

**Ústav aplikovanej informatiky  
Fakulta informatiky a informačných technológií  
STU v Bratislave**

**Vladimír Kvasnička**

**a**

**Ľubomír Varga**

# **Kognitívna veda pre informatikov**

**(predbežný text)**

**Bratislava 2011**

# Obsah

1. prednáška - Úvod do kognitívnej vedy
2. prednáška - Logické neuróny McCullocha a Pittsa
3. prednáška - Darwinovská evolúcia ako algoritmus
4. prednáška - Úloha učenia v evolúcii a memetika
5. prednáška - Učenie s odmenou a trestom a emergencia stratégie hry
6. prednáška - Holografická redukovaná reprezentácia pamäti
7. prednáška - Mentálne modely logiky v kognitívnej vede
8. prednáška - Teória argumentácie
9. prednáška - Revízia poznatkov
10. prednáška - Konštrukcia diagnózy chybného správania sa zložitých systémov
11. prednáška - Kognitívna logika a nemonotónne usudzovanie
12. prednáška - Riadenia zložitých systémov a výpočtové modely okolia
13. prednáška - Symbolická teória Hofstadterovho systému *Copycat* pre tvorbu analógie
14. prednáška - Logický model hry Sudoku
15. prednáška - Teória ohraničenej racionality

# Úvod

Tento predbežný učebný text k prednáške „Kognitívna veda“ bol tvorený za pochodu, keď v akad. r. 1997/98 bola zahájená táto výberová prednáška v inžinierskom štúdiu v rozsahu 2 prednášky a 2 hodiny cvičenia. Na webovskej stránke prednášky <http://www2.fiit.stuba.sk/~kvasnicka/CognitiveScience> každá prednáška má svoju prezentáciu (priesvitky) a riešené príklady na precvičenie preberanej látky z danej prednášky. ***Očakávam spätnú väzbu od študentov tým, že nás upozornia na prípadné chyby, nedôslednosti a preklepy v predkladanom texte (za ččo budú hodnotení pomocnými bodmi, ktoré sa započítajú do celkového záverečného hodnotenia).***

Niekoľko všeobecných poznámok k obsahu predkladaného textu. Niektorým čitateľom sa môže zdať veľmi inforaticky orientovaný na umelú inteligenciu, neurónové siete a evolučné algoritmy. Dôvod pre túto orientáciu je veľmi jednoduchý, prednáška je určená pre študentov inžinierskeho (magisterského) štúdia informatiky. Z týchto dôvodov boli zredukované všeobecné „filozofujúce“ partie kognitívnej vedy na minimum. V prezentovanom poňatí je kognitívna veda vlastne totožná s umelou inteligenciou, kde navrhované modely kognitívnych procesov sú mandatorne implementované na neurálnej úrovni (t. j. pomocou neurónových sietí).

U čitateľov (študentov) sa predpokladajú už určité základné vedomosti z matematickej logiky, umelej inteligencie, neurónových sietí a evolučných algoritmov. Pretože posledné dve uvedené oblasti sa prednášajú až v inžinierskom štúdiu ako nepovinné prednášky, snažili sme kapitoly, kde sa využívajú tieto moderné koncepcie (tzv. výpočtovej inteligencie) písať tak, aby boli zrozumiteľné aj pre absolventov bakalárskeho štúdia, ktorí (ešte) neabsolvovali výberové prednášky z týchto dvoch oblastí.

Predbežný učebný text obsahuje 15 prednášok, pričom štandardný semester obsahuje 10-11 týždňov. Z tohto pohľadu sú niektoré kapitoly nadbitočné. Väčší počet sme zvolili preto, aby prednášajúci v spolupráci so študentmi mal možnosť vybrať tie prednášky -kapitoly, ktoré sa javia v danom ročníku najviac zaujímavé. Ďalšia možnosť je použiť neprednesené kapitoly ako predmet prípadových štúdií - esejí, ktorých zaujímavé a tvorivé vypracovanie bude slúžiť ako podklad pre alternatívne absolvovanie skúšky .

August – september 2011

Vlado Kvasnička a Ľubo Varga

# 1. prednáška

## Úvod do kognitívnej vedy

---

### 1.1 Úvodné poznámky

Kognitívna veda je charakterizovaná ako teória mysle, ktorá metódami modernej vedy skúma fenomén ľudskej mysli. Používanie moderných, tak teoretických (logika, informatika, teória neurónových sietí, a pod.) ako aj experimentálnych (menovite použitie rôznych typov nukleárnej magnetickej rezonancie k zobrazeniu aktivít mozgu) metód kvantitatívnej psychológie k výskumu kognitívnych aktivít ľudí ju odlišuje od „klasického“ špekulatívno-esejistického prístupu, od „ad-hoc“ používania fenomenologických pojmov a koncepcií, ktoré nemajú oporu v súčasnej neurovede.

Kognitívna veda má silne interdisciplinárny charakter medzi filozofiou, psychológiou, neurovedou a umelou inteligenciou. Menovite umelá inteligencia má pre kognitívnu vedu zásadný význam, ktorej metódy (pojmy a koncepcie) sa prenášajú do kognitívnej vedy a tvoria jej teoretický základ. Môžeme si položiť otázku, aký je rozdiel medzi umelou inteligenciou a kognitívnou vedou. Na túto otázku možno jednoducho odpovedať tak, že modely (prístupy, koncepcie a pod) kognitívnej vedy musia byť kompatibilné s predstavami neurovedy, t. j. musia byť založené na jej základnej paradigme, že vysvetlenie aktivít ľudského mozgu musí byť „zakotvené“ na „neurónovom substráte“ (obvykle reprezentovanom neurónovými sieťami). Táto vlastnosť „neurónového zakotvenia“ sa nepožaduje od prístupov umelej inteligencie, kde je irelevantná podmienka realizácie kognitívnych aktivít ľudského mozgu. Jej snahou je zostrojiť formálny model pre vysvetlenie a modelovanie kognitívnych aktivít (napr. usudzovania, riešenia problémov a pod.), pričom požiadavka ich „neurónovej“ realizácie je (nemusí, ale obvykle je) ignorovaná.

Niekoľko poznámok k histórii kognitívnej vedy, ktorá má spoločné počiatky s umelou inteligenciou. Za jednu z prvých prác z kognitívnej vedy a z umelej inteligencie sa pokladá publikácia amerických vedcov Warena McCullocha a Waltera Pittsa z r. 1943 s názvom „A Logical Calculus of the Ideas Immanent in Nervous Activity“ v časopise *Bulletin of Mathematical Biophysics*. V tomto pionierskom článku formulovali základy teórie neurónových sietí s použitím binárnych logických neurónov. Dokázali univerzálnu vlastnosť týchto sietí, podľa ktorej každá Boolova funkcia môže byť vyjadrená pomocou neurónovej siete s binárnymi logickými neurónmi.

V súčasnosti je kognitívna veda uznávanou vednou disciplínou, má svoje stavovské organizácie, konferencie, časopisy, na mnohých významných univerzitách sú zavedené programy štúdia tejto vednej disciplíny, existujú univerzitné ústavy kognitívnej vedy. Kognitívna veda spoluexistuje veľmi efektívne s umelou inteligenciou, kognitívnou psychológiou a kognitívnou neurovedou.

### 1.2 Filozofia a kognitívna veda

Interakcia medzi kognitívnou vedou a filozofiou je veľmi bohatá a plodná, pretože problém mysle je predmetom skúmania filozofie už od čias gréckeho filozofa Aristotela., ktorý v *Etike*

špecifikoval ľudské bytosti ako *racionálne zvieratá*, pretože má telo s fyzickými potrebami a nutnosťou prijímať potravu a sú racionálnou bytosťou, pretože majú dušu, ktorej časťou je pamäť a schopnosť myslieť a usudzovať (t. j. to, čo môžeme nazvať myseľ). Cieľom tejto bytosti je dosiahnuť šťastie tak, že naplní zmysel svojej existencie. Podobne, ako zmyslom existencie noža je, aby krájala a zmyslom existencie očí je, aby videli, tak zmyslom ľudskej existencie – racionálneho zvieratá je, aby používalo svoju myseľ k usudzovaniu, t. j. správať sa racionálne pri vyhodnocovaní svojich životných skúseností. V tomto momente sa dostávame ku kľúčovému momentu gréckej filozofie, ktorá považovala za jeden z podstatných prejavov racionality ľudských bytostí schopnosť usudzovať pomocou zákonov logiky, ktorej základy boli taktiež vytvorené v starom Grécku.

Francúzsky filozof z prvej polovice 17. storočia René Descartes, vychádzajúci z dobových mechanistických predstáv o prírode, formuloval dualistickú teóriu ľudských bytostí. Aj keď príroda, teda aj ľudské telo, je materiálneho charakteru a je teda ovládané základnými princípmi mechaniky, ľudské bytosti sú v tom zvláštne, že sú kombináciou mentálnych substancií (myseľ) a materiálnych substancií (telo). Táto myšlienka tvorí základ Descartesovho *dualizmu*, ktorý môže byť preformulovaný do súčasného filozofického jazyka tak, že ľudia majú myseľ a telo. Základný problém spočívajúci v tejto paradigme je, akým spôsobom môže interagovať nehmotná myseľ s materiálnym telom.

Materialistický (fyzikalistický) pohľad na tento problém Descartesovho dualizmu je, že aj myseľ je materiálnej podstaty. Musíme však poznamenať, že aj tento materialistický pohľad na myseľ vyvoláva následné problémy. Fyzikálna myseľ je charakterizovaná špeciálnymi vlastnosťami, akými sú intencionalita (zámernosť alebo cielenosť) a vedomie, potom sa môžeme pýtať akým spôsobom vznikli tieto dve vlastnosti ľudskej mysle. K tomuto problému sa ešte vrátíme v ďalšej časti tejto prednášky.

Ďalší dôležitý problém filozofie mysle je otázka, čo to je myseľ, ako je odlišená od ostatných častí človeka, ako je štruktúrovaná a pod. Táto otázka má dva aspekty, jeden pochádza z filozofických snáh riešiť problém vzťahu medzi racionalizmom a empirizmom, ktorý bol veľmi aktuálny v sedemnástom a osemnástom storočí. Racionalizmus a empirizmus sú rôzne pohľady na podstatu a zdroje ľudských vedomostí. Podľa empirikov všetky naše poznatky pochádzajú z našich zmyslov, z empirických alebo experimentálnych interakcií s okolitým svetom. Racionalisti, naopak tvrdia, že existujú určité poznatky, ktoré nie sú odvoditeľné z empirie. Klasický racionalisti, akými sú Descartes alebo Leibniz, ale hlavne Kant tvrdia, že niektoré idey, napr. idea Boha alebo celých prirodzených čísel, nevyplývajú z našej experimentálnej skúsenosti, ale sú vrodenu súčasťou našej mysle. Empirici sa snažili v priebehu ostatných tristo rokov dokázať, ako aj takéto abstraktné idey – poznatky môže vyplývať z našej experimentálnej skúsenosti.

Rôzny pohľad racionalistov a empirikov na zdroje našich poznatkov spôsobil aj ich rôzny pohľad na myseľ. Podľa empirikov myseľ obsahuje len také idey – poznatky, ktoré sú primárne odvodené z našich zmyslov (včítane aj veľmi abstraktných a zložitých idey, akou je idea Boha). Z tejto paradigmy taktiež vyplýva, že myseľ nie je nič iné, ako idey odvoditeľné z pocitov pochádzajúcich z našich zmyslov. Potom kognitívne procesy sú založené na procesoch vnímania, t. j. kognícia a vnímanie sú v podstate identické pojmy, ktoré sa odlišujú len v tom, že kognícia je špecifický sled vnímaní. Podľa škótskeho filozofa Davida Humea, významného predstaviteľa empirizmu a osvietenstva, kognitívne procesy v našej mysli sú riadené asociáciami a podobnosťou. Myseľ je všeobecné zariadenie – orgán, kde všetky typy kognície sú realizované pomocou našej empirickej skúsenosti prostriedkami asociácie a analógie.

Racionalisti, tým, že neakceptujú paradigmu o zdroji našich poznatkov z našich pocitov, musia hľadať alternatívnu odpoveď na otázku, čo je zdrojom našich poznatkov. Tento problém vyriešili tak, že za zdroj našich poznatkov pokladajú vrodenosť za hlavný

zdroj poznatkov a štruktúry našej mysle. Ak idey sú vrodene (a teda nemusia byť odvodená z našej skúsenosti), potom štruktúra našej mysle je vrodene a tak vlastne aj ohraničuje použiteľnosť mysle pri spracovaní našich pocitov. Moderný pohľad na tento problém vrodeneosti (nativizmu) je, že len určité schopnosti našej mysle (napr. jazykové schopnosti, inštinkty a pod.) sú nám vrodene a preto sa nemusia učiť. Dôležitý aspekt tohto moderného prístupu k mysli je aj problém jej modularity, t. j. ľudská myseľ je chápaná ako systém navzájom poprepájaných modulov, z ktorých každý vykonáva určité špecifické kognitívne akcie a ktorých štruktúra je vrodene.

Vzťah medzi racionalizmom a empirizmom obsahuje aj ďalší aspekt mysle, a to jej štruktúru a miesto vedomia v nej. Problém vedomia patrí medzi najkomplikovanejšie záležitosti filozofie mysle a taktiež aj medzi tie vlastnosti mysle, ktorá „vzrušuje“ nielen filozofov a psychológov, ale aj informatikov aktívnych v umelej inteligencii a kognitívnej vede. Na čom sa všetci dohodnú v tejto zložitej oblasti je, že vedomie je integrálnou súčasťou mysle, ale taktiež aj na tom, že patrí medzi najzáhadnejšie vlastnosti mozgu. V súčasnej filozofii mysle existuje všeobecne akceptovaný názor, že mentálne stavy vedomia sú integrálnou súčasťou mysle dva krajné názory na vedomie. Každá mentálna (alebo kognitívna) aktivita mozgu je prejavom vedomia alebo je dostupná vedomiu (snáď môžeme povedať, že je jeho súčasťou). Ak však akceptujeme „komputačnú“ paradigmu podľa ktorej kognitívne procesy prebiehajúce v mozgu nie sú nič iné ako informačné procesy prebiehajúce v počítači – myslí, tak potom určitá časť týchto informačných procesov nie je dostupná vedomiu (napr. predspracovanie vstupnej vizuálnej informácie zo sietnice oka). Z tohto možno usudzovať, že predstava o tom, že každá mentálna aktivita je dostupná vedomiu je prehnaná, len určité mentálne aktivity sú dostupné nášmu vedomiu.

### 1.2.1 Myseľ v kognitívnej vede

Vzťah medzi mentálnym a fyzikálnym je centrálnym problémom filozofie mysle. Tento vzťah je aj základným problémom súčasnej kognitívnej vedy. Spôsob riešenia tohto vzťahu v kognitívnej vede sa nazýva funkcionalizmus, podľa ktorého mentálne stavy nie sú ekvivalentné s fyzikálnymi štruktúrami ktoré ich produkujú, ale prejavom-vlastnosťou ich usporiadania a kauzálnych súvislostí medzi nimi. Funkcionalizmus, ktorý vyjadruje určitú predstavu o vzájomnom vzťahu medzi mentálnym a fyzikálnym, je kompatibilný s materializmom (fyzikalizmom), myseľ (mentálne stavy) je prejavom materiálnych – fyzikálnych štruktúr, ktoré sú súčasťou tela a tvoria materiálny substrát (vehikulum) mysle.

Klasický príklad toho, čo nie je identické s tým z čoho je to urobené, sú peniaze. Peniaze nie sú totožné s papierom na ktorom sú vytlačené, ale majú určitú funkcionálnu úlohu v danom ekonomickom systéme. Materiálna reprezentácia peňazí nie je to podstatné, čo ich špecifikuje, či sú alebo nie sú peniaze; podstatné pre ne, je ich nemateriálna - funkcionálna úloha, ktorú hrajú v ekonomických systémoch ako platidlo, ktoré v niektorých prípadoch je reprezentované bankovkami, ale v prípade e-bankingu je reprezentované alfanumerickým reťazcom, ktorý sa prenáša internetom.

Dôležitý dôsledok tohto funkcionálneho pohľadu na myseľ spočíva v tom, že funkcionalistická paradigma umožňuje uvažovať aj nebiologické formy inteligencie. Pretože myseľ je „forma“ a nie „hmota“, ktorá špecifikuje psychické vlastnosti nejakého individua. To znamená, že z funkcionalistickej paradigmy vyplýva aspoň teoretická možnosť existencie umelej kreatúry (robot alebo počítač) s mentálnymi stavmi, ak majú tú správnu funkcionálnu štruktúru – organizáciu. Táto idea má ústredné postavenie v tradičnej umelej inteligencii, kde hlavnou víziou tejto informatickej oblasti je implementácia programu s takou algoritmickou štruktúrou, že umožňuje správať sa systému inteligentným spôsobom, ktorý je neodlíšiteľný od správania sa ľudskej bytosti. Tento redukcionistický pohľad na ľudskú inteligenciu je

v súčasnosti kritizovaný mnohými filozofmi a informatikmi, medzi najznámejšieho kritika patrí americký filozof John Searle, ktorý na jednoduchom príklade „čínskej izby“ ukazuje ako taká špičková kognitívna aktivita akou je preklad textu z čínštiny do angličtiny môže byť mechanicky realizovaná človekom, ktorý nevie po čínsky, ale má k dispozícii prekladové slovníky z čínštiny do angličtiny a text mechanicky prekladá znak po znaku, bez uvedomenia, akú činnosť vlastne vykonáva. Sumarizujúc, umelé inteligentné systémy nebudú mať vedomie a preto nebudú úplne rovnocenné ľudskej myšli. .

### 1.3 Kognitívna veda a psychológia

Psychológia je veda, ktorá sa zaoberá reprezentáciou a spracovaním informácie živočíchmi s vyššou nervovou aktivitou (a teda aj ľuďmi). Tieto sú schopné prijímať informáciu z okolia, vytvárať jej vnútornú reprezentáciu a transformovať ju s cieľom výberu a výkonu akcií. Inteligentné informačné procesy sú vyjadrené schopnosťou získavať a spracovávať informáciu z okolia, s cieľom vykonávania adaptívne akcií, ktoré zvyšujú schopnosť organizmu prežívať a rozmnožovať sa. V tejto súvislosti je dôležitý pojem „adaptívna akcia“, čo znamená, že proces vytvárania akcie je výsledkom učenia organizmu v danom prostredí s cieľom optimalizácie danej akcie. Ústredným problémom študovaným psychológiou sú informačné procesy, ktoré prebiehajú medzi vstupom informácie prostredníctvom zmyslov a výstupom informácie prostredníctvom motorických centier. Zložité formy inteligencie, ktoré sú pozorované u humanoidov, vyžadujú teórie mechanizmov výkonu kognitívnych akcií a vnútornej skúsenosti (existencia pamäti a schopnosti vytvárania vnútorných modelov prostredia). Tieto organizmy sú schopné výkonu zložitých kognitívnych akcií, akými sú rozhodovanie (angl. decision making), riešenie problémov (angl. problem solving) a plánovania (angl. planning); u ľudí pristupuje k týmto aktivitám schopnosť komunikácie prostredníctvom jazyka formálne reprezentovaného generatívnou gramatikou. Ľudia sú schopní najzložitejších a najvšeobecnejších foriem spracovania a transformácie, preto podstatná časť psychológie je orientovaná na pochopenie procesov spracovania informácie a inteligencie u ľudí s cieľom vytvárania kognitívnych architektúr a modelov týchto procesov (vyšších kognitívnych procesov).

Psychológia (menovite tá jej časť, ktorá sa zaoberá kognitívnymi procesmi, t. j. kognitívna psychológia) patrí do jadra kognitívnej vedy. Výskum kognitívnej vedy má vždy bezprostredný alebo sprostredkovaný význam v psychológii, Nie všetky výskumy o inteligentných informačných procesoch sú významné pre psychológiu. Musíme poznamenať, že odborný záber kognitívnej vedy je o mnoho všeobecnejší ako psychológia, preto nie všetky výsledky tohto nového vedného odboru sú relevantné pre psychológiu

Vznik informatiky v 50. rokoch minulého storočia znamenal nový impulz pre rozvoj psychológie, menovite jej odklon od striktného behaviorizmu, ktorý označoval za nevedecké snahy používať v psychológii paradigmu vnútorných mentálnych procesov. Počítače poskytujú jedinečné príklady o tom, že počítač transformuje vstup na výstup zložitou postupnosťou vnútorných procedúr, ktoré postupne transformujú vstupnú informáciu na výstupnú informáciu. Bezprostredne nasledujúci vznik umelej inteligencie znamenal ďalší stimul pre psychológiu, pretože snahou informatikou aktívnych v umelej inteligencii bolo nielen navrhnuť ale aj implementovať algoritmy, ktoré simulujú inteligentné činnosti, ktoré boli pokladané vtedy za výlučnú doménu ľudí (hrať dámu alebo šach, porozumenie prirodzenej reči, určenie diagnózy v medicíne a pod.). Americkí vedci (ekonómovia, psychológovia a informatici) Herbert Simon a Allan Newel sa zaslúžili o prienik metód umelej inteligencie do novovznikajúcej kognitívnej psychológie. Vznikla **komputačná**

**hypotéza**, podľa ktorej ľudský mozog pracuje ako biologický procesor, ktorý transformuje vstupné symboly pomocou sledu pravidiel typu *ak..., potom...* na výstupné symboly; táto jednoduchá idea je podstatou symbolického prístupu. Skoro súčasne s týmto prístupom vznikol alternatívny prístup nazývaný konekcionizmus (alebo neurónové siete), podľa ktorého ľudský mozog je neurónová sieť, ktorá obsahuje množstvo navzájom poprepájaných elementárnych procesorových jednotiek nazývaných neuróny. Základy tejto paradigmy už boli naformulované v r. 1943 už zmienanou prácou (pozri kapitolu 1.1) Warena McCullocha a Waltera Pittsa. Neurónové siete v súčasnosti patria medzi významnú časť počítačovo orientovanej umelej inteligencie, kde zaujali postavenie univerzálneho matematicko-informatického prístupu k štúdiu a modelovaniu procesov učenia, adaptácie umelých kognitívnych systémov založených na metafore ľudského mozgu.

## 1.4 Kognitívna neuroveda

„Neuroveda“ študuje nervový systém, jeho vlastnosti, architektúru, fyziológiu a biofyziku neurónov a spojov (synapsií) medzi nimi. Jej špecializovaná časť „kognitívna neuroveda“ študuje mechanizmy realizácie kognitívnych aktivít mozgu. Je to veda o informačných procesoch prebiehajúcich v mozgu: (1) akým spôsobom je táto informácia získavaná z našich zmyslov, (2) ako sú realizované procesy vnímania a rozpoznávania, (3) ako vzniká myslenie a vedomie, (4) ako vznikajú procesy rozhodovania, ktoré zahŕňajú predvídanie budúcich stavov nášho okolia a dôsledkov našej aktivity, (5) čo je to učenie a ako je realizovaná pamäť a konečne, (6) ako sú implementované procesy tvoriace náš jazyk? Tieto činnosti patria do fundamentálneho jadra kognitívnej vedy. Môžeme si položiť otázku, aký je rozdiel medzi kognitívnou vedou a kognitívnou neurovedou? Hľadať odpoveď na túto otázku nie je jednoduché v dôsledku veľkého presahu medzi týmito dvoma „kognitívnymi“ vednými disciplínami. Snáď, pomocou určitého zjednodušeného prístupu k tomuto problému, môžeme povedať, že medzi „kognitívnou neurovedou“ a „kognitívnou vedou“ existuje demrakčná čiara v predmete ich záujmu, predmetom záujmu kognitívnej neurovedy je „hardware“ (t. j. biologická realizácia procesov) a predmetom záujmu „kognitívnej vedy“ je software (t. j. informatická a/alebo algoritmická realizácia procesov). Tu musíme však poznamenať, že toto zakotvenie kognitívnej vedy v biologických vedách (na nervovom substráte) je hlavné kritérium odlíšenia kognitívnej vedy od umelej inteligencie, ktorá taktiež študuje algoritmickú realizáciu kognitívnych (alebo všeobecnejšie – inteligentných) procesov, ale nie je zakotvená na neurónovom substráte.

Veľmi zaujímavá je história neurovedy (a teda aj kognitívnej neurovedy). Za prvých experimentátorov s ľudským mozgom sú pokladaní Egypťania, ktorí už pred 5000 rokmi vykonávali chirurgické výkony na mozgu. O úspešnosti týchto zákrokov svedčia archeologické nálezy, kde boli nájdené lebky so zahojenou trepanáciou lebečnej kosti. Filozof Aristoteles veril, že centrom našej inteligencie je srdce (podľa neho mozog je orgán na ochladzovanie krvi). Tento názor pretrvával aj v starovekom Ríme, až do čias významného lekára a filozofa Galena (žil v druhom storočí po Kr.), ktorý na základe pozorovaní poranení hlavy gladiátorov usúdil, že sídlom inteligencie (alebo v dobovej terminológii duše) je mozog. K tomuto záveru prišiel pomocou experimentálnej skúsenosti s poraneniami hlavy gladiátorov, ktorí pri špecifických poraneniach strácali schopnosť niektorých kognitívnych akcií (videnia, reči a pod.).

Až počiatkom novoveku v 17. a hlavne v 18. storočí začali rôzni lekári a aj filozofi umiestňovať mentálne aktivity do rôznych častí mozgu. Ako klasický príklad uvedieme filozofa Descartesa, ktorý umiestnil myseľ/dušu v šišinke mozgovej (hypofýze). Táto snaha lekárov vyvrcholila v polovici 19. storočia, keď francúzsky lekár Paul Broca umiestnil centrum



reči do ľavého frontálneho laloku mozgu, čo bolo vykonané na základe jeho pozorovania lokálnej lézie mozgu, ktoré malo za dôsledok stratu reči. Tento prístup hľadať centrum rôznych činností mozgu pretrvával až do 30.-40. rokov 20. storočia, keď sa objavili prvé funkčné metódy vyšetřovania mozgu, akými boli elektrické stimulovanie mozgu. Tieto vyšetřovania boli mnohokrát vykonávané ako diagnostické vyšetřenie, pomocou ktorého sa hľadalo centrum ľažkej epilepsie, za účelom jej chirurgického odstránenia. Možno konštatovať, že v tomto pionierskom období boli umiestnené už skoro všetky kognitívne funkcie do rôznych oblastí frontálnych a temporálnych mozgových lalokov.

Pre kognitívnu vedu hrá kľúčovú úlohu vznik a rozvoj neurónovej paradigmy v neurovede. Aj keď už v 2. polovici 19. storočia bolo v podstate jasné, že existuje veľmi úzka korelácia medzi mikroskopickou štruktúrou mozgu a kognitívnymi aktivitami<sup>1</sup>, konkrétne vedomosti o mikroskopickej štruktúre mozgu boli veľmi malé. V tomto smere zohral kľúčovú úlohu taliansky lekár C. Golgi z 2. polovice 19. storočia, ktorý vynašiel metódu selektívneho zafarbenia buniek v mozgu pomocou dusičnanu strieborného, aplikovaním tejto metódy sa stali dobre rozpoznateľné mozgové mikroskopické štruktúry. Tento objav inicioval búrlivý rozvoj neurónovej paradigmy hlavne zásluhou španielskeho lekára S. E. y Cajal, ktorý na prelome 19. a 20. storočia vykonal dôležité výskumy pre etablovanie sa neurónovej paradigmy v neurovede; aplikoval Golgiho metódu k charakterizovaniu rôznych mikroštruktúr mozgu na rovnakom mieste (čím získal experimentálne podklady pre špecifikáciu mozgu ako neurónovej siete obsahujúcu neuróny, ktorá obsahuje dendritický systém (zložitý kanálový systém pre vstupnú informáciu), axón (kanál pre výstupnú informáciu) a somu (telo neurónu). Taktiež ukázal, že medzi rôznymi neurónmi existujú spoje – synapsie, kde axón z jedného neurónu sa „prilepí“ na dendrit druhého neurónu, pričom tieto spoje sa realizujú pomocou štruktúr nazývaných „trne“. Skutočnosť, že medzi neurónmi existujú synapsie, viedlo počiatkom 20. rokov minulého storočia k štúdiu elektrochemických mechanizmov prenosu signálov medzi neurónmi. V nasledujúcich dekádach minulého storočia došlo k prudkému rozvoju štúdia detailov elektrochemického mechanizmu prenosu signálu cez synapsie, čo viedlo k tvorbe detailného popisu aktivity neurónov. V tom istom období anglický lekár Sir E. A. Adrian, ktorý na tú dobu využil moderné prístrojové vybavenie (katódovú trubicu a elektronické zosilňovače), bol schopný zaznamenávať priebeh vnútorného potenciálu neurónu. Pomocou tejto techniky Adrian zistil, že neurón pracuje v móde „všetko alebo nič“: ak suma vstupných signálov (cez dendritický systém neurónu) prevýši prahovú hodnotu vnútorného potenciálu, potom neurón cez axón začne vysielaa výstupné signály.

## 1.5 Kognitívna veda a umelá inteligencia

Prv než pristúpime k definovaniu umelej inteligencie a špecifikácie jej významu pre kognitívnu vedu, obráťme našu pozornosť na dôležitý „pomocný“ pojem „inteligentný problém“, ktorý bude mať centrálny význam pri definovaní umelej inteligencie. **Inteligentný problém** ľudskej kognície je taký problém, ktorého riešenie vyžaduje tieto vyššie aspekty

---

<sup>1</sup> Americký filozof H. Spencer v knihe „*The Principles of Psychology*“ z r. 1855 prezentoval ideu, že pochopenie základných princípov fungovania nervového systému je nevyhnutným predpokladom pre korektnú interpretáciu procesov a fenoménov psychológie. Podobne, ďalší americký filozof W. James v knihe „*The Principles of Psychology*“ z r. 1890 vychádzal z postulátu, že psychologické fenomény sa musia vysvetľovať pomocou aktivít mozgu, a existuje úplný paralelizmus medzi analýzou fungovania nervov a analýzou mentálnych ideí, každá mentálna modifikácia musí byť sprevádzaná zmenami v nervovom systéme mozgu. Pomocou tejto idey rozpracoval teóriu asociatívnej pamäti, ktorá obsahuje jednotlivé idey a ktoré sú medzi sebou prepojené do paralelnej štruktúry tak, že vybavenie si (recall) jednej idey je sprevádzané súčasným pripomenutím si asociovaných ideí.

kognície: (1) percepciu prostredia, (2) vykonávanie akcií nad týmto prostredím, (3) učenie, (4) usudzovanie, (5) plánovanie, (6) rozhodovanie, (7) riešenie úloh, (8) komunikáciu, a pod. Samozrejme, nie vždy sa vyžaduje spoločný výskyt týchto aspektov, obvykle sa vyskytujú vo vzájomnej kombinácii len niekoľkých vyššie uvedených aspektov kognície. Potom, **umelá inteligencia je tá časť informatiky, ktorá sa zaoberá algoritmizáciou inteligentných problémov**. Pojem „inteligentný problém“ je historicky závislý, to čo sa nám javilo pred 30-40 rokmi ako inteligentný problém, v súčasnosti v dôsledku teoretického rozvoja matematiky a informatiky môže byť pomerne jednoducho algoritmizovateľný a nevyžaduje použitie výpočtových techník umelej inteligencie (napr. konštrukciu stromu riešení pomocou prehľadávania). Ako príklad môžeme uviesť program DENDRAL, vytvorený v 60. rokoch minulého storočia pod vedením laureáta Nobelovej ceny za medicínu Joshua Lederberga v spolupráci so známym stanfordským informatikom E. Feigenbaumom. Cieľom tohto programu bola rekonštrukcia vzorca molekuly zo známeho hmotnostného spektra (ktoré špecifikuje fragmenty vzniknuté z danej molekuly dopadom elektrónov). Tento problém bol algoritmizovaný použitím klasických metód vtedajšej umelej inteligencie. V súčasnosti je problém rekonštrukcie grafu na základe jeho podgrafov štandardný problém počítačovej teórie grafov a nepovažuje sa za „inteligentný problém“ vhodný pre umelú inteligenciu.

Kognitívna psychológia používa koncepcie umelej inteligencie a kognitívnej vedy k formulovaniu a testovaniu svojich hypotéz. Mentálne štruktúry a procesy sú popísané pomocou informatických koncepcií, čo umožňuje vykonávať počítačové simulácie modelov vytváraných v rámci kognitívnej psychológie. Otázka o tom, či ľudia vykonávajú danú kognitívnu aktivitu tak, ako je špecifikovaná jej kognitívnym modelom je fundamentálneho charakteru pre kognitívnu psychológiu, práve počítačové simulácie môže byť veľmi nápomocné pri hľadaní odpovede na túto otázku.

V umelej inteligencii už od jej vzniku opakovane diskutuje problém, ako najlepšie modelovať inteligenciu. Tento problém má v zásade dve riešenia, symbolickú umelú inteligenciu a konekcionistickú umelú inteligenciu. Symbolická metafora umelej inteligencie je založená na algoritmoch, ktoré pomocou formálnych pravidiel transformujú vstupné symboly na výstupné symboly. Jedná sa o sekvenčné algoritmy, ktoré sa vykonávajú krok po kroku, pričom v danom kroku vykonávania algoritmu využívame výsledky získané v predchádzajúcich krokoch vykonávania algoritmu. Konekcionistická metafora umelej inteligencie, ktorá je založená na používaní neurónových sietí obsahujúcej mnoho jednoduchých procesorov, ktoré sú medzi sebou poprepájané. Neurónovú sieť môžeme chápať ako paralelný počítač, kde aktivita vybraného neurónu v čase  $t$  závisí od aktivít neurónov v čase  $t-1$ , ktoré sú v sieti klasifikované ako „predchádzajúcich“ neuróny, t. j. sú spojené spojmi, ktoré z nich vychádzajú a vchádzajú do daného neurónu. Hybridne systémy sú kombináciou symbolického a konekcionistického prístupu k spracovaniu informácie. Tento prístup je v súčasnosti často aplikovaný v umelej inteligencii, kde nižšie kognitívne aktivity sú modelované neurónovou sieťou, zatiaľ čo vyššie kognitívne aktivity sú modelované symbolickým prístupom. Moderná umelá inteligencia sa zameriava na multiagentové systémy, ktoré obsahujú intencionálnych agentov (agenti, ktorý vykonávajú predpísanú aktivitu – intenciu).

### 1.5.1 Symbolické, konekcionistické a hybridné systémy

Konekcionistické modely mysle sú založené na metafore neurónovej siete obsahujúcej jednoduché procesné elementy – neuróny a jednosmerné spoje medzi nimi – synapsie. Poznamenajme, že konekcionistické počítanie (alebo počítače) sa diametrálne odlišujú klasického sekvenčného počítača modelovaného Turingovým strojom.

Často sa rozdiel medzi klasickým symbolickým prístupom a konekcionistickým prístupom (tiež nazývaným subsymbolickým) považuje za skoro neprekonateľný, ktorý rozdeľuje komunitu umelej inteligencie a/alebo kognitívnej vedy na dva „nezmieriteľné tábory“, čo je však veľký omyl. Tento rozdiel je však veľmi vágny, ba až delikátny. Nie všetky konekcionistické prístupy sú založené na subsymbolickom spracovaní informácie, a naopak, mnohé klasické prístupy umelej inteligencie sú založené na konekcionistickom spracovaní vstupnej informácie (menovite procesy videnia a pochopenie prirodzeného jazyka). Na prelome 50. a 60. rokov minulého storočia klasická umelá inteligencia sa vyvíjala podstatne rýchlejšie ako konekcionistické modely kognitívnych procesov. V roku 1969 Minsky spolu s Papertom vo svojej slávnej knihe „*The perceptron*“

Aj keď už na prelome 50-tych a 60-tych rokov bolo jasné, že perceptrón má určité obmedzenia, že nie je univerzálne použiteľný na klasifikáciu ľubovoľných obrazcov, až pomocou knihy "*Perceptron*" boli tieto obmedzenia presne opísané a špecifikované. Ukázalo sa a na rôznych príkladoch všeobecného charakteru sa demonštrovalo, čo dokáže perceptrón korektne klasifikovať a čo nie. Minsky a Papert zaviedli fundamentálny pojem *lineárnej separovateľnosti*, pomocou ktorého sa pomerne jednoducho (a tiež názorne) formulujú podmienky pre korektnú klasifikáciu danej triedy objektov. V dobe publikovania knihy bol jedným z najaktuálnejších problémov umelej inteligencie problém rozpoznávania a klasifikácie scén. Práve v tejto dobe vznikali základné "symbolické" algoritmy a postupy, ako riešiť problém analýzy a klasifikácie scén. Preto Minsky a Papert upriamili svoju pozornosť na riešenie týchto problémov pomocou perceptrónov. Zistili napríklad, že perceptrón je schopný rozlíšiť trojuholníky od štvorcov, ale nedokáže rozlíšiť súvislé geometrické telesá od nesúvislých. Neprekonateľné problémy má perceptrón aj s korektnou klasifikáciou nových obrazcov, ktoré vznikli z pôvodného obrazca aplikovaním translácie alebo rotácie. Ich záver sa dal preto očakávať: *perceptrón nie je univerzálne výpočtové zariadenie, existujú jednoduché úlohy, ktoré nie sú korektne klasifikovateľné perceptrónom*. V knihe boli navrhnuté aj rôzne zovšeobecnenia perceptrónu, a to použitie skrytých neurónov a neurónov vyššieho rádu. Autori ukázali, že tieto zovšeobecnenia perceptrónu sú schopné prekonať bariéru lineárnej separovateľnosti. Žiaľ, oba tieto prístupy boli spochybnené samými autormi tým, že nepovažovali za potrebné študovať metódu učenia sa pre nastavenie váhových koeficientov skrytých neurónov alebo neurónov vyšších rádov. Hlavné príspevky Minského a Paperta k teórii neurónových sietí možno zhrnúť takto:

(1) Minsky v knihe "*Computation: Finite and Infinite Machines*" pri štúdiu McCullochových a Pittsových neurónov zaviedol pojem rekurentnej neurónovej siete, ktorá bola zložená z prahových neurónov. Pomocou konštruktívneho dôkazu ukázal, že táto sieť simuluje konečný binárny automat. Podobnú problematiku, lenže na kvalitatívne vyššej úrovni, začali študovať až na prelome 80-tych a 90-tych rokov pomocou dynamických prístupov k teórii rekurentných neurónových sietí. Tieto siete sa chápu ako dynamické systémy so zložitými trajektóriami v stavovom priestore, ktoré môžu simulovať konečné automaty.

(2) Minsky a Papert v knihe "*Perceptron: An Introduction to Computational Geometry*" podrobne študovali výpočtové možnosti perceptrónu. Ukázali, že je schopný klasifikovať korektne len takú množinu objektov, ktorá je v zodpovedajúcom priestore vstupných aktivít lineárne separovateľná. Navrhli tiež dva možné prístupy, ako prekonať túto bariéru, a to buď použitím skrytých neurónov, alebo použitím neurónov vyšších rádov. Pretože nevenovali dostatočnú pozornosť procesu učenia sa takto zovšeobecných perceptrónov, uvedené prístupy označili za neefektívne, pretože nie je k dispozícii metóda učenia sa, ktorá by určila hodnoty váhových koeficientov priradených týmto neurónom.

Prečo sa venujeme tak podrobne vyčísleniu zásluh týchto dvoch informatikov pre rozvoj neurónových sietí? V literatúre o neurónových sieťach sa už notoricky opakuje tvrdenie, že Minsky s Papertom svojou druhou knihou "*Perceptron*" spôsobili, že nastal na niekoľko desaťročí útlm výskumu v oblasti neurónových sietí. Podľa mienky autorov tejto

kapitoly je tento názor nepravdivý, ba až smiešny. Táto vynikajúca kniha je stále primárnym zdrojom inšpirácie pre teóriu neurónových sietí a isto nespôsobila dlhodobý útlm rozvoja neurónových sietí. Hlavným zdrojom tohto útlmu bola skutočnosť, že vnútorné zdroje symbolickej umelej inteligencie neboli ešte zďaleka vyčerpané. Symbolická umelá inteligencia sa v tej dobe javila ako hlavný generátor nových trendov umelej inteligencie. Ďalším dôvodom bol odklon od problematiky rozpoznávania vzorov, v ktorom boli neurónové siete vtedy najlepšie, k problematike vyšších kognitívnych procesov, ako je napríklad riešenie problémov a dedukcia. Vtedy sa zdalo, že prostriedky klasickej umelej inteligencie dokážu tieto úlohy z vyšších kognitívnych procesov dobre riešiť. K tomu sa pridalo rozčarovanie z počiatočných neopodstatnených nádejí, keď priaznivci neurónových sietí, podobne ako ostatní vedci pracujúci v umelej inteligencii, prisudzovali vtedajším neurónovým sieťam aj kvality, ktoré tieto jednoducho nemali - čo sa skoro ukázalo. No aj tak boli neurónové siete aj naďalej rozvíjané hŕstkou nadšencov, ako bol napríklad Teuvo Kohonen a Stephen Grossberg. Situácia sa diametrálne zmenila až na prelome 80-tych a 90-tych rokov, keď si informatici začali čoraz naliehavejšie uvedomovať obmedzenosť metód symbolickej umelej inteligencie. Nepodarilo sa úplne uskutočniť žiadnu veľkú víziu symbolickej umelej inteligencie zo 60-tych a 70-tych rokov (napr. prekladač z prirodzeného jazyka, komunikácia s počítačom v prirodzenom jazyku, program pre šach na veľmajstrovskej úrovni,...).

## 1.6 Počítače a kognícia

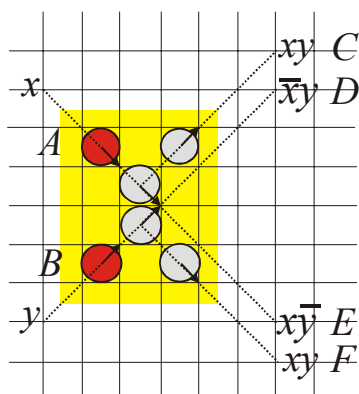
Ideové prekurzory umelej inteligenciu môžu byť vysledované až do 3. storočia pred Kr., keď Aristoteles vytváral svoju logiku sylogizmov a stoický filozof Chrysippos vytváral logiku, ktorú by sme z dnešného pohľadu nazvali výrokovou. Táto staroveká grécka logika v podstate s malými zmenami pretrvala až do novoveku. Zo súčasného pohľadu možno konštatovať, že veľmi pozitívnu úlohu pre rozvoj logiky zohral nemecký filozof a matematik Gottfried Wilhelm Leibniz, pokúsil sa vybudovať formálny systém, ktorý by nahradil verbálne metódy usudzovania manipuláciami s formulami. Postuloval formálny systém s dvoma časťami: (1) jazyk logiky *lingua characteristica*, pomocou ktorého je možné reprezentovať každý výrok a (2) kalkulus *calculus ratiocinator*, pomocou ktorého je možné uskutočňovať usudzovanie systematickým a matematicky presným spôsobom. Žiaľ, trvalo ešte ďalších 200 rokov než sa naplnila táto Leibnizova idea, keď v polovici 19. st. anglickí matematici A. de Morgan a G. Boole zostrojili „kalkulus“ – výrokovú logiku. Pre Descartesovho súčasník anglického filozofa T. Hobbesa myslenie už nebolo nič iné, ako len špeciálny druh výpočtu. Táto hypotéza, ktorá v 17. storočí znela veľmi neobvykle ba až exoticky, bola až v súčasnosti plne akceptovaná a realizovaná pomocou umelej inteligencie a kognitívnej vedy, kde má postavenie centrálnej paradigmy. Takto vzniknutú a formulovanú logiku možno považovať za formálnu metódu, ktorá algoritmizuje naše usudzovanie, čiže logika je integrálnou súčasťou umelej inteligencie, menovite tej časti, ktorá sa zaoberá usudzovaním.

Fundamentálna práca z r. 1937 (ktorá má zásadný význam pre vznik umelej inteligencie a kognitívnej vedy), ktorej autorom bol anglický matematik a logik (a snáď aj jeden z prvých britských informatikov) Alan Turing, riešila známy Hilbertov problém rozhodnuteľnosti o tom, či ľubovoľný matematický výrok je pravdivý alebo nepravdivý. Turing pri riešení tohto problému dospel k abstraktnému počítaču, ktorý je schopný vypočítať (alebo aj nevypočítať) funkciu charakterizujúcu daný matematický výrok<sup>2</sup>. Potom môžeme

---

<sup>2</sup> Napríklad vieme, že pravdivosť (alebo nepravdivosť) ľubovoľnej formuly výrokovej logiky môže byť pomocou tabuľkovej metódy (algoritmu) vypočítaná konečným počtom krokov. Môžeme teda povedať, že pravdivosť formúl výrokovej logiky je vypočítateľná, t. j. existuje algoritmus (aj keď s exponenciálnym nárastom zložitosti),

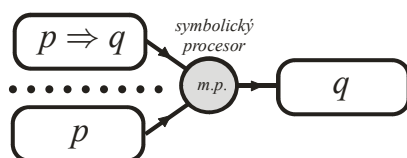
celú triedu prípustných funkcií rozdeliť na dve potriedy: funkcie, ktoré sú Turingovým počítačom vypočítateľné a funkcie, ktoré nie sú Turingovým počítačom vypočítateľné. K podobnému mu výsledku dospel súčasne aj americký logik Alonzo Church pomocou tzv. rekurzívnych funkcií, tieto dva výsledky sú spojené do Turingovej a Churchovej tézy, podľa ktorej pre každú funkciu, ktorá je algoritmizovateľná (pozri poznámku 2) je aj vypočítateľná pomocou Turingovho počítača; t. j. existuje veľmi tesná väzba medzi algoritmizovateľnosťou nejakej funkcie (problému) a jeho vypočítateľnosťou. Pretože veríme, že naše kognitívne aktivity sú algoritmizovateľné, tak môžu byť aj implementované na počítači, čiže naša myseľ a Turingov počítač sú „ekvivalentné“, t. j. že tak myseľ ako aj Turingov počítač, dávajú rovnaký výstup na rovnaký vstup, odlišujú sa len v spôsobe „vypočítania“ tohto výstupu.



Dokonala elastická zrážka dvoch biliardových gúl simuluje logickú bránu konjunkcie. Binárne premenné  $x$  a  $y$  špecifikujú, či v čase  $t = 1$  sa v daných východiskových polohách nachádza prvá resp. druhá guľa. K elementárnemu aktu zrážky dochádza v čase  $t = 2$ . Výsledné polohy v čase  $t = 3$  sú špecifikované súčinní (konjunkciami) binárnych premenných  $xy, \bar{x}y, x\bar{y}, \bar{x}\bar{y}$ .

V polovici 80. rokov minulého storočia anglický fyzik a informatik I. Deutsch formuloval fyzikálnu verziu Churchovej a Turingovej tézy: „každý konečný fyzikálny proces simuluje nejaký výpočet“. Podľa tejto verzie Churchovej a Turingovej tézy každý fyzikálny proces môžeme chápať ako špeciálny výpočet, problém je len v tom ako kódovať vstup do tohto fyzikálneho systému a dekódovať jeho výstup. Známy príklad takéhoto fyzikálneho počítania sú biliardové gule, ktoré bolo navrhnuté americkými fyzikmi Fredkinom a Toffolim v r. 1982.

Skoro súčasne s týmto dôležitým teoretickým Turingovým výsledkom bolo na konkrétnom príklade v neurovede ukázané, že počítače špeciálneho typu môžu byť zostrojené aj pomocou jednoduchých logických (alebo prahových) neurónov McCullochom a Pittsom (pozri kapitolu 1.1). Ukázali, že každá Boolova funkcia, ktorá obsahuje spojky konjunkcie, disjunkcie a negácie je vyjadrená pomocou doprednej neurónovej siete. Ak tieto siete obsahujú cykly (hovoríme, že sieť je rekurentného charakteru), potom sieť môže simulovať konečnostavový stroj schopný univerzálneho počítania.



Znázornenie symbolického procesoru, ktorý uskutočňuje pravidlo usudzovania výrokovej logiky „modus ponens“, ktoré je vyjadrené pomocou schémy usudzovania

$p \Rightarrow q$	1. predpoklad
$p$	2. predpoklad
$q$	dôsledok

kde  $p$  a  $q$  sú symboly – výrokové premenné.

pomocou ktorého sme schopný určiť pravdivostnú hodnotu výrokovej logiky. Tento problém už nemá jednoduché riešenie pre formuly predikátovej logiky, kde vo všeobecnosti neexistuje algoritmus, ktorý určí pravdivostnú hodnotu každej formuly predikátovej logiky; tento problém je riešiteľný len pre široké triedy formúl predikátovej logiky.

Na prelome 60. a 70. rokoch minulého storočia bola formulovaná **komputačná teória mysle** (taktiež nazývaná komputačná metafora alebo komputačná paradigma) kognitívnej vedy, podľa ktorej je myseľ špeciálny symbolický procesor, ktorý vstupné symboly transformuje na výstupné pomocou hierarchického systému pravidiel typu „ak..., potom...“. Je dôsledným prejavom funkcionalizmu, podľa ktorej myseľ nie je priamo totožná s mozgom, ale je prejavom jeho organizácie, štruktúr, ich vzájomného prepojenia. Táto hypotéza má tieto dôležité vedľajšie aspekty:

- (1) Výpočet prebieha nad syntakticky špecifikovanými symbolmi (t. j. symboly sú špecifikované pravidlami, ktoré taktiež špecifikujú ich kombinácie), ktoré majú svoj sémantický význam. Tak napríklad, súčet dvoch symbolov, ktoré majú význam – sémantickú interpretáciu – dekadické číslo.
- (2) Výpočet je reprezentovaný algoritmom, ktorý je konečnou krok-po-kroku procedúrou, z ktorej každá je uskutočnená počítačom – mozgom.
- (3) Potrebné výpočty neobsahujú len štandardnú aritmetiku, ale taktiež aj logické usudzovanie vrátane indukcie, abdukcie a rozhodovanie.
- (4) Výpočty pokrývajú relatívne samostatnú úroveň ich psychologickej explanácie, ktorá je nezávislá od neurofyziológie mozgu.

### 1.6.1 Syntax, sémantika a algoritmy

Základný význam pre pochopenie a interpretáciu komputačnej teórie mysle je rozdiel medzi syntaxou a sémantikou symbolického systému. Spôsob povolenej manipulácie so symbolmi je problém syntaktický a otázka čo znamenajú novovytvorené symboly je problém sémantický. Uvažujme tri rôzne notácie pre čísla: (1) dekadické čísla ‘1’, ‘2’, ‘3’,..., (2) rímske čísla ‘I’, ‘II’, ‘III’,... a (3) binárne čísla ‘1’, ‘10’, ‘11’,.... Tak napríklad, číslo „tri“ má tri rôzne syntaktické reprezentácie 3, III a 11. Pravidlá pre súčet čísel v rôznej reprezentácii sú diametrálne odlišné, ale ich výsledky musia mať rovnakú sémantickú interpretáciu. Keď sa dieťa učí jednoduchú aritmetickú operáciu napr. s celými číslami, potom sa učí používať vhodný algoritmus na realizáciu tejto operácie, pričom tento algoritmus musí odrážať ich sémantiku (menovite táto úloha „sčítať prirodzené čísla“ musí odrážať sémantickú interpretáciu čísel). Ak by sme vykonávali „sčítanie“ pre rímske alebo binárne čísla, musíme dostať rovnaké výsledky, aké dostávame pre dekadické čísla.

Algoritmus je “mechanická” krok-po-kroku procedúra, ktorá transformuje vstupné symboly na výstupné symboly, pričom tieto symboly majú presne špecifikovaný syntax, ktorý odráža ich sémantickú interpretáciu (význam). Turing navrhol všeobecnú konštrukciu tohto algoritmu (ktorý sa nazýva Turingov počítač), ktorá postupne transformuje vstupné symboly na výstupné symboly. Podľa Churchovej a Turingovej tézy, Turingov počítač môže počítať všetko, čo je je „intuitívne“ vypočítateľné. Táto idea bola inšpiráciou pre vznik súčasnej výpočtovej techniky. Komputačná teória mysle prenáša túto paradigmu do psychológie postulátom, že všetky inteligentné procesy môžu byť študované ako algoritmické procesy realizované na jednoduchých symbolických procesoroch.

### 1.6.2 Jednoduchý symbolický procesor

Turingov stroj obsahuje „pásku“ na ktorú sa zaznamenávajú symboly medzivýsledky, „skener“, ktorý je schopný s páskou pohybovať dopredu a dozadu a taktiež zapisuje a číta symboly na páske. Toto je však len jedna z možných architektúr Turingovho stroja, čo je dôležité pre stroj – počítač, že jeho elementárne operácie sú realizované pomocou nejakého

fyzikálneho procesu. K tomu, aby komputačný model mozgu bol aspoň vzdialene podobný ľudskému mozgu, použijeme prístup logických neurónových sietí McCullocha a Pittsa, ktoré obsahujú elementárne brány konjunkcie, disjunkcie a negácie. Budeme študovať elementárnu úlohu súčtu dvoch jednobitových čísel

$$\alpha_1 + \alpha_2 = \beta_1\beta_2$$

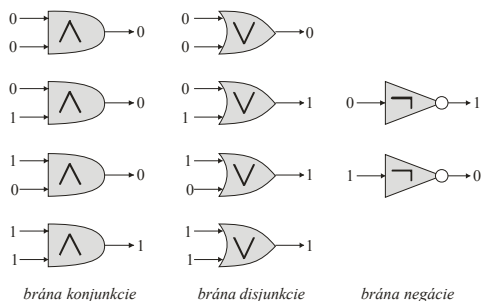
ktorej všetky možné prípady sú uvedené v nasledujúcej tabuľke

$\alpha_1$	$\alpha_2$	$\beta_1$	$\beta_2$
0	0	0	0
0	1	0	1
1	0	0	1
1	1	1	0

kde prvá dva stĺpce reprezentujú vstupné čísla – bity a posledná dva stĺpce reprezentujú výstupné čísla – bity. Z tejto tabuľky bezprostredne vyplýva, že výstupné bity  $\beta_1$  a  $\beta_2$  sú určené ako jednoduché funkcie vstupných bitov

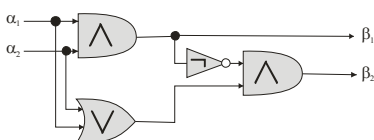
$$\beta_1 = \alpha_1 \wedge \alpha_2, \beta_2 = (\neg\alpha_1 \wedge \alpha_2) \vee (\alpha_1 \wedge \neg\alpha_2) = (\alpha_1 \vee \alpha_2) \wedge \neg(\alpha_1 \wedge \alpha_2) \quad (*)$$

kde pravá strana druhej formuly môže byť jednoducho verifikovaná aplikáciou de Morganovej formuly na konjunkcie  $\alpha_1 \wedge \alpha_2$  a potom následným „roznásobením“ oboch zátvoriek s disjunkciou. Týmto sme ukázali, že kognitívna úloha súčtu dvoch jednobitových čísel môže byť vyjadrená funkciou ktorá je jednoducho vypočítateľná (ako ukážeme v ďalšom texte, táto vypočítateľnosť je uskutočnená pomocou obvodu – neurónovej siete, ktorá obsahuje brány konjunkcie, disjunkcie a negácie.



Logické brány konjunkcie, disjunkcie a negácie, ktoré na ľavej strane majú vstupné „binárne“ hodnoty a na pravej strane výstupnú „binárnu“ hodnotu. Vstupné hodnoty obsahujú všetky možné rôzne prípady vstupov. Fyzikálna realizácia týchto brán je v súčasnosti rutinnou technologickou záležitosťou.

Ak tieto elementárne logické brány prepojíme do funkčného celku nazývaného neurónová sieť tak, aby boli „implementované“ Boolove funkcie (\*) pre jednotlivé komponenty výsledného súčtu  $\beta_1\beta_2$ .



Logický obvod (neurónová sieť) realizujúci súčet dvoch 1-bitových čísel  $\alpha_1$  a  $\alpha_2$  (tvoria vstup do siete) na výsledné 2-bitové číslo  $\beta_1\beta_2$  (jeho komponenty tvoria výstup siete). Tento obvod (neurónová sieť) môže byť formálne chápaný ako počítač realizujúcu úlohu súčtu dvoch 1-bitových čísel.

V rámci *konekcionistického prístupu* ku vzťahu medzi mozgom a myslou je tento problém riešený pomocou koncepcií neurónových sietí. Podľa jednej zo základných paradigiem kognitívnej vedy je mozog/mysel charakterizovaný ako počítač, (1) ktorý transformuje symboly pomocou syntaktických pravidiel na iné symboly, pričom (2)

myšlienky sú symbolické reprezentácie implementované pomocou jazyka myslenia, a (3) mentálne procesy sú postupnosti symbolov (medzi ktorými sú príčinné vzťahy) generované syntaktickými pravidlami. Použitie termínu „počítač“ obvykle evokuje predstavu sekvenčného počítača von neumannovskej architektúry (napr. personálne počítače majú túto architektúru), kde je možné striktné oddeliť hardware od software; kde na tom istom počítači – hardware môže byť vykonávaných nepreberné množstvo diametrálne rozdielnych programov – softvérov. Pre tieto počítače existuje striktná dichotómia medzi počítačom a programom – t. j. medzi hardvérom a softvérom. Žiaľ, metafora mysle ako počítača implikuje u mnohých ľudí predstavu, že je možné oddeliť mozog od mysle, ako dva „nezávislé“ fenomény, kde mozog hrá úlohu hardvéru, zatiaľ čo myseľ je softvér (vykonávaný na hardvéru – mozgu). V tejto súvislosti sa tiež hovorí o sociokultúrnom zdroji ľudskej mysle, kde mozog akoby bol len irelevantným „hardwarovým“ vehiklom mysle.

Obráťme našu pozornosť na súčasný prístup k chápaniu vzťahu medzi mozgom a myslou, ktorý je založený na konekcionistickom poňatí tak mozgu, ako aj mysle. Základná predstava o mozgu (založená na experimentálnych neurovedných poznatkoch) je, že je tvorený z neurónov navzájom poprepájaných pomocou jednosmerných synaptických spojov. Ľudský mozog vykazuje neobyčajnú plasticitu, v priebehu učenia neustále vznikajú (ale taktiež aj zanikajú) synaptické spoje. Architektúra mozgu je určená spojmi medzi neurónmi, ich inhibičným, alebo excitačným charakterom, a taktiež aj ich intenzitou. Možno konštatovať, že *schopnosť mozgu vykonávať nielen kognitívne aktivity, ale byť aj pamäťou, je plne zakódovaná do jeho architektúry*. Na základe týchto neurovedných poznatkov bazálneho charakteru môžeme konštatovať, že počítačové paradigma ľudskeho mozgu/mysle sa musí formulovať tak, že **mozog je paralelný distribuovaný počítač** (obsahujúci mnoho miliárd neurónov, elementárnych procesorov, ktoré sú medzi sebou poprepájané do zložitej neurónovej siete). Program v tomto paralelnom počítači je priamo zabudovaný do architektúry neurónovej siete, t. j. ľudský mozog je **jednoúčelový** paralelný počítač reprezentovaný neurónovou sieťou, ktorý nie je možné preprogramovať bez zmeny jeho architektúry. Z týchto všeobecných úvah vyplýva, že *mysleľ s mozgom tvoria jeden integrálny celok; myseľ je v tomto prístupe chápaná ako program vykonávaný mozgom, avšak tento program je špecifikovaný architektúrou distribuovanej neurónovej siete reprezentujúcej mozog*. Mozog a myseľ tvoria dva rôzne pohľady na ten istý objekt: (1) keď hovoríme o mozgu, myslíme tým „hardwarovú“ štruktúru, biologicky realizovanú neurónmi a ich synaptickými spojmi (formálne reprezentovanú neurónovou sieťou), v opačnom prípade, (2) keď hovoríme o myšli, myslíme tým kognitívne a iné aktivity mozgu, realizované výpočtami neurónovej siete reprezentujúcej mozog.

## Literatúra