

Neuronális automaták

TARJÁN REZSŐ

1. Bevezető megjegyzések

Több mint 50 év telt el azóta, hogy Leninnek az Empíriokriticizmusban a filozófiai idealizmusnak akkori képviselői ellen folytatott harcában még bizonyítania kellett, hogy az ember valóban az agyával gondolkodik. A természettudományok, különösen a biológia és a fiziológia fejlődése azóta ezt a kérdést is, mint sok más kérdést, a materializmus javára döntötte el. Ma már nem is vitatják, hogy valóban az agy a gondolkodásnak sui generis szerve. Korunk tudományos problémája többé már nem az, hogy az ember agyával gondolkodik-e vagy nem, — a kérdés inkább így hangzik: *hogyan* gondolkodik az ember, pontosabban, *hogyan működik a gondolkodás szerve?* Hogyan szerveződnek a részei egésszé, milyen fizikai és kémiai természetű folyamatok játszódnak le az agyban a komplex gondolkodási folyamat alatt?

Az utóbbi tíz évben ugyanis olyan bonyolult automatákat (elsősorban a modern elektronikus digitális* számológépekre gondolk) konstruáltak, amelyeknek már pusztán léte is bebizonyította, hogy bizonyos műveleteket, amelyeket csak azért, mert az ember agyával vitte véghez, „szellemi” munkának tekintettek, gépek is végre tudnak hajtani. A fejlődés azt is bebizonyította, hogy a modern digitális automaták nemcsak tisztán *numerikus* problémák megoldására alkalmasak. Az utóbbi időben nagy haladást értek el a nyelvek fordításának mechanizálása területén; legújabban pedig sikerült ilyen gépek segítségével matematikai-logikai és geometriai problémák bizonyítását heurisztikus módon megkonstruálni. Intenzív kutatás folyik az alakfelismerés problémáján, azzal a végcélal, hogy a kézzel írott számokat gépi úton olvassák le és a számoló automaták által közvetlenül feldolgozható formára hozzák.

Ilyen körülmények között nem csodálkozhatunk, hogy a szellemi munka, illetve a gondolkodás fogalmának a tartalma fokozatosan változik, és az utóbbi időben a természettudósok, különösen a számoló automaták specialistái érdeklődésének homlokterébe került. A. M. Turing, az ismert angol matematikus¹ gondos elemzés után megkísérelte megadni a gondolkodásnak úgynevezett „operatív”, azaz lényegében a magatartás által implicite meghatározott definícióját. Eszerint egy gépről akkor mondhatjuk, hogy gondolkodik,

* Az elnevezés az angol „digit” (számjegy) szóból származik.

¹ Turing, A. M.: Computing machinery and intelligence. Mind, Vol. LIX, No 236 (Oktober 1950), pp 433–460.

ha előírt kísérleti feltételek mellett ésszerű kérdésekre oly módon válaszol és ezáltal az embert úgy tudja utánozni, hogy a kérdéseket feltevő ember, aki csak az írógép által kiírt válaszokat látja, nem pedig magát a válaszoló gépet, az adott válaszokat egy ember által adottaknak tartja. Bár ennek a definíciónak, amint McCarthy és Shannon² megjegyzik, megvan az az előnye, hogy olyan metafizikus fogalmakat, mint az „Én” stb. nem használ, azonban nem tesz különbséget a valódi, eleven gondolkodás, (ahogyan ez valamilyen probléma megoldásának során történik) és aközött, hogy a helyes felelet között egy megfelelően nagy tárolóból (emlékezet) a gép egyszerűen kikeresi. Cobb³ az ismert angol fiziológus fiziológiai szempontból a gondolkodást úgy definiálja, mint „olyan események sorozatát, amelyek az agyban külső ingerek hatására (beleértve más agyakból jövő közléseket is), mint reakciók keletkeznek, és amelyek az agyvelő különböző részei közötti kölcsönhatásoktól függnék”. Habár ez a fiziológiai definíció is tisztán fenomenológiai, filozófiai szempontból azért érdekes, mert hangsúlyozottan materialista jellege van: a gondolkodást itt *reális* (azaz materiális) *események, külső ingerek következtében létrejövő esemény-sorozatnak* fogja fel.

1. 1. *Problémafeltevés.* Ezek valamint sok más hasonló vizsgálat a kérdésfeltevés általánosításához vezetett, a következő formában: *egyáltalán milyen szellemi munkák mechanizálhatók?* A *komplex* emberi gondolkodási folyamat mely részeit valósíthatják meg egészen vagy részben automaták? Milyeneknek kell az ilyen automaták szerkezeti elemeinek lenniük? Milyen logikai struktúrával kell bírniuk, hogy bizonyos problémákat meg tudjanak oldani? Melyek az ilyen automaták általános tulajdonságai? Ilyen és hasonló kérdésfeltevések tipikusak az automaták absztrakt általános elméletére nézve, amelynek része a neuronális automaták elmélete is.

Nyilvánvaló, hogy az ilyen és hasonló kérdésfeltevések nemcsak tisztán természettudományos, hanem filozófiai, különösen ismeretelméleti szempontból is nagy érdeklődésre tartanak számot. A következőkben megkíséreljük, amennyiben ez a fejlődés mai állapotában lehetséges, szisztematikusan ismertetni az automaták, különösen a neuronális automaták absztrakt elméletének fő eredményeit. A kifejtés szükségképpen egyoldalú: a tényanyagot főként a kutató mérnök szempontjából tárgyaljuk, akit elsősorban a gyakorlatilag is használható instrumentálás érdekel. Egy hipotézis, illetve elmélet helyes voltának legszigorúbb bíróját azonban éppen a gyakorlat kritériuma jelenti. Megkíséreljük a tényanyagot úgy bemutatni, hogy egyfelől, ami módszertanilag nagyon fontos, a kutatónak jellegzetes gondolkodási módja lehetőleg világosan kitűnjék, másfelől, azonban úgy, hogy a tulajdonképpeni filozófiai problémák kidolgozása, illetőleg az eredmények filozófiai általánosítása lehetőség szerint könnyebbé váljék.

1. 2. *Az automata absztrakt fogalma.* Az absztrakt elmélet tárgya érthető módon nem azonos a műszaki értelemben vett automatikus rendszerek tárgyával, amint azt az ipari szabályozó és vezérlési folyamatoknál kifejlesztették. Az automaták általános elmélete egy fokkal absztraktabb. Biológiai kifejezést használva minden műszaki automatában találhatunk receptor szerveket, amelyek a működés alapjául szolgáló információkat szállítják. Vannak továbbá

² *McCarthy—Shannon* : Automata studies, Princeton 1956.

³ *Cobb, St.* : On the nature and locus of mind. Archives for neurology and psychiatry, Vol. 67 (1952).

effektor szervek, amelyeknek feladata erőkifejtés, és amelyeket a receptorok által szállított információk alapján vezérelnek. A két szerv között mint összekötő *mindig* van egy *vezérlő szerv*, amelynek a feladata a receptorok által szállított *információk feldolgozása* és effektív vezérlő-jelek továbbítása az effektor szervhez. Egyszerű automatákban a vezérlő szerv rendszerint nincsen fizikailag elválasztva a gép többi részétől. Mint funkció azonban mindig jelen van és a modern komplex automatáknál általában mint fizikailag is elválasztott, külön egységet szerkesztik meg. A vezérlő szerv legtöbbször (nem mindig) *kapcsoló-elemek* (többé-kevésbé) *komplex hálózata*, amelyek elektromos impulzusokra reagálnak és ezeket bizonyos feltételek mellett továbbadják. A receptor, illetve effektor szervek az absztrakt elmélet számára érdektelenek; az automaták absztrakt elméletének tárgya a *központi vezérlő szerv funkciói és logikai struktúrája*, a vezérlő szervnek az egyes keresztülvendő feladatok számára szükséges, illetve elégséges tulajdonságai.

Ennek a szemléleti módnak megfelelően egy automata véges-sok, lehetőleg általánosan definiált tulajdonságú kapcsoló-elemből álló rendszer. A kapcsoló-elemek oly módon vannak egészszé szervezve, hogy a receptorokból kiinduló, jól definiált, időben és/vagy térben elrendezett bemenő jelzés-sorozatra (eseményekre) az effektoroknak időben és/vagy térben elrendezett jól definiált eredmény-sorozatot ad le; ezek az ún. kimenő jelzések a bemenő jelzéseken kívül az *automata belső állapotától is függenek*. Az automata belső állapotán az egyes kapcsoló-elemek állapotának összességét értjük, a bemenő események fellépésének időpontjában.

Az így definiált automata akkor *véges*, ha a lehetséges belső állapotok száma véges, más esetben végtelen végtelen automatának nevezzük. Az automatákat véges (vagy: végtelen) tárolóval („memória”) láthatjuk el. Két nagy osztályba oszthatók: a *logikai* automaták olyan kapcsoló-elemből állanak (döntési szerveknek is nevezik), amelyek két és csak két stabil állapotra képesek a $A \cdot \bar{B} \vee \bar{A} \cdot B$ kizáró-diszjunkció értelmében. A másik osztályba a *neuronális* automaták tartoznak, amelyeknek kapcsoló-elemei definíció-szerűen nagyobb vagy kisebb mértékben a valódi neuronok tulajdonságaival bírnak.

1. 3. *Az automaták elméletének keletkezése.* Az automaták elméletének keletkezését rendszerint A. M. Turing⁴ híres munkájának megjelenésétől számítják, amelyben egy (akkor még csak hipotetikus) automatikus számológép segítségével *egyáltalán kiszámítható számok* problémáját tárgyalta. Az eredmény az, hogy minden olyan szám kiszámítható, amelynek kiszámításához megadhatunk egy *effektív eljárást*, azaz a számolási műveletek olyan jól-definiált véges hosszú sorát, amit a gép effektíve végre tud hajtani. Érdekes megjegyezni, hogy a Turing-féle egyetemes gép absztrakt eszméje majdnem egy teljes évtizeddel megelőzte az első igazi automatikus digitális számológépet.

A következő lépést McCulloch és Pitts 1943-ban megjelent alapvető dolgozata⁵ jelenti, amelyben először kísérelték meg a központi idegrendszer matematikai értelemben véve szigorú elméletének kifejtését a matematikai

⁴ Turing, A. M.: On computable numbers, etc. Proc. London Math. Soc. Sc. 2. Vol. 42 (1937), pp 230—265.

⁵ McCulloch, W. S. and Pitts, W.: A logical calculus of the ideas immanent in neurons activity. Bull. Math. Biophysics, Vol. 5 (1943) pp 115—133.

logika segítségével. Az akkor ismert fiziológiai tényekből kiindulva a neuron absztrakt modelljét mint egy egyetemes kapcsoló-elemet a következő tulajdonságok segítségével definiálták: két és csak két stabil állapota van; meghatározott ingerküszöbvel rendelkezik, amelyen alul nem lehet ingerelni. Egy kimenő és több bemenő vezetéke van; utóbbiak vagy ingertkeltő vagy gátló vezetékek lehetnek, az ingerlés, illetve a gátlás fiziológiai tényének megfelelően. A neuron funkciója a kisülés aktusából áll, azaz meghatározott számú bemenő vezeték egyidejű ingerlésekor a kimenő vezetéken impulzus jelenik meg; ha viszont akár csak egyetlen tiltó vezetéket is inger ér, a neuron teljes gátlásba kerül. Az egyszerűség kedvéért felteszik továbbá, hogy az impulzusok időben egyenlő távolságuak és tartamúak. Ennek a modellnek alapján az idegrendszer a kapcsoló-elemek időben változatlan hálózatának lehet felfogni, és tulajdonságait a matematikai logika segítségével elméletileg lehet tanulmányozni.

A modern számoló automaták létrejötte, különösen vezérlési feladatokra való alkalmazásai nagy fellendülést okoztak az elmélet területén is. Az elmúlt évtized első fele óta Turing, illetve McCulloch—Pitts munkáit egyre növekvő számban követték más szerzők munkái is; jó részüket összegyűjtve is kiadták.⁶ Nem lehet feladata ennek a szükségképpen rövidre fogott referátumnak, hogy ennek az elméletnek minden érdekes eredményéről részletesen beszámoljon. A filozófiai szempontból érdekes fő eredményt azonban röviden tárgyalni kell.

1. 4. *A logikai automaták elméletének fő eredménye.* Lényegében a kiszámítható számok Turing-féle eredményének általánosításáról van szó; ezt tulajdonképpen McCulloch és Pitts eredeti munkája már tartalmazza, de a megfelelő matematikai-logikai szigorral először 1956-ban S. C. Kleene⁶ fogalmazta meg. Előzetesen azonban meg kell egyezni abban, hogy egy eseményt akkor nevezünk *definitnek*, ha egy adott időpontban (ami a véges múltban is lehet) egy neuron kisülése útján reprezentálható. A *szabályos* (reguláris) *esemény* olyan (komplex) esemény, amely véges sok definit esemény konjunkciójaként vagy diszjunkciójaként (az ítélet-kalkulus értelmében) állítható elő. A főeredmény mármost azt mondja ki, hogy a fent megadott tulajdonságokkal bíró univerzális kapcsolóelemekből álló véges hálózat segítségével *minden* reguláris eseményt és csakis ilyeneket reprezentálhatunk. Ez az eredmény, ami tulajdonképpen egy egzisztencia-tétel, alapvető jelentőségű. Amint ezt Neumann János a nemrég elhunyt magyar származású matematikus⁷ megjegyezte, lényegében azt jelenti, hogy minden egyes olyan funkciót, ami véges sok kijelentés segítségével definiálható, kapcsoló-elemek alkalmas hálózatával reprezentálhatunk, azaz realizálhatunk is.

Mai számoló automatáink könnyen belátható módon funkcionálisan ekvivalensek a fenti értelemben vett logikai automatával; a beépített utasítások (mint például az összeadás vagy a logikai összehasonlítás stb.) *funkcionálisan* megfelelnek a kapcsoló-elemeknek. Hogy a számoló automaták egyetlen *univerzális* utasítás helyett többféle utasítással bírnak csak *gyakorlatilag kényelmes*, de *nem jelenti az általánosság korlátozását*. Amint B. van der

⁶ Kleene, S. C.: Representation of events in nerve nets and finite automata. Mc Carthy—Shannon. Automata Studies, pp 3—41.

⁷ v. Neumann, J.: The general and logical theory of automata. The Hixon Symposium on cerebral mechanism in behaviour. Wiley, 1951, pp 1—41.

Poel⁸ bebizonyította, bármely utasítás *egyetlen* univerzális utasítás segítségével is (mikroprogramok formájában) kifejezhető. Újabban építettek olyan számológépeket is, amelyek olyan *tranzisztoros egyetemes kapcsoló-elemből* vannak felépítve, ami a $A|B = A \cap B$ Scheffer-féle egyetemes műveletnek felel meg. A Sheffer-féle egyetemes műveletből instrumentálható a van der Poel-féle egyetemes utasítás is.

A számoló automatákra vonatkozóan a McCulloch—Pitts-féle eredmény lényegében azt jelenti, hogy minden, ami egyértelműen és szigorúan megfogalmazható, röviden, hogy minden olyan tevékenység, amely egyértelmű szabályokba foglalható (algoritmizálható), a számoló automatákra is programozható. Ezzel a mechanizálható szellemi munkának a bevezetőben felvetett problémája a következő problémává alakult át: *milyen fizikai vagy szellemi folyamatok vagy tevékenységek számára találhatók egyértelmű szabályok?* Ezen a ponton a logikai automaták elmélete már átmegy az algoritmusok elméletébe, amely területen a szovjet matematikusok, elsősorban A. A. Markov⁹ végeztek úttörő munkát. Az algoritmusok elmélete az elméleti alapja az olyan, az első tekintetre talán meglepő alkalmazási lehetőségeknek, mint pl. a heurisztikus bizonyítás-konstrukció,¹⁰ amelyek olyan mértékben sikerülnek, amilyen mértékben az ismert formális-logikai szabályokon kívül a heurisztikus gondolkodásmód néhány empirikus szabályát is programozni tudjuk.

2. A logikai automaták elméletének jellemző nehézségei

2. 1. *A plauzibilis következtetés problémája.* Amint már előbb említettük, McCulloch és Pitts a matematikai logika segítségével a központi idegrendszer lehetőleg szigorú elméletét akarták kifejteni. Amint McCulloch később megjegyezte, a heurisztikus alapgondolat, amiből kiindultak, az volt, hogy a központi idegrendszert úgy tekintették mint egy Turing-féle gépet: a neuron absztrakt modelljének tulajdonságait, (illetve S. C. Kleene szigorúbb megfogalmazásában már az esemény fogalmát is) oly módon absztrahálták a valóságból, *hogy a matematikai logika alkalmazható legyen.* Ezáltal, bár nagy területen alkalmazható elméletet kapunk, azonban éppen az előfeltételek megválasztása vezet az elmélet alkalmazási területének korlátozásához. Az egyik következmény, hogy az elméletnek megfelelő automaták *tisztán logikai* gépek, abban az értelemben, hogy *csak olyan műveleteket képesek végrehajtani, amelyek kétértékű logikai műveletekre vezethetők vissza.* Olyan fontos gondolkodási folyamatok tárgyalása, mint például a plauzibilis következtetés G. Pólya¹¹ értelmezésében, azaz az induktív logika problémái, amint ezt pl. R. Carnap¹² tárgyalja, vagy az alakfelismerés, illetve a tanulás folyamata, ami éppen legújában lett gyakorlatilag is fontos, az alapfeltevések megválasztásának következtében csak nehézségek árán a formális logika kerülő útján, és csak nagyon megközelítőleg tárgyalhatók.

⁸ van der Poel, W. L. : The essential types of operations in an automatic computer. Nachrichten-technische Fachberichte. Bd. 4 pp 144—145.

⁹ A. A. Markov : Tyeoria Algoritmov.

¹⁰ Gelernter, H. : Realisation of a geometry theories proving machine. UNESCO Conference on Information Processing. Paper 1. 6. 6.

¹¹ Pólya, G. : Mathematics and plausible reasoning. I—II. Princeton. 1954.

¹² Carnap, R. : Logical foundations of probability. Routledge and Regan, London, 1954.

A plauzibilis következtetés alapformája, ami a klasszikus szillogisztika *modus tollens*-ének felel meg, Pólya (1. c.) szerint így ábrázolható:

A-ból következik B

B — igaz

A tehát inkább plauzibilis.

A „plauzibilis” szó itt a — *szubjektíve becsült* — valószínűség értelmében értendő. Ha egy pillanatra eltekintünk az itt alkalmazott valószínűség-fogalom szubjektív tartalmától, úgy, amint Pólya rámutatott, tényleg lehetséges, a fenti következtetési módnak megfelelő valószínűséget formálisan is felírni: az A kijelentés plauzibilitása egyszerűen az A kijelentés *feltételes valószínűségével* egyenlő, azzal a feltétellel, hogy a B kijelentés igaz. Ha a plauzibilitás *érzete* egy bizonyos küszöbértéket meghalad, levonjuk az *induktív következtést* és cselekszünk.

Plauzibilis munkahipotézisként tehát megkísérelhetjük, hogy a plauzibilis következtetést egy feltételes valószínűségi folyamat által közelítsük meg. Sholl¹³ anatómiai leletei mindenesetre arra utalnak, hogy legalábbis a központi idegrendszer bizonyos helyein, az egyes neuronok kapcsolatai jó megközelítéssel valószínűségi elosztással közelíthetők.

2. 2. *A hasonlóság problémája.* A valószínűségi folyamat feltevése azonban önmagában még nem elegendő a gondolkodási folyamat megértéséhez. A plauzibilitást ugyanis a valóságban majdnem sohasem csak egyetlen jel alapján becsülik meg, hanem *teljes helyzetek* hasonlósága alapján, amelyeket különböző *jelek halmaza jellemez*. Ez a fajtájú hasonlóság, amint ezt másutt¹⁴ részletesebben kifejtettük, lényegesen különbözik a geometriában alkalmazott hasonlósági fogalomtól. A hasonlóság a geometriában szigorúan logikai fogalom, amit nem *érzékelnek*, hanem *bizonyítanak*. A geometriai hasonlóság fogalma definíciószerűen adott egyértelmű jelekből áll, amelyek az összehasonlított geometriai objektumoknál (pl.: két síkháromszögnél) *vagy egyidejűleg fennállnak, vagy nem*. Az első esetben az objektumok hasonlóak, a második esetben nem. Két geometriai objektum hasonlóságának kijelentése ennek következtében csak vagy igaz, vagy hamis lehet.

Ezzel szemben a hasonlóság a biológiai rendszerekben általánosabb és inkább kvalitatív jellegű. Ebben az esetben a jelek olyan *végtelen* halmazáról van szó, amelyeknek egy része *lényeges*, másik része pedig *lényegtelen*, és ezek a reális külvilág összehasonlítható tárgyaiban csak részben fedik egymást. Ezért a biológiai rendszerek a reális külvilág objektumait *többé vagy kevésbé hasonlóknak érzékelhetik*, mint pl. az apa és fia hasonlóságának esetében. A biológiai hasonlóság-fogalom tehát általánosabb és a *mérték fogalmát* is tartalmazza. Axiomatikus tárgyalásnál a lenini visszatükrözési elvvel összhangban feltételezhetjük, hogy a külvilág objektumai a központi idegrendszerben mint tulajdonságaik halmaza tükröződnek: két objektum általánosított értelemben akkor hasonlít egymáshoz, ha tulajdonságaik halmazai nem idegenek, azaz van közös részük. Ebben az esetben definiálhatjuk a hasonlóság mértékét

¹³ Sholl, D. A. : Dendritic Organisation in the neurons of the visual and motor cortices of the cat. Journ. of Anatomy, Vol. 87 (1953) pp 387—406.

¹⁴ Tarján, R. : Neuronal Automata. Proc. II. Intern. Congress of Cybernetics. Namur, 1958.

is. A hasonlóságnak ez az általánosított fogalma az alakfelismerés céljaira szolgáló automaták elméletében és gyakorlatában nagy szerepet játszik.

2. 3. *A tanulás problémája.* A helyzet még bonyolultabb lesz, ha olyan automaták elméletére térünk át, amelyeknek *a tanulás folyamatát kell többé-kevésbé megközelíteniük.*

Pavlov óta tudjuk, hogy a tanulás alapja a feltételes reflexek. A központi idegrendszerben szóba jöhető fiziológiai mechanizmusok részleteit azonban a rendkívüli kísérleti nehézségek következtében csak nagyon pontatlanul ismerjük, hiszen a szükséges elektronikus kísérleti eszközök lényegükben csak a második világháború után fejlődtek ki. Mindenesetre bizonyos, hogy a feltételes reflexeknél a központi idegrendszernek az a képessége, hogy különböző részei között többé-kevésbé tartós kapcsolatokat tud létrehozni, szükséges feltétel. Hogy azonban ezek a kapcsolatok, illetve az irradáció fiziológiailag hogyan jön létre, még nem tudjuk.

Ilyen körülmények között nem csodálkozhatunk, hogy ma még nem beszélhetünk a neuronális automaták egységes elméletéről. A legnagyobb nehézséget az okozza, hogy a probléma túlságosan komplex, és sok esetben még a részproblémák sem eléggé pontosan fogalmazhatók meg. Mint korábban¹⁵ rámutattunk, *mai* automatáink a plauzibilis következtetésre *elvileg* sem alkalmasak, mert kapcsoló-elemeik két és csak két stabil állapottal rendelkező szervek: az impulzusok tartama vagy intenzitása teljesen irreleváns, feltéve, hogy egy bizonyos küszöbértéknél nagyobbak. *Mai* automatáink kapcsoló-elemei ezért csak az esemény *bekövetkezését* jelzik, minden kvalitatív jel nélkül. Ha olyan automatákat akarunk építeni, amelyek az előbb említett feladatokat akár csak bizonyos mértékig is teljesíteni tudják, olyan kapcsoló-elemekkel kell dolgoznunk, amelyek a valódi neuronoknak legalább legfontosabb tulajdonságaival rendelkeznek. Az ilyen automatákat érthetően már neuronális *automatáknak* kell nevezni.

2. 4. *A részleges modellek módszere.* Az említett nehézségek következtében az egyetlen helyes módszer az, hogy szisztematikus kísérletek útján annyi új tényanyagot gyűjtünk, amennyit csak lehetséges és az így összegyűjtött anyagot, illetve *a komplex gondolkodási folyamatnak egyes, legalább fenomenológiailag jól értett részeit ideiglenes munkahipotézisek alapján megkíséreljük rendezni.* Az érvényességi területet tudatosan a tárgyalt részfolyamatra korlátozzuk, ott azonban a hipotézisnek, amennyire csak lehetséges, összhangban kell lennie a kísérletileg megfigyelt tényekkel. A munkahipotézis alapján levont következtetéseket össze kell hasonlítani a tapasztalattal, és *a megfigyelt eltérések alapján addig kell korrigálni,* amíg (legalábbis gyakorlatilag) használható, ha elméleti szempontból még nem is egészen kifogástalan eredményekhez vezetnek. A nagy módszertani előny éppen az elmélet és a tapasztalat állandó kölcsönhatásában van. A *parciális modelleknek* ezt a módszerét a természettudományok más területein pl. az elméleti fizikában is széles körben használják, sőt újabban az elméleti közgazdaságtan területén is nagy sikerrel alkalmazzák a valóság fokozatos megközelítésére. Így például ma az atommagnak több matematikai modellje is van, (mint például a héjmodell vagy a cseppmodell), amelyek ugyan egyenként a jelenségeknek csak korlátozott csoportját tükrözik vissza, összességükben azonban azzal a nagy módszertani előnnyel járnak, hogy jó tájékozódást adnak a szisztematikus kísérletekhez.

¹⁵ *Tarján, R. : Neuronal Automata. Cybernetica. Vol. 1. (1958) No 3. pp 189—196.*

A fejlődés folyamán ezeket esetleg az atommag egységes elméletévé lehet majd összeolvasztani, hasonlóan mint annak idején az atomburok kvantumelméleténél is megtörtént.

A most következő második részben, a szűkebb értelemben vett neuronális automaták elméletének legfontosabb eredményeit ismertetjük. A régebbi eredmények miatt, amelyek inkább biofizikai jellegűek, utalunk N. Rashevsky¹⁶ kiváló könyvére; a továbbiakban a legújabb, főleg a rendszerelméleti vagy egy lehetséges instrumentáció értelmében orientált munkákra szorítkozunk.

3. A neuronális automaták elméletének alapvető problémái

3. 1. *Meghatározások.* A neuronális automaták szisztematikus elméletének a megfelelő fogalmak kibővített definíciójából kell kiindulnia. Először az *eseménynek* fogalmát kell kibővíteni, illetve újólagn definiálni, mivel S. C. Kleene definíciója a matematikai logika igényeire szabott, és éppen ezért a neuronális automaták elmélete számára nyilvánvalóan túlságosan szűk. Kifogástalan definíció egyelőre még nincs. Nyilvánvalóan lehetetlen a kívülvalódi eseményeit teljes változatosságukban, ahogy azok a központi idegrendszerben megjelennek, *egyetlen* definícióban megragadni. A bekövetkezés, illetve kimaradás (nem bekövetkezés) pusztán tényén kívül legalább *néhány* minőségi tulajdonságot mégis figyelembe kell venni. A legegyszerűbb minőségek kétségtelenül egy inger *tartama* és *intenzitása*; ezért kézenfekvő, hogy legalább ezt a két minőséget be kell vennünk a definícióba. Továbbá számításba kell vennünk azt a körülményt is, hogy a biológiai rendszerek több *különböző* *funkciójú* makroszkopikus receptorral bírnak (mint a szem és a fül), amelyek viszont igen nagyszámú egyes mikroszkopikus receptorokból (mint például a retinában a csapocskák és pálcikák), tevődnek össze. Ezek normális körülmények között *egyidejűleg* működnek; az általuk szállított jelzéseket a központi idegrendszer eredetüknek és összetételüknek megfelelően *diszkriminálja*. Következő lépésként egy univerzális kapcsoló-elem tulajdonságait kell lehetőleg szigorúan definiálni, úgy hogy egyfelől az előzetesen definiált eseményekre reagáljanak, másfelől pedig a valódi neuronok tulajdonságait, amennyire csak lehetséges megközelítsék. Végül az így definiált kapcsoló-elemeket olyan hálózattá kell szervezni, amelynek tulajdonságai szigorú eszközökkel vizsgálhatók, illetve meg kell mutatni, hogy egy adott organizáció egy ésszerűen felállított feladatot vagy feladatok egy osztályát végre tudja hajtani.

Végül, legalábbis munkahipotézisként magának a neuronális automatának fogalmát is kissé pontosabban kell meghatározni. A neuronális automatát úgy definiálhatjuk, mint nem biológiai természetű univerzális kapcsoló-elemek rendezett összességét, ami megfelelő receptor, illetve effektor szervekkel van ellátva, és úgy van szervezve, hogy a makroszkopikus receptorok által közvetített információkat tárolni és koordinálni tudja. A koordinálás az „események” kölcsönös idő- és intenzitás viszonyainak szabályozásából, valamint vezérlő jelzések létrehozásából áll, úgy, hogy az effektor szervek valamely egyértelműen meghatározott feladatot optimálisan teljesíteni tudjanak.

3. 2. *A neuron elméleti modelljei.* Az első lépést a neuronok valódi tulajdonságai feltevésének irányában W. K. Taylor¹⁷ tette meg, aki a fiziológiai

¹⁶ Rashevsky, N.: Mathematical Biophysics. Chicago, 1948.

¹⁷ Taylor, W. K.: Electrical stimulation of some nervous system functional activities. Colin chery : Information Theory, London, 1956.

tények elemzése után, bár a tulajdonságokat formálisan nem definiálta, olyan elektronikus modellt épített, aminek az volt a legfontosabb tulajdonsága, hogy az ingerre nem egy egyetlen impulzussal, hanem *impulzusok sorával* reagált. Az impulzusok frekvenciáját az inger intenzitása határozta meg. A küszöbérték helyébe az *érzékenység* fogalmát vezette be, amit úgy definiált, mint az egy másodperc alatt leadott impulzusoknak a beérkező impulzusok számához való viszonyának a növekedését. Közelebről le nem írt módon a modellbe bedolgozta az *akkomodáció* tulajdonságát, azaz a megszokás általi kifáradást is. Az elméleti várakozásoknak megfelelően a modell útján kapott kísérleti eredmények az izolált idegszálak (axonok) közvetlen elektromos ingerlése útján kapott fiziológiai kísérletekkel kvalitatíve jó egyezést mutattak.

Szerző korábbi dolgozataiban¹⁸ mesterséges neuronok tulajdonságait szintén a fiziológiai tényekre támaszkodva vizsgálta meg. Eszerint a mesterséges neuront olyan fizikai berendezésnek tekinthetjük, amelynek belső struktúráját nem kell közelebről ismerni és amelyet a következő, külsőleg megfigyelhető tulajdonságok jellemeznek: egy kimenete van, ami tetszés szerint szétágazhat és több bemenete, amelyek részben ingerlően, részben gátlóan hatnak. A neuron feladata abban áll, hogy a beérkező ingerekre adott válaszként, valahányszor a beérkező ingerek algebrai összege egy meghatározott, időtől függő küszöbértéket meghalad, a kimenő vezetékre impulzust ad le. Egy és csak egy stabil állapota van, a nyugalmi állapot, amelybe a kimenő impulzus leadása után spontán visszatér. Az ingerküszöb az impulzus leadását közvetlenül követő *abszolút* refraktér periódus alatt gyakorlatilag végtelen nagy; az utána következő *relatív* refraktér periódus alatt monoton módon visszatér a normális nívóra.

Ezeknek a tulajdonságoknak feltevése alapján kimutatható, hogy állandó ingernél az egy másodperc alatt leadott impulzusok száma arányos az inger intenzitásával, azaz az impulzusok frekvenciája az intenzitás mértéke. Mivel az impulzusok leadása a monostabilitás következtében csak addig tart, amíg maga az inger, az így definiált kapcsoló-elem jelzi az inger tartamát is, azaz megfelel az esemény kibővített definíciójának is. A reakció feltétele a küszöbérték következtében a térbeli, illetve az időbeli szummáció fiziológiai tényét is figyelembe veszi.

D. G. Willis¹⁹ legújabb dolgozatában szintén figyelembe veszi a summációs effektust, illetve az ingerküszöböt, de két fontos kiegészítő feltevést is tesz: először az ingerküszöb az egyes szinapszisokra vonatkozik, másfelől azonban ez az ingerküszöb (az úgynevezett szinaptikus érték) a befutó jelzések típusa szerint bizonyos állandó összeggel, *tartósan* csökkenthető vagy növelhető. Ennek az utóbbi feltevésnek két nagyon fontos következménye van: először is az, hogy bizonyos körülmények között minden egyes neuron több bitet (az információ bináris egysége) tárolhat, ami az ember hatalmas emlékező kapacitásának egyik lehetséges magyarázata lenne; másodsor pedig elméleti lehetőséget ad *állandó kapcsolatok automatikus kiépítésére*, ami a feltételes reflexek létrejötte számára, illetve a tanulás folyamatához feltétlenül szükséges. Kísérleti eredmények eddig még nem ismeretesek; a digitális számológépen történt modellezési kísérletek az elméleti várakozásokat meg-

¹⁸ Tarjan, R. : On the instrumentation of logical problems. Acta Technica Hungarica. Vol. 27 (1959), No. 3-4.

¹⁹ Willis, D. G. : Plastic neurons as memory elements. UNESCO Conference on Information processing, Paper J. 6-7.

erősítik. Még meg kell vizsgálni, hogy ezek a feltevések mennyire bizonyulnak helyeseknek fiziológiai szempontból is. Nemrégiben C. L. Coates²⁰ hasonló tulajdonságú mesterséges neuron modellt hozott nyilvánosságra, amit egy alakfelismerő gép (a perceptron) számára egyetemes kapcsoló-elemének szánt és amit tranzisztros kivitelben gyakorlatilag is megvalósított. A kísérleti eredmények a neuronok legfontosabb ismert tulajdonságaiival kielégítő módon megegyeznek.

Minden eddig leírt elméleti vagy gyakorlati modellezési kísérlet a neuronok tényleges tulajdonságainak csak nagyon kis részét veheti tekintetbe, amint ez másként nem is várható. Így például már az a feltevés, hogy a kapcsoló-elemeket csak a bemenő vezetéseken (amelyek a szinapszisoknak felelnek meg) lehet ingerelni, illetve gátolni, erős leegyszerűsítést jelent, ami fiziológiailag sok esetben nem helyes. Ahogyan V. Braitenberg²¹ jogosan rámutatott, egy neuront bizonyos körülmények között már az is ingerelhet, ha két neuron dendritjei elegendő közel vannak egymáshoz. De az ilyen finomságok figyelembevétele olyan elméleti nehézségekhez vezetne, amelyeket a mai elméleti eszközeinkkel még egyáltalában nem tudunk megoldani.

4. A neuron-hálózatok tulajdonságai

Ha a neuronok (kapcsoló-elemek) tulajdonságai rögzítve vannak, a neuronális automaták tulajdonságai, illetve a feladat, amelyet végre tud hajtani, csak szervezés módjától függ. Mivel egy neuron általában mind a bemenő, mind a kimenő oldalon több más neuronnal van összekapcsolva, a neuronok minden egyes konkrét organizációja gyakorlatilag ekvivalens egy hálózattal, úgy hogy az automata tulajdonságai a hálózat meghatározott tulajdonságainak felelnek meg. Mivel a hálózatok matematikai elmélete a gráf-elméletben meglehetősen jól ki van dolgozva, a neuronális automaták tulajdonságait ennek az elméletnek a segítségével szigorú matematikai módszerekkel tárgyalhatjuk. Ezt a lépést először N. Rashevsky²² és A. Rosen²³ tették meg. Kérdésfeltevésük még általánosabb mint a miénk; nemcsak a központi idegrendszert, hanem egy egész biológiai rendszert (organizmust) kísérlelnek meg elméletileg irányított gráffal reprezentálni. A metabolizmus esetében például a gráf útjainak irányai az anyagszállítás effektív irányait reprezentálhatják, amíg a sarokpontok aszerint, hogy milyen nívón szemléljük a rendszert, vagy makroszkopikus szerveket, vagy egyetlen sejtet is jelenthetnek. Ily módon biológiai rendszerekre vonatkozó általános elméleti tételekhez juthatunk, amelyeknek ismertetése azonban túlságosan messzire vezetne.

Gyakorlati szempontból az egyik legfontosabb feladat a neuronális automaták számára az alakfelismerés, különösen a kézzel vagy géppel írt

²⁰ Coates, C. L. und Fisch, E. A. : Design of a solid state neuron circuit for use in self-organizing system. 1960. International solid state circuits conference, 11–13 februar 1960, Philadelphia.

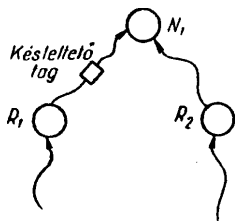
²¹ Braitenberg, V. : Morphology of nerve nets. Supplement No 2, Serie X del Nuovo Cimento.

²² Rashevsky, N. : General mathematical principles on biology. Transactions of the New-York Acad. of Sciences. Ser. II. Vol. 2, No 2 (Dec. 1958) pp 128–134; a dolgozatban további bibliográfia is található.

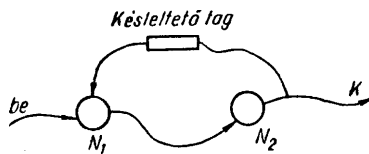
²³ Rosen R. : The representation of biological Systems from the standpoint of the theory of categories. Bull. of Mathem. Biophysics, Vol. 20 (1958), pp 317–341.

ábrák és betűk felismerése. Ilyen gép segítségével a modern adatfeldolgozó rendszerekben az átkódolás fáradságos munkáját mellőzhetjük, illetve a kézzel írott információk lyúkkártyákra való áttételét automatizálhatjuk. Elméletileg ehhez két funkció szükséges: először az elemi események térbeli vagy időbeli konfigurációinak *diszkriminációja*, illetve *klasszifikációja*, másodsor pedig a *memória*.

4. 1. *Diszkrimináció neuronális automatákban.* A diszkrimináció, illetve a klasszifikáció kérdését mind elméletileg, mind kísérletileg többen is vizsgálták. A. M. Uttley²⁴ azt feltételezi, hogy minden receptor csak két állapotban (aktív vagy inaktív) lehet, tehát vagy lead impulzusokat, vagy nem. (Binér receptorok). A térbeli klasszifikáció feladata lényegében annak a megállapítása, hogy a receptorok csoportjából melyek reagálnak egyidejűleg. Ehhez különleges asszociatív receptorok (koincidenca-neuronok) szükségesek, amelyek két-két, három-három stb. receptorból álló különböző kombinációk egyidejű reakcióját jelzik. Ha N közvetlen receptor (bemenő neuron) van, úgy, hogy minden lehető kombinációt jelezhessenek, összesen $(2^N - 1)$ neuron szükséges, amelyeket a lehetséges kombinációknak megfelelően *szisztematikusan előre kell összekötni*.



1. ábra



2. ábra

Az időbeni diszkrimináció esetében késleltető tagok szükségesek, amit a központi idegrendszerben az axonok mentén a különböző terjedési sebesség realizál. Az asszociatív neuronok részben *közvetlenül*, részben a késleltető tagokon keresztül közvetetten vannak a bemenő receptorokkal összekapcsolva, ahogyan ez az első ábrán két receptor egyszerű esetében látható. Az asszociatív neuron csak akkor válaszolhat, ha R_1 -t R_2 *előtt* ingerelték, mivel R_1 kimenete csak a késleltető tagon keresztül érheti el N_1 asszociatív neuront. Ennek a késleltetési elvnek természetesen komplexebb időbeni konfiguráció diszkriminálható. Az asszociatív neuronok módszerét a diszkriminációhoz W. K. Taylor is alkalmazta.

Uttley elméletileg legérdekesebb eredménye az, hogy térbeli, illetve időbeli klasszifikáció az asszociációs neuronoknak nemcsak preformált, szisztematikus (azaz tisztán logikai) kapcsolatai révén lehetséges, hanem abban az esetben is, ha a neuronok egymás között *véletlenszerűen* vannak összekapcsolva, ahogyan ez Sholl már említett anatómiai leleteinek megfelel. Ebben az eset-

²⁴ Uttley, A. M.: Temporal and spatial patterns in a conditional probability machine. Automata Studies, Princeton, 1956.

ben azonban az asszociatív neuron reakciója már nem egyértelmű, a diszkrimináció bizonytalanná válik.

Komplex események diszkriminációjához ami az alakfelismeréshez szükséges, ez az egyszerű mechanizmus önmagában még nem elégséges. A komplex eseményeket ugyanis az jellemzi, hogy *elemeik egymással meghatározott térbeli vagy időbeni kapcsolatokban állnak*, ami meghatározott információt jelent és éppen ez az információ fontos a komplexebb diszkriminációs folyamatok számára. Emellett a komplex eseményben tartalmazott információk, amint ezt a kísérleti tények mutatják, nem mind egyforma értékűek, és jó részüket mint *lényegtelen* elhagyhatjuk, anélkül, hogy az alakfelismerés lehetőségét lényegesen megzavarnók. A redundáns információknak ilyen kompressziója megy végbe például a retina és a látó cortex között.

A legegyszerűbb, ha a *relatív* idő- és intenzitás viszonyokból indulunk ki. A kísérleti eredményekből ugyanis kitűnik, hogy a központi idegrendszerben (különösen a látásnál) az inger kezdetének, illetve végének, tehát a változások jelzésére különleges neuronok szolgálnak. Ez összhangban áll a pszichológia tapasztalati eredményeivel is, ami szerint legalábbis a lényeges információkat nem egyedül az intenzitások, hanem inkább *relatív változásaik* reprezentálják, feltéve, hogy az ingerküszöböt meghaladják. Az ingerküszöb nagysága természetesen a mindenkori konkrét viszonyoktól függ és ezen kívül egyéni tényezők is meghatározzák.

A relatív intenzitásváltozások elvét kétféle módon alkalmazhatjuk a diszkriminációra, illetve az alakfelismerésre. Egyfelől tekintetbe vehetjük a *relatív intenzitásváltozások kölcsönös kapcsolatait* mint szemantikai információkat és, ha egy olvasó gépről van szó, megfelelő statisztika (például különböző kéziratok) kiértékelése útján minden egyes betű számára megállapíthatjuk a szükséges szemantikai információkat (mint például az A betű szárainak relatív hosszúsági arányait, görbületi viszonyokat stb.). Az így adódó konfigurációkat azután mintául későbbi összehasonlításokhoz tárolnunk kell. Ebben az esetben a helyzet hasonlít egy idegen nyelvű szó jelentésének szótárból való kikereséséhez. Azonban a folyamatot fiziológiailag jobban tudjuk a következő módon automatizálni, amint ezt legutóbb W. K. Taylor²⁵ tette: a relatív intenzitás-változásokat mint *részletet* definiálta. A közvetlen intenzitásokból ezeket a részleteket megfelelő elektronikus részletszűrőkön levezethetjük. Lényegesek definíciószerűen az olyan részletek, amelyeknél a relatív változás egy meghatározott küszöbértéket túlhalad, ezeket az összehasonlításhoz tárolják.

4. 2. *A memória problémája.* Amint már említettük, az alakfelismeréshez szükséges második funkció a memória. Az egyszerű tárolás funkcióját, ahogy a számoló automaták esetében is történik többféle módon is instrumentálhatjuk. McCulloch—Pitts, illetve Kleene neuron modelljének esetében, mivel az idő, feltevés szerint, egyforma intervallumokra van osztva, elegendő ha a kimenő vezeték megfelelő késleltető művonalon keresztül, ami az impulzust pontosan az időegység egész számú többszörösével késlelteti, visszavezetjük a bemenetre, amint ezt a második ábra bemutatja. Ha egy ilyen reverberatív lánc bemenetére egy impulzust vezetünk, úgy ez a késleltető vonalon végighaladva pontosan egy időegységgel később újból eljut a bemenetre. A folya-

²⁵ Taylor, W. K.: Pattern recognition by means of automatic analogue apparatus. Proc. J. E. E. Vol. 106. Pt. B, No 26 (März. 1959) pp 198—209.

mat ily módon határozatlan hosszú időn keresztül ismétlődik, amíg mesterségesen meg nem szakítják. Ha a lánc aktív, ez azt jelenti, hogy a bemenetre *valamikor a múltban* egyszer impulzust vezettek, — tehát az impulzust tárolja. Egy ilyen lánc, a késleltető művonal hosszától függően vagy egy impulzust, vagy egy hosszabb impulzus-sort is tárolhat. A központi idegrendszerben a késleltetést természetesen nem fizikailag elválasztott berendezések segítségével érjük el, mint a mesterséges automatákban, hanem úgy, hogy az axonok mentén a terjedési sebesség, illetve az egyes axonok hossza különböző. Ilyen reverberatív láncok létezését a központi idegrendszerben anatómiailag már régen bebizonyították és, ahogy Young²⁶ kísérletei az octopuson bebizonyították, legalábbis részben mint memória is funkcionálhatnak.

A tároló azonban, ahogyan a számoló automatákban alkalmazták, még nem *memória*. A különbséget rendszerint a „*plaszticitás*” szóban foglalják össze. Ha ezt egy kissé közelebről megvizsgáljuk, a következő eredményre jutunk: a mai számoló automatáink tárolóit a várakozási idő* szerint általában több fokozatban szervezik meg: az egyes fokozatok kapacitása annál kisebb, minél rövidebb a várakozási idő, amely kereken tíz mikromásodperctől több másodpercig is terjedhet. Hasonló lépcsőzetes szervezés figyelhető meg az emberi memóriánál is: az állandóan használt szavak rögtön eszünkbe jutnak, amíg a ritkábban használt szavakat rövidebb vagy hosszabb ideig „keresnünk kell” az emlékezetben. Amennyire azonban ma tudjuk, az emberi memóriának nincsenek különböző egységei, sőt az agyvelőben nem is lehet pontosan lokalizálni. Továbbá, a mai számoló automatáink memória-kapacitása még a legnagyobb berendezéseknél sem haladja meg a 10^8 bitet, ezzel szemben az emberi memória kapacitása több nagyságrendszel nagyobb mint mindaz, amit a mai, vagy a belátható jövő technikai eszközeivel meg tudunk valósítani: az emberi memória kapacitását különböző kutatók különböző módszerek segítségével 10^{15} — 10^{20} bitre becsülik.

A legnagyobb különbség azonban *minőségi*. Mai számoló automatáink tárolóit tisztá logikai módon, mereven kell megszervezni. A szelekció nem automatikus, hanem, amint a szakemberek mondják „címek” szerint történik. Ehhez járul még az a körülmény is, hogy a mechanikus tárolóba az információt egy munkaütem vagy *csak* beírja, vagy *teljesen* törli. A biológiai értelemben vett memória ebben a két vonatkozásban alapvetően másképpen funkcionál: *asszociációra képes*, aminek segítségével a szelekció *automatikusán, a tartalom alapján*, tehát egyértelmű cím megadása *nélkül történik*. A másik az a tény, hogy a memória tartalmakat nem lehet egyszerűen „beírni”, mint a tárolóknál, hanem *ismétlések útján fokozatosan vésődnek be*, azaz meg kell azokat tanulni. Ennek a két folyamatnak fiziológiai részletei még nagyon kevésbé ismertek. Alapos okunk van feltételezni, hogy az asszociációnál az *irradiáció effektusa* a tanuláznál pedig az *ismétlések által bekövetkező facilitálás* nagy szerepet játszik.

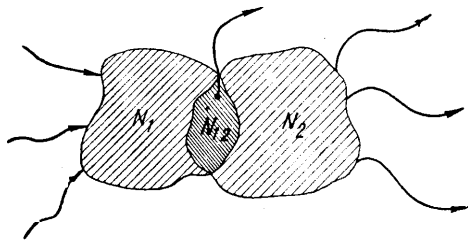
Így többé-kevésbé elfogadható munkahipotézisekre vagyunk utalva; ha azonban ezeket megfelelően választjuk ki, úgy a *logikai módszerekkel kombinálva* használható és technikailag is instrumentálható eredményekhez vezetnek. Az alakfelismerő gépek esetében ez azt jelenti, hogy a tárolt jegyekkel

²⁶ Young, J. Z. : Discrimination and learning in Octopus. Conference on Cybernetics, 1952. J. Macy Foundation.

* Várakozási időnek nevezik az elektronikus számológépeknél azt az időt, amire egy számnak a tárolóból való kikereséséhez vagy az oda történő elhelyezéséhez szükség van.

történő összehasonlítás előre kialakított, fixen beépített áramkörök segítségével történik, még abban az esetben is, ha a jegyeket, amikkel az összehasonlítás történik, maga a gép a reális fiziológiai viszonyokhoz hasonlóan automatikusan képezi, amint ez például W. K. Taylor (25) említett dolgozatában történik.

4. 3. *Irradiáció és alakfelismerés neuronális automatákban.* Ha az alakfelismerésnek biológiai értelemben véve kell automatikusan történnie, akkor a pavlovi értelemben vett irradiációra van szükség. Amint a már említett dolgozatban (18) megmutattuk, ehhez olyan kapcsoló-elemek szükségesek, amelyek tulajdonságait a 4. 2. pontban tárgyaltuk. Definíciószerűen felvesszünk egy N neuron-hálót, amely n számú olyan neuronból (kapcsoló-elemből) áll, amelyben minden egyes neuron minden más neuronnal szisztematikusan össze van kapcsolva. Mivel minden egyes neuronnak legalább egy bemenete van, és a kimenetek szétágazhatnak, a hálózatnak n bemenete és legalább



3. ábra

ugyanannyi kimenete van, amelyek útján a receptorokkal, illetve más hálózatokkal többszörösen is össze lehet kapcsolva. A gátlások megfelelő elosztása útján a hálózatból összesen $2^n - 1$ részhálózatot választhatunk ki, amelyek két-két, három-három stb. neuronból állanak; ezeknek mindegyike egy vagy több bit információt tárolhat: Minthogy a neuronok számát egyedül az agykéregben nagyságrendileg 10^{10} neuronra becsülik, ha az egyes neuronoknál csak *egyetlenegy* lehetséges küszöbértéket is tételezünk fel, kereken 10^{13} bit memóriakapacitást kapunk. Egy ilyen struktúra tehát a rendkívül nagy biológiai memóriakapacitást megközelítőleg megmagyarázhatja.

Megmutatható, hogy egy inger, amely egy N_1 részhálózatot ér, ezen keresztül más részhálózatokhoz is, pl. az N_2 , N_3 stb. részhálózatokhoz tovább jut; ezek azonban csak olyan mértékben válaszolnak, amilyen mértékben az inger a részhálózat *küszöbértékét* meghaladja. Az ingert tehát a hálózat primér ingerelt része diffúz módon más részhálózatokhoz is irradiálja, elsősorban olyan részhálózatok, illetve olyan irányok felé, amelyeknek küszöbértéke alacsonyabb, amit a mindenkori gátlások konfigurációja határoz meg.

Ez a struktúra, illetve az irradiáció képessége kézenfekvővé teszi, hogy *a külvilág egy objektumát fizikailag, illetve fiziológiailag olyan megfelelő nagyságú neuron hálózat segítségével reprezentáljuk*, amelyben az egyes jegyeket (elemi, illetve regurális eseményeket) megfelelő részhálózatok reprezentálnak. Vizsgáljunk meg két ilyen hálózatot, például N_1 és N_2 -t, amelyeknek egy közös N_{12} részhálózata van. Ha N_1 -t a megfelelő bemenetekenél az elemi eseményeknek meghatározott kombinációja ingerli, úgy az N_{12} közös részhálózat követ-

keztében N_2 hálózat is válaszol, éspedig annál erősebben, minél nagyobb az N_{12} közös részhálózat az N_1 és N_2 két hálózat összegéhez viszonyítva. Tehát az N_2 hálózatot a közös részhálózat *asszociatív* módon ingerli, ami fizikailag a közös asszociatív rész egy vagy több kimenetének aktivitásával jelezhető. Aszerint, hogy N_1 bemenetein az elemi események melyik konfigurációja van jelen, különböző részhálózatokat ér csak inger, amelyek a maguk részéről ismét más részhálózatokat ingerelhetnek. A mindenkori közös részhálózatok útján terjedő ingerlés révén a bemeneti ingerlés irradiálódik, amíg nem éri el a hálózatnak azokat a részeit, amelyek a szükséges memória tartalmakat tárolják. A szelekció tehát a közös asszociatív részhálózatokon keresztül automatikusan következik be, mintegy a tartalomnak megfelelően. Továbbá, ha N_1 és N_2 két hálózat a külvilág két objektumát fizikailag, illetve fiziológiailag reprezentálja, az N_{12} közös asszociatív részhálózat nagysága, illetve az ezzel arányos aktivitás megadja az általánosított értelemben vett hasonlóság mértékét, ami mind az asszociáció, mind az alakfelismerés számára rendkívül fontos.

5. A tanulás problémája

A legnagyobb nehézséget a logikai automatáknál a programozás szükségessége okozza: a logikai automaták *csak* olyan műveletet tudnak végrehajtani, amelyek programozhatók, illetve algoritmizálhatók. Sok gyakorlati esetben azonban sokkal egyszerűbb és bizonyos körülmények között célszerűbb is egy folyamatot *nem programozni* (mert ez túlságosan bonyolult lenne), hanem az automatának egyszerűen *megmutatni* és megtanultatni. A tanulás problémáját elvileg két egymással szemben álló módszerrel közelíthetjük meg. Az egyik tisztán fenomenológiai: *megkísérhetünk* egy matematikai modellt szerkeszteni, ami a makroszkopikus magatartást amennyire csak lehetséges visszatükrözi anélkül, hogy a fiziológiai részletekbe belemennénk. Ennek a modellnek alapján azután vagy programot készíthetünk egy számoló automata, vagy építhetünk egy speciális gépet. A másik módszer a tanulás fiziológiai alapjaiból (azaz a feltételes reflexekből) indul ki; megkísérli a neuron matematikai modelljének felállítását, illetve nem fiziológiai (legtöbbször elektronikus) eszközökkel instrumentálja és a mikroszkopikus folyamatokról a makroszkopikusok felé halad. Az első módszer a mi szempontunkból érdektelen; a következőkben csak a második módszer eredményeit ismertetjük. Az első módszer sok érdekes és fontos részlete miatt utalunk Bush és Mosteller ismert könyvére.²⁷

5. 1. *A tanulás és a feltételes reflexek.* Ismeretes, hogy a tanulás fiziológiai alapja a feltételes reflexek, amelyek amint elnevezésükből is következik a feltétlen reflexeken alapulnak. Ha csupán funkcionálisan tekintjük, a feltétlen reflex lényegében egy egyszerű, állandó struktúrájú szervomechanizmussal ekvivalens, a pillanatnyilag ható bemenő konfiguráció koincidenciája a korábbi specifikus ingerek alkalmas módon tartósan tárolt konfigurációjával, majdnem mechanikusan kiváltja a megfelelő reakciót. A feltételes reflex azért keletkezik, hogy a feltétlen reflexet kiváltó specifikus bemenő konfigurációt megelőzi egy másik, nem feltétlenül ugyanabból az érzékszervből származó, nem-specifikus konfiguráció. Az ismétlések folyamán a központi ideg-

²⁷ Bush, R. R. and Mosteller, F.: Stochastic models for learning, New-York, 1955

rendszerben többé vagy kevésbé tartós funkcionális kapcsolat létesül oly módon, hogyha az inger egy bizonyos *küszöbértéket* (amit természetesen egyéni tényezők is meghatároznak) meghalad, a reakciót a nem-specifikus konfiguráció egyedül is kiválthatja. Elméleti szempontból az ingerek (bemenő konfigurációk) természetesen azoknak a reguláris eseményeknek felelnek meg, amelyeket korábban definiáltunk. Néhány fontos momentumot, különösen a specifikus és nem-specifikus ingerek kölcsönös kapcsolatára vonatkozóan, a viszonyok gondos elemzése után W. Grey Walter²⁸ adott meg, aki az ismert *Machina Docilis*ben izolált feltételes reflexek instrumentálására technikailag funkcionáló, bár kissé körülményes módszert adott meg.

Az izolált feltételes reflexek önmagukban azonban még nem jelentik az intelligens tanulást. *Valamely elsajátított jártasság meghatározott külső ingerekre (események) jól definiált reakciósorozat* (műszaki szempontból: műveletek) *kivitelezésében jut kifejezésre, amelyek egy meghatározott cél elérésére irányulnak.* A sorozat irányát az a tény határozza meg, hogy a sorozat egyes tagjai szintén feltételes reflexek, amelyek *meghatározott időbeni kapcsolatban állanak egymással*: egy részreakció sikeres végrehajtása a következő részreakció kiváltó ingerévé lesz és a megelőző részreakciók sikerétől függ. Ennek a ténynek fontos következménye van: ha egy új feltételes reflexet akarunk kiépíteni, nem kell a sor kezdetére visszamennünk, hanem az új reflex ráépíthető a régiekre. Egy kissé pontatlanul, de plasztikusan azt mondhatjuk, hogy egy új reflex kiépítését megfelelő helyen valahol a sor közepén kezdhetjük el és így egy új elágazást kaphatunk, anélkül, hogy a sor előbbi tagjait érzékenyen megzavarnánk.

5. 2. *A tapasztalat szerepe.* A tanulás szisztematikus elméletének ezen kívül még sok fontos kísérleti tényről kell számot adnia; ezek közül a legfontosabb az, hogy a feltételes reflexeket *a siker (jutalom) megerősítheti, a sikertelenség (a büntetés) ezzel szemben gyengítheti, sőt teljesen eltörölheti.* Ebből a követelményből két következmény adódik, amelyek mind elméletileg, mind a gyakorlati instrumentálás számára is alapvető jelentőségűek.

Az első következmény az, hogy a neuronális automatának képesnek kell lennie arra, hogy *az általa végrehajtott műveletek sikerét, illetve sikertelenségét egy előre megadott normával való összehasonlítás révén* meg tudja állapítani, s ennek alapján az effektor szerveknek megfelelő vezérlő jelzéseket tudjon adni. Ezt sokféle módon, különleges eszközökkel vagy organizációs (logikai) eszközökkel instrumentálhatjuk. Ilyen összehasonlító szerveket kiterjedten alkalmaznak az ipari automatákban vezérlő jelzések létrehozására. A neuronális automatákban a vezérlő jelzések létrehozásán kívül még az a fontos funkciójuk is van, hogy (például a Willis (1. c.) által feltételezett mechanizmus útján) a helyes neuron-kapcsolatok képződését elősegítsék. A második, fontosabb következmény az, hogy a neuronális automatának, hogy tanulni tudjon, *a környezettel való állandó kölcsönhatás révén információkkal kell bírnia az általa végrehajtott műveletek eredményéről és ezeket a későbbi összehasonlításhoz tárolni kell tudnia.* A környezettel való állandó kölcsönhatás a negatív visszacsatolás (inverse feedback) formájában természetesen minden ipari szabályozó automatában megvan. Míg azonban az utóbbiak mindig csak a *pillanatnyi ingerekre és mindig ugyanolyan módon reagálnak*, minthogy a norma (referencia) egyszer és mindenkorra fizikailag be van építve, a neuronális

²⁸ Grey—Walter, W.: The living brain. London, 1953.

automatákban a művelet sikeréről szóló információkat egy *plastikus* memória tartalmával hasonlítjuk össze, amelynek tartalma az idők folyamán változik. A norma tehát mint *tárolt információ* van jelen. Ennek a körülménynek következtében a neuronális automatáknak kölcsönhatása a környezettel *funkcionálisan nyilvánvalóan ekvivalens a tapasztalattal*. A tapasztalat tehát nemcsak mint kísérleti tény, hanem mint *logikai szükségszerűség is adódik*, ahogyan ezt egyébként még a neuronális automaták elméletének létrejötte előtt a dialektikus materializmus állandóan állította.

5. 3. *A feltételes valószínűségek hipotézise*. Ezeket a követelményeket Uttley egy érdekes gondolat segítségével próbálja, legalábbis részben, teljesíteni. Abból a megfigyelésből indult ki, hogy a reakció még a jól kiépített feltételes reflexeknél is kicsi, de véges valószínűséggel kimaradhat, mivel a szervezet a specifikus inger fellépését, ami a reakciót tulajdonképpen kiváltja, tulajdonképpen *anticipálja*. Ezt *formálisan* úgy vehetjük tekintetbe, hogy a *specifikus inger fellépésének feltételes valószínűséget tulajdonítunk, azzal a feltétellel, hogy a nem-specifikus inger már fellépett*. Ha ez a feltételes valószínűség egy bizonyos küszöbértéket átlép, akkor történik meg az anticipáció. Így elméletileg ugyanaz a helyzet áll elő, amire a plauzibilis következtetéssel kapcsolatban már utaltunk. Uttley²⁹ valóban ezt tételezi fel és megvizsgálja az elméleti következményeket. Elméleti szempontból az a legfontosabb eredménye, hogy a bemenő kapcsoló szervek a magasabb szintű asszociációs szerveket nemcsak ingerelhetik, hanem gátolhatják is; így az egymás alá, illetve fölé rendelt egymást kölcsönösen ingerlő, illetve gátló kontrollszerveknek olyan hierarchiájához jut el, amely világosan a központi idegrendszer szervezésének az irányába mutat. A feltételes valószínűség nagysága, ami a halmaz-elméleti „közös rész” fogalmán alapul, mértékül szolgálhat az általánosított értelemben vett hasonlóság számára is. A feltételes valószínűségek hipotézise azért is érdekes, mert relatíve egyszerű eszközökkel instrumentálható, amit Uttley valóban meg is tett.²⁹ Ebben az utóbbi dolgozatban azt is bebizonyította, hogy a modell valóban meglehetősen jól megközelít sok fontos kísérleti tény, beleértve a kapcsolatok erősítését, illetve kioltását is. A megépített modell segítségével szerzett tapasztalatok — bizonyos megszorításokkal — megerősítik a várakozásokat, legalábbis olyan irányban, hogy az alapgondolat bonyolult folyamatokat megtanuló gépek (például egyszerű termelő folyamatok) felépítésére is felhasználható. Más változatban a gép bizonyos mértékig eltorzított betűket is felismerhet.

5. 4. *Önszervezés a tapasztalat útján*. Egy másik nagyon érdekes gondolat F. Rosenblatt³⁰ fízziológustól származik. A tervezési stádiumban levő és *Perceptron*-nak nevezett gép feladata a vizuális alakfelismerés, illetve a síkbeli ábrák diszkriminációja; ezért funkcionálisan az emlősök látási mechanizmusának organizációját utánozza.

A modell négy funkcionálisan különböző fő részből áll. Az optikai jelzések fotocellákból álló matrixra (receptor-mező) esnek, ami a retinának felel meg, ahonnan az elektromos jelzések egy *projekció mezőre* jutnak. Innen a jelzések egy *asszociációs mezőhöz* jutnak, amely két *rétgeből* áll: a projekciós

²⁹ Uttley, A. M.: Design of a conditional probability computer. Information and Control, Vol. 2. No 1 (1959. márc.) pp 1—24.

³⁰ Rosenblatt, F.: The perceptron, a probability computer for information storage and organization in the brain. Psychol. Rev. Vol. 65 (1958), pp 386—408.

mezőből a jelzések közvetlenül a felső réteghez jutnak el, amíg a továbbadandó kimenő jelzések az alsó rétegben jönnek létre és a *döntési mezőre* kerülnek, amely funkcionálisan a motor-cortexnek felel meg. Az asszociációs mező felső és alsó rétege között kölcsönös lokális visszacsatolások vannak; ezek tulajdonképpen a korábban említett reverberatív láncoknak felelnek meg s általuk a két réteg kölcsönösen befolyásolhatja egymást. A kimenő jelzéseken kívül a döntési mező visszafelé az asszociációs mezőhöz is ad jelzéseket, úgyhogy ezt egyfelől a projekciós mező, másfelől a döntési mező is befolyásolja.

Az egyes mezők eltekintve a fényérzékeny retinától, mesterséges neuronokból állanak, amelyeknek feltételezett tulajdonságai lényegében G. Willisenek a 3. 2. pontban leírt modelljével egyeznek; az egyetlen különbség az, hogy nem a küszöbérték, hanem a kimenő jelzés függ az ingerek és gátlások összegétől, de könnyen belátható, hogy a két feltevés funkcionálisan ekvivalens.

Az alapvető hipotézis mármint a következő: egyfelől a projekciós mező neuronjai és az asszociációs mező felső rétege között, másfelől az asszociációs mező alsó rétege és a döntési mező között a kapcsolatot kezdetben teljesen véletlenszerű, *azaz a rendszernek kezdetben egyáltalán nincsen organizációja*. A kísérletek folyamán, (amelyek egyszerű síkbeli ábrák bemutatásában állanak) a döntési mezőn keresztül jelzések jutnak visszafelé az asszociációs mezőhöz, olyan hatással, amely a véletlenül kialakult helyes (sikeres) kapcsolatokat megerősíti, a helyteleneket ellenben fokozatosan gyengíti. Mivel az asszociációs mező két rétege között lokális reverberatív láncok vannak, ezeket a kapcsolatokat a rendszer tárolja is. *Ily módon a kezdetben szervezetlen hálózat a kísérletek, azaz a tapasztalat során a megmutatott ábrának megfelelő módon fokozatosan önmagát szervezi meg.*

Amint említettük, a perceptron még csak tervezési stádiumban van. Tulajdonképpen programokból áll, amelyeket egy IBM 704 elektronikus számológép részére készítettek. Ennek a programnak segítségével a modell viselkedését utánozzák a számoló automatán. A különböző kísérleti feltételeknek megfelelően végrehajtott számítások elég biztató eredményeket mutattak és további finomításokra adtak utalást; ennek alapján legújában egy instrumentált kísérleti modell megépítéséről is beszámoltak. Az eredmények még nem ismeretesek; a tapasztalat műszaki értelemben vett organizáló hatásának alapgondolata azonban mindenestre nagyon figyelemre méltó és további vizsgálatokat érdemel.

5. 5. *Az elmélet kibővítésének szükségessége.* Habár Uttley feltételes valószínűségi hipotézise, valamint Rosenblatt hipotézise is gyakorlatilag használható *részeredményeket* ad, a valóságnak mégis csak egy kis részét tükrözik, amint ez másképpen nem is várható. Ha a valóságot az elméletben (és későbbben a gyakorlatban is) jobban vissza akarjuk tükrözni, úgy az 5. 1. és 5. 2.-ből adódó következményeket valahogyan be kell illeszteni az elméletbe. Ez pl. olyan módon történhet meg, hogy G. Willis neuron modelljének hipotéziseit és a szerző hipotéziseit úgy egyesítjük, hogy a két lényeges tulajdonságot (monostabilitás és a küszöbérték emelésének, illetve csökkentésének lehetőségét) egyidejűleg posztuláljuk; a kimenő jelzésekről az egyszerűség kedvéért feltételezhetjük, hogy arányosan oszlanak meg a kimenő vezeték egyes elágazásaira. Ily módon a neuron hálózatok tulajdonságait a lineáris algebra segítségével bizonyos egyszerű esetekben kvantitatív módon is tanulmányozhatjuk.

Igy olyan hálózatokhoz jutunk, amelyekben az egyes neuronok között automatikusan tartós kapcsolatokat hozhatunk létre, vagy a keletkezett kapcsolatokat ismételt gátlások útján kiolthatjuk, és amelyek elvileg intenzitásokra is érzékenyek. A „részleteket”, vagyis a primér intenzitások relatív változásait, ahogy azok W. K. Taylornál előfordulnak, relatíve egyszerű szervezési eszközök segítségével kiszűrhetjük, és mint a lényeges információk egy részét a primér receptoroktól tovább vihetjük. Bizonyos, a pontos fogalom meghatározás itt még jelentős nehézségekbe ütközik (például a kölcsönös relációk problémája), amelyek még megoldásra várnak. Ha még hozzá vesszük a Rosenblatt által ajánlott önszervezés érdekes módszerét is, olyan rendszert kapunk, amely az intenzitásokra érzékeny és amely a tapasztalat organizáló hatása révén tanulni is tud és ezáltal optimálisan alkalmazkodni tud az állandóan változó külvilághoz.

5. 6. *A reális fiziológiai viszonyokhoz való közeledés következményei.* Végül még utalni kell a neuronális automaták két olyan fontos tulajdonságára, amelyek a logikai automatákkal ellentétben a reális fiziológiai viszonyokhoz való közeledés következményeként lépnek fel. Egy logikai automata vagy jól vagy rosszul dolgozik. Ha a hibás program lehetőségét mint triviális esetet kizárjuk, a hibás működés mindig azt jelenti, hogy az automatának egyik alkatrésze meghibásodott. Ezzel szemben egy neuronális automata *tévedhet* is, azaz elkövethet hibákat akkor is, ha műszaki értelemben egyik alkatrésze sem hibás. Ez részben összefügg a neuronok feltételezett tulajdonságaival. Az éles különbség a nulla és az egy között, amelyen a tiszta logikai automaták működése alapszik, itt elmosódik. Az intenzitás változásokon alapuló diszkrimináció, különösen *nagy változások kis különbségei esetében* tisztán fizikai okok miatt sem történhet olyan biztosan, mint a logikai automatáknál, ahol az intenzitásvizonyok nem játszanak szerepet.

Egy másik, már itt is megmutatkozó fontos következmény az, hogyha a felismerendő alakok jellemző jeleit preformálva nem építik be a gépbe előre, hanem maga a gép állítja elő, akkor a felismerés tulajdonképpeni aktusa előtt a mintát néhányszor meg kell mutatni a gépnek, hogy a jeleket elő tudja állítani és meg tudja tanulni. Ez azonban azt jelenti, hogy a logikai automatákkal ellentétben, a neuronális automatákat előzetesen nem *programozni*, hanem *kondicionálni* kell.

6. *Befejező megjegyzések*

Amint már előljáróban megjegyeztük ez a tanulmány szükségképpen egyoldalú; főképpen a műszaki kutatók szempontjait tükrözi, bár törekedtünk utalni az elmélet nehézségeire is. Nem meglepő, hogy nehézségek vannak: *a gondolkodás mint filozófiai probléma* az emberiség legjobb koponyáit több mint háromezer év óta foglalkoztatja, *de a gondolkodást mint természettudományos vagy éppen mint műszaki problémát* csak az utóbbi tíz évben kezdték tanulmányozni, miután létrejöttek a szükséges, elsősorban elektronikus segéd-eszközök és kísérleti módszerek. Az eddigi elméleti és kísérleti tények mindenestre nagyon figyelemre méltóak.

A filozófiai általánosítás legbiztosabb bázisa mindenkor a természettudomány volt. Maga az a körülmény, hogy a neuronális automaták elmélete (és fokozatosan gyakorlata is) keletkezőben van, ömögában is már megerősíti a filozófiai materializmus álláspontját, ahogyan azt Marx a Szent Családban

megfogalmazta: „Nem választhatjuk el a gondolatot a gondolkodó anyagtól”. Valóban, a neuronális automaták elméletében még kétségtelenül meglevő nehézségek és ellentmondások ellenére filozófiai szempontból az a körülmény a legjellemzőbb, hogy az a gondolkodást mint objektív, inherensen materiális folyamatot szemléli, ami két alapon nyugszik: először az egyes neuronokban végbemenő komplex fizikai-kémiai folyamatokon és másodsor ezek kölcsönhatásán, vagyis azon, hogy hogyan szerveződnek a neuronok a központi idegrendszerben, különösen azonban az agyban egészzé. Mindkettőben jelentős része van az objektív külvilággal való állandó kölcsönhatásnak.

De az idealizmus képviselői is megkísérik a maguk módján filozófiailag általánosítani a természettudományok eredményeit. Amint Lenin „A harcoss materializmus jelentőségéről” című ismert munkájában rámutatott: „Éppen a modern természettudomány napjainkban végbemenő gyökeres átalakulásából születnek lépten-nyomon reakciós filozófiai iskolák és iskolácskák.” (Lenin, Művei 33. köt. 225. o.) Napjainkban ez még inkább igaz, mint Lenin idején, mert azóta a természettudományok fejlődési üteme példátlan módon megyorsult. Számos példát találunk arra, hogy a természettudományok, különösen a fizika eredményeit idealista módon interpretálják.³¹ Hűen a lenini örökséghez, a marxista filozófusok előtt az a feladat *is* áll, hogy a természettudományok más eredményei mellett a neuronális automaták elméletének és gyakorlatának eredményei *is* filozófiailag általánosítsák.

Befejezésül még egy megjegyzés: ma még nincsenek ilyen automatáink. Még nem is tudjuk pontosan, hogy mit kezdenénk velük, milyen célokra alkalmazhatjuk ezeket sikerrel. De akár filozófus, akár műszaki kutató az ember, soha nem szabad elfelejtenie, hogy a célt, amire minden gépet vagy automatát felhasználunk, végül is az azt felhasználó társadalmi rend határozza meg. A neuronális automaták, ahogy műszakilag tökéletesednek, szolgálhatnak, csak úgy, mint az atommag energiája, az emberiség javára, de egy automatizált nukleáris háború esetén kárára is. Így nemcsak az a fontos *elméleti feladat* áll előttünk, hogy a neuronális automaták elméletének eredményeit filozófiailag általánosítsuk, hanem az a nem kevésbé fontos *gyakorlati feladat* is, hogy mindent elkövessünk, hogy az ilyen automatákat az emberiség javára és *csak* is a javára használják fel.

НЕВРОНАЛЬНЫЕ АВТОМАТЫ

Режé Тарян

В статье дается обзор результатов абстрактной теории автоматов тесно связанной с важной проблемой философии — мышлением. В введении автором показано, что, по сути дела, вопрос не в том, способны-ли машины к мышлению, или нет, но скорее в том, какие умственные работы могут быть осуществлены автоматами. В статье дается определение абстрактного понятия дигитального автомата и обсуждается главный результат абстрактной теории конечных дискретных автоматов. Согласно этому такие автоматы могут осуществлять все типы операции, на которые можно дать однозначные правила. Однако, проблемы таких важных процессов, типичных для творческого мышления, как проблемы вероятных заключений, познания схем и учения, являются процессами, для которых неизвестны, по крайней мере в настоящее время, еще никакие правила. Согласно этому для решения проблем такого типа вместо конечных дискретных автома-

³¹ *Rosenblatt, F. : Perceptron stimulation experiments. Proc. IRE, März 1960 pp. 301—309.*

тов, представляющих собой, по существу, логические машины, нужны нейрональные автоматы, т. е. автоматы, состоящие из элементов, подобных нейронам.

Вторая часть статьи посвящена проблемам и результатам теории собственных нейрональных автоматов. Автором обсуждаются необходимые понятия и различные теоретические модели нейрона. После этого разбирается вопрос о свойствах нейрональных сетей, в первую очередь проблема дискриминации и роль пластичной памяти, нужных для познания схем. Автор показывает, как сеть искусственных нейронов может излучать и накапливать раздражения в такой форме, чтобы накопленная информация могла быть восстановлена процессом, аналогичным ассоциации. Потом обсуждается проблема учения. Показывается, что учение представляет собой управляемую цепь условных рефлексов, каждый из которых обусловлен удачей или неудачей предыдущих рефлексов, и может быть усилен или затухан опытом, имеющим функцию организатора. Между прочим, повидимому, два главных метода дают положительные результаты: метод условных вероятностей, предлагаемый Атли (Uttley) и перцептрон Розенблата, последний из которых является примером само-организующей системы. Из полученных до сих пор результатов расширение теории кажется необходимым. Важным последствием подхода к реальным физиологическим условиям будет то, что нейрональные автоматы вместо программирования должны будут подвергаться кондиционировке. Поскольку они гибкие, они могут не только неисправно функционировать, но и совершать ошибки. Наконец, показывается необходимость философского обобщения на основе диалектического материализма полученных до сих пор результатов.

NEURONAL AUTOMATS

Rezső Tarján

The paper surveys the result of the abstract theory of automata which is closely connected with the philosophically important problem of thinking. Introductory it is shown that, the correct question is not, whether machines can or cannot think, but rather what type of mental work can be performed by automata. The abstract notion of a digital automaton is defined and the main result of the abstract theory of finite discrete automata is discussed. According to this such automata can perform every type of operations for which unambiguous rules can be given. Problems of plausible inference, pattern recognition and learning however, which are important processes typical of creative thinking, are processes for which at least as yet no unique rules are known. Accordingly, instead of finite discrete automata, which are essential logical machines for such type of problems neuronal automata, that is automata consisting of neuron-like elements are necessary.

The second part of the paper deals with the problems and results of the theory of neuronal automata proper. The necessary concepts and the different theoretical models of the neuron are discussed. After these the properties of neuronal nets are dealt with, especially the problem of discrimination and the role of a plastic memory, necessary for pattern recognition. It is shown how a net of artificial neurons is capable to irradiate and store stimuli in such a fashion that, the information stored can be retrieved by a process similar to the association. Then, the problem of learning is discussed: it is shown that, learning consists of a directed chain of conditional reflexes, each being dependent on the succes, resp. failure of the previous ones and can be reinforced, resp. extinguished by the experience, which has the function of an organiser. Among others, two main methods seem to furnish promising results: the method of conditional probabilities proposed by Uttley and the perceptron of Rosenblatt, the later being an example of a self-organising system. From the results achieved so far, the extension of the theory seems to be necessary; an important consequence of the approximation of the real physiological conditions will be that, neuronal automata, instead of being programmed, must be conditioned. Flexible as they are, they can not only malfunction, but also make mistakes. Finally the necessity of the philosophical generalisation of the results achieved so far on the ground of dialectic materialism is pointed out.