

Computers and Automata

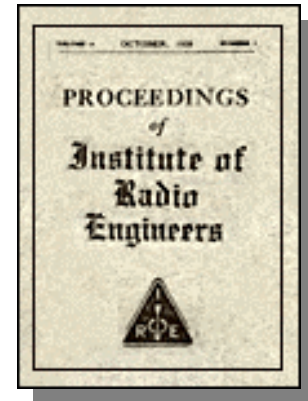
Shannon, C.E.

Bell Telephone Laboratories, Murray Hill, N.J.

Proceedings of the Institute of Radio Engineers, Oct. 1953.

Volume: 41 **Issue:**10, 1234-1241 pp.

ISSN: 0096-8390



Számológépek és automaták

Claude E. Shannon

ABSTRACT

Ez a dolgozat röviden áttekinti az automaták és a nem numerikus számítások terén mutatkozó fejlődés néhány fontosabb eredményét. Néhány tipikus gépet ír le, ideértve a logikai gépeket, a játszó gépeket és a tanuló gépeket. Néhány elméleti és fejlesztési kérdést tárgyal, mint például a számológépek és az agyvelő összehasonlítását, Turing-számológép definícióját és Neumann János önreprodukáló gépmodelljét.

BEVEZETÉS

Samuel Butler 1871-ben fejezte be az egyik legérdekesebb társadalmi satírának, az Erehonnak az írását. Ennek három fejezete, amely eredetileg Darwin a gépek között címen jelent meg, „A fajok eredeté”-nek szellemes paródiája. A satíra torz logikájával Butler úgy tekinti a gépeket, mint amelyek fokozatosan magasabb formákra fejlődnek. Tárgyalja a gépek osztályozását fajokra, nemekre és változatokra – táplálkozási szokásaikat, csökevényes érzékelő szerveiket, a reprodukáló és továbbfejlesztési mechanizmust (a nem hatékony gépek kényszerítik az embereket, hogy hatékonyabb gépeket tervezzenek), a degenerálódási tendenciákat, a csökevény szerveket, sőt „a gépek szabad akaratának” kérdését is.

Ha az ember az Erehont olvassa, azt találja, hogy a Gépek könyve meghökkenítő módon próféta. Valóban, a meglevő és a tervezett számológépek és vezérlőrendszerek az állatok és emberek egyre több funkcióját és képességét veszik fel, ma ténylegesen még sokkal nagyobb mértékben, mint ahogyan azt Butler előre elképzelte. A nagyméretű számológépek olyan simán oldanak meg bonyolult numerikus problémákat, mint ahogyan a kés szalad a vajban.

Sokunk részére azonban a számológépek legizgalmasabb lehetőségei éppen abban vannak, hogy nem numerikus műveletekre is képesek – logikai műveleteket tudnak végrehajtani, nyelveket fordítani, áramköröket tervezni, játékokat játszani, hogy koordináló és végrehajtó eszközöket tudnak működtetni – általában, hogy komplikált funkciókat látnak el, amelyek az emberi aggyal vannak kapcsolatban.

A nem numerikus számítás semmiképpen sem véletlen mellékhajtása a jobban ismert aritmetikai számításnak. A cipő inkább a másik lábon szorít. Több mint száz évvel ezelőtt Charles Babbage-nek egy selyembe szőtt portré – amelyeket az akkor már több mint ötven éve lyukasztott kártyákkal működő Jacquard-gépen szöttek – adta a ma is figyelemre méltóan

korszerű analitikus gép ötletét. A jelenleg létező legnagyobb és legmegbízhatóbb adatfeldolgozó gép még mindig az automatikus telefonrendszer. Gyáraink teli vannak szellemes és legnagyobbbrészt nem sokat emlegetett eszközökkel, amelyek szinte hihetetlen érzékelő-tulajdonságokkal rendelkeznek – továbbá sokféle adatfeldolgozó- és továbbító berendezésekkel. A vasutak és erőműrendszerek üzemi balesetek és emberi hibák elhárítására rendkívül bonyolult vezérlő és védelmi hálózatokat alkalmaznak.

Mindezek azonban speciális célú automaták. A nem numerikus számítások területének legjelentősebb elgondolása az általános felhasználású programozott számológép – olyan berendezés, amely képes arra, hogy a numerikus számológépekhez hasonlóan elemi utasítások hosszú sorozatát hajtsa végre. Ezek az elemi utasítások azonban nemcsak számokkal végzendő műveletekre vonatkozhatnak, hanem fizikai mozgásokra, szavakkal végzendő műveletekre, egyenletekre bejövő érzéki ingerekre vagy akármi más fizikai vagy fogalmi mennyiségre.

Ez a dolgozat röviden áttekinti a nem numerikus számításokra vonatkozó kutatások területét, és néhány idevágó problémát tárgyal. A terület jelenleg rendkívül aktív; ebben a rövid dolgozatban a legújabb fejlődésnek csupán néhány példáját lehet említeni.

AZ AGYVELŐ ÉS A SZÁMOLÓGÉPEK

Az agyvelőt – talán túlzott optimizmussal – sokszor összehasonlították már a számológéppel. Az agyvelőben körülbelül 10^{10} -en nagyságrendű, neuronnak nevezett aktív elem van. Minthogy a neuronok működése a „minden vagy semmi” elven alapszik, a neuronok működésükben némileg hasonlítanak a kettes számrendszerben dolgozó számológépek szerkezeti elemeihez: jelfogókhoz, elektroncsövekhez vagy tranzistorokhoz. Az elemek száma körülbelül hat nagyságrenddel nagyobb, mint a legnagyobb számológépben. McCulloch ezt szemléletesen úgy fejezte ki, hogy ha egy számológépben annyi elektroncső volna, mint neuron az agyvelőben, akkor elhelyezéséhez az Empire State Building épületére, az energiaszükséglet fedezésére a Niagara Erőműre, hűtéshez pedig a Niagara-folyóra volna szükség. Ha tranzistorokat használnánk, a számok az összehasonlításban lényegesen javulnának: az energiaszükséglet néhány száz kilowatt nagyságrendre (az agy kb. 25 W energiát ad le), a helyszükséglet pedig (szoros elhelyezés esetén) egy normális lakás színvonalára csökkenne. Azt is el lehet mondani, hogy az elektronikus eszközöknek kb. 10^3 nagyságrenddel nagyobb sebessége részben felér a nagyobb berendezési szükséglettel.

Az ilyesféle összehasonlításokat csak nagyon óvatosan szabad kezelni. Az agy működésére vonatkozó megállapítások – a nagyszámú, fontos és sok mindent megvilágító tudományos eredmények ellenére is – még mindig viszonylag nagyon primitívek. Még mindig nyílt kérdés például, hogy a neuron az alkalmas egység-e az agyvelő funkcionális elemzésére. A neuronális szint stochasztikus¹ szerkezete – a neuronok számát, elhelyezését és közös kapcsolatát illetően – azt a gondolatot veti fel, hogy a statisztikának ezen a szinten nagy szerepe van; következésképpen lehetséges a helyi sturktúrákat a matematikai modell megkonstruálása előtt átlagolni.

A számológépek és az agyvelő közötti hasonlóságokra már sokszor rámutattunk. Még többre derítenek azonban fényt a különbségek, minthogy ezek olyan fontos tulajdonságokra mutatnak rá, amelyek a jelenlegi legjobb agy modelljeinkből még hiányoznak. A legfontosabb ilyen különbségek a következők:

¹ Időben változó, véletlen jelenség – (A szerk.)

1. Méretkülönbségek. A szerkezeti elemek száma kb. hat nagyságrendben tér el, ami oly messze van mindennapi tapasztalatainktól, hogy a működésbeli extrapolációt gyakorlatilag majdnem értelmetlenné tesz.
2. Különbségek a felépítésben. Az ideghálózat szemmel láthatóan véletlen eloszlású, lokális struktúrája nagymértékben különbözik a mesterséges automaták precíz huzalozásától, ahol egyetlen rossz bekötés hibás működést okozhat. Az agyvelő valahogyan úgy van megkonstruálva, hogy általános működése nem túlzott mértékben függ a lokális struktúra pontos formájától.
3. A megbízhatósági organizáció különbsége. Az agyvelő évtizedeken keresztül minden komoly hiba nélkül működik szemben azzal az abrakadabrával, amit a számológép meghibásodás esetén produkál, annak ellenére, hogy az egyes szerkezeti elemek nem megbízhatóbbak, mint azok, amiket a számológépben használunk.
4. Különbségek a logikai organizációban. Ezek a különbségek sokkal nagyobbak, semhogy fel lehessen sorolni őket. Az agyvelő nagyjában és egészében önmagát organizáló jellegű. Tűrhetően jól alkalmazkodik a helyzetek rendkívül sokféle változataihoz. Figyelemreméltó emlékező és elérési készséggel rendelkezik, elraktározott adatokat rendkívül gyorsan képes behatárolni számos „koordinátarendszer” segítségével. Stabil szervo-rendszereket² tud organizálni, amelyek a szenzorikus³ bemenő és motorikus⁴ kimenő szervek között nagy könnyedséggel bonyolult viszonyokat valósítanak meg. Ezzel ellentétben a jelenlegi digitális számológépek olyanok, mint a tudós idióta. Hosszú aritmetikai műveleteknél a digitális számológép messze túlszárnyalja a legjobb emberi-számolókat. Ha azonban megkíséreljük a számológépet más tevékenységre programozni, az egész organizáció egyszerre alkalmatlanná és nehézkessé válik.
5. A kimenő és bemenő szervek különbségei. Az agyvelő esztétikailag is szépen tervezett bemenő szervekkel van felszerelve, elsősorban a füllel és a szemmel, amiken keresztül környezetének állapotát érzékeli. Mesterségesen létrehozott eszközeink legjobbika – mint például a Shepard-féle elemző-olvasó az írott betűk felismerésére és átírására, valamint az „Audrey” beszédfelismerő-rendszer, amely tíz számjegy hangzását tudja felismerni – összehasonlításként még csak figyelembe sem jöhet. A kimenő oldalon az agyvelő szákszámra vezérli a különböző izmokat és ízületeket. A két karnak és a kezeknek összesen körülbelül hatvan egymástól független szabadságfoka van. – Hasonlítsuk ezt össze a digitálisan vezérelt marógép manipulatív részével, amelyet a MIT-ben⁵ fejlesztették ki és amely csak három koordinátában tud mozogni. Valóban, legtöbb számológépünknek egyáltalán nincs közvetlen érzékelő vagy manipulatív kapcsolata a külvilággal, hanem csak számokból és számokon végzendő műveletekből álló absztrakt környezetben működik.

TURING-FÉLE GÉPEK

A digitális számológépek alapvető matematikai elméletét 1936-ban A.M. Turing fejlesztette ki, ma már klasszikussá vált *On Computable Numbers with an Application to the Entscheidungsproblem*⁶ című dolgozatában. Ebben definiálta a számológépeknek egy

² követő szabályozó rendszerek

³ érzékelő

⁴ mozgató – (A szerk.)

⁵ MIT – Massachusetts Institute of Technology, az Egyesült Államok legtekintélyesebb műszaki felsőoktatási és alapkutató intézménye. – (A szerk.)

⁶ Kiszámítható számok és alkalmazásuk az eldöntés problémára. – (A ford.)

osztályát, amelyet ma Turing féle gépeknek neveznek és amelyek elvileg egy végtelen távirószalagból és egy számológépszervezetből állnak. A számológépszeletnek véges számú belső állapota lehet, és alkalmas arra, hogy a szalagnak egy cellájáról leolvasson, illetve oda beírjon, továbbá hogy a szalagot egy cella szélességgel jobbra vagy balra elmozdítsa. A számoló elem egy adott időpontban meghatározott állapotban van, és leolvassa azt, ami a szalagon az éppen előtte levő cellába be van írva. A következő műveletet egyrészt a jelenlegi állapot, másrészt a szalagról leolvasott szimbólum határozza meg. Ez a művelet abból áll, hogy egy új állapotba megy át, és vagy egy új szimbólumot ír le (az éppen leolvasott helyett), vagy a szalagot jobbra vagy balra elmozdítja. Az ilyen típusú gépek alkalmasak arra, hogy számokat számítsanak ki azáltal, hogy a szimbólumok értelmezésére alkalmas kódot állapítanak meg. Így például Turing megfogalmazása szerint a gépek a szalag különböző celláiba kettes alapú számrendszerben végleges válaszokat nyomtatnak, míg a más cellákat közbenső számítások céljaira használnak.

Be lehet bizonyítani, hogy az ilyen gépek a számológépeknek rendkívül széles osztályát alkotják. Minden digitális számológép, amely nem tartalmaz véletlen vagy valószínűségszámítási elemeket, ekvivalens egy Turing-féle géppel. Minden olyan szám, amely ezeken a gépeken normális számítási folyamat útján kiszámítható, kiszámítható egy alkalmas Turing-féle géppel is. Vannak azonban – mint Turing rámutatott – bizonyos problémák, amelyeket a Turing-féle gépen nem lehet megoldani, illetve bizonyos számok, amelyeket a Turing-féle gép nem tud kiszámítani. Nem lehetséges például egy olyan Turing-gépet konstruálni, amely – ha ellátjuk egy másik Turing-gép alkalmas formában kódolt leírásával – meg tudja mondani, hogy a másik Turing-gép korlátlan ideig folytatja-e a szimbólumok lenyomtatását a végleges válasznak megfelelő cellákba, vagy nem. A számítások egy bizonyos pontján a közbenső számítások végtelen sorozatába bonyolódhat. Az, hogy ilyen típusú, gépileg meg nem oldható problémák léteznek, rendkívül érdekes a logicisták részére.

Turing kifejlesztette az univerzális Turing-gépek rendkívül érdekes elgondolását is. Ez olyan gép, amely – ha a szalagjára bármely más Turing-gép leírását alkalmasan kódolt formában felvisszük, továbbá ha a gépet egy alkalmas pontban és alkalmas állapotban indítjuk – úgy viselkedik, mint az a gép, amelyet leírtunk, vagyis kiszámítja (általában persze sokkal lassabban) ugyanazt a számot, mint a leírt gép. Turing bebizonyította, hogy lehet ilyen univerzális gépeket tervezni. Ezek a gépek természetesen alkalmasak arra, hogy bármely kiszámítható számot ténylegesen kiszámítsanak. A legtöbb digitális számológép – feltéve, hogy valamilyen típusú korlátlan memóriával rendelkeznek – ekvivalens az univerzális Turing-géppel és (legalábbis elvben) utánozhat bármely más számológépet, kiszámíthat bármely elvileg kiszámítható számot.

Turing munkáját többféle módon általánosították és átfogalmazták. Az egyik érdekes általánosítás az „A” kiszámíthatóság fogalma. Ez a Turing-gépeknek egy olyan típusa, amely a további tulajdonsággal rendelkezik, hogy a számítás menetének bizonyos pontjain kérdéseket tehet fel egy másik „jós” berendezésnek, és a válaszokat fel tudja használni a további számításoknál. Az ilyen jóstehetségű gépek például meg tudnák oldani a rendes Turing-gépek részére megoldhatatlan problémákat is; ennek következtében lehetővé tennék még szélesebb problémakör megoldását.

LOGIKAI GÉPEK

A Boole-féle algebra felhasználható arra, hogy a jelfogó és kapcsoló áramkörök tulajdonságait mechanikai módszerekkel tanulmányozzák. Másrészt viszont a Boole-féle algebra és a formális logika problémáit meg lehet oldani egyszerű jelfogós áramkörökkel. Ezt

a lehetőséget sok logikai gép felhasználja. Egy ilyen típusú gép – amelyeket McCallum és Smith ír le – hét változóig terjedő logikai relációkat tud feldolgozni. Az adott logikai probléma által megszabott relációkat megfelelő számú „konnektív egységek” megfelelő összeköttetésével állítják be: hatféle konnektív egységet alkalmaznak, nevezetesen „nem”, „és”, „vagy”, „kizáró vagy”⁷, „akkor és csak akkor” és végül „ha – akkor”. Ha az összeköttetések megtörténtek, a gépet bekapcsolják, mire a gép egymás után letapogatja a változóknak mind a $2^7 = 128$ kombinációját, és megáll mindazoknál a kombinációknál, amelyek kielégítik a feltételeket. A gép minden állapotban jelzi azt is, hogy mely változók „igazak”. McCallum és Smith a következő tipikus, a géppel megoldható példát adja:

Ismeretes, hogy a kereskedők mindig igazat mondanak, a mérnökök pedig mindig hazudnak. G és E kereskedők. C azt állítja, hogy D mérnök. A kijelenti, hogy B megerősítette, hogy C azt állította, hogy D azt mondta, hogy E kitartott amellett, hogy F tagadta, hogy G kereskedő. Ha A mérnök, hány mérnökről van összesen szó?

A gép rendkívül sokatmondó tulajdonsága a szelektív negatív visszacsatolási rendszer, amellyel a gép a logikai egyenletek egy partikuláris megoldását megtalálja anélkül, hogy kimerítően végigpróbálná az összes lehetséges kombinációkat. Ezt olyan elemek segítségével éri el, amelyek érzékelik, hogy egy partikuláris logikai reláció teljesül-e vagy sem. Ha nem, a szóban forgó logikai relációban szereplő logikai változók két lehetséges értékük között oszcillálnak. Így a nem teljesült relációkban szereplő változók állandóan változnak, míg azok, amelyek csak a teljesült relációkban szerepelnek, állandók maradnak. Ha egyszer valamennyi reláció egyidejűleg teljesül, a gép ennél a partikuláris megoldásnál megáll. Az a körülmény, hogy csak a nem teljesült relációkban szereplő változókat változtathatják, általában gyorsabban vezet egy partikuláris megoldáshoz, mint az összes lehetséges esetek rendszeres végigszámolása, de – mint ez általában előfordul, ha negatív visszacsatolást vezetünk be – a rendszer állandó lengésbe is jöhet. McCallum és Smith rámutat arra, hogy kívánatos volna a változóknak a negatív visszacsatolás útján történő megváltoztatását – amennyire csak lehet – véletlen jellegűvé tenni, ami lehetővé tenné, hogy a gép a jelfogók különböző periodikus állapotaiból folyó circulus vitiosusokból könnyebben ki tudjon kerülni.

A JÁTÉKOT JÁTSZÓ GÉPEK

A játékot játszó gépek tervezése rendkívül érdekes, és sok figyelmet fordítanak rá. A játékok szabályai olyan meghatározott „környezetet” jelentenek, amelyben a gép világosan meghatározott célzattal működhet. Az, hogy a legtöbb játék diszkrét természetű, jól illeszkedik a rendelkezésre álló digitális számítási technikához anélkül, hogy a többi manipulatív és érző gépeinknél a fizikai környezet átalakítása céljából szükséges nehézkes analóg – digitális átalakításhoz kellene folyamodni.

A játékot játszó gépeket – a növekvő bonyolultság sorrendjében – a következő típusokra lehet felosztani:

1. Szótár-típusú gépek. Itt a gép megfelelő lépése minden lehetséges helyzetre, amely a játék során előállhat, előre meg van határozva és egy „szótárban” vagy függvénytáblázatban van felsorolva. Ha egy meghatározott helyzet áll elő, a gép a szótárból egyszerűen kikeresi a szükséges lépést. A szükséges szélsőségesen nagy memória-követelmények miatt ez az elvileg érdektelen módszer csak kivételesen, egyszerű játékoknál alkalmazható.

⁷ szigorú diszjunkció. – (A ford.)

2. Olyan játékot játszó gépek, amelyekben szigorúan egyértelmű játékformulák vannak. Néhány játéknál, mint például a NIM-nél⁸, a szigorú matematikai elméletet is, amely alapján – elvileg egyszerű formula segítségével – bármely olyan pozícióban, amelyből a játék megnyerhető, a nyereshez vezető lépést ki lehet számítani. Ilyen játékoknál a formula mechanizálása tökéletes játékpartner ad.
3. Olyan gépek, amelyek csak megközelítően érvényes általános elveket alkalmaznak. A játékok legnagyobb részénél nem ismeretesek egyszerű, egzakt megoldások, de léteznek olyan különböző általános játékelvek, amelyek az előforduló pozíciók legnagyobb részében érvényesek. Ez igaz az olyan játékok esetében, mint például a sakk, bridge, póker és hasonlók. Lehet olyan gépeket tervezni, amelyek az előforduló szituációkra ezeket az elveket alkalmazzák. Minthogy azonban az elvek nem csálhatatlanok, a gépek sem azok, mint ahogy egyébként az emberek sem.
4. Tanuló gépek. Ezekben az esetekben a gépet csak a játékszabályokra és esetleg a játéknak valami elemi stratégiájára tanítják meg, továbbá módszert adnak arra, hogy a stratégiát tapasztalat útján javítani tudja. A tanulás megvalósítására javasolt sokféle módszer közül a legfontosabbak:
 - a) a próbálgatás módszere; ennél az eredményes módszert megtartják, az eredménytelen pedig elvetik;
 - b) egy jobban játszó ellenfél utánzása;
 - c) „tanítás” helyesítés vagy helytelenítés kifejezése útján, vagy úgy, hogy a gépet tájékoztatják hibáinak természetéről; és végül
 - d) a gép hibáit önmaga elemzi és így próbál általános elvekre jutni.

Az első típusú gépnek már sok példánya létezik; a harmadik típusból is néhány. A negyedik típushoz tartozó, tanulás útján játszó gépek rendkívül csábító terület a programozó és a gépkonstruktor részére.

A harmadik kategóriából – amelybe az általános elveket tartalmazó gépek tartoznak – két példa érdemel említést. Ezek közül az egyik gépet E. F. Moore és a szerző tervezték, egy Hex nevű, széles körben ismert játék céljaira. Ezt a játékot egy szabályos hatszögű mintázattal ellátott lapon játsszák; a két játékos felváltva fekete és fehér köveket tesz az üres hatszögekbe. Az egész tábla rombusz alakú: a fekete célja az, hogy a fekete kövek megszakítás nélküli sorával kösse össze a rombusz alapját és tetejét. A fehér célja az, hogy a rombusz két oldalát fehér kövekkel folyamatosan összekösse.

A játék tanulmányozása után arra a következtetésre jutottunk, hogy elfogadhatóan jó lépéshez juthatunk a következő eljárással: a játéktáblának megfelelő kétdimenziós potenciál-teret létesítünk, amelynél a fehér kövek a pozitív, a fekete kövek pedig a negatív töltéseket jelentik. A tábla alsó és felső része negatív, a két oldal pedig pozitív potenciálon van. A soron következő lépés ebben a térben egy bizonyos meghatározott nyeregpontra felel meg.

Hogy ezt a stratégiát kipróbáljuk, egy analóg eszközt készítettünk, amely egy ellenálláshálózatból és egy műszerből állt, amellyel a nyeregpontot meghatározhattuk. Az általános elv – kiegészítve néhány tapasztalati javítással – értelmes határok között használhatónak bizonyult. Ha az első lépést a gép tette, az emberi ellenféllel szemben játszmáinak körülbelül 70 százalékát megnyerte. Sokszor meglepte a tervezőket azzal, hogy látszólag értelmetlen lépéseket tett, amelyek azonban a részletes elemzés során helyesnek bizonyultak.

⁸ NIM – játék, amelyben a játszó felek valamilyen bonyolult ábrát raknak ki gyufaszálakból. Sorban egymás után szedik fel a gyufaszálakat, meghatározott irányokban haladva (pl. csak a fej felé), az elágazásoknál az utat szabadon választva. Az nyer (vagy veszít, megállapodásszerűen), aki az így választott irányokban az utolsó lehetséges gyufaszálat szedi fel. – (A szerk.)

Általában a számológépekről úgy gondolkozunk, mint amelyek kitűnően végeznek el hosszú számításokat, de értékek általános megítélésében gyengék. Paradox módon ez a gép kitűnően ítélte meg a helyzeteket. A leggyengébb pontja a kombinatorikus végjátékok megítélése volt. Az is figyelemre méltó, hogy a Hex játzó gép megfordította a szokásos számítási technikát, amennyiben egy alapvetően digitális problémát analóg gépen oldott meg.

A dáma játékot újabban – általános elvek felhasználásával – univerzális számológépre programozták. C. S. Strachey a dáma játék programozására hasonló módszert alkalmazott, mint amit a szerző a sakk programozására javasolt: nevezetesen néhány lépés összes lehetséges variációjának megvizsgálása s az eredményül kapott helyzetek minimax⁹ értékelése. A következő mintajátszmát ennek a programnak az alapján játszották, Strachey megjegyzéseivel.

(A fehér kockák – balról jobbra és alulról felfelé – folyamatosan 0-31-ig vannak számozva. A zárójelbe tett számok kiütést jelentenek.)

A gép	Strachey
11-15	23-18
7-11	21-17
8-12	20-16 a
12-21 (16)	25-16 (21)
9-14! b	18- 9 (14)
6-20 (16, 9) c	27-23
2- 7 d	23-18
5- 8	18-14
8-13 e	17- 8 (13)
4-13 (8)	14- 9
4- 5 f	9- 6
15-19	6- 1 (K)
5- 9	1- 6? g
0- 5! h	6-15 (10)
11-25 (22, 15)	30-21 (25)
13-17	21-14 (17)
9-18 (14)	24-21
18-23	26-22
23-27	22-17
5- 8 i	17-14
8- 13	14- 9
19-23	9- 6
23-26 j	31-22 (26)
27-31 (K)	6- 2 (K)

⁹ Minimax elv: a játékelmélet egyik alapelve, amely szerint úgy érdemes játszani, hogy az ellenfél maximális nyereséget biztosító stratégiája esetén is minimális legyen a veszteségünk.

7-10	2- 7
10-15	21-16? k
3-10 (7)	16- 9 (13)
10-14	9- 6
15-19	6- 2 (K)
31-27 m	2- 6
27-31 m	6- 10
31-26 m	10-17 (14)
19-23	29-25
26-31 p	

Megjegyzések:

- a) Kísérlet részemről. Az egyetlen szándékos áldozat, amelyet hoztam. Tévesen azt hittem, hogy teljesen biztos lépés.
- b) Nem láttam előre.
- c) Jobb mint 5-21 (9-17)
- d) Véletlen lépés (0 érték), konstruktív terv hiányát mutatja.
- e) Újabb 0 értékű véletlen lépés. Gyakorlatilag elég jó.
- f) Rossz; végső soron lehetővé teszi részemre, hogy egy királyhoz jussak. 10-14 jobb lett volna.
- g) Csúnya elnézés részemről.
- h) Teljesen kihasználja elnézésemet.
- i) Rossz; felszabadítja egy királynak az útját.
- j) Áldozat abból a célból, hogy egy királyt szerezzen.
- k) Ismét csúnya hiba részemről.
- l) Céltalan. A stratégia a végjátékban felmondta a szolgálatot.
- m) Túl késő.
- n) Céltalan. A játékot ebben a pontban beszüntettem, minthogy a kimenetel nyilvánvaló volt.

Noha a gép nyilvánvalóan nem világbajnok, mégis jobban játszik, mint sok ember. Strachey rámutat a program sok gyenge pontjára – elsősorban bizonyos végjáték pozíciókban – és különböző lehetséges javításokat is javasol.

TANULÓ GÉPEK

A tanulás fogalmát – csakúgy, mint a gondolkodását, öntudatát és más pszichológiai kifejezéseket – nehéz precízen meghatározni úgy, hogy a különböző érdekeltek részére egyaránt elfogadható legyen. Egy hozzávetőleges megfogalmazást a következőképpen lehet körvonalazni: tegyük fel, hogy lehetséges egy élő szervezetet vagy egy gépet elhelyeznünk vagy összekapcsolnunk környezetével: továbbá, hogy létezik a „sikeresség”-nek vagy a környezethez való „alkalmazkodásnak” bizonyos mértéke. Tételezzük fel továbbá, hogy ez a mérték olyan, hogy lehetséges a sikert a szervezet életéhez képest viszonylag rövid időtartamon belül mérni. Ha a sikerességnek ez a lokális mértéke a figyelembe a jövő környezetre vonatkoztatva időben javuló tendenciát mutat, mondhatjuk, hogy a szervezet vagy a gép a siker választott mértékszámának megfelelően megtanul alkalmazkodni a környezetéhez. A tanulás ilyen módon – azokhoz a környezetekhez viszonyítva, amelyekhez a gép alkalmazkodni tud – kvantitatív értelmet kap. Egy olyan sakkjátzó gép, amelynél a nyereség gyakorisága a működési tartam alatt nő, a fenti definíció értelmében úgy tekinthető,

mint amely sakkozni tanul: környezete azok a játékosok, akik az ellenfelei, az alkalmazkodás mértéke pedig a megnyert játszmák száma.

Egyszerű tanuló gépek konstruálására még egy sor kísérlet történt. A szerző például konstruált egy labirintusban eligazodó egeret. Ebben az eszközben egy 5 X 5 négyzetből álló területen tetszőleges labirintust lehet előállítani azáltal, hogy szomszédos négyzetek közé válaszfalakat helyeznek. A permanens mágnesezésű „egér”, amelyet a labirintusba helyezünk, bolyong, bolyong, próbálgat; közben beleütközik a különböző válaszfalakba, vagy zsákutcába kerül mindaddig, amíg véletlenül megtalálja az utat a „szalonnás kockába”. Ha már most másodszor is behelyezik a labirintusba, annak bármely részéből – amelyben korábbi próbálgatásai során már megfordult – közvetlenül megtalálja az utat a szalonnás kockába anélkül, hogy tévedne vagy hamis irányba menne. Ha a labirintus egy másik részébe helyezik, akkor addig bolyong, amíg egy korábban már megvizsgált részhez nem ér; innen már közvetlenül a célhoz megy. Eközben a memóriájába további információkat vesz fel erről a részről; ha ismét ugyanerre a pontra helyezik, egyenesen a célhoz tart. Így, ha egymás után a labirintus különböző, át nem vizsgált részeiből indítjuk, fokozatosan teljes információhálózatot épít fel, és bármely pontról közvetlenül el tudja érni a célt. Az egeret a labirintus különböző, még ki nem tapasztalt részeibe helyezve végül is értesüléseiből teljes képet épít föl, és bármely pontból el tud majd jutni a célhoz.

Ha már most a labirintust megváltoztatjuk, az egér először a régi úton próbálkozik, de minthogy válaszfalakba ütközik, új irányokat próbál ki, és átalakítja a memória tartalmát, mindaddig, amíg valamilyen úton célhoz nem ér. Ilyen módon – ha a probléma megváltozik – képes arra, hogy elfelejtse a régi megoldást.

Az egeret a valóságban egy elektromágnes hajtja, amely a labirintus alatt mozog. Az elektromágnes mozgását egy jelfogós áramkör vezérli, amely 110 jelfogóból áll, egy emlékező, valamint egy számolóáramkörbe van kapcsolva, körülbelül úgy, mint a digitális számológépeknél. Elmondhatjuk, hogy a labirintus-megoldó gép – rendkívül primitív fokon – a következő képességeket mutatja:

1. próbálgatás útján problémát old meg
2. hiba nélkül ismétli a megoldást
3. valamely partikuláris megoldáshoz hozzáadja és összekapcsolja vele az új információt
4. elfelejti a megoldást, ha tovább nem alkalmazható.

A mechanizált tanulás egy másik megközelítési módja az, hogy egy digitális számológépet alkalmas módon programozunk. A.E. Oettinger az EDSAC-számológép részére Cambridgeben, Angliában, két tanulási programot is dolgozott ki. Az elsőben ezek közül a gépet két részre osztották: az egyik a tanuló gép, a második pedig a környezet szerepét játszotta. A környezete – absztrakt formában – egy bizonyos számú üzletet reprezentált, amelyben különböző árucikkeket lehetett vásárolni; különböző üzletek különböző cikkeket tartottak raktáron. A tanuló gép azzal a problémával állt szemben, hogy megtanulja, melyik üzletben mit lehet vásárolni. Kezdetben semmi tudomása nem volt arról, hogy hol mit kaphat; a feladat az volt, hogy egy bizonyos árucikket kell megvennie. A gép találmásra kezdett keresgélni az üzletek között, míg az árucikkeket meg nem találta. Amikor ez megtörtént, megjegyezte magának, hogy hol találta a cikket. Ha másodszor ugyanarra a cikkre volt szüksége, már közvetlenül abba az üzletbe ment, ahol az árucikk raktáron volt. A programnak egy másik érdekessége az volt, hogy a gépet bizonyos mértékben „kíváncsivá” tették. Ha sikerült például a j típusú cikket egy bizonyos üzletben megtalálnia, akkor azt is megjegyezte, hogy lehet-e kapni az üzletben a $j-1$ és a $j+1$ típusú árucikket is.

Az Oettinger által leírt második tanulási program szorosabban alkalmazkodik az állatokon megfigyelhető feltételes reflexekhez. A géppel egy egész szám formájában változó intenzitású bemenő inger lehet közölni. Erre az ingerre a gép különböző módokon reagálhat, melyet ismét egy egész szám reprezentál. A kezelő a gép válasza alapján egy harmadik egész szám

beadása útján helyeslést vagy helytelenítést fejezhet ki. Amikor a gép működni kezd, az egyes ingerekre adott válaszai véletlen eloszlásúak. A helyeslés kifejezése növeli a közvetlenül megelőzőleg adott válasz előfordulási valószínűségét, a helytelenítés ezt a valószínűséget csökkenti. Továbbá, amilyen mértékben a gép egy bizonyos válaszadást a helyesléssel történő kondicionálás során megtanul, úgy csökken a kiváltáshoz szükséges inger intenzitása. Végül ha a gép által adott válaszokra közömbösséget kifejező ingereket kap, a válaszok fokozatosan zérusra csökkennek.

Az ilyen típusú programok további kifejlesztését csak a gép kapacitása, valamint a programozó energiája és tehetsége korlátozza. Sajnos, a legtöbb nagyméretű gépben rendelkezésre álló elemi utasítások kevéssé alkalmasak a tanulási programok logikai szükségleteinek céljaira, és így a gépeket rossz hatásfokkal lehet használni. A tanulási rutinban gyakran használt és logikailag egyszerű műveletek végzésére is sokszor egy tucat vagy még több utasítás is szükséges.

Egy más típusú tanuló gépet konstruált D. W. Hagelbarger. Ezt a gépet arra tervezték, hogy emberi ellenféllel szemben fej vagy írást játsszon. A gép előlapján egy indítógomb, plusz és mínusszal jelzett két lámpa és egy kétállású kapcsoló van, amelynek két állása szintén plusz és mínusszal van jelölve. Hogy a géppel játszani lehessen, a játékos a kapcsolót + vagy – állásba fordítja, és megnyomja az indítógombot. Erre a gép a két lámpa közül az egyiket felgyújtja. Ha a gép ugyanazt mutatja, mint a játékos, tehát azt a lámpát gyújtja ki, ami a játékos választásának megfelelt, a gép nyert, ellenkező esetben a játékos. Ha a játszma befejeződött, a játékos a kapcsoló megfelelő elfordításával jelzi, hogy milyen feldobást választott.

A gépet úgy tervezték, hogy az a játékos egymás után következő választásait elemzi,¹⁰ és ha megtalálta, megkísérli ezeket a szokásokat hasznosítani. Például egynémely játékosnak megvan az a szokása, hogy ha egy menetet megnyer, megismétli a korábbi lépéseit és ismét nyer, ezután pedig megváltoztatja a választását. A gép számon tartja ezeket a helyzeteket; ha ilyen tendenciák jelentkeznek, olyan módszer szerint játszik, hogy nyerjen. Ha ilyen tendenciákat nem észlel, a gép a véletlen eloszlás szerint játszik.

Azt találták, hogy a gép a játszmák 55-60 százalékát megnyeri, míg ha az ellenfél választásai tisztán a véletlenül múlnának, a játszmák 50 százalékát nyerné meg. Úgy látszik, hogy emberi lény részére rendkívül nehéz a + és – jeleket, tisztán véletlen eloszlásban produkálni (hogy a játékok elmélete alapján jogos 50 százalékos nyerési esélye legyen); még nehezebb a gépet ténylegesen legyőzni azáltal, hogy látszólagos szabályosságokat produkálva, hirtelen megváltoztatja szokásait.

Egy másik fej vagy írást játszó gépet tervezett a szerző is, amely lényegében véve ugyanezt az általános stratégiát követi, de más kritériumot használ annak eldöntésére, hogy mikor játsszon véletlenszerűen és mikor tételezze fel, hogy a látszólagos rendszer az ellenfél játékában jelentős. Miután sokáig vitattuk, hogy a két gép közül melyik győzi le a másikat, és eredménytelenül próbálkoztunk az akkor fellépő rendkívül komplikált stratégiai probléma matematikai megoldásával, ha a két gépet összekapcsoljuk, a megoldást a kísérletre bíztuk. Egy harmadik kis gépet terveztünk, amelynek a szerepe a közvetítés volt a két gép között, tehát oda és vissza továbbította az információt, hogy mikor kész a gép a lépésre, illetve, hogy mikor lépett. A három gépet ezután összekapcsoltuk és – a kísérletezők nagy mulatságára és kisebb fogadásai mellett – néhány óráig jártunk. Érdekes módon az derült ki, hogy a kisebb, meggondolatlanabb gép, következetesen legyőzte a nagyobb gépet, kb. 55-45 arányban.

Ismét egy más típusú tanuló gépet szerkesztett W. Ross Ashby, aki ezt Homeostatnak nevezte el. A homoesztázis – a szót Walter B. Cannon alkotta – az állatoknak arra a képességére vonatkozik, amellyel visszacsatolás útján olyan biológiai változókat, mint a testhőmérséklet, a

¹⁰ Megvizsgálja a stratégiáját. – (A. ford.)

véráram kémiai koncentrációja stb. stabilizálni tud. Ashby berendezése az önstabilizáló szervó-rendszernek egy fajtája. A Homeostat első modellje négy kölcsönösen összekapcsolt szervót tartalmazott. Ezek közös vezetőkei négy léptető kapcsolón, illetve az ezekhez csatlakozó ellenállásokon keresztül vezettek. Így a három hurok valamelyikének az egyensúlyból való kibillenése attól függően hat a negyedik hurokra, hogy a léptető kapcsoló milyen ellenállásokat kapcsol be. Ha bármelyik szervó az egyensúlyi helyzettől eléggé eltávolodott, egy megfelelő, határértéket jelző jelfogó jön működésbe, és a hozzá tartozó léptető jelfogót egy lépéssel tovább lépteti. Mármost egy négy szabadságfokú szervó-rendszer, amelyben véletlen eloszlású saját- és keresztérsítés szerepel, általában nem stabil. Ha ez az eset következik be, akkor egy vagy több léptető-kapcsoló elmozdul, más ellenállás-készletet kapcsol be, ami ismét más erősítési tényezőket jelent. Ha az így előálló új helyzet ugyancsak nem stabil, a kapcsolók mindaddig tovább lépnek, amíg egy stabil helyzetet nem találnak. A léptető-kapcsolókba bekötött ellenállások értékeit véletlen számok táblázata alapján választják meg. Gondoskodtak alkalmas segédberendezésekről is abból a célból, hogy a szervók állapotában önkényes változások, illetve megszorítások legyenek lehetségesek. Így például meg lehet fordítani a bekötések polaritását, két szervót össze lehet kapcsolni, egyet közülük állandó értékben tartani és így tovább. Mindezen feltételek mellett a mechanizmus tudott olyan stabil helyzetet találni, amelyben az összes szervók egyensúlyban voltak. Ha a gép céljának azt tekintjük, hogy a szervókat stabilizálja, a környezetet pedig a kezelő által előidézett különböző változások és kényszer-feltételek szolgáltatják, azt lehet mondani, hogy a Homeostat alkalmazkodik a környezetéhez.

A Homeostat bizonyos tulajdonságai a tanulógépek és agyvelő-modellek szempontjából rendkívül érdekesek. Úgy látszik, hogy bizonyos értelemben valamivel többet is csinál, mint amire tervezték. Így például olyan feltételek mellett is tudott stabilizálni, amelyekre a gép tervezése során nem is gondoltak. A véletlen eloszlás szerint választott ellenállások használata különösen érdekes, és erősen emlékeztet az agyvelő neuronjainak véletlen eloszlású kapcsolataira. Valóban Ashby lehetségesnek tartja, hogy a Homeostat szerkesztésének általános elve – amelyet Ő ultra-stabilitásnak nevez – az állati idegrendszer működésének is alapjául szolgál. Mint Ashby rámutat, az elmélet alkalmazásának egyik nehézsége abban áll, hogy a stabil megoldás megtalálásához szükséges idő a szabadságfokok számának növekedésével többé-kevésbé exponenciálisan nő. Ha mindössze csak húsz szabadságfokkal számolunk, a rendszer stabilizálódásához akkor is sok emberöltőre volna szükség. Azok a kísérletek, amelyek ennek a nehézségnek az áthidalására szolgáltak, rendkívül bonyolult elvi konstrukciókra vezettek, olyannyira, hogy nagyon nehéz még annak az eldöntése is, hogy egyáltalán működésképesek-e. Matematikai eszközeink – úgy látszik – nem elegendők ezeknek a problémáknak a megoldásához, úgyhogy a további kísérleti munka nagymértékben kívánatos.

ÖNREPRODUKÁLÓ GÉPEK

Az Erewthonban a szerző a gépek reprodukciós folyamatát úgy írja le, mint az emberek és gépek közötti szimbiotikus együttműködés egy fajtáját, amelyben a gépek az embert mint közvetítőt használják fel abból a célból, hogy új gépeket állítsanak elő, ha a régiék elhasználódtak. Az ember szerepe tehát hasonló ahhoz, amelyet a méhek a virágok megtermékenyítésében játszanak. Újabban Neumann János matematikai módszerekkel tanulmányozta a gépek valódi önreprodukáló tulajdonságait, és az ilyen „gépek” két különböző matematikai modelljét fogalmazta meg.

Az első körülbelül a következőképpen írható le. A modell „gépeit” kisszámú (kb. húszféle) elemi alkatrészből lehet megkonstruálni. Ezeknek az alkatrészeknek viszonylag egyszerű funkciójuk van: például tartószerkezet céljaira szolgáló profilvasak, elemi logikai egységek

számolási célokra – hasonlóan az egyszerűsített jelfogókhoz vagy neuronokhoz -, érzékelő elemek más elemek jelenlétének érzékelésére, összekapcsoló elemek (mint pl. a forrasztó páka) különböző elemek összekapcsolására és így tovább. Ezekből az elemekből különböző típusú gépeket lehet „konstruálni”: nevezetesen lehetséges egy univerzális konstruáló gépet szerkeszteni, hasonlóan a Turing-féle univerzális számológéphez. Az univerzális konstruáló gépet – hasonlóan a digitális számológép modelljéhez – el lehet látni egy utasítás-sorozattal, amely alkalmas kód segítségével „leírja”, hogyan lehet az elemi alkatrészekből bármely más gépet megépíteni. Az univerzális konstruáló gép ezek után környezetében meg fogja keresni a szükséges elemeket, és megépíti azt a gépet, amelyet az utasítás-sorozata leír. Ha az univerzális konstruáló gép utasítása magának az univerzális konstruáló gépnek a leírását tartalmazza, meg fogja építeni önmagának egy másodpéldányát; ezzel tehát egy önmagát reprodukáló géppé válna, kivéve azt az egy körülményt, hogy a másolat még nincs ellátva a szükséges utasításokkal. Ha az univerzális gépet kiegészítik egy utasításreprodukálóval és egy vezérlő szervvel, készen áll a valódi önreprodukáló gép. Most már az utasítások úgy írják le az eredeti univerzális gépet, hogy hozzáteszik a szalagmásoló és a vezérlő berendezés leírását is. A gép első műveletsorozata ezeknek az előállítására. A vezérlő szerv ezek után az eredeti utasítás-sorozatot reprodukálja a reprodukáló géppel, és a másolatot elhelyezi a második gépben. Végül bekapcsolja a második gépet, amely erre elkezd olvasni és végrehajtani az utasítást: megépít egy harmadik gépet és így tovább a végtelenségig.

Nemrégiben Neumann János ennél a többé-kevésbé mechanikus modellnél elvontabb önreprodukáló szerkezetet kezdett vizsgálni, melynek alapja egy kétdimenziós elemi „cellák”-ból álló hálózat. Minden cella belső szerkezete viszonylag egyszerű, körülbelül harmincféle lehetséges belső állapota van, és mindegyik cella csak a négy szomszédjával érintkezik közvetlenül. A következő megadott időpontban a cella állapota csak a cella jelenlegi állapotától és a négy szomszédnak az állapotától függ. Ezeknek az állapot-átmeneteknek az alkalmas megválasztásával el lehet jutni egy olyan rendszerhez, amely egy bizonyos fajta önreprodukáló szerkezettel rendelkezik. A szomszédos celláknak egy meghatározott csoportja mint egyetlen, szerves egység működhet és hathat a környező, nyugalomban levő cellákra oly módon, hogy ezeket azonos egységet jelentő csoporttá szervezze.

Ez a második modell mentes az alkatrészek helymeghatározásának, felismerésének és elhelyezésének azoktól a problémáitól, amelyek az első, mechanikus modellnél felmerültek. Ennek következtében matematikailag egyszerűbb megfogalmazáshoz vezet. Ezen túlmenően, bizonyos analógiákat mutat különböző kémiai és biológiai problémákhoz – mint például a kristálynövesztés vagy a gépreprodukció problémái -, míg az első modell inkább a nagyméretű állatok reprodukálásának problémájával rokon.

Mindkét modell a kritikus komplexitás fogalmához vezet, amely az önreprodukálás feltétele. Mindkét esetben csak eléggé komplikált „gépek” lesznek képesek önreprodukációra. Neumann becslése szerint nagyságrendileg több tízezer alkatrésze, illetve cellára van szükség ahhoz, hogy ez a tulajdonság fellépjen. Kevésbé komplikált szerkezetek csak egyszerűbb „gépeket” tudnak szerkeszteni, mint önmaguknál tökéletesebb, még komplikáltabb szerkezetek előállítására.

FELHÍVÁS AZ OLVASÓHOZ

Reméljük, hogy a nem numerikus számológépek fentiekben adott felsorolása eléggé felkeltette az olvasó érdeklődését ahhoz, hogy ezen a területen kutatásokhoz fogjon. Az a probléma, hogy hogyan működik az agyvelő és hogyan lehet olyan gépeket tervezni, amelyek az agyvelő működését utánozzák – bizonyára egyike a legfontosabb és legnehezebb problémáknak, amelyekkel a tudomány jelenleg szembenéz. Számptalan kérdés vár tisztázásra,

a kísérleti és fejlesztési munkától kezdve, egészen a tiszta matematikai kutatásig. Tudjuk-e a gépet olyan hierarchikusan szervezni, mint ahogy – a jelek szerint – az agyvelő szervezve van, nevezetesen úgy, hogy ha a gép tanul, fokozatosan magasabb szervezettséget ér el? Tudunk-e egy digitális számológépet úgy programozni, hogy például az általa végrehajtandó utasítások 99 százalékát maga a számológép állítsa elő, szemben azzal a néhány százalékkal, ami a jelenlegi programokban szerepel? Lehet-e olyan gépet építeni, amely megállapítja és kijavítja saját alkatrészeinek hibáit, beleértve a gép karbantartó alkatrészeket is? Mennyiben általánosítja a véletlen elem a Turing-féle gépet? Lehet-e olyan manipulatív és érzékelő eszközöket kifejleszteni, amelyek a kézzel és a szemmel összehasonlíthatók és amelyek működését egy számológép koordinálja? Lehetséges-e valamelyik Neumann-féle önreprodukáló gépet instrumentálni? Nem volna-e lehetséges a tanuláshoz valamilyen kielégítőbb elméletét megfogalmazni? Lehet-e olyan gépeket konstruálni, amelyek – tisztán az általános funkcionális jellegük alapján – más gépeket konstruálnak? Melyik az az utasítás-sorozat, amely a digitális számológépeket a legjobban alkalmassá teszi nem numerikus számításokra? Hogyan lehet a számológép memóriáját úgy megszervezni, hogy asszociáció útján tanuljon is, emlékezzen, körülbelül úgy, mint az emberi agyvelő?

Ezeket a tipikus kérdéseket, valamint az automaták egész területét felhívásnak szántuk az olvasó részére. A kutatásnak ez a területe megérett a tudományos feltárássra. Itt nem arról van szó, hogy régi dolgokat kell új életre kelteni: új, gazdag lelőhelyeket – és egyes esetekben talán felszínen heverő aranyrögöket – lehet találni.

SZÁMOLÓGÉPEK ÉS ÉLŐ SZERVEZETEK ÖSSZEHASONLÍTÁSA

Az élő szervezetek vegyes jellege.

Ha megvizsgáljuk a központi idegrendszert, benne mindkét eljárás – a digitális és az analóg – elemei megkülönböztethetők.

A neuron egy impulzust továbbít. Ez látszik elsődleges funkciójának, bár erről a funkcióról és kizárólagos vagy nem kizárólagos jellegéről még távolról sem mondtuk ki az utolsó szót. Az ideg-impulzus lényegében véve minden vagy semmi jellegű, hasonló egy bináris számhoz. Itt tehát egy digitális elemmel van dolgunk, de nyilván ez még nem minden. Ami a szervezetben végbemegy, annak nagy része nem így közvetítődik, hanem a véráram vagy más humorális közeg általános kémiai összetételétől függ.

Ismeretes, hogy a szervezetben különböző összetett funkcionális láncolatok vannak, amelyek különböző fokozatokon mennek át az eredeti ingertől a végső hatásig – egyes fokozatok neurálisak, azaz digitálisak, mások humorálisak, azaz analógok. Ezek a digitális és analóg részek az ilyen láncolatban váltakozva többször is előfordulhatnak. Ennek a típusnak bizonyos eseteiben a láncolat a valóságban önmagával visszacsatolásban lehet, azaz végső kimenetele ismét ingerelheti eredeti bemenetét.

Ismeretes, hogy az ilyen vegyes (részben neurális és részben humorális) visszacsatolt láncolatok rendkívül fontos folyamatokat idézhetnek elő. Így például az a mechanizmus, amely a vérnyomást állandó szinten tartja, vegyes jellegű. Az az ideg, amely érzékeli és jelzi a vérnyomást, egy neurális impulzus sorozattal, vagyis digitális módon teszi ezt. Azokat az izom összehúzódásokat, amelyeket ez az impulzusrendszer indít meg, még le lehet írni több digitális impulzus szuperpozíciójaként. De az ilyen összehúzódás befolyása a véráramra már hidrodinamikus jellegű, és ezért analóg. Az így keletkezett nyomás visszahatása arra az idegre, amely a nyomást jelzi, zárja a visszacsatolási láncot, és ezen a ponton az analóg folyamat ismét digitálisba megy át. Ezért az élő szervezet és a számológépek közti hasonlóság ezen a ponton bizony nem tökéletes. Az élő szervezetek rendkívül összetett, részben digitális, részben analóg mechanizmusok. A számológépek, legalábbis jelenlegi formájukban,

amelyekről jelen előadásomban beszélek, tisztán digitálisak. Ezért arra kell kérnem Önöket, fogadják el a rendszernek ezt a túlegyszerűsítését. Bár tudatában vagyok annak, hogy az élő szervezeteknek vannak analóg összetevői, és abszurd dolog volna tagadni fontosságukat, mégis a tárgyalás egyszerűsítése kedvéért eltekintek ettől a résztől. Úgy tekintem az élő szervezetet, mintha tisztán digitális automata volna.

Az egyes elemek vegyes jellege. Ezzel kapcsolatban úgy is lehet érvelni, hogy a neuron nem pontosan digitális szerv. Ezt a tényt ismételtelen hangsúlyozták, méghozzá erőteljesen. Ebben kétségtelenül sok igazság van, ha az ember a dolgot kellő részleteiben vizsgálja. Ebben a vonatkozásban fontos leszögezni, hogy a teljesen kifejlődött idegimpulzus, amelynek „minden vagy semmi” jelleget tulajdonítunk, nem elemi jelenség, hanem összetett. Ez egyszerűsített állapota annak a bonyolult elektrokémiai komplexusnak, amely a neuront jelenti, és amelyet működésének teljes elemzése folyamán analóg gépnek tekintünk. Valóban lehetséges a neuront gerjeszteni úgy is, hogy a „kisülés”, amely az idegingerületet megszakítja, nem következik be. A „küszöb alatti idegingerületeknek” ebben a tartományában (vagyis a leggyengébb ingerületeknél) először azt találjuk, hogy a válasz arányos az ingerrel, míg később (magasabb, de még mindig küszöb alatti ingereknél) a válaszok bonyolultabb, nem lineáris törvénynek engedelmesskednek, de mégis folytonosan változók, nem szakadásos jellegűek. Más bonyolult jelenségek is vannak a küszöbértéken innen és túl: fáradtság, összegeződés, hiányos önregzési formák stb.

Annak ellenére, hogy ezek a megjegyzések helytállóak, figyelembe kell venni, hogy ez a „minden vagy semmi” szerv fogalmának talán indokolatlanul szigorú kritikája. Az elektromechanikus jelfogó vagy az elektroncső kétségtelenül „minden vagy semmi” szervek, ha megfelelően használják őket. Sőt az ilyen szervek prototípusai. A valóságban mégis mind a kettő bonyolult analógiás mechanizmus, amelyek megfelelően szabályozott gerjesztésre folytonosan, lineárisan vagy nem lineárisan reagálnak, és csak egészen különleges jellegű működési feltételek mellett mutatnak „kisülési”, „minden vagy semmi” válaszokat.

Kevés a különbség e viselkedés és a neuronok fentebb leírt viselkedése között. Másképpen szólva: egyik sem kizárólagosan „minden vagy semmi” jellegű szerv (technológiai vagy fiziológiai tapasztalatunkban nagyon kevés tény mutat arra, hogy abszolút „minden vagy semmi” szervek egyáltalán léteznek): ez azonban nem fontos. „Mindén vagy semmi” szerven olyan szervet kell értenünk, amely eleget tesz a következő két feltételnek: először, bizonyos megfelelő működési feltételek mellett „minden vagy semmi” módjára működik; másodsor normálisan csakis ilyen működési feltételek között használják, ezek képviselik a dolgok normális állását azon a nagy organizmuson belül, amelynek a szerv része. Így nem az a fontos, hogy egy szerv szükségszerűen és minden körülmények között „minden vagy semmi” jellegű-e – ez valószínűleg sohasem fordul elő -, hanem az, hogy tulajdonképpen összefüggésében elsődlegesen úgy működik-e és előreláthatólag úgy szándékozunk-e működtetni, mint egy „minden vagy semmi” szervet. Belátom, hogy ez a definíció nemkívánatos kritériumokat vezet be, mint „tulajdonképpen” összefüggés, „előreláthatólag” és „szándékozunk”. De nem tudom, hogyan lehetne elkerülni ezeknek a használatát, és hogyan hagyhatnánk számításán kívül a józan ész alkalmazásunknál. Ennek megfelelően a továbbiakban azt a munkahipotézist használom, hogy a neuron „minden vagy semmi” jellegű digitális szerv. Tudom, hogy evvel kapcsolatban még nem mondtuk ki az utolsó szót, de remélem, hogy ez a kitérés a munkahipotézis korlátaira, és alkalmazásának okaira megnyugtatja Önöket. Csupán egyszerűsíteni szeretném gondolatmenetemet; nem kívánok előre állást foglalni semmiféle lényeges nyitott kérdésben.

Ugyanígyen értelemben megengedhetőnek tartom, hogy úgy beszéljek a neuronról, mint elektromos szervről. A neuron ingerlését, az impulzus kialakulását és továbbhaladását a szinapszisban és az impulzust ingerlő hatását elektromosan lehet leírni. A kísérő kémiai és más folyamatok azért fontosak, hogy megérthessük az idegsejt belső működését. Még

fontosabbak is lehetnek, mint az elektromos jelenségek. Sőt úgy látszik, nagyon is szükségesek ahhoz, hogy a neuront mint „fekete dobozt”, mint „minden vagy semmi” jellegű szervet írjuk le. A helyzet itt sem rosszabb, mint például az elektroncsőnél. A tisztán elektromos jelenségeket itt is a szilárd testek fizikájának, a termodinamikának, a mechanikának számos más jelensége kíséri. Mindezek fontosak ahhoz, hogy megértsük az elektroncső szerkezetét, de a legjobb kihagyni őket a tárgyalásból, ha az elektroncsövet sematikus leírással „fekete doboz”-ként akarjuk tárgyalni.

Kapcsolódó vagy jelfogó szerv fogalma. A neuron és az elektroncső, ha a fenti értelemben tekintjük őket, ugyanannak a fajtának két példánya, amelyet közkeletűen „kapcsolószerv”-nek vagy „jelfogó szerv”-nek szoktak hívni. (Az elektromechanikus jelfogó természetesen más.) Az ilyen szervet „fekete doboz”-nak definiáljuk, amely egy speciális ingerre vagy ingereknek a kombinációjára energetikailag függetlenül reagál. Vagyis azt kívánjuk, hogy a reakciónak legyen elég energiája ahhoz, hogy több hasonló jellegű ingert keltsen, mint amilyen őt keltette. Ezért a válasz nem nyerheti energiáját az eredeti ingertől. Más, ettől független energiaforrásból kell származnia. Az inger csak irányítja, vezérli ennek a forrásnak az energiáját.

(A neuron esetében ez a forrás a neuron általános anyagcseréje. Az elektroncső esetében pedig az az áramforrás, amely fenntartja a katód-anód potenciálkülönbséget, tekintet nélkül arra, hogy vezet-e cső vagy sem, valamint kisebb mértékben a fűtőáram, amely „kiforralja” az elektronokat a katódból. Az elektromechanikus jelfogó esetében az az áramforrás, amelynek útját a jelfogó zárja vagy nyitja.)

Az élő szervezetek alapvető kapcsoló elemei, legalábbis abban a mértékben, ahogyan mi itt vizsgáljuk őket, a neuronok. Az új típusú számológépek alapvető kapcsoló elemei elektroncsövek; a régebbieknél részben vagy egészben elektromechanikus jelfogók voltak. Lehetséges, hogy a számológép nem mindig lesz főként kapcsolószervek összessége, de fejlődésnek ez a formája még a távoli jövő kérdése. Sokkal közelebb van az a fejlődés, hogy az elektroncsövek helyett más kapcsolószerveket fognak használni a számológépekben. De még ez sem következik be néhány éven belül. Ezért a számológépeket kizárólag abból a szempontból fogom tárgyalni, mint elektroncsövekből álló kapcsolószervek összességét.¹¹

Nagy számológépek és élő szervezetek nagyságrendi összehasonlítása. Két jól ismert nagyon nagy, elektroncsövekből álló számológép van és működik. Mindkettő körülbelül 20 000 kapcsolószervből áll. Az egyik tisztán elektroncsöves gép. (Az Egyesült Államok Hadseregének Tüzérségi Osztálya Ballisztikai Kutató Laboratóriuma Aberdeen, Maryland, tulajdonában van, elnevezése „ENIAC”.) A másik vegyes – részben elektroncsöves, részben elektromechanikus jelfogós. (Az IBM Corporation tulajdona, New Yorkban van, elnevezése „SSEC”.) Ezek a gépek jóval nagyobbak, mint azok az elektroncsöves számológépek lesznek, amelyeket a következő néhány évben valószínűleg építenek és üzembe helyeznek. Valószínűleg ezek mintegy 2000-6000 kapcsolószervből fognak állni. (A csökkenés oka az, hogy másképpen oldják meg a „memóriát”, de erről most nem beszélek.) Lehetséges, hogy később ismét növekedni fog a gépek nagysága, de nem valószínű, hogy túllépik a 10 000 (vagy néhányszor 10 000) kapcsolószervet mindaddig, amíg a jelenlegi technika és szemlélet alapján állunk. Mindent összevéve, a számológép kapcsolószerveinek megfelelő nagyságrend mintegy 10^4 .

Ezzel szemben a központi idegrendszer neuronjainak a száma különböző becslések szerint körülbelül 10^{10} nagyságrendű. Nem tudom, mennyire pontos ez a szám, de feltehető, hogy a kitevő legfeljebb egy egységgel lehet magasabb vagy alacsonyabb. Így nagyon feltűnő, hogy a központi idegrendszer legalább milliószor akkora, mint a legnagyobb mesterséges automata, amiről jelenleg beszélhetünk. Rendkívül érdekes megvizsgálni, miért van ez így és milyen

¹¹ A ma gyártott számológépek elsősorban félvezetős, másodsorban mágneses kapcsolóelemeket tartalmaznak. Ez a cikkben tárgyaltakat nem befolyásolja lényegesen. – (A szerk.)

elvi kérdések következnek ebből. Azt hiszem, néhány világosan megfogalmazható elvi kérdés valóban összefügg ezzel.

Az elemek szignifikáns számarányainak meghatározása. Nyilvánvaló, hogy az elektroncső, amilyenek ma ismerjük, óriási az idegsejthez képest. Mérete és energia leadása is körülbelül milliárdszor nagyobb. (Ilyen számokat természetesen nem lehet egyértelmű érvényességgel megadni, de a fenti számok tipikusak.) Más szempontból összehasonlítva, a helyzet ellenkező. Elektroncsöveket rendkívül nagy sebességgel lehet működtetni a számológépektől eltérő, más alkalmazásokban, de ezek most nem érdekelnek bennünket. Számológépekben a maximum jóval kisebb, de még mindig igen tekintélyes. A fejlődés jelenlegi helyzetében általános vélemény szerint mintegy másodpercenként egymillió műveletet végeznek. Az idegsejt válaszai lényegesen lassabbak ennél, talán a másodperc 1/2000 részét teszik, és ami tulajdonképpen lényeges, a legrövidebb időköz, amely szükséges ahhoz, hogy az ingertől számítva tökéletesen helyreálljon és újabb ingert fogadjon, még hosszabb, mint ez – legjobb esetben is a másodperc 1/200-ad része. Ez 1:5000 arányt ad, amely azonban túlságosan kedvező az elektroncsőre nézve, mert az elektroncsövek, ha másodpercenként 1 000 000 lépés sebességű kapcsolószerként használják őket, sohasem működnek 100 százalékos kihasználással. Ezért 1:2000 arány megfelelőbbnek látszik. Így a körülbelül milliárdszoros energia befektetéssel működő elektroncső csak valamivel több mint ezerszeresen haladja túl a neuront. Ezért van valami igazság abban, amikor azt mondják, hogy nagyságrendileg mintegy milliószor kevésbé hatékony.

Az alapvető tény minden tekintetben a neuron kis terjedelme az elektroncsőhöz képest. Az arány mintegy egymilliárd, amint előbb kimutattunk. Mi ennek az oka?

A szélsőséges méretarány okainak elemzése. Ennek az eltérésnek az eredete az elektroncső alapvető vezérlő szervében, jobban mondva vezérlő elrendezésében keresendő, összehasonlítva a neuronéval. Az elektroncsőben a vezérlés kritikus területe a katód (ahol a hatáshordozók, az elektronok keletkeznek) és a rács (amely vezérli az elektronáramot) közötti tér. Ez a tér körülbelül egy milliméter átmérőjű. Ennek a neuronnál az idegsejt fala felel meg, a „membrán”. Vastagsága körülbelül egy mikron (1/1000 milliméter) vagy valamivel kevesebb. Tehát ezen a ponton a lineáris méretek aránya közelítőleg 1:1000-hez. Nos, ez a fő különbség. A vezérlő térben levő elektromos tér közelítőleg ugyanakkora az elektroncsőnél, mint a neuronnál. A potenciálkülönbségek, amelyekkel ezeket a szerveket megbízhatóan lehet vezérelni, az egyik esetben néhányszor 10-volt, a másik esetben néhányszor 10 millivolt. Arányuk ismét 1:1000-hez, és ezért a gradiensek (a téreőrök) közelítőleg egyenlők. Mármint 1:1000 arány a lineáris méretekben megfelel 1:1 000 000 000 aránynak a térfogatban. Így tehát a milliárdos eltérési tényező háromdimenziós terjedelemben (térfogatban) megfelel – amint ennek lenni kell – 1000-es eltérési tényezőnek lineáris méretekben, vagyis annak a különbségnek, amely az elektroncső egymilliméteres elektródák közti termélysége és a neuron mikronos membránvastagsága között van.

Érdemes megjegyezni, bár egyáltalán nem meglepő, hogyan vezet két olyan tárgy közti eltérés, amelyek mindkettő mikroszkopikusak és elemi komponensek belsejében helyezkednek el, észrevehető makroszkopikus különbségekhez a belőlük épült szervezetekben. Az egy milliméteres tárgy és az egy mikronos tárgy közti különbség az oka annak, hogy az ENIAC súlya 30 tonna és 150 kilowatt energiát ad le, míg az ember központi idegrendszere, amely funkcionálisan mintegy milliószor nagyobb, fél kiló nagyságrendű, és elfér az ember koponyájában. Az ENIAC súlyát és nagyságát tekintve, nem szabad elfelejteni, hogy ez az óriási készülék csak azért szükséges, hogy 20, egyenként tízjegyű számot, vagyis összesen 200 decimális jegyet dolgozzon föl, ami mintegy 700 bináris jegynek, azaz csupán 700 egyidejű „igen-nem” információnak felel meg!¹²

¹² A mai korszerű berendezések kb. 2-3 nagyságrenddel kisebbek a méretek és disszipált teljesítmények szempontjából, de a leküzdendő különbség még mindig igen nagy. – (A szerk.)

A nagy méretek okainak technológiai értelmezése. Ezekből a megfontolásokból világos, hogy jelenlegi technológiánk, amellyel nagysebességű és nagymértékben összetett információkat kezelünk, nagyon tökéletlen. Az erre készült apparátus rendkívül nagyméretű, mind terjedelem, mind energiaszükségletét tekintve.

Ennek a technológiának a gyengesége valószínűleg, legalábbis részben, az alkalmazott anyagokban van. Jelenlegi technikánk szerint fémeket használunk, amelyeket rendkívül kis távolság és bizonyos kritikus pontokon csak vákuum választ el egymástól. A közegeknek ez a kombinációja jellegzetes mechanikai instabilitást mutat, amely teljesen idegen az élő természettől. Ezen azt az egyszerű tényt értem, hogy ha egy élő szervezetet mechanikai sérülés ér, az igyekszik a hibát saját maga kiküszöbölni. Másrészt, ha egy kalapáccsal ráütünk egy ember készítése mechanizmusra, ilyen önhelyreállítási tendencia nincs. Ha két fém nagyon közel van egymáshoz, a körülvevő közegben mindig előfordulhat, hogy érintkezésbe kerülnek egymással. Ha elektromos potenciáljuk különböző volt, ez a rövidzárlat elsősorban azt eredményezheti, hogy elektromosan összeforrnak, és az érintkezés állandóvá válik. Ekkor ezen a ponton tényleges és állandó hiba lép fel. Ha megsértjük egy idegsejt falát, semmi ilyesmi nem következik be. Ellenkezőleg, a membrán rendszerint rövid időn belül helyreállítja önmagát.

Anyagainkban ez a mechanikai instabilitása akadályoz bennünket abban, hogy tovább csökkentsük a mértéket. Ez az instabilitás és más hasonló jellegű jelenségek teszik alkatrészeink viselkedését nem teljesen megbízhatóvá, még jelenlegi nagyságuk mellett is. Így az, hogy a mi anyagaink gyengébb minőségűek, mint azok amelyeket a természet használ, akadályoz bennünket abban, hogy a bonyolultság olyan magas fokát és olyan kis méreteket érijünk el, mint amilyenek a természetes szervezetekben megvalósulnak.

AZ AUTOMATÁK ELJÖVENDŐ LOGIKAI ELMÉLETE

További megfontolások azokról a tényezőkről, amelyek korlátozzák a mesterséges automaták jelenlegi nagyságát. Hangsúlyozzuk, mennyire korlátozott a mesterséges automaták bonyolultsága¹³ vagyis az a bonyolultság, amely mellett még különleges nehézségek nélkül lehet dolgozni, és amely mellett az automata előreláthatólag megbízhatóan működik. Már megadtunk két okot, amelyek ebben az értelemben korlátozzák a bonyolultságot. Ezek: a nagy méretek és a jobb híján felhasználandó szerkezetelemek korlátozott megbízhatósága, mindkettő annak a következménye, hogy olyan anyagokat használunk, amelyek egyszerűbb alkalmazásoknál teljesen kielégítőek, de alacsonyabb rendűek a természetes alkatrészeknél, és így a rendkívül bonyolult felhasználáskor éppen csak hogy megfelelnek. Van azonban még egy harmadik fontos korlátozó tényező is, és most erre fordítsuk figyelmünket. Ez a tényező intellektuális és nem fizikai jellegű.

Az automaták logikai elméletének hiányából folyó korlátozás. Még távolról sem vagyunk birtokában az automatákra vonatkozó olyan elméletnek, amely megérdemli ezt a nevet, vagyis valóban matematikai-logikai elmélet. Napjainkban a formális logikának, speciálisan a matematikára alkalmazott logikának jól kidolgozott rendszere ismeretes. Ennek a tudományágnak sok jó oldala van, de bizonyos komoly gyenge pontjai is vannak. Ez alkalomból nem térek rá a jó oldalakra, bár távol áll tőlem, hogy lekicsinyeljem őket. Nem megfelelő vonásairól a következőket mondhatom: Aki már dolgozott a formális logika területén, megerősítheti, hogy ez a matematikának technikailag egyik legmerevebb része. Ennek az az oka, hogy merev, minden vagy semmi fogalmakkal dolgozik, és nagyon kevés

¹³ Bonyolultság (komplexitás) az élő szervezetek és technikai berendezések jellegzetes tulajdonsága. Lényegében azt jelenti, hogy a rendkívül nagyszámú építőelem (sejtek, alkatrészek) sokszorosan összetett kölcsönhatásban van egymással. A bonyolultság fokának és mértékének egyelőre még nincs pontos matematikai meghatározása.

érintkezési pontja van a valós és komplex számok folytonos fogalmával, vagyis a matematikai analízissel. Pedig az analízis a matematika technikailag legsikeresebb és legjobban kidolgozott része. Így a formális logika, éppen a problémák megközelítésének természete miatt el van szigetelve a matematika legjobban kimunkált részeitől, és belekényszerül a matematikai területek legnehezebb részébe, a kombinatorikába.

A digitális, minden vagy semmi jellegű automaták elmélete, amelyet eddig tárgyaltunk, bizonyára a formális logika egy fejezete. Ezért úgy látszik, hogy osztoznia kell a formális logika előbb említett nem vonzó tulajdonságában. Matematikai szempontból inkább kombinatorikus lesz, mint analitikus.

A kombinatorikus elmélet valószínű jellemzői. Nos, azt hiszem, a valóságban nem így áll a dolog. Az automaták működését tanulmányozva nyilván figyelmet kell fordítani egy körülményre, amely a formális logikában eddig még sohasem jelent meg.

Az egész modern logikában egyetlen dolog fontos: vajon egy eredményt el lehet-e érni véges számú elemi lépéssel vagy sem. Viszont a szükséges lépések száma aligha érdekelte valaha is a formális logikát. Helyes lépéseknek minden véges sorozata elvileg ugyanolyan jó, mint bármelyik másik. Nem számít, hogy számuk kicsi vagy nagy, akár olyan nagy, hogy egy életem át, sőt az általunk ismert világegyetem várható élettartama alatt sem lehet elvégezni. Ha automatákkal dolgozunk, ezt az állítást lényegesen meg kell változtatni. Egy automata esetében nemcsak az a fontos, hogy bizonyos eredményt egyáltalán véges számú lépésekben tud-e nyújtani, de az is, hogy hány lépésre van szüksége ehhez. Ennek két oka van. Először is, automatákat azért készítenek, hogy bizonyos eredményeket bizonyos előre meghatározott idő alatt kapjanak, vagy legalábbis előre meghatározott nagyságrendű idő alatt. Másodszor felhasznált alkatrészek minden egyes műveletnél kicsi, de nem zérus hibavalószínűséggel dolgoznak. Elég hosszú művelet sor esetén ezeknek az egyedi hibavalószínűségeknek a felhalmozódó hatása (ha nem ellenőrzik) elérheti az egység nagyságrendjét, amikor is teljesen megbízhatatlanná válik. Az ebben szerepet játszó valószínűségek színvonala nagyon alacsony, de mégsem esik teljesen távol a megszokott technológiai tapasztalat területéről. Nem nehéz felbecsülni, hogy egy nagysebességű számológépnek egy tipikus probléma megoldása közben esetleg 10^{12} egyedi műveletet is el kell végeznie. Egy egyedi művelet még tűrhető hibavalószínűsége még ezért 10^{-12} -höz képest kicsi kell hogy legyen. Megemlítem, hogy egy elektromechanikus relé (egy telefon jelfogót) jelenleg akkor tartanak még elfogadhatónak, ha hibavalószínűsége egy egyedi műveletnél 10^{-8} nagyságrendű. Kitűnőnek tartják, ha ez a valószínűségi nagyságrend 10^{-9} . Így a nagysebességű számológépekben megkívánt megbízhatóság nagyobb, de nem elérhetetlenül nagyobb, mint azok, amelyek bizonyos ipari területeken a józan gyakorlatban előfordulnak. Azonban úgy látszik, hogy a jelenleg elérhető megbízhatóság nem áll nagyon távol a most említett minimális követelményektől. Ezért ezt feltétlenül szükséges kimerítően tanulmányozni és felállítani nem-triviális elméletét.

Így az automaták logikája két lényeges szempontban különbözik majd a formális logikától:

1. Tekintetbe kell venni a „gondolatsorok”, vagyis a műveletsorok tényleges hosszát.
2. A logikai műveleteket (szillogizmusokat, konjunkciókat, diszjunkciókat, megációkat stb., azaz az automatáknál szokásos terminológia szerint a különböző kapu-, koincidencia, antikoincidencia, záró stb. műveleteket) olyan eljárással kell tárgyalni, amelyek kicsi, de nem zérus valószínűséggel megengednek kivételeket (hibás működést).

Mindez olyan elméletekhez fog vezetni, amelyek sokkal kevésbé mereven „minden vagy semmi” természetűek, mint a formális logika a múltban és a jelenben. Ezek az elméletek sokkal kevésbé kombinatorikus és sokkal inkább analitikus jellegűek lesznek. Valóban számos jel készített bennünket arra a meggyőződésre, hogy a formális logikának ez az új rendszere közelebb jut egy másik tudományterülethez, amelynek a múltban a logikához kevés köze volt. Ez a termodinamika, elsősorban abban a formájában, amelyet Boltzmann alkotott

meg; az elméleti fizikának ez a része jut egyes vonásához. Technikája sokkal inkább analitikus, mint kombinatorikus; ez is illusztrálja azt az érvet, amelyet korábban megpróbáltam kifejteni. De túlságosan messzire vezetne, ha ez alkalomból mélyebben belemerülnék ebbe a témába.

Mindez ismét aláhúzza azt a konklúziót, amelyet korábban már jeleztem, vagyis azt, hogy szükség van az automaták és információk erősen matematikai jellegű, speciálisan analitikus elméletére. Jelenleg még csak az elmélet első jelei vannak kezünkben. Amíg – mint korábban említettem – mérsékelt nagyságú mesterséges automatákat értékeltünk, durva empirikus módon ilyen elmélet nélkül is elboldogulhattunk. Minden okunk megvan arra a feltételezésre, hogy fejlettebb automaták esetében ez már nem fog menni.

Hogyan befolyásolja a hibakezelési eljárásokat az a körülmény, hogy még kidolgozatlan az automaták logikai elmélete? Ez végül az utolsó, rendkívül fontos korlátozó tényező. Nem valószínű, hogy a mostaniaknál lényegesen bonyolultabb automatákat tudunk majd építeni az automaták és információk magasan fejlett, finom elmélete nélkül. Még kevésbé hihető ez olyan hallatlanul komplex automatáról, mint az ember központi idegrendszere.

Ez az intellektuális elégtelenség akadályoz bennünket abban, hogy sokkal messzebb jussunk, mint ahol most vagyunk.

Ennek a tényezőnek egyszerű megnyilvánulása jelenlegi viszonyunk a hibák ellenőrzéséhez. Élő szervezetekben is előfordul alkatrészek hibás működése. A szervezetnek nyilván megvan a módja a hibák felfedezésére és következményeik kiküszöbölésére. Könnyű felbecsülni, hogy a normális élettartam folyamán az idegsejtek működésének száma 10^{20} nagyságrendű kell legyen. Szemmel láthatóan az események ilyen lánc folyamán sohasem következik be olyan hibás működés, amelyet a szervezet maga ne tudna kijavítani, jelentősebb külső beavatkozás nélkül. Ezért a rendszer tartalmazza azt a szükséges elrendezést, amely a hibákat felismeri, amint bekövetkeznek, újra szabályozza a szervezetet, úgyhogy a tévedés hatását minimalizálja, és végül is kijavítja, vagy állandóan kirekeszti a hibás alkatrészt. A mi eljárásunk mesterséges automatáinkban fellépő hibás működéssel kapcsolatban teljesen más. A jelenlegi gyakorlati eljárás, amelyben egyetért a terület minden szakértője, valahogy így fest: Nagy erőfeszítéseket teszünk, hogy amint fellép, felfedezzük a hibát (matematikai vagy automatikus ellenőrzéssel). Ezután megkíséreljük elszigetelni a hibát okozó alkatrészt, amilyen gyorsan csak lehet. Ezt részben automatikusan lehet csinálni, de mindenesetre ennek a diagnózisnak jelentős része kívülről jövő beavatkozással történik. Mihelyt azonosítottuk a hibás alkatrészt, tüstént kijavítjuk vagy kicseréljük.

Figyeljük meg a két eljárás különbözőségét. A természetben a hibák elintézésének alapvető elve, hogy a hiba hatását, amennyire csak lehet, lényegtelené tegye, és javítás csak módjával kerül alkalmazásra, ha egyáltalán szükséges. Mi viszont mesterséges automatáinknál azonnali diagnózist követelünk meg. Ezért úgy igyekszünk berendezni az automatákat, hogy a tévedések minél szembetűnőbbé váljanak, és a beavatkozás és javítás közvetlenül kövesse a hibát. Más szóval a természetes szervezetek úgy épülnek föl, hogy a hibák minél észrevehetőbbek és ártalmatlanabbak legyenek. A mesterséges automatákat úgy tervezik, hogy a hibák minél feltűnőbbek, minél katasztrofálisabbak legyenek. Ennek a különbségnek az ésszerűséget nem kell messze keresni. A természetes organizmusok elég jól szervezetek ahhoz, hogy még akkor is működjenek, ha hiba lép fel bennük. Működnek a hibák ellenére is, és következő törekvésük az, hogy kiküszöböljék a hibás működést. Egy mesterséges automatát kétségtelenül lehet úgy megszerkeszteni, hogy korlátozott számú hiba ellenére bizonyos korlátozott területeken még normálisan működjenek. De minden hiba azzal a jelentékeny kockázattal jár, hogy valami általános degenerálódási folyamat indul meg a gépben. Ezért szükséges az azonnali beavatkozás, mert az a gép, amely már egyszer hibásan kezdett működni, ritkán javítja önmaga ki a hibát, sőt valószínűleg egyre rosszabbá válik. Mindez egy dologra vezethető vissza. Mesterséges automatáinkkal sokkal inkább sötétben

tapogatózunk, mint ahogy szemmel láthatólag a természet teszi a szervezetekkel. Egy elszigetelt hiba és az ebből következő hibás működés nagymértékben „rémületbe ejt” bennünket, és úgy látszik, hogy – legalábbis most – így is kell lenni. Magatartásunk a tudatlanságból eredő túlzott óvatosság.

Az egyetlen hiba elve. Az eljárás egy kevésbé fontos oka az, hogy úgyszólván minden hibafelismerő technikánk azon a feltevésen alapszik, hogy a gépben csak egy hibás alkatrész van. Ebben az esetben a gép ismételt részekre osztása lehetővé teszi, hogy megállapítsuk, melyik részben van a hiba. Mihelyt fennáll az a lehetőség, hogy a gépben több hiba is lehet, ezek a nagyon eredményes, felosztó diagnosztikai módszerek nem érnek semmit. A hiba megállapítás ettől a ponttól kezdve egyre reménytelenebb feladattá válik. Nagy eredménynek tekintjük, ha a feltárandó hibák számát egyre vagy mindenesetre a lehető legkisebbre szorítjuk le: ez ismét csak ebben a tárgykörben való tudatlanságunkat illusztrálja, és egyik fő oka annak, miért kell a hibákat a lehető legfeltűnőbbé tenni, hogy felismerhessük őket, és miért kell lehetőleg fellépésük után azonnal kiküszöbölni őket, mielőtt további hibák fejlődnének belőlük.

DIGITALIZÁLÁSI ELVEK

Folytonos mennyiségek digitalizálása: a digitális kifejtés, illetve a leszámlálás módszere. Tekintsük egy természetes szervezet digitális részét; speciálisan, tekintsük az idegrendszert. Szemmel láthatóan igazolva van az a feltevésünk, hogy ez a digitális mechanizmus üzeneteket továbbít, amelyek minden vagy semmi jelekből állnak. (Lásd a korábbi vizsgálódást a 70. oldalon.) Más szóval minden elemi jel, minden impulzus egyszerűen vagy van, vagy nincs, további árnyalatok nélkül. Ezt a tényt különlegesen jól illusztrálják azok az esetek, amikor a szóban forgó probléma ellenkező jellegű, vagyis amikor az idegrendszer tényleges feladata folytonos mennyiségek továbbítása. Jellegzetesen ilyen eset az, amikor az idegnek egy nyomás nagyságát kell jeleznie.

Tegyük fel például, hogy nyomást (tehát folytonos mennyiséget) kell továbbítani. Tudjuk, milyen „fogással” történik ez. Az ezt végző ideg csak egyes, minden vagy semmi jellegű impulzusokat továbbít. Hogyan fejezhető ki ezeknek az impulzusoknak, vagyis számjegyeknek a segítségével a nyomás folyamatosan numerikus értéke? Más szóval: hogyan kódolja át az ideg a folytonos mennyiséget digitális jelzésre? Bizonyos, hogy nem úgy, hogy a szóban forgó számot a szokásos értelemben, decimális (vagy bináris vagy bármely más alapú) számjegyekké kódolja át. Láthatóan az történik, hogy az impulzusokat olyan frekvenciával viszi át, amely bizonyos határok közt változhat és arányos a szóban forgó folytonos mennyiséggel, általában annak monoton függvénye. Az a mechanizmus tehát, amely ezt az „átkódolást” végzi, lényegében véve frekvencia-modulációs rendszer.

A részleteket ismerjük. Az idegnek véges feléledési ideje van. Más szóval, miután egyszer már vezetett, véges időnek kell eltelnie addig, amíg ismét vezetni tud, és ez az idő a következő (megkísérelt) inger erősségétől függ. Így, ha az ideg folytonos inger hatása alatt áll (amely szüntelenül jelen van, mint a nyomás, amelyről itt szó van), az ideg periódikusan reagál; a két egymás után következő inger közti periódus, az előbb említett feléledési idő az állandó inger erősségének a függvénye (jelen esetben a nyomásé). Így, nagy nyomás alatt az ideg esetleg minden 8 milliszekundumban tud válaszolni, vagyis másodpercenként 125 impulzust tud továbbítani; míg kisebb nyomás hatása alatt esetleg csak minden 14 milliszekundumban tud ismételni, azaz 71-szer tud vezetni másodpercenként. Világos, hogy egy sajátosan igen vagy nem szerv viselkedik így, egy digitális szerv. De nagyon tanulságos, hogy a „leszámlálás” és nem a „decimális kifejtés” (vagy „bináris kifejtés” stb.) módszerét alkalmazza.

A két módszer összehasonlítása. Az élő szervezetek a leszámhlálási módszert részesítik előnyben. Hasonlítsuk össze ennek a két módszernek az előnyeit és hátrányait. A leszámhlálási módszer kétségtelenül nem olyan hatékony, mint a kifejtés módszere. Ha egy millió körüli számot kell kifejezni leszámhlálással (vagyis egy olyan fizikai mennyiséget, amelynek millió megkülönböztethető fokozata van), millió impulzust kell átvinni. Ugyanilyen nagyságrendű szám kifejezéséhez a kifejtés módszerével 6 vagy 7 decimális számjegy kell, vagyis körülbelül 20 bináris számjegy. Ezért ebben az esetben csak 20 impulzusra van szükség. Ezért a kifejtés módszere sokkal gazdaságosabb a jelölésben, mint a leszámhlálási módszer, amelyhez a természet folyamodik. Másrészt a leszámhlálási módszer rendkívül stabil és biztonságos a hibákkal szemben. Ha egy milliós nagyságrendű számot leszámhlálással fejezünk ki és elvétünk egy számjegyet, az eredmény csak lényegtelenül változik. De ha (decimális vagy bináris) kifejtéssel fejezzük ki, egyetlen számjegyben tévedés meghamisíthatja az egész eredményt. Így számológépeink nemkívánatos tulajdonságai jelentkeznek a digitális kifejtés módszerében; valóban a két dolog szorosan összefügg, és részben egyik következik a másiktól. Másrészt a természetes szervezetek nagymértékű stabilitása és majdnem hibamentes jellege tükröződik a leszámhlálási módszerben, amely ebben az esetben láthatóan alkalmazást nyer. Mindez egy általános szabályt tükröz; növelhetjük a hibabiztonságot azáltal, hogy csökkentjük a jelölés hatékonyságát, vagy – pozitíven kifejezve – azáltal, hogy megengedjük a jelölés redundanciáját.¹⁴ Nyilvánvaló, hogy a biztonság redundancia útján való növelésének legegyszerűbb módja az, hogy az önmagában véve meglehetősen bizonytalan digitális kifejezést használjuk, de az egyes üzeneteket elég sokszor megismételjük. A tárgyalt esetben a természet nyilvánvalóan egy még redundánsabb és ennek következtében még biztonságosabb módszerhez folyamodott.

Természetesen, valószínűleg egyéb okai is vannak, amiért az idegrendszer a leszámhlálási és nem a digitális kifejtési módszert alkalmazza. Azok a kódoló és dekódoló eszközök, amelyekre az első módszerhez szükség van, sokkal egyszerűbbek, mint amelyekre a második módszernél lenne szükség. Az is igaz azonban, hogy a természet a bonyolultság irányában láthatóan sokkal messzebb hajlandó és képes is elmenni, mint mi; pontosabban – mint ameddig mi magunknak megengedhetjük. Felmerül ennek következtében a gyanú, hogy ha a digitális kifejezési rendszernek az egyetlen hátránya csak a nagyobb logikai bonyolultság, a természet ennek az egy oknak a következtében nem vetette volna el. Ennek ellenére tény, hogy a természetes szervezetekben sehol sem találjuk jelét annak, hogy alkalmazást nyer. Nehéz megmondani, hogy az ilyen megfigyelésnek mennyi „végső” érvényességet kell tulajdonítani. A szempont azonban mindenképpen figyelmet érdemel, elsősorban az idegrendszer működésének eljövendő vizsgálatánál.

FORMÁLIS NEURÁLIS HÁLÓZATOK

A formális neurális hálózat McCulloch-Pitts-féle elmélete.

Erről a tárgyról logikai és organizatorikus szempontból még sok mindent lehetne mondani; ez azonban túl messzire vezetne. Ehelyett az eddig elért, minden bizonnyal legjelentősebb eredményt fogom axiomatikus úton tárgyalni. A McCulloch-Pitts-féle figyelemre méltó tételekre gondolok, amelyek a logikának és neurális hálózatoknak a kapcsolatára vonatkoznak.

Ennél a tárgyalásnál, mint mondtam, szigorúan axiomatikus szempontok szerint járok el. Ezért a neuront úgy tekintem, mint egy „fekete dobozt”, meghatározott számú bemenő vezetékkel, amelyeken ingerelhető, és egy kimenő vezetékkel, amelyen a kimenő impulzusok

¹⁴ Redundancia: terjengősség. Információelméleti fogalom, mely azt jelenti, hogy valamit nem a lehető legrövidebb formában fejezünk ki.

jelennek meg. Részletesebben; fel fogom tételezni, hogy a bemenő vezetékek kétféle típusúak lehetnek, ingerlő és gátló jellegűek. Maguk a „fekete dobozok” szintén két típusúak lehetnek; egy-, illetve kétnívójú küszöbvel rendelkezhetnek. Ezeket a fogalmakat a következő definíciók kapcsolják össze, illetve írják körül: az ilyen szerv ingerléséhez az szükséges, hogy az ingerlő bemenetein egyidejűleg legalább annyi ingert kapjon, amennyi a küszöbértéknek megfelel, viszont a gátló bemeneteken egyetlen ingert sem. Ha a szervet ily módon ingerelték, meghatározott idő múlva (amelyről feltételezzük, hogy mindig ugyanakkora, és amelyet az időegység definíciójára is felhasználunk) egyetlen kimenő impulzust bocsát ki. Ezt a kimenő impulzust megfelelő összeköttetések révén más neuronok akárhány bemenő vezetékére (esetleg bármelyik saját bemenő vezetékére is) kapcsolhatjuk; ebben az esetben mindegyik ugyanolyan bemenő ingert jelent, mint amit fentebb leírtunk.

Mindez természetesen a neuron valóságos működésének nagyfokú leegyszerűsítése. Korábban már tárgyaltam az axiomatikus módszer jellegzetességeit, korlátait és előnyeit. Valamennyi megjegyzés itt is érvényes, és a következőkben erre figyelemmel kell lennünk.

McCulloch és Pitts ezeket az egységeket arra használták fel, hogy kombinált hálózatot építsenek fel, amelyeket „formális neurális hálózat”-nak nevezhetünk. Egy ilyen rendszer tetszőleges számú egységből épül fel úgy, hogy a bemeneteket és a kimeneteket alkalmas módon, tetszőleges bonyolultságban összekapcsolták. A hálózat „működése” úgy definiálható, hogy a bemenő és kimenő vezetékek közül néhányat kiválasztunk, és leírjuk, hogy a bemenő vezetékekre adott ingerlő impulzusok a kimenő vezetékeken milyen végső impulzusokat hoznak létre.

A McCulloch-Pitts-féle elmélet fő eredménye.

McCulloch és Pitts elméletének fontos eredménye az, hogy a fenti értelemben vett bármely olyan működés, amelyet logikailag szigorúan és egyértelműen véges számú „szó” segítségével egyáltalán definiálhatunk, ilyen formális neurális hálózattal megvalósítható is.

Célszerű ennél a pontnál kissé megállni, és megfontolni, hogy mik a következmények. Gyakran állítjuk azt, hogy az emberi idegrendszer tevékenysége és funkciói olyan bonyolultak, hogy azokat semmiféle szokásos mechanizmus nem tudja elvégezni. Azt is megkísérelték, hogy olyan specifikus funkciókat nevezzenek meg, amelyek ezt a korlátozást természetüknél fogva mutatják. Megkísérelték bebizonyítani, hogy ilyen specifikus funkciók, noha logikailag teljes mértékben leírhatók, önmagukban nem alkalmasak arra, hogy mechanikusan, neurálisan realizálhatók legyenek. A McCulloch-Pitts-féle eredmény ennek véget vet. Bebizonyítja, hogy minden, amit kimerítően és egyértelműen szavakba lehet foglalni – alkalmas véges neurális hálózattal ipso facto realizálható is. Minthogy az állítás megfordítása nyilvánvaló, állíthatjuk, hogy bármely reális vagy elképzelt, teljesen és egyértelműen szavakba foglalható viselkedési mód leírásának a lehetősége és ugyanennek a véges formális neurális hálózattal való megvalósításának a lehetősége között nincs különbség. A két fogalom kiterjedése egyenlő. Bármely típusú viselkedésnek ilyen hálózattal történő megvalósításánál csak akkor merül fel elvi nehézség, ha a viselkedést nem tudjuk teljes mértékben leírni. Így a következő két probléma marad: először, ha egy bizonyos viselkedési módot véges neurális hálózattal elvileg realizálni lehet, még mindig fennáll az a kérdés, hogy a hálózat a gyakorlatban elfogadható méretben is realizálható-e, speciálisan: elfér-e a szóban forgó szervezet fizikai határai között. Másodszor: felmerül a kérdés, vajon ténylegesen lehet-e minden létező viselkedési módot teljes mértékben és egyértelműen szavakba foglalni.

Az első probléma természetesen az idegphysiologia végső kérdése, és meg sem kíséreltem, hogy vele itt mélyebben foglalkozzam. A második kérdés ettől eltérő természetű, és érdekes logikai vonatkozásai vannak.

A fenti eredmény elemzése.

Semmi kétség, hogy minden elképzelt viselkedési formának bármely fázisát szavakkal „teljesen és egyértelműen” le lehet írni. Ez a leírás esetleg hosszadalmas, de mindig

lehetséges. Ezt tagadni annyit jelent, mint ragaszkodni a logikai miszticizmus egy formájához, ami bizonyára legtöbbször távol áll. Mégis fontos korlátozás, hogy az elmondottak külön minden egyes elemre vonatkoznak, de távolról sem világos, hogyan vonatkozik az egész viselkedésre együttvéve. Pontosabban kifejezve; nem merül fel nehézség annak leírásában, hogyan tud egy szervezet például két derékszögű háromszöget, amelyek a retinán megjelennek, úgy azonosítani, hogy mindkettő a „háromszög” kategóriájába tartozik. Ezen túlmenően abban sincs nehézség, hogy a szabályszerűen rajzolt derékszögű háromszögeken kívül más tárgyakat, mint háromszögeket, amelyeknek oldalai görbék; amelyeknek oldalai nincsenek teljesen kirajzolva; amelyeket csak belsejüknek többé-kevésbé homogén árnyékoltsága jelez stb. Minél teljesebben akarunk leírni mindent, ami általánosan ebbe a kategóriába esik, annál hosszabb lesz a leírás. Valószínűleg meglesz az a bizonytalan és kellemetlen érzésünk, hogy egy ilyen típusú teljes katalógus nemcsak rendkívül hosszú, hanem határait illetően elkerülhetetlenül bizonytalan is. Ennek ellenére is, a művelet lehetséges. Mindez azonban csak egy kis része a hasonló geometriai alakok azonosítása általánosabb fogalmának; ez viszont csak egy mikroszkopikus darabja az analógia általános fogalmának. Senki sem kísérelné meg, hogy bármely, gyakorlatilag elfogadható terjedelmen belül leírja és meghatározza az analógia általános fogalmát, amely a látás interpretációjában a legfontosabb dolog. Semmi alapunk nincsen ahhoz, hogy valami határozottat állítsunk, például azt, hogy egy ilyen vállalkozás ezer vagy egymillió, vagy még ennél is több, gyakorlatilag használhatatlanul sok kötetet igényelne. Mármost tökéletesen lehetséges, hogy a látási analógia meghatározásának a legegyszerűbb és az egyetlen gyakorlati útja éppen abban áll, hogy megadjuk az agyvelő vizuális központjában levő összeköttetések leírását. Ezen a területen a logikának olyan részével van dolgunk, amelyben gyakorlatilag semmi korábbi tapasztalattal nem rendelkezünk. A bonyolultságnak ez a foka túlmegy minden arányon, bármihez képest, amit eddig ismertünk. Nincs jogunk feltételezni, hogy azok a logikai fogalmak és eljárások, amelyeket eddig használtunk, tárgyunknak erre a részére is alkalmazhatók. Egyáltalán nem biztos, hogy ezen a területen nem maga a reális tárgy adja önmagának a legegyszerűbb leírását: másként szólva, bármely kísérlet, hogy a szokásos irodalmi vagy formális-logikai módszerekkel leírjuk, olyasvalamihez vezethet, ami sokkal kevésbé kezelhető és sokkal bonyolultabb. Ténylegesen a modern logikának bizonyos eredményei arra látszanak utalni, hogy ha valóban komplikált dolgokhoz érünk, ehhez hasonló jelenségeket várhatunk. Ennek következtében egyáltalán nem valószínű, hogy reménytelen a vizuális analógia pontos szóbeli leírására pontos logikai fogalmat keresni. Lehetséges, hogy ennek az elvnek a legegyszerűbb logikai kifejezése vagy definíciója magának a vizuális centrumnak a kapcsolási hálózata.

Nyilvánvaló, hogy ezen a szinten a McCulloch-Pitts-féle eredményből már nincs hasznunk. Az eredmény ezen a ponton már csak egy további illusztrációt ad ahhoz a helyzethez, amelyet az előbb már körvonalaztunk. A logikai elvek és a neurális hálózat alakjában történő megvalósítás között ekvivalencia áll fenn; míg egyszerűbb esetekben a hálózat egyszerűsített kifejezéseit ezek az elvek adhatják, teljes mértékben lehetséges, hogy a szélsőségesen bonyolult esetekben az ellenkező igaz.

Mindez nem változtatja meg azt a meggyőződésünket, hogy egy új, lényegileg logikai elméletre van szükség ahhoz, hogy a nagymértékben komplikált automatákat és különösen a központi idegrendszert megértsük. Lehetséges azonban, hogy ennek folyamán a logika kénytelen lesz látszólag a neurológia irányába tolni, sokkal inkább mint az ellenkező irányba. Az előző tárgyalás azt mutatja, hogy egyike azoknak a lényeges dolgoknak, amiket a központi idegrendszer elméletét illetően pillanatnyilag tehetünk az, hogy megállapítjuk, milyen irányok azok, amelyekben az igazi probléma nem keresendő.

A KOMPLIKÁLTSÁG FOGALMA; ÖNREPRODUKÁLÁS

A komplikáltság fogalma.

Az eddigi tárgyalás megmutatta, hogy az automatákra vonatkozó bármely elméleti erőfeszítésnél fontos szerepe van a nagy komplexitásnak, és hogy ez a fogalom – noha első látásra kvantitatívnak látszik, valójában mégis valahogy kvalitatív jellegű – egy elv szerepét tölti be. A tárgyalás további részében ennek a fogalomnak egy távolabbi következményét fogom vizsgálni, azt, amely a fogalom természetének kvalitatív vonásait még jobban kiemeli. Van a természetnek egy circulus vitiosus típusú, élesen nyilvánvaló vonása; legegyszerűbb kifejezése az a tény, hogy nagyon bonyolult organizmusok képesek önmagukat reprodukálni. Mindnyájan hajlunk – bizonytalanul – annak feltételezésére, hogy a „komplikáltság” fogalma létezik. Ezt a fogalmat és vélt tulajdonságait sohasem fogalmazták meg világosan. Ha egy automata bizonyos műveleteket végez, azt kell elvárunk, hogy ezek eredménye alacsonyabb komplikáltságú fokon lesz, mint maga az automata. Különösen pedig, hogyha egy automata arra képes, hogy egy másikat építsen, a komplikáltságnak csökkenni kell, ahogy a szülőtől az utódig megyünk. Azaz: ha A elő tudja állítani B -t, akkor A -nak valamilyen módon tartalmaznia kellett B teljes leírását. Hogy feladatát be tudja tölteni, A -ban szükség van továbbá különböző elrendezésekre, amelyek gondoskodnak ennek a leírásnak az értelmezéséről, valamint arról, hogy a mindenkori szükséges konstruktív műveletek ténylegesen megtörténjenek. Ebben az értelemben úgy látszik tehát, hogy amikor egy automata egy másik automatát csinál, bizonyos degeneratív tendencia, a komplexitás csökkenése várható.

Noha van ebben valami meg nem határozott plauzibilitás, a természetben végbemenő legfőbb folyamattal nyilvánvaló ellentmondásban van. A szervezetek reprodukálják önmagukat, azaz olyan új szervezeteket produkálnak, amelyeknek komplexitása nem csökken. Ehhez járul még, hogy hosszú fejlődési periódusok léteznek, amelyek során a komplexitás még nő is. Az egyes organizmusok közvetve olyan más organizmusoktól származnak, amelyeknek komplexitása kisebb volt.

Ilyen módon a valószínűség és a nyilvánvalóság között legalábbis nyílt konfliktus áll fenn, ha ugyan nem rosszabb. Erre való tekintettel érdemes megvizsgálni, hogy van-e itt olyasmiről is szó, amit szigorúan meg lehet fogalmazni. Mindaddig kissé bizonytalan és homályos voltam, nem is egészen szándék nélkül. Nem hiszem, hogy másképpen helyes képet lehetne adni az itt jelentkező helyzetről. Legyen szabad most erről részletesebben szólni.

A számoló automaták Turing-féle elmélete.

Turing angol logicista körülbelül tizenkét évvel ezelőtt a következő problémát vizsgálta. Általános definícióját kívánta adni annak, hogy mit értünk számoló automatán. A formális definíció a következő:

Az automata olyan „fekete doboz”, amelynek részleteit pontosan nem ismerjük, de a következő tulajdonságokkal rendelkezik. Véges számú állapota lehetséges, amelyeket első látásra elegendő, ha egy sorszám, például n megadásával jellemzünk és megfelelő módon felsorolunk: $1, 2, \dots, n$.

Az automata lényeges működési jellegzetessége abban áll, hogy leírjuk, hogyan készítjük állapotának megváltoztatására, tehát arra, hogy az i állapotból j állapotba menjen át. Ez az átmenet a külvilággal való valamilyen kölcsönhatást igényel, amelyet a következő módon lehet „szabványosítani”. Ami a gépet illeti, az egész külvilágot egy hosszú papírszalag képviselheti. Legyen ez a szalag például 1 centiméter széles és tételezzük fel, hogy 1 cm hosszú mezőkre (négyszögekre) van osztva. A szalag minden egyes mezejébe teszünk vagy nem teszünk jelet, például egy pontot, és feltételezzük, hogy a pontot nemcsak beírni, hanem törölni is lehet. Azt a mezőt, amelyet egy ponttal jelöltünk, 1-nek nevezzük; ha a mező nem

tartalmaz pontot, 0-nak fogjuk nevezni. (Megengedhetünk többféle jelölési módot is, de Turing megmutatta, hogy ez lényegtelen és az általánosság tekintetében semmi lényeges előnyhöz nem vezet.) A szalagnak az automatahoz viszonyított helyzetét vizsgálva feltételezzük, hogy az automata közvetlenül a szalagnak egy meghatározott mezejét vizsgálja, és alkalmas arra, hogy a szalagot előre vagy hátra, egyszerre például egy mezővel elmozgassa. Hogy ezt specifikáljuk, legyen az automata például az i ($=1\dots n$) állapotban és észleljen a szalagon egy e -t ($=0,1$). Erre át fog menni a j ($=0,1\dots n$) állapotba, p mezővel előremozgatja a szalagot [$p=0,+1,-1$] ($a+1$ egy lépéssel előre, $a-1$ egy lépéssel hátra történő mozgást jelent), az eléje kerülő új mezőbe pedig beír egy f -et ($=0,1;0$ beírása törlést, 1 pedig egy pont beírását jelenti). Az ilyen automata működésének teljes definíciója az, ha j -t, p -t és f -et megadjuk, mint i és e függvényét. Turing gondosan elemezte, hogy az ilyen típusú automaták milyen matematikai műveleteket tudnak elvégezni. Ebben az összefüggésben különböző tételeket bizonyított be a logika klasszikus „eldöntés” problémájával kapcsolatosan, de ezekre itt nem térek ki. Emellett azonban bevezette és elemezte az „univerzális automata” gondolatát. Ez az a része a tárgynak, amely a jelen összefüggésben lényeges.

Az $e(=0,1)$ jegyekből álló végtelen sorozatok a matematika alapvető fogalmaihoz tartoznak. Ha úgy tekintjük őket, mint egy bináris kifejtést, lényegében véve ekvivalens a reális szám fogalmával. Ennek következtében Turing meggondolásaiban ilyen sorozatokkal foglalkozott. Megvizsgálta azt a kérdést, hogy az ilyen típusú automata milyen sorozatokat tud konstruálni. Azaz: ha adva van egy ilyen sorozat meghatározott képzési szabálya, azt kérdezte, hogy milyen automatákat lehet e szabályon alapuló sorozatok képzésére felhasználni. A „sorozatképzés” fogalmán a következőket kell érteni. Egy automata akkor képes arra, hogy egy meghatározott sorozatot „képezzen”, ha meg lehet adni egy alkalmas módon megjelölt, véges hosszúságú szalagrészt úgy, hogy ha a szalagot betápláljuk a szóban forgó automatába, az automata ennek alapján a sorozatot a szalagnak a megmaradó (végtelen hosszúságú) folyamata természetesen határozatlan ideig tart. Itt a következőkről van szó: az automata határozatlan ideig fog működni és ha elég hosszú ideig várunk, a (végtelen) sorozat tetszőleges (de természetesen véges) részét le fogja írni. A szalagnak a véges, előre jelzett része az automatáknak a problémára vonatkozó „utasítása”.

Egy automata akkor „univerzális”, ha bármely olyan sorozat, amelyet egyáltalán automata elő tud állítani, előállítható a szóban forgó automata útján is. E célból természetesen általában egy, a többitől eltérő utasításra van szüksége.

A Turing-elmélet fő eredménye.

A priori azt várhatnánk, hogy ez lehetetlen: hogyan lehetséges egy olyan automata, amely legalább olyan hatékony, mint bármely más elképzelhető automata, beleértve például olyant is, amely kétszer olyan nagyméretű és komplexitású, mint ő maga?

Turing azonban bebizonyította, hogy ez lehetséges. Noha konstrukciója rendkívül bonyolult, az alapul szolgáló elv mégis egészen egyszerű. Turing megfigyelte, hogy véges számú szó segítségével meg lehet adni bármely elképzelhető automata teljesen általános leírását (az előbbi definíciónak megfelelően). Ez a leírás meghatározott üres szakaszokat fog tartalmazni – azokat, amelyek a korábban említett függvényekre vonatkoznak (j , p , f , mint i -nek és e -nek a függvénye) és amelyek az automata tényleges működését specifikálják. Ha ezeket az üres szakaszokat kitöltjük, egy specifikus automatával van dolgunk; ha azonban üresen maradnak, a szkéma az általános automata általános definícióját jelenti. Így lehetségessé válik leírni egy olyan automatát, amely az ilyen definíciót interpretálni tudja. Más szavakkal kifejezve: az automata, ha betápláljuk azokat az utasításokat, amelyek a fent leírt értelemben a specifikus automatát meghatározzák, a továbbiakban úgy viselkedik, mint a leírt objektum. A képesség, hogy ez megtörténjék, semmivel sem rejtélyesebb, mint az, hogy szótárt és nyelvtant

használunk és a szavak jelentését és kombinációs elveiket illetően követjük az utasításokat. Az automata, amelyet arra konstruáltak, hogy leolvasson egy leírást és utánozza a leírt objektumot, megfelel a Turing értelmében vett univerzális automatának. Ahhoz, hogy reprodukáljon bármely olyan műveletet, amelyet bármely más automata el tud végezni, elegendő, ha ellátják a szóban forgó automata leírásával, továbbá utasításokkal, hogy hogyan végezze el a figyelembe jövő műveleteket.

A program kiterjesztése olyan automatákra, amelyek automatákat állítanak elő.

Ahhoz a problémához, amivel én itt foglalkozom, ti. az automaták önreprodukciójához, Turing eljárása csak egy vonatkozásban nem elég általános. Turing automatái tisztán számológépek. Termékük egy papírszalag, amelyeken 0-ák és 1-esek vannak. Ahhoz a konstrukcióhoz, amelyre utaltam, olyan automata szükséges, amelynek végterméke egy másik automata. Semmi elvi nehézsége nincs azonban annak, hogy ezt az általánosabb gondolatot tárgyaljuk és Turing eredményével ekvivalens eredményt vezessünk le.

Az alapvető definíciók.

Mint az első esetben, most is elsőrendű fontosságú, hogy szigorú definícióját adjuk annak, ami a vizsgálat szempontjából automatának tekinthető.

Mindenekelőtt meg kell adni mindazon elemi részek teljes listáját, amelyeket fel lehet használni. Ennek a listának nemcsak teljes felsorolást, hanem minden egyes elemi alkatrész teljes működési definícióját is tartalmaznia kell. Viszonylag könnyű dolog ilyen listának a felfektetése, azaz olyan gépalkatrészek katalógusának a leírása, amely elegendő ahhoz, hogy az itt szükséges mechanizmusok sokféle változatát megkonstruálhassuk és amely az ilyen természetű megfontolásokhoz szükséges axiomatikus szigorúsággal rendelkezik. A listának nem is kell túlságosan hosszúnak lenni. Tulajdonképpen akár tetszőlegesen hosszú, akár tetszőlegesen rövid lehet. Meghosszabbítható azáltal, hogy elemi alkatrészekként olyanokat is felvesszünk, amelyek mások kombinációjából is előállíthatók. Meg is rövidíthető – ténylegesen egyetlen egységre csökkenthető –, ha minden elemi alkatrészt sokféle tulajdonsággal és funkcióval ruházunk fel. Az elemi alkatrészek számára vonatkozó minden állítás tehát végső fokon a józan ész diktálta kompromisszumot jelent, amelyben egyetlen elemi alkatrészből sem várunk el semmi túlságosan komplikált dolgot, viszont egyetlen elemi alkatrész sem végez el sokféle, nyilvánvalóan különböző funkciót. Ebben az értelemben bebizonyítható, hogy körülbelül egy tucat elemi alkatrész elegendő. Az önreprodukálás problémáját ezek után körülbelül így fogalmazhatjuk meg; lehetséges-e ilyen elemekből olyan szerkezetet építeni, amely, ha betesszük egy nagy tartályba, amelyben ugyanezek az alapelemek elegendő nagy számban megvannak, elkezd más szerkezeteket konstruálni, amelyek mindegyike a végén pontosan ugyanolyan automata lesz, mint az eredeti? Ez lehetséges; az elv, amelyen alapul, szorosan kapcsolódik a fent körvonalazott Turing-féle elvhez.

Az önreprodukálási tétel levezetésének körvonalai.

Mindenekelőtt részletes leírást lehet adni mindarról, amiből az itt tárgyalt értelemben vett automata áll. Ezt a leírást eléggé általánosnak kell elképzelni, tehát üres helyeket fog tartalmazni. Ezeket az üres helyeket azokkal a függvényekkel kell kitölteni, amelyek az automata tényleges struktúráját leírják. Mint az első esetben, a kitöltött és kitöltetlen szakaszok közötti különbség felel meg a specifikus automata leírása és az általános automata leírása közötti különbségnek.

Nincs elvi nehézség abban, hogy leírjuk a következő automatát:

- a) Az A automata olyan, hogyha ellátjuk bármely más automatának a megfelelő függvényekben kifejtett leírásával, ezt az utóbbi automatát meg fogja konstruálni. A

leírást ebben az esetben nem lyukasztott papírszalag formájában kell megadni, mint Turing tette, minthogy rendes körülmények között a papírszalagot nem fogjuk szerkezeti elemnek választani. Meglehetősen könnyű azonban leírni a szerkezeti elemek olyan kombinációját, amelyek a jelölhető mezőkkel ellátott papírszalag tulajdonságával rendelkeznek. Az ilyen értelmű leírást utasításnak fogjuk nevezni és I betűvel jelöljük.

A „konstruálást” ugyanolyan értelemben kell érteni, mint az imént. Feltételezzük, hogy a konstruáló automata egy nagy tartályban van elhelyezve, amelyben minden szükséges elemi alkatrész nagy számban jelen van; a tényleges konstruálást ebben a környezetben végzi el. Nem szükséges azon törni a fejünket, hogyan tud egy ilyen típusú rögzített automata olyan más automatákat létrehozni, amelyek nagyobbak és bonyolultabbak, mint ő maga. Ebben az esetben a konstruálandó tárgy nagyobb méretei és nagyobb bonyolultsága előreláthatóan abban tükröződik, hogy a I utasítások, amelyekkel az automatákat el kell látni, még nagyobbak. Ezek az utasítások – mint rámutattunk – tulajdonképpen az elemi részek halmazai kell, hogy legyenek. Ebben az értelemben szerepel a folyamatban egy olyan egység, amelynek méreteit és bonyolultságát a tervezendő objektum méretei és bonyolultsága határozzák meg. A továbbiakban minden automata, amelynek tervezésére A tulajdonságait használjuk fel, ezekben a tulajdonságokban A-val osztozni fog. Mindegyikben hely lesz az I utasítás számára, tehát egy olyan hely, ahová az utasítást be lehet táplálni. Ha egy ilyen automatát leírunk (például azáltal, hogy megadjuk a megfelelő utasítást), feltételezzük, hogy annak a helynek a megadása, ahová az előbbi értelemben értett I utasítást be kell táplálni, magának az utasításnak egy része. Ennek következtében minden további magyarázat nélkül beszélhetünk arról, hogy „egy adott utasítást betáplálunk egy adott automatába”.

- b) A B automata olyan, amely el tudja készíteni a betáplált I utasítás másolatát. Az I utasítás elemi alkatrészeknek összessége abban az értelemben, ahogy azt a)-ban körvonalaztuk és amely a papírszalagot helyettesíti. Ezt a lehetőséget akkor használjuk, ha I egy másik automata leírását szolgáltatja. Más szavakkal: ez az automata semmivel sem bonyolultabb, mint egy „reproduktor” – például az a gép, amely leolvasson egy lyukasztott szalagot, és előállt egy másik, az előzővel azonos lyukasztott szalagot. Megjegyezzük, hogy ez az automata is tud olyan tárgyakat termelni, amelyek nagyobbak és komplikáltabbak, mint ő maga. Megjegyezzük azt is, hogy mindebben semmi meglepő sincs. Minthogy mindössze másolni tud, pontosan ugyanolyan méretű és komplexitású bemenettel kell ellátni, mint a végtermék. Ezután az előkészítés után áttérhetünk a döntő lépésre.
- c) Kombináljuk az A és B automatákat egymással és egy C vezérlő-mechanizmussal. Lássuk el A-t egy I utasítással (megint az a), illetve b) értelemben). Ekkor C először A-val is megkonstruáltatja azt az automatát, amelyet a betáplált I utasítás leír. Ezután a C a B automatával lemásoltatja a fent említett I utasítást, és betáplálja abba az automatába, amelyet A automata az imént konstruált. Végül C elválasztja ezt a konstrukciót az A+B+C rendszertől, és szabadjára bocsátja, mint független egységet.
- d) Jelöljük a teljes A+B+C összességet D-vel.
- e) Hogy működni tudjon, a $D=A+B+C$ összességet el kell látni a fent leírt utasítással. Mint fentebb kimutattuk, ezt az utasítást A-ba kell táplálni. Mármost alkossunk egy I_D utasítást, amely leírja a D automatát, és tápláljuk be az I_D utasítást a D automatán belül az A automatába. Az így kapott automatát nevezzük E-nek.

Világos, hogy E önreprodukáló lesz. Figyeljük meg, hogy nem forog fenn circulus vitiosus. A döntő lépés E-ben történik akkor, amikor a D automatát leíró I_D utasítást konstruáljuk és betápláljuk D-be. Amikor az I_D konstruálására (másolására) van

szükség, D már létezik és az I_D konstruálása egyáltalán nem módosítja. I_D -t egyszerűen hozzáadjuk, hogy E-t képezzük. Így tehát létezik egy meghatározott időbeli és logikai sorrend, amely szerint D-t és I_D -t ki kell alakítani; az egész folyamat a logika törvényei szerint szabályos és helyénvaló.

Az eredmény és közvetlen kiterjesztésének értelmezése.

Az E automata fenti leírásának további érdekes oldalai vannak, amelyekre azonban ez alkalommal nem térhetek ki részletesen. Így például teljesen világos, hogy az I_D utasítás nagyjából ugyanazt a szerepet tölti be, mint a gén. Az is nyilvánvaló, hogy a B másolómechanizmus a reprodukciónak, a genetikai anyag duplikálásának alapvető tevékenységét végzi el, ami nyilvánvalóan az élő sejtek szaporodásának az alapvető művelete. Azt is könnyű belátni, hogy az E rendszer, különösen pedig az I_D utasítás önkényes megváltoztatásai bizonyos, a mutációval kapcsolatos jellemző vonásokat mutathatnak; ezek a mutációk általában nem életképesek, de megvan annak a lehetősége is, hogy a reprodukció változott jellegzetességekkel folytatódjék. Természetesen az is ugyanilyen világos, hogy milyen pontokon szűnik meg az analógia érvényessége. A természetes gén valószínűleg nem tartalmazza annak a tárgynak a részletes leírását, amelynek létrejöttét éppen a jelenléte váltja ki. Valószínűleg csak általános jellegzetességeket, általános „utasításokat” tartalmaz. Abban az általánosságban, amelyben az előző megfontolások mozognak, ezt az egyszerűsítést meg sem kíséreltem. Mindazonáltal világos, hogy ez és az ehhez hasonló egyszerűsítések önmagukban is nagy, kvalitatív fontosságúak. Nagyon távol kerülünk az élő természetben lejátszódó folyamatok reális megértésétől, ha nem próbálunk meg behatolni az ilyen egyszerűsítő elvekbe.

Az előbb vázoltak kis megváltoztatása azt is lehetővé teszi, hogy olyan automatákat konstruáljunk, amelyek egyrészt önmagukat reprodukálják, másrészt ezen felül más automatákat is konstruálnak. (Az ilyen automata specifikusabban valósítja meg azt, ami valószínűleg a géneknek egyik – ha ugyan nem a legfontosabb – tipikus funkciója, t. bizonyos specifikus enzimek önreprodukcióját; ugyanakkor pedig ezek keletkezését vagy keletkezésének kiváltását.) Valóban elegendő, ha az I utasítást, egy I_{D+F} utasítással helyettesítjük, ami a D automatát, plussz egy másik adott F automatát ír le. Tápláljuk be az I_{D+F} utasítást már az A-ba, és jelöljük a D automatát E-vel. Az E automata nyilvánvalóan rendelkezik a most leírt tulajdonsággal. Reprodukálja önmagát és emellett konstruálja F-et. Figyeljük meg, hogy E_F -nek a „mutációja”, amely az E-ben szereplő I_{D+F} utasítás F-részében szerepel – életképes.

Ha ez a mutáció F-t egy F' -vel helyettesíti, az E_F automatát $E_{F'}$ automatává változtatja, tehát a „mutált változat” még mindig önreprodukáló, a melléktermék azonban F-ről F' -re változott. Ez a tipikus életképes mutáció.

Mindezek erősen kezdetleges lépések az automaták rendszeres elméletének irányában. Ezen felül mindössze egy speciális irányt jelentenek. Mint fentebb már jeleztem, ez az irány a „komplikáltság” fogalom tartalmának szigorú megfogalmazása. A kísérlet azt illusztrálja, hogy a „komplikáltság” alacsonyabb szinten valószínűleg degeneratív jellegű, tehát minden olyan automata, amely egy másik automatát tud építeni, csak önmagánál kevésbé komplikált automaták építésére képes. Van azonban egy bizonyos minimális szint, amelyen ez a degeneratív tendencia már nem általános. Ezen a szinten lehetségessé válnak olyan automaták, amelyek önmagukat reprodukálni képesek, vagy éppen magasabb rendű szervezeteket tudnak konstruálni. Az a tény, hogy a komplikáltság csak úgy, mint a szervezethez, egy bizonyos minimális szint alatt degeneratív jellegű, ezen túl azonban önmagát fenntartó vagy éppen növekvő jellegűvé válik, nyilvánvalóan a téma bármely eljövendő elméletében fontos szerepet fog játszani.

VITA

Dr. McCulloch:

Be kell vallanom, hogy semmit sem irigyelek jobban Dr. Neumantól, mint azt a tényt, hogy azokhoz a gépekhez, amelyekkel neki dolga van, kezdettől fogva rajzok állanak rendelkezésre arról, hogy a gép milyen műveleteket és hogyan hajt végre. Sajnos, a biológiai tudományokban – vagy legalábbis a pszichiátriában – ismeretlen vagy éppen ellenséges mechanizmusokkal van dolgunk. Nem tudjuk pontosan, hogy a gépnek mit kellene csinálni, és persze, nincsen rajzunk róla. A pszichiátriában felmerülő problémáknál csak azt tudjuk, hogy a gép helytelen válaszokat produkál: tudjuk, hogy a gép a saját maga által okozott hiba következtében ámokfutóvá vált. Azt azonban, hogy milyen természetű nehézségek állnak fenn magában a gépezetben, egyáltalán nem könnyű meghatározni.

Ahogy én látom, az, amire elsősorban és legfőképpen szükségünk van, nem egy korrekt elmélet, hanem egyáltalán valamilyen elmélet, amelyből kiindulhatunk és amelynek segítségével remélhetjük, hogy olyan kérdéseket tudunk feltenni, amelyekre választ kapunk – legalábbis olyan értelemben, hogy kérdésünk teljesen téves volt. Legtöbbször még olyan formában sem tudjuk feltenni a kérdést, hogy egyáltalán választ kapjunk.

Mínt hogy a tárggyal szoros kapcsolatban van, ha megengedik, elmondom, hogy történelmileg hogyan kezdtem érdeklődni ez iránt a speciális probléma iránt. Érdeklődésem eleinte főleg a filozófia és a matematika területéhez fűzött; a pszichológiához úgy jutottam el, hogy felmerült a kérdés: hogyan jöhetett létre egyáltalán olyasvalami, mint a matematika, miről is van szó? Így történt, hogy fokozatosan eltolódtam a pszichológia felé – s mínt hogy ismételt kísérletek ellenére sem sikerült a szignifikáns változókat megtalálni – belekényszerültem a neurofiziológiába. Rendkívül kemény dió az a kísérlet, hogy egy olyan területen építsünk ki elméletet, mint amiről most szó van, mégpedig úgy, hogy az valamilyen módon igazolható is legyen. Utólag elmondva, tréfásan hangzik, de 1919 körül teljesen helytelen úton indultam el, amennyiben megpróbáltam megkonstruálni a tárgyas igék logikáját. Erről kiderült, hogy a modális logikának egy eszköze. Csak amikor Turing dolgozatát láttam (1937-ben. – A ford.), kezdtem a helyes útra térni és Pitts segítségével megfogalmazni a szükséges logikai kalkulust. Amit csináltunk (és azt hiszem, eléggé sikeresen), lényegében az volt, hogy az agyvelőt úgy tekintettük, mint egy Turing-gépet, azaz mint olyan eszközt, amely olyan funkciókat tud végrehajtani, mint az agyvelő – ha meghibásodik és pszichotikussá válik. Részünkre az volt a lényeges, hogy alkalmaztunk valamilyen logikát, ugyanakkor azonban figyelembe vettük a jelek előfordulásának az időpontját is (ami, ha úgy tetszik, nem más, mint állítások időbeli sorozata). Erre azért volt szükség, hogy kielégítő elméletet tudjunk konstruálni annak megmagyarázására, hogyan tud az idegrendszer egyáltalán valamit csinálni. A kellemes az egészben az, hogy az egyébként megfelelő lehető legegyszerűbb feltevések elegendők annak kimutatására, hogy az idegrendszer minden logikailag meghatározható számot ki tud számítani: ha úgy tetszik, ugyanolyan típusú, mint a Turing-féle gép.

Az agyvelő bizonyos funkcióival kapcsolatban azonnal felmerült a kérdés, hogyan tudja az agyvelő ezeket a funkciókat ellátni. A jelenlegi elméletek semmivel sem tudják jobban megmagyarázni azt, hogy egy speciális művelet hogyan történik, mint azt, hogy milyen típusú idegrendszerben megy végbe, vagy azt, hogy egy számológép melyik részében megy végbe. Ezért meg kell legyen a számológép elvi kapcsolása vagy pedig az egymásba kapcsolódó kerek leírása.

Ez annyit jelent, hogy az anatómiát kellett tanulmányozni és olyan dolgokat kérdezni az anatómustól, amilyeneket eddig ritkán írt le elegendő részletességgel. Míg az orvosi egyetemen voltam, neuroanatómiát tanítottam; de egészen az utolsó egy-két évig nem voltam

abban a helyzetben, hogy egy neuroanatómustól valamilyen struktúra precíz részletei iránt érdeklődhettem volna. Semmiféle fiziológiai ok nem volt arra, hogy ilyen természetű információt kérjek. Most azonban már kezd szükségünk lenni rá.

Dr. Gerard:

Több különböző alkalommal volt részem abban a szerencsében, hogy dr. Neumannt hallgathattam. Ilyenkor mindig abban a kellemes, de nehéz szerepben találom magamat, mint aki egy léggömb kötelébe kapaszkodik. Noha követni tudom a gondolatmenetét, nem sok önálló gondolatot tudok produkálni az előadás kapcsán. Ennek ellenére szeretnék egy kérdést feltenni, amiről azt gyanítom, hogy a többiek fejében is megfordult. Előadásának több pontján gondosan kifejtette, hogy mindaz, ami szóbeli formában szóbeli kérdésekkel kifejezhető, realizálható is. Nincs ebben valami csapda? Mi a kérdés ilyen módon való korlátozásának a következménye?

Dr. Neumann:

Megpróbálok válaszolni, de azt hiszem, válaszom nem lesz eléggé teljes. Ha valamely problémával, különösen pedig ha a központi idegrendszer valamely funkciójával foglalkozunk, az első feladat, amely felmerül az, hogy egyértelműen megfogalmazzuk, szigorú értelemben szavakba foglaljuk. Ha egy olyan komplikált rendszerről van szó, mint a központi idegrendszer, felmerül az a további feladat is, hogy ezt a „megfogalmazást”, „szavakba öntést” értelmes határokon belül maradó szómennyiséggel hajtsuk végre – például olyan mennyiségű szóval, amit egy ember élete során el tud olvasni. Ebben rejlik a tulajdonképpeni nehézség.

Másképp szólva: azt hiszem, hogy teljesen lehetséges, hogy valaki a központi idegrendszernek kívülről látható funkcióiról emberileg lehetséges időtartamon belül tisztán deskriptív jellegű leírást adjon. Ez az időtartam lehet tíz vagy húsz év, ez ugyan hosszú idő, de ez nem akadály. Ezután a McCulloch-Pitts-féle eredmények alapján, ugyancsak plauzibilis határidőn belül meg lehet rajzolni azt a fiktív „idegrendszeri hálózatot”, amely mindezeket a funkciókat elvégzi. Azt gyanítom, azonban, hogy ez a hálózat sokkal nagyobbak fog bizonyulni, mint az, amivel ténylegesen rendelkezünk. Lehetséges, hogy kiderül, hogy sokkal nagyobb, mint az egész fiatal univerzum. És akkor? Nem tévesztettük-e az eljárás során szem elől a valódi problémát?

Ezért célszerűbb, ha a problémát nem úgy fogjuk fel, hogy a központi idegrendszer funkcióit egyáltalán valamiféle hálózattal utánozni kell, hanem úgy, hogy olyan hálózattal kell a funkciókat megvalósítani, amely az emberi agyvelő tényleges térfogatában elfér. Vagy még jobb, ha olyan hálózatra gondolunk, amelyet az anyagcsere tényleges tápfeszültségével üzemben lehet tartani és amelyet a tényleges genetikai vezérlő-mechanizmusainkkal meg lehet építeni, illetve meg lehet szervezni.

Összegezve: azt hiszem, hogy a problémák első fázisát – azt a tisztán formális problémát, hogy egyáltalán valamely „ekvivalens hálózatot” találjunk – McCulloch és Pitts megoldották. Azt hiszem továbbá, hogy annak a „rossz érzésnek” a legnagyobb része, amelyet akkor érzünk ha megpróbáljuk „megmagyarázni” a központi idegrendszert, ehhez a fázishoz tartozik, és ennek következtében túlhaladottnak tekintendő. Ennek ellenére bőségesen elegendő rossz érzés jut a probléma következő fázisára, amelyben arról van szó, hogy olyan „ekvivalens hálózatot” kell találni, amely a lehetséges, vagy akár csak plauzibilis térfogatnak és (anyagcsere- és genetikai) követelményeknek megfelel.

A probléma tehát nem az, hogy hogyan hajt végre a központi idegrendszer valamely tetszőleges partikuláris feladatot, hanem az, hogy hogyan tudja végrehajtani teljes komplexitásában mindazt, amit ténylegesen végrehajt. Mik a szervezési elvei? Hogyan kerül

el a valóban komoly, tehát végzetes kimenetelű hibás működéseket olyan hosszú időtartamon át, amelyek átlagában több évtizedre tehetők?

Dr. Gerard:

Véleménye szerint tehát vannak meg sem fogalmazott problémák?

Dr. Neumann:

Lehetséges, hogy vannak olyan problémák is, amelyeket a jelenlegi logikai eszközeinkkel meg sem lehet fogalmazni.

Dr. Weiss:

Feltételezem, hogy a központi idegrendszernek csak egy elképzelhető és logikailag helytálló, de nem szükségképpen valóságos mechanizmusát tárgyaljuk. A valóságos idegrendszer bármely elméletének azonban meg kell tudni magyarázni a szabályozás tényeit – tehát azt, hogy a mechanizmus ugyanazt vagy lényegében hasonló eredményt produkál akkor is, ha az összeköttetések hálózata sokféle, előre nem látható módon megváltozott. Dr. Neumann szerint lehetséges olyan gépet tervezni, amely biztosítékokat tartalmaz a hibák ellen, illetve alkalmas berendezéseket a hibák kijavítására, ha előfordulnak. Ebben az esetben a gép tervezésénél a jövőbeli eshetőségeket figyelembe vették. Az idegrendszer esetében a fejlődésnek kellett a szükséges korrektív eszközöket megépíteni. Minthogy a tényleges behatások és a természetben előforduló változatok, továbbá a kísérletező neurofiziológusok rendkívül nagyszámú eltéréseket okoztak, felmerül a kérdés: elképzelhető-e olyan mechanizmus, amely ezt a számtalan lehetőséget előre látja és amelyben a megfelelő korrektív eszközök be vannak építve?

Dr. Neumann:

Természetesen meg sem tudom kísérlni, hogy választ adjak arra a kérdésre, hogyan jutott el a fejlődés egy adott pontig. Szeretnék azonban néhány megjegyzést tenni arról a sokkal jobban körülhatárolt kérdésről, amely a hibákra, ezek előrelátására, felismerésére és kijavítására vonatkozik.

Egy mesterséges gépet fel lehet szerelni olyan szervekkel, amelyek a hibákat automatikusan felismerik és kijavítják. Ténylegesen majdnem minden jól tervezett gép tartalmaz olyan szerveket, amelyeknek éppen ez a feladatuk – természetesen mindig bizonyos meghatározott és korlátozott területeken. Továbbá: ha adva van bármely egyedi gép, mindig lehetséges olyan második gépet konstruálni, amely „figyeli” az elsőt és érzékeli, esetleg ki is javítja az elsőnek a hibáit. A baj azonban az, hogy most már a második gépnek a hibái maradnak ellenőrzés nélkül, azaz, „quis custodiet ipsos custodes”¹⁵? Ha egy harmadik, negyedik stb. gépet építünk a másodrendű, harmadrendű stb. hibák ellenőrzésére, a probléma csak áttolódik. Ezen túlmenően pedig az első és másodrendű gép együttesen több hibát produkál, mint az első egyedül, hiszen több alkatrésze van.

Ennek ellenére valamely ilyesféle eljárásnak – persze szerényebb méretekben – mégis van értelme. Egy meghatározott gépen vagy gépcsoporton szerzett statisztikai tapasztalatból tudjuk, melyek általában azok az alkatrészek, amelyek a leggyakrabban meghibásodnak és megtehetjük azt, hogy csak ezeket „figyeljük meg”, stb.

A problémának egy másik lehetséges megközelítési módja, amely általánosabb kvantitatív értékelést is lehetővé tesz, a következő: tételezzük fel, hogy egy olyan gépről van szó, amelynél annak a valószínűsége, hogy egy elemi művelet hibás: 10^{-10} , általánosságban tehát minden 10^{10} művelet közül egyszer követ el hibát. Tételezzük fel, hogy ennek a gépnek olyan problémát kell megoldania, amelyhez 10^{12} elemi műveletre van szükség. Normális, „felügyelet nélküli” működés során ennek megfelelően a probléma feldolgozásánál általánosságban 100 hiba fordul elő, vagyis teljesen használhatatlan lesz.

¹⁵ Ki őrizze magukat az öröket?

Kapcsoljunk össze azonban három ilyen gépet oly módon, hogy az eredményeket minden egyes elemi lépés után összehasonlíttjuk, majd a következőképpen járunk el: (a) ha mindhárom gép eredménye egyezik, további vizsgálat nélkül folytatják a működést, (b) ha bármely két gép eredménye egyező és csak a harmadik különböző, a műveletsort a többségi eredmény alapján folytatják, (c) ha azonban a három közül nincs két egyező eredmény, mindhárom gép megáll.

Ez a rendszer helyes eredményeket ad, kivéve, ha a számítás valamely pontján a három gép közül legalább kettő egyidejűleg hibázik. Annak a valószínűsége, hogy két gép egy adott műveletnél egyidejűleg hibát vét, $10^{-10} \times 10^{-10} = 10^{-20}$. Annak a valószínűsége, hogy az adott műveletnél bármely két gép a három közül egyidejűleg hibázik, 3×10^{-20} (mert három lehetséges párt lehet képezni a három egyed /gép/ közül). Annak a valószínűsége, hogy ez egyáltalán megtörténik (ti. bárhol) az egész probléma folyamán, ennek megfelelően $3 \times 10^{12} \times 10^{-20} = 3 \times 10^{-8}$, tehát körülbelül 33 millió működés közül egy. Ilyen módon 33 millió működés közül csak egy eshetőség van arra, hogy egy ilyen géphármas hibásan oldja meg a problémát, jól lehet a géphármas bármelyik tagjának önmagában alig van valami esélye a helyes megoldásra.

Figyeljük meg azonban, hogy ez a géphármas, tekintet nélkül arra, hogy a hibavizsgálat milyen bonyolult módon történik, logikailag még mindig hibásan működhet, noha természetesen csak rendkívül kis valószínűséggel. A hiba előfordulás, illetve ennek valószínűsége azonban mégis lényegesen csökkent és tulajdonképpen ez az, amire szükség van.

Dr. Weiss:

Hogy a lényegét kidomborítsuk, szeretném megismételni: ha ismerjük azokat a leggyakoribb hibatípusokat, amelyek egy meghatározott gépben előfordulhatnak, a gép tervezése során lehetséges a hibát korrigáló segédberendezéseket is tervezni. Az idegrendszer egyik legfontosabb tulajdonsága azonban az, hogy a hibát nyilvánvalóan olyan helyzetekben is ki tudja javítani, amelyeket nem lehetett előre látni. (Azoknak a mesterséges beavatkozásoknak a száma, amelyeket különböző készülékekkel a központi idegrendszerre gyakorolni lehet anélkül, hogy a szervezet biológiailag hasznos reakcióját befolyásolnánk, végtelen nagy.) A neurális automata elgondolásnak nemcsak az idegrendszer normális működéséről kell tudni számot adni, hanem a legkülönbözőbb abnormis helyzetekben mutatott viszonylagos stabilitásról is.

Dr. Neumann:

Nem értek egyet ezzel a következtetéssel. Az alkalmazott érvelés túlságosan kockázatos és nagy óvatosságot igényel. Ennek ellenére lehetséges elhárítani olyan hibákat, amelyeket specifikusan nem látunk előre. Íme néhány példa annak illusztrálására, amire gondolok:

Lehetséges olyan elektromos automatát tervezni és építeni, amely mindaddig helyesen fog működni, amíg minden egyes beépített ellenállás a tervezési értéktől legfeljebb 10 százalékkal tér el. Mármost meg lehet próbálni ezt a gépet kísérleti úton „megzavarni” olyan módszerekkel, amelyek az ellenállások értékét megváltoztatják (például azáltal, hogy a gép bizonyos részeit túlfűtik.) Mindaddig, amíg egyetlen ellenállás értéke sem változik többet, mint 10 százalékkal, a gép helyesen működik tekintet nélkül arra, hogy mennyire komplikált, mennyire „előre nem látott” is a zavaró kísérlet.

Vagy például: ki lehet képezni egy olyan páncéllemezt, amely legfeljebb meghatározott erejű becsapódásokkal szemben áll ellen. Ha vizsgálat alá veszik, mindaddig eredményesen kiállja a vizsgálatokat, amíg a szilárdsági határt el nem éri, tekintet nélkül arra, hogy mennyire újszerű a kísérletnél felhasznált fegyver, lőpor vagy lövedék koncepciója stb.

Világos, hogy hogyan lehet ezeket a példákat a neurális és genetikus helyzetekre átvinni.

Összefoglalva: elegendő, ha a hibákat és hibaforrásokat eredetük szerint, tehát bizonyos döntő vonásaikban látjuk előre, nem pedig specifikusan, vagyis teljes részletességükben. Azok a

területek, amelyekre ezek az általánosságok vonatkoznak, rendkívül nagyok lehetnek, teli előre nem látott, előre nem gyanított, végeredményben azonban nem lényeges részletekkel.

