

# Kognitívna sémantika komplexných kategórií založená na rozlišovacích kritériách

Martin Takáč<sup>1</sup>

**Abstrakt.** V článku prezentujeme sémantiku vhodnú na reprezentáciu rôznych typov konceptov (objektov, vlastností, vzťahov, zmien, udalostí a situácií) v dynamických a otvorených prostrediach, pre kognitívne modely vývoja a osvojovania jazyka. Sémantika je založená na perceptuálne ukotvených štruktúrach – rozlišovacích kritériách, ktoré sú osvojiteľné na základe interakcií s prostredím a inými agentmi.

## 1 Úvod

Osvojovanie si jazyka sa tradične chápe ako problém získania správnych asociácií medzi externe pozorovateľnými formami (slovami, vetami, gestami) a internými význammi [23]. Významy, ako mentálne koncepty, sa buď považujú za vrodené [10], získané interakciou so svetom [5] alebo vytvorené čisto vplyvom samotného jazyka [40]. Ak by sme prijali extrémny názor, že všetky koncepty sú vrodené, osvojovanie si jazyka by spočívalo iba v naučení sa nových slov (nálepiek) pre existujúce významy. Ak by sme prijali opačný extrém, bez jazyka by nemohlo existovať žiadne myslenie.

Empirické výskumy napovedajú, že deti majú rôzne poznatky o svete omnoho skôr, ako začnú hovoriť [15], osvojujú si rôzne zručnosti a senzomotorické schémy [28] a rozlišujú medzi rôznymi objektmi a situáciami [29]. Na evolučnej scéne sa jazyk objavuje pomerne neskoro, napriek tomu sú mnohé živočíšne druhy schopné kategorizovať (napr. identifikovať objekty vhodné na zjedenie alebo vyhodnotiť situáciu ako nebezpečnú). Slovo koncept, resp. kategória<sup>2</sup> budeme odteraz používať na označenie abstrakcie schopnosti rozlišovať medzi objektmi, situáciami, stavmi atď. [31] (teda ak napr. zvieratá dokážu isté objekty identifikovať ako jedlo, povieme, že disponuje kategóriou jedla). Z takejto definície vyplýva, že (aspoň niektoré) koncepty môžu existovať aj nezávisle od jazyka. V súvislosti s jazykom však hrajú dôležitú úlohu, pretože slúžia ako významy jazykových výrazov a poskytujú tak potrebné ukotvenie symbolov [16]. V tomto zmysle sú koncepty stelesnené neverbálne štruktúry späté s perceptuálnou skúsenosťou a fyzicky korešponujú s aktiváciami neurálnych okruhov korelujúcich s vnímaním, konaním, predstavovaním si alebo rozprávaním a počúvaním o obsahu, ktorý reprezentujú [3]. Kognitívna sémantika sa zaoberá skúmaním a formalizáciou takýchto štruktúr [11, 23, 35].

Podľa nášho názoru je dôležité, aby sémantika spolu s návrhom reprezentácie konceptuálnych štruktúr opisovala aj mechanizmy ich osvojovania. Takto poňatá sémantika môže nájsť uplatnenie nielen v kognitívnych modeloch, ale aj v industriálnych aplikáciách. Moderné technológie založené na autonómnych agentoch, ktoré navzájom vyjednávajú, vymieňajú si dáta a koordinujú svoje aktivity v otvorených a dynamických prostrediach, kde žiadna preddefinovaná ontológia či jazyk nemôže postihnúť vopred všetky potenciálne komunikačné témy, si vyžadujú agenty schopné

<sup>1</sup> Katedra aplikovanej informatiky FMFI UK, Mlynská dolina, 842 48 Bratislava, E-mail: takac@ii.fmph.uniba.sk

<sup>2</sup> Výrazy koncept a kategória budeme v texte voľne zamieňať, používame ich v rovnakom význame.

vytvoriť si a kontinuálne aktualizovať ontológie relevantné prostrediu a vzájomným interakciám a vytvoriť si vlastný komunikačný systém.

Väčšina multiagentových modelov emergencie jazyka (prehľad v [22, 30]) však neposkytuje dostatočné prostriedky na reprezentáciu dynamických konceptov ako sú procesy, udalosti, akcie a situácie, ktoré sú dôležité pre sémantiku slovies. V kognitívnej sémantike konceptuálnych priestorov [11] sú koncepty reprezentované ako oblasti v mnohorozmernom geometrickom priestore. Riešenie sémantiky slovies pridaním časovej dimenzie je však problematické [12, 31].

V tomto článku prezentujeme novú sémantiku, ktorá sa snaží prekonať vyššie opísané nedostatky. Sémantika je založená na perceptuálne ukotvených konceptuálnych štruktúrach – rozlišovacích kritériách, ktoré sú jednotiacim formálnym rámcom pre reprezentácie objektov, vlastností, vzťahov, zmien, udalostí a situácií. Významy slovies sú reprezentované ako medzikategoriálne asociácie kritérií. V článku definujeme a analyzujeme vzájomné vzťahy perceptuálnej, reprezentačnej, jazykovej a pragmatickej úrovne ako aj dôsledky spriahnutia procesov formovania kritérií a osvojovania jazyka.

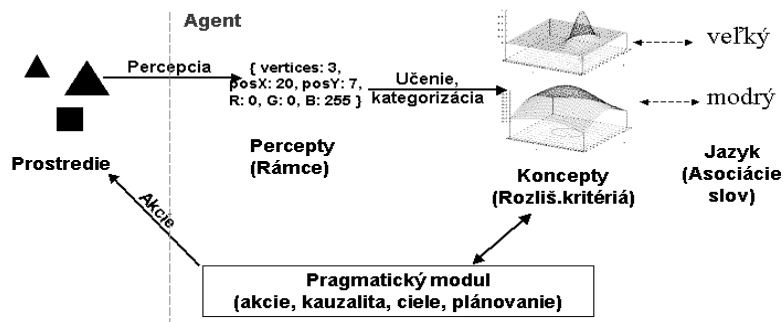
## 2 Od percepcie ku kategóriám a správaniu v prostredí

Akýkoľvek agent – améba, webovský robot alebo človek, reagujúci na isté stavy svojho prostredia istým spôsobom, potrebuje rozlišovať medzi stavmi prostredia tie, ktoré sú relevantné jeho cieľom. Táto elementárna schopnosť považovať niektoré stavy prostredia za identické (vzhľadom na nejaké kritériá) a odlišovať ich od iných stavov, je jadrom každej kategorizácie.

V najjednoduchšom prípade čisto reaktívnych agentov chýba akákoľvek vnútorná reprezentácia a celé správanie je realizované formou priamych spojov „stimul-reakcia“ asociujúcich perceptuálne vstupy s behaviorálnym výstupom. V tomto článku sa chceme venovať zložitejším kognitívnym agentom vykonávajúcim úlohy vyžadujúce flexibilitu, učenie, vzájomnú kooperáciu a komunikáciu v dynamických prostrediach. V ďalšom texte použijeme štyri úrovne opisu (obr. 1):

- **Perceptuálna úroveň.** Táto úroveň tvorí rozhranie medzi externým prostredím agenta a vyššími úrovňami. V stelesnených agentoch zodpovedá signálom zo sensorov agenta predspracovaným nízkoúrovňovými perceptuálnymi rutinami. V softvérových agentoch zodpovedá vstupným dátam, na ktorých agent operuje, opísaným formou, ktorá je spracovateľná na reprezentačnej úrovni.
- **Reprezentačná úroveň.** Toto je úroveň kategórií, resp. konceptov. Každý koncept je reprezentovaný rozlišovacím kritériom – funkciou, ktorá vracia hodnotu pravdepodobnosti, že jej vstup je inštanciou reprezentovaného konceptu.
- **Jazyková úroveň.** Rozlišovacie kritériá agenta sú súkromné a nie sú priamo prenositeľné ostatným agentom. Agenty komunikujú pomocou konvenčne ustanovených signálov jazykovej úrovne. Význammi jazykových signálov sú perceptuálne ukotvené kritériá z reprezentačnej úrovne. Agenty môžu úspešne komunikovať, iba ak sú ich súkromné významy dostatočne podobné. To môže nastať, pokiaľ agenty používajú podobné mechanizmy formovania konceptov na základe podobných skúseností v zdieľanom prostredí.
- **Pragmatická úroveň.** Na tejto úrovni agenty konajú v prostredí, plánujú a dosahujú ciele. Potrebné kauzálne znalosti o vlastných akciách a ich dôsledkoch sú reprezentované medzikategoriálnymi asociáciami kritérií, ciele sú reprezentované

kritériami želaných situácií a plány sekvenciami akcií, ktoré by mali viesť z aktuálnej situácie do cieľa [31].



**Obrázok 1.** Kognitívna architektúra agenta zahŕňa percepčnú, reprezentačnú, jazykovú a pragmatickú vrstvu.

### 3 Perceptuálna úroveň

Budeme predpokladať, že dynamické prostredie agentov sa mení v diskrétnych krokoch. V každom kroku každý agent vníma svoje okolie, aktualizuje si vnútornú reprezentáciu a môže komunikovať alebo vykonávať akcie v prostredí, podľa konkrétnej aplikácie. Perceptuálny vstup agenta v danom čase – *scénu* – opíšeme množinou rámcov zložených z párov atribút-hodnota, napr.

{vrcholy:3; veľkosť:5; posX:12; posY:18}.

Rámce reprezentujú napr. fyzické objekty v okolí agenta, prichádzajúce dáta, alebo agentove „proprioceptívne“ vnemy (hodnoty vnútorných stavových premenných, parametre vykonávaných akcií, pozíciu ramena atď.). Formálne je perceptuálny rámec  $f$  charakterizovaný množinou atribútov  $A_f$  a selektorom atribútov  $h_f: A_f \rightarrow \mathbf{R}$  ( $\mathbf{R}$  označuje množinu reálnych čísel). V texte budeme namiesto  $h_f(a)$  používať zaužívanejšiu notáciu  $f.a$ .

Používanie rámcov alebo iných arbitrárnych amodálnych symbolov na sémantickú reprezentáciu kritizoval Barsalou [3]. V našom modeli nepoužívame rámce ako reprezentáciu; sú len výpočtovým opisom štruktúr, ktoré vznikajú ako výsledok nízkoúrovňového predspracovania senzorického vstupu agenta. Môžu opisovať napr. pole intenzít zo siete alebo web-kamery alebo segmentáciu scény na objekty a ich perceptuálne charakteristiky. Ak dva agenti vnímajú rovnakú scénu, dostávajú identické množiny vstupných rámcov, ďalej ich však môžu spracovávať, kategorizovať a reprezentovať rôzne.

### 4 Reprezentačná úroveň

Agent sa postupne učí rozlišovať vlastnosti prostredia a združuje do kategórií entity, ktoré sú z nejakého hľadiska podobné. Každá kategória je reprezentovaná tzv. rozlišovacím kritériom – aktivačnou funkciou, ktorej hodnota vyjadruje stupeň príslušnosti vstupu funkcie k reprezentovanej kategórii. Vstupom funkcie môže byť rámec objektu (u kritérií objektov a vlastností), viacerých objektov (u kritérií vzťahu), rámce toho istého objektu v rôznych časoch (u kritérií zmeny) aj výsledky iných rozlišovacích kritérií (u kompozičných kritérií situácií a udalostí). Na začiatku agent

nemá žiadne kritériá – tie sú budované na základe príkladov extrakciou ich spoločných štatistických vlastností.

Extrakciu štatistických vlastností z príkladov a ich použitie na kategorizáciu nových príkladov zabezpečujú tzv. *lokálne reagujúce detektory*, ktoré sú jadrom každého rozlišovacieho kritéria. Lokálne reagujúci detektor dostáva ako vstupný argument jediný rámec a vracia hodnotu z uzavretého intervalu  $[0, 1]$  vyjadrujúcu, do akej miery je rámec inštanciou reprezentovaného konceptu (1 znamená najlepší príklad – prototyp).

#### 4.1 Geometrický pohľad na kategórie

Lokálne detektory majú intuitívnu geometrickú interpretáciu založenú na konceptuálnych priestoroch [11]. Konceptuálny priestor je geometrický priestor s dimenziami zodpovedajúcimi atribútom reprezentovaných entít. Keďže nie všetky atribúty sú relevantné pre každú reprezentovanú entitu, dimenzie sú organizované do podprieštrov – domén. Konkrétny objekt je reprezentovaný ako bod v priestore zloženom z jednej alebo viacerých domén a podobnosť medzi dvoma objektmi je nepriamo úmerná vzdialenosti ich bodov (aby bolo možné vypočítať vzdialenosť, body musia zdieľať podpriestor so spoločnou metrikou). Reprezentácia prirodzených kategórií je založená na predpoklade konvexnosti: ak dva body reprezentujú objekty, ktoré sú dobrým príkladom nejakej kategórie, aj všetky body ležiace medzi nimi by mali byť dobrým príkladom tejto kategórie [11]. Prirodzené kategórie sú teda reprezentované konvexnými oblasťami v konceptuálnom priestore. Aktivita lokálne reagujúceho detektora bude tým vyššia (bližšia k jednej), čím je jeho vstup lepším príkladom reprezentovanej kategórie. Lokálne reagujúci detektor môže reprezentovať kategórie s neostrými (fuzzy) hranicami, ale z praktických dôvodov môžeme stanoviť rozhodovací prah detektora  $\theta$  pre vymedzenie hranice kategórie. Oblasť vstupného priestoru  $D$ , pre ktorú detektor  $r$  vracia nadprahovú aktivitu, nazveme jeho *receptívnym polom*  $\Psi_\theta(r) = \{x \in D \mid r(x) > \theta\}$ . Lokálne reagujúce detektory majú vysokú neurálnu a biologickú plauzibilitu [2, 17, 26].

#### 4.2 Indukcia lokálne reagujúceho detektora z príkladov

V tejto časti opíšeme rôzne verzie mechanizmu vytvorenia detektora na základe extrakcie štatistických vlastností súboru vstupných príkladov. Čitateľ, ktorý sa nezaujíma o matematické detaily, alebo je preňho použitý matematický aparát príťažlivý, môže preskočiť na časť 4.3.

Predpokladajme, že agent má odvodiť reprezentáciu neznámeho konceptu z príkladov, teda vytvoriť lokálne reagujúci detektor zo sekvencie rámcov  $\{f^{(1)}, \dots, f^{(N)}\}$ .<sup>3</sup> Každý rámec  $f^{(i)}$  zodpovedá bodu  $(f^{(i)}.a_1, \dots, f^{(i)}.a_k^{(i)})$  v konceptuálnom priestore s dimenziami definovanými atribútmi  $a_j \in A_f^{(i)}$ , kde  $k^{(i)} = |A_f^{(i)}|$  je ich počet. Indukciu založíme na štatistických vlastnostiach hodnôt atribútov spoločných pre všetky rámce, teda najprv urobíme projekciu všetkých rámcov do spoločného podpriestoru  $A$  definovaného prienikom množín atribútov  $\bigcap_{i=1}^N A_f^{(i)}$ . Atribúty, ktoré sa nevyskytujú vo všetkých

<sup>3</sup> V praxi je to zložitejšie, agent v priebehu času vníma zmiešanú sekvenciu inštancií rôznych konceptov. Dokonca jeden perceptuálny rámec môže byť dobrou inštanciou viacerých konceptov, napr. jedno konkrétne jablko môže byť príkladom konceptov označených slovami „červený“, „okrúhly“, „malý“, „jablko“. Problémom určenia, ktorý percept je inštanciou ktorého konceptu, sa zaoberáme v časti 5.

rámcoch, budeme považovať za irelevantné pre príslušnosť ku kategórii. Preto môžeme teraz príklady kategórie reprezentovať ako množinu vektorov  $\vec{x}^{(i)} = (x_1^{(i)}, \dots, x_n^{(i)})$  pre  $i = \overline{1, N}$  v spoločnom priestore  $A$  s dimenzionalitou  $n = \left| \bigcap_{i=1}^N A_{f^{(i)}} \right|$ .

Za najlepší príklad – prototyp kategórie budeme považovať geometrické centrum, resp. priemer všetkých príkladov

$$\vec{p} = \frac{1}{N} \sum_{i=1}^N \vec{x}^{(i)}.$$

Lokálne reagujúci detektor  $r_p$  založený na prototypu vráti aktiváciu exponenciálne klesajúcu s vzdialenosťou od prototypu

$$r_{\vec{p}}(\vec{x}) = \exp(-k \cdot d(\vec{p}, \vec{x})), \quad (1)$$

kde  $k$  je kladná konštanta,  $d(\vec{p}, \vec{x})$  je vhodne zvolená metrika a  $\vec{x}$  je projekcia vstupného rámca do priestoru  $A$  (ak sa rámec nedá projikovať, pretože v ňom chýbajú niektoré atribúty z  $A$ , detektor vráti nulu).

Tvar receptívneho poľa  $\Psi_{\theta}(r) = \{ \vec{x} \in A \mid r_{\vec{p}}(\vec{x}) > \theta \}$  v priestore  $A$  závisí od použitej metriky. V prípade Euklidovskej metriky  $d_{L_2}$  je to  $n$ -rozmerná hyperguľa so stredom  $\vec{p}$  a polomerom  $\theta$ .

Pri použití jediného rozhodovacieho prahu  $\theta$  by mali receptívne polia všetkých detektorov rovnaký polomer bez ohľadu na distribúciu príkladov, z ktorých boli odvodené. To nie je žiaduce, ak potrebujeme reprezentovať kategórie s rôznou mierou všeobecnosti. Najjednoduchším vylepšením je použiť funkciu vzdialenosti váhovanú prevrátenou hodnotou rozptylu  $\sigma^2$  hodnôt všetkých atribútov v množine príkladov

$$d_{L_2, \sigma}(\vec{p}, \vec{x}) = \sqrt{\sum_{i=1}^n \frac{1}{\sigma^2} (x_i - p_i)^2} = \frac{1}{\sigma} \cdot d_{L_2}(\vec{p}, \vec{x})$$

Teraz majú kategórie s väčším rozptylom väčšie receptívne pole. Ak povolíme nekonečne veľké váhy a dodefinujeme  $\infty \cdot 0 = 0$ , receptívne pole kategórie s nulovým rozptylom bude jediný bod (prototyp  $\vec{p}$ ). Takéto kategórie reprezentujú indivíduá.

Prírodným rozšírením predchádzajúceho prípadu je zaznamenávať rozptyly zvlášť pre každú dimenziu a použiť váhovanú euklidovskú vzdialenosť  $d_{L_2, \vec{w}}$  s vektorom

$$\text{váh } \vec{w} = \left( \frac{1}{\sigma_1^2}, \dots, \frac{1}{\sigma_n^2} \right), \text{ teda } d_{L_2, \vec{w}}(\vec{p}, \vec{x}) = \sqrt{\sum_{i=1}^n \frac{(x_i - p_i)^2}{\sigma_i^2}}.$$

Receptívne pole je v tomto prípade  $n$ -rozmerná hyperelipsa (elipsoid) so stredom  $\vec{p}$  a dĺžkami osí úmernými  $\sigma_i$ , pričom osi sú rovnobežné so súradnicovými osami priestoru  $A$ . Pre príslušnosť ku kategórii majú teda väčšiu dôležitosť atribúty, v hodnotách ktorých sa vstupné príklady veľmi nelíšia, ako atribúty s veľkým rozptylom. Ak sa v hodnote niektorého atribútu zhodujú všetky príklady (teda má nulový rozptyl), táto hodnota sa stane pre príslušnosť ku kategórii povinnou (akákoľvek iná dá nekonečnú vzdialenosť a teda nulovú aktivitu detektora). Receptívne pole bude degenerovaný elipsoid s príslušnou osou nulovej dĺžky.

Váhované dimenzie sa používajú aj v pôvodnej sémantike konceptuálnych priestorov [11] na vyjadrenie kontextovej dôležitosti jednotlivých dimenzií. Rozdiel je v tom, že v našom prístupe sú váhy lokálne pre každý detektor, čím sa posudzovanie podobnosti stáva asymetrickým, teda typický príklad jednej kategórie môže byť lepšou

inštanciou druhej kategórie, ako naopak<sup>4</sup> (vyjadrené matematicky, pre dva detektory  $r_{\bar{p}_1}$  a  $r_{\bar{p}_2}$  sa môžu hodnoty  $r_{\bar{p}_1}(\bar{p}_2)$  a  $r_{\bar{p}_2}(\bar{p}_1)$  líšiť).

Detektory zohľadňujúce dôležitosť jednotlivých atribútov na základe rozptylov sa nedokážu naučiť korelácie medzi atribútmi, zatiaľ čo ľudia túto úlohu úspešne zvládajú [27]. Napríklad pre odvodenie konceptu pre štvorec je nutné naučiť sa nielen povinnú hodnotu atribútu `pocet_vrcholov:4`, ale aj to, že atribúty `strana1` a `strana2` musia byť rovnaké. Toto dokáže zvládnuť detektor používajúci Mahalanobisovu metriku

$$d_{\Sigma^{-1}}^2(\bar{p}, \bar{x}) = (\bar{x} - \bar{p})^T \Sigma^{-1} (\bar{x} - \bar{p}),$$

kde  $\bar{x}$  a  $\bar{p}$  sú stĺpcové vektory a  $\Sigma^{-1}$  je inverzná matica ku kovariančnej matici  $\Sigma$  množiny príkladov konceptu.

$$\Sigma = (\sigma_{ij})_{i,j=1,n}, \text{ kde } \sigma_{ij} = \frac{1}{N-1} \sum_{k=1}^N (x_i^{(k)} - p_i)(x_j^{(k)} - p_j).$$

Keďže príklady kategórie nie sú obvykle k dispozícii všetky naraz, ale postupne, detektor bude používať na inkrementálny výpočet priemeru a kovariančnej matice iteratívne formuly. Nech  $N-1$  je počet doteraz videných príkladov a  $\bar{x}^{(N)}$  je nový ( $N$ -tý) príklad. Potom:

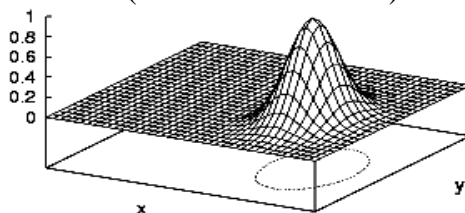
- pre  $N=1$  je  $\bar{p}^{(1)} = \bar{x}^{(1)}$  a  $\Sigma^{(1)} = (0)_{n \times n}$  alebo  $\sigma_0^2 I_n$ , kde  $\sigma_0^2$  je nejaký počiatočný odhad rozptylu,
- pre  $N > 1$  použijeme formuly

$$\bar{p}^{(N)} = \frac{N-1}{N} \bar{p}^{(N-1)} + \frac{1}{N} \bar{x}^{(N)}$$

$$\Sigma^{(N)} = \frac{N-2}{N-1} \Sigma^{(N-1)} + \frac{N}{(N-1)^2} (\bar{x}^{(N)} - \bar{p}^{(N)}) (\bar{x}^{(N)} - \bar{p}^{(N)})^T.$$

Množina atribútov určujúca priestor  $A$  sa tiež aktualizuje iteratívne prienikom s množinou atribútov nového príkladu. Ak sa takto odstránia niektoré dimenzie z  $A$ , príslušný riadok a stĺpec sa odstráni z kovariančnej matice, resp. z prototypového vektora.

Ak v rovnici (1) položíme  $k=0.5$ , dostaneme detektor, ktorého aktivačná funkcia má tvar mnohorozmerného (nenormalizovaného) Gausiánu (obr. 2).



**Obrázok 2.** Lokálne reagujúci detektor v 2-rozmernom vstupnom priestore, s aktivačnou funkciou v tvare mnohorozmerného Gausiánu. Receptívne pole detektora pre prah  $\theta = 0.1$  je zobrazené v rovine  $(x, y)$ .

Keďže  $\Sigma$  je štvorcová symetrická kladne semidefinitná matica, možno ju rozložiť na súčin matíc

$$\Sigma = UDU^T, \tag{2}$$

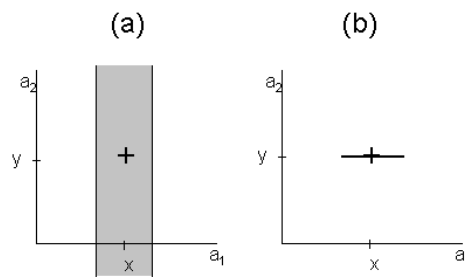
kde  $U$  je ortonormálna matica rotácie a  $D = \text{diag}(\lambda_1, \dots, \lambda_n)$  je diagonálna matica

<sup>4</sup> Ľudia tiež posudzujú podobnosti asymetricky, napr. tvrdia, že Tel Aviv sa podobá na New York viac, ako sa podobá New York na Tel Aviv [38].

vlastných hodnôt matice  $\Sigma$ , pričom  $\lambda_1 \geq \dots \geq \lambda_n$  [20]. Potom vzdialenosť  $d_{\Sigma^{-1}}^2(\vec{p}, \vec{x}) = (\vec{x} - \vec{p})^T \Sigma^{-1} (\vec{x} - \vec{p}) = (U(\vec{x} - \vec{p}))^T D^{-1} (U(\vec{x} - \vec{p}))$  je vlastne vzdialenosť medzi vektormi  $\vec{x}$  a  $\vec{p}$  v priestore zrotovanom podľa  $U$ , kde rozdiely v jednotlivých dimenziách sú škálované podľa  $D^{-1}$ . Keďže rotácia nemení tvar ani veľkosť receptívneho poľa, vidíme, že receptívnym poľom bude hyperelipsa, ktorej osi majú dĺžky úmerné  $\lambda_i$ , ale nemusia byť rovnobežné so súradnicovými osami priestoru  $A$ . Metriky  $d_{L_2}$ ,  $d_{L_2, \sigma}$ , resp.  $d_{L_2, \vec{w}}$  sú špeciálnymi prípadmi metriky  $d_{\Sigma^{-1}}$  s kovariančnými maticami  $I_n$ ,  $\sigma^2 I_n$ , resp.  $\text{diag}(\lambda_1, \dots, \lambda_n)$ .

Ak sa v množine príkladov vyskytuje atribút s nulovým rozptylom, matice  $\Sigma$  bude singularná, teda neinvertovateľná. V praxi sa namiesto inverznej matice  $\Sigma^{-1}$  používa Moore-Penrosova pseudoinverzná matice  $\Sigma^+$ , ktorej výpočet je založený na rozložení na singularne hodnoty (Singular Value Decomposition – SVD) matice  $\Sigma$ , ktoré má podobu (2). V prípade singularnej matice  $\Sigma$  je  $D = \text{diag}(\lambda_1, \dots, \lambda_k, 0, \dots, 0)$  pre nejaké  $k < n$ . Potom  $\Sigma^+ = U D^+ U^T$ , kde  $D^+ = \text{diag}(1/\lambda_1, \dots, 1/\lambda_k, 0, \dots, 0)$ .

Detektor používajúci  $\Sigma^+$  bude atribúty konštantné v celej vzorke ignorovať, namiesto aby ich považoval za povinné: ich váha bude nulová a zodpovedajúce osi degenerovanej hyperelipsy budú mať nekonečnú dĺžku (obr. 3). To je proti filozofii zachytávania regularít vo vstupnej vzorke. Napr. konštitutívnu vlastnosťou kategórie *trojuholník* je „mať 3 vrcholy“. Ak však budú mať všetky príklady vo vstupnej vzorke 3 vrcholy, kovariančná matice bude singularná a atribút *početVrcholov* sa bude ignorovať, lebo v  $D^+$  dostane nulovú váhu. Preto budeme namiesto  $D^+$  používať v SVD forme matice  $\Sigma$  štandardnú inverznú matice  $D^{-1} = \text{diag}(1/\lambda_1, \dots, 1/\lambda_k, \infty, \dots, \infty)$  s povolením nekonečných prvkov  $1/\lambda_i = \infty$  pre  $\lambda_i = 0$ . Výsledok funkcie vzdialenosti bude konečný, len ak bude mať otočený vektor  $U(\vec{x} - \vec{p})$  na osiach zodpovedajúcich nekonečným prvkom  $D^{-1}$  nulové súradnice.



**Obrázok 3.** Príklad kategórie s prototypom  $(x, y)$  a nulovým rozptylom v atribúte  $a_2$ . Detektor používajúci pseudoinverznú matice bude považovať atribút  $a_2$  za nepodstatný (a), zatiaľ čo detektor používajúci inverznú matice s nekonečnými prvkami bude považovať hodnotu  $a_2 = y$  za povinnú (b).

SVD dekompozícia je užitočná aj pre zovšeobecnenie výberom konštitutívnych črt kategórie. Napríklad pri odvodzovaní pojmu „okrúhly“ môže pozorovateľ zistiť, že príklady okrúhlych vecí majú v porovnaní s tvarom veľkú variabilitu farieb, veľkostí a iných vlastností, na základe čoho ich môže považovať za nepodstatné. V našom modeli tento typ zovšeobecnenia zodpovedá nájdeniu tých prvkov matice  $D^{-1}$ , ktoré sú v porovnaní s ostatnými veľmi malé, a nahradeniu ich nulami. Pretože  $1/\lambda_1 \leq \dots \leq 1/\lambda_n$ , môžeme nájsť najväčšie  $L$  také, že

$$\sum_{i=1}^L \frac{1}{\lambda_i} < b, \quad (3)$$

kde  $b$  je nejaký percentuálny prah, napr.  $b = 10\%$ . Nahradením prvých  $L$  prvkov  $D^{-1}$  nulami zanedbáme komponenty prispievajúce k celkovej vzdialenosti len veľmi málo. Toto je komplementárny proces k metóde *Principal Component Analysis* (PCA), v ktorej sa dátový priestor redukuje na komponenty s najväčšími rozptylmi [18].

Z rovnice (3) vyplýva, že ak matica  $D^{-1}$  obsahuje nekonečné prvky (zodpovedajúce úplne invariantným vlastnostiam vstupnej vzorky), všetky konečné sa zanedbajú. To je správne pre väčšinu konceptov, avšak nekonečné prvky môžu byť len umelou chybou – artefaktom spôsobeným niektorými konštantnými, ale úplne irelevantnými atribútmi (napr. ak hľadáme medzi predmetmi položenými na stole invariantné vlastnosti všetkých jabĺk a zaradíme medzi ne aj to, že všetky sú na stole). Odlíšenie figúry od pozadia je spoločným problémom všetkých indukčných algoritmov, ktoré sa učia iba z pozitívnych príkladov [14]. Vlastnosť zdieľaná všetkými pozitívnymi príkladmi by sa mala považovať za irelevantnú, ak ju zdieľajú aj negatívne príklady (ignorovaníu spoločných vlastností zodpovedá už zmienené použitie pseudoinverznej matice s  $D^+$ ). V našom modeli sa informácia o negatívnych príkladoch nevyužíva.

### 4.3 Znamienkové lokálne detektory

V každodennom živote ľudia často abstrahujú od numerických hodnôt a rovníc a používajú jednoduchší kvalitatívny kalkul založený na ordinálnych vzťahoch a invariantných znamienkových štruktúrach [32]. Znamienková štruktúra atribútov je konštitutívna aj pre mnohé vzťahové koncepty, napr. relácia  $vacsiAko(f_1, f_2)$  je založená na vlastnosti  $f_1.velkost > f_2.velkost$ , resp.  $sgn(f_1.velkost - f_2.velkost) = +1$ .

Na rozlišovanie znamienkovej štruktúry atribútov môžeme zaviesť znamienkový lokálne reagujúci detektor. Podpriestor  $A$  bude definovaný množinou tých atribútov, ktoré sa vo všetkých príkladoch vstupnej vzorky vyskytujú s rovnakým znamienkom a projekcia do podpriestoru  $A$  bude zložená s operátorom  $sgn$ . Spoločná znamienková štruktúra príkladov vstupnej vzorky bude zaznamenaná v prototypu  $\vec{p} = (p_1, \dots, p_n)$ , kde  $p_j = sgn(x_j^{(i)}) \quad \forall i = \overline{1, N}$ . Znamienková štruktúra sa zaznamená iba raz, po zhladnutí prvého príkladu. Neskoršie aktualizácie kritéria iba odstraňujú z  $A$  atribúty, ktoré majú v nových príkladoch iné znamienko ako to, čo je zaznamenané v prototypu. Detektor aplikovaný na nejaký vstup potom reaguje binárnou aktivitou: ak sa projekcia vstupu do  $A$  zhoduje s prototypom vo všetkých zložkách, vráti hodnotu 1, inak vráti hodnotu 0.<sup>5</sup>

### 4.4 Implementácia rozlišovacích kritérií pomocou detektorov

Rozlišovacie kritériá individuálnych objektov, napr.  $Bratislava(f)$ , vlastností objektov, napr.  $modrý(f)$ ,  $veľký(f)$  alebo  $zadĺžený(f)$ , a tried objektov, napr.  $strom(f)$ ,  $dieťa(f)$ , alebo

<sup>5</sup> Znamienkový detektor možno chápať ako špeciálny prípad detektora s metrikou  $d_{L_2, \vec{w}}$ , v ktorom sú hodnoty (znamienka) všetkých atribútov povinné (keďže majú vo vstupnom súbore nulový rozptyl).



*počítač(f)*,<sup>6</sup> môžeme priamo implementovať lokálne reagujúcim detektorom, pretože ich vstupným argumentom je jediný rámeček. Z formálneho hľadiska nie je medzi kritériami individuí, vlastností a tried žiadny rozdiel.

Rozlišovacie kritériá, ktoré majú viac vstupných argumentov (napr. vzťahy a zmeny) sa dajú tiež redukovať na lokálne reagujúce detektory transformáciou ich vstupu na jeden rámeček. Binárne vzťahy medzi objektmi, napr. *väčšíAko*( $f_1, f_2$ ) alebo *blízko*( $f_1, f_2$ ), sa dajú implementovať lokálne reagujúcim detektorom, ktorého vstupom bude rámeček  $f = \Delta(f_1, f_2)$  rozdielov hodnôt atribútov prítomných v oboch rámečkoch:  $A_f = A_{f_1} \cap A_{f_2}$  a  $f.a = f_1.a - f_2.a$  pre každé  $a \in A_f$ .

Na reprezentovanie vzťahov založených na magnitúde rozdielov (napr. *blízko*, *omnohoViac*, *takmerRovnaký*), je vhodné použiť detektor založený na kvantitatívnych štatistických vlastnostiach vstupného súboru, napr. priemere a kovarianciách. Na reprezentáciu kvalitatívnych vzťahov založených na znamienku rozdielov atribútov, resp. usporiadaní ich hodnôt je vhodné použiť znamienkový lokálny detektor (pozri časť 4.3) aplikovaný na vstup  $\Delta(f_1, f_2)$ .

Kritériá na rozlišovanie situácií zahŕňajúcich komplexnejšie vzťahy medzi viac ako dvoma objektmi alebo vlastnosti celej scény sa dajú vybudovať ako hierarchické siete lokálne reagujúcich detektorov. Kritériá na najnižšej úrovni hierarchie rozlišujú objekty, ktoré majú byť prítomné na scéne a ich požadované vzájomné vzťahy. Vektor výstupných aktivít týchto kritérií slúži ako vstup pre hierarchicky vyššie kritériá, ktoré môžu rozlišovať požadované hodnoty aktivít jednotlivých elementárnejších kritérií s rôznou dôležitosťou, detekovať vzájomné korelácie aktivít a podobne.

#### 4.5 Reprezentácia dynamiky prostredia: kritériá zmeny

Pre agenty operujúce v dynamických prostrediach je veľmi dôležitá schopnosť vnímať a reprezentovať zmeny. V spojitom svete dokážu už štvormesačné deti sledovať objekty pohľadom a osvojiť si princíp *kontinuity objektov* [21], ktorý je nutnou podmienkou pre vnímanie zmien vlastností objektov.

Keďže v našom modeli sa čas mení v diskretných krokoch, musíme zabezpečiť, aby agent nevnímal scény v jednotlivých krokoch ako nezávislé, ale ako množiny s korešpondenciami medzi perceptuálnymi rámečkami toho istého objektu v rôznych časoch. Z tohto dôvodu predefinujeme *scénu* – perceptuálny vstup agenta v čase  $t$  ako množinu dvojíc perceptuálnych rámečkov a ich podôb v predchádzajúcom časovom kroku:

$$S_t = \left\{ (f_1^{(t)}, f_1^{(t-1)}), \dots, (f_k^{(t)}, f_k^{(t-1)}) \right\}.$$

Ak sa objekt objavil na scéne až v kroku  $t$ , jeho percept  $f^{(t-1)}$  bude mať špeciálnu hodnotu  $\perp$ . Ak bol objekt na scéne v kroku  $t - 1$  a teraz zmizol,  $f^{(t)} = \perp$ .<sup>7</sup> Kritériá rozlišujúce zmenu vlastností objektu v čase sú realizované ako detektor aplikovaný na rámeček  $\Delta(f_1^{(t)}, f_1^{(t-1)})$ .

<sup>6</sup> Použité identifikátory rozlišovacích kritérií majú čitateľovi napovedať, aký koncept je reprezentovaný kritériom, nejedná sa však o zápis jazykového výrazu. Jazykové výrazy budeme v texte od označenia rozlišovacích kritérií odlišovať typom písma a úvodzovkami, napr. „veľký“ je slovo (a jeho významom môže byť rozlišovacie kritérium *veľký*).

<sup>7</sup> Na detekovaní  $\perp$  na príslušnej pozícii vstupného páru sú založené špeciálne kritériá zmeny *objavilSa* a *zmizol*.

## 4.6 Sémantika slovies

Zachytenie dynamiky prostredia kritériami zmeny je dôležitou súčasťou sémantiky slovies. Niektoré zmeny sa dajú zachytiť čisto kvalitatívnymi reláciami, napr. *narástol* môže byť reprezentované ako  $\text{sgn}(\Delta(f_1^{(t)}, f_1^{(t-1)}).vel'kost) = +1$ , iné môžu obsahovať predpoklady o typických veľkostiach zmien, napr. kritériá pohybu *plazíSa*, *kráča*, *beží* sa môžu líšiť v priemerných hodnotách  $\Delta(f_1^{(t)}, f_1^{(t-1)}).pozícia$ . Kritériá s nulovým znamienkom zmeny niektorých atribútov môžu reprezentovať stav, resp. trvanie vlastnosti, napr. *stojí*.

Komplexnejšia reprezentácia sémantiky slovies je prepojená s akciami a zahŕňa spôsoby ich vykonania, napr. reprezentácia *bežania* môže obsahovať prepojenie na nedeklaratívnu procedurálnu reprezentáciu invariantov motorického stereotypu spolu s rámcom variabilných parametrov akcie, napr. rýchlosti alebo natočenia končatín [1].

Keďže stelesnená reprezentácia významov vzniká v senzomotorických interakciách s prostredím, obsahuje aj situovanú kauzálnu znalosť o podmienkach úspechu akcie a jej možných dôsledkoch. Napríklad akcia *zdvihni* vykonaná rovnakým spôsobom (s rovnakou silou) môže viesť k rôznym výsledkom podľa toho, či zdvíhaným objektom je pingpongová loptička alebo 200 kilový kus nábytku. Agent preto využíva tie isté reprezentačné štruktúry na porozumenie významu slov, usudzovanie a plánovanie vlastných akcií v prostredí.

Vhodnou reprezentačnou štruktúrou pre zachytenie propozičných znalostí sú asociácie typu (*podmienky*, *akcia*, *dôsledky*), kde *podmienky* sú okolnosti (alebo objekt) akcie reprezentované kritériami situácií (alebo objektu), *akcia* kritériom reprezentujúcim typ a parametre vykonávanej akcie a *dôsledky* kritériami zmien, ktoré má akcia spôsobiť [33]. BDI<sup>8</sup> agent [6] môže reprezentovať svoje ciele ako kritériá želaných situácií a plánovať sekvencie akcií, ktoré by mohli viesť z aktuálnej situácie do cieľovej.

Takáto reprezentácia je v súlade s poznatkami o ľudskej organizácii reprezentácie znalostí. Podľa Gibsona [13] dokážeme v prostredí priamo vnímať možnosti akcií a manipulovania s objektmi: po zemi sa dá chodiť, dvere sa dajú otvoriť, jablko sa dá zjesť atď. Vnímateľné interakčné vlastnosti predmetov Gibson nazval *afordanciami*. V našom modeli im zodpovedajú dôsledky asociované s objektovými kategóriami. Podľa Tomasella [36] je mentálny lexikón dieťaťa spočiatku organizovaný okolo konkrétnych slovies s konkrétnymi scenármi vykonávania. Takéto štruktúry Tomasello nazval *slovesnými ostrovmi*. V našom modeli slovesné ostrovy tvoria objekty a dôsledky asociované s akciami.

## 5 Osvojovanie si kritérií

V predošlom texte sme opísali, ako môže agent odvodiť rozlišovacie kritérium nejakého konceptu indukciou z množiny jeho príkladov. V skutočnosti však agent vníma zmes rámcov, ktoré sú príkladmi rôznych kategórií a musí určiť, ktoré z existujúcich kritérií má pomocou perceptu aktualizovať (alebo či má vytvoriť nové kritérium).

Združovanie do kategórií môže prebiehať rôznymi spôsobmi, podľa dostupnosti spätnej väzby z prostredia. Ak agent dostáva iba vstupy z prostredia, ale nemá žiadnu spätnoväzobnú slučku, môže prirodzeným spôsobom vytvoriť kategórie tak, aby

---

<sup>8</sup> belief, desire, intention

odrážali štatistické rozdelenie vlastností prostredia. Prirodzená kategorizácia maximalizuje medzikategoriálne rozdiely a minimalizuje rozdiely v rámci kategórie. Na vytvorenie prirodzenej kategorizácie možno použiť klasterizačné algoritmy ako *Distributed Clustering Algorithm* [19], *1-nearest neighbour* a iné [8].

Prirodzená kategorizácia nemusí byť pre ciele systému postačujúca: napr. živočích, ktorý sa živí hubami (a chce prežiť), sa musí naučiť vnímať jemné odlišnosti vo vzhl'ade aj veľmi podobných húb, pokiaľ niektoré sú jedlé a iné jedovaté. Tu je vhodným typom kategorizácia na základe pragmatickej spätnej väzby. V článku [33] opisujeme model, v ktorom agent aktívne interagoval s objektmi v prostredí a zoskupoval ich do kategórií na základe ich afordancií. Agent sa pokúšal objekty zdvíhať alebo položiť a objekty tvorili jednu kategóriu, ak vykonanie niektorej akcie s nimi viedlo k rovnakému výsledku, napr. kategóriu objektov „príliš ťažkých na to, aby sa dali zdvihnúť“.

Tretím typom je kategorizácia s pomocou jazykových inštrukcií externého učiteľa. Externým učiteľom agenta môžu byť ostatné agenty, ktoré už zdieľajú spoločný jazyk, alebo človek. Napríklad dieťa si môže vyformovať kategóriu objektov, ktoré jeho rodič pomenúva tým istým menom – takto sa deti učia väčšinu nadradených kategórií ako zvierat, či nábytkov [39]. Kategórie osvojené podľa inštrukcií v prirodzenom jazyku sú kultúrne podmienené – známym príkladom je existencia rôznych kategorizácií farieb vo svetových jazykoch [4]. V ďalšej časti sa zameriame na formovanie významu pojmov pomocou jazyka. Model nemá zabudované žiadne kultúrne špecifické predpoklady a môže si rovnako dobre osvojovať kategórie z inštrukcií v umelom jazyku.<sup>9</sup> Proces formovania kritérií je takto spriahnutý s procesom osvojovania jazyka a my budeme skúmať ich vzájomnú interakciu.

## 6 Spriahnuté osvojovanie si jazyka a kategórií

Agent dostane v každom časovom kroku na vstupe scény, spolu s ňou jazykový výraz opisujúci niektoré jej aspekty a tiež tzv. fókus. Fókus je neverbálna referencia upriamujúca pozornosť agenta na aspekty scény pomenované jazykovým výrazom. Pri osvojovaní jazyka u detí zodpovedá ukazovaniu, sledovaniu pohľadu, či zdieľanej pozornosti dieťaťa s matkou [37]; v neskoršej fáze, keď už agent rozumie niektorým slovám, kontext referencie môže byť zúžený na základe známych častí výrazu, napr. výraz „veľký X“ môže zúžiť kontext pre X na veľké objekty na scéne. Formálne je fókus projekciou scény na  $k$ -ticu rámcov ( $k=1$  pre objektovú kategóriu,  $k=2$  pre vzťahovú kategóriu alebo kategóriu zmeny; v prípade kategórie zmeny je jeden z rámcov podobou objektu v predošlom časovom kroku). Výsledok projektívnej funkcie budeme nazývať *referentom* fókusu.

V najjednoduchšej učiacej situácii agent vníma scénu  $S_t$ , jednoslovný jazykový výraz  $w$  a fókus  $\phi$ . Agentove jazykové znalosti sú ukladané vo forme tabuľky asociácií slov (postupností písmen) a ich významov. Významom slova je rozlišovacie kritérium odvodené indukciou zo všetkých kontextov (referentov), s ktorými bolo slovo použité.

Hoci prirodzené jazyky obsahujú slová s viacerými význammi (homonymá) a viaceré slová s rovnakým významom (synonymá), v prípade spriahnutej akvizície

---

<sup>9</sup> Jediným zabudovaným predpokladom je konvexnosť prirodzených kategórií.

jazyka a konceptov bude veľmi užitočné začať bez synonymie a homonymie.<sup>10</sup> V skutočnosti práve toto sú hlavné princípy, na ktorých je v našom modeli založené osvojovanie:

1. *Žiadna pravá homonymia*. Každé slovo má iba jeden význam, aj keď je použité s viacerými referentmi (vtedy sa referenty pokladajú za inštancie/príklady tej istej kategórie).<sup>11</sup> V praxi sa tento princíp využíva tak, že ak agent už pozná nejaký význam slova  $w$  (má s ním asociované kritérium  $r$ ) a teraz je slovo  $w$  použité s novým referentom  $\phi(S_t)$ ,  $r$  je aktualizované pomocou  $\phi(S_t)$  (namiesto vytvorenia nového kritéria).
2. *Žiadna pravá synonymia*. Rôzne slová majú rôzny význam, aj keď zdieľajú ten istý referent (v tom prípade označujú rôzne aspekty referentu). V praxi sa tento princíp využíva tak, že ak agent počuje neznáme slovo  $w$  v kontexte referentu  $\phi(S_t)$ , vytvorí sa nové kritérium na základe prvého príkladu  $\phi(S_t)$  a v lexikóne sa asocjuje so slovom  $w$ .

**Príklad:** Uvažujme agenta, ktorý žije vo svete geometrických tvarov na mriežke  $50 \times 50$  so súradnicami  $(1, 1)$  v ľavom dolnom a  $(50, 50)$  v pravom hornom rohu. Ak agent vníma percept

```
f = { vertices: 3; size: 18; color: 3; posX: 1; posY:23 }
```

označený slovami „vľavo“, „veľký“, „trojuholník“, vytvorí si tri rozlišovacie kritériá, ktoré sú na začiatku identické a reprezentujú „fotku“ percipovaného objektu. Postupom času agent počuje tieto slová s ďalšími objektmi a kritériá sa začínajú diferencovať. Napr. počiatočná „fotka“ asociovaná so slovom „trojuholník“ sa aktualizuje pomocou perceptuálnych rámcov rôznych objektov najrôznejších farieb, pozícií, veľkostí a iných vlastností. Rámce všetkých objektov však budú mať spoločné to, že majú tri vrcholy. Atribúty, ktoré sa nevyskytujú v každom príklade sa z kritéria vylúčia, iné s veľkým rozptylom budú mať malú váhu. Rozhodujúcou (a nutnou) pre príslušnosť v kategórii asociovej so slovom „trojuholník“ sa stane vlastnosť mať tri vrcholy (keďže počet vrcholov má vo vzorke priemernú hodnotu 3 a rozptyl nula). Takisto slovo „vľavo“ bude agent počuť v kontexte najrôznejších objektov, ktoré všetky budú mať veľmi malú hodnotu atribútu *posX* atď. Čím viac je rôznych kontextov použitia slova, tým väčšia je pravdepodobnosť, že referenty sa budú líšiť v hodnotách atribútov nepodstatných pre význam slova. Avšak keby boli napr. všetky trojuholníky v agentovom svete veľké, vlastnosť byť veľkým by sa stala súčasťou významu slova „trojuholník“.

Princípy používané pri osvojovaní si významov jednotlivých slov sa dajú aplikovať aj na viacslovné menné frázy: predpokladá sa, že každé slovo vo fráze označuje iný aspekt referentu. V modeloch akvizície bez gramatiky je poradie slov vo fráze nepodstatné a indukcia významov z frázy „veľký trojuholník vľavo“ (alebo jej ľubovoľnej permutácie) má rovnaký efekt ako tri postupné indukcie z jednotlivých slov tak, ako to bolo opísané v príklade v predchádzajúcej časti.

---

<sup>10</sup> Deti osvojujúce si jazyk používajú *predpoklad vzájomnej výlučnosti* pomenovaní, teda predpoklad, že nové slová nemôžu označovať objekty, ktoré už majú nejaké iné meno [25]. Inou formuláciou je *princíp kontrastu* [7], ktorý predpokladá, že každej odlišnosti vo forme zodpovedá odlišnosť vo význame.

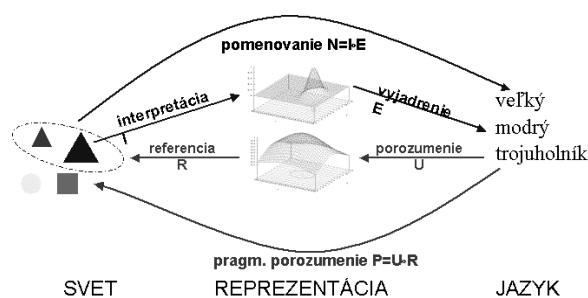
<sup>11</sup> Ak bolo slovo použité v zjavne úplne odlišných kontextoch (napr. referenty nemajú nič spoločné), agent by mohol detekovať homonymiu a asociovať so slovom viaceré kritériá; toto však v našom modeli nie je implementované.

V reálnych gramatických jazykoch môže rôzne poradie slov dávať celej fráze rôzne významy.<sup>12</sup> Takisto význam adjektív v mennej fráze môže byť kontextovo závislý od významu podstatného mena, napr. slovo „malý“ označuje rôzne veľkosti vo frázach „malý zajac“ a „malý slon“. Proces osvojovania, tak ako sme ho opísali vyššie, by jednoducho aktualizoval kritérium asociované so slovom „malý“ pomocou perceptov slona a zajaca, čím by sme dostali kritérium s priemernou veľkosťou medzi zajacom a slonom a veľmi veľkým rozptylom. Rozumnejšie by bolo vytvoriť pre každý menný kontext iné kritérium pre „malý“ a aktualizovať každé zvlášť, to by si však už od agenta vyžadovalo istú znalosť gramatickej štruktúry osvojovaného jazyka (na rozlíšenie, ktorá časť mennej frázy je podstatné meno a ktoré sú rozvíjajúce adjektíva).

Bez znalosti pravidiel gramatickej kompozície agent nemôže pochopiť ani význam fráz ako „mačka na horúcej plechovej streche“, aj keby si jednotlivo osvojil správne významy slov „mačka“, „na“, „horúca“, „plechová“, „strecha“.<sup>13</sup>

## 7 Pragmatická úroveň

Opísali sme mechanizmy, akými si agent môže interakciou s prostredím vybudovať rozlišovacie kritériá reprezentujúce rôzne koncepty a ukotviť v nich významy jazykových výrazov. To, ako bude agent naučené kritériá a osvojené jazykové výrazy ďalej používať, závisí od konkrétnej aplikácie. Pre komunikáciu je však užitočné definovať niekoľko všeobecných pragmatických funkcií zabezpečujúcich zobrazovanie medzi prostredím, repertoárom významov agenta a jazykom (obr. 4). Význam jednotlivých funkcií by mal byť z obrázku zrejmy, ich formálne definície čitateľ nájde v [34]. Dodáme len, že manipuláciou s rozhodovacím prahom  $\theta$  určujúcim receptívne polia kritérií dosiahneme porozumenie kontrastívnemu použitiu slov. Ak napríklad agent počuje slovo „veľký“ so scénou obsahujúcou iba malé objekty, znížením prahu kritérií mu funkcia referencie vráti objekty, na ktoré sa popis najlepšie hodí, teda najväčšie z malých objektov prítomných na scéne.



**Obrázok 4.** Pragmatické funkcie zabezpečujú prepojenie prostredia, vnútorných reprezentácií a jazyka.

<sup>12</sup> Deti pri osvojovaní jazyka využívajú nielen poradie slov, ale aj ďalšie gramatické informácie na obmedzenie možných významov neznámych častí frázy (teória *syntaktického bootstrappingu* [24]). Obohatenie nášho modelu o citlivosť na gramatickú štruktúru je dôležitou témou pre budúci výskum.

<sup>13</sup> V slovenčine sú ďalšou komplikáciou, od ktorej abstrahujeme, rôzne gramatické tvary toho istého slova, napr. „horúca“ a „horúcej“. Avšak ani v angličtine, kde sú tvary rovnaké, na určenie významu frázy „cat on a hot tin roof“ nestačí poznať významy jednotlivých slov (bez znalosti ich vzťahov daných gramatikou by sme význam uvedenej frázy nevedeli odlíšiť napr. od významu fráz „hot tin cat on a roof“ alebo „tin roof on a hot cat“).

## 8 Záver

Rozlišovacie kritériá poskytujú jednotný reprezentačný formalizmus pre koncepty rôznych typov a ich osvojovanie. Proces formovania kritérií je ukotvený v percepcii a tesne spriahnutý s jazykom, čím sa navrhnutý formalizmus stáva dobrým kandidátom na kognitívnu sémantiku jazyka. Keďže kritériá sú citlivé na rozptyl a korelácie hodnôt atribútov vrámci kategórie, sú flexibilnejšou a efektívnejšou reprezentáciou ako „čisté“<sup>14</sup> prototypy v konceptuálnom priestore, čo sme aj experimentálne ukázali v našom implementovanom modeli interakčného osvojovania kritérií [33]. Rozlišovacie kritériá sú implementované pomocou lokálne reagujúcich detektorov a ich hierarchických sietí, čo je biologicky plauzibilné.

Ak by sme chceli pomocou formalizmu rozlišovacích kritérií realistickejšie reprezentovať sémantiku slovies, bolo by ho treba rozšíriť o štruktúry reprezentujúce perspektívu, spôsob, roly, aspekt a ďalšie parametre slovies. Indukčný proces v súčasnej podobe nie je robustný voči chybám a šumu (jeden nesprávny príklad môže nevratne ovplyvniť množinu atribútov kritéria), preto by ho bolo rozumné nahradiť stochastickou indukciou. V budúcnosti tiež treba dopracovať detaily kritérií situácií a ich indukciu a rozšíriť jazykovú úroveň o viacslovné frázy a nejakú formu gramatiky.

V každom prípade považujeme navrhnutú sémantiku za solídnu základňu pre kognitívne modely osvojovania jazyka, BDI architektúry a aplikácie v otvorených umelých prostrediach, akým je napr. world wide web.

**PodĎakovanie:** Tento výskum bol financovaný z grantu č. APVV-20P04805 Agentúry na podporu výskumu a vývoja a z grantu VEGA 1/3105/06.

## Literatúra

- [1] Bailey, D., Feldman, J., Narayanan, S., Lakoff, G.: Modeling Embodied Lexical Development. In *Proceedings of the 19th Cognitive Science Society Conference* (pp. 19–24) 1997.
- [2] Balkenius, C.: Are There Dimensions in the Brain? In *Spinning Ideas, Electronic Essays Dedicated to Peter Gärdenfors on his Fiftieth Birthday*. 1999.  
<http://www.lucs.lu.se/spinning/categories/cognitive/Balkenius/Balkenius.pdf>
- [3] Barsalou, L. W.: Perceptual Symbols Systems. *Behavioral and Brain Sciences* **22** (1999) 577–660.
- [4] Berlin, B., Kay, P.: *Basic Color Terms: Their Universality and Evolution*. Berkeley, CA: University of California Press 1969.
- [5] Bloom, P.: *How Children Learn the Meanings of Words*. Cambridge, MA: MIT Press 2000.
- [6] Bratman, M.: *Intention, Plans, and Practical Reason*. Cambridge, MA: Harvard University Press 1987.
- [7] Clark, E. V.: The Principle of Contrast: A Constraint on Language Acquisition. In MacWhinney, B. (Ed.): *Mechanisms of Language Acquisition*. London: Erlbaum 1987.
- [8] Everitt, B. S., Landau, S., Leese, M.: *Cluster Analysis*. London: Arnold 2001.
- [9] Feldman, J.: *From Molecule to Metaphor: A Neural Theory of Language*. Cambridge, MA: MIT Press 2006.

---

<sup>14</sup> S euklidovskou vzdialenosťou.

- [10] Fodor, J.: *Representations: Philosophical Essays on the Foundations of Cognitive Science*. Cambridge, MA: MIT Press 1981.
- [11] Gärdenfors, P.: *Conceptual Spaces*. Cambridge, MA: MIT Press 2000.
- [12] Geuder, W., Weisgerber, M.: Verbs in Conceptual Space. In Katz, G., Reinhard, S., Reuter, P. (Eds.): *Sinn und Bedeutung 6, Proceedings of the Sixth Meeting of the Gesellschaft für Semantik*, volume 1, Osnabrück: Publications of the Institute of Cognitive Science 2002.
- [13] Gibson, J. J.: *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin 1979.
- [14] Gold, E. M.: Language Identification in the Limit. *Information and Control*, **10(5)** (1967) 447–474.
- [15] Gopnik, A., Meltzoff, A. N., Kuhl, P. K.: *The Scientist in the Crib: What Early Learning Tells Us About the Mind*. New York: HarperCollins 2000.
- [16] Harnad, S.: The Symbol Grounding Problem. *Physica D* **42** (1990) 335–346.
- [17] Hassoun, M. H.: *Fundamentals of Artificial Neural Networks*. Cambridge, MA: MIT Press 1995.
- [18] Haykin, S.: *Neural Networks: A Comprehensive Foundation (Second ed.)*. Upper Saddle River, NJ: Prentice Hall 1999.
- [19] Hulth, N., Grenholm, P.: A Distributed Clustering Algorithm. *Lund University Cognitive Studies* **74** (1998).
- [20] Ientilucci, E. J.: Using the Singular Value Decomposition. 2003.  
<http://www.cis.rit.edu/~ejipci/Reports/svd.pdf>
- [21] Johnson, S. P., Amso, D., Slemmer, J. A.: Development of Object Concepts in Infancy: Evidence for Early Learning in an Eye Tracking Paradigm. *Proceedings of the National Academy of Sciences USA* **100** (2003) 10568–10573.
- [22] Kirby, S., Hurford, J.: The Emergence of Linguistic Structure: an Overview of the Iterated Learning Model. In Parisi, D., Cangelosi, A. (Eds.): *Computational Approaches to the Evolution of Language and Communication*. Berlin: Springer-Verlag 2001.
- [23] Langacker, R.: *Concept, Image and Symbol: The Cognitive Basis of Grammar*. Mouton de Gruyter 1991.
- [24] Lidz, J., Gleitman, H., Gleitman, L. R.: Kidz in the 'Hood: Syntactic Bootstrapping and the Mental Lexicon. In Hall, D. G., Waxman, S. R. (Eds.): *Weaving a Lexicon*. Cambridge, MA: MIT Press 2004.
- [25] Markman, E.: Constraints on Word Learning: Speculations About Their Origins and Domain Specificity. In Gunnar, M. R., Maratsos, M. (Eds.): *Modularity and Constraints in Language and Cognition*. Hillsdale, NJ: Lawrence Erlbaum Associates 1992.
- [26] Martin, J. H.: Coding and Processing of Sensory Information. In Kandel, E. R., Schwartz, J. H., Jessel, T. M. (Eds.): *Principles of Neural Science* (pp. 329–340). New York: Elsevier 1991.
- [27] Medin, D. L., Altom, M. W., Edelson, S. M., Freko, D.: Correlated Symptoms and Simulated Medical Classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **8** (1982) 37–50.
- [28] Piaget, J., Inhelder, B.: *La Psychologie de L'enfant*. Paris: PUF 1966.
- [29] Spelke, E. S.: Principles of Object Perception. *Cognitive Science* **14** (1990) 29–56.
- [30] Steels, L.: Language as a Complex Adaptive System. In Schoenauer, M. (Ed.): *Proceedings of PPSN-VI*, Berlin: Springer-Verlag 2000.

- [31] Šefránek, J.: Kognícia bez mentálnych procesov. In Beňušková, Ľ. et al. (Eds.): *Kognitívne vedy*. Bratislava: Kaligram 2002.
- [32] Takáč, M.: *Kvalitatívne modelovanie a simulácia*. Bratislava: Vydavateľstvo UK 2003.
- [33] 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*. Trieste: Edizione Goliardiche 2006.
- [34] Takáč, M.: Kognitívna sémantika rozlišovacích kritérií. In J. Kelemen, V. Kvasnička (Eds.): *Kognice a umělý život VI*. Opava: Slezská univerzita 2006.
- [35] Talmy, L.: *Toward a Cognitive Semantics*. Cambridge, MA: MIT Press 2000.
- [36] Tomasello, M.: *First Verbs: A Case Study of Early Grammatical Development*. Cambridge: CUP 1992.
- [37] Tomasello, M., Farrar, J.: Joint Attention and Early Language. *Child Development* **57** (1986) 1454–1463.
- [38] Tversky, A.: Features of Similarity. *Psychological Review* **84** (4) (1977) 327–352.
- [39] Waxman, S. R., Braun, I.: Consistent (but not Variable) Names as Invitations to Form Object Categories: New Evidence from 12-Month-Old Infants. *Cognition* **95** (2005) B59-B68.
- [40] Whorf, B. L.: *Language, Thought and Reality: Selected writings of Benjamin Lee Whorf*. Cambridge, MA: MIT Press 1956.