

Opice, psy, sémantika a logika

Dana Retová, Jarmila Šillová, Ján Šefránek

Univerzita Komenského Bratislava
[dana.retova, jarmila.sillova]@gmail.com, sefranek@fmph.uniba.sk

Abstrakt

Mnohé experimenty a pozorovania svedčia o tom, že také kognitívne schopnosti ako chápanie a usudzovanie možno pripísať aj vyšším živočíchom. V tomto článku sa pokúšame skonštruovať sémantický rámec, v ktorom sa dá upresniť predstava o usudzovaní a chápaní agentov s predjazykovým správaním. Významami sú v tomto rámci rozlišovacie kritériá. Zvláštnu pozornosť tu venujeme rozlišovacím kritériám situácií a udalostí. Situáciu alebo udalosť pochopil agent, ktorý disponuje zodpovedajúcim rozlišovacím kritériom. Behaviorálny test takéhoto pochopenia možno založiť na teste usudzovania, Pravidlá usudzovania sú špeciálne rozlišovacie kritériá. V našom príspevku ich ponímame nasledovne: nejakým rozlišovacím kritériám situácií alebo udalostí (významom predpokladov pravidla) priradia nejaké rozlišovacie kritériá (významy záveru).

1 Úvod

Kontext. Charakterizácia usudzovania je od konca 19. storočia takmer výsadnou doménou logiky. Matematická logika sformulovala ideálny konštrukt formálneho, na intuícii a obsahu nezávislého dôkazu, umožnila pochopenie fundamentálnych vlastností teórií, pre ktoré je dôkaz jedinou prijateľnou metódou. Okrem axiomatickej metódy pochopila aj algoritmickejšiu vypočítateľnosť. Hlboké výsledky matematickej logiky ovplyvnili dokonca aj ponímanie vedeckej metódy. Podstatným spôsobom determinovali a determinujú prístup k skúmaniu usudzovania. Aj v umelej inteligencii, v súčasnej teórii reprezentácie znalostí a usudzovania dominuje prístup založený na logike. Optika

a metódy matematickej logiky umožnili prenikavý vhľad a mnohé dôležité výsledky pri skúmaní jazykov na reprezentáciu znalostí a usudzovania.

Problém. Napriek tomu však paradigma matematickej logiky nedáva dostatočné prostriedky na formuláciu a riešenie mnohých problémov reprezentácie znalostí a usudzovania. Matematická logika dokonca neštuduje ani matematické usudzovanie - jej doménou je idealizovaná komunikácia (prezentácia) matematických výsledkov [1]. V umelej inteligencii sa ukázalo, že sú potrebné nové logiky a nové spôsoby ako robiť logiku, aby sa dali formulovať a riešiť dôležité problémy. Preto vznikli nemonotónne logiky. V tejto oblasti je však mnoho otvorených dôležitých problémov a dokonca ešte pravdepodobne neboli sformulované viaceré základné problémy. Výskumy v psychológii ukázali, že je veľká bariéra medzi logickým prístupom k usudzovaniu a tým, ako ľudia usudzujú. Veľkou výzvou je preklenúť túto bariéru a pozitívne ovplyvniť obe strany [6]. Pre kognitívnu vedu sú dôležité aj biologické korene usudzovania a reprezentácie, doteraz sú však na okraji pozornosti (najmä sémantických a logických výskumov).

Cieľ. Pokúsime sa navrhnúť pojmovú aparatúru vhodnú na to, aby sme mohli konzistentným, priehľadným a kontrolovateľným spôsobom hovoriť o reprezentácii, pochopení a usudzovaní agentov nedisponujúcich jazykom v bežnom zmysle slova. Domnievame sa, že takáto pojmová aparatúra môže prispieť aj k návrhu zaujímavých experimentov, čo by mohlo ovplyvniť prekonávanie bariéry medzi formálnymi modelmi usudzovania a ich empirickým výskumom.

Náčrt riešenia. Naš model je sémantický. Ide o koncepciu významov a sémantiky, nezávislú na nejakom jazyku, umožňujúcu hovoriť o významoch a reprezentácii agentov s predjazykovým správaním. Naša sémantika

je sémantikou rozlišovacích kritérií [8-14]. Predbežné nápady o rozlišovacích kritériách situácií, udalostí a pravidiel z [8] však potrebujeme upresniť.

Hlavné výsledky. Zdefinovali sme rozlišovacie kritériá situácie, udalosti, pravidla. Na tomto základe sme vymedzili, čo znamená pochopiť a usudzovať (bez použitia jazyka). Ukázali sme, ako možno opísať experimenty z [2] pomocou tohto pojmového aparátu.

2 Motivácia

Vychádzame z princípu, ktorý sformuloval Max Debrück a ktorý často cituje L. Kováč: *konkrétny biologický fenomén treba študovať na najjednoduchšom organizme, u ktorého tento fenomén existuje.*

Usudzovanie je nesporne biologický fenomén. Niektoré biologické druhy (dokonca aj človek) sú schopné rozhodovať sa na základe informácií, ktoré nezískavajú priamo z prostredia.

Nadväzujeme na pozorovania a experimenty, ktoré registrujú a analyzujú fenomén usudzovania u živočíchov.

Konrad Lorenz [3] píše o psovi, ako „pochopil situáciu“, „pochopil prekvapivo rýchlo“, „pochopil presne“, prejavil sa uňho „záblesk poznania“.

V experimentoch, opísaných v [2], skúmali vedci z Inštitútu Maxa Plancka pre evolučnú antropológiu inferenčné schopnosti psa domáceho a šimpanza (pre nedostatok vhodných jedincov šimpanza bonobo bola vzorka rozšírená aj o jedincov druhu šimpanz učennivý).

Autori experimentu vychádzali z hypotézy, že schopnosti usudzovania týchto dvoch druhov sa budú líšiť v závislosti od prostredia, v ktorom sa druhy vyvíjali; domestikovaný pes by mal byť schopný lepšie spracovať náznaky od človeka, zatiaľ čo šimpanz by mal byť schopný usudzovania na základe kauzality. Na overenie tejto hypotézy použili experiment, ktorý testoval schopnosť jedinca rozpoznať miesto ukrytej potravy podľa troch typov náznakov: komunikatívnych, behaviorálnych a kauzálnych.

Úlohou pokusného zvierata bolo vybrať si z dvoch nádob obrátených hore dnom tú,

pod ktorou sa nachádza potrava na základe indikácie jedným zo spomínaných typov náznakov.

Pri komunikatívnych náznakoch išlo o ukazovanie na nádobu s potravou a o náznak pohľadom v smere umiestnenia nádoby.

Behaviorálne náznaky zahŕňali naznačený pokus experimentátora načiahnuť sa po nádobu alebo jeho (neúspešný) pokus ju otvoriť.

Kauzálne náznaky boli realizované prostredníctvom fyzickej manipulácie s nádobou. Išlo o sluchové a vizuálne indikácie umiestnenia potravy, ako napríklad hrkanie nádobou, na základe ktorých si mohlo zviera nepriamo odvodiť kauzálne vzťahy medzi lokalizáciou zvuku a umiestnením potravy.

Šimpanzy podľa očakávania veľmi dobre reagovali na kauzálne náznaky, no nevenovali veľkú pozornosť ukazovaniu. Psy naopak spoľahlivo nasledovali komunikačné náznaky poskytnuté človekom, ale umiestnenie potravy neboli schopné odvodiť na základe kauzality.

Môžeme to okomentovať aj tak, že informácie o kauzalite a o význame ľudských signálov nemožno získať priamym pozorovaním prostredia, ale na ich získanie sú potrebné také kognitívne operácie ako chápanie a usudzovanie. V experimentoch sa ukázalo, že šimpanzy sú slabšie v chápaní významov signálov, ktoré vysiela človek a v usudzovaní na ich základe. Podobne je to s psami a kauzalitou.

Všimnime si teraz niektoré formulácie z [2]. „Different species are adapted for making inferences of different kind, using different perceptually based cues, in different domains of activity“, „Apes ... seem to *understand* something of the causal connections between objects, movements, and the noises that object movements can produce in different circumstances, and they *use this understanding* to locate hidden food“.

Sme pri jadre nášho problému - podľa citovanej práce (a mnohých ďalších) má zmysel hovoriť o tom, že zvieratá **rozumejú a usudzujú, odvodzujú dôsledky z toho, čo pochopili.**

Vedy, ktoré sa tradične venujú skúmaniu usudzovania, chápania (a súvisiacich fenoménov ako je *poznanie* a *význam*),

napríklad logika alebo teória reprezentácie znalostí, sa sústreďujú na poznatky a úsudky, sformulované v *jazyku*. Chápeme ako veľkú výzvu práve pre tieto vedy skúmať, čo je usudzovanie a pochopenie ako *biologický fenomén*, ktorý sa vyskytuje v prírode v *predjazykovom štádiu*.

Toto úsilie sa môže zdať príliš ambiciózne pretože, ako poznamenáva Dennett [15], ak skúmaný subjekt neuvažuje o svojom okolí pomocou jazyka schopného určitých rozlíšení, potom nemôžeme po úžasných rozlišovacích schopnostiach nášho jazyka chcieť, aby tieto jeho určité myšlienky, spôsoby myslenia alebo druhy citlivosti vyjadrili. Domnievame sa však, že jazyk a teda aj určité formálne modely môžeme nepriamo použiť na to, aby sme tieto špecifickosti popísali.

Veríme, že takto postavený cieľ pomôže hlbšie, mnohotvárnejšie a v inom svetle pochopiť ľudské usudzovanie, poznanie, reprezentáciu poznania a významov. Spomenieme tu myšlienku K. Lorenza, podľa ktorého jeho *fylogenetická metóda* výskumu umožňuje vidieť to podstatne ľudské.

Navyše, o tom, že sa treba usilovať o *znovupochopenie* spomínaných fenoménov svedčia niektoré nedávne (aaalebo súčasné) výskumy ľudského usudzovania.

V Stanforde sa pri vyučovaní logiky používa softvér.¹ Práca s ním výrazne zlepšila efektívnosť vyučovania logiky a výkon študentov. Najväčší a neočakávaný prínos však spočíval v tom, že pri práci s týmto softvérom používali študenti rôzne netradičné (z hľadiska logiky) postupy, ktoré viedli jeho designérov a špičkových logikov, Barwisa a Etchmendyho, k zmene pohľadu na povahu logiky. Počítač zrevolucionizoval ich chápanie najzákladnejších pojmov logiky [1]. Ich koncepcia usudzovania a reprezentácie prekročila tradičný logický rámec.

V umelej inteligencii viedla požiadavka, až nutnosť, implementovať usudzovanie z neúplných báz znalostí k odklonu od monotónnych logík k logikám nemonotónnym, ku „skákaniu k záverom“ [5, 6]. Táto oblasť je v súčasnosti oblasťou intenzívneho bádania a zďaleka ju nemožno považovať za uzavretú. Stále je veľmi nejasné, ako opísať a realizovať rýchle

nemonotónne usudzovanie (skutočné skákanie k záverom). Objavujú sa aj pokusy priblížiť logiku a kognitívnu vedu, preklenúť bariéru medzi kognitívnu vedou/psychológiou na jednej strane a logikou/na logike založenou teóriou reprezentácie znalostí na strane druhej [7, 8].

Hlbšie porozumenie procesov zvieracieho usudzovania nám môže ukázať cestu, ktorou by sme sa mali vydať pri skúmaní ľudského usudzovania, ktoré ani zďaleka podľa viacerých súčasných objavov nie je totožné s bežnou logickou inferenciou. Prostredie, v ktorom sa pohybujeme je tak komplexné a dynamické a dôsledky našich akcií v ňom natoľko nepredpovedateľné, že zväziť všetky alternatívy a logickým spôsobom odvodiť najvhodnejšie akcie je v reálnom čase prakticky nemožné. Akýkoľvek agent konajúci v reálnom prostredí sa pri usudzovaní stretáva s niekoľkými obmedzeniami; v prvom rade je to obmedzený čas, ktorý má na výber akcie, keďže prostredie sa rýchlo mení a problém, ktorý práve rieši, by po nejakom čase už vôbec nemusel byť relevantný. Ďalej sú to obmedzené informácie o vstupoch aj o výstupoch, o možných dôsledkoch akcie a pod.

Podľa neurovedca Antonia Damasia [16] tkvie podstata spomínaného „skákania k záverom“ práve v procesoch spojených s emóciami, ktoré môžu prebiehať na vedomej alebo nevedomej úrovni. Podľa tejto teórie „filtrujú“ emócie veľký počet alternatív, ktoré sú neadaptabilné vzhľadom na vrodené alebo naučené preferencie organizmu a umožňujú zamerať pozornosť tam, kde má biologicky najväčší prospech.

V kontexte experimentu opísaného v [2] by sme teda mohli tvrdiť, že zvieratá boli schopné zamerať svoju pozornosť len na kritériá, ktoré mali pre ne priamu biologickú dôležitosť. Je prirodzené, že pes porozumel ľudským komunikačným náznakom, pretože toto nadanie bolo vypestované u tohto druhu počas tisícročí spolužitia s človekom. Porozumenie tohto druhu sa teda stalo pre psa životne dôležitým. Naproti tomu, šimpanz, ktorému v jeho prirodzenom prostredí pomáha pochopenie určitých základných princípov kauzality vo svete vyhľadávať potravu, prejavoval väčšiu

¹Pozri www.sli.stanford.edu/php.

citlivosť na tieto podnety. U týchto dvoch druhov sa vyvinul biologicky odlišný systém hodnôt, ktorý sa líši vrozenými preferenciami dôležitými pre prežitie.

Je zrejmé, že obom živočíšnym druhom môžeme prisúdiť určité prvky usudzovania, ktoré by sme mohli opísať napríklad pomocou defaultových pravidiel. V ich reprezentáciách nie sú však tieto pravidlá formulované explicitne, ale skôr implicitne vo forme nutkania na určitú akciu, ktoré je spôsobené naučenými príjemnými alebo nepríjemnými pocitmi pri odhadovaní budúcnosti, ktoré sa vyvinuli na základe druhových predispozícií a na základe skúseností.

3 Reprezentácia

V teórii, založenej na logike, možno reprezentáciu znalostí definovať ako trojicu (J, E, Cn), kde J je nejaký (formálny) jazyk na reprezentáciu znalostí, E je množina formúl tohto jazyka, báza explicitne sformulovaných znalostí a Cn je operátor odvodenia, matematická idealizácia inferenčného stroja. Sémantický výskum je orientovaný na charakterizáciu jazyka a operátora odvodenia (definovanie relácie vyplývania).

Ak však máme vytvoriť formálny model usudzovania a pochopenia, vhodný pre agenty, ktoré nepoužívajú jazyk (v nám známom zmysle slova), musíme konštruovať reprezentáciu a sémantiku bez/mimo jazyka. Iba tak sa dá charakterizovať pochopenie a usudzovanie v štádiu predjazykového správania.

Jednou z možností je použiť sémantiku rozlišovacích kritérií [8], jej spresnenie a rozpracovanie pre potreby výpočtového modelovania evolúcie a akvizície významov a jazyka možno nájsť v [9-14].

Vychádzame z toho, že reprezentáciu a významy treba predpokladať aj u organizmov, ktoré nedisponujú jazykom. Reprezentácia prostredia a/alebo vlastných cieľov je tvorená z významov. Významy modelujeme ako rozlišovacie kritériá: ak nejaký organizmus dokáže rozlíšiť, čo je jedlé (nebezpečné, v pohybe a pod.) a čo nie, osvojil si akýsi význam. Malé deti, skôr než si osvoja nejaké slovo, napríklad pes, musia vedieť rozlíšiť

psov od ne-psov (i keď sa môže stať, že za psov považujú aj holuby).

Pochopenie stotožňujeme s *rozpoznávaním* a *tvorením významov (rozlišovacích kritérií)*. Ak organizmus (agent) dokáže používať nejaké rozlišovacie kritérium pri rozpoznávaní objektov, vlastností (atd.), alebo keď vytvára nové rozlišovacie kritériá, aby sa zorientoval v prostredí, môžeme hovoriť, že chápe (vlastnosť, vzťah, situáciu, udalosť, závislosť, pravidlo atd.).

Rozlišovacie kritériá budeme deliť do dvoch veľkých tried - *dlhodobé* (long-term) a *krátkodobé* (short-term) rozlišovacie kritériá. Prvé z nich reprezentujú významy, ktoré sú použiteľné opakovane (individuum - moja mama; vlastnosť - žltý atd.), druhé iba jednorázovo (práve táto situácia). Prvé z nich budeme chápať ako funkcie, priradujúce nejakému objektu hodnotu z intervalu $\langle 0, 1 \rangle$ (poznáme však, že je tu viac možností - napríklad iba 0 alebo 1, nejaká čiastočne usporiadaná množina, zväz atď.).

Druhé z nich budú tiež funkcie, ale definované na abstraktnom objekte *focus*, ktorý reprezentuje momentálne zameranie pozornosti agenta na prostredie. Hodnoty týchto funkcií budú reprezentáciami situácií alebo udalostí. Poznamenajme, že tieto hodnoty sú závislé na agentovi a na kontexte. Reprezentácia situácie alebo udalosti závisí od schopností agenta a aj od jeho momentálneho stavu.

3.1 Situácie

Základná myšlienka: Rozlišovacie kritérium situácie zobrazí focus na ohodnotený graf. Vrcholy i hrany môžu byť ohodnotené rozlišovacími kritériami.

Graf schematicky² reprezentuje situáciu. Vrcholy zodpovedajú individuám alebo

² Reprezentované sú iba tie detaily (individuá, vlastnosti, vzťahy, ...), na ktoré je zameraná pozornosť. Graf, ktorý tu používame, je v hlavných črtách sémantická sieť. Od väčšiny koncepcií sémantických sietí sa však líšime tým, že nevyžadujeme acyklickosť grafov, ktorými reprezentujeme situácie a hierarchia konceptov nie je kľúčovou. Charakteristickou črtou našej reprezentácie je ohodnotenie rozlišovacími kritériami.

triedam, Hrany vzťahom. Používame dve špeciálne ohodnotenia hrán - *is* a *ISA*. Prvé z nich reprezentuje príslušnosť individua v triede, druhé vzťah podtriedy.

V tomto texte budeme reprezentovať iba binárne vzťahy. Ak by prichádzala do úvahy aj reprezentácia viacrozmerných vzťahov, miesto grafu by sme použili hypergraf - hyperhrany sú množiny, nie dvojice, vrcholov.

Príklad: Ak pes vníma situáciu, v ktorej je nejaký človek vedľa nejakej mačky, pričom pes oboch pozná, vie ich odlíšiť od iných ľudí a iných mačiek, reprezentujeme to dvoma vrcholmi a hranou medzi nimi. Vrcholy a hrana sú ohodnotené príslušnými rozlišovacími kritériami. Ak ide o individua pre nášho psa neznáme, na reprezentáciu môžeme použiť 4 vrcholy a 3 hrany. Nové vrcholy reprezentujú triedu ľudí a triedu mačiek a sú označené príslušnými rozlišovacími kritériami. Nové hrany majú označenie *is*. □

Operácie na týchto grafoch: *refine*, *zoom-in*, *zoom-out*, *abstract*.

Operácia *refine* je aplikovateľná na vrcholy a jej výsledkom je graf. Intuitívne to znamená, že každý vrchol grafu môže „ukrývať“ ďalší graf.

Príklad: Ak je vrchol ohodnotený rozlišovacím kritériom konkrétnej lúky, môžeme sa v navrhovanej reprezentácii prepnúť do iného grafu, kde lúka je reprezentovaná viacerými vrcholmi a hranami (rastlinami, stromami, potokom, topologickými vzťahmi atd.). □

Takýto graf budeme nazývať *viacvrstvomým grafom*.

Operácie *zoom-in* a *zoom-out* pridávajú (odoberajú) grafu vrcholy a/alebo hrany. Intuitívne to reprezentuje rozširovanie alebo zužovanie pozornosti.

Abstract je vlastne taký *zoom-out*, ktorý zabúda na vrcholy ohodnotené rozlišovacími kritériami individuí a na hrany z nich vychádzajúce. Komplikovanejšie abstrakcie vyšších úrovní tu nereprezentujeme. Pre doménu zvieracieho pochopenia a usudzovania nám stačí to, čo sme zaviedli.

Pochopenie situácie modelujeme tým, že k focusu je priradený (viacvrstvomý) graf.³ Interpretujeme to tak, že k danej situácii si agent priradil akýsi význam. Uvidíme, že

³ Bez neznámej hodnoty, pozri nižšie.

v tomto poňatí majú aj problémy dobre definovaný význam.

3.2 Problémy

Niektoré vrcholy alebo hrany grafu môžu byť ohodnotené hodnotou *unknown* (*neznáma hodnota*). Vtedy hovoríme, že sme v danej situácii identifikovali *problém*. Situácia nie je úplne pochopená. Riešením problému je situácia, ktorú z danej situácie dostaneme tak, že všetky ohodnotenia *unknown* nahradíme nejakým rozlišovacím kritériom.

Príklad: Niektorý vrchol v nejakom grafe môže byť ohodnotený rozlišovacím kritériom konkrétnej *nádoby*, od neho vedie hrana ohodnotená rozlišovacím kritériom vzťahu *obsahuje* a susedný vrchol je ohodnotený ako *unknown*. Riešením problému v tomto prípade môže byť iný graf, kde *unknown* nahradíme rozlišovacím kritériom *potravy*. □

3.3 Udalosti

Udalosť je pre nás postupnosť situácií, meniacich sa v diskretnom čase. Rozlišovacie kritérium udalosti však nemožno navrhnúť jednoducho ako postupnosť rozlišovacích kritérií situácií, pretože príslušné situácie musia na seba prirodzene nadväzovať. Využijeme tu konštrukt rozlišovacieho kritéria *zmeny* [9,10,11]. Speciálnym prípadom rozlišovacieho kritéria *zmeny* je rozlišovacie kritérium *akcie* (ktorá zmenu spôsobí).

Rozlišovacie kritérium udalosti zobrazí focus na postupnosť tvaru $s_0, z_1, s_1, \dots, z_k, s_k$, kde si sú rozlišovacie kritériá situácií a zi rozlišovacie kritériá zmeny s_{i-1} na s_i .

Príklad: Predstavme si, že nádoba s neznámym obsahom hrká. Reprezentujeme to trojicou - rozlišovacie kritérium situácie s nádobou, ktorej obsah je neznámy, rozlišovacie kritérium zmeny pôvodnej situácie na inú (zodpovedá akcii pohybovania nádobou), rozlišovacie kritérium novej situácie, kde bude reprezentované, že obsah nádoby vydáva nejaké zvuky. □

Operácie: *refine*, *condense* a *abstract*. *Refine* rozloží dvojicu s, z na postupnosť viacerých situácií a zmien, Intuitívne ide o detailnejší

pohľad na príslušnú udalosť/zmenu. *Condense* naopak danú postupnosť situácií a zmien nahradí kratšou postupnosťou, Intuitívne ide o kompaktnější, zhrňujúci pohľad na udalosť. *Abstract* sa vyskladá z abstrakcií situácií, Ide teda o zovšeobecnenie, abstrakciu od individuálnych aspektov východných udalostí. Výsledky abstrakcie sú vlastne typy situácií a/alebo typy udalostí. Neskôr si však zdefinujeme tieto typy ako rozlišovacie kritériá, ktoré pre všetky inštancie takýchto abstrakcií dajú hodnotu (blízku) 1.

3.4 Pravidlá

Príklad: V experimente z [2] sa ukázalo, že opice rozumejú udalosti s nádobou, ktorá pri pohybovaní vydáva nejaké zvuky tak, že nádoba môže niečo obsahovať. □

Opice dokážu *usudzovať* v nasledujúcom zmysle slova. Disponujú rozlišovacím kritériom, ktoré k rozlišovaciemu kritériu nejakého typu udalostí dokáže priradiť rozlišovacie kritérium zmeny (akcie), ktorá udalosť daného typu zmení v inú udalosť. Túto konštrukciu potrebujeme motivovať a spresniť.

Najprv poznámka o tom, či a ako možno predpoklad o tom, že opice (alebo iné vyššie živočíchy) usudzujú, potvrdiť pozorovaním, t.j. ako možno operacionalizovať hypotézu o usudzovaní opíc (a pod.). Samozrejme, inú možnosť ako cez pozorovanie správania nemáme. Pozorujeme správanie opíc (alebo ľubovoľných agentov v predjazykovom štúdiu) v (takmer) konštantných podmienkach, počas udalosti toho istého typu. Predpokladajme, že za týchto podmienok má opica na výber z rôznych možných správání. Ak v štatisticky významnej miere si vyberá jedno z nich a jej rozhodovanie sa zmení pri určitej zmene daných podmienok, môžeme prijať, že sa to udialo na základe usudzovania. Dôležité je dodať, že presnejšia špecifikácia bude s najväčšou pravdepodobnosťou závislá na konkrétnej doméne a konkrétnom rozhodovaní (podobne sú aj defaultové pravidlá doménoovo špecifické).

Vyzerá to teda tak, že tento typ usudzovania možno prilietavo charakterizovať v štýle defaultových pravidiel. Dané podmienky

môžeme reprezentovať nejakým rozlišovacím kritériom typu udalosti (alebo situácie) v predpoklade pravidla. Dôsledkom pravidla bude rozlišovacie kritérium nejakej akcie. Táto akcia povedie k zmene pôvodnej udalosti. V stráži pravidla budú nejaké rozlišovacie kritériá, ktorých cieľom je eliminovať výnimočné, nežiadúce podmienky (ktoré nevhodným spôsobom zmenia podmienky z predpokladu pravidla). Samotné pravidlo je rozlišovacím kritériom (skonštruovaným z rozlišovacích kritérií).

Definície:

Rozlišovacie kritérium *typu udalosti* berie ako vstup udalosť (t.j. hodnotu rozlišovacieho kritéria udalosti) a jej výstupom je hodnota z intervalu $\langle 0, 1 \rangle$.

Príklad: Teda, ak má agent rozlišovacie kritérium (typu udalosti) hrkania, potom priraduje na základe tohto kritéria hodnotu 1 tým reprezentáciám udalostí, ktoré plne zodpovedajú tomuto (jeho) kritériu.

Zavedieme dva druhy pravidiel - akčné a situačné. Dôsledkami prvého z nich sú rozlišovacie kritériá akcie, u druhých rozlišovacie kritériá typu situácie. Intuitívne, prvé vedú agenta ku konaniu, druhé k rozpoznaní stavu prostredia.

Akčné pravidlo je rozlišovacie kritérium (funkcia), ktoré na vstupe dostane rozlišovacie kritérium typu udalosti alebo situácie U (predpoklad), okrem toho môže, ale nemusí dostať na vstupe ďalšie rozlišovacie kritérium typu udalosti alebo situácie V (stráž) a jeho výstupom bude rozlišovacie kritérium zmeny (akcie) W (dôsledok) vtedy, keď U dáva hodnotu blízku 1 a V hodnotu blízku 0. Samozrejme, pri detailnom formálnom rozpracovaní alebo pri aplikácii treba upresniť, čo znamená blízku.

Poznamenali sme už, že ide o konštruované inšpirované pojmom defaultového pravidla. Nepožadujeme však globálnu charakterizáciu korektného a vyčerpávajúceho aplikovania množiny defaultových pravidiel, ktorý sa v defaultovej logike vyjadruje pojmom extenzie.

Situačné pravidlo sa líši od akčného pravidla tým, že jeho výstupom je rozlišovacie kritérium typu situácie.

Navrhujeme hovoriť, že nejaký agent pochopil situáciu S (alebo udalosť U) vtedy, keď disponuje rozlišovacím kritériom typu

situácie (alebo typu udalosti), ktoré aplikuje na S (alebo U) a dostane hodnotu blízku 1.

3.5 Riešenie problémov

Problém máme reprezentovaný ako situáciu s neúplnou informáciou (niektoré vrcholy alebo hrany sú označené hodnotou unknown). Tradičné na logike založené koncepcie reprezentácie znalostí v typickom prípade zúplňovanie informácie redukujú na usudzovanie. Treba k tomu však dodať, že často ide o hypotetické usudzovanie, generujúce nejaké hypotézy, zúplňujúce reprezentáciu (napríklad predpoklad uzavretého sveta alebo abdukcia).

V našom kontexte akciami na doplnenie informácie budú popri usudzovaní (aplikovaní pravidiel) aj prieskum (exploration) prostredia, v ktorom agent rozlíšil (zameril pozornosť na) nejakú situáciu alebo udalosť.

4 Prípadová štúdia

Z experimentov opísaných v [2] si všimnime nasledovný. Experimentátori skúmali usudzovanie šimpanzov tak, že zvieratá si mali po pomocnom náznaku experimentátora vybrať z dvoch ponúknutých nádob, pričom jedna z nich obsahovala potravu. Jedna časť experimentu sa sústredila na kauzálne zvukové naznačenia prítomnosti potravy v nádobe. Išlo o zatrasenie nádoby s potravou sprevádzané prirodzenou produkciou zvuku, ďalej o trasenie prázdnej nádoby, ale aj o produkovanie zvuku umiestnením mobilného telefónu, ktorý zvonil počas experimentu, spolu s potravou do nádoby.

Výsledkami bola 79% úspešnosť šimpanzov pri hrkaní potravou v nádobe, 55% úspešnosť pri trasení prázdnu nádobou a iba 43% úspešnosť v prípade nádoby so zvoniacim mobilným telefónom. Úspešnosť sa merala nájdením potravy (identifikovaním nádoby, obsahujúcej potravu).

Priebeh a výsledok experimentov budeme teraz analyzovať so zameraním na pozorované inferencie. Tie budeme interpretovať ako použitie nejakých pravidiel. Prvé, čo si treba uvedomiť je, že z mnohých

potenciálne použiteľných pravidiel, ktoré majú šimpanzy pravdepodobne zabudované vo svojich reprezentáciách, sa aktivujú iba niektoré. Šimpanzy (a všetky usudzujúce biologické organizmy) zrejme spontánne selektujú na základe **dynamických preferencií** tie pravidlá, ktoré sa viažu na danú situáciu.

V prvom prípade môžeme hovoriť o (vcelku úspešnom v danej populácii) aplikovaní pravidla, ktorého predpokladom je rozlišovacie kritérium nádoby, ktorá pri pohyboch vydáva prirodzené zvuky (typ udalosti), dôsledkom rozlišovacie kritérium akcie, vedúcej k identifikácii nádoby, ktorá niečo obsahuje.

V druhom prípade môžeme správanie šimpanzov interpretovať ako aplikáciu dvoch pravidiel. Prvé z nich je situačné, jeho predpokladom je rozlišovacie kritérium typu udalosti, v ktorej nádoba nevydáva zvuk pri pohyboch a záverom rozlišovacie kritérium typu situácie s nádobou, ktorá nič zaujímavého neobsahuje. Druhé pravidlo je akčné (a eliminačné): ak rozlíšime situáciu s dvoma nádobami, z ktorých jedna nič zaujímavého neobsahuje, potom uzavrieme, že druhá z nich stojí za preskúmanie (akciu). Výsledky experimentu ukazujú, že takéto pravidlá v danej populácii neboli všeobecne uplatňované.

Samozrejme, podobne (ešte horšie) to bolo v treťom prípade. Ten by sme mohli interpretovať tak, že u šimpanzov neprevládla tendencia aplikovať pravidlo z prvého prípadu - počuté zvuky nerozlíšili ako zvuky prirodzene vydávané nádobou s nejakým obsahom.

Môžeme uvažovať aj o verziách pravidiel so strážami. Zaujímavý je tu tretí prípad. Predpokladom pravidla so strážou môže byť vydávanie zvuku počas pohybov nádoby. Strážou môže byť rozlišovacie kritérium toho, že zvuk je vydávaný pohybujúcou sa nádobou. Stráž je narušená vtedy, keď sa zdroj zvuku identifikuje niekde inde. Zrejme je to obtiažne - pozornosť je sústredená na nádobu, s ktorou experimentátor manipuluje. Interpretácia, podľa ktorej v uvedenom experimente úspešné jedince rozpoznali narušenie stráže takéhoto pravidla, je pomerne vieryhodná. Naznačuje to i fakt, že 9 opíc zo 16 reagovalo nervozitou a ústupom (nevedeli si poradiť s netypickou situáciou).

Opäť však treba zdôrazniť, že nehovoríme o reprezentácii (pravidiel), ktorá je explicitná a dostupná introspekcii. Predstavujeme si to tak, že tieto pravidlá sú zabudované (a naučené, prípadne niektoré aj vrodené) ako mechanizmy správania.

6 Záver

Zhrnutie. Experimenty z [2], v ktorých sa skúmala inferencia šimpanzov a psov sme analyzovali v sémantickom rámci rozlišovacích kritérií. Definovali sme rozlišovacie kritériá situácií, udalostí, typov situácií, typov udalostí, problémov a pravidiel. Pravidlá sú rozlišovacie kritériá, ktoré rozlišovacím kritériám situácií a udalostí priradujú nejaké rozlišovacie kritériá. Napokon sme sa pokúsili rekonštruovať inferencie pozorované v experimentoch ako aplikácie nejakých pravidiel (rozlišovacích kritérií zabudovaných v reprezentáciách). Podľa našich vedomostí ide o prvý pokus charakterizovať v nejakom sémantickom rámci inferencie, ktoré neprebiehajú v úrovni jazyka.

Otvorené problémy. Okrem nutnosti doladiť a zjemniť pojmovú aparáturu, použitú v tomto článku, bude treba analyzovať viac zaznamenaných prípadov zvieracieho usudzovania, detailne (oveľa detailnejšie) ich opísať pomocou rozlišovacích kritérií.

Podnety pre teóriu reprezentácie znalostí a usudzovania. Konceptulný aparát rozlišovacích kritérií je flexibilný a možno ho použiť od opisu veľmi jednoduchých prípadov rozlišovania až po rozlišovanie opreté o používanie syntakticky bohato štruktúrovaného jazyka, charakterizovateľného teoreticko-modelovou sémantikou. Domnievame sa, že analýza, formalizácia a implementácia rýchleho usudzovania v jazyku môže významne ťažiť z pokusov pochopiť zvieracie usudzovanie. Na rozlišovacie kritériá situácií, udalostí pravdepodobne možno naviazať rozmanité, heterogénne typy reprezentácií (od propozičných po grafické vizualizácie). Heterogénne reprezentácie a usudzovanie, ktoré ich využívajú považujeme za zaujímavú výzvu pre teóriu reprezentácie znalostí a usudzovania (pozri [1]). Od dôkladného štúdia

zvieracieho usudzovania možno urobiť dôležité kroky pre pochopenie špecificky ľudského usudzovania (a predpokladáme, že tak možno dosiahnuť podstatne mnohotvárnejšie charakterizácie, než sú tie, ktoré dnes študuje logika).

Literatúra

- [1] J. Barwise, J. Etchemendy: Computers, visualization, and the nature of reasoning. In: *The Digital Phoenix: How Computers are Changing Philosophy* (T.W. Bynum, J.H. Moor, eds.), 1998, pp. 93-116.
- [2] J. Bräuer et al.: Making Inferences about the Location of Hidden Food: Social Dog, Causal Ape. In: *Journal of Comparative Psychology* 2006, vol. 120, No. 1, 38-47
- [3] K. Lorenz: *Man Meets Dog*. 1954
- [4] M. Minsky: *A Framework for Representing Knowledge*, MIT-AI Laboratory Memo 306, June 1974.
- [5] J. Mc Carthy: Circumscription - a Form of Jumping to Conclusions. *Artificial Intelligence* 13 (1980), 27-39
- [6] K. Stenning, M. van Lambalgen: *Human reasoning and cognitive science*. MIT 2007 (to appear)
- [7] M. van Lambalgen, F. Veltman: *The role of nonmonotonicity in cognition*. Project proposal. <http://staff.science.uva.nl/~michiell/docs/WebversieOnderzoek.pdf>
- [8] J. Šefránek: *Kognícia bez mentálnych procesov*. In: J. Rybár, L. Beňušková, V. Kvasnička (eds.): *Kognitívne vedy*. Kaligram Bratislava, 2002.
- [9] Takáč, M.: Konštrukcia významov a jej dynamika v procese iterovaného učenia. Accepted to *Kognícia a umelý život VII*. Smolenice, June 2007.
- [10] Takáč, M.: Cognitive Semantics for Dynamic Environments. In: Hitzler, P., Schärfe, H., P. Øhrstrøm (eds.): *Contribution to ICCS 2006 - 14th International Conference on Conceptual Structures*. Aalborg University Press, Aalborg, Denmark, 2006, pp. 202-215.
- [11] Takáč, M.: Kognitívna sémantika rozlišovacích kritérií. In: Kelemen, J., Kvasnička, V. (eds.): *Kognícia a umelý*

život VI. Slezská univerzita, Opava, 2006, pp. 363-372.

- [12] Takáč, M.: Categorization by Sensory-Motor Interaction in Artificial Agents. In: Fum, D., Del Missier, F., Stocco, A. (eds.): *Proceedings of the 7th International Conference on Cognitive Modeling*. Edizioni Goliardiche, Trieste, Italy, 2006, pp. 310-315.
- [13] Takáč, M.: Autonomous Construction of Ecologically and Socially Relevant Semantics. *Cognitive Systems Research*, under review.
- [14] Takáč, M.: Kognitívna sémantika komplexných kategórií založená na rozlišovacích kritériách. In: Kvasnička, V., Trebatický, P., Pospíchal, J., Kelemen, J.: *Myseľ, inteligencia a život*. Vydavateľstvo STU, Bratislava, 2007.
- [15] Dennett, D. C.: *Druhy myslí. K pochopení vědomí*. Archa, Bratislava, 1997, ISBN 80-7115-140-8
- [16] Damasio, A. R.: *Descartes' error: emotion, reason, and the human brain*. A Grosset/Putnam Book, New York, 1994, ISBN 0-399-13894-3.