

Kognitívna veda a umelá inteligencia

Kolektív autorov:

Igor Farkaš

Barbora Cimrová

Ján Rybár

Martin Takáč

Andrej Lúčný

Katedra aplikovanej informatiky /
Centrum pre kognitívnu vedu
Fakulta matematiky, fyziky a informatiky
Univerzita Komenského v Bratislave

Vydané pre Univerzitu Hradec Králové
v rámci projektu "Informační, kognitivní a interdisciplinární podpora
výzkumu" (reg. č. CZ.1.07/2.3.00/20.0001)

Február 2013

Obsah

- 1 Paradigmy a metódy výskumu v kognitívnej vede
(I. Farkaš)
- 2 Neuroveda – strategická oblasť výskumu v kognitívnej vede
(B. Cimrová)
- 3 Fundamentálne aspekty kognície: percepcia a pamäť
(J. Rybár)
- 4 Reprezentácia významov v živých a umelých systémoch
(M. Takáč)
- 5 Postkognitivistický prístup k percepcii a akcii v mobilnej robotike
(A. Lúčný)

Kapitola 1

Paradigmy a metódy výskumu v kognitívnej vede

IGOR FARKAŠ

„Veda sa vyvíjala opačným smerom, než aký by sa bol dal očakávať. To, čo bolo od nás najviac vzdialené, sme opísali pomocou prírodných zákonov ako prvé, a potom postupne to, čo k nám bolo bližšie: najprv vesmír, potom Zem, svet zvierat a rastlín, potom ľudské telo, a napokon (hoci stále veľmi neúplne) ľudskú myseľ.“

(B. Russell, 1935)

1.1 Úvod

Snahy o pochopenie fungovania ľudskej mysle siahajú až do antických čias, keď boli výsostnou doménou filozofov. Už tí si kládli otázky, v akom vzťahu je myseľ (duša) k telu, ktoré znalosti sú vrodené, a ktoré sa učíme, a ako ich získavame, či fyzikálne a abstraktné objekty existujú nezávisle od nás alebo sú výplodom ľudskej mysle atď. Až koncom 19. storočia sa od filozofie oddelila psychológia, postupným zavádzaním nových vedeckých metód skúmania. Ďalším významným míľnikom v histórii bola polovica 20. storočia, keď tzv. *kognitívna revolúcia*, predznamenaná prvým Hixonským sympóziom konanom v Kalifornii v roku 1948, viedla k vzniku *kognitívnej vedy*, ako vedy o poznaní (z lat. *cognitio*). Táto udalosť bola vyvrcholením paralelných snáh o skúmanie ľudskej mysle (*mind*) a mentálnych procesov (Gardner, 1987). Tieto snahy sa týkali materských disciplín kognitívnej vedy (časť 1.2), čím sa vytvorili predpoklady pre zrod tejto novej vednej disciplíny. Na myseľ a kognitívne procesy sa môžeme pozeráť z rôznych perspektív, využívajúc pritom odlišné metódy skúmania. To si vyžaduje samotný objekt skúmania – myseľ, kvôli svojej komplexnosti, ako aj „špeciálnemu“ statusu. Na jednej strane, myseľ je privátna (perspektíva prvej osoby) a vlastnému subjektu prístupná, vďaka čomu ten dokáže vypovedať o (niektorých) svojich mentálnych stavoch. Na druhej strane, mozog, ako predpokladaný fyzikálny substrát mysle, je objektívne pozorovateľný (perspektíva tretej osoby), takže ho môžeme skúmať metódami prírodných vied, no človek samotný k nemu nemá prístup (nie sme si vedomí aktivity vlastného mozgu).

Aj keď myseľ nie je dobre definovateľná entita (môžeme povedať, že neexistuje v tomto smere v odbornej literatúre jednotný názor), snažíme sa ju skúmať vedeckými metódami, aj keď nepriamo, napríklad pozorovaním správania alebo pomocou rôznych fyziologických meraní. Tento proces skúmania je zložitý, pretože schopnosti a prejavy človeka sú veľmi rozmanité a komplexné. Možno aj preto v súčasnosti stále neexistuje konsenzus, či kognitívna veda je samostatnou vednou disciplínou alebo stále ostáva súborom viacerých vedných smerov, z ktorých každý nezávisle napreduje so svojimi metódami a špecifickými aspektmi skúmania mysle (Greco, 2012).

Ďalej budeme postupovať nasledovne. V časti 1.2 uvedieme krátky prehľad materských disciplín kognitívnej vedy a spomenieme i niektoré kľúčové pojmy. Časť 1.3 je „preletom ponad“ hlavné paradigmy kognitívnej vedy. V častiach 1.4 a 1.5 konkrétnejšie vysvetľujeme dva pojmy – reprezentáciu a počítanie. V časti 1.6 uvedieme argumenty v prospech paradigmy ukotvenej kognície a časť 1.7 je stručným záverom.

1.2 Materské disciplíny kognitívnej vedy

Mysel' je predmetom skúmania viacerých vedeckých disciplín. Samozrejme historicky najstaršou disciplínou je *filozofia*, založená na vnútornom pozorovaní (introspekcii) myšlienok, pocitov, predstáv či želaní. Introspektívne poznanie tvorilo podľa viacerých filozofov základnú metódu, ako sa oboznamovať s vlastnou skúsenosťou. Okrem toho, tento spôsob sa používal aj pri uvažovaní o mentálnych stavoch druhých. Aj keď filozofiu možno vnímať v porovnaní s prírodovednými metódami ako menej exaktnú, úloha filozofie mysle bola a ostáva veľmi dôležitá. Namiesto odpovedí sa filozofi mysle snažia spochybňovať teórie a formulovať otázky. Dokážu sa však pozeráť na svet s nadhľadom a nútiť expertov k reflexii o vlastnom snažení. Ich filozofické otázky sa obyčajne týkajú ontológie (povahy existencie niečoho) a epistemológie (podstaty poznávania). Príkladom ontologickej otázky môže byť to, či povaha existencie nejakého fyzikálneho objektu (povedzme kameňa alebo mozgu) je rovnaká ako v prípade nejakého abstraktného objektu (napríklad prvočísel, alebo mentálnych predstáv). Epistemologická otázka sa môže týkať napríklad toho, či je niečo poznateľné, alebo aký je vzťah znalostí k pravde, či presvedčeniam.

Keď sa (vedecká) *psychológia* odčlenila od filozofie, prevzala od nej síce introspekciu ako metódu skúmania, ale hlavne ju obohatila o iné metódy – skúmanie pomocou behaviorálnych experimentov, ktorých dizajn, realizácia a vyhodnocovanie sú podložené štatistickou teóriou. Kognitívni psychológovia robia experimenty s ľudskými subjektmi, alebo zvieratami, formulujú teórie a vedecké hypotézy, ktoré potom overujú pomocou ďalších experimentov. Psychológia predstavuje veľmi dôležitú súčasť kognitívnej vedy a v súčasnosti sa kognitívna psychológia stáva dokonca dominantnou z hľadiska počtu publikácií (Gentner, 2010). Experimentálnych poznatkov týkajúcich sa rôznych kognitívnych funkcií (percepcia, pamäť, plánovanie, rozhodovanie, konanie a iné) je v súčasnosti veľmi veľa, no stále nie dost, pretože i v rámci psychológie existuje nejednoznačnosť v interpretácii výsledkov. To znamená, že namerané dáta sa dajú vysvetliť rôznymi spôsobmi a teda môžu byť niekedy konzistentné s viacerými teóriami.

Ďalšia materská disciplína je *lingvistika*, pretože jazyková schopnosť človeka si zasluhuje špecifické postavenie. Jazykovedci využívajú fakt, že jazyk je akousi bránou do mysle človeka (čo v prípade zvierat neplatí, na rozdiel od niektorých iných kognitívnych funkcií, ktoré človek so zvieratami zdieľa). Lingvistika, ako veda o jazyku, má svoje bohaté vnútorné členenie. Tu spomeňme len dve podoblasti, dôležité z pohľadu kognitívnej vedy. Psycholingvistika, t.j. psychológia jazyka, sa zameriava na psychologické experimenty zamerané na skúmanie jazykového správania. Neurolingvistika si všíma súvislosti medzi poškodeniami mozgu a jazykovým správaním, čo umožňuje hľadať tzv. neurálne koreláty jazykových funkcií.

Nemenej zásadnú rolu pri konštituovaní kognitívnej vedy zohrala *umelá inteligencia* a vznik digitálneho počítača, ktorý vniesol do kognitívnej vedy novú paradigmu (ale i nové otázky). Metafora mozgu ako počítača, a mysle ako jeho softvéru, navodila intuitívne zaujímavú predstavu (v duchu funkcionalizmu) o tom, ako by mohla myseľ fungovať. Okrem toho, snaha modelovať ľudské myslenie a správanie pomocou formálnych metód (matematických a algoritmických postupov) dodala skúmaniu kognície väčšiu exaktnosť. Navrhnuť inteligentné systémy, resp. autonómne robotické systémy vybavené počítačom (umelým mozgom), je dlhodobým cieľom umelej inteligencie, ktorý napriek viacerým nezdarom a sklamaniam je stále živý, aj vďaka novým, prírodou inšpirovaným prístupom, napríklad v rámci kognitívnej robotiky (Asada a spol., 2009).

V rámci súčasných empirických vied sa *kognitívna neuroveda* snaží dávať do súvisu kognitívne funkcie (mysle) s rôznymi časťami mozgu, Spočiatku to bolo hlavne pomocou behaviorálnych experimentov. K tomuto skúmaniu od 80. rokov minulého storočia výrazne napomáhajú zobrazovacie metódy mozgu, ktoré nám poskytujú pomerne detailnú informáciu o činnosti rôznych častí mozgu pri vykonávaní kognitívnych úloh. Ukazuje sa, že je dôležité skúmať mozog, ak chceme spoznať myseľ (pozri kapitolu 2). Opierame sa pritom o principiálny predpoklad, že všetky myšlienkové pochody charakterizujúce myseľ majú svoj biologický základ (korelát).

Napokon, k materským disciplinám kognitívnej vedy sa radí aj *kognitívna antropológia*, pretože človek je sociálna bytosť, a tak jeho poznanie a s tým súvisiaci vývin mysle závisí aj od kultúry. Kognitívni antropológovia sa teda pozerajú na ľudský druh ako na spoločnosť jedincov, v rámci ktorej vzniká a šíri sa poznanie a kultúrna inovácia. Kognitívna antropológia sa zaoberá otázkami týkajúcimi sa toho, čo ľudia z rôznych komunít vedia a ako ich znalosti ovplyvňujú ich vnímanie a interakciu vo svete okolo nich.

1.2.1 Interdisciplinarita

Jednou zo základných otázok kognitívnej vedy je, či je multidisciplinárna alebo interdisciplinárna. Multidisciplinarita je neintegratívna, čo znamená, že každá disciplína si zachováva svoje metodológie skúmania ako aj predpoklady, z ktorých vychádza. Na druhej strane, interdisciplinarita prekračuje tradičné hranice medzi disciplínami. V oboch prípadoch možno kognitívnu vedu metaforicky ilustrovať ako snahu o skúmanie veľkého objektu (napríklad slona) prostredníctvom skupinky nevidiacich vedcov, z ktorých každý môže preskúmať svojimi dotykmi len nejakú časť. Ich vzájomná komunikácia by však mohla vyústiť v skupinové poznanie celého objektu.

Každá zo spomínaných disciplín teda prináša so sebou svoje tehličky do mozaiky poznania kognitívnych procesov. Empirické metódy sú zdrojom informácií pre vznik teórií a výpočtových modelov, ktoré sa snažia vysvetliť namerané dáta. Tým jednak ponúkajú mechanistické vysvetlenia,¹ ako by myseľ mohla fungovať, a na druhej strane

¹ Mechanistický prístup k vysvetľovaniu nejakého fenoménu, založený na redukcionizme, netreba chápať nejakou pejoratívne, či znevažujúco. Ide o štandardný vedecký prístup, ktorého podstatou je pochopiť fenomén rozkladom na (predpokladané) časti (redukcionizmus) ale aj na interakcie medzi nimi (rekonštrukcia).

ponúkajú nové, testovateľné predikcie, ktoré inšpirujú nové experimenty. Tento cyklický proces poznávania si bude určite vyžadovať ešte veľa iterácií, a to najmä kvôli nerozhoditeľnosti medzi rôznymi teóriami ako aj nedostatku dobrých a overiteľných výpočtových modelov (Addyman a French, 2012).

1.2.2 Dôležité pojmy kognitívnej vedy

Ak sa pozrieme na rôzne pokusy o definíciu kognitívnej vedy, môžeme si v nich všimnúť niekoľko dôležitých pojmov. Napríklad filozof John Searle (1984) uvádza, že „úlohou kognitívnej vedy je charakterizovať mozog, no nie na úrovni neurónov, ani na úrovni vedomých mentálnych stavov, ale na úrovni jeho fungovania ako systému spracovania informácií.“ Jeden z otcov umelej inteligencie, Herbert Simon (1989) píše, že „kognitívna veda je štúdiom inteligencie a inteligentných systémov, so zvláštnym zreteľom na inteligentné správanie ako počítanie.“ Kognitívny vedec Paul Thagard (2001) zase zužuje pozornosť na myslenie a tvrdí, že „myslenie možno najlepšie vysvetliť v termínoch reprezentačnej štruktúry mysle a výpočtových procedúr, ktoré pracujú s týmito štruktúrami“ (Thagard, 2001). V týchto definíciách (a mnohých ďalších) sa vyskytujú viaceré kľúčové pojmy, ako *mentálne stavy*, *reprezentácie*, *spracovanie informácie*, *výpočtové operácie*, *úrovne abstrakcie*, a iné. Môžeme teda povedať, že *reprezentačno-výpočtový pohľad* na myseľ hrá v kognitívnej vede významnú úlohu. Názory sa však odlišujú v tom, aké reprezentácie a aké výpočty s nimi v myslí prebiehajú. Týmto pojmom sa preto budeme bližšie venovať v ďalšom texte.

1.2.3 Úrovne analýzy

V snahe o poznanie mozgu/mysle hrá tiež dôležitú úlohu úroveň abstrakcie. Významný britský neurovedec a psychológ David Marr (1982) predstavil svoju teóriu, v ktorej postuloval existenciu troch nezávislých, komplementárnych úrovní abstrakcie, analýzy (vysvetlenia nejakého fenoménu) – výpočtovej, algoritmickej a implementačnej. Táto teória bola rozpracovaná v kontexte Marrovhovho modelu spracovania vizuálnej informácie, a pomocou nej zdôraznil podobnosť mozgu/mysle s počítačom. *Výpočtová* úroveň definuje, aká informácia sa počíta a prečo, bez bližšej špecifikácie. *Algoritmická* úroveň už vyžaduje znalosť o používaných reprezentáciách, uchovávajúcich informácie, a o algoritmických výpočtoch s nimi. *Implementačná* úroveň sa týka opisu konkrétneho algoritmu realizovateľného na konkrétnom fyzickom substráte. Tieto úrovne sú považované za vzájomne nezávislé a platí medzi nimi vzťah, že jedno vysvetlenie na vyššej úrovni sa dá transformovať na viacero vysvetlení na nižšej úrovni (princíp viacnásobnej realizovateľnosti). Ako ilustračný príklad uvažujme násobenie dvoch viacciferných čísel. To je teda cieľ výpočtu, ktorý sa dá dosiahnuť rôznymi algoritmami, napríklad takým, ktorý človek bežne používa, a v ktorom sa medzivýsledky násobenia (jedného čísla číslicou druhého čísla) zapíšu pod seba a potom sa sčítajú. Napokon, implementačná úroveň už predstavuje konkrétnu realizáciu tohto algoritmu v nejakom fyzickom médiu (počítač, pero a papier, a i.).

1.3 Paradigmy kognitívnej vedy

Od svojich počiatkov prešla kognitívna veda za 60 rokov svojej existencie vývojom, ktorý sprevádzali štyri významné paradigmy – symbolizmus, konekcionizmus, dynamické systémy a pravdepodobnostné prístupy. Zrejme by bolo dobré najprv objasniť význam pojmu *paradigma*. Významný americký filozof vedy T. S. Kuhn zadefinoval vedeckú paradigmu ako „všeobecne uznávaný vedecký pokrok, ktorý na nejaký čas ponúka modelové problémy a riešenia pre komunitu výskumníkov“ (Kuhn, 1996, s. 10). Každá paradigma teda prináša nový pohľad na študovaný fenomén, podložený nejakou teóriou, prípadne empirickými poznatkami, a býva často spojená s kontrastom voči predchádzajúcej paradigme. Kuhn okrem iného zaviedol aj termín *posun v paradigme* (paradigm shift), ktorý vyjadruje prechod z jednej paradigmy na druhú, aj keď napríklad v prípade kognitívnej vedy môžeme povedať, že prechody medzi paradigmami sú pozvoľné, a viaceré paradigmy stále koexistujú, pretože majú svojich stúpencov (aj keď ich význam sa s časom mení).

Všetky štyri vyššie spomínané paradigmy sa dotýkajú výpočtových prístupov pri formalizácii mysle a mentálnych procesov, no majú aj svoje filozofické aspekty týkajúce sa vzťahu mysle a mozgu. Formálne výpočtové prístupy, ako bolo spomínané, ponúkajú explicitné mechanizmy fungovania mysle/mozgu, čo principiálne nemožno dosiahnuť pomocou teoretických modelov (s krabičkami a šípkami). V týchto paradigmách vystupujú vyššie spomínané pojmy, no nie vždy sa používajú s rovnakým významom, čo sa pokúsime objasniť.

1.3.1 Symbolizmus

Vznik symbolizmu (známeho aj ako *kognitivizmus* alebo *klasická paradigma*) bol stimulovaný vynájdением moderných digitálnych počítačov, ktoré realizujú diskkrétne výpočty so symbolmi. Výpočet v počítači prebieha za pomoci dvoch odlišných kľúčových komponentov: procesora a pamäte. Procesor sériovým spôsobom spracováva symboly, uložené v pamäti a vykonáva pritom inštrukcie podľa programu uloženého v inej časti pamäte. Interakcia počítača s prostredím je na periférii záujmu a prebieha prostredníctvom vstupno-výstupných podsystémov, ktoré (vhodnou transformáciou) sprostredkovávajú vstupy pre centrálny procesor a pre výstupy z neho. Klasická paradigma poníma *mysel ako výpočtový stroj*, oddeliteľný od prostredia, ktorý manipuluje s internými symbolmi podľa logických pravidiel, podobne ako to robí počítač. *Výpočtovú teóriu mysle* vystihuje hypotéza *o fyzikálnom symbolovom systéme*, ktorý podľa jej autorov „disponuje nutnými a postačujúcimi prostriedkami na všeobecné inteligentné konanie“ (Newell a Simon, 1976).² Významným počiatočným impulzom pre symbolizmus bol návrh operacionalizácie testu procesu myslenia, známeho ako Turingov test (Turing, 1950), podľa ktorého by počítač, pokiaľ by prešiel úspešne testom, vykazoval inteligentné správanie, neodlišiteľné od správania človeka.

² Čo to presne je inteligentné konanie/správanie nás teraz nemusí trápiť. Postačí, že čitateľ má o tom intuitívnu predstavu. Viac na túto tému možno nájsť napríklad v práci Šefránek, Takáč a Farkaš (2008).

Vychádzajúc zo skorších prác symbolikov (Turing, 1950; Newell a Simon, 1976; Fodor, 1983), Harnad (1990) zrekonštruoval výstižnú *definíciu symbolového systému*, kde kľúčovú úlohu hrá syntax, pričom sémantika sa dodá do systému „zvonka“ dizajnérom systému. Harnad (1990) taktiež sformuloval problém ukotvenia symbolov, ktorý možno považovať je jeden z kľúčových v oblasti kognitívnej vedy a akvizície významov (pozri kapitolu 4).

Z pohľadu symbolizmu je teda podstatné, že myslenie (a tým aj konanie) človeka je vnútorne riadené algoritmami, ktoré sú realizované programami, s využitím vhodných reprezentácií (napr. predikáty alebo propozície), nejako implementovanými v mozgu. Ako, to symbolizmus nerieši, pretože sa to považuje len za vec implementácie. Pri počítačoch môžeme ľubovoľný algoritmus naprogramovať v nejakom programovacom jazyku a potom ho implementovať vo zvolenej hardvérovej platforme. Podobná predstava prevláda i v prípade mozgu. Nezáleží ako sú v mozgu implementované mentálne procesy.

Symbolový (algoritmický) pohľad na myseľ je dodnes viacerými odborníkmi považovaný za tú správnu úroveň abstrakcie, a to najmä z troch dôvodov (Barsalou, 2008): (1) Poskytuje elegantné, transparentné, expresívne (silné) formalizmy na reprezentáciu znalostí, (2) Zachytáva dôležité ľudské intuície týkajúce sa symbolového charakteru kognície. Je možné, že ide o jednu z početných kognitívnych predispozícií človeka (cognitive biases) smerom k diskretným formalizmom, pretože operácie používané na symboloch sú blízke ľudskému mysleniu. (3) Dá sa dobre implementovať v umelej inteligencii.³ Symbolizmus vníma kogníciu ako manipuláciu tzv. amodálnych symbolov (t.j. nezávislých od modalít ako je zrak, sluch a iné zmysly) algoritmickým spôsobom, čo vedie ku klasickému počítaniu, ktoré možno implementovať v štandardných počítačoch. Symbolizmus sa používa najmä na vysvetľovanie procesov tzv. *vyššej kognície*, ako je usudzovanie, plánovanie, a tiež používanie jazyka (napr. Fodor, 2000; Šeřránek, 2002; Pinker 2009).

1.3.2 Konekcionizmus

Konekcionizmus predstavuje spektrum metód, ktoré vznikli v rámci umelej inteligencie, a ktoré boli inšpirované architektúrou a fungovaním mozgu. Môžeme teda hovoriť o inšpirácii „zdola.“ Konekcionistický systém pozostáva z množiny jednoduchých prvkov – umelých neurónov, navzájom komunikujúcich cez váhované spojenia, ktoré predstavujú dlhodobú pamäť systému. Sila konekcionistického systému – *umelej neurónovej siete* – nie je v samotných neurónoch, ale v ich *vzájomnom prepojení a interakcii*. Paralelné spracovanie a distribuovanosť aktivity predstavujú základný architekturný rozdiel v porovnaní so symbolovým systémom, pretože každý neurón je súčasne procesorom aj pamäťou (aj keď elementárnou). Povaha komunikácie medzi neurónmi má *numerický* a nie symbolový charakter, preto hovoríme v prípade sietí o tzv. subsymbolových reprezentáciách.

Neurónová sieť môže pracovať so spojitými hodnotami, čo je veľmi vhodným nástrojom na modelovanie kognície, pretože empirické dáta nasvedčujú tomu, že mnoho

³ Umeľá inteligencia a výpočtová kognitívna veda majú veľa spoločného, no dá sa zhruba povedať, že v umelej inteligencii menej záleží na biologickej prijateľnosti modelu.

kognitívnych procesov má spojité prejavy. Napríklad, nielenže perceptuálne vstupy (napr. obraz na sietnici) a motorické výstupy bývajú spojité (pohyby, zvuková podoba reči), ale podobne môžeme argumentovať aj v prípade modelovania vyšších procesov, napr. rozhodovania (Spivey, 2007). Spomeňme si, aké ťažké môže byť niekedy rozhodnúť sa medzi dvoma alternatívami. Tento proces „súperenia“ alternatív sa dá vysvetliť pomocou neurónovej siete so spojitými aktiváciami. Neurónová sieť teda spracováva numerické dáta v podobe vektorov. Operácie nad takýmito vektorovými dátami sú (nelineárnymi) *transformáciami*, čo má svoje implikácie. Je zrejmé, že takéto operácie nie sú vôbec transparentné, a preto sú človeku málo zrozumiteľné, a aj preto si neurónové siete vyslúžili prívlastok „čierna skrinka.“ Môžeme však použiť techniky zhukovania a vizualizácie vysokorozmerných dát, ktoré nám pomáhajú odhaliť, čo sa v neurónovej sieti deje (O'Brien a Opie, 2006).

Dôležitou vlastnosťou umelých neurónových sietí je to, že na rozdiel od symbolovej umelej inteligencie, kde znalosť do systému vkladá dizajnér, ponúkajú biologicky inšpirované mechanizmy akvizície znalostí, *učenia*, aplikovateľné na rôznych úrovniach abstrakcie (Haykin, 2008; Kvasnička a spol., 1997). To znamená, že parametre siete (váhy spojení) nemusíme explicitne nastaviť (čo sa prakticky ani nedá), no treba sieti poskytnúť vhodný mechanizmus učenia. Podobne ako človek, neurónová sieť sa učí na príkladoch. Učenie spôsobuje zmenu váh tak, aby sa po naučení dosiahlo požadované správanie siete, čo formálne odpovedá hľadaniu extrému (minima, resp. maxima) nejakej hodnotovej funkcie.

Na rozdiel od symbolizmu, konekcionizmus zahŕňa aj implementačnú úroveň analýzy. Niektorí filozofi zastávajú názor, že toto nie je podstatné, pretože vysvetľované fenomény treba oddeliť od mechanizmov ich vzniku (Abrahamsen a Bechtel, 2006), avšak v konekcionistických modeloch sú tieto dva aspekty spojené. To znamená, že existuje závislosť medzi implementovanými mechanizmami (danými neurálnymi interakciami) a kognitívnymi fenoménmi (O'Reilly a Munakata, 2000). Inými slovami, špecifikáciu neurónovej siete môžeme vytvoriť na výpočtovej úrovni (prostredníctvom funkcie, ktorú chceme realizovať), no elementárne funkcie (týkajúce sa neurónov) už umožňujú priamu implementáciu v hardvéri, s relatívne malou abstrakciou od biologických neurónových sietí.

V prípade neurónových sietí tiež bežne hovoríme, že počítajú a spracovávajú informácie. Počítajú však rovnako ako symbolové systémy? Detailnejší pohľad na existujúce modely neurónových sietí nám napovedá, že tie nie sú homogénnou skupinou, ale naopak, spektrom metód z pohľadu typov aktivácií neurónov a ich dynamiky. Modely, ktoré pracujú v spojitom priestore, a niektoré dokonca v spojitom čase, sú matematickými modelmi, ktoré sa dajú presne opísať len pomocou (nelineárnych) diferenciálnych rovníc, a nie pomocou diskretných krokov výpočtu (počítačová implementácia ponúka len aproximáciu). Zahrnutie takýchto modelov si vyžaduje voľnejšiu definíciu počítania, o čom pojednávame ďalej. Od modelov neurónových sietí pracujúcich v spojitom čase sa priamo dostávame k paradigme dynamických systémov.

1.3.3 Dynamické systémy

Čo klasická paradigma a konekcionizmus zdieľajú, je ich izolovanosť od okolitého prostredia. V prípade symbolového modelu je to zjavné, pretože nezávislosť od prostredia je jedným z predpokladov. Neurónová sieť sa síce učí, čiže interaguje s prostredím, no bežným predpokladom je, že všetky tréningové dáta máme hneď dispozícii, takže interakcia nie je „plnohodnotná,“ pretože výstupy siete nemajú dopad na prostredie.

Takýto pohľad na kogníciu sa stal trňom v oku niektorým kognitívnym vedcom, ktorí prišli s myšlienkou kognitívneho systému ako dynamického systému, ktorý predstavuje proces interakcie s prostredím v spojitom čase a priestore (Thelen a Smith, 1994; Port a van Gelder, 1995). Kognitívne procesy teda nie sú chápané izolovane od prostredia, ale v úzkej previazanosti s ním (coupling). Kognitívne procesy ako aj konanie človeka sú aktivitou v čase, ktorú môžeme opísať pomocou systému nelineárnych diferenciálnych rovníc. Táto paradigma, postavená na teórii nelineárnych dynamických systémov predpokladá existenciu dvoch kľúčových, vzájomne prepojených, typov premenných. Jedným sú tzv. *kolektívne premenné*, ktoré zahŕňajú vzťahy medzi interagujúcimi zložkami dynamického systému, a vysvetľujú správanie systému. Druhým typom sú *riadiace parametre*, ktorých kvantitatívne zmeny môžu spôsobovať kvalitatívne zmeny v správaní systému (tzv. fázové prechody). Príkladom takéhoto dynamického systému môže byť pohyb koňa, ktorého rytmus pohybu (vyjadrený periodickými priebehmi kolektívnych premenných) závisí od rýchlosti pohybu (riadiaci parameter). Kôň sa môže pohybovať v každom čase len v jednom rytme, (fázové) prechody medzi nimi nemusia nutne nastávať pri tých istých rýchlostiach, a závisia aj od stavu prostredia (napr. hrbolatá cesta). Súčasťou takéhoto dynamického opisu kognitívneho procesu sa teda priamo stáva aj prostredie, čo rozširuje konceptuálne chápanie mysle. Bressler a Kelso (2001) argumentujú, že dynamický pohľad je relevantný aj pre charakterizáciu samotných mentálnych procesov (vnútri kognitívneho systému), hoci identifikácia kolektívnych premenných a riadiacich parametrov v tomto prípade je zatiaľ tvrdým orieškom. Okrem toho, zmeny parametrov bývajú nastavované zvonka (dizajnérom), čo je tiež slabou stránkou dynamických systémov (McClelland, 2009).

Dynamický pohľad na spoznávanie vonkajšieho sveta predstavuje pilier *zjednávacieho (enaktívneho) prístupu* (Varela, Thompson a Rosch, 1991), kde sa dôraz z vnútorných reprezentácií (vopred daného) vonkajšieho sveta presúva na *vnímanie a jednanie vo svete*, ktorý sa takto spoluvytvára. Tým vzniká prostredníctvom spätnej väzby kruhový proces, pretože zjednávaný svet je prostredím pre zjednávanie (enactment), ale zjednávanie je súčasne predpokladom zjednávaného sveta. V rámci tohto pohľadu si každý subjekt (kognitívny systém) zjedná svoj svet sám v procese dynamickej interakcie so svetom (vrátane iných jedincov), a ten bude jedinečný. Zjednávací prístup je teda založený na konceptoch ako skúsenosť, autonómia, emergencia a vytváranie významov (sense-making) prostredníctvom skúsenosti. Táto perspektíva je predmetom aj niektorých ostatných kapitol (4 a 5).

1.3.4 Pravdepodobnostné modely

Pravdepodobnostný prístup⁴ sa stal populárny v kognitívnej vede najmä v ostatnej dekáde rokov (Perfors, Tenenbaum, Griffiths a Xu, 2011). Tento teoreticky podložený Bayesovský prístup ktorý využíva široké spektrum reprezentácií (stromy, vektory, logické pravidlá atď.), kombinuje ich so štatistickým učením a tvorením inferencií (t.j. nových poznatkov na základe dedukcie) za prítomnosti neurčitosti, poskytuje vysvetlenia na výpočtovej úrovni (v zmysle Marrovej taxonómie). Implementačná úroveň je považovaná za menej dôležitú, nechávajúc priestor pre hľadanie potenciálnych odpovedajúcich neurálnych mechanizmov. Na rozdiel od konekcionizmu, ktorý začína zdola nahor definovaním mechanizmov, pravdepodobnostný prístup ide smerom zhora nadol, začínajúc funkciou, ktorú chceme vysvetliť, hľadajúc optimálnu reprezentáciu na vysvetlenie dát. Tento prístup zahŕňa aj dimenziu vrodene–získané (nature–nurture) prostredníctvom tzv. indukčných výchylok (inductive biases), čo sú apriórne distribúcie na množine hypotéz (vyplývajúce zo znalosti domény). Tie tiež vstupujú do výpočtov posteriorných distribúcií, t.j. tých, ktoré zohľadňujú aj pozorované dáta. Hypotézy si môžeme predstaviť ako rôzne riešenia alebo kategórie (napr. pri klasifikácii dát), ktoré máme zistiť na základe apriórnych znalostí (o svete), ako aj pozorovaných dát. Bayesovský prístup de facto kalkuluje s rôznymi riešeniami paralelne a s prichádzajúcimi dátami ich aktualizuje, čo je v podstate tiež forma učenia na vyššej úrovni abstrakcie.

Pravdepodobnostný prístup dokázal vysvetliť rôzne prejavy ľudského správania (pozri článok Griffiths, Kemp a Tenenbaum, 2010, a tamojšie referencie). Vysvetlenia na výpočtovej úrovni sú považované nielen za výpočtovú abstrakciu skrytých mechanizmov, ale ich zástancovia tvrdia, že majú nezávislú výpovednú hodnotu (platnosť). McClelland (2009) však načrtáva niektoré otázky súvisiace s týmto prístupom. Pravdepodobnostný prístup vychádza z presvedčenia, že ľudskú kogníciu je možné chápať ako sériu optimálnych odpovedí (čo vedie k racionálnemu správaniu), vzhľadom na obmedzenia, ktoré človek v danej situácii a kontexte. Vieme však, že človek niekedy koná neracionálne, aj pod vplyvom emócií (teória ohraničenej racionality). Okrem toho, problémom Bayesovských prístupov je ich výpočtová neúnosnosť v prípade modelovania zložitejších problémov. Na druhej strane možno predpovedať, že zvýšený nárast o tieto prístupy v ostatných rokoch predznamenáva ich možnú dôležitosť v budúcnosti. Bude zaujímavé objavovať potenciálne linky medzi pravdepodobnostnou výpočtovou úrovňou a neurálnou implementačnou úrovňou pri snahe pochopiť ľudskú myseľ.

1.4 Reprezentácia

Ako sme avizovali vyššie, pozrime sa teraz bližšie na pojem *reprezentácia*. Tento pojem sa bežne používa nielen v umelej inteligencii, ale aj v kognitívnej psychológii, či lingvistiky. Jazykový pôvod tohto pojmu napovedá, že ide o re-reprezentovanie niečoho

⁴ Na rozdiel od predchádzajúcich troch paradigiem, pravdepodobnostný prístup zrejme nenájdeme medzi paradigmami v bežnej učebnici kognitívnej vedy. V literatúre niektorí hovoria o modelovacom prístupe, no používajú sa aj spojenia Bayesovská paradigma. Nazdávam sa, že Bayesovský prístup sa dosť odlišuje od ostatných paradigiem, preto medzi ne tiež patrí.

vonkajšieho (vo svete) niekde inde, vo vnútri nejakého systému (živého alebo umelého) alebo i na papieri. V kognitívnej psychológii sa hovorí o *mentálnych reprezentáciách*, čo je v podstate prvý teoretický konštrukt v kognitívnej vede (Pitt, 2008). Mentálna reprezentácia je základným konceptom výpočtovej teórie mysle, podľa ktorej sú kognitívne stavy a procesy konštituované výskytom, transformáciou a uchovávaním (v myšli/mozgu) reprezentácií (t.j. štruktúr nesúcich informáciu) nejakého typu. Korene pojmu mentálna reprezentácia však siahajú až do antických čias, keď širšie chápanie tohto pojmu medzi filozofmi nemalo výpočtový charakter.⁵ Až neskorší predpoklad, že reprezentácia je (abstraktný) objekt so sémantickými vlastnosťami (pozri aj kapitolu 4), dal tomuto konceptu výpočtový charakter, pretože s takto definovanými objektmi by sa už podľa predpokladu dalo manipulovať.

S nástupom umelej inteligencie sa hlavným cieľom stala *reprezentácia znalostí*, t.j. snaha o formalizáciu znalostí takým spôsobom, aby sa dali uložiť a používať v počítači (symbolovom systéme). Tu zohrali dôležitú úlohu lingvistika a logika, ktoré ponúkli existujúce formalizmy na rôzne implementácie reprezentácie znalostí. Môžeme povedať, že všetky znalosti (tzv. báza znalostí) sú z pohľadu klasickej paradigmy *explicitné a symbolové*, a keď systém disponuje pravidlami odvodzovania (inferenčné mechanizmy), dokáže si odvodiť ďalšie znalosti. Dôležitou vlastnosťou reprezentácií v symbolovom systéme je, že sú stabilné. Môžu sa síce v čase meniť (ak je to súčasťou programu), ale sú väčšinou kontextovo nezávislé. Napríklad, koncept⁶ *strom* je zväčša reprezentovaný rovnakým amodálnym symbolom (alebo štruktúrou) vo všetkých prípadoch, a ak nie, tak ku symbolu reprezentujúceho strom treba pridať iný symbol (alebo štruktúru), ktorý daný kontext špecifikuje.

V prípade konecionizmu majú reprezentácie inú povahu. Keďže vieme, že dlhodobou pamäťou natrénovanej siete sú spojenia medzi neurónmi i so svojimi hodnotami, znalosť je v sieti distribuovaná, a reprezentovaná *implicitne a numericky*. Analýzou natrénovanej siete však vieme zväčša identifikovať, čo konkrétne neuróny a spojenia medzi nimi zabezpečujú, aj keď komplikáciou je práve distribuovanosť dlhodober pamäti. Krátkodobá pamäť je realizovaná pomocou aktivácie neurónov, ako odpovede na nejaký podnet (vonkajší alebo vnútorný). Miera distribuovanosti predstavuje kontinuum, od úplnej distribuovanosti, cez riedke reprezentácie (sparse coding) až po lokalistické reprezentácie (ktoré odpovedajú symbolovej reprezentácii).

V doméne dynamických systémov existuje radikálna téza, ktorá nepredpokladá existenciu reprezentácií (Thelen a Smith, 1994). Dynamické procesy pracujú v spojitom čase, a preto v žiadnom časovom okamihu nemôžeme predpokladať, že niečo z vonkajšieho sveta by mohlo byť aktuálne reprezentované vnútri systému, pretože všetko sa stále mení. Takýto postoj sa objavil aj v umelej inteligencii, keď Brooks (1986)

⁵ No už bolo možné hovoriť o reprezentačnej teórii mysle, ktorá pracovala s pojmami ako sú *myšlienky, presvedčenia, túžby*, či *obrazy*, a ktorá sa v kognitívnej vede voľne zamieňa s výpočtovou teóriou mysle.

⁶ V literatúre neexistuje konsenzus o definícii termínu *koncept*. Náš pohľad je konzistentný s perspektívou kognitívnej sémantiky, t.j. koncept chápeme ako mentálny konštrukt (vzťahujúci sa obyčajne na nejakú kategóriu), nezávislý od jazyka. Tým sa líši od *pojmu*, ktorý sa týka prirodzeného jazyka. Oba termíny spolu súvisia: aby bol pojem pojmom, musí byť k nemu priradený nejaký význam vyjadrený konceptom. Kognitívnej sémantike je venovaná kapitola 4.

navrhol tzv. *subsumpčnú architektúru* inteligentného systému v robotike ako alternatívu k tradičnej symbolovej architektúre. Subsumpčná architektúra je hierarchická, symbolová, no správanie robota je čisto reaktívne, čiže nevyžaduje vnútorné reprezentácie (pozri kapitolu 5). Tento trend v robotike mal snahu priblížiť umelú inteligenciu ku kognitívnej vede, no nezachytáva aspekty vyššej kognitívnej činnosti, ktorá sa nezaobíde bez vnútorných mentálnych reprezentácií (Gärdenfors, 1996). Preto reprezentačno-výpočtový prístup je kompatibilný s mäkkšou verziou dynamickej paradigmy (Kelso, 1995; Port a van Gelder, 1995), resp. s pohľadom na mozog ako na samoorganizujúci sa systém (Farkaš, 2006).

Pravdepodobnostné modelovanie využíva celé spektrum reprezentácií (symbolové, numerické, grafové), a principiálne nemá žiadne obmedzenia v tomto smere, na rozdiel od ostatných paradigiem. Znalosť v systéme je daná pomocou pravdepodobností a tie nadobúdajú spojité hodnoty. Podobnosť s konekcionizmom spočíva v schopnosti učenia a tiež v tom, že niektoré formy konekcionistického učenia majú prirodzenú Bayesovskú interpretáciu (napr. Mackay, 1996). Okrem toho, niektoré tzv. generatívne konekcionistické modely využívajú pravdepodobnostný rámec, pričom si zachovávajú svoj „konekcionistický štýl“ (Bengio, 2009).

1.5 Počítanie z pohľadu paradigiem

Všetky spomínané paradigmy sa síce dotýkajú počítania, ale bližší pohľad na tento koncept nám prezradí, že ho musíme poňať inak, ako nám napovedá intuícia a tradičné používanie tohto pojmu. Taktiež, úloha výpočtových modelov v kognitívnej vede je považovaná za dôležitú (McClelland, 2009; Farkaš, 2012) no kontroverznú, práve z tohto dôvodu. Neexistuje konsenzus o význame kľúčových pojmov ako sú počítanie alebo spracovanie informácie (Piccicini a Scarantino, 2011).

Vznik moderného digitálneho počítača v 50. rokoch minulého storočia bol jednoznačne dôležitým míľnikom, ktorý stimuloval pokrok v mnohých vedných disciplínach, vrátane umelej inteligencie a kognitívnej vedy. Avšak, svet počítania vznikol dávno pred konceptom Turingovho stroja (Turing, 1936). Sloman (2002) uvádza historický kontext, ktorý tomu predchádzal prostredníctvom návrhu rôznych typov strojov a zariadení, u ktorých by sa dalo diskutovať či počítajú alebo nie (napríklad tkáčske stroje, rôzne regulátory, triediace stroje a iné). Jedine v prípade univerzálneho digitálneho počítača nikto nepochybuje, že počíta (pretože na ten účel bol skonštruovaný), v prípade ostatných systémov sa názory rôznia (Farkaš, 2012). Pochybnosti sa týkajú samozrejme i biologických systémov, vrátane mozgu, aj keď si intuitívne možno myslíme, že mozog je určite výpočtový systém (Churchland a Sejnowski, 1992). Táto otázka sa dotýka aj filozofie, pretože jedna vec je, čím mozog je (ontologicky), a druhá (epistemologická), ako ho môžeme a chceme poznávať (napríklad ako výpočtový systém). Výpočtový pohľad v kognitívnej vede je však napriek tomu kľúčový, lebo poskytuje pridanú hodnotu, ako sme už spomínali vyššie (Farkaš, 2012).

Vysvetleniu pojmu *počítanie* a jeho využiteľnosti v kognitívnej vede môže napomôcť rozšírenie obsahu tohto pojmu. V užšom slova zmysle sa počítanie spája s diskretnými výpočtami v diskretnom čase (klasické počítanie), zatiaľ čo v širšom slova

zmysle zahŕňa neklasické počítanie, ako napríklad analógové, kvantové alebo pravdepodobnostné výpočty (Piccicini a Scarantino, 2011). Motiváciou pre tento posun sú spomínané paradigmy kognitívnej vedy. Čo sa týka konekcionizmu, počítanie v neurónovej sieti nemusí byť to isté ako počítanie v symbolovom systéme, ktorý vykonáva program, realizujúci nejaký algoritmus sériovým spôsobom. Pracuje v diskretnom čase, s diskretnými entitami, symbolmi. Neurónové siete však nie sú homogénnou skupinou metód, pretože niektoré modely neurónových sietí počítajú klasicky, niektoré neklasicky a niektoré nepočítajú vôbec (Piccicini, 2008). To závisí od toho, či pracujeme v spojitom alebo diskretnom priestore a čase. Debata ohľadne rozdielov medzi konekcionistickými a symbolovými prístupmi možno nájsť v početnej literatúre. V kognitívnej vede prevláda názor, že ľudská myseľ realizuje výpočty, a teda spracováva informácie, pričom my sa tieto procesy snažíme vysvetliť pomocou formálnych modelov. Viacerí zástancovia konekcionizmu zastávajú názor, že mozog vykonáva výpočty a že neurálne počítanie vysvetľuje ľudskú kogníciu (napr. Feldman a Ballard, 1982; Churchland a Sejnowski, 1992; Bechtel a Abrahamsen, 2002; O'Brien a Opie, 2006). V čom sa líšia názory konekcionistov a symbolikov je otázka, či konekcionistické modely sú lepšími výpočtovými modelmi než klasické symbolové modely. Väčšina konekcionistov tvrdí, že áno, symbolisti zase, že nie, lebo obe vysvetlenia sú vzájomne transformovateľné (opierajúc sa o spomenuté dôkazy o ekvivalencii). Niektorí symbolici tvrdia, že neurónové siete nepočítajú vôbec, alebo ak, tak len vtedy, keď implementujú symbolový systém (Fodor a Pylyshyn, 1988). Niektorí prívrženci konekcionizmu zase zastávajú názor, že mozog nie je výpočtovým systémom (napr. Edelman, 1992; Freeman, 2001), a že naše výpočtové modelovanie je len jeho diskretnou aproximáciou.

Čo sa týka ostatných paradigiem, v zmysle rozšírenej definície počítania dynamické systémy nepočítajú vôbec (radikálna téza), alebo počítajú neklasicky (reprezentačno-výpočtová pozícia). Pravdepodobnostné modely počítajú neklasicky, pretože pracujú (obyčajne) v diskretnom priestore spojitých pravdepodobností. Vo všetkých prípadoch však počítač môžeme použiť ako nástroj na aproximáciu týchto modelov.⁷

1.6 Smerom k ukotvenej kognícii

Vývoj v kognitívnej vede možno pozorovať aj v kognitívnej psychológii, ktorá prechádzala rôznymi štádiami. Venujme ešte pozornosť hlavným prúdom a teoretickým koncepciám vo vývoji kognitívnej vedy, ktoré prispeli k posunu v dominantnosti jednotlivých paradigiem (Gentner, 2010). Už v 70. rokoch minulého storočia sa v kognitívnej psychológii objavovali skeptické pohľady na centrálnosť (vnútorných) reprezentácií a symbolového spracovania informácie v ľudskej kognícii. Gibson (1970), hlavný predstaviteľ prúdu známeho ako *ekologická psychológia*, zdôrazňoval, že „nie je dôležité, čo je vo vnútri hlavy, ale vnútri čoho (akého prostredia) sa hlava nachádza.“

Tento prúd plynulo prešiel v 80. rokoch do teórie *situovanej kognície*, ktorá bola postavená na predpokladoch, že ľudská inteligencia je fundamentálne interaktívna

⁷ To že diskretný počítač úspešne používame na modelovanie kognitívnych procesov nehovorí nič o tom, aká je povaha ľudskej kognície.

(Hutchins, 1996), a že ľudská kognícia je fundamentálne neoddeliteľná od kontextu (Lave, 1988). Podľa prvého predpokladu človek bezprostredne využíva prostredie (iných ľudí, zariadenia, situácie, atď.) pre svoje inteligentné konanie, a je s nimi spojený v rámci spätnoväzbovej slučky. Druhý predpoklad viedol k rozšírenému chápaniu mysle a kognície v zmysle jej umiestnenia aj do prostredia – myseľ človeka nesídlila len v jeho mozgu, ale jej súčasťou sú aj prvky prostredia (napr. knihy, obrazy, počítač, rôzne poznámky a i.).

Dôraz na interakciu s prostredím ostal v centre pozornosti aj naďalej, no v 90. rokoch sa začal klásť dôraz na reprezentácie, ktoré túto interakciu umožňovali, a to v rámci prúdu *stelesnenej kognície*. Stelesnená (embodied) kognícia, podobne ako tvrdil Gibson, vychádza z predpokladu, že pochopenie mentálnych procesov nemožno oddeliť od tela človeka, konkrétne jeho zmyslov a motoriky. Stelesnená kognícia súčasne predpokladá, že informácia vstupuje do tela organizmu v podobe modálnych reprezentácií a podieľa sa na ďalších, vyšších procesoch kognície (Barsalou, 1999). Zmysly určujú povahu perceptuálnych vstupov, a motorické ústrojenstvo zase určuje spektrum akcií, ktoré človek môže vykonať, a ktoré tiež predstavujú obmedzenia pre vznik a modifikáciu vnútorných reprezentácií. Napríklad, mentálna reprezentácia objektu *kladivo* nezahŕňa len jeho perceptuálne vlastnosti, ale aj možnosti použitia tohto objektu (ak s ním dotýčný subjekt má skúsenosť).

V komunite kognitívnej vedy sa objavilo viacero rôznych no súvisiacich pohľadov na stelesnenie a situovanosť kognície (Gibbs, 2006; Wilson, 2002). Barsalou (2008) navrhol termín *ukotvená kognícia*, ktorý zjednocuje rôzne aspekty stelesnenia a situovania v prostredí. Silnú podporu pre túto paradigmu v kognitívnej vede možno nájsť v narastajúcej empirickej evidencii (za ostatných 20 rokov), či už z behaviorálnych štúdií alebo zobrazovacích metód (pozri početné referencie v Barsalou, 2008). Teórie ukotvenej kognície sú síce opisné, podobne ako klasické teórie, no ponúkajú testovateľné hypotézy pre ďalší empirický výskum (napr. Lakoff a Johnson, 1999; Smith, 2005) ako aj priestor pre výpočtové modely. Zaujímavou predikciou týchto teórií je to, že všetky kognitívne funkcie (vrátane tých vyšších) sú redukovateľné na (nízkoúrovňové) senzomotorické procesy. Inými slovami, neurálne reprezentácie vyšších a nižších kognitívnych procesov majú prienik. Tento pohľad na myseľ má svoju relevantnosť aj ohľadne jednej z kľúčových otázok v kognitívnej vede, ktorou je otázka významov (kapitola 4). Táto paradigma sa v literatúre nazýva aj postkognitivistická (kapitola 5).

1.7 Záver

V rámci prehľadu paradigiem kognitívnej vedy sme sa snažili objasniť ich podstatné charakteristiky, podobnosti a rozdiely. Spomenuli sme kľúčové pojmy kognitívnej vedy – reprezentácia a počítanie, a vysvetlili sme, že pojem počítanie musíme chápať širšie než ponúka počítačová metafora mysle v duchu klasického počítania. Výpočtové modelovanie v kognitívnej vede možno považovať za nevyhnutný metodologický komponent, ktorý ponúka explicitné mechanizmy vysvetlenia ľudskej kognície, a teda procesov mysle.

Literatúra

- Abrahamsen A., Bechtel W. (2006). Phenomena and mechanisms: Putting the symbolic, connectionist, and dynamical systems debate in broader perspective. In R. Stainton (Ed.), *Contemporary Debates in Cognitive Science*. Oxford: Basil Blackwell.
- Addyman, C., French, R. (2012). Computational modeling in cognitive science: A manifesto for change. *Topics in Cognitive Science*, 4: 332-341.
- Asada M. a spol. (2009). Cognitive developmental robotics: a survey. *IEEE Transactions on Autonomous Mental Development* 1(1), 12-34.
- Barsalou L. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22(4), 577-660.
- Barsalou L. (2008). Grounded cognition. *Annual Reviews of Psychology*, 59, 617-645.
- Bengio, Y. (2009). Learning deep architectures for AI. *Foundations and Trends in Machine Learning*, 2(1): 1-127.
- Bressler S., Kelso J.S. (2001). Cortical coordination dynamics and cognition. *Trends in Cognitive Sciences*, 5, 26-36.
- Brooks R. (1986). A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation*, 2(1), 14-23.
- Churchland P., Sejnowski, T. (1992). *The Computational Brain*. Cambridge, MA: MIT Press.
- Edelman G. (1992). *Bright Air, Brilliant Fire: On the Matter of the Mind*. New York: Basic Books.
- Farkaš I. (2006). Samoorganizácia ako hybná sila dynamických vzorcov aktivít v mozgu a mysli. *Kognice a umělý život VI*, Opava: Slezská univerzita, 143-148.
- Farkaš I. (2011). Konekcionalizmus v náručí výpočtovej kognitívnej vedy. In Kvasnička V. et al. (eds.), *Umělá inteligencia a kognitívna veda III*, 19-62.
- Farkaš I. (2012). Indispensability of computational modeling in cognitive science. *Journal of Cognitive Science*, 13, 211-235.
- Feldman J., Ballard D. (1982). Connectionist models and their properties. *Cognitive Science*, 6, 205-254.
- Fodor J. (1983). *The Modularity of Mind*. Cambridge, MA: MIT Press.
- Fodor J. (2000). *The Mind Doesn't Work That Way*. MIT Press.
- Fodor J., Pylyshyn Z. (1988). Connectionism and cognitive architecture. *Cognition*, 28, 3-71.
- Freeman W. (2001). *How Brains Make Up Their Minds*. New York: Columbia University Press.
- Gardner H. (1997). *The Minds New Science: A History of the Cognitive Revolution*. New York: Basic Books.
- Gärdenfors P. (1996). Cued and detached representations in animal cognition. *Behavioural Processes*, 36, 263-273.
- Gentner D. (2010). Psychology in cognitive science: 1978-2038. *Topics in Cognitive Science*, 2, 328-344.
- Gibbs R. (2006). *Embodiment and Cognitive Science*. New York: Cambridge University Press.

- Gibson J. (1970). *Ecological Approach to Visual Perception*. Boston: Houghton-Mifflin.
- Greco A. (2012). Cognitive science and cognitive sciences. *Journal of Cognitive Science*, 13, 471-485.
- Griffiths T., Kemp C., Tenenbaum J. (2008). Bayesian models of cognition. In R. Sun (Ed.), *Cambridge Handbook of Computational Psychology*, New York: Cambridge University Press, 59-100.
- Harnad S. (1990). The symbol grounding problem. *Physica D*, 42, 335-346.
- Haykin S. (2008). *Neural Networks and Learning Machines*, Prentice Hall.
- Hutchins E. (1996). *Cognition in the Wild*. MIT Press.
- Kelso J. (1995). *Dynamic Patterns: The Self-Organization of Brain and Behavior*. The MIT Press, Cambridge, MA.
- Kuhn, T. (1996). *The Structure of Scientific Revolutions* (3. vyd.), Chicago: University of Chicago Press.
- Kvasnička V., Beňušková E., Pospichal J., Farkaš I., Tiňo P., Král' A. (1997). *Úvod do teórie neurónových sietí*. Iris, Bratislava.
- Lakoff G., Johnson M. (1999). *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. New York: Basic Books.
- Lave J. (1988). *Cognition in Practice: Mind, Mathematics and Culture in Everyday Life*. New York: Cambridge University Press.
- Marr D. (1982). *Vision*. San Francisco, CA: W.H. Freeman.
- McClelland J. (2009). The place of modeling in cognitive science. *Topics in Cognitive Science*, 1(1), 11-38.
- O'Brien G., Opie J. (2006). How do connectionist networks compute? *Cognitive Processing*, 7, 30-41.
- O'Reilly R., Munakata Y. (2000). *Computational Explorations in Cognitive Neuroscience: Understanding the Mind by Simulating the Brain*. MIT Press.
- Piccinini G. (2008). Some neural networks compute, others don't. *Neural Networks*, 21(2-3), 311-321.
- Pinker S. (2009). *How the Mind Works*. W.W. Norton and Co.
- Pitt D. (2008). Mental representation, *The Stanford Encyclopedia of Philosophy*, <http://plato.stanford.edu/archives/fall2008/entries/mental-representation>.
- Port R., van Gelder T., Eds. (1995). *Mind as Motion: Explorations in the Dynamics of Cognition*. Cambridge, MA: MIT Press.
- Searle J. (1980). Minds, Brains, and Programs. *Behavioral and Brain Sciences*, 3, 417-458.
- Smith L. (2005). Cognition as a dynamic system: principles from embodiment. *Developmental Review*, 25, 278-98.
- Spivey M. (2007). *The Continuity of Mind*. Oxford: Oxford University Press.
- Šefránek J. (2002). *Inteligencia ako výpočet*. IRIS Bratislava.
- Šefránek J., Takáč M., Farkaš I. (2008). Vznik inteligencie v umelých systémoch. V knihe Magdolen, D. (zost.): *Hmota, život, inteligencia: Vznik*. VEDA, Bratislava, 245-270.
- Thelen E., Smith L. (1994). *A Dynamic Systems Approach to the Development of Cognition and Action*. Cambridge, MA: MIT Press.

- Turing, A. (1936). On computable numbers, with an application to the Entscheidungs problem. *Proceedings of the London Mathematical Society*, 2(42): 230-65.
- Turing A. (1950). Computing machinery and intelligence. *Mind*, 59, 433-460.
- Varela F., Thompson E., Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT Press.
- Wilson M. (2002). Six views of embodied cognition. *Psychonomics Bulletin Review*, 9(4), 625-636.
- Ziemke T. (1999). Rethinking grounding. In: Riegler A. et al. (eds.), *Understanding Representation in the Cognitive Sciences*, New York: Plenum Press, 177-190.