

6 Kognícia bez mentálnych procesov

Ján Šefránek

6.1 Úvod

Táto kapitola má v kontexte celej knihy svoju parketu – má prezentovať tzv. symbolový prístup ku kognitívnej vede. Napriek tomu, že v závere kapitoly sa pokúsime odmietnuť toto označenie (ako zahmlievajúce a nedostatočne výstižné), začneme stručnou charakterizáciou „symbolového“ prístupu.

Podľa Gärdenforsa (2000) symbolový prístup vychádza z predpokladu, že kognitívny¹ systém možno považovať za Turingov stroj² a kogníciu za výpočet, manipulujúci so symbolmi. Tradične sa za príklady symbolového prístupu považujú formálne modely, založené na logike alebo na formálnych jazykoch a aj implementácie, vychádzajúce z týchto modelov. Štúdiu a konštrukcii takýchto modelov sa venuje časť umelej inteligencie, označovaná (tiež) prívlastkom symbolová.

Jedným z našich cieľov je ukázať, ako umelá inteligencia a logika prispievajú k chápaniu kognície. Dá sa to povedať aj silnejšie: chceme ukázať, že logika a na logike založená umelá inteligencia sú o kognícii (nie iba o „vymyslených umelých jazykoch“ alebo o mystických „strojoch, ktoré myslia“). Často budeme hovoriť striedavo o umelej

¹ Slovo kognícia nie je v slovenčine zaužívané. Napriek tomu ho budeme používať a orientačne sa budeme opierať o výstižnú charakterizáciu Oxfordského výkladového slovníka angličtiny „**cognition:** *knowing, perceiving, or conceiving as an act or faculty distinct from emotion and volition; result of this*“. Zdôraznime tu predovšetkým rozlišovanie aktu, schopnosti a výsledku. V centre pozornosti kapitoly budú výsledky (kognitívnych aktov, procesov). Pod kognitívnym systémom rozumieme – veľmi voľne a liberálne - systém, ktorého správanie alebo schopnosti alebo výsledky sa podobajú (v dobre definovanej miere) kognícii, pozorovanej v živej prírode. Čitateľa však upozorňujeme na to, že slovo kognícia budeme používať trochu slobodnejšie. Potrebné terminologické poznámky umiestnime do príslušného kontextu.

² Turingov stroj je matematický konštrukt, vytvorený s cieľom presne definovať intuitívny pojem algoritmu. Tu sa uspokojíme s intuíciou algoritmu – je to predpis, určujúci, ako mechanicky vykonať nejaký výpočet.

inteligencii a kognitívnej vede. Budeme pritom vyslovovať veci, ktoré sa týkajú oboch disciplín.³

Azda treba okomentovať sformulovaný cieľ (ukázať, že umelá inteligencia a logika prispievajú k pochopeniu kognície). Tento komentár je potrebný vzhľadom na módny, až kultový trend považovať neurovedu (prípadne konekcionizmus) za dominantný, až výlučný prístup k štúdiu kognície. Podrobnejšie pozri v (PYLYSHYN 1995, PYLYSHYN 1998). Navyše, veľmi často aj tí, ktorých prístup sa vymyká neurovede alebo konekcionistickej paradigme, stotožňujú alebo takmer stotožňujú kognitívnu vedu s psychológiou. Preto optika tejto kapitoly bude veľmi okrajová v kontexte súčasnej kognitívnej vedy. Túto okrajovosť sa usilujeme zdôrazniť aj provokatívnym názvom kapitoly - kognícia bez mentálnych procesov. Budeme sa usilovať pochopiť, čo z kognície možno opísať, ak abstrahujeme od mentálnych procesov. Pokúsime sa zabudnúť na ľudský mozog a venovať sa usudzovaniu.⁴ Máme to šťastie, že o tejto ľudskej schopnosti naše súčasné poznanie mozgu takmer nič nevytvorilo. Na druhej strane, nechceme tento pohľad vnucovať. Netvrdíme, že ide o jediný možný pohľad. Naopak, pokúšame sa zistiť, aké sú jeho hranice a obmedzenia.

Ešte ďalšia výhrada prípravného charakteru. Nebudeme prijímať silné (až metafyzické) predpoklady typu „kognícia je výpočet“, „kognícia je Turingov stroj“, „v mysli (v mozgu) sú lokalizované symboly“, „podstatou kognície je spracovanie symbolov“, „v mysli je inferenčný mechanizmus“. Takéto predpoklady považujeme za príliš silné, nedostatočne zdôvodnené a nepotrebné (aspoň pre naše ciele). Za oveľa produktívnejšie a presvedčivejšie považujeme, ak sa podobné (ale slabšie) tvrdenia skonštruujú z jednoduchších východísk.

³ V tejto kapitole máme o kognitívnej vede nasledujúcu predstavu: Je to pokus o integráciu poznania, týkajúceho sa kognície, podnietený vznikom počítačov, informatiky a umelej inteligencie. Aplikuje optiku týchto disciplín na modelovanie kognície. Ambície kognitívnej vedy, kladené otázky (a neodoláme pokušeniu dodať, že aj emocionálne, energetické podhubie) sú skôr filozofické ako vedecké.

⁴ Naša centrálna téma je usudzovanie. Budeme používať aj termín odvodzovanie. Vieme, že možno hľadať významové odlišnosti medzi týmito dvoma slovami a že niekedy je potrebné ich dôsledne dodržiavať. Pri jemnosti analýzy tejto kapitoly to však nie je potrebné.

Prejdeme k Searlovej (1984) metafore čínskej izby. Máme na to dva dôvody. Po prvé, čitateľ môže získať dobrú predstavu o tom, čo je symbolový prístup. Po druhé, o kritickú analýzu tejto metafory oprieme základný postoj a ciele našej kapitoly.

Predstavme si, že ktosi (napríklad Searle) je zamknutý v izbe. Jeho materinským jazykom je angličtina a vôbec neovláda čínštinu. Má prístup k trom sadám čínskych znakov: Prvá z nich obsahuje čínske písmená. Druhá z nich podáva nejaký príbeh, tretia z nich kladie otázky, týkajúce sa tohto príbehu. Okrem toho má Searle k dispozícii množstvo pravidiel (sformulovaných v angličtine), ktoré určujú, akými kombináciami čínskych písmen (z prvej sady) má odpovedať na aké otázky (na základe daného príbehu). Jeho odpovede sú pozorovateľné zvonku.

Predstavme si aj, že takáto hra umožní odpovedať čínsky na čínske otázky tak dokonale, že vonkajší pozorovateľ nie je schopný odlíšiť práve opísané odpovedanie od odpovedí toho, kto ovláda čínštinu.

Dôsledky podľa Searla. Searle rozlišuje slabú a silnú umelú inteligenciu (SEARLE 1984, GARDNER 1987). Slabá pristupuje k výpočtovému modelovaniu ako k nástroju, využiteľnému pri štúdiu ľudského správania a kognície, k nástroju, umožňujúcemu rigorózne testovanie hypotéz. Môže tým prispieť k pochopeniu kognície. Naproti tomu, pozíciu silnej umelej inteligencie Searle charakterizuje ako presvedčenie o tom, že vhodne naprogramovaný počítač skutočne je kogníciou, že takýto počítač môže rozumieť a zažívať ďalšie kognitívne stavy. Svoju argumentáciu zameral proti takémuto poňatiu umelej inteligencie.

Searle vidí markantný rozdiel medzi porozumením čínštinu a angličtine v našom (jeho) príklade. Na rozdiel od angličtiny (ktorej Searle perfektne rozumie), odpovede v čínštinu sú iba výsledkom manipulácií neinterpretovaných formálnych symbolov. Searle tvrdí, že takéto manipulácie neumožňujú pochopiť čínske vety. Tak, ako človek z nášho príkladu, ani počítač nedokáže rozumieť, pochopiť, čo robí. Aj počítač iba manipuluje neinterpretované formálne symboly.

„Filozofia“, ktorá je za tým: Pokiaľ niekto alebo niečo postupuje mechanicky podľa nejakých pravidiel, pokiaľ vykonáva iba syntaktické operácie, nemožno povedať, že skutočne rozumie. Počítač - na rozdiel od človeka - vykonáva iba formálne, syntaktické

operácie, manipuluje so symbolmi bez väzby na nejakú sémantiku, na poznanie sveta a bez zámeru dosiahnuť isté výsledky. Podľa Searla pochopenie je výlučne schopnosťou ľudského mozgu, ktorý je schopný vyvolať (zapríčiniť) mentálne procesy.

Zhrňme:

- programy sú iba manipulácie neinterpretovaných symbolov,
- syntax nepostačuje na vytvorenie sémantiky
- kognícia kauzálne závisí od (ľudského) mozgu,
- iba kognícia vytvára mentálne obsahy (sémantiku).

Z toho:

- každý kognitívny systém (systém, schopný vytvoriť kogníciu), musí mať aspoň takú kauzálnu silu ako (ľudský) mozog,
- programy nestačia, principiálne nestačia na vytvorenie sémantiky a kognície.

Iný pohľad na Searlovu metaforu:

Searle ponúka bezstarostné a jednoduché riešenie - iba ľudský mozog produkuje kogníciu (na základe definície). Toto tvrdenie patrí do triedy nevyvrátiteľných. Ak povieme, že kognícia kauzálne závisí od ľudského mozgu (a nič viac nepovieme), každý pokus o tvrdenie, že kogníciu má niekto iný (niečo iné) ako človek, možno odmietnuť. Vysvetlenie bude jednoduché – bude nám chýbať kauzálna závislosť od ľudského mozgu. Teda, na základe definície, kogníciu má iba človek. Nemáme sa ďalej o čom baviť, všetko, čo sa na túto tému dá povedať, sme už povedali. Mohli by sme azda adaptovať (s výrazným posunom významu) Wittgensteinov záverečný aforizmus z Traktátu (1968): Ak bolo povedané všetko, treba mlčať.

V tejto chvíli by sme mali vysvetliť, prečo sa teda Searlovi venujeme. Podľa Patricka Hayesa, jedného zo zakladateľov umelej inteligencie, jadro kognitívnej vedy by sa dalo vyjadriť (zhrnúť) ako „starostlivé a detailné vysvetlenie, prečo je Searlov argument čínskej izby hlúpy, skutočne hlúpy“ (1982). Naša konštrukcia kognície bez mentálnych

procesov chce byť jedným z možných vysvetlení (a teda – neskromne – jedným z možných priblížení sa k jadrú kognitívnej vedy).

Natočme teraz pozornosť na dva problémy (domnievame sa, že - v porovnaní so Searlom - produktívnejším smerom). Možno toto natočenie niečo povie aj o Searlovom argumente, aj o Hayesovskej charakterizácii jadra kognitívnej vedy.

1. Pokúsime sa pochopiť, čo metafora čínskej izby znamená pre „slabú“ umelú inteligenciu. I keď Searle svoju argumentáciu nezameriava proti slabej (opatrnnej) umelej inteligencii, dôsledky budú zaujímavé. Budú relevantné aj pre analýzu kognície a porozumenia, prípadne produktívnych abstrakcií nad nimi.

2. Ambície „silnej“ umelej inteligencie nás nezaujímajú. Zaujíma nás však, či možno vytvoriť produktívnu abstrakciu nad kogníciou, ale zároveň aj nad výpočtovou mašinériou, prípadne nad inými fenoménmi (abstraktnými či konkrétnymi, živými či neživými). Naším cieľom bude, aby táto abstrakcia umožnila prejsť od pseudoprotblému k vážnemu problému. Ako pseudoprotblém hodnotíme napríklad otázku, či vhodne naprogramovaný počítač je kogníciou, či môže rozumieť alebo naopak, či kognícia je „turingovský“ výpočet. Ako vážny problém hodnotíme napríklad otázku, čo z kognície môžeme pochopiť, ak k nej pristupujeme **ako** k výpočtu. Budeme (ostro) rozlišovať mentálne procesy a významy. Mentálne procesy nebudeme potrebovať a významy budeme konštruovať ako entity nezávislé od mozgu a oddeliteľné od neho. Významy pokladáme za produktívnu abstrakciu nad kogníciou. Pokúsime sa zdôvodniť, že táto abstrakcia je užitočná pre pochopenie kognície (i porozumenia). Domnievame sa, že takýmto spôsobom možno skonštruovať univerzálnejší (a zaujímavejší) pojem. Bude ním *reprezentácia*.⁵ Práve reprezentácia je pre nás kogníciou bez mentálnych procesov. Za vážnu otázku považujeme, čo z kognície možno pochopiť štúdiom výpočtov nad reprezentáciami. Pochopiteľne, musí nás aj zaujímať, aké sú hranice tejto abstrakcie (čo ona znemožňuje pochopiť).

K prvému problému. Medzi výskumné ciele, ktoré možno zaradiť do slabej umelej inteligencie, patrí množstvo úloh z výpočtovej lingvistiky: bádanie v oblasti automatického prekladu, v oblasti komunikácie s počítačom v prirodzenom jazyku,

⁵ Samozrejme, neumožní postihnúť všetko, čo intuitívne spájame s mysľou.

v oblasti získavania zhrnutí z obsiahlejších dokumentov napísaných v prirodzenom jazyku a ďalšie, vrátane základného problému – *pochopiť*⁶ prirodzený jazyk, i pochopiť súvislosť používania jazyka s poznaním a so schopnosťou usudzovať.

Je však možný taký detailný opis čínštiny a dokumentov napísaných v čínštine (pomocou nejakých pravidiel), aby bol Searlov príbeh uskutočniteľný? Predpokladalo by to extrémne detailný opis:

- lexiky, morfológie, syntaxe a sémantiky čínštiny,
- spôsobov, ako sa môže meniť tvar a význam viet prirodzeného jazyka v najrozmanitejších kontextoch,
- sveta (na pochopenie viet prirodzeného jazyka sú zväčša potrebné nejaké znalosti o svete),
- odvodzovacích schopností aspoň takých silných, ako je tzv. zdravý ľudský rozum (vďaka nemu sme schopní rozumieť prirodzenému jazyku, v ktorom je otvorený priestor pre tvorivosť, metaforu, inovácie, učenie, ale aj chyby, odchýlky od normy a zdravého rozumu).

Iba na takomto základe by sa pomocou pravidiel dalo odpovedať na otázky v čínštine spôsobom, neodlíšiteľným od odpovedí človeka, ktorého materinskou rečou je čínština.

Predpokladáme, že sa dá seriózne vyvrátiť možnosť tak dôkladného opisu. Tu sa uspokojíme iba s konštatovaním, že je to nesmierne ťažká úloha, nad hranicami súčasného poznania. Odporúčame čitateľovi vyskúšať, ako odpovedá na otázky v prirodzenom jazyku (v angličtine) systém START, vytvorený na špičkovom svetovom pracovisku, MIT, možno ho nájsť na <http://www.ai.mit.edu/projects/infolab/start-system.html>

Zhrňme a zovšeobecňme: Opis komplikovanejších a komplexnejších reálnych kognitívnych schopností a inteligentných správaní, ktorý by bol taký detailný, že by

⁶ V súvislosti s umelou inteligenciou i kognitívnou vedou sa často zdôrazňuje hľadisko a cieľ konštruovať umelé výtvory (roboty, programy), ktoré „myslia“, prejavujú kognitívne schopnosti, správajú sa inteligentne. Nechceme zaznávať takéto ciele, i keď vidíme, že pootvárajú dvere, najmä v polohe „teórie“ aj prúdom, ktoré patria skôr do science fiction alebo do podradnej filozofie. (S nekvalitnou filozofiou tieto prúdy spája aj to, že ich závery sú založené predovšetkým na populárnovedeckej literatúre.) Chceme však zdôrazniť, že základnou úlohou kognitívnej vedy je predovšetkým prehľbovať poznanie, pochopenie kognície. Umelá inteligencia tiež môže a má prispieť k tomuto pochopeniu.

umožnil modelovanie pomocou počítača, je ťažká úloha, spravidla na hraniciach súčasného poznania (alebo nad nimi). Preto konštrukcia takých modelov, ktoré túto hranicu posúvajú, prispieva k *pochopeniu* kognitívnych schopností. Napríklad: úloha opísať jazyk tak dôkladne, aby program mohol zabezpečiť preklad z jedného prirodzeného jazyka do iného, vedie k presnejšiemu a podrobnejšiemu pochopeniu jazyka v porovnaní s pochopením, ktorým ľudia⁷ disponovali pred vznikom paradigmy výpočtových modelov. Túto paradigmu môžeme formulovať takto: *pochopiť znamená opísať do takého detailu, že zodpovedajúce procesy (modely) možno realizovať i počítačom*. Otázku, či počítač rozumie tomu, čo robí, nechajme tým, ktorých zaujíma. Veríme však, že jeho program stelesňuje hlboké pochopenie (ak je výpočtový model ozaj kvalitný). Je teda čas prejsť k analýze pochopenia a k druhému natočeniu našej optiky. Pokúsime sa ukázať, prečo a ako možno vybudovať koncept *reprezentácie*. Reprezentácia nám bude predstavovať *kogníciu bez mentálnych procesov*. Najprv „prečo“ – pokúsime sa motivovať túto abstrakciu (dôraz bude na abstrakcii *od*). V ďalšej časti vysvetlíme „ako“ – pokúsime sa o konštrukciu reprezentácie (dôraz bude na abstrakcii *smerom k*⁸).

6.2 Abstrakcia od mentálnych procesov

Naším programovým cieľom je abstrahovať od mentálnych procesov a sústrediť pritom pozornosť na výsledky, obsahy kognície⁹. Rozhodli sme sa ostro rozlišovať mentálne procesy a významy. Významy budeme považovať za obsahy kognície. Naše rozlišovanie mentálnych procesov a významov korešponduje Pylyshynovmu rozlíšeniu kognitívnej architektúry a reprezentácie, poznania (1995). Podľa Pylyshyna využívame dva druhy zdrojov, keď myslíme (spracúvame informácie). Jedným z nich je naša kognitívna

⁷ Iba poznamenávame, že ich pochopenie bolo menej presné a podrobné napriek tomu, že sa mohli pýšiť mozgom a mentálnymi procesmi.

⁸ Možno je potrebné podoprieť toto neobvyklé jazykové spojenie: cieľom abstrakcie (od niečoho) je totiž dopracovať sa *k čomusi*.

⁹ Terminologická poznámka: Opreli sme sa o „oxfordskú“ charakterizáciu kognície ako štruktúrovaného pojmu – rozlíšili sme schopnosti, procesy a výsledky (obsahy). Odteraz si budeme všímať predovšetkým obsahy (kognície).

architektúra – schopnosti, ktoré nám poskytujú mozog. Druhým z nich sú naše poznatky. Od prvého zdroja budeme abstrahovať.¹⁰ Túto abstrakciu budeme označovať striedavo ako abstrakciu od mentálnych procesov, neurálnych procesov, neurálnych a mentálnych procesov – tak, ako sa nám to bude hodiť v danom kontexte.

Teraz zdôvodníme, prečo je táto abstrakcia dôležitá.

Vráťme sa k metafore čínskej izby. Čo hovorí o kognícii? Špeciálne – čo o pochopení? Iba na okraj, medzi hlavné otázky, ktoré môžu viesť k pochopeniu kognície, patrí: čo vlastne je pochopenie, rozumenie?

Samozrejme, ľudská schopnosť chápať sa opiera o naše neurobiologické vybavenie, bez neho by nebola možná. Je nesporné aj to, že neurobiologické vybavenie zapríčiňuje isté mentálne procesy a mentálne procesy sú aktivované aj vtedy, keď niečo chápeme (alebo sa usilujeme pochopiť).

Podstatné však je, že (iba) neurálne a mentálne procesy nestačia na vysvetlenie pochopenia. Pochopenie je vždy pochopením **niečoho**. Keď sa chceme naučiť čínsky (a chápať čínske vety) toto „niečo“ predstavujú lexika, morfológia, syntax, sémantika, „správanie“ jazykových konštrukcií v kontexte, podľa všetkého aj relevantné znalosti o svete vzdelaného Číňana. Analyzujeme nasledujúci príklad.

Príklad. Na roentgenovej snímke každý (nehandicapovaný) človek niečo vidí. Naše vnemy sú výsledkom neurálnych a mentálnych procesov, ktorými prirodzene disponujeme. Ale – iba niektorí z nás rozumejú tomu, čo vidia. Rozumejú vtedy, keď vedia, čo *znamenajú* dôležité črty, pozorovateľné na snímke (externé vzhľadom na naše mentálne procesy: vlastnosti kostí a ich zobrazenie na snímke podľa všetkého nie sú diktované mentálnymi procesmi). Rozumejú vtedy, keď majú adekvátnu *reprezentáciu*

¹⁰ Už sme poznamenali, že naším cieľom je aj pochopiť hranice a únosnosť abstrakcie od mentálnych procesov. Tu to môžeme formulovať takto: nie je isté, či „pylyshynovské“ zdroje sú rozlíšiteľné. Niektoré naše poznatky totiž možno považovať za vrodené, „implementované“ naším mozgom. Možnosti a hranice nami navrhovanej abstrakcie treba detailne analyzovať práve (aj) z hľadiska poznania „zabudovaného“ do našej kognitívnej architektúry. I tak sa však nazdávame, že dôležitá a produktívna je abstrakcia od kognitívnej architektúry, upriamanie pozornosti na výsledky, obsahy kognície, na úlohu „externého“ prostredia pri vytváraní kognície.

potrebných oblastí.¹¹ Obsahy kognície (reprezentácia) sú vytvárané z významov. Významy sú *externé* vzhľadom na naše mentálne procesy.

Potrebu abstrahovať od mentálnych procesov budeme motivovať ešte podrobnejšie. Pokračujme ďalším príkladom.

Príklad. Predstavme si *kogníciu1* a *kogníciu2*.¹² Obe riešia dva problémy,

- $12 + 7$,
- treba nájsť loptu, ktorá odletela do kríkov.¹³

Kognícia1 nevie ani jedno (hoci vie, že $2+7=9$, $10+2=12$, $10+9=19$), ani druhé (nenájde loptu, nesystematicky sa moce).

Kognícia2 odvodí z $2+7=9$, že $12+7=19$ (pomocou $10+2=12$ a $10+9=19$) a systematicky prehľadá priestor, loptu nájde.

Analýza príkladu:

Neurálna úroveň - nie je známe, že by na tejto úrovni bol pozorovaný relevantný rozdiel medzi úspešnými a neúspešnými (pri riešení takýchto úloh) kogníciami. Čo je však podstatné: aj keby takýto rozdiel existoval, správny výsledok, ktorý dosiahla *kognícia2*, je nezávislý od neurálnych procesov v tomto zmysle: *Dôvodom*, že $12+7=19$, sú pravidlá externé vzhľadom na *kogníciu2*, a nie to, že prebieha čosi v neurónovej sieti. Podobne, loptu môže nájsť aj robot, ktorý systematicky prehľadá priestor a rozpozná loptu. Táto *metóda* zaručí, že loptu nájde.

Na druhej strane, nesprávny výsledok, ktorý dosiahla *kognícia1*, je výlučne závislý od

¹¹ Pripomeňme a modifikujme si Baconovo podobenstvo o mravcoch, pavúkoch a včelách: Pavúk nech je symbolom „tkania“ obsahu kognície „zvnútra“, z mentálnych procesov. Mravec nech je symbolom postupného vytvárania obsahov kognície zvláčaním komponentov týchto obrazov z vonkajšieho prostredia (napríklad, skladaním roentgenovej snímky z jej jednotlivých častí). Napokon, včela nech je symbolom vytvárania obsahov kognície tak, že materiál, získaný z prírodného prostredia pomocou nášho mentálneho vybavenia, pretvorí na novú entitu, reprezentáciu.

Zámerné provokujeme: v reprezentácii je viac ako vo vonkajšom prostredí (v roentgenovej snímke), ale aj viac ako vo vnútornom prostredí (v mozgu a v neurónovej sieti) – sú tam významy. Dodajme však, že naše ambície nie sú metafyzické: nezáleží nám na tom, kde „existujú“, kde sú lokalizované významy. Chceme iba skonštruovať teóriu významov a reprezentácie, ktorá nepotrebuje odvolávať sa na mozog a neurónovú sieť, na symboly v mozgu a chceme ukázať, že takáto teória má čo povedať.

¹² Ďalšia terminologická poznámka: Niekedy budeme personifikovať, slovo kognícia budeme používať aj na hovorenie o organizme (systéme) disponujúcom kogníciou. Iste to zjednoduší vyjadrovanie (i čítanie).

¹³ Tieto dva – dosť odlišné – príklady sú zvolené zámerné. Prvý z nich predstavuje hľadanie tzv. matematických právd a schopnosť usudzovať. Druhý z nich predstavuje empirické „zisťovanie“, schopnosť poznávať vonkajšie prostredie.

mentálnych procesov a nemá žiadny objektívny (externý) dôvod.

Psychologická úroveň (analýzy príkladu). To, že *kognícia*² korektne usudzuje, sa dá posúdiť iba s ohľadom na nejaké (externé) dôvody a metódy. Mentálne procesy, stavy jej vedomia, ani jej správanie nie sú podstatné.¹⁴

Naša hlavná otázka je: čo z kognície možno opísať, ak abstrahujeme od mentálnych procesov? Vôbec nechceme tvrdiť, že mozog možno transplantovať a nahradiť ho počítačom. Chceme iba pochopiť, ako a do akej miery načrtnutá abstrakcia dokáže vysvetliť a charakterizovať kogníciu, najmä usudzovanie (tým aj pochopiť niečo z možností, resp. obmedzení umelej inteligencie).

Pripájame dôležité upozornenie: Zamýšľame abstrahovať od mentálnych procesov, zároveň však chceme, aby naša teória bola relevantná pre reálne (živé) kognície. Preto sa budú v celom texte vyskytovať príklady a motivácie z ľudského a živočíšneho prostredia. Nepovažujeme to za konflikt. Ide o to, aby naša abstrakcia mala dôležité vlastnosti.

Záver: Významy, dôvody a metódy sú externé vzhľadom na mentálne procesy (kognitívnu architektúru). Môžeme abstrahovať od mentálnych procesov, ak chceme analyzovať, pochopiť, ako významy, dôvody a metódy tvoria obsah kognícií. V tejto časti sme sa sústredili predovšetkým na to, aby sme ukázali, prečo možno a treba abstrahovať od mentálnych procesov. Používali sme veľmi predbežné pojmy ako reprezentácia, význam, dôvod, metóda. Systematickejšia predstava o obsahu kognícií bude nasledovať. V nasledujúcej časti si doplníme dôležité východiskové predstavy:

- zdôrazníme nedeterminizmus kognície,
- zasadíme kogníciu do nejakého prostredia a dohodneme sa na základných črtách jej vzťahov s prostredím,
- napokon, vytvoríme si predbežnú predstavu o tom, čo sú významy.

¹⁴ Ani teória, podľa ktorej sú mentálne procesy výpočtami, nemení situáciu, pokiaľ neumožňuje odlišiť (nesprávne) výpočty kognície¹ od (správnych) výpočtov kognície². Ak bude schopná takéhoto rozlišovania, bude musieť (sémanticky) špecifikovať korektné výpočty. Dostali sme sa tým (opäť) ku kľúčovej úlohe sémantiky.

Potom, v ďalšej časti, podáme náčrt kognitívnej teórie reprezentácie. Reprezentácia bude stelesňovať kogníciu bez mentálnych procesov. Prívlastok „kognitívna“ má zdôrazniť, že táto teória reprezentácie aspiruje na to, aby bola relevantná pre biologický svet, pre kognície, ktoré sa reálne vyskytujú v živočíšnej ríši. Dôvod, prečo treba tento prívlastok používať: existujú rôzne teórie reprezentácie. V umelej inteligencii možno hovoriť minimálne o dvoch prístupoch k teórii reprezentácie. Jeden z nich skúma nejaké logické jazyky, rôzne druhy odvoditeľnosti z množín viet takýchto jazykov (reprezentujúcich nejaké poznatky) a predovšetkým sa zaujíma o to, čo je vyjadriteľné a odvoditeľné v danom jazyku. Druhý sa zaoberá údajovými štruktúrami, ktoré sa používajú na vyjadrenie, kódovanie poznatkov v reálnych programových systémoch a manipuláciami (výpočtami) na týchto štruktúrach¹⁵. Ani jeden z týchto prístupov tu nebude predmetom nášho hlavného záujmu.

6.3 Východiská konštrukcie teórie reprezentácie

Skôr než prejdeme k teórii reprezentácie, pripravíme si ďalšie dôležité piliere nášho prístupu (prvým je, zdôraznime, abstrakcia od mentálnych procesov). Ďalšími sú: *nedeterminizmus* kognície, interakcia kognície s *prostredím* a, napokon, naša abstrakcia smeruje (od mentálnych procesov) k *významom*.

6.3.1 Nedeterminizmus

Vyštartujme z deterministickej predstavy kognície, t.j. z predstavy kognície ako funkcie. Zdá sa, že nie sú závažné prekážky, aby sme na slnečnicu nazerali ako na funkciu.¹⁶ Aj k

¹⁵ Prvému prístupu je venované (ŠEFRÁNEK 2002), s druhým sa možno stretnúť v (POPPER, KELEMEN 1989).

¹⁶ Uvedomujeme si, že táto veta môže iritovať botanika. Celkom určite by do hry mali vstúpiť štatistické závislosti. Podstatné – pre nás – je, že u slnečnice neočakávame reprezentáciu a ani potrebu tento konštrukt zavádzať.

človeku možno pristúpiť podobne. Podľa laplaceovského pohľadu na svet, ak by sme poznali všetky (nevypočítateľné) okolnosti, ktoré na človeka pôsobia, dokázali by sme určiť „odpoveď“ človeka na tieto okolnosti (jeho správanie). Takýto pohľad nemožno ani potvrdiť, ani vyvrátiť, ide o „pozaempirickú“ konštrukciu. Bez ambícií na metafyzicky záväzné vyjadrenia prijmemé nedeterministickú pozíciu.

Samozrejme, máme na to nejaké dôvody. Po prvé, naše rozhodnutie je motivované emocionálne. Človek ako funkcia, ale aj mačka ako funkcia sú predstavy, ktoré radšej odmietneme.¹⁷ Menej emocionálny pohľad môže využiť analógiu z vývoja psychológie. Behaviorizmus v princípe pristupoval k psychike ako k funkcii. Skúmanie správania ako reakcie na podnety sa stalo príliš obmedzujúce (a príliš slabé z hľadiska schopnosti vysvetliť správanie organizmov). Miller, Galanter a Pribram zaviedli do hry dva dôležité konštrukty – *obraz* a *plán* (1960). Reakcie na podnety sú sprostredkované, modifikované obrazom.¹⁸ Obraz spolu s podnetom však nepostačuje na vysvetlenie reakcií. Organizmus, človek (kognícia) nie je iba pozorovateľom, ktorý si vytvára obraz, je aj (a predovšetkým) aktérom, ktorý v prostredí uspokojuje svoje potreby, zámery. Preto správanie nemožno do detailov opísať len ako reakcie na podnety (sprostredkované obrazom). Na vysvetlenie reakcií konajúceho aktéra je užitočný aj pojem plánu.

Považujeme teda za primerané predpokladať, že reakcie na podnety sú modifikované obrazom a plánom. Obraz a plán však dávajú organizmu isté možnosti voľby. Obraz (presvedčenia, poznanie) často obsahuje *alternatívy*. Plán je založený na *výbere* cieľov. Preto je vhodné prejsť od predstavy reakcií determinovaných podnetmi k predstave nedeterministického výberu reakcie na daný podnet (predpokladáme pritom, že podnet aktivuje nejaké alternatívne obrazy a plány).

Analogicky (k myšlienke obrazu a plánu) postúpime od predstavy kognície ako funkcie ku kognícii vybavenej reprezentáciou. Úlohu obrazu a plánu preberie reprezentácia.

¹⁷ Toto rozhodnutie má aj svoj estetický (človek ako funkcia je nevkusný konštrukt) a aj etický (funkcia nezodpovedá za nič) obsah. Za upozornenie na etický aspekt takéhoto rozhodnutia ďakujeme Ladislavovi Kováčovi. Samozrejme, týmto rozhodnutím nikomu nebránime nazerať na mačku alebo človeka ako na funkciu.

¹⁸ Nebudeme podrobnejšie vysvetľovať, čo sa v (MILLER 1960) myslí pod obrazom, iba zdôrazníme, že nejde o obrazné predstavy, ale o presvedčenia (a – prípadne – poznanie), ktoré si kognícia vytvára o prostredí. Neskôr sa od obrazu dostaneme k reprezentácii.

Predpoklad o existencii reprezentácie je v istom zmysle spriahnutý s nedeterministickým chápaním kognície. Vychádzame z toho, že reprezentácia spravidla obsahuje alternatívy. A naopak, ak by sme kogníciu pohodlne mohli chápať ako funkciu (ako deterministické zariadenie), nepotrebovali by sme konštruovať reprezentácie.

Nebudeme tvrdiť, že konanie a rozmyšľanie živočíchov („prirodzených kognícií“) nie je ničím ovplyvnené. Prijmeme však, že existujú situácie, v ktorých kognície majú na výber rôzne možnosti a z týchto možností vyberajú. Vyberajú jednak ciele, ktoré usmerňujú ich činnosť: Zdá sa, že už mačka alebo kojenec¹⁹ si v niektorých situáciách vyberajú ciele z viacerých možností. To vôbec neznamená, že v iných situáciách ich cieľ nemôže byť jednoznačne determinovaný (toto platí aj pre dospelého človeka). Vyberajú si aj riešenia. Ak je daný cieľ alebo problém, možno vo všeobecnosti riešenie vybrať z viacerých možných riešení.

Poznamenávame, že je vhodné predpokladať výber cieľa (úlohy, problému) z *otvoreného* priestoru možností. Ak má kognícia vybraný problém (úlohu, cieľ), môže jeho riešenia (realizácie cieľa) vybrať opäť z nejakého *otvoreného priestoru možných riešení*. Uzavreté a fixované priestory sa nám nezdajú byť vhodné na modelovanie kognície, ktorej dôležitou črtou je adaptabilita na nové podmienky a tvorivosť, kladenie nových otázok a nachádzanie nových riešení.

Tu používaný pojem nedeterminizmu (nedeterministického výberu) zodpovedá nedeterministickému výpočtovému modelu. Nedeterministický algoritmus pracuje vo dvoch fázach. V prvej „*odhadne*“ riešenie problému (vyberie jednu z možností z akéhosi priestoru potenciálnych riešení), v druhej *preverí*, či tento výber je skutočne riešením. Hovoríme, že nedeterministický algoritmus *rieši* nejaký problém vtedy, keď *existuje* taký výber, o ktorom sa vo fáze preverovania ukáže, že je riešením.

Základná je tu predstava metódy pokusov a omylov: výber možného riešenia a následné testovanie. Samozrejme, uskutočnený výber môže byť niečím determinovaný, môže závisieť od individuálnych skúseností. Dôležitejšie je však toto: V reprezentácii („obraz a plán“) vygenerujeme (vďaka mechanizmu, od ktorého tu abstrahujeme) možné riešenie. Potom ho preverujeme (v interakcii reprezentácie a prostredia). Opakujeme to dovedy,

¹⁹ Nehovoriac o dojčati alebo lezúňovi.

kým riešenie nenájde alebo nerezignuje. Z výpočtového hľadiska: operácie nad reprezentáciami sú nedeterministické výpočty.

Veľmi dôležité budú v našej koncepcii usudzovacie operácie. Aj tie možno (a treba) charakterizovať bez odvolávania sa na kognitívnu architektúru nejakého organizmu. Adekvátne je chápať ich ako (nedeterministické) operácie na reprezentáciách.

6.3.2 Kognícia a prostredie

Pred vlastnou konštrukciou predstavy o reprezentácii si potrebujeme vyjasniť, v akom prostredí reprezentáciu (t.j. kogníciu bez mentálnych procesov) chceme mať.

Predpokladáme nejaké spoločenstvo **organizmov**. Každý organizmus vykonáva nejaké **akcie** (správanie) v nejakom **prostredí**. Organizmy disponujú **reprezentáciou**, majú nejaké **potreby** a **ciele**, ktoré sa usilujú uspokojovať akciami. Reprezentácie sú vytvorené z **významov** (zatiaľ nevieme, čo významy sú). Organizmus je schopný usudzovať: obohacovať svoju reprezentáciu o nové významy aj bez interakcie s prostredím.

Pre každý organizmus sú všetky ostatné organizmy súčasťou prostredia. Organizmy komunikujú. Spracovanie významov v procese komunikácie je pre nás kľúčovou témou. Organizmus, ktorý vysiela správy, tvorí a vysiela vlastne nejaké významy. Podobne, ak organizmus prijíma správu, rozpoznáva významy. Veľkú pozornosť budeme venovať tomu, ako spolu súvisia rôzne typy jazykov, v ktorých sa komunikuje, a im zodpovedajúce typy reprezentácií.

Organizmus, prostredie a reprezentáciu spájajú nejaké relácie (nie funkcie):

- organizmus prijíma z prostredia nejaké vstupy (medzi nimi môžu byť aj jazykové výrazy) a priradzuje im (viaceré možné) významy v reprezentácii, túto reláciu budeme značiť ako **I** (interpretácia),²⁰

²⁰ **I** predstavuje jednak rozpoznanie (známych) významov, ale aj tvorenie (nových) významov. Operácia tvorenia nových významov býva spravidla náročná na čas a úsilie.

- organizmus je aktívny pri výbere vstupov: zameriava pozornosť, selektuje tie vstupy, ktoré sú preňho zaujímavé, relevantné, dôležité (v danej situácii), reláciu selekcie budeme značiť ako **S**,
- relácia **S** je závislá od kontextu, organizmus má nejaké potreby, ciele, zámery, problémy, tie ovplyvňujú (aktuálnu) podobu relácie **S**,
- organizmus v prostredí koná – relácia **A** priradzuje niektorým významom (predovšetkým významom, zodpovedajúcim cieľom, zámerom, potrebám, problémom) nejaké akcie; medzi akcie rátame aj jazykové výpovede, aj usudzovanie; zdôrazníme: v reprezentácii nosíme aj problémy.

Poznamenajme, že selekcia nemusí byť (a nebýva) založená na interpretácii. Prostredie nám ponúka nevyčerpatelné množstvo vstupov. Značnú časť z nich vieme (ako-tak :-)) interpretovať. Nie všetky z nich však v ľubovoľnej situácii selektujeme ako dôležité. Mechanizmy selekcie relevantných vstupov bývajú zabudované v organizmoch na veľmi elementárnej úrovni.

Okrem relácií, spájajúcich organizmus a jeho reprezentáciu s prostredím (môžeme ich nazvať oknami), sú tu aj nejaké bariéry, ktoré komplikujú orientáciu organizmu v prostredí:

- (i) Pre funkciu **I** sú rozlíšiteľné iba niektoré vstupy (obrazne: nie všetko z prostredia je vnímateľné, poznateľné).
- (ii) Predpokladáme, že organizmus koná a pohybuje sa v čase a priestore. Každá jeho situácia, určená trojicou [akcia, čas, miesto] obmedzuje množinu vstupov, ktoré **I** môže registrovať.
- (iii) Relácia **I** reaguje iba na tie vstupy, ktoré sú vyselektované reláciou **S** - niektoré z principiálne rozlíšiteľných vstupov organizmus nerozlišuje, lebo nezameriava na ne pozornosť. Teda, aj keď je organizmus schopný isté vstupy rozlišovať, nerozlišuje ich, lebo na ne nezameriava pozornosť - rozlišuje iba to, čo považuje za hodnotné, relevantné, dôležité, zaujímavé. Môže pritom „vyberať“ aj zo vstupov, ktoré sú v istom okolí

vyselektovaných vstupov. (Intuitívne ide o náhodné vstupy, na ktoré bola pozornosť zameraná mimovoľne alebo kvôli tomu, že sa vyskytovali „blízko“ relevantných vstupov.)

(iv) Interpretácia vstupov reláciou **I** produkuje diskkrétne významy, ktorými reprezentuje (často alebo zväčša) spojité vstupy (diskrétna reprezentácia spojitých vstupov).²¹

Naša abstrakcia *od* mentálnych procesov *smereje k* významom. Významy vytvárajú reprezentáciu. Je najvyšší čas venovať sa im podrobnejšie.

6.3.3 Významy

Predovšetkým, nepredpokladáme, že významy sú priradené výlučne jazykovým výrazom. Naopak, podstatnou črtou našej teórie je presvedčenie, že významy a na nich založenú reprezentáciu treba predpokladať aj u organizmov, ktoré nedisponujú jazykom. Navyše, domnievame sa, že pri používaní jazyka sa dajú pozorovať príklady nevyjadriteľnosti a nezrozumiteľnosti a ich výskyt možno vysvetliť tým, že aj u človeka komunikujúceho v jazyku je doména jazyka a doména významov oddeliteľná.

Napriek deklarácii z predchádzajúceho odseku začneme – kvôli motivácii – významami vo vzťahu k jazyku.²² Keď niečo vypovedáme v jazyku, spravidla *označujeme* nejaké objekty (materiálne alebo ideálne). Významy však nebudeme stotožňovať s týmito *označovanými* objektami. Po Fregem (1892) nemožno (bez prijateľného alebo neprijateľného zjednodušenia) sémantiku budovať iba na objektoch, *označených* jazykovými výrazmi. Nasledujú závažné dôvody, ktoré to podporujú.

- Ten istý objekt²³ možno označiť jazykovými výrazmi, ktorých význam je rôzny. Preto je dôležité identifikovať práve tieto odlišnosti, určujúce rôznosť významov.

²¹ Tejto dôležitej otázke sa ďalej nevenujeme.

²² Čitateľ môže nájsť podrobnejší a dôkladnejší výklad v (CMOREJ 2001).

²³ Nemáme na mysli iba individuálne, konkrétne objekty, ale aj vlastnosti, vzťahy, udalosti a rôzne zložito štruktúrované objekty, konkrétne i abstraktné.

- Objekt, označený jazykovým výrazom, nemusíme poznať. Napriek tomu výrazu môžeme rozumieť (poznať jeho význam).
- Ďalej, existujú jazykové výrazy, ktoré neoznačujú nijaký objekt, ale majú význam, rozumieme im.

Význam (jazykového výrazu) môžeme – veľmi voľne, ale výstižne - charakterizovať ako **kritérium**, ktoré dovoľuje rozhodnúť o tom (rozlíšiť), či nejaký objekt (ľubovoľne štruktúrovaný alebo abstraktný) je označený týmto jazykovým výrazom alebo o tom, že výraz bol adekvátne použitý v nejakej situácii. Samozrejme, nepredpokladáme, že toto kritérium je „ostré“, vieme, že pre našu hru potrebujeme „rozmazané, neostré“ kritériá. Nejakému jazykovému výrazu rozumieme vtedy, keď disponujeme takýmto „rozlišovacím“ kritériom.

Príklad. Výrazu „do čerta“ rozumieme, keď máme kritérium, dovoľujúce posúdiť, v akej situácii ho možno adekvátne použiť. Asi by sme mali veľké problémy, keby sme chceli určiť objekt označený týmto výrazom.

Ďalej budeme predpokladať, že kritériá môžu byť iba predbežné a že sa môžu postupne vyvíjať. Pripomíname, že naše konštrukcie sú o svete, v ktorom prevláda nedokonalé chápanie (porozumenie). Významy sú schematické v tom zmysle, že ich rozlišovanie nie je detailné, ale je orientované iba na niektoré črty, ktoré možno považovať pre daný význam za relevantné. Pre rozlišovacie kritérium kopca nie je podstatné mravenisko (ktoré môže, ale nemusí byť na kopci). Schematickosť vedie k predbežnosti, nedefinitívnosti, približnosti. Ďalej poznamenávame, že sémantika, budovaná nad kritériami, má holistické črty: kritérium kopca nekonštruuje kopec z pixelov.

Teraz rozšírime chápanie významu tak, aby sa s jeho pomocou dalo charakterizovať aj predjazykové správanie. Základnú myšlienku však zachováme. Významy budeme charakterizovať pomocou schopnosti rozlišovať.²⁴

Schopnosť *rozlišovať* je schopnosťou základnej dôležitosti pre živé organizmy. Natočenie slnečnice smerom k zdroju denného svetla je jedným z mnohých príkladov tejto schopnosti. Významy však nebudeme chcieť spojiť s ľubovoľnou schopnosťou rozlišovať.

²⁴ Menej antropomorfisticky môžeme hovoriť o rozlišovacom mechanizme alebo o rozlišovacom kritériu. Ďalej uprednostníme kritérium (mechanizmus nám pripomína mentálne procesy). V motivačnej časti však budeme hovoriť o schopnosti rozlišovať.

Tá je totiž vlastná každému reaktívnemu správaniu. O významoch budeme hovoriť až vtedy, keď schopnosť rozlišovať má aspoň tieto črty:

A. Dokáže v danej situácii rozlišovať aj objekty, ktoré nie sú dostupné priamemu vnímaniu. Pripomíname, že slovo „objekty“ chápeme stále vo veľmi všeobecnom zmysle.

Príklad. Starý známy šimpanz, ktorý sa usiluje dosiahnuť banány a dokáže rozlíšiť, že sa mu k tomu zide palica, ktorá nie je práve v dosahu jeho zmyslov, je príkladom takejto schopnosti rozlišovať. Fabulujme: dokáže rozlíšiť vlastnosti, ktoré by asi mal mať vhodný nástroj.

B. Dôležitou vlastnosťou rozlišovania je jeho potenciálny nedeterminizmus. (Táto vlastnosť sa môže uplatniť, ale nemusí sa uplatniť v každej situácii.)

Príklad. Aj tu máme starého známeho, je ním potkan hľadajúci cestu z bludiska. Predpokladáme, že je viac možných ciest z bludiska. Jednu z nich potkan nájde (rozlíši) po sérii pokusov a omylov.

Uvedené príklady ukazujú, že okrem vlastností A a B treba predpokladať, že organizmus sa dokáže učiť. Uzatvárame, že významy má zmysel predpokladať až vtedy, keď

C. organizmus na základe pozorovania prostredia a experimentovania v prostredí dokáže konštruovať nové významy (nové rozlišovacie kritériá),

D. organizmus dokáže odvodzovať (z daných rozlišovacích kritérií dokáže odvodiť nové odvodzovacie kritériá).

Príklad. Zdá sa, že šimpanz skôr odvodzoval a potkan skôr konštruoval nový význam.

Skôr než budeme význam explicitne charakterizovať pomocou schopnosti rozlišovať (rozlišovacieho kritéria), všimneme si dva z prístupov k sémantike. Jedným z nich je tradičná logická intenzionálna sémantika. Aj táto sémantika zodpovedá našej intuícii kritéria. Význam nejakého jazykového výrazu je – veľmi približne - funkcia, ktorá tomuto výrazu priradzuje označený objekt. Približne²⁵ preto, lebo detailná analýza, potrebná pre dôkladnú sémantiku ukazuje, že je potrebné zaviesť do hry možné stavy vecí (možné svety) a určitým spôsobom štruktúrovať svet objektov (zaviesť nejaké typy objektov). Prípadne sa zavádzajú aj ďalšie konštrukty. Takýchto funkcií však možno

²⁵ Opäť odkazujeme na (CMOREJ 2001).

konštruovať ľubovoľne veľa, pričom ich obrovská väčšina nezodpovedá nijakým „prirodzeným“ významom. Okrem toho, táto sémantika sa nezaujíma o ďalšie vlastnosti, ktoré sú dôležité z hľadiska štúdia kognície. Napríklad nehovorí o prototypoch, o podobnosti (predstavme si pojem červenosti, farba niektorých objektov je prototypovo červená, odtiene červenej sa viac alebo menej podobajú). Ďalej nedokáže adekvátne konštruovať zložené významy z elementárnejších (pozri príklad bieleho vína z časti 6.5.2).

Druhým zo spomínaných prístupov je kognitívna sémantika Petra Gärdeforsa (2000). Jej cieľom je vytvoriť rámec, ktorý je zaujímavý z hľadiska štúdia kognície a nemá nedostatky, o ktorých sme sa zmienili vyššie. Gärdenfors zavádza konštrukt konceptuálneho priestoru. Tento priestor je určený množstvom dimenzií. Niektoré z nich sú vrodené, predovšetkým tie, ktoré sú späté s našim zmyslovým vnímaním. Príkladmi takýchto dimenzií sú teplota, váha, výška, šírka, hĺbka, jas, odtieň, sýtosť (farby), sladkosť, horkosť, kyslosť atď. Ďalšie dimenzie sú naučené v priebehu „prirodzeného“ života organizmov, ďalšie sú vytvorené kultúrnym prostredím alebo vedou. Niektoré z dimenzií sú neseparovateľné: napríklad hodnoty jasú, sýtosti a odtieňa farby môžeme určiť iba súčasne. Neseparovateľné dimenzie definujú domény (farba, zvuk, tvar, chuť atď.).

Dimenzie vytvárajú nejaký mnohorozmerný priestor. Každému individuu zodpovedá nejaký bod v tomto priestore. Bod je určený n -ticou hodnôt, získaných pre toto individuum v každej z dimenzií n -rozmerného priestoru. Predpokladajme, že takýto priestor sa dá skonštruovať, alebo – slabšie – že pre nejakú konkrétnu aplikáciu sme vytvorili model tohto typu. Predpokladajme ďalej, že na tomto priestore sa dá definovať vzdialenosť medzi bodmi. Na základe vzdialenosti si definujeme podobnosť (čím sú individua vzdialenejšie, tým sú menej podobné, v (GÄRDENFORS 2000) je táto závislosť exponenciálna). Prirodzeným pojmom zodpovedajú konvexné oblasti v konceptuálnom priestore: ak vezmeme ľubovoľné dve inštancie prirodzeného pojmu, potom každý bod, ktorý je medzi nimi, je tiež inštanciou daného pojmu. (Ak máme dva body, zodpovedajúce oranžovým objektom, potom každý bod medzi nimi zodpovedá tiež nejakému oranžovému objektu.)

Naše hlavné výhrady voči Gärdenforsovej sémantike:

- nedáva kritérium pre identifikovanie prirodzených pojmov: nie každý konvexný podpriestor konceptuálneho priestoru možno označiť za prirodzený pojem,
- samotná konštrukcia konceptuálneho priestoru sa zdá byť nerealizovateľným projektom (už iba vďaka rozsahu potrebnej práce),
- navyše, mnohé dôležité dimenzie – zdá sa – majú kvalitatívne hodnoty, bez možnosti definovať na nich vzdialenosť (metriku), u ďalších možno hodnoty nanajvyššie usporiadať, niekedy iba čiastočne; to znamená, že akákoľvek konštrukcia metriky a podobnosti v takto heterogénnom súradnom systéme bude spôsobovať vážne až neriešiteľné komplikácie,
- nedáva vhodné prostriedky na priradenie významu slovesám, a tým aj vetám.

Teraz načrtneme našu predstavu (dôkladný a podrobný výklad takejto sémantiky presahuje možnosti tejto kapitoly).

6.4 Kognitívna sémantika

Významy založíme na *rozlišovacom kritériu*, na abstrakcii schopnosti rozlišovať. Pre naše chápanie rozlišovacieho kritéria sú dôležité vlastnosti A – D z predchádzajúcej časti.

Poznámka. Nebudeme zafažovať formálnymi detailami pri našej konštrukcii kognitívnej sémantiky (výklad bude veľmi neformálny). Predpokladáme však, že niektorým čitateľom to nebude vyhovovať, preto aspoň načrtneme, ako možno rozlišovacie kritérium zaviesť formálnejším spôsobom. Rozlišovacie kritérium možno chápať ako čiastočnú funkciu, ktorá objektu priraduje hodnotu z nejakej čiastočne usporiadanej množiny. **Príklad.** Rozlišovacie kritérium potraviny môžeme modelovať ako čiastočnú funkciu, ktorá nejakým objektom nepriradí nič, iným objektom priradí hodnotu z množiny *určite jedlé, asi jedlé, asi nejedlé, určite nejedlé, neviem–vyskúšam*. Množina je čiastočne usporiadaná, hodnota *neviem–vyskúšam* je neporovnateľná s každou z ostatných.

Pripomíname, že máme na mysli veľmi rôznorodé a rozmanito štruktúrované objekty. Keďže funkcia je čiastočná, pre niektoré objekty nebude definovaná (myšlienky nemôžu byť zelené, ľavy nemôžu byť diametrálne atď.)

Chceme sa však zaoberať bez pojmu funkcie (hoci niekedy v poznámke naznačíme, ako sa dá pomocou pojmu funkcie realizovať práve diskutovaná konštrukcia).

Poznamenávame, že rôzne organizmy môžu „používať“ rôzne rozlišovacie kritériá toho istého objektu.

Príklad. Rozlišovacie kritériá potravy dvoch rôznych organizmov (kognícií :-) môžu tomu istému objektu priradiť rôzne hodnoty.

Ďalej postupujeme tak, že „rozlišovacie kritérium **niečoho**“ budeme postupne, po prípadoch vysvetľovať. Vždy v závislosti od toho, čo toto „niečo“ je. Najprv zavedieme nejaké elementárne (atomárne) rozlišovacie kritériá. Potom načrtneme spôsoby konštrukcie komplikovanejších rozlišovacích kritérií.

Elementárne rozlišovacie kritériá rozlišujú *predmety* (individuá), *prirodzené triedy* predmetov, *prirodzené vlastnosti* predmetov, *prirodzené vzťahy* medzi predmetmi.²⁶

Príklad. Predstavme si niektoré rozlišovacie kritériá - predmetov (Liptovský Hrádok, Jožo z 2.C ...), tried (kamene, medvede, ...), vlastností (oranžový, ťažký, jedlý, ...), vzťahov (pod, vpravo, vyšší, ...).

Veľmi dôležitá je konštrukcia prirodzenej triedy. To, že ju máme medzi elementárnymi pojmami, znamená, že významy predpokladáme iba u takých organizmov, ktoré majú schopnosť zovšeobecňovať. Popri rozlíšení nejakého konkrétneho kameňa pri nohe dokážu rozlíšiť, čo je kameň (vo všeobecnosti). Samozrejme, chyby pri rozlišovaní nevyklúčujeme.

Pod *prirodzeným pojmom* rozumieme elementárne rozlišovacie kritériá.

Podobne ako Gärdenfors, nie sme schopní podať kritérium prirodzeného pojmu. Domnievame sa, že relácia selekcie **S** by mohla byť užitočným nástrojom pri špecifikácii prirodzených pojmov. Nebudeme sa však do toho púšťať. Preto iba predpokladáme, že možno vytvoriť zoznam elementárnych rozlišovacích kritérií. Tento zoznam je

²⁶ Vieme si predstaviť kritérium *zmeny* (bude o ňom reč neskôr) ako elementárne. Všetky ostatné kritériá by sa dali (možno) konštruovať nad ním. Zvolili sme však pohodlnejšiu a vychodenejšiu cestu.

modifikovateľný. Organizmus si časom môže vytvárať nové a ďalšie rozlišovacie kritériá ďalších a ďalších predmetov, prirodzených tried, prirodzených vlastností a prirodzených vzťahov. Navyše, existujúce rozlišovacie kritériá môže časom modifikovať.

Príklad. Príkladom nie-prirodzeného pojmu môže byť „zelený do 31.12.2016 a modrý od 1.1.3000“. Netvrdíme, že by sme mali ťažkosti s určením rozlišovacieho kritéria takéhoto pojmu, ale rozhodne nemienime takéto kritérium považovať za elementárne – nezaradili by sme ho do spomínaného zoznamu.

Popri elementárnych rozlišovacích kritériách predpokladáme nejaké pravidlá, umožňujúce konštruovať zložitejšie, neelementárne, odvodené rozlišovacie kritériá z jednoduchších, už skonštruovaných.

Z rozlišovacích kritérií predmetov, tried, vlastností a vzťahov budeme konštruovať rozlišovacie kritériá *situácií*. Najprv uvedieme príklady situácií.²⁷

Príklad. Nejaký predmet má nejakú vlastnosť (toto jablko je jedlé, tamto je nezrelé), Nejaké predmety (prípadne s nejakými špecifikovanými vlastnosťami) sú v nejakom vzťahu (toto oranžové jablko je na tomto dubovom poškodenom stole).

Každá situácia je jedinečná (podobne ako predmet), preto jej rozlišovacie kritérium funguje ako (neostrý) identifikátor. Poznávanie si predstavujeme (aj) ako konštrukciu alternatívnych rozlišovacích kritérií situácií a výber toho kritéria, ktorého „výstup“ je v danej situácii „najlepší“ (ako výstupy rozlišovacích kritérií situácií si môžeme predstaviť čiastočne usporiadané množiny nejakých pravdivostných hodnôt – radšej by sme sa však v tejto chvíli vyhli až takejto podobnosti na logiku).

Poznámka. V reči funkcií môžeme situácie definovať ako aplikácie funkcií na hodnoty iných funkcií (so zachovaním takej typovej disciplíny, ktorá by dovoľovala udržať rozumný zmysel týchto konštrukcií).

Môžeme ďalej rozlišovať vlastnosti vlastností, vzťahy medzi vlastnosťami, vlastnosti situácií. Tejto téme sa však radšej vyhneme. Jednak by nás mohla viesť dosť ďaleko od rámca prirodzených pojmov, okrem toho, z technického hľadiska sú tieto konštrukcie dobre zvládnuté.

²⁷ Zavádzame jednoduchý, technický pojem situácie (nenašli sme iné, vhodnejšie slovo). Žiaľ, slovo „situácia“ používané aj vo voľnejšom, netechnickom, intuitívnom zmysle. Dúfame, že to nepovedie k vážnym nedorozumeniam.

Príklad. Nebezpečie možno chápať ako vlastnosť situácie.

Schopnosť zovšeobecňovať umožňuje rozlišovať dôležitejšie **typy** situácií.

Príklad. Z opakovaného rozlíšenia toho, že toto jablko je červené a je aj chutné, tamto je ešte zelené a je nezrelé, môžeme zovšeobecňovať (možno prirýchlo a nie celkom správne). Získame tak pravidlá: ak je jablko červené, tak je chutné (a mnoho ďalších).

Budeme preto hovoriť o *rozlišovacích kritériách typov situácií* a nazveme ich *pravidlami*.

Pravidlo dovoľuje rozpoznať nejakú zložitejšiu situáciu na základe nejakých symptómov tej situácie.

Príklad. Niektoré pravidlá sa neosvedčia. Pravidlo *ak je jablko zelené, tak je nezrelé* sa časom ukáže ako nespoľahlivý identifikátor situácií na základe niektorých symptómov. Na tomto mieste opäť zdôrazníme, že diskvalifikácia tohto pravidla je opretá o externé dôvody, nie o mentálne procesy.

Na rozlišovacích kritériách možno zaviesť ďalšie operácie. Najjednoduchšie z nich sú booleovské. Vo všeobecnosti môžeme povedať, že z jednoduchších rozlišovacích operácií vytvárame komplexnejšie tak, že z jednoduchších vytvárame nejaký štruktúrovaný celok (z rozlišovacieho kritéria kopca rozlišovacie kritérium strmého a zalesneného kopca). Budeme hovoriť, že významy (rozlišovacie kritériá) majú vlastnosť kompozicionality – možno ich skladať (komponovať) do celkov z jednoduchších častí a naopak, celky možno rozkladať (dekomponovať) na časti.

V nasledujúcich častiach budeme postupne prechádzať od jednoduchších reprezentácií ku komplikovanejším a tieto prechody budú umožnené práve zavádzaním nových konštrukcií s vyššou silou.

V kognitívnej sémantike je dôležitým pojmom *prototyp*. S niektorými pojmami (oranžový, stolička, pes ...) možno prirodzene spájať nejaký prototyp. S inými, podľa všetkého, nie (nábytok, ovocie, potrava ...).

Poznámka. Ak rozlišovacie kritérium definujeme ako funkciu do čiastočne usporiadanej množiny a táto množina má maximum (alebo nejakú podmnožinu maximálnych prvkov), zrejme môžeme identifikovať prototyp. Na čiastočne usporiadanej množine sa dá - za istých podmienok - definovať aj podobnosť (prípadne relácia „x sa podobá na y aspoň tak ako na z“).

Predpokladáme, že rozlišovacie kritériá (a operácie na nich) možno definovať a používať tak, aby umožňovali identifikovať prototyp tam, kde to je možné a aby aj umožňovali odvodzovať dôsledky o podobnosti.

Doposiaľ sme sa zmienili o rozlišovacích kritériách *predmetov, prirodzených tried, prirodzených vlastností, prirodzených vzťahov, situácií, typov situácií, ľubovoľných vlastností a vzťahov*. Ak odhliadneme od produktov vysokej schopnosti abstrakcie, ľubovoľných vlastností a vzťahov, ide hlavne o stavebné kamene pomerne jednoduchej kognitívnej schopnosti, klasifikácie a o orientáciu v teréne (hlavne tú zabezpečujú jednoduché vzťahy). Mali by sme sa však venovať aj komplexnejším rozlišovacím kritériám. Kľúčové budú rozlišovacie kritériá, ktoré umožnia postrehnúť, pochopiť dynamiku, zmenu.

Opakovane sme sa už zmieňovali o cieľoch organizmov. Našou ambíciou je zabudovať ciele do (kognitívnej) sémantiky. Domnievame sa, že medzi veľmi dôležité manipulácie na významoch patria tie, čo vedú k formulácii cieľov.

Intuitívne, cieľ je želaná situácia. Doposiaľ sme hovorili o kritériách, ktoré rozoznávajú (rozlišovali) reálne situácie. Teraz potrebujeme nejakú schopnosť, ktorá dokáže *generovať* nejaké (imaginárne) situácie. Túto schopnosť konštruujeme pomocou nejakých operácií, konštruktorov na významoch. Dekompozíciou rozlišovacieho kritériá nejakej situácie a následnou kompozíciou (nahradením) môžeme prísť k rozlišovaciemu kritériu (želanej, imaginárnej) situácie, ktorá sa líši od existujúcej v nejakom dôležitom ohľade.

Naše organizmy interagujú s nejakým prostredím. Okrem iného rozoznávajú (aspoň približne a schematicky) nejaké situácie. Želané situácie sa líšia od rozpoznaných situácií tým, že existujú iba ako významy. Preto s cieľmi môžeme spájať schopnosť (aspoň elementárnu) *sebareflexie (metareprezentácie)*. Môžeme teda predpokladať rozlišovacie kritériá, ktoré rozlišujú rozlišovacie kritériá. Tejto možnosti však nebudeme teraz venovať väčšiu pozornosť. Podobne sa nebudeme zaoberať významami *problémov*.

Veľmi dôležitou kognitívnou schopnosťou je schopnosť *rozpoznať zmenu* v prostredí. Budeme preto predpokladať *rozlišovacie kritériá zmeny* v prostredí. Ďalej si treba všimnúť schopnosť *projektovať* zmenu v prostredí. Sémantiku zámernej (projektovanej) zmeny v prostredí nám poskytnú *plány*.

Príklad. Predstavme si líšku v situácii, keď je hladná a pozerá sa cez plot do záhrady, v ktorej je kurník. Podľa všetkého nie je ťažké uveriť, že cieľom líšky je iná situácia – napríklad taká, že vlečie do priľahlého lesa kuru, ktorej krk má pevne stisnutý v zuboch. Rovnako sa dá dobre predstaviť, že líška dokáže „vytvoriť“ nejaké plány na dosiahnutie tohto cieľa v danej situácii. Navyše, tieto plány splňajú podmienky A – D: Nemusia byť nevyhnutne späť s práve vnímanou scénou (líška si niektorý z týchto plánov môže vymyslieť hlboko v lese). Nie sú deterministické (môže to skúsiť v susednej záhrade alebo v lese). Môžeme sa dohodnúť, že za tým všetkým sú nejaké inferencie a že líška je schopná tvoriť nové plány – pravdepodobne každá líška sa dostane do situácie, v ktorej nikdy predtým nebola a musí ju nejakým spôsobom zvládnuť. Ak sa čitateľovi uvedená predstava zdá trochu divoká, odvolávame sa na analýzu inštinktov ako plánov v časti 6.5.1.

Elementárnym sémantickým požadím (rozlišovacím kritériom) zmeny v prostredí je schopnosť dekomponovať nejakú (relatívne pevnú) scénu tak, že jednu jej súčasť možno (a treba) nahradiť inou. Vyspelejším stupňom je rozpoznanie, aké akcie viedli k tejto zmene (čo takúto zmenu zapríčinilo, zapríčiňuje). Napokon, je tu projektovanie, generovanie akcií, ktoré povedú k takejto zmene. *Plán* budeme považovať za rozlišovacie kritérium nejakých organizovaných (usporiadaných) akcií, ktoré od danej situácie povedú k inej, zmenenej situácii.

Metódy môžeme považovať za plány, ktoré sa osvedčili (nejakej skupine organizmov) a ktoré možno rutinne, až mechanicky, znovu aplikovať.

Dostali sme sa k veľmi dôležitej téme, k téme odvodzovania (inferencie). Plán môžeme považovať za pravidlo nasledujúceho typu: ak sme v nejakej situácii a vykonáme akcie podľa plánu, potom sa dostaneme do (požadovaným spôsobom) zmenenej situácie. Pripomíname, že sme opäť v situácii, keď niečo rozlišujeme na základe iba niektorých symptómov. Preto ide o hypotetickú inferenciu. Pravidlá tohto typu nezaručujú, že cieľ dosiahneme. Niekedy (často) treba vyskúšať iný plán, keď predchádzajúci zlyhal. Pri plánovaní sa opäť uplatňuje metóda pokusov a omylov.

Sémantiku organizmu sme dostali do súvislosti s jeho konaním, akciami. Tým sme si pripravili východisko pre zvládnutie slovík v kognitívnej sémantike.

Už sme poznamenali, že sémantika sloviess je slabým miestom Gärdenforsovho prístupu. Ako dostať zmenu, akcie do Gärdenforsovho konceptuálneho priestoru? Môžeme ho – aspoň teoreticky – „rozbehať, animovať“. Necháme ho vyvíjať sa v čase. Sú s tým späté minimálne dva problémy. Prvý je verziou problému rámca: obrovská časť priestoru by sa v čase príliš nemenila. Preto navrhovaný „animačný“ opis by bol veľmi nevýstižný, neprehľadný a neprimerane rozsiahly. Druhým je budúcnosť: práve tá je z hľadiska modelovania kognície dôležitá. Šimpanz chce v (blízkej) budúcnosti dočiahnuť banány. Ako to vyjadriť v konceptuálnom priestore? Zaznamenať v ňom všetky možné budúcnosti?

Poznámka. Predstavme si teraz, ako analogický problém riešiť v našej teórii. Majme skonštruovaný význam (rozlišovacie kritérium) nejakej situácie (sú tam banány, ale sú nedosiahnuteľné), máme nejaký cieľ (tomu opäť môžeme priradiť nejaký význam, rozlišovacie kritérium situácie, v ktorej sú banány dosiahnuteľné). Potrebujeme teda ešte akési *transformácie* rozlišovacích kritérií. Ak by sme to išli formalizovať, boli by to funkcie definované na funkciách. Vedeli by sme si predstaviť aj rozšírenie gärdenforsovskej sémantiky o transformácie konceptuálnych priestorov.

Spomínané transformácie možno robiť na celých triedach situácií. Keď však zafixujeme konkrétny vstup a výstup takejto transformácie, môžeme hovoriť o *udalosti* (tento potkan teraz našiel cestu v tomto bludisku). Z danej transformácie môžeme získať jej rôzne verzie pridávaním nejakých parametrov (čas, miesto, spôsob, nástroj atď. – čitateľ iste vidí, že smerujeme k zložito štruktúrovaným slovesným frázam; tie nás však ešte nezaujímajú).

Podobne, ako sme sa dostali od situácií k typom situácií, môžeme sa dostať od udalostí k *typom udalostí*. Poznamenajme, že plány sa dajú konštruovať na báze typov situácií a typov udalostí.

Reprezentácia. V tejto chvíli sme schopní vytvoriť si iba predbežný pojem reprezentácie. Vytvoríme si ho však, už ho potrebujeme. Napriek tomu, že reprezentáciu nestotožňujeme s výrazmi nejakého jazyka (naše presvedčenia nie sú formuláciami v nejakom jazyku, naše presvedčenia sú významy), plný obraz o reprezentácii si budeme môcť vytvoriť až nad syntakticky bohato štruktúrovaným jazykom. Takýto jazyk (spolu

s reprezentáciou a inferenciou, ktoré si vynucuje) dovoľuje plné rozvinutie kognície – tak, ako ju poznáme.

Zatiaľ môžeme povedať toto: Reprezentácia je konštruovaná z významov. Významy stotožňujeme s kritériami rozlišovania (kritériá predmetov, tried, vlastností, vzťahov, situácií, cieľov, problémov, plánov a udalostí). Naša koncepcia reprezentácie obsahuje aspoň tri vrstvy:

- konceptuálny aparát (bázu významov) – kritériá rozlíšenia predmetov, tried, vlastností, vzťahov,
- bázu poznatkov (kritériá rozlíšenia situácií, udalostí a ich typov)
- inferenčný aparát (plány).

Uvedomujeme si, že predloženú koncepciu treba dopracovať, zjemňovať a rozširovať.

Všetky tri vrstvy považujeme za dynamické, sú otvorené rozširovaniu a modifikáciám.

6.5 Reprezentácia: od biologického pozadia k syntaktickým štruktúram

6.5.1 Biologické pozadie

V tejto časti zavedieme vlastne iba nové slová, pomocou ktorých budeme hovoriť o tom, čo sme si už pomerne podrobne pripravili. Opakovane sme zdôrazňovali, že rozlišovacie schopnosti, rozlišovacie kritériá, inferenciu a reprezentáciu (minimálne nejaké ich zárodočné podoby) vidíme a hľadáme aj u organizmov s predjazykovým správaním. Kognícia je zakotvená hlboko v biologickej ríši.

Skôr než sa malé dieťa naučí používať nejaké slovo (mama, havo, voda, biely), dokáže rozlíšiť mamu, psa, vodu, bielu farbu (i keď s nejakými chybami). Podobne mačka, líška alebo cap dokážu vo svojom prostredí rozlišovať dôležité predmety, situácie a udalosti. Príklady o šimpanzovi alebo potkanovi sme už uviedli. Pochopenie, rozmýšľanie majú

nesporne korene v rovine biologickej.

Zrekapitulujeme teda pomocou nových slov, čo sme už povedali alebo naznačili o reprezentácii, sémantike, inferencii a komunikácii v prostredí organizmov s predjazykovým správaním. Budeme hovoriť o protoreprezentácii, protosémantike, protoinferencii a protokomunikácii. Všimneme si hlavne ich previazanosť.

Protosémantika, protoinferencia, protokomunikácia. Na „proto“-úrovni predpokladáme iba *rozlišovacie kritériá* pozorovateľných predmetov, vlastností, vzťahov, situácií a udalostí. Okrem toho prijmeme aj predpoklad o niečo silnejší: v protosémantike predpokladáme rozlišovacie kritériá tried (či príslušnosti do triedy) pozorovateľných predmetov. Takéto kritériá sú späté so schopnosťou zovšeobecňovať. (O spomínaných kritériách môžeme hovoriť ako o *protovýznamoch*.)

Už sme zdôrazňovali, že pri tvorení a identifikácii významov hrá dôležitú úlohu (proto)inferencia, uvádzali sme aj viaceré príklady zo sféry predjazykového správania. Pripojíme ďalší.

Príklad. Keď sa dravec priblíži k mláďatám tetra, matka predstiera, že má poranené krídlo (usiluje sa tým odvieť jeho pozornosť od mláďat). Podobne, keď mačka zbadá, že myš sa jej stratila pod závesom, očakáva ju na opačnej strane závesu. Dokáže aj striechnúť pri myšacej diere, i keď nemá nijaký priamy signál o prítomnosti myši na scéne.²⁸

Očividne v pozadí je nejaký plán, nejaký druh jednoduchej inferencie. Výhrada, že správanie tetra je inštinktívne, nič podstatné nemení. Miller a jeho kolegovia (1960) chápali inštinkty ako špeciálny druh plánov. Inštinkty sú podľa nich také plány, ktoré sú vrodené, nemožno preusporiadať ich komponenty a nemožno ich meniť v závislosti od toho, ako vplývajú na organizmus.

Organizmus s protoinferenciou môže používať lesť, môže sa učiť z neúspechov (i z úspechov). Protoinferencia predpokladá často štruktúrovanú reprezentáciu - napríklad účinná lesť je možná preto, že organizmus má reprezentované niečo z reprezentácie iného organizmu (toho, ktorého chce preľstiť). Nadväzne na to môžeme predpokladať, že vo vlastnej reprezentácii má organizmus reprezentovanú aj svoju vlastnú reprezentáciu,

²⁸ Všetky príklady sú z (GÄRDENFORS 1996). Podobné príklady uvádza (DENNETT 1997) a mnohí ďalší.

môžeme teda hovoriť o reprezentácii s reflexiou, s metareprezentáciou (GÄRDENFORS 1996).

Zdôraznime previazanosť protoinferencie a protosémantiky: Plány sa aplikujú na isté situácie (na danú situáciu a na skonštruovanú, cieľovú situáciu). Aplikácia plánu je možná vďaka (proto)sémantickej vybavenosti organizmu, jeho schopnosti tvoriť a používať rozlišovacie kritériá. Na druhej strane, vďaka (proto)inferencii sa cibrí, tríbi rozlišovacia schopnosť, aj schopnosť konštruovať (proto)významy a modifikovať staré.

Organizmy sa v prostredí združujú do skupín. Efektívne fungovanie týchto skupín predpokladá skupinovú koordináciu (rôznych typov). Na úrovni organizmov s protoreprezentáciou môžeme zaviesť predstavu protokomunikácie - myslíme na vysielanie a prijímanie signálov zodpovedajúcich (aspoň niektorým) protovýznamom.²⁹ Podobne konflikty medzi skupinami, ich súperiace záujmy sa prejavujú a riešia vysielaním a prijímaním signálov. Opäť narážame na užitočnosť (a potrebu) metareprezentácie. Možnosti vzájomnej výmeny správ sú účinnejšie vtedy, keď komunikujúce organizmy majú vo svojej reprezentácii istý obraz reprezentácie partnera v komunikácii.

Zhrňme: V (proto)reprezentácii „do seba zapadajú“ a vzájomne si „vynucujú“ zodpovedajúcu silu (proto)sémantika (rozlišovanie pozorovateľného), (proto)inferencia („plánovanie“ akcií, transformujúcich pozorovateľné situácie) a (proto)komunikácia. – vysielanie a prijímanie signálov, zodpovedajúcich niektorým (proto)významom. Testom tejto konštrukcie môže byť model, ktorý skonštruuje rozlišovacie kritériá spomínaných typov a s nimi späté akcie. Pritom výpočty nad takouto reprezentáciou budú produkovať očakávané správanie.

6.5.2 Jazyk

Pod *2jazykom* rozumieme jazyk, ktorého výrazmi sú krátke kombinácie malého počtu slov. Jeho častou podobou sú kombinácie dvojíc slov, odtiaľ náš názov. Používajú ho

²⁹ Odhliadame od emocionálnych a ritualizovaných aspektov komunikácie.

malé deti, cudzinci v ľubovoľnom jazykovom prostredí. Porozumieť mu (jeho fragmentom) dokážu šimpanzy, ale aj psy. Nebude nás zaujímať, ako takýto jazyk (evolučne) vznikol.³⁰ Bude nás zaujímať, aké „kognitívne vybavenie“ predpokladá (vyžaduje). Podobne, takýto jazyk nás nebude zaujímať ako reálny, v nejakom prostredí fungujúci komunikačný prostriedok, bude nás zaujímať iba ako nejaký abstraktný typ, vyžadujúci istú sémantiku, inferenciu a umožňujúci istý typ komunikácie.

Prejdeme preto k reprezentácii, ktorá by (schematicky) mohla zodpovedať jazyku tohto typu. V „biologickom podloží“ sémantických schopností sme predpokladali iba kritériá rozlišovania pozorovateľných predmetov, tried takýchto predmetov, vlastností pozorovateľných predmetov, vzťahov medzi nimi, situácií a udalostí. Komplikovanejšie a abstraktnejšie rozlišovacie kritériá (napríklad vzťahov medzi udalosťami) budeme predpokladať až tu. Prepracujeme sa však k nim od úrovne elementárnych rozlišovacích kritérií pomocou určitých konštrukcií, umožňujúcich potrebné odvodzovanie.

Na používanie 2jazyka treba v reprezentácii nejaké elementárne *konštruktory* na kritériách rozlišovania. Najjednoduchšie z nich sú *booleovské*. Predstavme si, že máme kritériá dvoch vlastností. Môžeme potom zaviesť napríklad kritérium vlastnosti, spočívajúcej v tom, že nejaký predmet má súčasne obe spomínané vlastnosti.

Uplatnenie booleovských operácií si však nepredstavujeme tak, že ľubovoľné rozlišovacie kritériá voľne spájame s ich pomocou. Skôr si predstavujeme, že tieto operácie organizmy, používajúce 2jazyk, aplikujú limitovane, iba na význačné, vyselektované, dôležité dvojice rozlišovacích kritérií.

Komplikovanejšie konštrukcie zaviedol Gärdenfors v (2000). Predstavme si niečo také ako pojem bieleho vína. Iste ho nedostaneme ako prienik bielej farby a vína. V konceptuálnom priestore sa dá robiť akýsi *posun* celého farebného spektra do oblasti, ktorá zodpovedá farbám vína. Výsledkom tohto posunu od bielej je farba, typická pre biele víno.³¹ Schopnosť takýchto „posunov“, *transformácií* je niekde v pozadí za tvorbou metafor, jazykových hier, je to predpoklad kreativity.

Veľmi dôležité pre silu 2jazyka sú konštrukcie dovoľujúce spájať rozlišovacie kritériá

³⁰ Nesporne, vznik takéhoto jazyka niekedy v evolúcii (a všetkých jeho prerokvizít – schopnosti artikulovať hlásky, vytvárať z nich väčšie celky a najmä: rozmýšľať istým spôsobom) treba považovať za fascinujúci.

³¹ Podobné operácie „posunu“ môžeme definovať aj nad rozlišovacími kritériami.

predmetu (prípadne triedy) a rozlišovacie kritériá zmeny.

Príklad. Takáto konštrukcia nám môže spojiť rozlišovacie kritérium medveďa a rozlišovacie kritérium zmeny živého na už nie živé. Dostávame tak rozlišovacie kritérium mŕtveho (zabitého) medveďa.

Takéto konštrukcie možno chápať ako (sémantický) predpoklad tvorenia holých viet. V podstate ide o špecializáciu kritéria zmeny, ktorá sa začne viazať na daný predmet alebo triedu. Pripomeňme, že plány sú špeciálnym prípadom kritérií (projektovanej) zmeny a že na nich možno založiť inferencie, vedúce k tvorbe nových významov.

Prejdeme k poslednému typu konštrukcie, ktorému sa tu budeme venovať. Ide o tému kľúčovej dôležitosti. Výrazy 2jazyka majú potenciál podporovať abstraktnejšie rozmýšľanie.³² Takéto výrazy sa môžu stať (spolu s významami, čo im zodpovedajú) objektom ďalších analýz.

Príklad. Fabulujme: Konštrukcia *zabiť medveď* sa po istom rozmýšľaní ukáže nepostačujúcou. Predpokladajme, že organizmus potrebuje dať do vzťahu dve udalosti – udalosť zabitia medveďa a udalosť, čo sa práve odohráva. Ide tu o *vzťah medzi udalosťami*. Rozlíšenie, že nejaká udalosť sa stala skôr ako iná, je podstatným zosilnením sémantickej výbavy. Vďaka nej možno v 2jazyku vysloviť niečo ako *včera zabiť medveď*. 2jazyk je rozširovateľný – možno ho postupne obohacovať o ďalšie dvojice, trojice, Súbežne s týmto rozširovaním ide rozširovanie triedy významov (kritérií rozlišovania, ktorými organizmus disponuje). Uplatňujú sa pritom učenie, zjemňovanie relácie podobnosti, ktorá je spriahnutá s rozlišovaním. Zväčšujú sa však nároky na analýzu viacznačností.

Príklad. *Zabiť medveď* môže znamenať, že zabil medveď, ale aj, že medveď bol zabitý. Predlžovaním postupností slov (dvojice, trojice, ešte dlhšie) sa stále naliehavejšie ukazuje, že takéto postupnosti musia podliehať (vnútornej) štruktúre, aby boli zvládnuteľné, pochopiteľné.

Zhrňme: Reprezentácia, zodpovedajúca 2jazyku, disponuje konštruktormi, ktoré

³² Podľa Dennetta (1997) živočíchy (ľudia) zosilňujú svoje schopnosti tým, že nechávajú stopy v prostredí. Tieto stopy môžu byť pachové, orientačné tabule na cestách, ale možno za ne považovať aj knihy, experimentálne zariadenia alebo počítače. Takouto stopou (pozorovateľnou, v istom zmysle slova materializovanou) môže byť aj výraz 2jazyka.

z jednoduchších rozlišovacích kritérií kombinujú zložitejšie. Zosilňuje sa tak inferenčná výbava tejto reprezentácie. Tá dovoľuje konštruovať nové abstraktnejšie významy. Zosilňovanie sémantiky a inferencie vedie k tomu, že vyjadrovacie možnosti 2jazyka sa stávajú pritesné. Implementácia takejto reprezentácie a jej analýza je veľkou výskumnou výzvou.

6.5.3 Syntax a propozičná reprezentácia

Predpokladajme, že sme dotiahli do detailného a konzistentného tvaru vymedzenie konštruktorov 2jazyka. Opísali by sme tak takmer úplnú výbavu na zavedenie menných fráz³³. Chýbala by nám však možnosť rozvíjať menné frázy pomocou vedľajších viet. Už z toho vidíme, že posun od 2jazyka k bohatému syntakticky štruktúrovanému jazyku nie je možný s prostriedkami, ktoré sme načrtli v súvislosti s reprezentáciou, potrebnou pre 2jazyk.

Najcitlivejším miestom je tu sémantika slovies. Bez jej podstatného zosilnenia nemôžeme pokročiť.

Terminologická poznámka: Budeme hovoriť o *jazyku so syntaxou* a budeme mať pritom na mysli, že zložené výrazy takéhoto jazyka majú pomerne pevne definovanú stavbu. Pre túto stavbu sú podstatné nasledujúce črty: Postupne možno **vnárať** niektoré syntaktické komponenty do komponentov toho istého typu. Do vety *Muž ukradol vrece* možno „vnoriť“ ďalšiu vetu napríklad takto *Muž, ktorý sa dopoludnia opaloval, ukradol vrece*. Takáto syntax sa zvykne nazývať *rekurzívnou*. Principiálne môžeme vytvárať neobmedzene dlhé výrazy. Ďalej, syntax známych prirodzených jazykov je *závislá od kontextu*. Napríklad, vetu, ktorú možno akceptovať v niektorých kontextoch, nemožno akceptovať vo všetkých kontextoch. Napríklad: *Máte čokoládu? A revolver?*

Ak použijeme *termín jazyk so syntaxou*, budeme mať na mysli syntax, ktorá je *rekurzívna* a *závislá od kontextu*.

³³ Mennú frázu dostávame z vlastných mien, podstatných mien alebo zámen. Môžeme ich kombinovať (rozvíjať) s inými slovnými druhmi. Napríklad: *skala* → *podivne tvarovaná skala*.

Podobne ako Gärdenfors si myslíme, že kooperácia hrala kľúčovú úlohu v evolúcii reprezentácie a kognície (2002). Na rozdiel od jeho konštrukcie sa však domnievame, že toto hľadisko treba do kognitívnej teórie reprezentácie vniesť (predovšetkým) cez sémantiku slovíes (a tým aj slovesných fráz a viet). Naším cieľom nebude vysvetliť, ako sa to dialo v evolúcii. Pokúsime sa načrtnúť predstavu, založenú na sémantike.

Dôležité predpoklady pre charakterizáciu sémantiky slovíes už máme pripravené. Zaviedli sme rozlišovacie kritérium zmeny a projektovanej zmeny (plán). Ukázali sme súvislosť plánov s inferenciou. Treba sa nám už len venovať konštruktorom na týchto rozlišovacích kritériách (a tým aj na plánoch). S ich pomocou dostaneme *kooperatívne plány*.

Kooperácia v skupine umožňuje vykonávať zložitejšie a náročnejšie akcie. Plánovanie takejto kooperácie si vynucuje bohatšie a sofistikovanejšie štruktúry: kto zo spolupracujúcich organizmov bude robiť čo, kedy, kde, ako, v akej nadväznosti a súčinnosti ... *Konštruktory rolových rámcov* sú dôležité pre reprezentáciu koordinácie akcií v spoločenstve. Prv než to vysvetlíme, ukážeme si na príklade, čo sú rolové rámce.

Príklad. Predpokladajme, že reprezentácia obsahuje rozlišovacie kritérium zabitia medveďa, že obsahuje kritérium včerajška, poludnia, muža. Môžeme skonštruovať rozlišovacie kritérium udalosti zabitia, ktorej agentom bol (nejaký) muž, objektom medveď, časom včerajšie poludnie. Takéto kritérium, ktoré dovoľuje odlíšiť agenta, objekt, čas (a mnoho ďalších rolí, aspektov) udalosti, nazvime *rolovým rámcem*. Pridávaním ďalších rolí dostávame čoraz komplikovanejšie rámce.

Rolové rámce sú užitočné nástroje na reprezentáciu akcií, slovíes. Tu nás zaujímajú hlavne koordinované a štruktúrované akcie. Práve konštruktory rolových rámcov dovoľujú z jednoduchších rámcov konštruovať komplikovanejšie, štruktúrovanejšie.

Keď nejaký organizmus vytvorí veľmi zložitý štruktúrovaný význam, založený na koordinovanom pláne a na konštruktoroch rolových rámcov, nemôže ho už spravidla komunikovať v 2jazyku. Po prekročení istej hranice zložitosti sa stáva nevyhnutným zavedenie takých štruktúr do jazyka, ktoré by umožnili jednoznačnejšie vyjadrenie komplexných a štruktúrovaných významov. Ale aj naopak: tvorba zložito štruktúrovaných významov je nemysliteľná bez podpory jazyka so syntaxou. Opäť sme

pri prenikavej Dennettovej myšlienke o tom, že kognícia do prostredia umiestňuje nástroje (znaky), ktoré jej pomáhajú rozmýšľať (1997, s.128 a nn).

Od sémantiky slovies sa ľahko dostaneme k sémantike slovesných fráz a k sémantike viet. Každá slovesná fráza a každá veta je vlastne nejakým rozvinutím rolového rámca slovesa tejto vety. Sloveso je kľúčom k myšlienke vety, k sémantike vety.

Pomocou viet môžeme modifikovať, dopĺňať, rozširovať členy viet, môžeme vnárať vety do viet. Ak pochopíme sémantiku vety, dokážeme sémanticky charakterizovať spomínané modifikácie.

Venujme ešte pozornosť odvodzovaniu (inferencii), predovšetkým hypotetickému, ktoré je kľúčové pre identifikovanie a tvorbu významov. V sémantike slovies a viet je implicitne prítomná hypotetická inferencia. Plány sme si predstavovali aj ako nejaké odvodzovacie pravidlá: ak sme v nejakej situácii a urobíme nejaké akcie, dostaneme inú situáciu. Hypotézy, vygenerované pomocou plánov, treba preverovať. Najlepšie pred uskutočnením zodpovedajúcich akcií, ak sa chceme vyhnúť nepríjemným následkom. Cenným nástrojom preverovania je deduktívne odvodzovanie. Spätne, na báze silnejšieho odvodzovania, môžeme konštruovať zložitejšiu sémantiku a reprezentáciu.

Jazyk so syntaxou umožňuje tvoriť významy, ktoré nie sú konštruovateľné bez neho. Takéto významy sú zavádzané vetami ako „Muž, ktorý včera na poludnie zabil medveďa“ alebo – preskočíme veľkú priepasť – „Číslo, ktoré je deliteľné iba 1 a sebou“ a ďalšími a ďalšími. Niektoré vety zavádzajú isté „jazykové hry“. Sémantiku takejto jazykovej hry možno tiež charakterizovať nejakým rozlišovacím kritériom – takým, ktoré dovoľuje rozlíšiť situácie³⁴, v ktorých je slovo (spojenie slov) použiteľné.

Tvorba a analýza syntakticky správnych viet i významov, prislúchajúcich k nim a k ich častiam, vyžaduje silnú konštrukčnú a inferenčnú výbavu. Vyžaduje vyvinuté hypotetické myslenie, schopné generovať relevantné hypotézy, odmietať a revidovať ich, keď sa ukáže, že nie sú prijateľné. Reálne používanie takéhoto jazyka iba aproximuje syntaktický i sémantický ideál. Preto vyžaduje aproximatívne usudzovanie vo veľmi širokom zmysle - schopnosť pracovať s približnými, „plávajúcimi“ významami

³⁴ Upozorňujeme, že na tomto mieste slovo „situácia“ používame vo voľnom, netechnickom, veľmi všeobecnom zmysle slova.

a konštrukciami, schopnosť tvoriť a chápať metafory, analógie, nové konštrukcie. Vyžaduje schopnosť vidieť to, čo je relevantné a čo nie. Schopnosť objavovať, tvoriť, učiť sa.

Nazdávame sa, že všetky tvrdenia o vrodennom jazyku, či univerzálnej gramatike a argumenty na ich podporu treba posudzovať veľmi opatrne. Je totiž ťažko odlíšiteľné, čo je vlastne vrodené – či skutočne univerzálna gramatika alebo schopnosť usudzovať. Aktívne i pasívne zvládnutie ľubovoľných gramatických konštrukcií, ako aj tvorba a identifikácia významov sú založené na sofistikovanom usudzovaní a učení sa s vhladom.

Podobne ako jazyk so syntaxou dovoľuje tvoriť nové významy, umožňuje aj skonštruovať model ideálnej inferencie. Tomuto modelu sa budeme venovať v nasledujúcej časti, je založený na sémantike, ktorú by sme mohli nazvať propozičnou. Táto sémantika buduje svoje konštrukcie tak, že priradzuje výrazom (idealizovaného) jazyka so syntaxou nejaké objekty. Kľúčové je, že vetám priradzuje pravdivostné hodnoty. Toto zjednodušenie sémantického „obrazu“ je zdôvodnené cieľom, pre ktorý je propozičná sémantika budovaná. Je ním vytvorenie kritérií správneho usudzovania. Hlavnou myšlienkou je pritom zachovávanie pravdivosti: cieľom usudzovania je zaručiť pravdivosť záverov, ak sú pravdivé predpoklady (v prípade hypotetického usudzovania je situácia trochu komplikovanejšia).

Otázka je, prečo sme k téme usudzovania nepristúpili priamo cez budovanie propozičnej sémantiky a formálnych kalkulov. Nazdávame sa, že z hľadiska kognitívnej vedy takýto postup je menej produktívny. Ak chceme pochopiť kogníciu (a usudzovanie), dôležité je pochopiť systém vzťahov od biologického sveta, od protoreprezentácie cez jazyk so syntaxou až k formálnym kalkulom.

Vtip je v tom, že kognície si vytvárajú nástroje, aby si uľahčili činnosť. Sofistikované usudzovanie je jedným z týchto nástrojov. Nie je však mysliteľné bez jazyka so syntaxou. Ten nie je možný bez zodpovedajúcej reprezentácie a inferencie. Opäť sa nám uzatvára vzájomne previazaný kruh komunikácie (v jazyku so syntaxou) reprezentácie a inferencie (založených na sémantike slovesných fráz a koordinovaných plánov). Formálny kalkul a

propozičná sémantika, nástroje na zavedenie kritérií korektnej, ideálnej argumentácie³⁵, sa stávajú nad týmto pozadím nástrojmi ešte viac kultivujúcimi sofistickované ľudské usudzovanie

6.6 Usudzovanie

V súvislosti s plánovaním sme hovorili o usudzovaní spočívajúcom v jednoduchom aplikovaní pravidiel tvaru: *ak daná situácia splňa isté podmienky a vykonáš túto postupnosť akcií, potom výsledná situácia bude vyzerať tak a tak*. Je množstvo skúseností, ktoré umožňujú oceňovať takýto typ odvodzovania, predvídania. Sú však aj skúsenosti, ktoré svedčia o tom, že pravidlá spomínaného druhu nás vedú k chybným záverom.

Ľudia často usudzujú tak, že ani nepotrebnú žiadne pravidlá, ktorými sa riadia. V nejakej situácii jednoducho vidia, čo platí. Skáču k záverom (MC CARTHY 1977). Opäť - sú skúsenosti s tým, že takýto druh usudzovania je užitočný, ale aj množstvo svedectiev o chybných záveroch.

Reálne ľudské usudzovanie málokedy používa zdĺhavé a podrobné odvodzovanie, argumentáciu. Človek veľmi často vie rýchlo prehliadnuť situáciu a prijať adekvátny záver, prípadne rozumnú hypotézu. Dokáže však prejsť k detailnej argumentácii, ak to je potrebné.

Potrebné to býva vtedy, keď narazí na to, že jeho (naivné, nedôkladné) usudzovanie je chybné a vedie ho (pri konaní alebo rozmyšľaní) nesprávnym smerom. Analýza nesprávneho usudzovania vedie k pochopeniu sily a užitočnosti cenného nástroja - *protipríkladu*.

Príklad. Predstavme si, že z viet (predpokladov) „Ak pršalo, tak ulice sú mokré“ a „Ulice sú mokré“ odvodíme „Pršalo“. Z toho odvodíme, že nemusíme ísť do záhrady poliať čerstvo vysadené kvety. Keď kvety zvädnú, môžeme začať pátrať, kde sa stala

³⁵ Poznamenajme, že ide o kritérium *rozlišujúce* korektnú argumentáciu. Vidíme, že kognícia postupne expanduje svoju sémantiku (sémantiky!). Kritériá rozlišovania sa postupne zmocňujú ďalších a ďalších domén.

chyba. Môžeme sa napríklad dozvedieť, že nepršalo, a pochopiť, že z predpokladov, ktoré považujeme za pravdivé, sme odvodili nepravdivý záver. Získali sme tak protipríklad k tejto *schéme usudzovania*: z predpokladov „ak p, tak q“ a „q“ odvod' záver „p“. Vidíme, že pomocou takejto schémy sa dá dostať k nepravdivému záveru z pravdivých predpokladov, teda nejde o spôsob usudzovania, ktorý zaručuje zachovanie pravdivosti našich presvedčení.

Od myšlienky protipríkladu sa môžeme dostať k myšlienke všetkých možných situácií (všetkých možných svetov): Protipríklad ukazuje, že nejaký spôsob usudzovania nie je spoľahlivý (korektný). Preto – ak **nemožno** nájsť k nejakej schéme usudzovania protipríklad, potom musí ísť o korektnú schému, ktorá nám zaručí, že od pravdivých predpokladov sa nemôžeme dostať k nepravdivému záveru. Ak nemôžeme nájsť protipríklad k nejakej schéme usudzovania, znamená to toľkoto: Nech svet vyzerá akokoľvek, nenájedme pre danú schému príklad, kde by všetky predpoklady boli pravdivé, ale záver nepravdivý.

Teda, schéma usudzovania je korektná, ak *v každom možnom svete*, v ktorom sú pravdivé predpoklady, je pravdivý aj záver. Tým sme sa dostali k dedukcii. Deduktívne usudzovanie je korektné, zaručuje pravdivosť záverov, ak sú predpoklady pravdivé.

Zdôraznime, že obrovskú silu získala kognícia práve tým, že začala brať do úvahy všetky možné situácie. Reprezentácia, obohatená o *konštrukt všetkých možných situácií*, umožňuje dôkladnú argumentáciu, *jemné rozlišovanie spoľahlivých a nespoľahlivých argumentov*.

Môžeme hovoriť o *rozlišovacom kritériu správneho (korektného) usudzovania*. Toto kritérium predstavuje význam (sémantiku) dedukcie (deduktívneho usudzovania). Dobrým zvykom sa stalo kritériá tohto druhu vyjadrovať presnými matematickými prostriedkami. My si tento luxus nedoprajeme, napriek tomu sa o tému aspoň obtrieme.

6.6.1 Charakterizácia hľadiska

Dôkladné pochopenie usudzovania je extrémne komplikovaný problém. Ide skutočne o zamotaný komplex problémov, ktoré sú (alebo mohli by byť) predmetom viacerých vied, pokúšajúcich sa o pochopenie kognície. Logika si všíma pomerne jednoduchú stránku tohto komplexu - vety, vyjadriteľné v nejakých jazykoch a podmienky ich pravdivosti v (kombinatoricky) možných svetoch.

V istom zmysle slova sa môžeme oprieť o pozorovateľné javy. Môžeme pozorovať, že človek disponuje nejakými poznatkami, niečo vie. Dá sa aj pozorovať, že z toho, čo človek vie, dokáže odvodiť dôsledky (niekedy triviálne, niekedy zaujímavé, niekedy významné, niekedy nesprávne).

Budeme abstrahovať od fyziologického, materiálneho podložia usudzovania (mozog, nervová sústava) a aj od evolúcie tejto schopnosti. Je to úplne legitímny postoj -- napríklad, ani tatranské kamzíky nás nemusia (ale môžu) zaujímať ako fyzikálne systémy alebo ako produkty evolúcie.

Na čo zameriame pozornosť, sú

- kritériá, umožňujúce odlíšiť „správne“ usudzovanie, argumentáciu od „nesprávneho“,
- aby sme zvládli prvú úlohu, potrebujeme adekvátne odlíšiť rôzne druhy usudzovania,
- napokon, budeme zvedaví na hranice, limitujúce formálne štúdium správneho usudzovania: ako detailne sa dá opísať usudzovanie; môže takýto opis reálne fungovať?

Hľadisko, z ktorého pristúpime k analýze usudzovania, charakterizujeme takto:

- opierame sa o konštrukciu detailných formálnych modelov,
- preferujeme sémantickú špecifikáciu týchto modelov,
- sémantická špecifikácia je založená opäť na nejakom formálnom modeli.

Formálne modely považujeme za užitočné nástroje, ktoré používame, aby sme pochopili komplikované procesy, objekty a situácie. Samozrejme, schopnosť poodstúpiť od takéhoto modelu a skúmať hranice jeho použiteľnosti, je tiež užitočnou vlastnosťou kognície.

Najprv pripravíme východiskové predstavy a základné pojmy. Ukážeme si, že už veľmi jednoduchá predstava o dedukcii vedie k nepríjemne zložitým (nezvládnuteľným, prípadne až nemožným) výpočtom. Predstavíme si aj ďalšie dôvody, prečo dedukcia nestačí na modelovanie a pochopenie usudzovania. Potom sa budeme venovať hypotetickému usudzovaniu a revíziám. Na záver sa budeme pýtať (pokúsime sa zhrnúť), čo nám o kognícii môže povedať tu použitý prístup.

6.6.2 Dedukcia

Už sme videli, že základná myšlienka je kúzelne jednoduchá. Pojem dedukcie a aj reálne používané deduktívne usudzovanie, korektné a nevyvrátiteľné usudzovanie je založené na nemožnosti nájsť protipríklad.

Načrtneme veľmi jednoduchý formálny jazyk – jazyk predikátovej logiky prvého rádu (prvorádový jazyk). Opiera sa o jednoduchú predstavu o svete. Svet tvoria

- indivíduá (z množiny U),
- vlastnosti objektov a vzťahy medzi objektami.

Nad touto predstavou sa konštruuje zodpovedajúci jazyk:

- konštanty (označujú objekty),
- predikátové symboly (označujú vlastnosti a vzťahy),
- premenné.

Rozšírenie takéhoto jazyka („o logiku“), obsahuje *spojky* (nie, ak–tak, a, alebo), *kvantifikátory* (pre každý objekt platí, pre niektorý objekt platí).

Z uvedených prvkov jazyka sa konštruujú zložitejšie výrazy. Kľúčové sú *formuly* – za istých podmienok zodpovedajú vetám (výrokom) obvyklých jazykov, môžu byť pravdivé alebo nepravdivé. Najjednoduchšie formuly sa nazývajú *atomárnymi*.

Jazyk výrokovej logiky je ešte jednoduchší. Obsahuje iba výrokové premenné a spojky.

Úlohou logiky je špecifikovať, čo znamená, že nejaká veta v vyplýva z množiny viet V , a navrhnúť metódu, ako to preveriť.

Už poznáme základnú myšlienku, sémantickú špecifikáciu vzťahu deduktívneho (vyplývania). Dedukciou vždy (v každom možnom stave sveta) získame z pravdivých predpokladov pravdivé závery.

Niekoľko technickejších odsekov. Naše konštrukcie však budú nenáročné a nie celkom dôkladné. *Interpretácia* je funkcia, ktorá priradzuje každej vete jazyka nejakú pravdivostnú hodnotu. Kvôli jednoduchosti budeme predpokladať v tejto chvíli iba dve hodnoty -- pravdu (**t**) a nepravdu (**f**). *Model* množiny viet **V** je taká interpretácia, ktorá priradzuje každej vete z **V** hodnotu **t**. Modelom vety **v** je model množiny viet, obsahujúcej iba **v**, t.j. model množiny { **v** }. Ak **V** je množina viet a **v** veta, potom „**v** vyplýva z **V**“ má tento význam: každý model množiny **V** je aj modelom vety **v**.

Možno navrhnuť čisto syntaktický systém pozostávajúci z neinterpretovaných symbolov a manipulácií s nimi tak, aby sme dostali syntakticky definovaný vzťah *odvodenia*. Navyše, pri vhodnej konštrukcii syntaktického systému presne pokryjeme sémanticky definovaný vzťah vyplývania. Hovoríme, že syntakticky realizované odvodenie je *úplné* práve vtedy, keď pre každú dvojicu **v**, **V**, kde **v** vyplýva z **V**, dokážeme syntaktickými prostriedkami odvodiť **v** z **V**. Odvodenie je *korektné* vtedy, keď pre ľubovoľné **v** a **V** také, že **v** je odvoditeľné z **V**, platí, že **v** aj vyplýva z **V**. Prvorádové odvodenie je aj korektné, aj úplné.

Zamontujme teraz do toho výpočty. Možno napísať taký program, ktorý by bol schopný rozhodnúť pre každú dvojicu (**V**, **v**), či **v** vyplýva z **V**? Žiaľ, pre prvorádové jazyky to nie je možné. Ide o *nerozhodnuteľný* problém - dokázateľne neexistuje algoritmus, ktorý by ho riešil. Tento problém je polorozhodnuteľný: Existuje algoritmus s nasledujúcim správaním. Ak dostane vstup (**v**, **V**), pre ktorý platí, že **v** vyplýva z **V**, tento algoritmus rozhodne, že **v** je odvoditeľné z **V** (ak pre ne tento vzťah neplatí, algoritmus nemusí zastaviť). Aj oveľa jednoduchší problém, problém vyplývania vo výrokovej logike, spôsobuje vážne ťažkosti z hľadiska výpočtového. Existuje síce algoritmus, ktorý ho umožní rozhodovať (známa tabuľková metóda), výpočet podľa tohto (ale aj podľa ľubovoľného iného) algoritmu *nie je zvládnuteľný* v rozumnom čase pre vstupy väčšieho rozsahu.

Teda, pre veľmi jednoduchú predstavu o svete a jazyku, vhodnom na opis takéhoto sveta, sa dostávame do vážnych problémov s vypočítateľnosťou dedukcie. Táto správa je nepríjemnejšia pre doslovné chápanie Searlovej metafory ako pre paradigmu konštrukcie formálnych modelov usudzovania. Uvedené negatívne charakterizácie (nerozhodnuteľnosť a nezvládnuteľnosť) treba považovať za veľmi cenné hodnotenia náročnosti, zložitosti problémov spojených s usudzovaním. Tento vhľad, toto pochopenie by nebolo možné bez konštrukcie detailného formálneho modelu.

Ak nepatrne skomplikujeme spomínanú predstavu o svete a o jazyku, vhodnom na jeho opis, situácia bude ešte horšia. Budeme predpokladať aj vlastnosti vlastností, vzťahy medzi vlastnosťami, vlastnosti vzťahov a vzťahy medzi vzťahmi. V jazyku budeme mať predikátové symboly (druhého rádu), ktorých argumentami môžu byť predikátové symboly (prvého rádu). Kvantifikovať môžeme aj predikátové symboly (prvého rádu). V takomto jazyku sa možno pokúsiť o axiomatizáciu aritmetiky (sme v ňom schopní vyjadriť princíp indukcie). Platí však veľmi nepríjemný výsledok: nie je možné skonštruovať množinu axióm, z ktorých by sa dali dokázať všetky vety pravdivé pre prirodzené čísla. To znamená, že nie je možné vytvoriť úplný axiomatický systém aritmetiky.

V ďalších častiach textu sa budeme pýtať, či existujú také zjednodušenia predstavy o dedukcii, ktoré sú výpočtovo zvládnuteľné. Budeme sa aj pýtať, či dedukcia stačí na všetky problémy, ktoré chceme riešiť usudzovaním. Zavedieme pojem hypotetického usudzovania. Vytváranie hypotéz (domnienok) a ich revidovanie patrí medzi kľúčové kognitívne schopnosti. Čitateľa, ktorý sa zaujíma o podrobnejšie informácie o problémoch, čo nasledujú, odkazujeme na (ŠEFRÁNEK 2000). Napokon, pokúsime sa o istý nadhľad, o analýzu usudzovania z hľadiska (reálnej) kognície.

6.6.3 Limitované logické jazyky

Každá databáza obsahuje okrem explicitne zaznamenaných viet (riadkov, priamo obsiahnutých v nejakej tabuľke), aj implicitné znalosti. Tie si dokážeme z explicitne reprezentovaného obsahu odvodiť.

Príklad. Predstavme si, že v databáze máme iba tabuľku, ktorá zaznamenáva dvojice rodič, dieťa. Informácia o tom, kto je čí starý rodič, nie je v databáze priamo (explicitne) obsiahnutá, je tam však implicitná, dá sa z uvedenej tabuľky vypočítať, odvodiť.

V relačných databázach sa používa jednoduchý, limitovaný jazyk, pomocou ktorého môžeme formulovať požiadavky na získavanie (odvodzovanie) implicitných znalostí. Výhodou za toto „uskromnenie sa“ je výpočtová efektívnosť, odpovede na otázky adresované databáze sa dajú vypočítať v rozumnom čase. Relačné databázy možno považovať za aplikáciu fragmentu logiky, v ktorom sa dajú realizovať efektívne výpočty. Z nášho hľadiska je dôležité všimnúť si, že takéto odpovedanie na otázky nie je založené na vzťahu vyplývania. Nepožadujeme, aby odpoveď na otázku bola pravdivá vo všetkých modeloch databázy, stačí nám preveriť, či platí v danej (aktuálnej) databáze.

Ide o dôležitý trend pri implementácii niektorých typov usudzovania: niekedy nepotrebujeme zisťovať, či nejaká veta v vyplýva z nejakej množiny viet V , stačí nám zistiť, či v je pravdivá v (aktuálnom) modeli množiny V . Hovoríme, že miesto odvodzovania (dokazovania) stačí testovať modely.

Iným užitočným ohraničením jazyka logiky je *hornovská logika*. Táto logika obsahuje iba jednoduché pravidlá typu $A \leftarrow A_1, A_2, \dots, A_n$. Tieto pravidlá obsahujú iba atómy A, A_i . Pod názvom *problém splniteľnosti* budeme rozumieť tento problém: je daná množina V viet výrokovej logiky. Treba rozhodnúť, či existuje interpretácia (priradenie hodnôt t alebo f všetkým atómom), pre ktorú každá veta z V nadobúda hodnotu t .

Kým vo všeobecnosti je problém splniteľnosti formúl výrokového počtu nezvládnuteľný, pre hornovskú logiku (pre množiny pravidiel) existuje algoritmus, ktorý tento problém rozhodne v rozumnom čase (vzhľadom na rozsah vstupov). Za tento zisk z hľadiska výpočtovej efektívnosti sa platí nižšou odvodzovacou (vyjadrovacou) silou: v jazyku hornovskej logiky neodvodíme (nevyjadríme) všetky dôsledky, ktoré možno odvodiť vo

všeobecnom prípade. Napríklad: z p neodvodíme $p \vee q$, z pravidla $p \leftarrow q$ neodvodíme pravidlo $\neg q \leftarrow \neg p$.

6.6.4 Hypotetické usudzovanie

V počiatkoch umelej inteligencie vládla predstava, že k úspechu sa dá dostať touto cestou: Stav vecí vyjadríme (nejakú oblasť reprezentujeme) v jazyku logiky (dostaneme tak množinu viet \mathbf{V}). Problémy z tejto oblasti riešime tak, že logickú reprezentáciu problému, označme ju ako \mathbf{v} , sa pokúsime dokázať z \mathbf{V} . Skoro sa ukázalo, že táto cesta nepovedie. Jednou z príčin boli negatívne výpočtové vlastnosti takto postavenej úlohy (už o nich vieme). Tak sa objavila predstava, že namiesto dedukcie, logického dôkazu treba hľadať metódy rýchleho skákania k záverom.

Ďalšou z príčin hľadania alternatívneho spôsobu usudzovania (namiesto dedukcie) je obvyklá neúplnosť poznania. Ak máme rozhodnúť o pravdivosti vety \mathbf{v} (alebo dokonca množiny viet \mathbf{V}) odvodzovaním, málokedy máme k dispozícii všetky relevantné poznatky, ktoré by nám to umožnili. Preto dedukciu treba popreplietat s generovaním (navrhovaním) hypotéz. Navyše, ani človek neusudzuje tak, že by jeho závery boli „vynútené“ akceptovanými predpokladmi. Dedukcia je skôr okrajový prípad ľudského usudzovania. Spravidla pri našich úsudkoch berieme do úvahy iba kontexty, ktoré považujeme za relevantné, nie všetky možné kontexty (všetky modely). Budeme sa teda venovať hypotetickému usudzovaniu, usudzovaniu, pri ktorom záver nemusí byť nevyhnutne pravdivý, ak sú pravdivé všetky predpoklady. Tento druh usudzovania využíva tzv. common sense (mohli by sme to preložiť ako zdravý sedliacky rozum, ďalej však budeme používať anglický termín).

Zhrnieme základné inutície, ktoré motivujú štúdium takéhoto druhu usudzovania:

- neúplnosť (našich) poznatkov,

- dôležité je analyzovať usudzovanie na pozadí evolúcie poznatkov; poznatky majú svoju dynamiku,
- pre evolúciu poznatkov sú typické revízie; (naše) poznatky sú obvykle revidovateľné,
- závery, ku ktorým prichádza „common sense“, často závisia od kontextu,
- pri usudzovaní často treba zohľadňovať výnimočné prípady,
- usudzovanie často vedie k chybným záverom,
- dôležitou črtou „common sensu“ je zvládanie, riešenie rozporov (nekonzistentností),
- pri usudzovaní je dôležité rešpektovať približnosť, aproximatívnosť, neurčitost', vágnosť.

Dá sa povedať, že existujú dva základné prístupy k štúdiu hypotetického usudzovania. Prvý spočíva v rozlišovaní medzi relevantnými a menej relevantnými kontextami. Definuje sa pomocou nejakej relácie *preferencie* na modeloch. Predpokladá sa také usporiadanie na základe preferencie, ktoré dovoľuje vyčleniť maximálne preferované modely (neexistujú modely, ktoré sú preferovanejšie). Namiesto (deduktívneho) vyplývania sa pracuje s *preferenčným* vyplývaním.

Druhý prístup spočíva v realizácii nejakej predstavy „poupratovania“ množiny presvedčení - definuje sa taká množina presvedčení, ktorá je v istom zmysle slova nasýtená (nič viac ani nič menej nemôže obsahovať) a je aj v dobre definovanom zmysle slova koherentná.

Obom prístupom sa stručne budeme venovať.

6.6.4.1 Preferenčné vyplývanie

Príklad. Predstavme si informácie o odchodoch vlakov. Môžeme si ich vyjadriť množinou viet v jazyku logiky. $V = \{\text{odchod(kúty,4:50), odchod(brno,5:20), odchod(žilina,6:35), odchod(košice,6:58), odchod(zvolen,7:45), odchod(kúty,8:11), odchod(budapest,8:50), odchod(praha,9:09)}\}$. Zdravý rozum („common sense“) na základe informácie vo V usudzuje, že neplatí veta odchod(kúty,8:00). V jazyku logiky si tento dôsledok zapíšeme ako $\neg\text{odchod(kúty,8:00)}$. Akceptujeme ho napriek tomu, že nevyplýva z uvedenej množiny viet V . Ľahko nájdeme model M množiny viet V , ktorý

nie je modelom vety \neg odchod(kúty,8:00). Takýto model môže napríklad priradiť hodnotu t všetkým vetám z V a aj vete odchod(kúty,8:00). Našli sme protipríklad, teda uvedený záver nemôžeme získať dedukciou z V . Napriek tomu, s takýmto typom usudzovania nemáme (spravidla) problémy.

Práve opísaný spôsob usudzovania preferuje niektoré kontexty (modely) pred inými. Model M z nášho príkladu je v istom zmysle slova príliš silný. Pre našu situáciu nie je dôležité, aby veta odchod(kúty,8:00) bola pravdivá.

Každú interpretáciu si môžeme reprezentovať množinou atómov, neobsahujúcich žiadne premenné. Myšlienka: uvedená množina atómov je pravdivá pri danej interpretácii a žiadne ďalšie atómy pri tejto interpretácii nie sú pravdivé. Odteraz môžeme interpretácie porovnávať. Niektoré interpretácie sú podmnožinami iných interpretácií. Z dvoch interpretácií, ktoré sú porovnateľné, je preferovanejšia tá, ktorá je podmnožinou druhej. Spomedzi všetkých modelov budú najpreferovanejšie tie, ktoré sú minimálne (vzhľadom na vzťah podmnožiny). V našom príklade je presne jeden taký model. Predpokladá iba toľko, koľko stačí na to, aby všetky vety z V boli pravdivé a nič viac nepredpokladá.

Zavedieme teraz pojem *preferenčného vyplývania*: veta v preferenčne vyplýva z množiny viet V práve vtedy, keď v je pravdivá v každom najpreferovanejšom modeli množiny V . Príklad s odchodmi vlakov predstavuje spôsob usudzovania, ktorý je známy pod názvom predpoklad uzavretého sveta (closed world assumption, CWA). Ak vychádzame z množiny viet V (z databázy DB), potom $\neg v$, negácia vety v , sa prijíma vtedy, keď v nevyplýva z V (v neplatí v DB).

Ide o elementárny príklad hypotetického usudzovania: hypotézu $\neg A$ prijímame, ak nie je známe, že by bola vyvrátená (že by platilo A). Na príklade CWA si môžeme priblížiť niektoré dôležité vlastnosti hypotetického usudzovania.

Ak doplníme cestovný poriadok o nový spoj, o vetu odchod(a,t), potom hypotézu \neg odchod(a,t), ktorú sme mohli odvodiť z predchádzajúceho stavu „databázy“, už viac nemôžeme odvodiť. Tejto vlastnosti sa hovorí *nemonotónnosť*.

Všeobecne: Ak máme množinu viet A , potom pomocou $C_n(A)$ označme množinu všetkých dôsledkov množiny A (pri spôsobe odvodzovania C_n , o ktorom v tejto chvíli nič

nepredpokladáme). Teraz môžeme definovať *nemonotónny operátor odvodenia*. Ak existujú množiny viet A, B také, že $A \subseteq B$ (A je podmnožinou B) a neplatí, že $C_n(A) \subseteq C_n(B)$, potom hovoríme, že C_n je nemonotónny.

Ďalšia dôležitá vlastnosť: s každým nemonotónnym operátorom odvodenia je v dobre definovanom zmysle slova asociovaná nejaká revízia: To, že platí $A \subseteq B$ a súčasne neplatí $C_n(A) \subseteq C_n(B)$ znamená, že existuje veta v , ktorá patrí do množinového rozdielu $B - A$ (v patrí do B , ale nepatrí do A) a existuje veta w , ktorá patrí do $C_n(A) - C_n(B)$. Môžeme to chápať ako vkladanie nových viet do A , čím získame B , a toto vkladanie nám vynucuje odstraňovanie nejakých viet z $C_n(A)$.

6.6.4.2 Stabilné množiny presvedčení

Druhý z triedy prístupov k štúdiu hypotetického usudzovania možno charakterizovať ako hľadanie stabilných množín presvedčení.

Príklad. Cestovný poriadok platí v obvyklých prípadoch. V mimoriadnych situáciách (pred a po sviatkoch, pri výnimočných udalostiach na trati) nemusí platiť, sú možné nejaké výnimky, zmeny. Vtedy známu (staršiu) informáciu o odchodoch vlakov berieme s rezervou a usilujeme sa zistiť nové informácie. Tento spôsob usudzovania formalizujú defaultové teórie.

Defaultové teórie používajú tzv. defaultové pravidlá. Ak chceme napríklad vyjadriť, že odchod do Kútov v obvyklom prípade nie je o 8:00, urobíme to takto:

$$\begin{aligned} &: \neg \text{odchod}(\text{kúty}, 8:00) \\ &\neg \text{odchod}(\text{kúty}, 8:00) \end{aligned}$$

Uvedený zápis môžeme čítať takto: pokiaľ nie je vyvrátené, že vlak do Kútov neodchádza o 8.00, potom akceptujeme, že vlak do Kútov neodchádza o 8:00.

Všeobecný tvar *defaultového pravidla*:

$$\underline{\quad} :$$

Defaultové pravidlo hovorí: vtedy, keď platí A a nie je známe, že by neplatilo $\neg A$ (nemáme nijaké argumenty proti $\neg A$, nevieme dokázať $\neg A$), akceptujeme A . Veta „Prednáška z kognitívnych vied býva obvykle vo štvrtok“ sa dá pomocou defaultového pravidla vyjadriť takto:

: predn-CogSc-stv

predn-CogSc-stv

Ak nemáme nijakú informáciu, ktorá by vyvrátila, že prednáška z kognitívnych vied bude vo štvrtok, tak prednáška z kognitívnych vied bude vo štvrtok. Defaultové pravidlá však nemôžeme chápať ako pravidlá odvodzovania. Defaultové pravidlá sa používajú na charakterizáciu hypotetického usudzovania pomocou stabilných množín presvedčení. Intuitívne tomu môžeme hovoriť „poupratovanie“ presvedčení.

Myšlienku upratovania postihuje pojem extenzie (defaultovej teórie). *Defaultová teória* (v jazyku J) je dvojica (W, D) , kde W je množina viet jazyka J a D je množina defaultových pravidiel, v ktorých sa vyskytujú vety z J . Množina W predstavuje vety, ktorým veríme, ktoré sú overené. D predstavuje nejaký „generátor hypotéz“. Extenzia je taká množina viet E , pre ktorú platí:

1. $W \subseteq E$ (všetky vety z W patria do extenzie),
2. E je uzavretá vzhľadom na dedukciu (obsahuje všetky svoje deduktívne dôsledky),
3. napokon kľúčová vlastnosť: ak je nejaké defaultové pravidlo aplikovateľné na E , potom jeho dôsledok patrí do E .

Extenzia je *stabilná množina presvedčení* v tomto zmysle slova: Nech E je extenzia, potom E nezmeníme, ak k nej pridávame deduktívne dôsledky a /alebo dôsledky aplikovateľných defaultových pravidiel.

Dôležité je aj upozorniť, že jedna defaultová teória môže mať viac extenzií, niektoré však nemajú žiadnu extenziu. Existenciu viacerých extenzií si môžeme vykladať ako alternatívnosť hypotetického usudzovania. Neexistenciu žiadnej extenzie ako „neupratateľnosť“, ako nemožnosť nájsť rozumnú množinu presvedčení.

Nemonotónne usudzovanie sa intenzívne začalo študovať s cieľom pochopiť, opísať a implementovať rýchle usudzovanie, podobné na „common sense“. Žiaľ, aj preferenčné

vyplývajúce, aj hľadanie stabilných množín presvedčení je výpočtovo ešte ťažšie ako dedukcia.

V prípade preferenčného vyplývania treba kontrolovať reláciu preferencie na modeloch, v prípade stabilných množín presvedčení treba generovať množiny presvedčení a preverovať, či sú stabilné. V oboch prípadoch treba okrem operácií, používaných pri dedukcii, robiť ešte niečo (nákladné na vynaložené zdroje) navyše. To platí vo všeobecnosti. Iba poznamenáme, že existujú nejaké špeciálne prípady, jednoduché formalizmy s limitovaným jazykom, ktoré dovoľujú zvládnuť výpočet (aj v jednom, aj v druhom prístupe).

Usudzovanie (dôkaz) je fundamentálne nedeterministický koncept (PAPADIMITRIOU 1993). Problémy späté s hypotetickým usudzovaním patria medzi ťažké problémy pre nedeterministické výpočtové zariadenia. To, že inferenciu možno z výpočtového hľadiska charakterizovať pomocou nedeterministického výpočtového zariadenia, zosilňuje našu dôveru v nedeterministické chápanie kognície.

Zdá sa teda, že hypotetické usudzovanie nemôže spĺňať naraz tri požiadavky:

- bude používať pomerne silný jazyk (aspoň výrokovologický),
- bude sa usilovať o istý druh dôkladnosti (koherenciu, úplnosť, nasýtenosť, preferovanie relevantnej informácie),
- bude efektívne vypočítateľné.

Čitateľ si iste všimol, že naša charakterizácia formalizácií hypotetického usudzovania obsahovala pomerne vágne, metaforické všeobecné charakterizácie na jednej strane, a konkrétne formalizmy (CWA a defaultové teórie) na strane druhej. Teraz si predstavíme všeobecný rámec, ktorý unifikuje rôzne formalizácie nemonotónneho usudzovania. Je ním *abstraktný argumentačný rámec*. Čitateľ, ktorý nemá chuť na symboly (i v nepatrnom počte) môže časť 6.6.4.3 preskočiť.

6.6.4.3 Abstraktný argumentačný rámec

Keď (metódou pokusov a omylov) generujeme nejaké hypotézy, tie „zapadajú“ do nejakého kontextu. Hypotézy sa opierajú o nejaké argumenty, zdôvodnenia. Niektoré

z týchto argumentov hovoria v prospech (nejakých hypotéz), niektoré v neprospech. Človek sa pýta: zapadá to do seba? Čo z toho do seba zapadá a čo nie? Vytvára si alternatívne obrazy prostredia (sveta). Vytvára si predstavu o tom, ktoré argumenty (v prospech alebo v neprospech hypotézy) sú akceptovateľné a ktoré nie.

Naším cieľom je teraz podať jednoduchú predstavu o akceptovateľných argumentoch. Budeme najprv definovať prijateľnú množinu argumentov. Zdôraznenie množiny je kľúčové – ide skutočne o to, či argumenty „zapadajú do seba“. Cieľom teórie usudzovania (argumentácie) nemôže byť určenie pravdivých argumentov.

Komplikovanejšie konštrukcie možno nájsť napríklad v (BONDARENKO 1993) My sa tu uspokojíme s najjednoduchším prístupom, ktorý prezentoval Dung (1995) .

Budeme potrebovať iba dve veci – množinu argumentov **AR** a dvojicu argumentov, ktoré sa vzájomne atakujú (t.j. máme neprázdnu binárnu reláciu atakovania: existujú argumenty $a, b \in \mathbf{AR}$ také, že a atakuje b). Ďalej sa dohodneme, že množina argumentov S atakuje argument b , ak existuje argument $x \in S$ taký, že x atakuje b . Množinu argumentov S považujeme za bezkonfliktnú, ak neexistujú argumenty $a, b \in S$ také, že a atakuje b .

Budeme hovoriť, že argument $a \in \mathbf{AR}$ je akceptovateľný vzhľadom na množinu argumentov $S \subseteq \mathbf{AR}$, ak proti každému argumentu $b \in \mathbf{AR}$, atakujúcemu a , možno skonštruovať kontraatak v S (S atakuje b).³⁶

Množinu argumentov S považujeme za prijateľnú, ak

- je bezkonfliktná,
- každý argument $a \in S$ je akceptovateľný vzhľadom na S .

Druhá z podmienok hovorí, že každý atak na S možno odraziť.

Príklad. Predstavme si, že $\mathbf{AR} = \{i, j, a\}$, atakuje = $\{(i, a), (a, i), (j, a)\}$. Inými slovami, predpokladáme, že a, i sa vzájomne atakujú, okrem toho j atakuje a . Konfliktné množiny argumentov sú $\{i, a\}$, $\{j, a\}$, \mathbf{AR} . Bezkonfliktné sú všetky ostatné: \emptyset (prázdna množina), $\{i\}$, $\{j\}$, $\{a\}$, $\{i, j\}$. $\{a\}$ nie je prijateľná: j ju atakuje, ale ona neatakuje j . Na druhej strane, \emptyset , $\{i\}$, $\{j\}$, $\{i, j\}$ sú prijateľné. $\{i, j\}$ je maximálna prijateľná.

³⁶ Pripomíname opäť, že cieľom tejto teórie nemôže byť definovať akceptovateľný argument (sám osebe). Nejaký argument môže všeobecná teória argumentácie vyhlásiť za akceptovateľný iba vzhľadom na nejakú množinu argumentov (vzhľadom na nejakú podporu).

Možno stojí za zmienku, že prázdna množina argumentov je vždy prijateľná – nič ju nemôže atakovať (kto nič nehovorí ... :-).

S je preferovaná extenzia, ak je maximálna prijateľná. Poznamenajme, že každý argumentačný rámec má aspoň jednu preferovanú extenziu.

S je stabilná extenzia, ak je

- bezkonfliktná,
- S atakuje každý argument a , ktorý nepatrí do S.

Inými slovami, stabilná extenzia má „vyhranené stanovisko“ ku každému argumentu: ak $a \notin S$, tak S atakuje a , ak S neatakuje a , potom $a \in S$. (Vyššie sme charakterizovali stabilnú množinu presvedčení ako nasýtenú a v dobrom zmysle slova koherentnú množinu – tu definovaná stabilná extenzia je presne ten prípad.)

Základné známe nemonotónne formalizmy možno predstaviť ako argumentačné rámce. Pre každý z nich sa vhodne sformuluje množina argumentov (spravidla ide o nejaké dvojice, kde jeden ich prvok slúži ako podpora pri zdôvodnení, odvodení druhého). Okrem toho sa definuje relácia atakovania: nejaká dvojica (A, a) , kde A je podporou pre a , atakuje dvojicu (B, b) , ak B obsahuje $\neg a$.

Sémantická charakterizácia takto definovaných argumentačných rámcov sa zhoduje s obvykle definovanými sémantikami nemonotónnych formalizmov.

Uzavrieme týmto: V predchádzajúcej časti sme sa trochu vymkli štýlu, ktorý sme uplatňovali doposiaľ. Predviedli sme niekoľko formálnejších – i keď veľmi jednoduchých – definícií. Za zaujímavé pokladáme dve veci. Po prvé, ukazuje sa, že hypotetické usudzovanie možno chápať ako sériu argumentov a kontraargumentov. Po druhé, zdá sa, že reprezentáciu (alebo aspoň jej časť) je vhodné vidieť tak, že spolu s nejakými významami (presvedčeniami) obsahuje aj argumenty, dôvody v ich prospech alebo neprospech. Tieto argumenty môžu byť aj alternatívne.

6.6.4.4 Ďalšie typy hypotetického usudzovania

Nasleduje iba letmá informácia o niektorých typoch hypotetického usudzovania.

Usudzovanie o hierarchiách. Človek svoje poznatky často organizuje do hierarchií. Niekedy sú tieto hierarchie pomerne komplikované - pripúšťajú výnimky, niektoré hierarchické závislosti sú „mäkké“ (neplatia vždy), vyskytujú sa v nich konflikty. Pri štúdiu odvodzovania v hierarchiách sa uplatnili oba tu spomínané prístupy. Po prvé, preferenčný. Ak sa vyskytne v hierarchii konflikt a jedno z konfliktných tvrdení je opreté o špecifickejšiu informáciu, preferuje sa toto tvrdenie. Rozhodnutie o preferencii však nie je jednoduché. Ak by argumentácia, opierajúca sa o špecifickejšiu informáciu obsahovala aj neutralizované tvrdenie, nemožno ju považovať za platnú. Rozhodnutia o vzájomnej neutralizácii a o preferovaní špecifickejšej informácie možno však urobiť až na základe všetkých možných argumentácií, opierajúcich sa o danú hierarchiu. To znamená, že sa robí istý druh globálneho „upratania“ (vytvára sa stabilná množina presvedčení, v ktorej nie sú vzájomne neutralizované presvedčenia).

Indukcia. Zovšeobecňovanie z pozorovaní - v zásade ide o preferenčné vyplývanie, spomedzi všetkých zovšeobecnení danej množiny pozorovaní sa preferujú tie najslabšie.

Abdukcia. Vysvetľovanie - ak je daná množina pozorovaní a nejaká teória, cieľom je hľadať hypotézy, ktoré po pridaní k danej teórii umožnia odvodiť dané pozorovania.

Mnohým ďalším dôležitým typom usudzovania (analógia, usudzovanie za prítomnosti neurčitosti a vágnosti, usudzovanie založené na prípadoch) sme sa na danom priestore nemohli venovať.

6.6.5 Reálne usudzovanie

Prejdime k podstatnej otázke. Celkom prirodzene a oprávnené sa dá spýtať, čo majú spoločné s kogníciou teoretické prístupy k štúdiu usudzovania, založené na nejakých formalizmoch a ako umožňujú pochopiť kogníciu.

Vychádzali sme z toho, že pre pochopenie kognície je dôležitá sémantická špecifikácia „správneho“ usudzovania. Táto špecifikácia nemá byť založená na mentálnych procesoch a stavoch, ale na sfére, ktorá je intersubjektívna, externá vzhľadom na mentálne procesy.

Načrtli sme základné prístupy k takto chápanej sémantickej špecifikácii usudzovania (rôznych jeho typov).

Idealizáciu, ktorej sme sa tu venovali, môžeme nazvať *ideálnym usudzovačom* (aby sme opakovane nepoužívali toto nešikovné pomenovanie, vystačíme so skratkou IU).

Samozrejme, treba sa opýtať, nakoľko sa táto idealizácia podobá na reálne usudzovanie. Núka sa tu isté pozorovanie. Aj veľmi racionálne rozmyšľajúci človek rozmyšľa inak ako IU. Aplikuje nejaké šablóny, vzory. Rýchlo, skutočne rýchlo akceptuje užitočné závery. Niekedy, keď narazí na problém, začne argumentovať. Niekedy mu stačí niekoľko málo argumentov, aby modifikoval vzory, šablóny, ktoré používa, a ďalej úspešne aplikoval tieto modifikácie. Výnimočne je nútený podrobne, zdĺhavo a detailne odvodzovať.

Môžeme teda vysloviť hypotézu i výskumný cieľ. Adekvátne teórie reálnej kognície by sa dali budovať tak, že budú charakterizovať *reálneho usudzovača* (skratka RU). Jeho základnou ambíciou nie je okamžitá dôkladnosť a úplnosť usudzovania, skôr rýchlosť, získaná vďaka používaniu overených vzorov, šablón. Tieto vzory, šablóny možno reprezentovať napríklad okrem iného stabilnými množinami presvedčení (alebo ich fragmentami), podopretými nejakými dobre overenými, „naučenými“ „teóriami“ (bázami poznatkov). V prípade neúspechu, omylov, ťažkostí, problémov pri aplikovaní týchto šablón RU využíva schopnosť krátkej argumentácie. Keď ani to nestačí, „volá“ IU a s jeho pomocou urobí dôkladnú argumentáciu. Táto predstava umožní rekonštruovať rýchle usudzovanie a súčasne udržať „dotyk“ s dôkladným, sémanticky špecifikovaným (ideálnym) usudzovaním.

Určite je však dôležité vedieť posúdiť, či RU usudzuje správne, alebo sa aspoň v rozumnej miere približuje k správne usudzovaniu. Ak teda chceme aproximovať správne usudzovanie pomocou RU, dôležité je vedieť, čo chceme aproximovať, teda ako vyzerá IU.

Práve v oblasti vzťahov medzi rýchlym usudzovačom a ideálnym usudzovačom možno naraziť na hranice abstrakcie od mentálnych procesov. Dá sa predpokladať, že rýchlosť usudzovania je zabudovaná v kognitívnej architektúre. Nie je však vylúčené, že rozdiely v efektívnosti sú spôsobené iba formuláciou problémov: problémy vyplývajú,

preferenčného vyplývania, stability množín presvedčení, konzistentnosti a ďalšie sú síce dôležité problémy, ale pri reálnom usudzovaní to nie sú prvoradé problémy.

6.7 Záver

Pokúsili sme sa demonštrovať, že k pochopeniu kognície prispieva aj symbolová umelá inteligencia, založená na logike. Sústredili sme sa pritom na analýzu a rekonštrukciu usudzovania, najmä hypotetického.

Schopnosť usudzovať sme zasadili do kontextu kognitívnej teórie reprezentácie. Absolvovali sme príliš dlhú cestu k jadrú našej kapitoly – k téme usudzovania. Zaplatili sme tým daň za to, že – aspoň sa nám zdá – nie je dostatočne pochopený význam logiky pre poznanie kognície. Uvedomujeme si, že tento rámec, ktorý sme vytvorili (kognitívna teória reprezentácie), je vágny a bude si vyžadovať ešte veľa práce. Domnievame sa však, že základy, ktoré sme postavili, sú falzifikovateľné a modifikovateľné a, navyše, veríme, že abstrakciou od mentálnych procesov a načrtnutým chápaním kognitívnej sémantiky a na nej založenej predstave posunu od biologických základov významov k syntakticky bohato štruktúrovanému jazyku (z ktorého mohli vyrásť logiky) sme urobili dôležitý krok. Hneď v prvých riadkoch sme avizovali, že s označením *symbolová paradigma* nie sme spokojní. Charakterizácie tejto paradigmy pomocou symbolov a pomocou manipulácií so symbolmi považujeme za zahmlievajúce a nevýstižné. Termín symbol bez dodania sémantického pozadia je príliš univerzálny. Okolnosť, že so symbolmi manipulujú v istom zmysle slova aj neurónové siete, je dosť nepríjemná.

Na tzv. symbolovej paradigme sú totiž hlavné okolnosti, ktorým sa pri tradičných vymedzeniach nevenuje pozornosť. Po prvé, nepostačujúca je redukcia na výpočty realizovateľné na Turingovom stroji. Usilovali sme sa ukázať, že usudzovanie (a reprezentáciu) možno adekvátnejšie charakterizovať pomocou nedeterministického výpočtového zariadenia (napríklad nedeterministického Turingovho stroja). Pripomínáme ďalší zaujímavý formálny model, vhodný na charakterizáciu kognície - interaktívny

a evolutívny Turingov stroj (WIEDERMANN 2002, VAN LEUWEN, WIEDERMANN 2001). Iba pre zaujímavosť dodajme, že výrazné črty interaktívnosti a evolutívnosti má aj naša koncepcia kognície bez mentálnych procesov (koncepcia reprezentácie).

Druhou výraznou charakteristickou črtou tzv. symbolového prístupu je propozičná sémantika. Symboly prvoradej dôležitosti, ktorými sa v rámci tzv. symbolovej paradigmy manipuluje, sú – zo sémantického hľadiska – propozície.

Zaujímali nás v texte tejto kapitoly sémantické základy kognitívnej vedy. Kľúčovou úlohou v tejto súvislosti je vytvoriť všeobecnú (kognitívnu) teóriu reprezentácie. Pokúsili sme sa ukázať, že reprezentáciu vytvára sémantika a inferencia. Rôzne typy sémantiky vyžadujú rôzne typy inferencie a dovoľujú zodpovedajúce typy komunikácie. Zdôrazňujeme, že tieto závislosti a prepojenosť sú mnohostranné. Plný obraz o reprezentácii sa vytvára až nad jazykom so syntaxou. Otvorenou úlohou ostáva dopracovanie toho obrazu.

Na záver si pripomenieme otázku: *kde sú hranice abstrakcie od mentálnych procesov?* Odpovedať na ňu nebudeme, predpokladáme však, že navrhnutý prístup ich umožní identifikovať, ak bude realizovaný v detailných interdisciplinárnych výskumoch. Na viacerých miestach sme narazili na body, kde cítime hranice navrhutej abstrakcie.³⁷

Literatúra

- [1] BACON, F. (1966): *Nové organon*. V: *Antológia z diel filozofov. Humanizmus a renesancia*. Bratislava, s.253-306
- [2] BONDARENKO, A., TONI, F., KOWALSKI, R. (1993): An assumption-based argumentation framework for non-monotonic reasoning. V: *Logic Programming and Nonmonotonic Reasoning. Proc. of Second Int. Workshop*, 171-190. MIT Press.
- [3] CMOREJ, P. (2001): *Úvod do logickej syntaxe a sémantiky*. IRIS, Bratislava
- [4] DENNETT, D. (1997): *Typy myslí. K poznaniu vedomia*. Archa, Bratislava

³⁷ Za množstvo pripomienok k tejto kapitole ďakujeme Martinovi Kanovskému. Veľmi podnetné boli postrehy Dezidera Kamhala. Komentármi k textu pomohli aj Ľubica Beňušková, Ladislav Kováč, Ján Habdák a Jiří Pospíchal.

- [5] DUNG, P.M. (1995): On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. V: Artificial Intelligence Journal, vol. 77, No 2, 321-358
- [6] FREGE, G.(1892): *Über Sinn und Bedeutung*. Zeitschrift für Philosophie und Philosophische Kritik, 100
- [7] GARDNER, H. (1987): *The Mind's New Science. A History of the Cognitive revolution*. Basic Books Inc. New York
- [8] GÄRDENFORS, P. (1996): *Cued and detached representations in animal cognition*. Behavioural Processes 36, 263-273
- [9] GÄRDENFORS, P. (2000): *Conceptual Spaces*. MIT Press
- [10] GÄRDENFORS, P. (2002): *Cooperation and the evolution of symbolic communication*.
- [11] HAYES, P. (1982): *Introduction*. V: P.J.Hayes, M.M.Lucas (eds.), *Proceedings of the Cognitive Curricula Conference*, vol. 2, Rochester
- [12] JOHNSON-LAIRD, P. (1993): *The Computer and the Mind*. Fontana Press, London
- [13] MC CARTHY, J. (1977): *Epistemological problems of artificial intelligence*. Proc. IJCAI'77, 1038-1044
- [14] MILLER, G.A., GALANTER, E., PRIBRAM, K.H. (1960): *Plans and the Structure of Behavior*. New York
- [15] PAPADIMITRIOU, C. (1993): *Computational Complexity*. Addison-Wesley
- [16] POPPER, M., KELEMEN, J. (1989): *Expertné systémy*. Alfa, Bratislava
- [17] PYLYSHYN, Z. (1998): *What's in your mind?* V: Lepore, E., Pylyshyn, Z. (eds.), Rutgers Lectures in Cognitive Science. Oxford: Blackwell
- [18] PYLYSHYN, Z. (1995): *What's in the mind: constraints on mental structures*. ESCOP/ECONA Talk, September 1995
- [19] SEARLE, J.R. (1984): *Minds, Brains, and Programs*. Harvard University Press
- [20] ŠEFRÁNEK, J. (2000). *Inteligencia ako výpočet*. IRIS, Bratislava
- [21] *The Oxford Dictionary of Current English*. Oxford University Press. 1984

- [22] VAN LEUVEN, J., WIEDERMANN, J. (2001): *Beyond the Turing limit: evolving interactive systems*. V: Pacholski, L., Ružička, P. (eds.) SOFSEM'01: Theory and Practice of Informatics. LNCS 2234, Springer, Berlin, 90-109
- [23] WIEDERMANN, J. (2002): *Výpočetní meze kognitivních systému*. V: Kelemen, J., Kvasnička, V. (editori) *Kognice a umělý život II*. Opava, 253-264
- [24] WITTGENSTEIN, L. (1968) *Tractatus logico-philosophicus*. V: *Antológia z diel filozofov. Logický empirizmus a filozofia prírodných vied*. Bratislava, 139-176

