-Fei és munkatársaik 2009-ben mutatták be az ImageNetet, amely a képeket a WordNet fogalmi hierarchiájához kapcsolta. [170] A későbbi ImageNet Large Scale Visual Recognition Challenge, az ILSVRC közös mérési feladatokat és adathalmazokat adott a kutatóknak. Így az eredményeket nem csupán különböző laboratóriumok saját példáin lehetett megítélni. [171]

A képosztályozás egyik fő feladata ezer előre megadott kategória megkülönböztetése lett. Egy ilyen mérés pontosan meghatározza, mit kérünk a rendszertől. Ugyanakkor szűkíti is a sikerről szóló állítást: a kategóriák felismerése nem jelenti egy teljes jelenet minden részletének megértését, és nem bizonyít hibátlan működést más kamerák, szokatlan nézőpontok vagy új környezet esetén.



88. kép Adatok, számítás és tanítási módszerek együtt alakítják a nagyobb kísérletek lehetőségét. Saját fogalmi ábra; az összefüggést szemlélteti, nem a tényezők számszerű hozzájárulását méri.

A grafikus processzor más munkát kap

A grafikus processzorok, a GPU-k sok hasonló számtani művelet párhuzamos végrehajtására alkalmasak. A neurális hálózatok tanítása pedig nagy mennyiségben tartalmaz vektorokkal és mátrixokkal végzett számításokat. Ez kedvező találkozás volt. A megfelelően megírt program az eredetileg képi megjelenítésre fejlesztett hardvert a hálózat súlyainak tanítására használhatta.

A gyorsulás nem pusztán kényelmet jelentett. Ha egy kísérlet hamarabb véget ér, több felépítést és beállítást lehet kipróbálni. A nagyobb memória több súlynak és köztes eredménynek ad helyet. A hardver lehetőségeit azonban a szoftvernek is ki kell használnia; az adatmozgatás vagy a rosszul szervezett számítás visszafoghatja az előnyt.

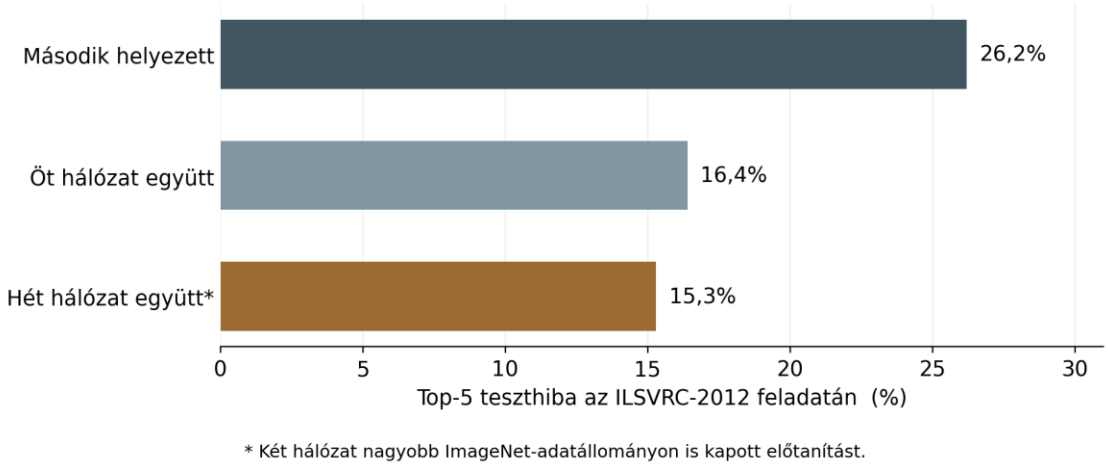


89. kép EVGA gyártmányú Nvidia GeForce GTX 580 videokártya, 2023-as fényképen. A típus a GPU-val végzett korai mélytanulás korszakát idézi; az AlexNet tanulmányában is GTX 580-asok szerepelnek. Fotó: TheStriker, CC BY-SA 4.0.

Az AlexNet eredménye

Alex Krizhevsky, Ilya Sutskever és Geoffrey Hinton 2012-es munkája, amelyet AlexNet néven ismerünk, mély konvolúciós hálózatot tanított GPU-kon. A cikkben ismertetett hálózat tanítása két GTX 580 processzoron öt-hat napot vett igénybe. A szerzők a gyors számítás mellé a tanítást segítő megoldásokat és a túlilleszkedést mérséklő eljárásokat kapcsoltak. [172]

Az ILSVRC-2012 versenyre készített legjobb változatuk 15,3 százalékos top-5 teszthibát ért el, míg a második helyezett 26,2 százalékot. A top-5 mérésben akkor számít hibának az eset, ha a helyes kategória az öt legelőrébb sorolt válasz között sincs. A legjobb eredmény több hálózat együtteséből származott; egy részük a nagyobb ImageNet-adatállományon is előtanítást kapott. [172]



90. kép Top-5 teszthibák az ILSVRC-2012 feladatán. A legjobb együttes nagyobb adathalmazon előtanított hálózatokat is tartalmazott. Saját grafikon a [172] 2. táblázatából.

A tanítás javításának egyik útja a megfelelő nemlinearitás kiválasztása volt. Másik fontos feladat lett a túlilleszkedés visszafogása. A tanítópéldák értelmet megőrző változtatásával, például kisebb képkivágásokkal, ugyanaz a tárgy többféle megjelenésben mutatható be. A rendszer így kevésbé támaszkodhat a kép egyetlen esetleges elrendezésére. Ez továbbra is gondosan megtervezett tanítás: egy olyan változtatás, amely eltünteti vagy másra cseréli a felismerendő tárgyat, már hibás példát teremthet.

A különbség azért volt meggyőző, mert közös feladaton mutatkozott meg. A korábbi fejezetekben bemutatott, kézzel kialakított képi jellemzők mellé erős alternatíva került: a rendszer a felismeréshez hasznos belső leírás jelentős részét a tanítás során alakította ki. A fordulatot a mérhető eredmény tette széles körben követhetővé.

A növekedésnek arányai is vannak

A következő években egyre fontosabb lett, hogyan osszuk el a rendelkezésre álló számítást a modell mérete és a tanítóadat mennyisége között. Hoffmann és munkatársai 2022-es, Chinchillához kapcsolódó kutatása szerint a vizsgált körülmények között több nagy nyelvi modellt túl kevés szövegen tanítottak a méretéhez képest. [173] A nagyobb paraméterszám tehát nem önmagában jelentett jobb felhasználást ugyanakkora számítási keretből.

A tanítás költségét ráadásul meg kell különböztetni a betanított modell használatának költségétől. A ritkán végzett, hosszú tanítást követheti nagyon sok rövidebb válaszgenerálás. A működő szolgáltatás értékét így az is alakítja, hány felhasználót kell kiszolgálnia, milyen gyorsan, és mekkora modellel. A tudományos eredményből itt is külön munkával lesz fenntartható műszaki rendszer.

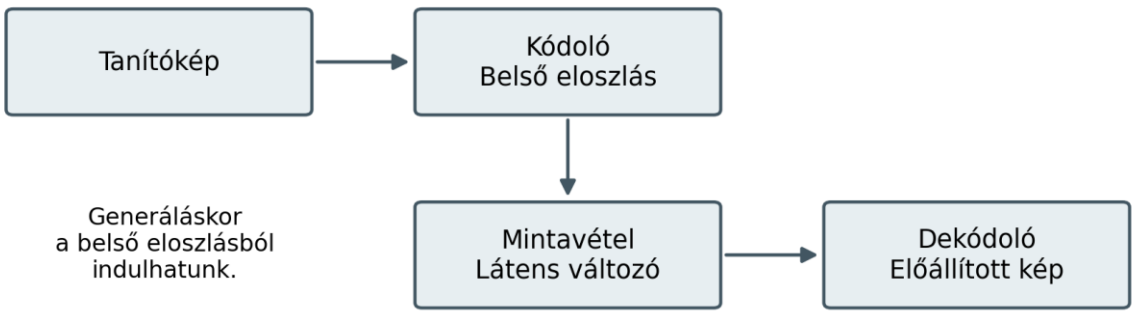
33. A gép megtanul alkotni

A képfelismerő rendszer egy megadott képet osztályoz. A generatív modell új mintát állít elő: képet, hangot, szöveget vagy más adatsort. Ehhez a tanítóanyagban megjelenő összefüggések olyan leírását kell kialakítania, amelyből új változatokat lehet létrehozni. Az „alkotás” itt először ezt a működést jelenti. A szándék, az esztétikai ítélet és a tudatos élmény kérdése külön vizsgálatot kíván.

Új minták egy megtanult térből

A generálás gondolata jóval régebbi a látványos képkészítő alkalmazásoknál. A valószínűségi modellekből is lehet mintát venni. A neurális megoldások abban hoztak új lehetőségeket, hogy nagy és összetett adatok belső szerkezetét is megtanulhatták. Az előállított képhez így nem kellett minden ecsetvonást vagy tárgyrészletet külön programutasításban megadni.

Diederik Kingma és Max Welling 2013 végén közölt, 2014-ben bemutatott variációs autoenkódolója, a VAE fontos irányt képviselt. A modell egy valószínűségi belső leírást tanul, amelyből mintát véve a dekódoló adatot állíthat elő. [174] A belső változókat gyakran látens változóknak nevezzük: közvetlenül nem a kép látható pontjai, hanem a modellben használt leírás elemei.

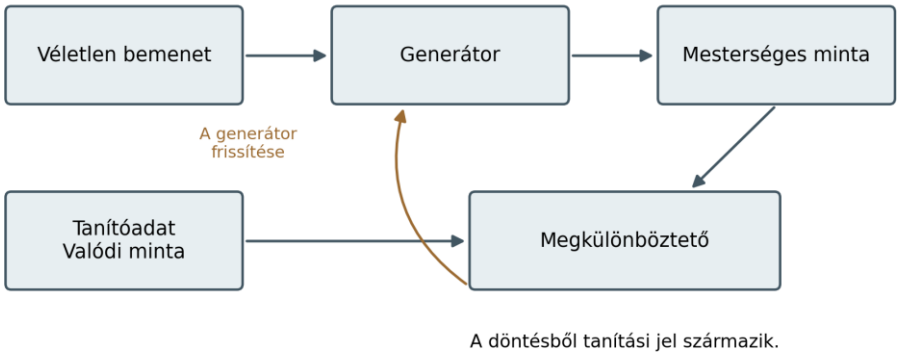


91. kép Egy variációs autoenkódoló működésének egyszerűsített magyarázata. A kódoló eloszlás paramétereit adja meg, a dekódoló a belső változó mintájából állít elő adatot. Saját ábra a [174] témájához.

Ez a gondolat segít megérteni, miért nem szükségképpen egyetlen tárolt kép előhívása történik. A mintavétel a tanult eloszlásból indul. Ettől még a modell megjegyezhet konkrét példákat is; az általánosítás és a megjegyzés arányát külön kell vizsgálni.

Két hálózat tanítja egymást

Ian Goodfellow és munkatársai 2014-ben mutatták be a generatív ellenséges hálózatokat, a GAN-okat. A generátor mesterséges mintákat készít, a megkülönböztető hálózat pedig azt tanulja, hogyan válassza el őket a tanítóadatok mintáitól. A generátor a megkülönböztető jelzéséből kap tanítási útmutatást. [175]

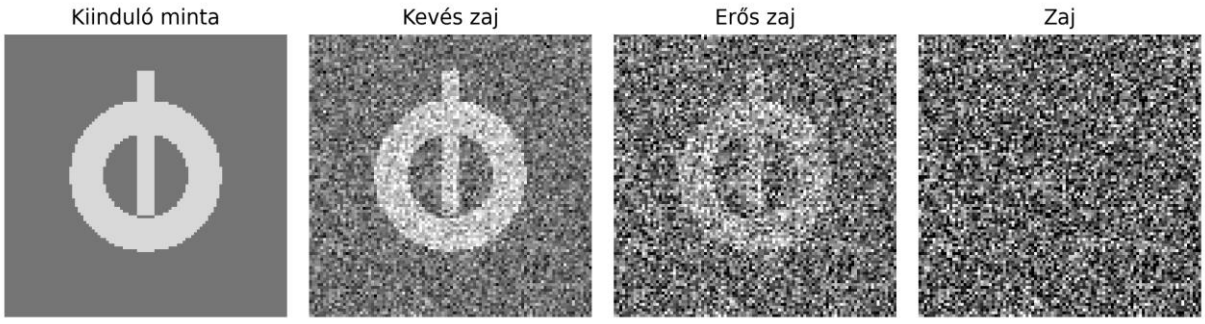


92. kép A GAN alapelve. A megkülönböztető a tanítóadat és a generátor mintáit kapja; jelzéséből a generátor számára is tanítási információ származik. Saját magyarázó ábra a [175] alapján.

A hasonlat szerint az egyik fél egyre hihetőbb képet készít, a másik egyre ügyesebben keresi az eltérést. Valójában két számítási eljárás paraméterei változnak meghatározott célfüggvények alapján. A megkülönböztető nem esztétikai zsűri, és nem ellenőrzi, hogy az ábrázolt esemény megtörtént-e. A tanítás egyensúlya is nehéz lehet: a generátor például a változatosság helyett kevés, ismétlődő mintatípusra szorítkozhat.

Képkészítés zajból

A diffúziós modellek más utat követnek. A tanításban ismert adatokhoz fokozatosan zajt adunk, és a modell egy fordított irányú, zajt csökkentő folyamatot tanul. Jascha Sohl-Dickstein és munkatársai 2015-ben már bemutatták e megközelítés egyik alapmunkáját. [176] Jonathan Ho, Ajay Jain és Pieter Abbeel 2020-as tanulmánya a diffúziós képgenerálás újabb fontos eredménye lett. [177]



93. kép Fokozatos zajhozzáadás egy saját geometriai mintához. A képsor matematikai szemléltetés, nem egy tanított diffúziós modell futása. A generatív modell a fordított irányú, zajt csökkentő lépéseket tanulja; generáláskor nem kell ugyanezt a kiinduló mintát visszaadnia. A módszerhez lásd [176–177].

Generáláskor megfelelő zajmintából indulhatunk, és a tanult eljárás több lépésben hoz létre képszerű szerkezetet. Ez nem egy tetszőleges zajfolt mögött elrejtett eredeti fénykép visszakeresése. A szöveges feltétel, például egy táj leírása, irányíthatja, milyen mintázatok alakuljanak ki. Más induló zaj más képet eredményezhet ugyanahhoz a leíráshoz is.

Robin Rombach és munkatársai 2022-es látens diffúziós modellje a zajcsökkentő folyamatot tömörített belső képreprezentációban végezte. A dekódoló ebből állította elő a részletes képet. [178] A számítás így nem minden lépésben a teljes felbontású képpontokon zajlott. A szöveges irányítás és a hatékonyabb számítás hozzájárult a módszer szélesebb használhatóságához.

A hihető kép és az eredeti tartalom

Egy könyvborítóhoz kért kép esetén a felhasználó leírhatja a témát, a kompozíciót és a hangulatot, majd válogathat a változatok között. A folyamatban emberi döntés marad, melyik kép fejezi ki jól a könyv gondolatát. A szép eredmény nem bizonyítja azt sem, hogy a modell pontosan betart minden műszaki vagy térbeli feltételt. Egy látványos szerkezet rajza még nem működő gép terve.

A „minden kép új” kijelentés ugyancsak túl erős volna. Carlini és munkatársai 2023-as vizsgálata diffúziós modellekből tanítóképekhez nagyon közeli példák kinyerését is bemutatta. [179] A generálás lehet új összeállítás és általánosítás, de tartalmazhat megjegyzett részleteket is. A történeti eredmény abban állt, hogy a gép használható új mintákat tudott előállítani; az eredetiség és a megbízhatóság megítélését ez nem tette feleslegessé.

34. A transzformer és a nyelvi modellek felemelkedése

A nyelv különösen gazdag tanítóanyag. Szövegekben együtt jelennek meg a szavak alakjai, a világra vonatkozó állítások, magyarázatok és példák. Egy következő szövegrészlet jó előrejelzéséhez a modellnek gyakran távoli összefüggéseket is figyelembe kell vennie. Ki beszél, mire utal egy névmás, milyen témában járunk, és milyen válasz illik egy kérdéshez?

A tanítójel a szövegben van

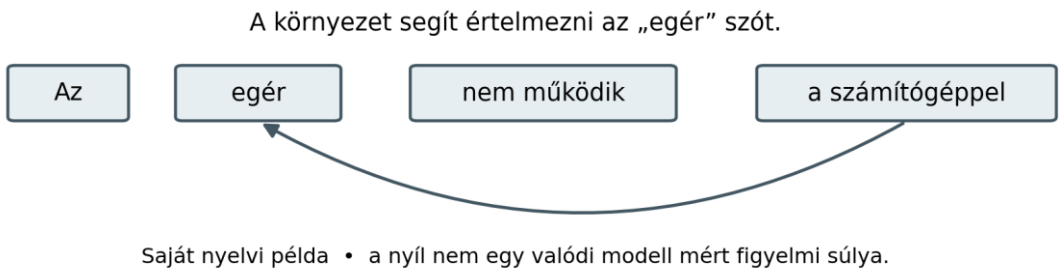
A következő token előrejelzésekor a tanító szöveget elvágjuk egy ponton: az előzmény a bemenet, a tényleges következő részlet pedig a cél. A token nem mindig teljes szó. Lehet szórész, írásjel vagy más, a szöveg kódolásához használt egység; pontos határait az adott rendszer tokenizálója határozza meg.

Így nagy mennyiségű tanítópélda állítható elő anélkül, hogy minden mondathoz ember írna külön kérdést és helyes választ. Ezt a megközelítést önfelügyelt tanulásnak nevezzük. Az emberi munka ettől nem tűnik el: valaki létrehozza az eredeti szövegeket, kiválasztja és tisztítja az adatállományt, valamint megtervezi a tanítást.

A nyelv egyszerre tartalmaz sokféle feladatmintát. Levelek, párbeszédek, programok és magyarázó szövegek más-más folytatási szabályosságot kínálnak. A gyakoriság azonban nem azonos az igazsággal. A modell megalapozott állításokat, elavult információkat, előítéleteket és kitalált történeteket is találhat ugyanabban a tanítóanyagban.

Figyelmi súlyok az előzmények között

A figyelmi mechanizmus nem a transzformerrel született. Bahdanau, Cho és Bengio 2014-ben közölt, 2015-ben bemutatott fordítási munkájában a rendszer a kimenet előállításakor a bemenet különböző részeihez eltérő súlyt rendelhetett. [180] Így nem kellett a teljes bemeneti mondat minden hasznos részletét egyetlen változatlan összefoglalásba szorítani.



94. kép Az „egér” jelentését a szövegkörnyezet is alakítja. Saját nyelvi szemléltetés; a szavak elrendezése nem egy konkrét tokenizáló kimenete, és a nyíl nem mért figyelmi súlyt mutat.

A belső figyelem, az önfigyelem során egy szövegpozíció leírása más pozíciók leírásaival kerül kapcsolatba. A rendszer tanult átalakításokból számítja ki, mely részek mennyire járuljanak hozzá az új belső reprezentációhoz. Több figyelmi fej párhuzamosan különböző kapcsolatokat kezelhet. A matematikai „figyelem” elnevezés nem állít tudatos összpontosítást, és a súlyok önmagukban nem adnak teljes magyarázatot egy válasz okára.

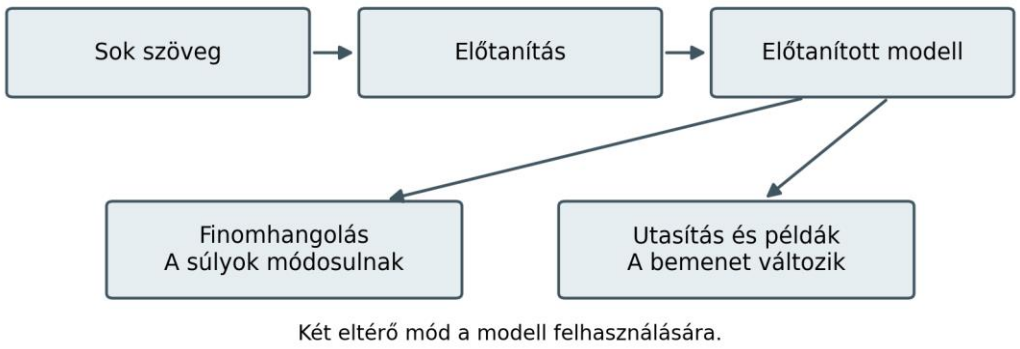
A 2017-es transzformer

Ashish Vaswani és hét szerzőtársa Attention Is All You Need című 2017-es tanulmánya a fordításra épített transzformerarchitektúrát mutatta be. Az eredeti modell kódoló és dekódoló részt tartalmazott, figyelmi műveletekkel és további számítási rétegekkel. [181] A pozíciókra vonatkozó információ segített a sorrend kezelésében, hiszen ugyanazok a szavak más sorrendben más állítást jelenthetnek.

Az architektúra a tanító mondatokon belül sok művelet párhuzamosítását tette lehetővé, elhagyva a korábbi visszacsatolt hálózatok kötelező, pozícióról pozícióra haladó feldolgozását. Ez a tanítást kedvezően alakította. Egy balról jobbra szöveget generáló nyelvi modell válasza ugyanakkor továbbra is egymás után előállított tokenekből épülhet fel. A párhuzamos tanítás és a sorozatos szövegkibocsátás nem ellentmondás.

Előtanítás és feladatra igazítás

Alec Radford és munkatársai 2018-as GPT-munkája a generatív nyelvi előtanítást feladatokra végzett finomhangolással kapcsolta össze. [183] A BERT 2018-ban közölt, 2019-ben konferencián megjelent változata más irányt hangsúlyozott: kitakart szövegrészek előrejelzésével kétoldali környezetből tanult reprezentációkat. [182] Az eredeti BERT ezért nem egyszerűen egy korábbi beszélgető GPT volt.



95. kép Előtanítás után a modell különböző módon használható. A finomhangolás a súlyokat, az utasítás és a bemeneti példa elsősorban az aktuális szövegkörnyezetet változtatja. Saját egyszerűsített magyarázó ábra.

Az előtanítás során tanult súlyok nem egy hagyományos enciklopédia betűrendes szócikkei. Sok paraméter együttesen vesz részt a válasz előállításában, és ugyanazok a kapcsolatok többféle feladathoz is hozzájárulhatnak. Emiatt a modell egy kérdésre új megfogalmazásban válaszolhat, de egy konkrét állítás forrását nem feltétlenül tudja azonosítani. A kereshető dokumentumtár és a megtanult paraméterekben tárolt tudás különböző módon szolgálja a felhasználót.

A nagy nyelvi modellben sok feladathoz felhasználható belső leírás alakulhat ki. Ezt külön tanítással egy célra igazíthatjuk, de a felhasználás történhet a bemeneti szövegben megadott útmutatással és néhány példával is. A két esetet fontos megkülönböztetni. Egy utasítás vagy példa a beszélgetésben rendszerint nem írja át közvetlenül a modell súlyait; az aktuális válasz feltételeit változtatja meg.

A GPT-3 2020-as tanulmánya a 175 milliárd paraméteres változaton is vizsgálta a néhány bemeneti példával elvégezhető feladatokat. A fordítás, a kérdésválasz és más tesztek mellett a korlátokat és a nagy internetes tanítóanyagból adódó mérési problémákat is tárgyalta. [184] A nagy méret és az előtanítás sokoldalú eszközt adott, de a hasznos, következetes asszisztensi viselkedéshez további munka kellett.

35. A ChatGPT-pillanat

2022. november 30-án az OpenAI nyilvános kutatási előzetesként elérhetővé tette a ChatGPT-t. A kezdeti szolgáltatás a GPT-3.5 modellcsalád egyik tagjának beszélgetésre igazított változatára épült. [186] A nagy nyelvi modellek ekkor már léteztek; a változás abban is állt, hogyan találkozott velük a felhasználó. Hétköznapi mondatokban kérhetett magyarázatot, szöveget vagy programot, majd az eredményre visszakérdezhetett.

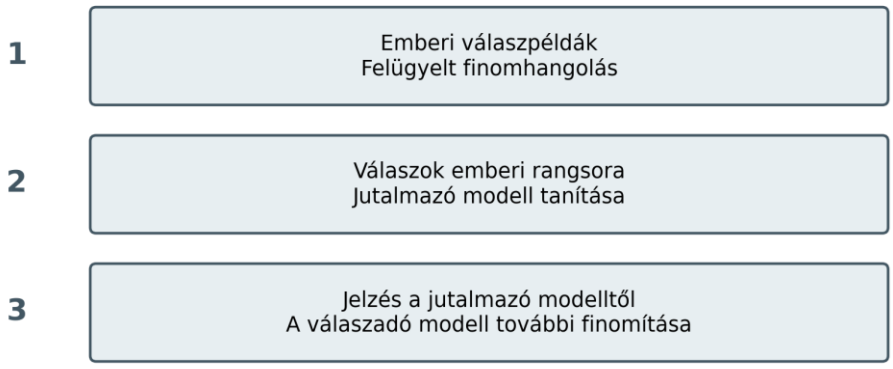
A beszélgetés használható munkafelületté válik

Egy hagyományos programban a készítő előre meghatározza a menüpontok és mezők jelentős részét. A beszélgető felületen a kérés maga is rugalmasabb. „Írd át közérthetőbben”, „adj rá példát”, „rövidítsd a felére” vagy „ezt a szót más értelemben használtam” új lépést indíthat ugyanabban a munkában. A felhasználó fokozatosan pontosíthatja, mire van szüksége.

Ez különösen akkor hasznos, ha a feladat első megfogalmazása még bizonytalan. Egy levél vázlata közben derülhet ki, milyen hangnemet szeretnénk. Egy műszaki magyarázatnál az első válasz után látszik, hol maradt hiány a megértésben. A párbeszéd az elkészítendő eredmény alakításának eszköze lett. Ehhez azonban a programnak az utasítások jellegét is követnie kellett, nem csupán egy szövegrészlet valószínű folytatását előállítania.

Emberi visszajelzésből tanított viselkedés

Az InstructGPT-ről Ouyang és munkatársai 2022-ben közölt tanulmánya emberi példákkal és értékelésekkel végzett utólagos tanítást írt le. Először megfelelő válaszokat mutattak a modellnek, majd emberek válaszváltozatokat rangsoroltak. A rangsorokból tanított jutalmazó modell jelzése segítette a további finomítást. [185] A folyamat az emberi visszajelzésből történő megerősítéses tanulás, az RLHF fontos példája lett.



96. kép Az InstructGPT emberi visszajelzésre építő tanításának fő szakaszai. A harmadik szakaszban a jutalmazó modell ad jelzést a válaszadó modell további tanításához. Saját összefoglalás a [185] alapján.

A ChatGPT bemutatásakor az OpenAI hasonló, párbeszédes példákat is alkalmazó eljárásról számolt be. [186] A tanítás célja az volt, hogy a rendszer jobban kövesse a felhasználó szándékát és használhatóbb választ adjon. A válaszok rangsorolása azonban emberek és értékelési szempontok

választásait tükrözi. Nem egy minden állítást a világ tényeivel összevető, tévedhetetlen ellenőrző rendszer épül belőle.

A gördülékenység félrevezető is lehet

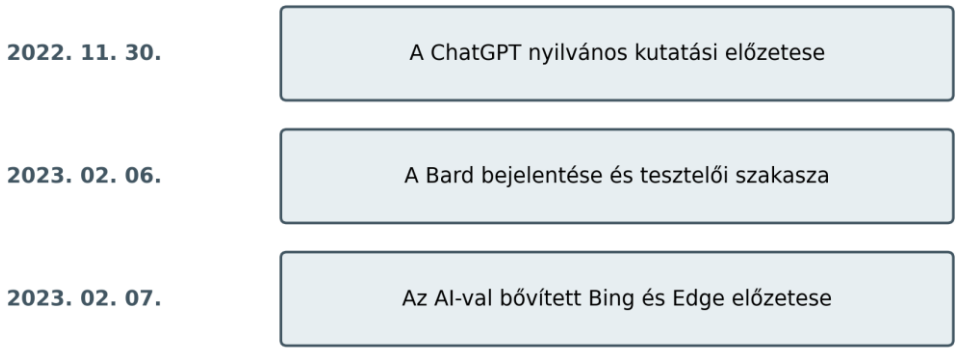
A nyilvános bemutató már a kezdetkor jelezte, hogy a ChatGPT hihetően hangzó, de hibás válaszokat is adhat. [186] Egy összefüggő bekezdés, egy pontosnak látszó dátum vagy egy szabályosan formázott hivatkozás ezért önmagában kevés. A rendszer olyankor is folytathat egy mintázatot, amikor az állításhoz nincs megfelelő alapja.

A hétköznapi munkában ebből kettős tapasztalat adódik. Egy vázlat, magyarázat vagy megfogalmazási változat gyorsan elkészülhet, miközben a nevek, számok és források ellenőrzése megmarad. A helyes munkamegosztás függ a feladattól. Egy születésnapi köszöntőnél más hiba számít súlyosnak, mint egy mérnöki számítás vagy ügyfélminősítés előkészítésében.

Mérhető haszon egy meghatározott feladaton

Shakked Noy és Whitney Zhang 2023-as kísérlete szakmai szövegalkotási feladatokon vizsgálta a ChatGPT hatását. A vizsgált feladatokban csökkent az átlagos elkészítési idő, és a független értékelők szerint javult a szövegek minősége. [187] Ez dokumentált eredmény egy meghatározott környezetben. Nem jelenti azt, hogy minden foglalkozás minden munkafolyamata ugyanennyit nyer, vagy hogy az ellenőrzés ideje mindig elhanyagolható.

A szolgáltatások köre gyorsan bővült. A Google 2023. február 6-án jelentette be a Bard kísérleti beszélgető szolgáltatást, kezdetben kiválasztott tesztelőknek. [188] A Microsoft február 7-én mutatta be az AI-val bővített Bing és Edge előzetesét, amely a keresést beszélgetéssel is összekapcsolta. [189] A versenyben a felület és a szolgáltatásba építés ugyancsak fontossá vált.



97. kép Három nyilvános bejelentés a beszélgető szolgáltatások történetéből. A Bard ekkor még tesztelői szakaszban indult; a dátumok nem azonos körű hozzáférést jelentenek. Saját időrendi összefoglaló a [186, 188–189] alapján.

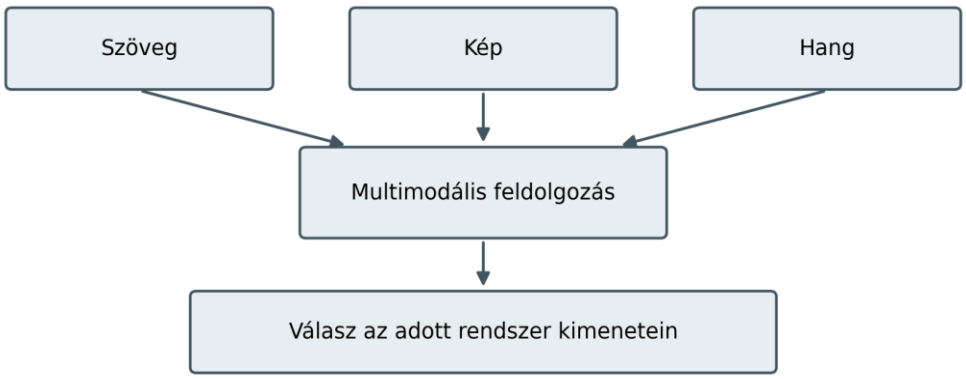
A „ChatGPT-pillanat” ezért technikai és használati fordulatot egyaránt jelöl. A felhasználó közvetlenül találkozott a generatív modellel, és saját feladataival próbálhatta ki. A korábbi, többnyire háttérben működő AI-alkalmazások mellé olyan eszköz került, amellyel beszélgetni lehetett a munkáról. Ettől kezdve a képességekről és a korlátokról szóló vita is személyes tapasztalatokkal bővült.

36. Egy rendszer már lát, hall és beszél

A beszélgető gép eleinte elsősorban szövegmezőn keresztül vált ismertté. A következő lépésben képet, hangot és más adatot is be lehetett vonni ugyanabba a feladatba. Egy képernyőképről feltett kérdés, egy hangosan elmondott utasítás vagy egy lefényképezett dokumentum már többféle bemenet összekapcsolását igényli. Az ilyen rendszereket nevezzük multimodálisnak.

Többféle adat összekapcsolása

A multimodalitásnak több műszaki megoldása lehet. Egy alkalmazás külön beszédfelismerőt, nyelvi modellt és beszédszintetizátort kapcsolhat össze. Más rendszer többféle bemenet és kimenet kezelését szorosabban egy modellben tanulja meg. A felhasználó mindkét esetben beszélgetést tapasztalhat, miközben a késleltetés, a megőrzött információ és a hibák útja eltér.



98. kép Többféle bemenet összekapcsolása. A rajz fogalmi séma: a ténylegesen kezelt bemenetek, kimenetek és a belső felépítés rendszerenként eltérnek. Saját ábra; nem valamelyik termék teljes architektúrája.

A CLIP 2021-es munkája képek és a hozzájuk tartozó szövegek összekapcsolásából tanult átvihető vizuális reprezentációkat. [190] A GPT-4 2023-as bemutatása képi és szöveges bemenetet kezelő, szöveget kibocsátó modellről szólt. [191] A képolvasás tehát nem volt azonos azzal, hogy ugyanaz a modell képet is rajzol vagy hangot is előállít.

A 2024. május 13-án bemutatott GPT-4o a szöveg, a kép és a hang szorosabb összekapcsolásának újabb állomása lett. A fejlesztő egyetlen hálózattal, végponttól végpontig tanított rendszerként írta le. [192] A bemutatott modellképességeket és az alkalmazásban ténylegesen elérhető funkciókat itt is külön kell kezelni: a bevezetés fokozatos volt. A történeti bemutató nem minden későbbi termékváltozat képességlistája.

Egy hibaüzenet fényképe alapján adott jó magyarázat hasznos teljesítmény. A „látás” hétköznapi szavából azonban nem következik automatikusan, hogy a rendszernek emberi látásélménye van. Ebben a fejezetben a mérhető bemenetkezelésről és működésről beszélünk; a tudat kérdését ezek önmagukban nem döntik el.

A modell külső információhoz jut

Az előtanított modell súlyaiban őrzött összefüggések mellett egy alkalmazás dokumentumokból is visszakereshet részleteket. Patrick Lewis és munkatársai 2020-as RAG-munkája az előtanított generálást visszakereshető információval kapcsolta össze. [193] Ilyenkor a válaszhoz adott szövegkörnyezet változhat anélkül, hogy minden új dokumentum miatt a teljes modellt újratanítanánk.

A visszakeresés ugyanakkor csak akkor segít, ha a megfelelő, megbízható részleteket találja meg, és a modell azokat helyesen használja. Egy forrás megjelenítése nem bizonyítja, hogy az alátámasztja a mellette álló mondatot. A forráshoz kötött válasz ellenőrizhetőbb lehet, de az ellenőrzés feladata megmarad.

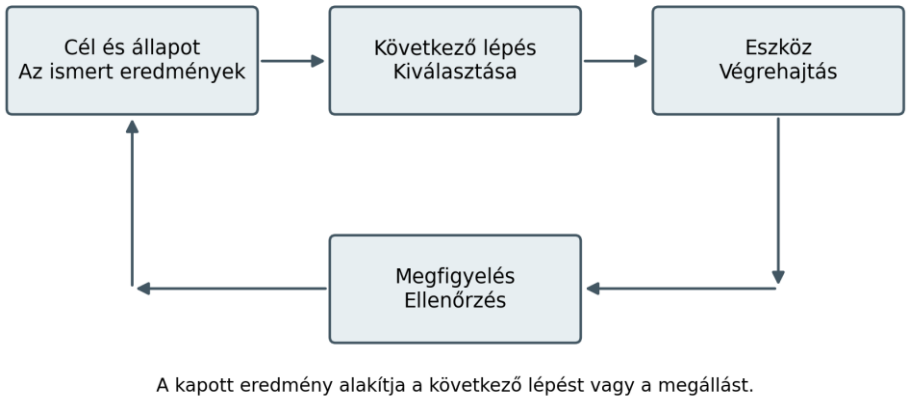
A válasz eszközhívást indíthat

A modell számológépet, keresőt vagy adatbázis-lekérdezést is kérhet. Az OpenAI 2023. júniusi függvényhívási bejelentése olyan megoldást mutatott be, amelyben a modell strukturált adatként adja meg egy külső művelet szükséges paramétereit. [194] A kérést a környező szoftver értelmezi és hajtja végre. A művelet sikeréről visszaadott eredmény ezután új információt ad a modellnek.

Ez lényeges határ. Egy levél szövegének megírása önmagában nem küld el üzenetet, és egy útiterv nem foglal szállást. A külső művelethez megfelelő kapcsolat és jogosultság kell. A modell válaszának minősége és a végrehajtás biztonságos megszervezése egymáshoz kapcsolódó, de külön műszaki feladat.

Az ügynök újra megvizsgálja az eredményt

Az AI-ügynök kifejezés olyan rendszerre utalhat, amely egy cél érdekében több lépést választ, eszközt használ, majd a kapott eredmény alapján folytatja vagy módosítja a munkát. A Yao és munkatársai által 2022-ben közölt, 2023-ban bemutatott ReAct megközelítés a feladat közbeni következtetést és cselekvést váltogatta. [195] Az ügynökök fogalma régebbi a nagy nyelvi modelleknél; itt ezekkel összekapcsolt változatuk válik fontossá.



99. kép Eszközhasználó ügynök egyszerűsített visszacsatolási folyamata. A végrehajtás megfigyelt eredménye bekerül a következő lépés állapotába. Saját szemléltetés a [195–196] témájához; nem jelenti, hogy a modell súlyai minden lépésben módosulnak.

Saját példánkban egy rendszer egy mappában található számlákról készít összesítést. Megkeresi a dokumentumokat, kiolvassa az összegeket, számol, és összeveti az eredményt a forrásokkal. Ha egy fájl nem olvasható, más eljárást próbálhat vagy jelezheti a hiányt. Az elkészült táblázat így egy több lépésből álló munkafolyamat eredménye. A folyamat közben összegyűlt megfigyeléseket a rendszer munkakörnyezetében is tárolhatja. Ez nem feltétlenül a nyelvi modell újratanítása. A következő eszközhívás már gazdagabb állapotleírásból indulhat akkor is, ha a modell súlyai végig változatlanok maradnak.

Az OpenAI 2025. július 17-i ChatGPT agent bemutatója már böngészővel és más eszközökkel végzett összetettebb feladatokat ismertetett. [196] A hosszabb folyamat azonban több hibalehetőséget is hoz. Egy tévesen kiolvasott összeg későbbi számításokban is megjelenhet, egy külső dokumentum pedig félrevezető utasítást rejthet. Az elérhető jogosultságokat, az ellenőrzést és a leállítás lehetőségét ezért a működő rendszerben is rendezni kell.

A válaszadás és a végrehajtás összekapcsolása új fejezetet nyitott a használatban. Egyre több feladatnál már nem csupán az a kérdés, helyes-e a gép mondata, hanem az is, mit tett a nevében eljáró szoftver, milyen adatok alapján, és hogyan ellenőrizhető az eredmény. A kibernetika korai visszacsatolási gondolata itt új környezetben tér vissza: a következő lépést a korábbi cselekvés megfigyelt következménye is alakítja.

Források és további olvasmányok

A hivatkozások és képek számozása az előző részeket folytatja. A tanulmányok mellett a termékbejelentések a korabeli bemutatás és bevezetés dátumát dokumentálják. A fejlesztők állításait nem tekintjük a mindenkori szolgáltatás független értékelésének. A magyarázó rajzok saját szerkesztések; a zajos képsor nem generatív modell futási eredménye.

[170] Jia Deng – Wei Dong – Richard Socher – Li-Jia Li – Kai Li – Li Fei-Fei: ImageNet: A Large-Scale Hierarchical Image Database. CVPR, 2009, 248–255. DOI: 10.1109/CVPR.2009.5206848. Eredeti tanulmány

[171] Olga Russakovsky és munkatársai: ImageNet Large Scale Visual Recognition Challenge. International Journal of Computer Vision, 115, 2015, 211–252. DOI: 10.1007/s11263-015-0816-y. A közös mérési feladatok és adatok története. A szervezők oldala · Tanulmány

[172] Alex Krizhevsky – Ilya Sutskever – Geoffrey E. Hinton: ImageNet Classification with Deep Convolutional Neural Networks. NIPS, 2012, 1097–1105. Különösen a 6. szakasz és a 2. táblázat. Eredeti tanulmány

[173] Jordan Hoffmann és munkatársai: Training Compute-Optimal Large Language Models. 2022, arXiv:2203.15556. A Chinchilla-kísérletek és a modellméret, adatmennyiség, számítási keret arányai. Tanulmány

[174] Diederik P. Kingma – Max Welling: Auto-Encoding Variational Bayes. ICLR, 2014; első előzetes közlés 2013. december 20. arXiv:1312.6114. Eredeti tanulmány és verziók

[175] Ian Goodfellow és munkatársai: Generative Adversarial Nets. Advances in Neural Information Processing Systems, 27, 2014, 2672–2680. Eredeti tanulmány

[176] Jascha Sohl-Dickstein – Eric Weiss – Niru Maheswaranathan – Surya Ganguli: Deep Unsupervised Learning using Nonequilibrium Thermodynamics. ICML, PMLR 37, 2015, 2256–2265. Eredeti tanulmány

[177] Jonathan Ho – Ajay Jain – Pieter Abbeel: Denoising Diffusion Probabilistic Models. NeurIPS, 2020. arXiv:2006.11239. Eredeti tanulmány

[178] Robin Rombach – Andreas Blattmann – Dominik Lorenz – Patrick Esser – Björn Ommer: High-Resolution Image Synthesis with Latent Diffusion Models. CVPR, 2022, 10684–10695. Eredeti tanulmány

[179] Nicholas Carlini és munkatársai: Extracting Training Data from Diffusion Models. USENIX Security, 2023, 5253–5270. A tanítóanyagból megjegyzett képek kinyerésének kutatási vizsgálata. Tanulmány és előadás

[180] Dzmitry Bahdanau – Kyunghyun Cho – Yoshua Bengio: Neural Machine Translation by Jointly Learning to Align and Translate. ICLR, 2015; első előzetes közlés 2014. arXiv:1409.0473. Eredeti tanulmány és verziók

[181] Ashish Vaswani – Noam Shazeer – Niki Parmar – Jakob Uszkoreit – Llion Jones – Aidan N. Gomez – Łukasz Kaiser – Illia Polosukhin: Attention Is All You Need. NIPS, 2017. arXiv:1706.03762. Eredeti tanulmány

[182] Jacob Devlin – Ming-Wei Chang – Kenton Lee – Kristina Toutanova: BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. NAACL, 2019, 4171–4186; első előzetes közlés 2018. DOI: 10.18653/v1/N19-1423. Eredeti tanulmány

[183] Alec Radford – Karthik Narasimhan – Tim Salimans – Ilya Sutskever: Improving Language Understanding by Generative Pre-Training. OpenAI, 2018. A fejlesztők közlése és tanulmány

[184] Tom B. Brown és munkatársai: Language Models are Few-Shot Learners. NeurIPS, 2020. arXiv:2005.14165. A GPT-3 és a bemeneti példákból történő feladatkövetés vizsgálata. Eredeti tanulmány

[185] Long Ouyang és munkatársai: Training Language Models to Follow Instructions with Human Feedback. NeurIPS, 2022. arXiv:2203.02155. InstructGPT, emberi válaszpéldák, rangsorok és jutalmazó modell. Eredeti tanulmány

[186] OpenAI: Introducing ChatGPT. 2022. november 30. A nyilvános kutatási előzetes bemutatása, a tanítás és az akkori korlátok leírásával. Korabeli fejlesztői bejelentés

[187] Shakked Noy – Whitney Zhang: Experimental Evidence on the Productivity Effects of Generative Artificial Intelligence. Science, 381, 2023, 187–192. DOI: 10.1126/science.adh2586. Szakmai írásfeladatokon végzett kísérlet; az eredményeket ehhez a környezethez kötjük. Tanulmány és kiadási adatok

[188] Sundar Pichai: An Important Next Step on Our AI Journey. Google, 2023. február 6. A Bard bejelentése és a kiválasztott tesztelőkkel induló szakasz. Korabeli fejlesztői bejelentés

[189] Yusuf Mehdi: Reinventing Search with a New AI-Powered Microsoft Bing and Edge, Your Copilot for the Web. Microsoft, 2023. február 7. A kereső és a böngésző AI-val bővített előzetese. Korabeli fejlesztői bejelentés

[190] Alec Radford és munkatársai: Learning Transferable Visual Models from Natural Language Supervision. ICML, PMLR 139, 2021, 8748–8763. CLIP. Eredeti tanulmány

[191] OpenAI: GPT-4 Technical Report. 2023, arXiv:2303.08774. Képi és szöveges bemenetet kezelő, szöveges kimenetű modell. Műszaki jelentés · 2023. március 14-i bemutató

[192] OpenAI: Hello GPT-4o. 2024. május 13. A multimodális modell bemutatója, a működés és a fokozatos bevezetés leírásával. Korabeli fejlesztői bejelentés

[193] Patrick Lewis és munkatársai: Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. NeurIPS, 2020, 9459–9474. arXiv:2005.11401. Eredeti tanulmány

[194] OpenAI: Function Calling and Other API Updates. 2023. június 13. A külső művelethez szükséges paraméterek strukturált előállítása és az eszközkapcsolat. Korabeli fejlesztői bejelentés

[195] Shunyu Yao és munkatársai: ReAct: Synergizing Reasoning and Acting in Language Models. ICLR, 2023; első előzetes közlés 2022. arXiv:2210.03629. Eredeti tanulmány · Szerzői projektoldal

[196] OpenAI: Introducing ChatGPT Agent: Bridging Research and Action. 2025. július 17. Történeti bemutató a böngészőt és egyéb eszközöket használó rendszerről; nem a mindenkori termékkínálat leírása. Korabeli fejlesztői bejelentés

Képforrások és felhasználási adatok

88. kép Saját fogalmi ábra az adat, számítás és módszer együttes szerepéről.

89. kép TheStriker: EVGAGeforceGTX580.jpg. 2023. február 9-i fénykép, 2010-es kártyatípusról. A letöltött Commons-változat további módosítás nélkül. CC BY-SA 4.0. Kép és metaadatok · Licenc

90. kép Saját grafikon a [172] 2. táblázatából. Az együttesek eltérő tanítóadatot is használnak.

91. kép Saját egyszerűsített magyarázat a variációs autoenkódoló belső mintavételéről; lásd [174].

92. kép Saját működési séma a generátor és a megkülönböztető tanításáról; lásd [175].

93. kép Saját geometriai minta és programmal hozzáadott zaj. Nem tanított modell kimenete, és nem valódi rekonstrukciós kísérlet. A módszerhez [176–177].

94. kép Saját magyar mondat és szemléltető kapcsolat. Nem valódi modellből kinyert figyelmi térkép.

95. kép Saját fogalmi ábra az előtanítás, finomhangolás és a bemeneti útmutatás különbségéről.

96. kép Saját tanítási összefoglaló a [185] módszeréhez; nem állítja, hogy minden későbbi modell pontosan ezt az eljárást használja.

97. kép Saját időrendi felsorolás a [186, 188–189] bejelentéseinek dátumaiból; nem felhasználószám- vagy terjedési grafikon.

98. kép Saját fogalmi séma a szöveges, képi és hangbemenet összekapcsolásáról.

99. kép Saját visszacsatolási séma az eszközhasználó rendszer működéséről; lásd [195–196].

VIII. rész – A történet elér a jelenhez

Az eddigi történetben a gépek egyre többféle feladatot tanultak meg. A záró rész azt vizsgálja, mit jelent ez a fizikai világban, mennyit árul el a teljesítmény a megértésről, és milyen anyagi feltételei vannak a fejlődésnek. Végül eljutunk azokhoz a rendszerekhez, amelyek már a tudományos kutatást és az újabb gépek fejlesztését is segítik. A jelenről itt a 2026. november 8-ig ellenőrzött források alapján beszélünk; az előrejelzéseket különválasztjuk a megfigyelt eredményektől.

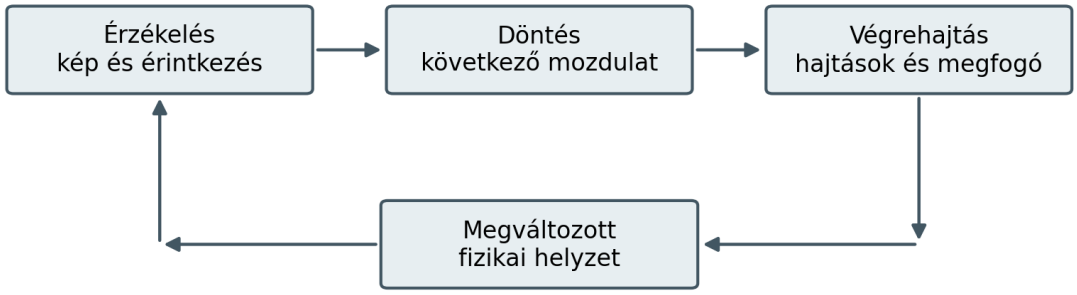
37. Az intelligencia visszakapja a testét

A robotika végigkísérte a mesterséges intelligencia történetét. Grey Walter teknősei és Shakey már jóval a nyelvi modellek előtt kapcsolatot teremtettek az érzékelés és a cselekvés között. A fejezet címében szereplő visszatérés ezért a kutatás súlypontjának újabb eltolódását jelenti: a tanuló rendszerek a képernyőről ismét egyre többféle testbe kerülnek.

Egy robotnak ugyanabban a pillanatban kell figyelembe vennie a környezetét és saját lehetőségeit. A pohár helyzetének felismerése még nem mondja meg, hol szabad megfogni, mennyire kell megszorítani, és hogyan kerülhető el a szomszédos tárgy meglökése. A mozdulat megváltoztatja a helyzetet, ezért a robotnak az eredményt is érzékelnie kell.

A szavaktól a mozdulatokig

A 2023-ban bemutatott RT-2 a képi és nyelvi előtanítást robotokkal gyűjtött tapasztalatokkal kapcsolta össze. A rendszer kimenetében a mozdulatok is kódolható egységként jelentek meg. A cél az volt, hogy a webes tanulásból származó tárgyismeret segítse a cselekvés kiválasztását is. A kísérletek nem jelentettek minden helyzetben működő háztartási robotot. [197]



A mozgás eredménye a következő döntés bemenete.

100. kép A testtel rendelkező rendszer érzékelés és cselekvés közötti kapcsolata. Saját egyszerűsített ábra; a visszacsatolás nem feltétlenül módosítja a modell súlyait.

A Physical Intelligence 2024 októberében ismertetett π0 modellje különböző robotokkal gyűjtött adatokat használt, és folytonos mozgásparancsok előállítására készült. [198] A 2025 szeptemberében bemutatott Gemini Robotics 1.5 család ugyancsak a látás, a nyelvi utasítások, a tervezés és a robotvezérlés összekapcsolását célozta. [199] Ezek a munkák azt mutatták, hogyan kerülhet át

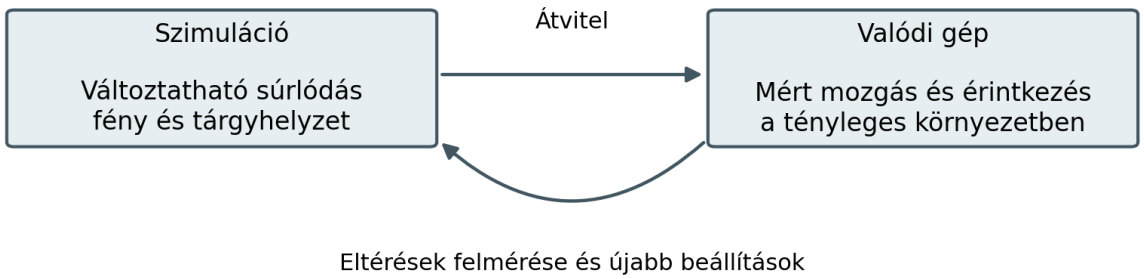
többféle feladatra használható előtanítás a robotikába. A közölt bemutatók mellett továbbra is számít, hány próbálkozásból sikerül a feladat, és mennyi segítséget kap a rendszer.

A fizikai világ nem várja meg az új választ

Egy szöveges válasz rendszerint újrafogalmazható. A leejtett pohár azonban már eltört, és a mozgó jármű akkor is halad, amikor a számítógép még értékeli a helyzetet. A robot irányításában ezért a pontosság mellett a válaszidő és a hibák következménye is meghatározó.

A tapintás különösen nehezen pótolható pusztán kameraképpel. Egy tárgy súlyát, csúszósságát vagy rugalmasságát gyakran csak érintés közben lehet megbecsülni. A ruhadarab, a kábel és a papírzacskó alakja a megfogástól is változik. A robotnak egyszerre kell kezelnie a bizonytalan észlelést, a mechanika korlátait és a tárgy viselkedését. Ez a 15. fejezetben tárgyalt humanoidprobléma mai folytatása.

A szimuláció sok gyakorlást tesz lehetővé anélkül, hogy minden hibánál valódi gép sérülne. A számítógépes környezetben változtatható a súrlódás, a megvilágítás vagy a tárgyak helyzete. A valódi anyagok és érzékelők azonban eltérhetnek a szimuláció feltételezéseitől. A sikeres virtuális tanulást ezért a tényleges gépen végzett ellenőrzésnek kell követnie.



101. kép A szimuláció és a valódi gépen végzett próba egymást kiegészítő szerepe. Saját fogalmi ábra; nem egy konkrét robot tanításának teljes folyamata.

Az önvezetésnek is van működési tartománya

Az önvezető járművek jól példázzák a képesség és az alkalmazási feltétel kapcsolatát. A Waymo biztonsági módszertana egy meghatározott működési tartományhoz köti a rendszer értékelését. Ebbe a terület, az utak és a környezeti feltételek is beletartoznak. [200] Egy adott városban igazolt működésből nem következik automatikusan, hogy a jármű minden országút, hóvihar és útépítés között is ugyanúgy helytáll.

A humanoid forma azért vonzó, mert az épületeket és eszközöket többnyire emberi testhez alakítottuk. Ugyanez a forma járási, egyensúlyozási és energiaellátási feladatokat is hoz. Egy kerekes gép vagy egy rögzített robotkar bizonyos munkákban célszerűbb lehet. A hasznosságot a feladat teljesítése dönti el, nem az emberhez való külső hasonlóság.

A vezérlés több időléptékben dolgozhat. A magasabb szintű tervező kiválasztja a következő részfeladatot, míg a gyors szabályozó a hajtások és az egyensúly pillanatnyi állapotát kezeli. A biztonsági funkciókat is ehhez kell tervezni: az ütközés elkerülése és a leállítás lehetősége nem függhet attól, hogy a társalgó rendszer éppen helyesen értette-e a mondatot. Az egész gép megbízhatósága a részek együttműködésén múlik.

Egy négyéves gyerek hétköznapi ügyessége mögött élő test, fejlődő idegrendszer és folyamatos saját tapasztalatszerzés áll. A gyerek kipróbálja, mi történik, segítséget kér, és a látványt az érintéssel, a mozgással meg a társas helyzettel együtt tanulja. A robotok egyes részfeladatokban már túlszárnyalhatják, de ez nem adja meg magától a gyerek sokféle helyzetre kiterjedő alkalmazkodását. Az általános ügyességhez a sikeres mozdulatok mellett az is kell, hogy a rendszer felismerje, mikor bizonytalan, és mikor kell megállnia.

38. Mit ért meg a mesterséges intelligencia?

Ha egy gép helyesen magyaráz el egy fizikai jelenséget, megold egy feladatot és válaszol az ellenvetésekre, indokolt komolyan venni a teljesítményét. A megértés kérdéséhez azonban azt is vizsgálnunk kell, hogyan boldogul megváltozott feltételekkel. Képes-e különválasztani a lényeges összefüggést az ismerős megfogalmazástól? Észreveszi-e, ha a kérdés hibás feltevést tartalmaz?

A teljesítménynek többféle bizonyítéka lehet

Egy bemagolt mondat visszaadása és egy új következmény levezetése eltérő teljesítmény. A modell értékelésénél ezért hasznosak az új példák, az átfogalmazott kérdések és az olyan feladatok, amelyekben a körülmények egy részét megváltoztatjuk. Egyetlen jó válasz kevesebbet mond, mint több, egymással összefüggő próba.

Emily Bender és Alexander Koller 2020-as állásfoglalása arra figyelmeztetett, hogy a nyelvi forma feldolgozását nem szabad automatikusan azonosítani a jelentés megértésével. [201] Ez egy elméleti álláspont, nem minden későbbi, érzékelőkkel és eszközökkel összekapcsolt rendszer végleges minősítése. A vita egyik tétje éppen az, milyen kapcsolat kell a jelek használata és a világban szerzett tapasztalat között.

Intelligencia	Milyen feladatokat és milyen szinten old meg?
Megértés	Hogyan használ összefüggéseket új helyzetben?
Tudat	Van-e szubjektív élménye?

102. kép Az intelligencia, a megértés és a tudat vizsgálata eltérő kérdéseket tesz fel. Saját szerkesztésű fogalmi összefoglaló; a sorok nem fejlődési fokozatok vagy rangsorok.

A tanult belső reprezentációk lehetővé tehetik fogalmi kapcsolatok és feladatmegoldási minták használatát. Az, hogy a rendszer számításokat végez, önmagában nem cáfolja ezt: az emberi gondolkodásnak is van fizikai megvalósulása. Ugyanakkor a gördülékeny beszéd önmagában nem igazolja az emberével azonos megértést. A pontosabb kérdés az, mely képességet, milyen körülmények között és milyen bizonyítékkal tulajdonítjuk a gépnek.

A tudás a modellben nem feltétlenül elkülönült, visszakereshető tételek soraként van jelen. Ugyanaz a tanult összefüggés többféle válaszban is szerepet kaphat, és ugyanarra a kérdésre az eltérő megfogalmazás más választ hívhat elő. Egy adatbázissal összekapcsolt rendszer részben másképp működik: ott a friss dokumentum tartalma is a válasz bemenetévé válhat. A tárolt dokumentum, a modell tanult képessége és a válaszban tett állítás ezért három külön ellenőrzési pont.

Következtetés és a következtetés magyarázata

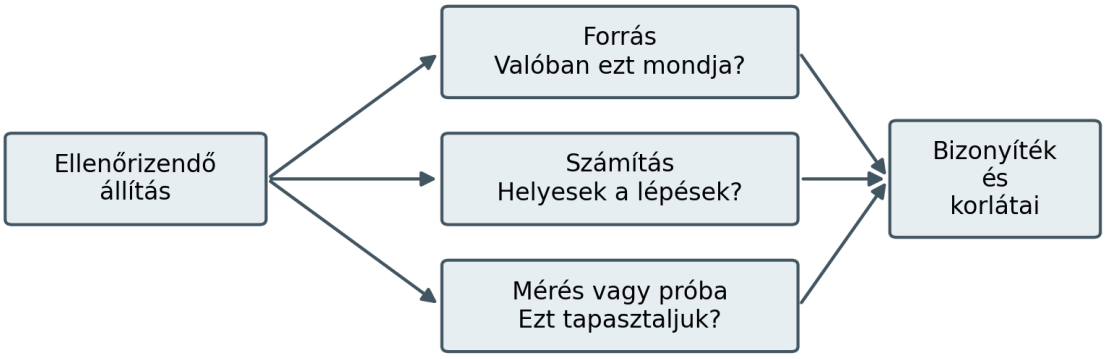
A következtető modellek fejlődésének egyik iránya a több köztes számítás és az ellenőrizhető feladatokra adott jutalmazás. A DeepSeek-R1 2025-ös munkája megmutatta, hogy megerősítéses tanulással javítható a modellek teljesítménye ilyen feladatokon. [202] Ez műszakilag vizsgálható eredmény, amelyhez nem szükséges előbb eldönteni a tudat kérdését.

A közölt gondolatmenet viszont nem feltétlenül hű jegyzőkönyve a válasz kialakulásának. Turpin és munkatársai 2023-ban olyan kísérleteket közöltek, amelyekben a bemenet torzító részletei befolyásolták a választ, a magyarázat mégsem nevezte meg őket. [203] Egy részletes levezetés akkor értékes, ha a lépései külön is ellenőrizhetők. A magyarázat hosszúsága nem helyettesíti ezt.

Miért születhet hihető tévedés?

A hallucináció kifejezést az AI esetében többnyire olyan állításra használjuk, amely megalapozottnak látszik, de hibás vagy nincs megfelelően alátámasztva. Nem kell hozzá feltételeznünk emberi értelemben vett érzékcsalódást. A modell olyan folytatást is előállíthat, amely nyelvileg jól illeszkedik a kérdéshez, miközben az állításnak nincs valós megfelelője.

Vegyünk egy saját példát: a rendszer egy régi magyar műszaki könyvhöz hitelesnek látszó oldalszámot ad, mert a kérdés azt sugallja, hogy ott szerepel egy idézet. Ha nem látta a könyv megfelelő oldalát, a pontos szám pusztán a pontosság látszatát kelti. A hiba javításához az eredeti kiadás ellenőrzése kell, nem ugyanannak a válasznak a magabiztosabb megismétlése.



103. kép Egy állítás ellenőrzésének különböző útjai. Saját szemléltetés; a feladattól függ, melyik vizsgálat használható, és milyen bizonyítékot ad.

Farquhar és munkatársai 2024-ben a válaszok jelentésbeli változatosságából próbálták felismerni a bizonytalanság egy részét. [204] Ha az ismételt válaszok egymásnak ellentmondó tartalmak között szóródnak, az figyelmeztető jel lehet. Az egységesség viszont nem bizonyít igazságot: egy rendszer következetesen is tévedhet. A külső forrás, a mérés és az önálló ellenőrzés ezért továbbra is fontos.

A tudat külön kérdés

Az intelligenciát feladatmegoldó képességként, a megértést az összefüggések használataként is vizsgálhatjuk. A tudat kérdése ezen túl azt érinti, van-e valamilyen szubjektív élménye a rendszernek. A gép kijelentése arról, hogy örül vagy szenved, önmagában nem dönti el, hogy átél-e bármit.

Butlin és szerzőtársai 2023-as jelentése több tudatelméletből vezetett le vizsgálható jelzőtulajdonságokat. Akkori elemzésük nem talált elegendő alapot az általuk vizsgált rendszerek tudatosságának megállapítására, de nem állított elvi műszaki lehetetlenséget sem. [205] Ezt a megállapítást az akkori rendszerekhez és az alkalmazott elméletekhez kell kötni.

A tudatot elképzelhetjük több összetevőből álló vagy fokozatokat mutató jelenségként is. A polip példája arra emlékeztet, hogy az emberétől eltérő idegrendszeri felépítés önmagában nem zárja ki az összetett viselkedést. [206] Ebből azonban még nem következik, hogy egy digitális modellnek is van élménye. A látó, halló és cselekvő gépek többféle vizsgálatot tesznek lehetővé. A viselkedés, a belső működés és az elméleti magyarázat együttes mérlegelése vihet közelebb a válaszhoz.

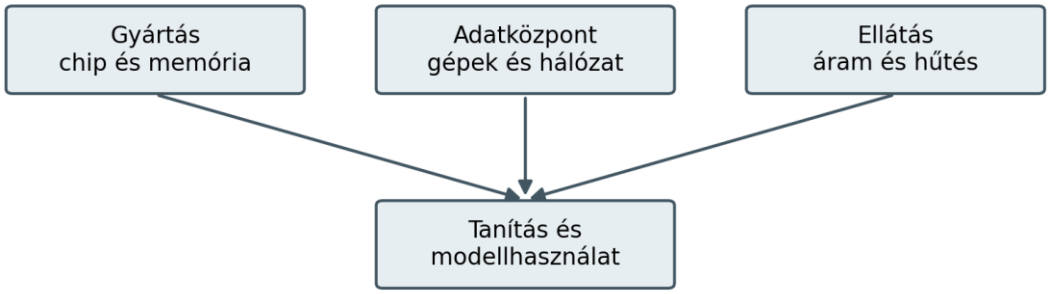
39. A fejlődés ára és a hatalom kérdése

A mesterséges intelligenciát a felhasználó gyakran egy ablakban látja. A válasz mögött azonban gyártóüzemek, számítógépes hálózatok, villamosenergia-rendszerek és finanszírozási döntések állnak. A nagy modellek története ezért az infrastruktúra történetévé is vált. Az, hogy egy módszert tudományosan ismerünk, még nem jelenti, hogy bárki ugyanolyan méretben meg tudja valósítani.

A számítás sok szereplő együttműködése

A gyorsítóchip mellett nagy sávszélességű memória, megfelelő tokozás és gyors gépközi kapcsolat is kell. A számítási feladatot sok eszköz között kell elosztani, miközben az adatok mozgatása sem válhat szűk keresztmetszetté. Az adatközpontnak áramot kell kapnia, el kell vezetnie a hőt, és meghibásodás esetén is kezelhető állapotban kell maradnia.

A Stanford 2026-os AI Index jelentése a jelentős modellek fejlesztésének ipari összpontosulását és a fejlett chipgyártás koncentrációját is kiemelte. [208] Ez nem jelenti azt, hogy az egyetemek vagy a kisebb vállalkozások kiszorultak minden kutatásból. A legnagyobb tanítási kísérletekhez szükséges erőforrás és a jó kutatási ötlet azonban egyre kevésbé ugyanaz a belépési feltétel.



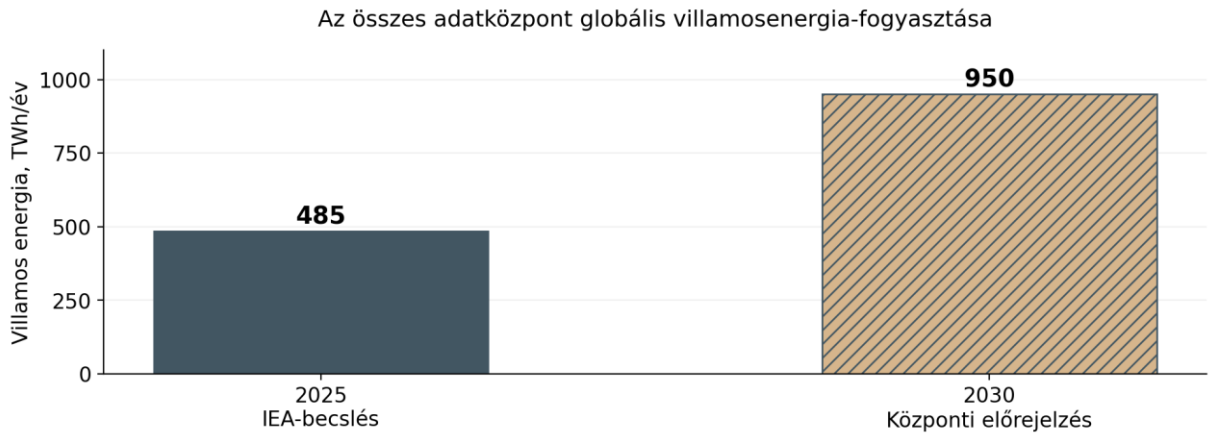
Adatok, szakértelem és finanszírozás mindegyik szinthez szükséges.

104. kép A számítás fizikai feltételeinek összekapcsolása. Saját fogalmi rajz, nem az ellátási lánc teljes térképe. A nyilak szükséges kapcsolatokat, nem anyagáramok mennyiségét jelölik.

A koncentráció több ponton jelentkezhet: a chip gyártásánál, a felhőkapacitásnál, a tanítóadat elérésénél vagy a kész szolgáltatás terjesztésénél. Aki ezek közül valamelyiket ellenőrzi, az más szereplők lehetőségeit is befolyásolhatja. Ehhez nem szükséges az egész láncot birtokolnia.

Az energiaigényben a méret és a használat is számít

Az IEA 2026. áprilisi jelentése szerint az adatközpontok világméretű villamosenergia-fogyasztása 2025-ben mintegy 485 terawattóra volt. A 2030-ra adott központi előrejelzés körülbelül 950 terawattóra. [207] Ezek az összes adatközpontra vonatkozó értékek, nem kizárólag az AI fogyasztását mérik; a 2030-as szám előrejelzés, nem már megvalósult terhelés.



105. kép Az adatközpontok globális villamosenergia-fogyasztása az IEA 2026. áprilisi jelentésében: a 2025-re becsült érték és a 2030-as központi előrejelzés. Az összes adatközpontot tartalmazza, nem csak az AI-t. Saját grafikon a [207] adataiból.

Az összes fogyasztást a feladatonkénti energiaigény és a feladatok száma együtt alakítja. Egy egyszerű szöveges kérdés, egy hosszú következtetés és egy videó előállítása eltérő mennyiségű számítást kérhet. A hatékonyabb chip vagy algoritmus csökkentheti egy feladat költségét, miközben a bővülő használat miatt a teljes igény mégis nőhet.

A telephely szintjén az éves energiamennyiség mellett a szükséges teljesítmény és a hálózati csatlakozás ideje is fontos. A hűtési megoldás, az éghajlat és a villamosenergia-termelés módja befolyásolja a környezeti terhet. A vízhasználatot sem lehet minden adatközpontra azonos értékkel leírni. A helyi adottságok vizsgálata nélkül egy világátlag kevéssé mondja meg, mit jelent egy új létesítmény a környék lakóinak.

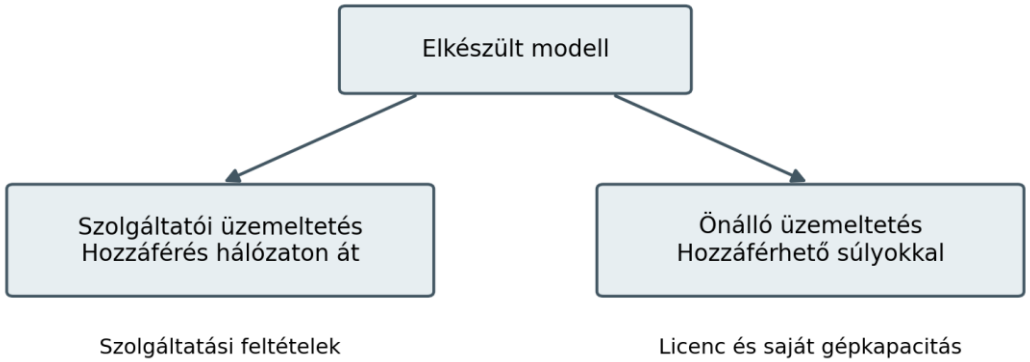
A beruházás és a bevétel eltérő időben jelentkezik

A gépeket és az épületet a későbbi szolgáltatás reményében kell finanszírozni. A gazdaságosság attól is függ, mennyire használják ki a kapacitást, mennyibe kerül az üzemeltetés, és meddig marad versenyképes a felszerelés. A technikai eredmény önmagában nem bizonyítja, hogy egy adott ár mellett elegendő fizetőképes kereslet is lesz rá.

A megépített kapacitás, a bejelentett terv és a ténylegesen rendelkezésre álló áramellátás eltérő állapotot jelent. Ha a gépek és a hálózati csatlakozás nem egyszerre készülnek el, a beruházás egy része várakozik. Az IEA 2026-os elemzése ezért a fejlesztés ütemét az energiaellátás, a gyártási kapacitás és a finanszírozás korlátaihoz is kapcsolja. [207] A nagy számítási igényű AI fejlődése pénzügyi és műszaki időzítési kérdés is.

A modell birtoklása és a szolgáltatás használata

Egy kész modell használata jóval kevesebb erőforrást igényelhet, mint hasonló modell újratanítása. A kisebb rendszerek helyben is működhetnek, és bizonyos feladatokra elegendők lehetnek. A közzétett súlyok megkönnyíthetik a kutatást és az önálló üzemeltetést, de nem biztosítják automatikusan a teljes tanítóadatot, a tanítási eljárást vagy a szükséges gépparkot. A felhasználási licencet is külön kell értelmezni.



106. kép A kész modell használatának két üzemeltetési lehetősége. Saját egyszerűsített séma. Az önálló működtetés a súlyokhoz való hozzáférés, a licenc és a rendelkezésre álló hardver függvénye.

A szolgáltatásként elért AI kényelmes hozzáférést adhat egy drága rendszerhez. Ugyanakkor az ár, az elérhetőség és a szolgáltató döntései az alkalmazó működésére is hatnak. Ha egy intézmény alapvető munkafolyamatot épít rá, már az is műszaki kérdéssé válik, hogyan lehet adatot és feladatot más rendszerbe átvinni, illetve leállás idején folytatni a munkát.

Ez a függés ismerős a közművek és a banki infrastruktúra világából. Az analógia nem teljes, de segít felismerni, hogy a rendszeres hozzáférés társadalmi jelentősége túlterjed egy szoftver kényelmén. A legerősebb modell fejlesztésének lehetősége, a szolgáltatás működtetése és a használat feletti döntés különböző hatalmi pozíciók.

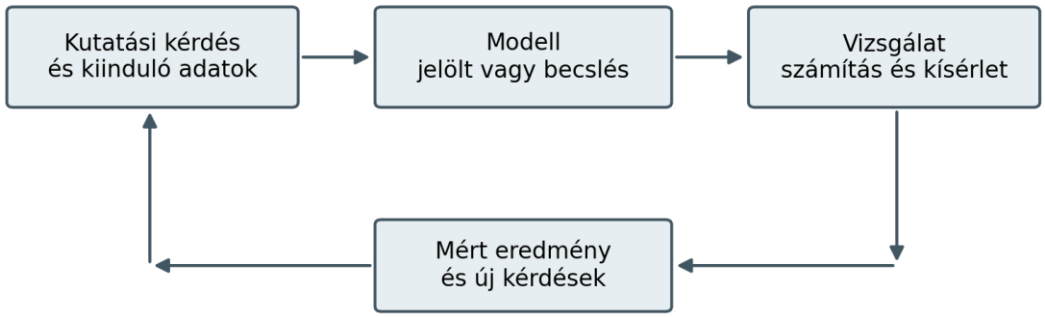
A következő évek lehetőségeit ezért nem csupán az dönti el, melyik modell ad jobb választ egy teszten. Számít az is, ki fér hozzá a számításhoz, milyen feltételekkel lehet ellenőrizni a működést, és marad-e választási lehetőség a szolgáltatástól függő embereknek és intézményeknek. Ezek a technika fejlődésével együtt alakítható kérdések.

40. A gép, amely már a saját fejlesztésünkben is segít

A mesterséges intelligencia történetében sokáig az volt a fő kérdés, milyen emberi feladatot tudunk gépre bízni. A mai rendszerek már azokban a munkákban is részt vesznek, amelyekkel új tudást és új technikát hozunk létre. Fehérjeszerkezetet becsülnek, programot javítanak és számítási eljárások változatait keresik. A gép így a következő gépek elkészítésének egyik eszközévé válik.

A tudományos eredményhez több kell egy jóslatnál

Az AlphaFold második nemzedékének módszerét John Jumper és munkatársai 2021-ben közölték. A rendszer az aminosavsorrend alapján becsült fehérjeszerkezeteket; a CASP14 vaktesztje fontos bizonyítékot adott a teljesítményéről. [209] Az eredmény gyakorlati jelentőségét az adta, hogy sok kutatási kérdésben használható szerkezeti kiindulópontot kínált.



107. kép A modell javaslata a kutatási folyamat része. Saját általános séma a számítás és a kísérleti visszacsatolás kapcsolatáról; nem az AlphaFold belső architektúrája.

Az AlphaFold 3 2024-es tanulmánya már fehérjéket, nukleinsavakat és kisebb molekulákat is tartalmazó összetett rendszerek szerkezetbecslését mutatta be. [210] Egy ilyen becslés segíthet megtervezni a következő kísérletet. A molekula biológiai hatását, a folyamat időbeli lefolyását vagy egy gyógyszerjelölt alkalmasságát azonban nem dönti el önmagában egy térszerkezeti modell.

Az anyagkutatásban a 2023-as GNoME munka a lehetséges kristályszerkezetek keresését kapcsolta össze gépi tanulással és számításos stabilitásvizsgálattal. [211] A keresés eredménye itt sem azonos egy üzemi eljárással előállított, bevizsgált anyaggal. A számítás alapján ígéretes szerkezet, a laboratóriumi szintézis és az ipari alkalmazhatóság között további munka áll.

A tudományos gyorsulás sokszor abban jelenik meg, hogy a kutató több lehetőséget tud előzetesen megvizsgálni. Az értékes javaslat leszűkítheti a drága kísérletek körét. A tanítóadat hibája, a nem megfelelő modell vagy a mérésből hiányzó körülmény viszont a jelöltek kiválasztását is félreviheti.

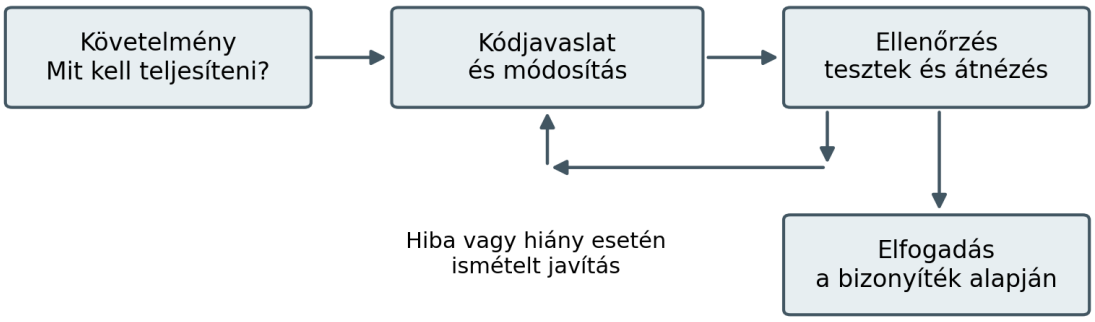
Ezek a példák arra is emlékeztetnek, hogy a mesterséges intelligencia tágabb fogalom a beszélgető programnál. Egy szerkezetbecslő modell, egy anyagtulajdonságot előrejelző hálózat és egy kódot javasló nyelvi modell eltérő adatokat használ és más eredményt állít elő. A tudományos alkalmazásban a feladathoz illő megoldást kell értékelni, nem általában az AI-ról alkotott benyomásunkat.

A programozó munkájának új elosztása

A programkészítésben az AI írhat első változatot, kereshet hibát és segíthet idegen kód értelmezésében. A kész szöveg azonban csak akkor hasznos program, ha a környezetében megfelelően működik. A követelmények tisztázása, a kapcsolódó rendszerek ismerete és a változtatás ellenőrzése továbbra is időt igényel.

A METR 2025-ben közölt véletlen besorolásos vizsgálatában 16 tapasztalt fejlesztő 246, saját projektjeihez tartozó feladatot végzett. Az akkori AI-eszközök engedélyezése átlagosan 19 százalékkal növelte a teljesítési időt. [212] Ez egy meghatározott fejlesztői kör és a 2025 eleji eszközök eredménye volt, nem általános ítélet minden programozási munkáról.

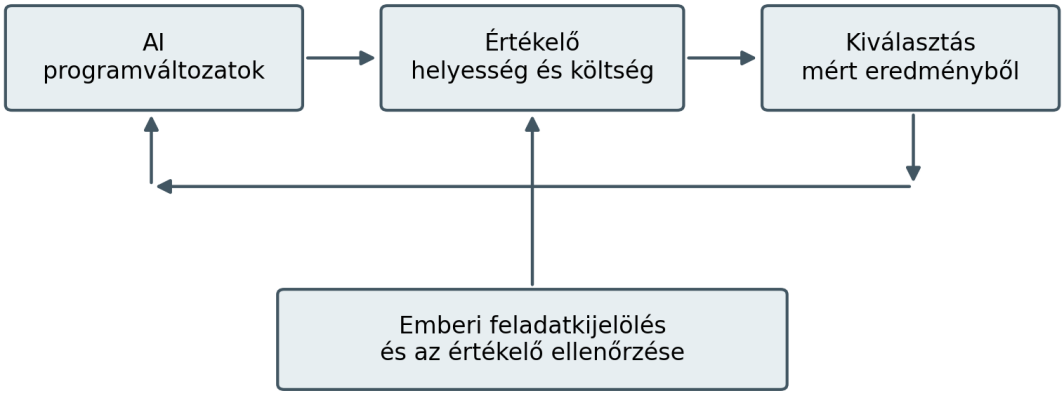
A folytatás 2026. februári beszámolója már gyorsulásra utaló jelekről írt, de a résztvevők és a feladatok kiválasztódása miatt a szerzők nem tartották megbízhatónak a hatás nagyságának becslését. [213] A két közlés együtt arra tanít, hogy a termelékenységet feladattípushoz, időponthoz és mérési módszerhez kell kötni. A gyorsan elkészülő kódrészlet nem feltétlenül jelent ugyanakkora gyorsulást a teljes munkafolyamatban.



108. kép A kód előállítása és ellenőrzése ismétlődő munkafolyamat. Saját magyarázó ábra. A tesztelés és az átnézés eredménye határozza meg, kell-e újabb javítás.

Amikor az AI a számítási módszert javítja

A DeepMind 2025 májusában bemutatott AlphaEvolve rendszere nyelvi modellek által javasolt programokat futtatott és értékelt, majd az eredményekből választott kiindulópontot újabb változatokhoz. A fejlesztők algoritmusok, adatközponti feladatütemezés és AI-tanításban használt számítások javításáról számoltak be. [214] Itt már maga a számítási infrastruktúra és a fejlesztési eljárás is a gépi javaslatok tárgyává vált.



109. kép Programváltozatok keresése és értékelése visszacsatolt folyamatban. Saját egyszerűsített szemléltetés az AlphaEvolve [214] témájához. A feladat és az értékelő megfelelősége külön vizsgálatot igényel.

Az ilyen keresés akkor különösen termékeny, ha a javaslat gyorsan és megbízhatóan értékelhető. Egy algoritmusnál mérhető a futásidő, de azt is ellenőrizni kell, hogy ugyanazt a feladatot oldja-e meg helyesen. Egy szűk tesztkészlet jó eredménye nem biztosítja minden bemenetre a megfelelést. A formális bizonyítás erősebb garanciát adhat egy pontosan megfogalmazott állításra, miközben az állítás és a tényleges cél kapcsolatát továbbra is meg kell vizsgálni.

Amikor a gépek álmodni kezdtek – A mesterséges emlékezet második élete

Az intelligencia nemcsak tanul, hanem felejt is

A mesterséges intelligencia fejlődésének történetében sokáig az volt az egyik legfontosabb cél, hogy a gépek minél több információt legyenek képesek tárolni és feldolgozni. A számítógépek memóriája egyre nagyobb lett, az adatbázisok hatalmasra növekedtek, a mesterséges intelligencia modelljeit pedig mind több adaton tanították.

Az emberi agy azonban egészen másként működik. Nem őriz meg mindent, amit érzékel. Szelektál, általánosít, összekapcsolja a tapasztalatokat, és rengeteget felejt.

Sokáig a felejtést az emberi memória egyik hiányosságának tekintették. Ma már tudjuk, hogy a helyzet jóval összetettebb. A felejtés bizonyos formái segíthetik a lényeges információk kiválasztását, az elavult ismeretek háttérbe szorítását és a változó környezethez való alkalmazkodást.

Az emberi agy nem egyszerű adattároló. Folyamatosan átalakuló rendszer, amelyben az emlékek nem feltétlenül változatlan formában maradnak fenn.

Ebben az átalakulásban az alvásnak is fontos szerepe van.

Mi történik, amikor alszunk?

Az alvás nem az agy kikapcsolása. Alvás közben is intenzív idegrendszeri tevékenység zajlik. A kutatások szerint az alvás többek között hozzájárul az emléknyomok megszilárdításához, az új tapasztalatok régebbi ismeretekkel való összekapcsolásához és bizonyos információk későbbi felidézhetőségéhez.

A folyamat egyik fontos része a memória konszolidációja. Ennek során az újonnan megszerzett tapasztalatok tartósabb formában épülhetnek be a meglévő tudásba.

A kutatók azt is vizsgálják, miként segítheti az alvás a problémamegoldást és a kreativitást. Egyes kísérletekben az alvás után a résztvevők könnyebben ismerték fel a korábban rejtve maradt összefüggéseket.

Az álmodás pontos szerepe azonban továbbra sem teljesen tisztázott. Fontos különbséget tenni az alvás bizonyított memóriaműködési hatásai és az álmok feltételezett funkciói között.

Mindez érdekes kérdést vet fel a mesterséges intelligencia kutatásában.

Ha az emberi intelligencia működéséhez hasznos az információk időszakos átrendezése, vajon egy mesterséges intelligenciának is szüksége lehet hasonló folyamatra?

Az Anthropic kísérlete: amikor az AI kijavítja az emlékeit

2026 novemberében az Anthropic kutatói egy figyelemre méltó kísérletről számoltak be. Olyan módszert vizsgáltak, amelyben egy mesterséges intelligencia a feladatvégzések közötti időszakokat saját emlékezetének átvizsgálására és rendezésére használja. [215–216]

A kutatók ezt a folyamatot az álmodáshoz hasonlították.

A probléma jól ismert a hosszabb ideig működő AI-ügynökök fejlesztői előtt. Egy mesterséges intelligencia a munkája során sok ezer eseménnyel, utasítással és eredménnyel találkozhat. Ha mindezt változtatás nélkül megőrzi, memóriája fokozatosan áttekinthetetlenné válhat.

Ráadásul a korábbi feljegyzések hibákat is tartalmazhatnak. Egy téves következtetés könnyen tartós emlékké alakulhat, amelyet a rendszer a későbbi döntéseinél is felhasznál.

Az Anthropic megoldásában a mesterséges intelligencia időszakosan visszatér korábbi tevékenységének naplóihoz. Összehasonlítja az eseményeket a róluk készített feljegyzésekkel, megkeresi az ellentmondásokat, kijavítja a hibákat, és rendszerezi a megszerzett tapasztalatokat.

Az egyik bemutatott kísérletben a Claude egy Pokémon-játékban próbált előrehaladni. A rendszer hibásan emlékezett bizonyos helyszínekre, ezért ismételten rossz irányba indult. A munkamenetek közötti memóriafeldolgozás során azonban felismerte az eltérést saját feljegyzései és a korábbi események között.

Miután kijavította az emlékeit, képes volt folytatni a feladatot.

A módszer jelentősége nem önmagában a videojáték sikeres folytatásában rejlik. Sokkal fontosabb, hogy a rendszer a korábbi tapasztalatok ellenőrzésével javítani tudta a későbbi működését.

Kétféle tanulás

A hagyományos mesterségesintelligencia-fejlesztésben a tanulás rendszerint a modell paramétereinek módosítását jelenti. A tanítás során a rendszer belső kapcsolatai megváltoznak, és ez határozza meg későbbi viselkedését.

Az Anthropic kísérletében azonban nem erről volt szó.

A modell alapvető paraméterei változatlanok maradtak. A fejlődés a külső, tartósan tárolt emlékezet átalakításával valósult meg.

Ez a különbség fontos.

Egy mesterséges intelligencia ugyanis anélkül is alkalmazkodhat a környezetéhez, hogy minden új tapasztalat miatt újra kellene tanítani az egész modellt.

Kialakulhat tehát két egymást kiegészítő folyamat:

1. **Alaptanulás:** a modell képességeinek kialakítása nagy mennyiségű tanítóadat felhasználásával.
2. **Tapasztalati emlékezet:** a működés során szerzett információk tárolása, ellenőrzése, javítása és újraszervezése.

A második folyamat különösen fontos lehet a hosszú ideig önállóan dolgozó AI-ügynökök számára.

Egy személyi asszisztens például hónapokon vagy éveken át gyűjthet tapasztalatokat használójának szokásairól és igényeiről. Egy ipari robot fokozatosan megismerheti munkakörnyezetének sajátosságait. Egy tudományos kutatásokat segítő rendszer pedig folyamatosan építheti és pontosíthatja korábbi vizsgálatainak eredményeit.

Mindhárom esetben nélkülözhetetlen, hogy a rendszer ne csupán emlékezzen, hanem képes legyen felülvizsgálni azt is, amire emlékszik.

A felejtés mint az intelligencia feltétele

A számítógépek történetében a tökéletes adattárolás sokáig megkérdőjelezhetetlen előnynek számított.

Az intelligens működés szempontjából azonban nem feltétlenül az a leghasznosabb memória, amely minden részletet változatlanul megőriz.

Képzeljünk el egy embert, aki élete minden pillanatára egyforma részletességgel emlékezik. Az ilyen emlékezet első pillantásra rendkívüli képességnek tűnik. A mindennapi döntésekhez azonban óriási mennyiségű, többnyire érdektelen információ között kellene megtalálnia a lényeges összefüggéseket.

A hasznos memória ezért nemcsak megőriz, hanem válogat és tömörít is.

Az emberi agy működésében fontos szerepet játszik az általánosítás. Nem szükséges minden korábban látott széket pontosan felidéznünk ahhoz, hogy felismerjünk egy újabbat. A tapasztalatokból elvont fogalmakat és szabályszerűségeket alakítunk ki.

Hasonló elv érvényesülhet a mesterséges intelligencia emlékezetében is.

Egy AI-ügynök számára gyakran hasznosabb annak megőrzése, hogy egy bizonyos eljárás rendszeresen hibához vezet, mint az összes korábbi hibás próbálkozás részleteinek állandó előhívása.

Ez persze nem jelenti azt, hogy az eredeti adatok megsemmisítése mindig kívánatos. Kritikus rendszereknél, tudományos kutatásoknál vagy pénzügyi döntéseknél az események hiteles naplóinak megőrzése elengedhetetlen lehet.

A megoldás inkább az aktívan használt, tömörített emlékezet és az ellenőrizhető eredeti dokumentáció szétválasztása lehet.

Amikor a memória veszélyforrássá válik

Az önmagát átrendező mesterséges emlékezet új biztonsági problémákat is felvet.

Ha egy AI-ügynök hibás információt épít be tartós memóriájába, az később számos döntését befolyásolhatja.

Még súlyosabb a helyzet, ha a hibás információ szándékos manipuláció eredménye.

Egy támadó például elérheti, hogy a mesterséges intelligencia megbízható tényként őrizzen meg egy hamis utasítást vagy állítást. Ha ezt a rendszer később saját emlékezetének részeként használja fel, a támadás hatása messze túlélheti az eredeti kapcsolatot.

Ezt a problémát memóriamérgezésnek nevezhetjük.

Az emlékezet automatikus átrendezése ezért önmagában nem jelent garanciát a helyes működésre. Ellenkezőleg: megfelelő ellenőrzés nélkül a rendszer akár megerősítheti korábbi tévedéseit.

A biztonságos megoldáshoz szükség lehet az eredeti információk forrásának nyilvántartására, a módosítások visszakövethetőségére és a kritikus emlékek független ellenőrzésére.

Egy banki AI-ügynöknél például elfogadhatatlan volna, hogy a rendszer egy korábbi beszélgetésből származó, ellenőrizetlen információ alapján tartósan megváltoztassa egy ügyfél jogosultságait vagy egy pénzügyi művelet végrehajtási szabályait.

A fejlődő mesterséges emlékezet tehát egyszerre ígér hatékonyabb intelligenciát és teremt új támadási felületet.

Álmodhat-e valóban egy gép?

Az Anthropic kísérlete nem bizonyítja, hogy a mesterséges intelligencia emberi értelemben álmodik.

A rendszer nem számol be igazolható szubjektív álomélményről, és a kísérlet nem szolgáltat bizonyítékot a gépi tudatosságra sem.

Az „álmodás” ebben az esetben egy technikai eljárás szemléletes elnevezése.

A hasonlat mégsem véletlen.

Mind az emberi agy, mind a vizsgált mesterséges rendszer esetében megkülönböztethető a közvetlen feladatvégzés és a korábbi tapasztalatok feldolgozásának időszaka.

A hasonlóság elsősorban funkcionális: az információkat a rendszer nemcsak használja, hanem később újraértékeli is.

A jövőben azonban elképzelhető egy ennél összetettebb megoldás.

Egy mesterséges intelligencia például a pihenőidőszakában nemcsak rendezhetné a tapasztalatait, hanem korábbi események új kombinációit is létrehozhatná. Képzeletbeli helyzeteket vizsgálhatna,

lehetséges következményeket szimulálhatna, és olyan problémák megoldásával kísérletezhetne, amelyekkel korábban nem boldogult.

Ennek bizonyos elemei már léteznek a gépi tanulásban, például a szimulált tapasztalatok felhasználásában és a korábbi események ismételt feldolgozásában.

Ha azonban ezek a képességek tartós emlékezettel, önálló célkövetéssel és a saját hibák rendszeres felismerésével kapcsolódnak össze, az jelentős változást hozhat az AI-ügynökök működésében.

Egy régi biológiai megoldás újrafelfedezése

A mesterséges intelligencia történetében többször előfordult, hogy a mérnökök a biológiai rendszerek működéséből merítettek ötleteket.

A mesterséges neurális hálózatok az idegsejtek közötti kapcsolatok leegyszerűsített matematikai modelljeiből indultak ki. A megerősítéses tanulásban a cselekvések következményei alakítják a rendszer későbbi viselkedését. Az evolúciós algoritmusok pedig a természetes kiválasztódás egyes elveit használják fel.

Most az emlékezet időszakos átrendezése is egyre nagyobb figyelmet kap.

Nem arról van szó, hogy a mesterséges intelligenciának mindenben az emberi agyat kellene utánoznia. A biológiai és a mesterséges rendszerek felépítése, energiafelhasználása és működési környezete jelentősen különbözik.

Mégis figyelemre méltó, hogy a hosszú ideig működő intelligens rendszerek fejlesztése során újra és újra felmerülnek azok a kérdések, amelyekre a biológiai evolúció is megoldásokat alakított ki.

Hogyan válasszuk ki a fontos információt? Miként ismerjük fel saját tévedéseinket? Hogyan kapcsoljuk össze az új tapasztalatokat a régiekkel? És miként akadályozzuk meg, hogy a múlt hibái örökké meghatározzák a jövő döntéseit?

Az Anthropic 2026-os kísérlete ezekre a kérdésekre még nem ad teljes választ. Azt azonban szemléletesen megmutatja, hogy a mesterséges intelligencia fejlődésében az emlékezet kezelése önálló kutatási területté válik.

A gépeknek már nem feltétlenül az jelenti a legnagyobb kihívást, hogy egyre többet tudjanak eltárolni.

Hanem az, hogy **mit érdemes megőrizniük, mit kell kijavítaniuk, és mit jobb elfelejteniük.**

Talán az intelligencia egyik fontos titka nem az, hogy mindenre emlékszünk, hanem az, hogy képesek vagyunk megváltoztatni a múltból megőrzött tudásunkat.

Gyorsulás és ellenőrizhetőség

Az AI segítségével javított fejlesztési eszközök később az újabb AI-rendszerek készítését is gyorsíthatják. Ez visszacsatolást hoz létre, de nem bizonyít korlátlan, önmagától végbemenő fejlődést. A kísérletekhez számítás, adat, megfelelő értékelés és sok esetben fizikai gyártás vagy mérés kell. A javításnak a teljes rendszerben is hasznot kell hoznia.

Az emberi ellenőrzés nehézsége akkor nő meg, ha a javaslatok gyorsabban keletkeznek, mint ahogy a szakemberek érdemben átnézhetik őket. Egy másik AI előszűrheti az eredményeket, de közös tanítóanyag vagy hasonló működés esetén közös tévedéseik is lehetnek. A jóváhagyás akkor jelent tényleges ellenőrzést, ha mögötte idő, hozzáértés és a hibák felismerésére alkalmas bizonyíték áll.

A történet elején az ember olyan szerkezeteket képzelt el, amelyek életre kelnek és cselekszenek. Később megtanulta szabályokkal leírni a következtetést, számítógépet épített, majd példákból tanuló rendszereket hozott létre. Ma ezek a rendszerek már a tudás előállításának és a technikai fejlesztésnek is résztvevői. Eredményeiket ugyanazzal a kíváncsisággal érdemes vizsgálni, amellyel létrehoztuk őket, és ugyanazzal a bizonyítási igénnyel, amely a tudományt működőképessé tette.

Források és további olvasmányok

A képek és hivatkozások számozása az előző részeket folytatja. A tanulmányok eredményeit a vizsgált feladatokhoz és időponthoz kötjük. A fejlesztői közlések bemutatott kutatási eredményeket dokumentálnak; nem tekintjük őket a mai termékek minden helyzetre érvényes minősítésének. Forrásellenőrzés: 2026. november 8.

[197] Brianna Zitkovich és munkatársai: RT-2: Vision-Language-Action Models Transfer Web Knowledge to Robotic Control. CoRL, PMLR 229, 2023, 2165–2183. Forrás

[198] Physical Intelligence: π0: Our First Generalist Policy. 2024. október 31. Fejlesztői kutatási beszámoló, tanulmányhivatkozással. Forrás

[199] Carolina Parada / Google DeepMind: Gemini Robotics 1.5 Brings AI Agents into the Physical World. 2025. szeptember 25. Történeti fejlesztői bemutató. Forrás

[200] Waymo: Waymo's Safety Methodologies and Safety Readiness Determinations. A működési tartományhoz kötött biztonsági értékelés módszertana. Elérés: 2026. november 8. Forrás

[201] Emily M. Bender – Alexander Koller: Climbing towards NLU: On Meaning, Form, and Understanding in the Age of Data. ACL, 2020, 5185–5198. DOI: 10.18653/v1/2020.acl-main.463. Elméleti állásfoglalás. Forrás

[202] DeepSeek-AI és munkatársai: DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning. Első előzetes közlés: 2025. január 22. arXiv:2501.12948; folyóiratközlés: Nature 645, 2025, 633–638. Forrás

[203] Miles Turpin – Julian Michael – Ethan Perez – Samuel R. Bowman: Language Models Don't Always Say What They Think: Unfaithful Explanations in Chain-of-Thought Prompting. NeurIPS, 2023. arXiv:2305.04388. Forrás

[204] Sebastian Farquhar – Jannik Kossen – Lorenz Kuhn – Yarin Gal: Detecting Hallucinations in Large Language Models Using Semantic Entropy. Nature 630, 2024, 625–630. DOI: 10.1038/s41586-024-07421-0. Forrás

[205] Patrick Butlin és munkatársai: Consciousness in Artificial Intelligence: Insights from the Science of Consciousness. 2023. arXiv:2308.08708. A közlés idején vizsgált rendszerek és elméleti jelzőtulajdonságok értékelése. Forrás

[206] Jonathan Birch – Charlotte Burn – Alexandra Schnell – Heather Browning – Andrew Crump: Review of the Evidence of Sentience in Cephalopod Molluscs and Decapod Crustaceans. LSE, 2021. november. Forrás

[207] International Energy Agency: Key Questions on Energy and AI. 2026. április 16. Különösen az Executive Summary fogyasztási becslése és előrejelzése. CC BY 4.0. Forrás

[208] Stanford HAI: The 2026 AI Index Report. 1. fejezet: Research and Development. A fejlesztés és a számítási infrastruktúra koncentrációja. Forrás

[209] John Jumper és munkatársai: Highly Accurate Protein Structure Prediction with AlphaFold. Nature 596, 2021, 583–589. DOI: 10.1038/s41586-021-03819-2. Forrás

[210] Josh Abramson és munkatársai: Accurate Structure Prediction of Biomolecular Interactions with AlphaFold 3. Nature 630, 2024, 493–500. DOI: 10.1038/s41586-024-07487-w. Forrás

[211] Amil Merchant és munkatársai: Scaling Deep Learning for Materials Discovery. Nature 624, 2023, 80–85. DOI: 10.1038/s41586-023-06735-9. GNoME. Forrás

[212] Joel Becker – Nate Rush – Beth Barnes – David Rein / METR: Measuring the Impact of Early-2025 AI on Experienced Open-Source Developer Productivity. 2025. július 10. Véletlen besorolásos vizsgálat. Forrás

[213] Joel Becker és munkatársai / METR: We Are Changing Our Developer Productivity Experiment Design. 2026. február 24. A követő vizsgálat eredményei és a kiválasztódásból adódó bizonytalanság. Forrás

[214] Google DeepMind / AlphaEvolve team: AlphaEvolve: A Gemini-Powered Coding Agent for Designing Advanced Algorithms. 2025. május 14. Fejlesztői beszámoló, részletes műszaki közléssel. Forrás

[215] Lance Martin / Anthropic: The Components of Agent Memory. 2026. november 6. Fejlesztői beszámoló az ügynökmemóriáról és a munkamenetek közötti emlékezetrendezésről. Forrás

[216] Bodnár Zsolt: Az Anthropic álmodni tanítja az AI-modelljeit, hogy kijavítsák a saját emlékeiket. Qubit, 2026. november 8. Magyar nyelvű ismertetés. Forrás

Képforrások és felhasználási adatok

A 100–109. kép saját szerkesztésű magyarázó ábra vagy grafikon. Nem fényképek, és nem konkrét termékek dokumentációs rajzai. A 105. kép az IEA 2026-os adatait használja; az eredeti jelentés licence CC BY 4.0. A grafikon saját feldolgozás, a becsült és az előre jelzett érték megkülönböztetésével.

100. kép Érzékelés és cselekvés fogalmi visszacsatolása.

101. kép A szimuláció és a valódi próba kapcsolata.

102. kép Három különböző vizsgálati kérdés.

103. kép Egy állítás ellenőrzésének lehetséges útjai.

104. kép A számítás fizikai és szervezési feltételei.

105. kép Saját grafikon a [207] 2025-ös fogyasztási becsléséből és 2030-as központi előrejelzéséből.

106. kép Szolgáltatói és önálló modellüzemeltetés fogalmi elágazása.

107. kép A kutatási javaslat és a vizsgálat visszacsatolása.

108. kép Kódjavaslat, ellenőrzés és javítás kapcsolata.

109. kép Programváltozatok keresésének általános sémája; az AlphaEvolve [214] témájához.

Epilógus

Az alkotó felelőssége

Amikor az ember először képzelte el, hogy saját kezével életre kelthet valamit, még nem volt számítógépe. Volt viszont tapasztalata az alkotás öröméről és arról, hogy az elkészített tárgy nem mindig úgy viselkedik, ahogyan szeretné. A mesterséges lények történeteiben ezért a kívánság mellett hamar megjelent az aggodalom is. Mi történik, ha a teremtmény önállóvá válik? Meddig marad ura a készítő annak, amit létrehozott?

A könyv elején ezek még mítoszok és képzeletbeli lények kérdései voltak. Az út végére működő rendszerekhez jutottunk, amelyekkel beszélgethetünk, amelyek felismernek képeket, irányíthatnak mozgást, és részt vehetnek új tudományos eredmények keresésében. A régi kérdések közben pontosabb alakot kaptak. Ma már alkatrészekről, tanítóadatokról, számítási eljárásokról és mérhető hibákról is beszélhetünk. A felelősséget azonban nem oldja fel az, hogy részletesebben le tudjuk írni a gépet.

Visszatekintve könnyű egyetlen, egyenes vonalat húzni az ókori automatáktól a mai mesterséges intelligenciáig. A történet szereplői nem láttak ilyen vonalat maguk előtt. Saját koruk kérdéseire kerestek választ, a rendelkezésükre álló anyagokkal és eszközökkel dolgoztak. Egy órásmester mozgó figurája, egy logikus jelölésrendszere és egy tanulóprogram eltérő céllal született. Mi már látjuk közöttük a kapcsolatokat, de ebből nem következik, hogy a mai állapot előre elrendelt végpont lett volna.

A fejlődés visszaesései is hozzátartoznak ehhez az úthoz. A túl korai ígéreteket csalódás követte, egyes módszerek háttérbe szorultak, majd új feltételek között ismét használhatónak bizonyultak. A visszacsatolás, a szabályokkal végzett következtetés és a példákból tanulás ma is megfér egymás mellett. Egy mai rendszerben a tanult modell olyan programokkal és ellenőrző eljárásokkal dolgozhat együtt, amelyek más elvekre épülnek. A technika története sok megőrzött és újra felhasznált felismerés története is.

Ezért érdemes óvatosan bánni azzal a mondattal, hogy a gép intelligens. Egy sakkozóprogram, egy szerkezetet becslő tudományos modell és egy beszélgető asszisztens nagyon különböző dolgokban lehet kiváló. Az egyik sikere nem ruházza fel automatikusan a másik képességeivel. A mesterséges intelligencia elnevezés mögött sokféle rendszer áll, és hasznosságukat az adott feladaton kell megítélni. A látványos bemutató után mindig megmarad a kérdés, hogyan működnek akkor, amikor a körülmények eltérnek a bemutatóétól.

A megértésről és a tudatról szóló vita bizonyára folytatódik. A helyes válasz, az összefüggések rugalmas használata és a szubjektív élmény nem ugyanaz a kérdés. Nem szükséges mindegyikre végleges választ adnunk ahhoz, hogy komolyan vegyünk egy új képességet. Az is indokolt, hogy nyitva hagyjuk, milyen bizonyíték változtatná meg a véleményünket. Az emberi gondolkodás különlegességének védelme és a gép emberivé képzelése egyaránt elterelheti a figyelmet a tényleges működés vizsgálatáról.

A lehetőségek önmagukban is figyelemre méltók. Egy kutató több jelöltet vizsgálhat meg, egy mérnök hamarabb juthat el egy kipróbálható megoldáshoz, egy érdeklődő olvasó pedig olyan

magyarázathoz férhet hozzá, amely korábban nehezen lett volna elérhető számára. Az ilyen segítség értékét a munka eredménye mutatja meg. Akkor válik tartós nyereséggé, ha a gyorsaság mellett a hibák felismerésére és a megszerzett tudás megőrzésére is jut figyelem.

A hozzáférés feltételei ugyanakkor befolyásolják, ki élhet ezekkel a lehetőségekkel. A képernyőn megjelenő válasz mögött fizikai infrastruktúra és szervezeti döntések állnak. A felhasználó szabadságához ezért az is hozzátartozik, hogy megértse a szolgáltatástól való függését, és legyen módja más megoldást választani. A technika társadalmi helyét a kutatólaboratóriumok mellett az alkalmazó intézmények és a használó emberek döntései is alakítják.

Ebben a könyvben régi szerkezetekkel és elfeledettnek tűnő nevekkel is találkoztunk. A magyar kibernetika kutatói, a szegedi katicabogár alkotója és a tudományt közérthetően bemutató szerzők ugyanennek a történetnek a részei. Egy régi kibernetikai könyv lapjain különös élmény felismerni egy ma is eleven kérdést. A műszaki válasz azóta megváltozhatott, a kérdés mégis segít megérteni, mire törekedtek az előttünk járók. Emlékük megőrzése a mai eredmények arányainak felismerésében is segít.

Ez a kötet maga is egy megváltozott munkamód példája: elkészítésében mesterséges intelligencia vett részt. A téma kiválasztása, a fontosnak tartott összefüggések keresése és a szöveg alakítása így ember és gép együttműködésében történhetett. Ettől a leírt állítások még ugyanúgy ellenőrzésre szorulnak. A közös munka értékét végül az dönti el, mit ért meg belőle az olvasó, és mennyire tudja megkülönböztetni a dokumentált eredményt az értelmezéstől vagy a nyitott kérdéstől.

Az emberi ellenőrzés sem merülhet ki egy jóváhagyó mozdulatban. Aki dönt, annak lehetőséget kell kapnia a megértésre, és fel kell ismernie, ha ehhez segítségre van szüksége. Minél több munkát bízunk a gépekre, annál fontosabbá válik annak tudása, mit érdemes megkérdezni tőlük, milyen eredményt fogadhatunk el, és mely döntéseket kell más bizonyítékokra is alapozni. Ez tanulható gyakorlat, amelyet minden új alkalmazásnál tovább kell alakítani.

A történet itt véget ér, a tárgya tovább változik. Egy későbbi olvasó bizonyára talál majd olyan mai korlátot, amelyet időközben átléptek, és olyan ígéretet, amely továbbra is ígéret maradt. Remélhetőleg azt is felismeri, hogy a bizonytalanságot érdemes volt megőrizni a szövegben. A jövő megértéséhez az is hozzátartozik, hogy pontosan emlékezzünk arra, mit tudtunk és mit nem tudtunk egy adott pillanatban.

Az alkotás vágya, amellyel ez a történet elindult, velünk marad. Újabb gépeket fogunk építeni, és munkájuk révén saját képességeinkről is többet tudhatunk meg. Hogy mire fordítjuk a segítségükkel megszerzett időt, tudást és erőt, arról továbbra is nekünk kell döntenünk. Az alkotó felelőssége abban is áll, milyen életet kíván lehetővé tenni azzal, amit létrehozott.

Függelék: Rövidítések és fogalommagyarázat

Ez a függelék a könyv olvasását segítő rövidítéseket, idegen eredetű szavakat és fontosabb szakkifejezéseket magyarázza. A történeti programnevek külön jegyzékben szerepelnek, mert nem minden nagybetűvel írt név betűszó. A magyar és az angol elnevezéseket együtt adjuk meg, ahol ez megkönnyíti a visszakeresést. A szócikkek közérthető eligazítást adnak; egyes fogalmaknak más szakterületen eltérő jelentésük is lehet.

Rövidítések és jelölések

Jelölés	Eredeti elnevezés	Magyar jelentés és szerep
AAAI	Association for the Advancement of Artificial Intelligence	A mesterséges intelligencia kutatásának nemzetközi tudományos szervezete. Korábbi nevében az American szó szerepelt.
AGI	Artificial General Intelligence	Általános mesterséges intelligencia: sokféle területen rugalmasan alkalmazható gépi képesség megnevezése. Nincs mindenki által elfogadott egységes mércéje.
AI / MI	Artificial Intelligence / mesterséges intelligencia	A könyv központi gyűjtőfogalma. A tanuló rendszerek mellett szabályalapú megoldásokat is magában foglal.
ALPAC	Automatic Language Processing Advisory Committee	Az automatikus nyelvfeldolgozást vizsgáló amerikai tanácsadó bizottság; 1966-os jelentése fontos a gépi fordítás történetében.
ANN	Artificial Neural Network	Mesterséges neurális hálózat. Matematikai műveletekkel összekapcsolt feldolgozóegységekből álló modell.
API	Application Programming Interface	Alkalmazásprogramozási felület: meghatározza, hogyan kérhet egy program szolgáltatást vagy adatot egy másiktól.
BERT	Bidirectional Encoder Representations from Transformers	Kétoldali szövegkörnyezetből tanuló transzformeralapú nyelvi modellcsalád. Az eredeti változat kitakart szövegrészek előrejelzését is tanulta.
CASP	Critical Assessment of protein Structure Prediction	A fehérjeszerkezet-becslés független értékelési kezdeményezése. A CASP14 a tizennegyedik fordulót jelöli.
CC0	Creative Commons CC0 Public Domain Dedication	A jogosult a jogilag lehetséges mértékben lemond szerzői jogairól. A megjelölés nem minden esetben rendez minden más érintett jogot.
CC BY	Creative Commons Attribution	Olyan felhasználási licenc, amely a feltételek betartása mellett megengedi a megosztást és átdolgozást, a szerző megjelölésével.
CC BY-SA	Creative Commons Attribution–ShareAlike	Nevezd meg! – Így add tovább! licenc. Az átdolgozás továbbadására is előír meghatározott licencfeltételeket.
CLIP	Contrastive Language–Image Pre-training	Képek és szövegek párosításából tanuló megközelítés, illetve modell. A kétféle adat összetartozó leírásait kapcsolja össze.
CNN	Convolutional Neural Network	Konvolúciós neurális hálózat. Különösen képfeldolgozásban használt hálózattípus, amely helyi mintázatokat keres.

Jelölés	Eredeti elnevezés	Magyar jelentés és szerep
CPU	Central Processing Unit	Központi feldolgozóegység, általános feladatú processzor.
DEC	Digital Equipment Corporation	A számítógépgyártó vállalat, amelyhez az XCON szakértői rendszer üzleti alkalmazása kapcsolódott.
DL	Deep Learning	Mélytanulás. Több feldolgozási rétegből felépülő neurális modellek tanítása.
DOI	Digital Object Identifier	Digitális objektumazonosító. Tudományos közlemények és más dokumentumok tartós azonosítására használják.
EDSAC	Electronic Delay Storage Automatic Calculator	Korai brit tárolt programú számítógép. Nevében a késleltetővonalas tárolásra utaló kifejezés szerepel.
EDVAC	Electronic Discrete Variable Automatic Computer	Korai amerikai számítógép; tervezete fontos szerepet játszott a tárolt program elvének történetében.
ENIAC	Electronic Numerical Integrator and Computer	Korai amerikai elektronikus, általános célú számítógép. Első változatának programozása jelentős fizikai átkapcsolást igényelt.
GAN	Generative Adversarial Network	Generatív ellenérdekű hálózat. A generátor mintákat készít, a megkülönböztető valódi és generált példákon tanul; értékelése a generátort is tanítja.
GPS	General Problem Solver	Általános problémamegoldó. Newell, Shaw és Simon korai AI-programja; itt nem a műholdas helymeghatározást jelenti.
GPT	Generative Pre-trained Transformer	Generatív előtanított transzformer. Nyelvi modellek ismert elnevezése; nem minden nyelvi modell GPT.
GPU	Graphics Processing Unit	Grafikus feldolgozóegység. Sok párhuzamos művelet végzésére alkalmas, ezért a neurális hálózatok számításait is gyorsíthatja.
HBM	High Bandwidth Memory	Nagy sávszélességű memória. A gyorsítók és a memória közötti nagy adatforgalmat támogatja.
HMM	Hidden Markov Model	Rejtett Markov-modell. A megfigyelt jelek mögött közvetlenül nem látható, egymást követő állapotokat feltételező valószínűségi modell.
IAS	Institute for Advanced Study	A princetoni kutatóintézet neve. A könyvben az ott épített korai számítógép megjelölésében is szerepel.
IBM	International Business Machines	Számítástechnikai vállalat; a könyvben egyebek mellett a Deep Blue fejlesztőjeként szerepel.
IEA	International Energy Agency	Nemzetközi Energiaügynökség. Az adatközpontok energiaigényéről idézett elemzések egyik forrása.
ILSVRC	ImageNet Large Scale Visual Recognition Challenge	Az ImageNet-adatokra épülő nagy képfelismerési verseny, illetve értékelési feladatsor.
LLM	Large Language Model	Nagy nyelvi modell. Nagy mennyiségű nyelvi adaton tanított modell; képessége és megbízhatósága feladatonként eltérhet.

Jelölés	Eredeti elnevezés	Magyar jelentés és szerep
MIT	Massachusetts Institute of Technology	Amerikai műszaki és tudományos egyetem, a korai AI-kutatás egyik fontos helyszíne.
ML	Machine Learning	Gépi tanulás. A rendszer a példákból vagy visszajelzésekből alakítja ki működésének bizonyos összefüggéseit.
MTA	Magyar Tudományos Akadémia	A magyar kibernetika és a korai hazai számítástechnika intézménytörténetében szereplő tudományos szervezet.
NJSZT	Neumann János Számítógép-tudományi Társaság	Magyar szakmai szervezet, amely a számítástechnikai örökség megőrzésében is szerepet vállal.
NLP	Natural Language Processing	Természetesnyelv-feldolgozás. Emberi nyelvek számítógépes elemzése és előállítása.
OCR	Optical Character Recognition	Optikai karakterfelismerés. A dokumentum képéből géppel feldolgozható szöveget állít elő.
ODD	Operational Design Domain	Tervezett működési tartomány. Azok a feltételek, amelyekre egy automatizált rendszer működését tervezték.
RAG	Retrieval-Augmented Generation	Visszakereséssel kiegészített generálás. A modell a válaszadáshoz külső dokumentumrészleteket is kap.
RL	Reinforcement Learning	Megerősítéses tanulás. A rendszer a cselekvések után kapott jutalom alapján javítja a viselkedési stratégiáját.
RLHF	Reinforcement Learning from Human Feedback	Emberi visszajelzésből történő megerősítéses tanulás. A visszajelzések a modell viselkedésének utólagos alakítását segítik.
RT	Robotic Transformer	Robotvezérlésre fejlesztett transzformermodellek megnevezése; a könyvben például RT-2.
SRI	Stanford Research Institute	A Shakey robotot fejlesztő intézet történeti neve. Későbbi neve SRI International.
STRIPS	Stanford Research Institute Problem Solver	Korai tervezőrendszer, illetve az állapotok és cselekvések leírásának hozzá kapcsolódó módja.
TPU	Tensor Processing Unit	Neurális hálózatok számításaira tervezett, a Google által fejlesztett gyorsítóchip-család.
TWh	terawattóra	Energiaegység: egymilliárd kilowattóra. Nem teljesítményegység; éves villamosenergia-fogyasztás megadására is használható.
VAE	Variational Autoencoder	Variációs autoenkódoló. Belső valószínűségi reprezentáció segítségével új minták előállítására is használható modell.
VLA	Vision–Language–Action	Látást, nyelvet és cselekvést összekapcsoló modell. Robotikai alkalmazásban mozgási parancsot is előállíthat.
VLM	Vision–Language Model	Képi és nyelvi adatokat együttesen kezelő modell.
XOR	Exclusive OR	Kizáró VAGY: két logikai bemenet esetén akkor igaz, ha pontosan az egyik bemenet igaz.

Idegen szavak és szakkifejezések

A magyar kifejezés mellett az angol megfelelő is szerepel, ahol az segíti a keresést. Az önállóan elterjedt idegen szavak saját alakjukban is megtalálhatók. A szócikkek betűrendben következnek.

Adathalmaz — dataset. Meghatározott célra összegyűjtött adatok együttese. A tanításhoz, a fejlesztés közbeni ellenőrzéshez és a végső értékeléshez lehetőleg elkülönített adathalmazokat használunk.

Adatközpont — data centre. Sok számítógép működtetésére kialakított létesítmény, hálózattal, energiaellátással és hűtéssel. Nem minden adatközpont foglalkozik mesterséges intelligenciával.

Aiolipilé — aeolipile. Hérón forgó gőzgömbje. A kiáramló gőz reakcióereje forgatja; az ókori automaták egyik ismert szerkezete.

AI-tél — AI winter. Olyan visszaeső időszak megnevezése, amikor az AI-hoz fűzött várakozások, a támogatások vagy az üzleti befektetések csökkennek. Az egyes részterületeket nem egyformán érinti.

AI-ügynök — AI agent. Olyan rendszer, amely egy cél érdekében lépéseket választ, műveleteket végezhet, és az eredményt felhasználva folytatja a munkát. Önállóságának mértéke a kialakítástól és a jogosultságaitól függ.

Algoritmus — algorithm. Pontosan meghatározott eljárás egy feladat végrehajtására. Egyes eljárások csak bizonyos bemenetekre állnak meg; a mindig befejeződő megoldóeljárás ennél szigorúbb követelmény.

Állapottér — state space. Egy feladatban megkülönböztetett lehetséges helyzetek összessége. A keresőeljárás az állapotok közötti megengedett átmeneteket vizsgálhatja.

Általánosítás — generalization. A tanulás során megismert összefüggések használata új példákon. A tanítóadat jó visszaadása önmagában nem bizonyít megfelelő általánosítást.

Analitikus gép — Analytical Engine. Babbage tervezett, általános célú mechanikus számítógépe. Terve a műveletek vezérlését és az adatok tárolását is elkülöníthető feladatként kezelte.

Analóg és digitális. Analóg leírásban egy folytonosan változó mennyiség képviseli az információt. Digitális ábrázolásban elkülöníthető jeleket használunk, például bitek sorozatát.

Annotáció — annotation. Adatok kiegészítése címkékkel vagy magyarázó jelölésekkel. Például annak megadása, milyen tárgy látható egy képen.

Architektúra — architecture. Egy gép vagy modell felépítésének és fő kapcsolatainak terve. A tanítási eljárás és a megtanult paraméterek ettől különböző fogalmak.

Attention. Figyelmi mechanizmus. A modell az aktuális feldolgozáshoz különböző súllyal kapcsolja össze az adat egyes részeinek belső leírását. A megnevezés nem állít tudatos figyelmet.

Automata — automaton. Olyan szerkezet, amely meghatározott működést közvetlen emberi beavatkozás nélkül végez. Az önműködés önmagában nem jelent tanulást vagy általános intelligenciát.

Autonómia — autonomy. A működés önállóságának mértéke egy adott feladatban és környezetben. Egy önállóan mozgó gép is szigorúan korlátozott működési feltételekhez kötődhet.

Axióma — axiom. Egy formális rendszer bizonyítás nélkül elfogadott kiinduló állítása. A rendszer tételeit ezekből és a következtetési szabályokból vezetik le.

Backpropagation. Visszaterjesztés. Eljárás annak kiszámítására, hogy a hálózat paraméterei hogyan befolyásolják a hibát. A kapott jelből egy optimalizáló módszer határozhatja meg a paraméterváltoztatást.

Bayes-háló — Bayesian network. Valószínűségi modell, amely irányított, körmentes gráfban ír le változókat és feltételes függőségeiket. Bizonyítékokból más változókra is következtethet.

Bayes-i következtetés — Bayesian inference. A bizonytalanság leírásának módja: az új bizonyíték alapján módosítjuk a korábbi valószínűségi becslést. Az eredmény a kiinduló feltevésektől és a bizonyíték modelljétől is függ.

Benchmark. Összehasonlító teszt vagy egységes értékelési feladatsor. Eredménye a megadott feladatokra és mérési feltételekre vonatkozik.

Bináris — binary. Két lehetséges értékre épülő ábrázolás. A kettes számrendszer két számjegyet, 0-t és 1-et használ.

Bit. Bináris számjegy, amely két érték egyikét veheti fel. Információmennyiség egységeként is használják, meghatározott matematikai értelemben.

Bizonyossági tényező — certainty factor. Egyes szakértői rendszerekben az állítás támogatottságát kifejező szám. Nem minden ilyen tényező értelmezhető közvetlenül valószínűségként.

Boole-algebra — Boolean algebra. Logikai értékekkel és műveletekkel foglalkozó algebrai rendszer. Alapvető szerepe van a digitális áramkörök leírásában.

Chatbot. Beszélgető program. Felépülhet egyszerű szabályokból vagy tanuló modellekből is; a társalgó felület nem mondja meg önmagában a belső működést.

Church–Turing-tézis. Az a feltevés, hogy a mechanikusan, hatékony eljárással elvégezhető számításokat a Turing-gép modellje megragadja. Nem egy fizikai gép sebességéről szól, és nem szokásos matematikai tétel.

Dedukció — deduction. Következtetés, amelyben az elfogadott előfeltevésekből a szabályok szerint szükségszerűen adódik az eredmény. A helyes forma önmagában nem biztosítja az előfeltevések igazságát.

Dekóder — decoder. Egy belső kódolt leírásból kimenetet előállító egység. Képgeneráló rendszerben képet, nyelvi modellben szöveges kimenetet is létrehozhat; pontos szerepe architektúránként változik.

Determinisztikus — deterministic. Olyan működés, amely azonos teljes kiinduló állapot és bemenet esetén azonos eredményt ad. A bonyolult működés is lehet determinisztikus.

Diffúziós modell — diffusion model. Generatív modellcsalád, amely tipikusan fokozatos zajhozzáadáshoz kapcsolódó fordított, zajcsökkentő lépéseket tanul. Új minta készítésekor nem egy eredeti kép visszakeresése a feladata.

Döntési fa — decision tree. Elágazó kérdésekből álló modell. Egy útvonal kérdések sorozatán át vezet az eredményhez; a kérdések példákból is megtanulhatók.

Előtanítás — pretraining. Szélesebb körű adatokon végzett kezdeti tanítás, amely későbbi feladatokhoz használható képességeket alapozhat meg.

Embedding. Beágyazás: adat, például szó vagy kép tanult számvektoros reprezentációja. A vektorok közötti kapcsolatok a feldolgozási feladat szempontjából hasznos hasonlóságokat tükrözhetnek.

Emergens viselkedés — emergent behaviour. Több egyszerű elem vagy szabály kölcsönhatásából kialakuló összetettebb viselkedés. A kifejezés önmagában nem magyarázza meg a kialakulás mechanizmusát.

Enkóder — encoder. Az adatot feldolgozható belső reprezentációvá alakító egység. Lásd még: dekóder.

Entrópia — entropy. Az információelméletben egy eloszlás bizonytalanságának mértéke. Nem azonos az adat jelentésével vagy igazságtartalmával.

Eszközhasználat — tool use. Amikor az AI-rendszer külső programot, keresőt vagy más szolgáltatást vesz igénybe. A modell által kért műveletet a környező szoftver hajtja végre.

Felhőszolgáltatás — cloud service. Hálózaton elért számítási vagy tárolási szolgáltatás. A feladat távoli gépeken is futhat; a hozzáférés és az adatkezelés feltételeit a szolgáltatás határozza meg.

Felügyelet — oversight. A működés megfigyelése és ellenőrzése. Nem azonos a felügyelt tanulással, amely a tanítópéldákhoz rendelt célválaszokat jelenti.

Felügyelet nélküli tanulás — unsupervised learning. Mintázatok vagy szerkezet keresése olyan adatokban, amelyekhez nincs minden példára előre megadott célválasz.

Felügyelt tanulás — supervised learning. Tanulás olyan példákból, amelyekhez megadott célérték vagy helyes válasz tartozik. Például képekhez rendelt tárgynevekből.

Finomhangolás — fine-tuning. Már tanított modell paramétereinek további módosítása egy adathalmaz vagy cél alapján. Egy beszélgetésbe írt utasítás önmagában rendszerint nem finomhangolás.

Formális rendszer — formal system. Jelekből, megalkotási szabályokból és következtetési szabályokból álló pontos rendszer. Megkülönböztethető benne a jelképes levezetés és az állítás értelmezése.

Foundation model. Alapmodell. Szélesebb adatkörön tanított modell, amely többféle feladathoz felhasználható vagy tovább alakítható. A kifejezés nem garantál általános megbízhatóságot.

Generatív modell — generative model. Olyan modell, amely a tanult összefüggések alapján új adatmintát állít elő, például szöveget vagy képet. Az új kimenet nem feltétlenül teljesen eredeti vagy tényszerűen helyes.

Gólem. Zsidó legendákban ember által formált és megelevenített lény. A könyvben az alkotás, az irányítás és a készítő felelősségének kultúrtörténeti példája.

Gödel nemteljességi tételei. Megfelelő feltételek mellett egy kellően erős, ellentmondásmentes és hatékonyan megadott formális rendszerben maradnak eldönthetetlen állítások, és a rendszer nem bizonyíthatja saját ellentmondásmentességét. Ez nem minden matematikai kérdés megoldhatatlanságát jelenti.

Gradiens — gradient. A függvény változásának irányát és mértékét több paraméter szerint leíró matematikai mennyiség. A tanításban a hiba csökkentéséhez adhat irányt.

Hallucináció — hallucination. Az AI szaknyelvében megalapozottnak tűnő, de hibás vagy nem alátámasztott kimenet megnevezése. Nem feltételez emberi értelemben vett érzékcsalódást.

Hardver — hardware. A számítógép és kapcsolódó eszközeinek fizikai részei. Lásd még: szoftver.

Hebb-féle tanulás — Hebbian learning. Tanulási elképzelés, amelyben az idegsejtek együttes aktivitása a köztük lévő kapcsolatot alakítja. Gépi változatai ennek egyszerűsített matematikai modelljei.

Heurisztika — heuristic. A keresést vagy döntést segítő, gyakran tapasztalati szabály. Hasznos lehet anélkül, hogy minden esetben a legjobb megoldást garantálná.

Homeosztát — homeostat. Ashby alkalmazkodó szerkezete. Beállításainak változtatásával keresett olyan működést, amelyben a fontos változók elfogadható határok között maradtak.

Homeosztázis — homeostasis. Fontos belső változók működőképes tartományban tartása. A fogalom az élettanból került át a szabályozás és a kibernetika gondolkodásába.

Homunculus. Szó szerint emberke; a mesterségesen létrehozott ember képzete az alkímiai hagyományban. Itt kultúrtörténeti motívum, nem létező technikai megoldás.

Humanoid. Emberhez hasonló testfelépítésű robot vagy alak. A külső forma nem jelzi automatikusan az intelligencia vagy az önállóság szintjét.

Indukció — induction. Egyedi példákból általánosabb összefüggés keresése. Az eredmény új adatokkal módosulhat; nem ugyanolyan szükségszerűségű, mint a deduktív levezetés.

Inference. Gépi tanulási környezetben rendszerint a tanított modell használata, kimenetének kiszámítása. Logikában és szakértői rendszereknél következtetést jelent. Nem azonos a tanítással.

Intelligencia — intelligence. A tanulás, következtetés, alkalmazkodás és feladatmegoldás képességeinek összefoglaló neve. Pontos tartalma az alkalmazott elmélettől és mérési módtól is függ.

Jutalmazó modell — reward model. A lehetséges kimeneteket tanult szempontok alapján értékelő modell. Pontszáma nem szükségképpen az igazságot vagy a felhasználó tényleges érdekét méri.

Kibernetika — cybernetics. Az irányítás és kommunikáció összefüggéseit vizsgáló tudományos megközelítés, élőlényekben és gépekben. A visszacsatolás az egyik központi fogalma.

Kombinatorikus robbanás — combinatorial explosion. A vizsgálandó lehetőségek számának igen gyors növekedése a feladat méretével. A lehetséges lépések felsorolása emiatt hamar gyakorlatilag kezelhetetlenné válhat.

Konnekcionizmus — connectionism. Olyan megközelítés, amely sok összekapcsolt feldolgozóegység működéséből és kapcsolataiból épít fel képességeket. A neurális hálózatokhoz kötődik.

Kontextus — context. A feldolgozáskor rendelkezésre álló szöveg vagy más környezeti információ. A nyelvi modell aktuális kontextusa nem azonos a tanítás során kialakult paramétereivel.

Kontextusablak — context window. Az a korlátozott adatmennyiség, amelyet a modell egy feldolgozáskor közvetlenül figyelembe vehet. Nem azonos a teljes tanítóanyaggal vagy egy szolgáltatás minden tárolt adatával.

Konvolúció — convolution. Matematikai művelet, amely helyi mintázatok feldolgozására is használható. Neurális képfeldolgozásban a tanult szűrők a kép több helyén is alkalmazhatók.

Korpusz — corpus. Összegyűjtött szöveganyag. Nyelvészeti vizsgálatra vagy nyelvi modellek tanítására is használható.

Következtetés — reasoning. Kiinduló információkból új állítás vagy megoldási lépés kialakítása. Megvalósulhat kifejezett logikai szabályokkal vagy tanult mintázatok használatával; az eredmény helyességét külön kell vizsgálni.

Következtető gép — inference engine. Szabályok és rendelkezésre álló tények alapján következtetéseket kereső programrész. A szakértői rendszer tudásbázisát használhatja.

Küszöb — threshold. Olyan határérték, amelynek átlépése eltérő választ vagy döntést vált ki. Egy egyszerű mesterséges neuron is alkalmazhat küszöbszabályt.

Lambda-kalkulus — lambda calculus. Függvények megadására és alkalmazására épülő formális rendszer. Church kiszámíthatósági munkáinak egyik alapja.

Latens tér — latent space. A modell belső, közvetlenül nem a megfigyelt adatok formájában megadott reprezentációs tere. A benne kezelt változók tanult összefüggéseket fejezhetnek ki.

Licenc — licence. A felhasználás feltételeit meghatározó engedély. A hozzáférhető kód vagy modellsúlyok felhasználási lehetőségeit is külön feltételek szabhatják meg.

Lisp. Korai, szimbolikus feldolgozásra és listakezelésre alkalmas programozási nyelv. Fontos szerepet játszott az AI-programok és a Lisp-gépek történetében.

Lisp-gép — Lisp machine. A Lisp programozási környezet hatékony használatára kialakított speciális számítógép. Piacának visszaesése a második AI-tél történetéhez kapcsolódik.

Lyukkártya — punched card. Lyukak elrendezésével információt rögzítő hordozó. A szövőgépek vezérlésében és később az adatfeldolgozásban is használták.

Megállási probléma — halting problem. Annak eldöntése, hogy egy tetszőleges program egy megadott bemeneten megáll-e. Nincs minden lehetséges esetre helyesen döntő általános algoritmus.

Megerősítéses tanulás — reinforcement learning. A rendszer cselekvések és jutalmak alapján tanul viselkedési stratégiát. A jutalom meghatározása döntően befolyásolja, mit tanul meg.

Megértés — understanding. Összefüggések felismerése és felhasználása, például új helyzetben vagy megváltozott feltételek mellett. A gépi megértés megítélésében a viselkedés és a belső működés kapcsolata is vitatott.

Mélytanulás — deep learning. Több feldolgozási rétegű neurális modellek tanítása. A mélység az egymásra épülő átalakításokra utal.

Mesterséges neuron — artificial neuron. Bemeneteket összegző és átalakító matematikai feldolgozóegység. A biológiai idegsejtnek csak bizonyos vonásait modellezi.

Mikrovilág — microworld. Gondosan korlátozott, pontosan leírható környezet, amelyben az AI feladata vizsgálható. Az ottani siker nem terjed át automatikusan a hétköznapi világra.

Modell — model. Egy rendszer vagy összefüggés számításra használható leírása. Gépi tanulásban magában foglalhatja a választott felépítést és a megtanult paramétereket.

Moravec-paradoxon — Moravec's paradox. Megfigyelés arról, hogy a gépeknek gyakran nehezebbek az ember számára ösztönös érzékelési és mozgási feladatok, mint egyes tudatos, szabályos szellemi feladatok. Nem matematikai lehetetlenségi tétel.

Multimodális — multimodal. Többféle adattípust, például szöveget, képet és hangot összekapcsoló működés. Nem minden ilyen modell képes minden adattípust előállítani is.

Neurális hálózat — neural network. Összekapcsolt mesterséges feldolgozóegységekből álló modell. A tanulás rendszerint a kapcsolatok paramétereinek módosításával jár.

Nyílt súlyú modell — open-weight model. Olyan modell, amelynek betanított paraméterei hozzáférhetők. Ez nem jelenti automatikusan a tanítóadat, a teljes tanítási kód vagy minden felhasználási jog hozzáférhetőségét.

Optimalizálás — optimization. Egy megadott célfüggvény szerinti kedvező megoldás keresése. A gép a kijelölt mércét javítja; a mérce és a valódi cél megfelelőségét külön kell vizsgálni.

Osztályozás — classification. Adat besorolása előre meghatározott kategóriákba, például annak eldöntése, milyen tárgy látható egy képen.

Önfelügyelt tanulás — self-supervised learning. A tanítás céljelét magából az adatból állítják elő. Például a szöveg egyik részéből a következő vagy kitakart részletet kell előrejelezni.

Önjáték — self-play. A rendszer saját maga vagy másolatai ellen játszva gyűjt tanulási tapasztalatot. A játék szabályai és a jutalmazás továbbra is meghatározott keretet adnak.

Paraméter — parameter. A modell működését meghatározó mennyiség, például egy kapcsolat súlya. Tanításkor sok paraméter módosulhat; használatkor rendszerint rögzítettek.

Perceptron. Egyszerű tanuló osztályozó modell, illetve történeti modellcsalád. Az egyszerű egyrétegű változat csak bizonyos módon elválasztható minták közötti döntést tud megtanulni.

Pneumatika — pneumatics. A levegő vagy más gázok nyomását és áramlását felhasználó technika. Hérón szerkezeteinek megértéséhez is fontos.

Predikció — prediction. Előrejelzés vagy becslés megadott adatokból. Nem feltétlenül jövőbeli eseményre utal: hiányzó adat vagy ismeretlen tulajdonság becslése is lehet.

Prompt. A modellnek adott bemenet vagy útmutatás. Szöveg mellett képet és más adatot is tartalmazhat, a rendszer képességeitől függően.

Prompt injection. Olyan támadás vagy manipuláció, amelyben a rendszerhez kerülő adatba utasítást rejtenek, hogy eltérítsék a feladattól. Az adat és a végrehajtandó utasítás összekeverését használja ki.

Prototípus — prototype. Egy elképzelés kipróbálására készült korai változat. Működése még nem bizonyít tartós, gazdaságos és biztonságos üzemi alkalmazhatóságot.

Q-learning. Megerősítéses tanulási eljárás, amely állapotokhoz és lehetséges cselekvésekhez becsült hosszú távú értéket tanul. A Q nem egyszerűen a pillanatnyi jutalmat jelenti.

Regresszió — regression. Számérték becslése adatokból, például egy mennyiség várható nagyságának meghatározása. Az osztályozástól eltérően nem feltétlenül kategóriát választ.

Rejtett Markov-modell — hidden Markov model. A jelek sorozatát rejtett állapotok változásával magyarázó valószínűségi modell. Korai beszédfelismerő rendszerekben is fontos szerepet játszott.

Rekurzió — recursion. Olyan meghatározás vagy eljárás, amely önmaga egyszerűbb eseteire hivatkozik. Egy rekurzív program helyes befejeződéséhez megfelelő leállási feltételek is szükségesek.

Reprezentáció — representation. Az adat vagy tudás belső ábrázolásának módja. Lehet kifejezett jelkép, logikai állítás vagy tanult számvektor is.

Robot. Fizikai környezetében műveleteket végző gép. Lehet előre programozott, érzékelőkkel szabályozott vagy tanuló rendszerrel irányított; nem szükségképpen ember alakú.

Súly — weight. A hálózat kapcsolatának vagy műveletének tanulható erősségét meghatározó paraméter. A tanult súly és a használat közben számított figyelmi súly nem ugyanaz.

Szakértői rendszer — expert system. Meghatározott szakterület tudását és következtetési eljárásait használó program. Klasszikus változatában a tudás gyakran kifejezett szabályokból áll.

Szemantika — semantics. A jelentéssel foglalkozó terület, illetve egy jelrendszer értelmezése. Lásd még: szintaxis.

Szillogizmus — syllogism. Klasszikus logikai következtetési forma: megadott előfeltevésekből szabályos szerkezetben adódik a következtetés. Arisztotelész logikájának központi fogalma.

Szimbolikus AI — symbolic AI. Jelképeket, szabályokat és kifejezett tudásleírásokat használó AI-megközelítés. A következtetési lépések ezek műveleteiként jelennek meg.

Szimuláció — simulation. Egy rendszer működésének modellben történő utánzása. A kapott eredmény a modell feltevéseitől és pontosságától függ.

Szintaxis — syntax. A jelek megengedett elrendezésének szabályai. Egy szintaktikailag helyes mondat vagy program tartalmilag még lehet hibás.

Szoftver — software. Programok és a működésükhöz kapcsolódó leírások együttese. A fizikai eszközöktől, vagyis a hardvertől különböztetjük meg.

Sztochasztikus — stochastic. Véletlen változókat vagy valószínűségi mintavételt alkalmazó működés. Azonos szöveges kérdésre emiatt eltérő kimenet is születhet.

Tanítóadat — training data. A modell kialakításához felhasznált példák együttese. Minősége és kiválasztása hat a megtanult képességekre és hibákra.

Tárolt program — stored program. A végrehajtandó utasítások gépi formában, memóriában is tárolhatók. Ez a program megváltoztatását elválasztja a teljes gép fizikai átépítésétől.

Tenzor — tensor. A gépi tanulás gyakorlatában többdimenziós számtömb. A számvektor és a mátrix egyszerűbb példája; sok gyorsító ezek műveleteire készül.

Token. A modell által kezelt kódolási egység. Szövegben lehet szó, szórész vagy írásjel is; nem mindig azonos egy teljes szóval.

Top-5 hibaarány — top-5 error rate. Képosztályozási mérce: hibának számít, ha a helyes címke nincs az öt legmagasabbra rangsorolt válasz között. Nem azonos az első helyre tett válasz hibaarányával.

Transzformer — transformer. Figyelmi mechanizmusokra épülő neurális architektúracsalád. Nemcsak nyelvi modellekben használható; felépítése többféle lehet.

Tudásbázis — knowledge base. Egy rendszer által használható, rendezett ismeretek gyűjteménye. Tartalmazhat szabályokat, tényeket és kapcsolataikat; nem azonos automatikusan a neurális modell súlyaival.

Tudásszerzés — knowledge acquisition. A feladathoz szükséges szaktudás összegyűjtése és feldolgozható alakba öntése. A szakértői rendszerek egyik nehéz és költséges feladata volt.

Tudat — consciousness. Itt a szubjektív élmény kérdését jelöli: van-e valamilyen átélés a működés mögött. Nem következik közvetlenül egy tesztpontszámból vagy a rendszer önmagáról tett kijelentéséből.

Túlilleszkedés — overfitting. Amikor a modell a tanítóadat sajátosságaihoz túl szorosan igazodik, és emiatt új adatokon rosszabbul teljesít. Jó tanítási eredmény mellett is felléphet.

Turing-gép — Turing machine. Elméleti számítási modell, amely szalagot, jeleket és pontos átmeneti szabályokat használ. A kiszámíthatóság vizsgálatának alapvető eszköze.

Turing-teszt. Turing utánzási játékából kialakult elnevezés a gépi viselkedés megítélésére. A beszélgetés emberinek ható jellege önmagában nem bizonyít igazmondást, általános intelligenciát vagy tudatot.

Univerzális gép — universal machine. Olyan számítási modell, amely megfelelően kódolt program alapján más gépek számítását is utánozhatja. Nem jelent minden probléma megoldására képes gépet.

Validálás és tesztelés — validation and testing. A validálás gyakran a modell fejlesztés közbeni kiválasztását segíti; a külön tartott tesztadat a végső értékelést. Az elnevezések szakterületenként eltérhetnek.

Vektor — vector. Rendezett számsor, amellyel több jellemző egyszerre írható le. A modellek szavakhoz, képekhez vagy más adatokhoz is rendelhetnek vektorokat.

Vezérlés és szabályozás. Műszaki megkülönböztetésben a szabályozás a mért eredményt visszacsatolva módosítja a beavatkozást. A nyílt hurkú vezérlésben nincs ilyen eredmény-visszacsatolás; a hétköznapi szóhasználat tágabb lehet.

Visszacsatolás — feedback. A működés eredménye visszakerül a rendszer további működését befolyásoló bemenetként. A negatív visszacsatolás eltérést csökkenthet, a pozitív erősíthet; egyik jelző sem erkölcsi minősítés.

Történeti programok és rendszerek nevei

Az alábbi nevek a könyvben bemutatott programokat, modelleket vagy gépeket jelölik. Többségük tulajdonnév; önmagában a nagybetűs írásmód nem bizonyítja, hogy rövidítésről van szó.

AlexNet. A 2012-es ImageNet-versenyben kiemelkedő eredményt elérő mély neurális képosztályozó hálózat. A GPU-k, az adatok és a tanítási módszerek együttes jelentőségét mutatta.

AlphaEvolve. Nyelvi modellekkel programváltozatokat javasló és értékelőkkel összehasonlító kutatási rendszer. Algoritmusok keresésére és optimalizálására használták.

AlphaFold. Fehérjék, későbbi változataiban összetettebb biomolekuláris rendszerek térszerkezetének becslésére fejlesztett modellcsalád. Nem beszélgető program.

AlphaGo, AlphaGo Zero és AlphaZero. Tanulást és keresést összekapcsoló játékos rendszerek. Az AlphaGo emberek játszmáiból is tanult; a Zero-változatok tanulásában az önjáték kapott központi szerepet.

BKG 9.8. Hans Berliner backgammonprogramja, amely 1979-ben bemutatómérkőzésen legyőzte a világbajnokot. Ez nem jelentette a játék matematikai megoldását.

ChatGPT. Beszélgető AI-szolgáltatás neve. A szolgáltatás, az alatta működő nyelvi modell és a hozzá kapcsolt külső eszközök különböző fogalmak.

Chinook. Dámajátékprogram és kutatási projekt. Az angol dáma alapállásának gyenge megoldása azt igazolta, hogy hibátlan játékkal a játszma döntetlen.

Deep Blue. Az IBM sakkgépe, amely 1997-ben mérkőzésen legyőzte Garri Kaszparovot. Nagy számítási teljesítményt, keresést és sakkismeretet kapcsolt össze.

DENDRAL. Korai szakértői rendszer molekulaszerkezetek vizsgálatára. Kémiai ismeretekkel és mérési adatokkal szűkítette a lehetséges szerkezetek körét.

ELIZA. Joseph Weizenbaum mintázatok és átfogalmazási szabályok alapján működő beszélgető programja. Ismert DOCTOR változata egy terapeuta visszakérdező stílusát utánozta.

GNoME. Gépi tanulást használó anyagkutatási rendszer kristályszerkezetek keresésére. A számítás alapján ígéretes jelölt nem azonos a laboratóriumban előállított és bevizsgált anyaggal.

ID3. Példákból döntési fát kialakító algoritmus. A kérdések kiválasztásához információelméleti mércét is használ.

Logic Theorist. Korai szimbolikus tételbizonyító program. A megengedett lépések között keresett bizonyítást, nem pusztán kész bizonyításokat tárolt.

M–3. A korai magyar számítástechnika egyik meghatározó elektronikus számítógépe. A név géptípus-jelölés, nem egy mai AI-modell neve.

MYCIN. Fertőzésekhez kapcsolódó döntéseket és antibiotikum-választást vizsgáló klasszikus szakértői rendszer. Szabályokat és bizonyossági tényezőket használt; itt történeti kutatási rendszerként szerepel.

Nimatron és Nimrod. A nim nevű matematikai játék gépi megvalósításai. Szabályok alapján valóban játszottak; nem rejtett emberi játékos mozgatta őket.

π0 — pi-null. Különböző robotok tapasztalataiból tanított robotikai modell. Képeket, utasításokat és mozgási adatokat kapcsol össze.

Shakey. Korai mozgó kutatórobot, amely az érzékelést, a tervezést és a cselekvést kapcsolta össze. A STRIPS tervezőrendszer is a kutatásához kötődik.

SHRDLU. Terry Winograd nyelvi utasításokat értelmező programja egy korlátozott virtuális kockavilágban. A név nem a képességeit felsoroló betűszó.

XCON és R1. A DEC számítógépes megrendeléseinek összeállítását segítő szakértői rendszer nevei. A történeti R1 megjelölés nem azonos a jóval későbbi DeepSeek-R1 modell nevével.

Tájékozódási források

A magyarázatok a kötet fejezeteihez és azok forrásjegyzékeihez kapcsolódó, önálló magyar összefoglalások. A történeti rövidítések és egyes jelölések ellenőrzéséhez az alábbi eredeti közlések, intézményi gyűjtemények és hivatalos ismertetők is használhatók. A licencjelölések rövid leírásai nem helyettesítik az adott licenc teljes feltételeit.

National Research Council: Language and Machines, 1966. Az ALPAC jelentésének eredeti címlapja és szövege. Forrás

SRI: 75 Years of Innovation: Shakey the Robot. A STRIPS elnevezése és működésének ismertetése. Forrás

Smithsonian National Museum of American History: ENIAC Accumulator #2. Az ENIAC név feloldása. Forrás

Computer History Museum: 1949: EDSAC Computer Employs Delay-Line Storage. Az EDSAC név és tárolási elv. Forrás

Computer History Museum: DENDRAL. Eredeti tanulmányok és történeti dokumentumok gyűjteménye. Forrás

AAAI: About AAAI. A szervezet neve és korábbi elnevezése. Forrás

DOI Foundation: What Is a DOI? A digitális objektumazonosító ismertetése. Forrás

Creative Commons: a licencek áttekintése és a CC0 leírása. Forrás

Creative Commons: CC0 1.0 Universal. Forrás

NIST CSRC Glossary: AI. A mesterséges intelligencia különböző szakmai meghatározásai. Forrás

Idővonal

A GONDOLKODÓ GÉP NYOMÁBAN

Az élet utánzása

01

Ókor – 1948

Kr. u. I. sz.



Hérón automatái

Víz, levegő és mechanika mozgatja az ókor csodáit.
A szerkezet még előre meghatározott működést követ.

1206



Al-Dzsazarí szerkezetei

Vízórák és önműködő figurák: a mechanikus automaták építése kifinomult mesterséggé válik.

1738



Vaucanson automatái

A fuvolázó figura az ember mozgását és hangképzését utánozza mechanikus eszközökkel.

1770



Kempelen sakkozó Törökje

A látszólag gondolkodó gépben ember rejtőzik.
A bemutató mégis felveti a gépi intelligencia kérdését.

1804



A programozható szövőszék

Jacquard lyukkártyái a szövés mintáját vezérlik.
A működés megváltoztatható utasítássorral.

1837



Babbage analitikus gépe

Megszületik egy általános célú, programozható számológép terve; teljes egészében nem épül meg.

1936



Turing: mi számítható ki?

Az elméleti Turing-gép megmutatja, hogyan írható le az algoritmikus számítás, és hol vannak a határai.

1943



A mesterséges neuron modellje

McCulloch és Pitts az idegsejtek működését egyszerű logikai elemekkel közelíti meg.

1948



Wiener: Kibernetika

A visszacsatolás, az irányítás és a kommunikáció az élőlények és a gépek közös vizsgálati témája lesz.

Az automata a működést utánozza; a program a működés menetét írja le.

Források: Science Museum Group; MIT Press; Turing (1936); McCulloch–Pitts (1943).

Válogatott mérföldkövek • Az időközök nem méretarányosak • Lezárás: 2026. november 5.

Az intelligencia nyomában

02

1948 – 1993

1948–1949



Grey Walter robotteknősei

Egyszerű érzékelők és áramkörök segítségével fényt keresnek, akadályokat kerülnek meg.

1950



Turing és az utánzási játék

A gép gondolkodásának kérdését a megfigyelhető nyelvi viselkedés felől közelíti meg.

1956



Dartmouth: nevet kap az AI

Az 1955-ös kutatási terv után az 1956-os nyári projekt az új kutatási terület jelképes kezdete.

1956–1959



A magyar kibernetika műhelyei

Muszka Dániel szegedi katicabogara; az M–3 megépítése. Nemes Tihamér és Tarján Rezső munkája.

1961



Unimate a gyárban

Ipari robot áll munkába a General Motorsnál. Az automatizált mozgás önmagában még nem AI.

1966–1972



ELIZA és Shakey korszaka

A társalgóprogram és a környezetében tervező robot az AI két eltérő lehetőségét mutatja meg.

kb. 1973–1980



Az első AI-tél

A nagy ígéreteket lassabb haladás követi. Több kutatási irány támogatása visszaesik.

1980-as évek



A szakértői rendszerek fellendülése

A szaktudást szabályokba foglaló programok az üzleti alkalmazásokban is megjelennek.

1986



A neurális hálózatok új lendülete

Rumelhart, Hinton és Williams munkája segíti a visszaterjesztés széles körű elterjedését.

kb. 1987–1993



A második AI-tél

A szakértői rendszerek korlátai és a piaci kudarcok újabb kiábrándulást hoznak. A kutatás folytatódik.

A robotika, a szabályalapú AI és a tanuló hálózatok egymás mellett fejlődnek.

Források: Computer History Museum; Stanford AI100; NJSZT; Rumelhart és mtsai. (1986).

Válogatott mérföldkövek • Az időközök nem méretarányosak • Lezárás: 2026. november 5.

A GONDOLKODÓ GÉP NYOMÁBAN

A tanuló és alkotó gép

03

1997 – 2026

1997

Deep Blue legyőzi Kaszparovot

Az IBM sakkgépe mérkőzésen győzi le a világbajnokot. Kiemelkedő teljesítmény egy körülhatárolt feladatban.

2012

AlexNet és a mélytanulás áttörése

A neurális hálózat nagy előnnyel nyer az ImageNet képfelismerési versenyén; fontos szerepet kap a GPU.

2016

AlphaGo győzelme

A mélytanulás és a keresés együtt hoz áttörést: az AlphaGo 4:1-re legyőzi Lee Sedolt.

2017

Megjelenik a transzformer

Az Attention Is All You Need új architektúrát mutat be. Erre épül a későbbi nagy nyelvi modellek jelentős része.

2020

AlphaFold 2: AI a tudományban

Jelentős előrelépés a fehérjék térszerkezetének előrejelzésében a CASP14 megmérettetésen.

2022

Generatív AI és a ChatGPT

A képgenerálás és a novemberben megnyíló ChatGPT széles közönséghez viszi el a generatív modelleket.

2023–2024

Szöveg, kép és hang találkozik

Terjednek a multimodális modellek és az eszközhasználat. Az AI már többféle információt dolgoz fel együtt.

2025–2026

Ügynökök és tanuló robotok

A modellek eszközökkel, tervezéssel és robotvezérléssel kapcsolódnak össze. A megbízhatóság továbbra is kihívás.

A történet folytatódik: a látványos képesség és a megbízható működés külön kérdés.

Források: IBM; AlexNet (2012); Google Research; Google DeepMind; OpenAI (2022).

Válogatott mérföldkövek • Az időközök nem méretarányosak • Lezárás: 2026. november 5.

A gondolkodó gép nyomában

Az automatáktól a mesterséges intelligenciáig

Az ember jóval a számítógépek előtt elképzelte a mesterséges lényt. Automatákat épített, gépesítette a számolást, majd megpróbálta szabályokba foglalni a gondolkodást.

A kötet az ókori mítoszoktól és az óraszerkezetektől a kibernetikán, a gépi tanuláson és az AI telein át vezet el a mai mesterséges intelligenciáig. Bemutatja a magyar úttörőket, a játékprogramok fordulópontjait és a nyelvi modellek felemelkedését.

Közérthető, képekkel kísért technikatörténet arról, hogyan jutottunk a mozgást utánzó gépektől a tudományos kutatást is segítő rendszerekig, és milyen kérdéseket hagy nyitva ez a fejlődés.



Nagy Károly - ChatGPT