

# Konštrukcia významov a jej dynamika v procese iterovaného učenia

Martin Takáč

Katedra aplikovanej informatiky, FMFI UK  
Mlynská dolina, 842 48 Bratislava  
takac@ii.fmph.uniba.sk

## Abstrakt

V minulosti sme modelovali autonómnu konštrukciu významov na základe individuálnej interakcie agenta s prostredím (kognitívny konštruktivizmus). V tomto príspevku sa zameriame na vplyv sociálnej/jazykovej interakcie na formovanie významových kategórií (sociálny konštruktivizmus a Sapir-Whorfova téza) a na dynamiku (premeny) významov v procese iterovaného medzigeneračného prenosu. Príspevok prináša výsledky počítačových simulácií a ich analýzu.

## 1 Úvod

Kognitivistický prístup k sémantike (jazyka) sa zakladá na predpoklade, že významy nie sú entity objektívne existujúce vo svete, ale sú individuálne ukotvené v subjektívnej skúsenosti. Často sa na ne zároveň nahliada ako na vnútorné, či „mentálne“ reprezentácie [9]. Takýto prístup sa musí vysporiadať s dvoma problémami.

Prvým je *problém ukotvenia symbolov* [12] spočívajúci v tom, že vnútorné reprezentácie samy o sebe nemajú inherentný význam. V tomto prípade treba odpovedať na otázku, ako vnútorné reprezentácie vznikajú a aký je ich vzťah k reálnemu svetu. My sa ňou budeme zaoberať v časti 2.

Ak sú významy individuálne, vzniká druhý problém: ako sa komunikujúci jedinci môžu navzájom dorozumieť, ak každý z nich má vlastné (potenciálne rôzne) vnútorné reprezentácie? Zrejme sa treba pýtať, aké sú ohraničujúce podmienky procesu tvorby významov spôsobujúce nie príliš veľkú odlišnosť individuálnych reprezentácií, a aké sú mechanizmy koordinácie takýchto reprezentácií v spoločenstve. Tento problém sa nazýva *externým (sociálnym) ukotvovaním symbolov* [3]. Vplyv sociálnych (najmä jazykových) interakcií na podobu významov je hlavnou témou tohto článku. V časti 3 rozoberieme niektoré teoretické aspekty problému, v nasledujúcich častiach predstavíme výpočtový model konštrukcie významov na základe verbálnej inštrukcie a výsledky experimentov s týmto modelom.

## 2 Individuálne významy

V tejto časti sa budeme zaoberať individuálnymi význammi jedincov vo vzťahu k prostrediu, v ktorom sa nachádzajú. Jedince, ktoré sú vnorené v nejakom prostredí a interagujú s ním pomocou svojich senzorov a aktuátorov za účelom dosahovania svojich cieľov budeme v ďalšom texte nazývať *agenty* [13]. V najelementárnejšom zmysle agenty pripisujú významy prostrediu tým, že v ňom rozpoznávajú tie aspekty, ktoré sú relevantné ich cieľom [18] a diferencovane na ne reagujú. V tomto zmysle sú významy predverbálne: stelesnenú znalosť pochádzajúcu z vnímania a konania v prostredí môžeme nájsť u živých organizmov ešte pred objavením sa jazyka, či už vo fylogenéze [15] alebo ontogenéze [21, 25].

Gärdenfors [7] rozlišuje medzi tzv. *vyvolanými* a *oddelenými* reprezentáciami. Vyvolaná reprezentácia je vždy spustená niečím, čo sa aktuálne nachádza v prítomnej situácii. Každý organizmus, ktorý diferencovane reaguje na stavy svojho okolia (vrátane svojho vlastného stavu), napr. konzumuje objekty rozpoznané ako potravu a vyhýba sa objektom rozpoznávaným ako predátory, vykonáva kategorizáciu a disponuje vyvolanými reprezentáciami príslušných kategórií. Príslušné správanie sa však spustí (vyvolá) vždy iba v prítomnosti reprezentovaného podnetu.

Oddelená reprezentácia môže ovplyvniť správanie systému aj v neprítomnosti svojho referentu. Napríklad šimpanz, ktorý nevie dočiahnuť banán, odíde hľadať palicu (ktorú naokolo nevidí) a keď ju nájde, vráti sa a strhne pomocou nej banán, musí mať oddelenú reprezentáciu palice aj možnosť jej použitia.

Oddelené reprezentácie sa vo fylogenéze objavujú približne spolu s vývojom neokortexu [8]; v ontogenetickom vývine korešpondujú so sformovaním *trvalého predmetu* [21]. Disponovanie oddelenými reprezentáciami je podmienkou vyšších kognitívnych funkcií ako sú plánovanie, predstieranie, sebauvedomenie a jazyk [7].

Vyvolané reprezentácie pozorovateľné ako automatické behaviorálne reakcie sú vrodené a vyvinuli sa fylogeneticky. Nás budú v tomto článku zaujímať reprezentácie získané v priebehu individuálnej

ontogenézy. Náš prístup je blízky *kognitívnemu konštruktivismu* [20]: každý agent si aktívne konštruuje svoje vlastné významy pomocou kognitívnych procesov založených na minulých skúsenostiach a na interakciách so svetom. Základné konkrétne pojmy sú priamo ukotvené v telesnej skúsenosti (senzomotorických interakciách), abstraktné pojmy vznikajú ako nadstavba nad konkrétnymi pomocou metaforických a metonymických zobrazení [16]. Empirických dôkazov o stelesnenej povahe našej kognície je neúrekom [11]. Významy sú ukotvené v percepcii a akcii [1, 19] a neurálne aktivity merateľné pri porozumení jazykovým výpovediam korelujú s aktivitami vyskytujúcimi sa pri pozorovaní alebo priamej fyzickej interakcii s reprezentovanými obsahmi výpovedí [6, 22, 24]. V tomto zmysle sú vnútorné reprezentácie štruktúrne spriahnuté s okolitým svetom [17, 32], čo im dáva inherentný význam. Nie je pritom dôležité, či interakcie prebiehajú v skutočnom svete, alebo vo virtuálnom. Podstatná je existencia interakčnej spätoväzobnej slučky medzi agentom a prostredím, pomocou ktorej si agent konštruuje vlastné významy.

Ukotvenie významov na základe interakcie s reálnym či simulovaným prostredím je súčasťou mnohých výpočtových modelov vzniku a osvojovania jazyka, napr. [2, 27, 32], vrátane našej predchádzajúcej práce [29].

### 3 Sociálne koordinované významy

Ľudský jazyk je komunikačný systém umožňujúci externalizáciu a prenos oddelených reprezentácií. To znamená, že umožňuje komunikovanie o veciach mimo tu a teraz, dokonca aj o veciach, ktoré nie sú fyzicky možné (napr. ježibaba, či štvorcové trojuholníky). Evolučným tlakom, ktorý formoval jazyk, bola podľa Gärdenforsa potreba koordinovať spoluprácu spoločenstva pri dosahovaní vzdialených cieľov [10]. Takáto spolupráca si vyžaduje sociálne zdieľané významy. Ak sú významy konštruované individuálne, čo zabezpečí vzájomné dorozumenie členov spoločenstva?

Po prvé, významy nesmú byť úplne odlišné, inak by úspešná komunikácia nebola možná [28]. Podobnosť významov je zabezpečená tým, že jedinci zdieľajú spoločné prostredie, rovnaké učiace mechanizmy a podobné interakčné skúsenosti.

Po druhé, komunikačná situácia inherentne obsahuje významovú nejednoznačnosť [23], ktorú účastníci komunikácie prekonávajú inferenciou z pragmatického kontextu.

Po tretie, individuálne významy sa vzájomne koordinujú v procese samoorganizácie. Globálne koherentný jazyk emerguje ako výsledok lokálnych interakcií používateľov jazyka na princípe pozitívnej spätnej väzby: používatelia jazyka adaptujú svoje

lingvistické správanie na základe úspešnosti v predchádzajúcich interakciách [26]. Spoločnými resp. sociálnymi význammi sú (pohyblivé) ekvilibriá tohto komplexného dynamického systému [9].

Pri koordinovaní jazyka z iniciálneho stavu bez akejkoľvek spoločnej slovnej zásoby<sup>1</sup> je kľúčové najprv ustanoviť zdieľané významy tzv. referenčných výrazov, ako sú vlastné a všeobecné podstatné mená a prídavné mená [10]. To sa deje pomocou pomenovávacích rečových aktov, pričom významnú úlohu majú neverbálne pomôcky ako ukazovanie, sledovanie pohľadu, či zdieľaná pozornosť [31].

Pri osvojovaní si jazyka u detí má pomenovávanie dospelými významný vplyv na ich konceptuálnu organizáciu a facilituje aj objavovanie nových konceptov. Používanie rôznych pomenovaní pre rôzne objekty motivuje 12-mesačné deti k hľadaniu *odlišností* a podporuje individuáciu, zatiaľ čo pomenovávanie rôznych objektov rovnakým menom motivuje k hľadaniu *podobností* a podporuje vytváranie kategórií [36]. Tieto empirické zistenia podporujú (opatrnějšíu verziu) Sapir-Whorfovej tézy [37], že jazyk má vplyv na podobu konceptuálnej reprezentácie. Filozoficky je tento prístup blízky sociálnemu konštruktivismu [34]: významy sú naďalej konštruované individuálne, avšak tento proces je ovplyvnený sociálnym kontextom. Túto hypotézu sme sa rozhodli testovať výpočtovým modelom, ktorému sú venované nasledujúce časti.

### 4 Konštrukcia sociálne relevantných významov

Experiment, ktorý opíšeme, je zameraný na skúmanie vplyvu verbálnej inštrukcie (pomenovania) na proces formovania kategórií. Model pozostáva z dvoch agentov situovaných v simulovanom prostredí. Prvý agent (učiteľ) pomenováva rôzne aspekty súčasnej situácie, druhý agent (žiak) odvodzuje významy učiteľových slov na základe podobností medzi referentmi rovnakých slov v rôznych situáciách.

Na reprezentáciu významov používame formalizmus rozlišovacích kritérií, ktorý sme podrobne opísali v [30], preto sa už jeho detailom v tomto článku nebudeme venovať. Rozlišovacími kritériami je možné reprezentovať objekty, ich vlastnosti, vzájomné vzťahy, zmeny v čase, situácie a udalosti. Všetky kritériá sú osvojiteľné resp. skonštruovateľné z príkladov.

Odvodzovanie významov v modeli s jazykovou inštrukciou je založené na predpokladoch *žiadnej pravej homonymie* a *žiadnej pravej synonymie* [30]. Prvý predpoklad znamená, že rôzne slová majú rôzne

<sup>1</sup> Angličtina má pre takúto konštrukciu „z ničoho“ termín *bootstrapping*.

významy, aj keby zdieľali rovnaký referent (v tomto prípade každé z nich opisuje iný aspekt referentu).<sup>2</sup> Druhý predpoklad je kľúčový pre postupnú elimináciu viacznačnosti významov: všetky referenty toho istého slova v rôznych situáciách sa považujú za príklady konceptu pomenovaného týmto slovom a z nich sa indukuje význam pomocou štatistických metód. V číms viacerych rôznych kontextoch je slovo použité, tým je väčšia pravdepodobnosť, že vlastnosti irelevantné pre význam slova sa budú vyznačovať veľkou variabilitou.

#### 4.1 Model

Simulované prostredie v experimente pozostávalo z plošných geometrických tvarov charakterizovaných piatimi atribútmi: počet vrcholov ( $v \in [2,5]$ ), súradnice stredu, resp. ťažiska ( $x, y \in [0,50]$ ) a veľkosť opísaného obdĺžnika ( $sX, sY \in [0,25]$ ). Počítačové hodnoty atribútov boli rovnomerne náhodne generované celé čísla z príslušných intervalov. Prostredie sa dynamicky menilo: v každom časovom kroku mohli náhodne vybrané objekty zmeniť veľkosť alebo polohu, alebo mohli byť z prostredia odstránené. Takisto mohli pribudnúť nové objekty (obvykle sa na scéne naraz vyskytovali 2-4 objekty). V jednom časovom kroku mohlo prebehnúť aj viac zmien súčasne. Agent – učiteľ bol vybavený preddefinovanou ontológiou a lexikónom, pomocou ktorých komentoval súčasnú situáciu (vrátane prebiehajúcich zmien) žiakovi. Učiteľov lexikón obsahoval 2 podstatné mená, 3 prídavné mená a 2 slovesá (Tabuľka 1).

Experiment bežal 5000 učiacich epoch (časových krokov). Žiak používal na reprezentáciu významov rozlišovacie kritériá schopné zachytávať kovariancie atribútov (detektory s Mahalanobisovou metrikou, pozri [30]) s rozhodovacím prahom 0,1 a zanedbanie komponentov v SVD rozklade na hladine 10% [30].

Tab. 1. Preddefinovaný lexikón a významy učiteľa.

Slovo	Význam
štvorec	$v = 4 \wedge sX = sY$
trojuholník	$v = 3$
veľký	$sX > 15 \wedge sY > 15$
štíhly	$sX < 0.2 sY$
malý	$sX < 10 \wedge sY < 10$
narástol	$sX(t) > sX(t-1) \wedge sY(t) > sY(t-1)$
zmenšil sa	$sX(t) < sX(t-1) \wedge sY(t) < sY(t-1)$

<sup>2</sup> Tento predpoklad zodpovedá princípu kontrastu [5] známemu zo štúdia akvizície jazyka u detí.

#### 4.2 Výsledky

Pre účely kvantitatívneho vyhodnotenia vernosti prenosu významov situáciu v každom časovom kroku komentoval aj žiak a jeho opis scény bol porovnaný s učiteľovým. Merali sme *korektnosť* a *úplnosť* žiakovho opisu a celkovú *podobnosť opisov* ako aritmetický priemer korektnosti a úplnosti.

Korektnosť žiakovho opisu je definovaná ako  $1-w/L$ , kde  $w$  je počet nesprávnych slov v žiakovom opise a  $L$  je celkový počet slov v žiakovom opise (slovo, ktoré žiak použil na opis nejakého referentu sa považuje za nesprávne, ak v učiteľovom opise toho istého referentu chýba).

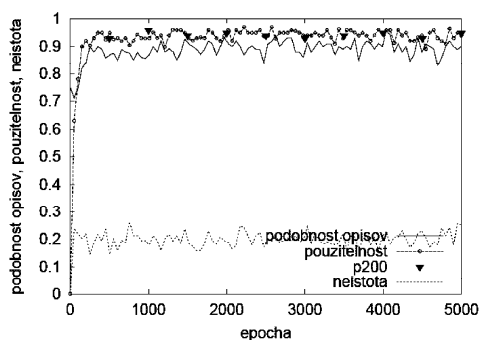
Úplnosť žiakovho opisu je definovaná ako  $1-m/T$ , kde  $m$  je počet učiteľových slov chýbajúcich v žiakovom opise referentov, s ktorými ich použil učiteľ.  $T$  je celková dĺžka (počet slov) učiteľovho opisu.

Tiež sme merali pragmatickú *použitelnosť* žiakovej ontológie v tzv. hádacích hrách. V každom časovom kroku učiteľ slovami opísal náhodne vybraný referent a žiak hádal, ktorý to bol. Žiakovou odpoveďou bola množina  $O_L$  referentov, ktoré podľa jeho porozumenia vyhovujú učiteľovmu opisu. Ak množina  $O_L$  neobsahovala referent, ktorý mal na mysli učiteľ, *použitelnosť* bola nula. Inak sa množina  $O_L$  porovnávala s množinou  $O_T$  referentov, ktoré by pod daným opisom pochopil sám učiteľ (ten nemusel referent opísať jednoznačne, rovnaký opis mohol sedieť na viac objektov). Použitelnosť sa počítala ako  $1/(1+|O_L-O_T|)$ , čo znamená, že v prípade správneho hádania bola hodnota použiteľnosti znížená počtom nadbytočných referentov, ktoré nemohol mať učiteľ na mysli (ak neboli žiadne nadbytočné, použiteľnosť bola rovná 1).

Keďže učiteľove opisy nemuseli byť jednoznačné, merali sme inherentnú *neistotu* učiteľovho opisu definovanú ako  $1-1/|O_T|$ . Ak by napríklad učiteľ opísal zvolený objekt slovom „trojuholník“ a na scéne by bol ešte jeden ďalší objekt, ktoré by mohol kategorizovať a opísať rovnako, neistota učiteľovho opisu by bola 50%. V prípade jediného možného referentu by bola neistota nulová.

Agenty hrali hádacie hry priebežne a ešte odohrali ďalších 200 hier po každých 500 učiacich epochách. Hádacie hry sa hrali iba za účelom kvantitatívneho vyhodnocovania experimentu a nemali žiaden vplyv na učiaci proces.

Výsledky simulácií na obr. 1 ukazujú, že učenie na základe medzi-situačných podobností je rýchle a spoľahlivé: žiak si veľmi skoro osvojí ontológiu a lexikón dostatočnej kvality, ktoré ostanú stabilné počas ďalšieho behu simulácie.



**Obr. 1.** Kvalita lexikónu získaného v rámci jednej generácie. Miera „p200“ je priemer z použiteľností v 200 hádacích hrách odohraných po každých 500 epochách. Každý údaj v grafe bol priemerovaný v časovom okne posledných 30 krokov. Uvedené výsledky sú priemerom z 10 náhodných simulačných behov.

## 5 Dynamika významov v iterovanom medzigeneračnom prenose

Ukázali sme, ako si učiaci sa agent môže na základe pomenovaní skonštruovať významy dostatočne podobné učiteľovým. Teraz nás bude zaujímať, či významy skonštruované týmto spôsobom ostanú stabilné, alebo sa menia, ak necháme osvojovací proces iterovať viacgeneračne. Predchádzajúci experiment zasadíme do rámca modelu iterovaného učenia (ILM) opísaného v [14].

Model iterovaného učenia zahŕňajúci vertikálny kultúrny prenos jazyka medzi generáciami bol pôvodne navrhnutý na skúmanie emergencie kompozičných štruktúr (gramatiky). Jazyk sa v rámci ILM vyvíja opakovanými transformáciami medzi osobnou jazykovou kompetenciou a externalizovanými výpoveďami pomocou akvizície a produkcie. Učiaci sa agent (žiak) si vytvára vnútornú reprezentáciu jazyka na základe externého vstupu – výpovedí svojho učiteľa. Neskôr je učiteľ odstránený a žiak sa stane v ďalšej generácii učiteľom pre nového žiaka atď.

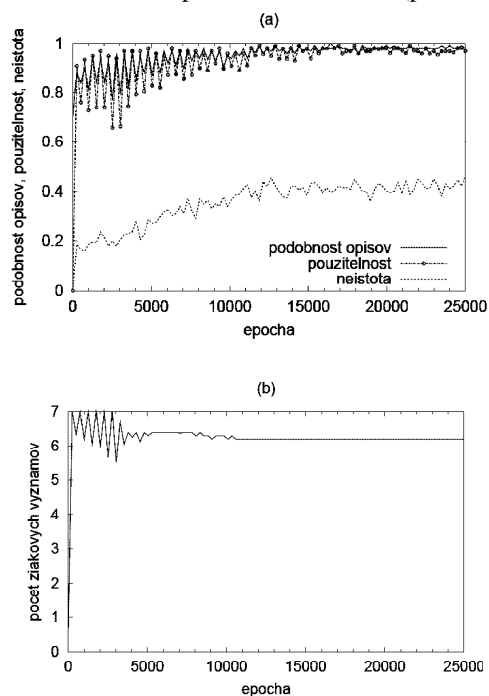
### 5.1 Model

V prvej generácii experiment prebiehal identickým spôsobom, ako sme opísali v predchádzajúcej časti. Po určitom počte epoch bol učiteľ s preddefinovanou ontológiou odstránený a jeho žiak, ktorý si ontológiu skonštruoval sám, sa stal učiteľom pre nového žiaka s prázdnu ontológiou a lexikónom. Tento proces sme nechali iterovať 50 generácií.

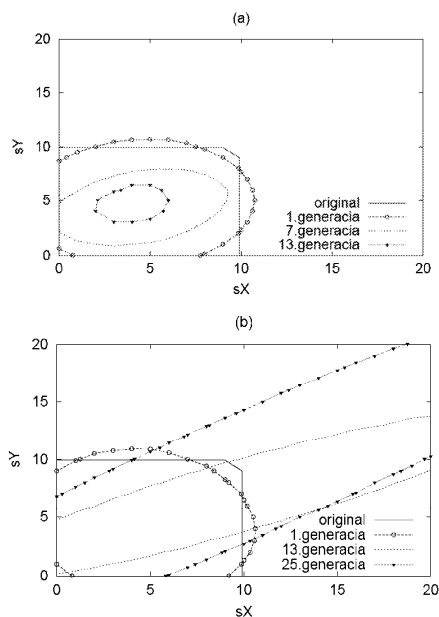
## 5.2 Výsledky

Na obr. 2 vidíme výsledky experimentu, ktorý bežal 50 iterácií (každá iterácia mala 500 krokov). Ako je vidno, kvalita lexikónu sa po každom poklese spôsobenom výmenou generácií rýchlo obnoví.

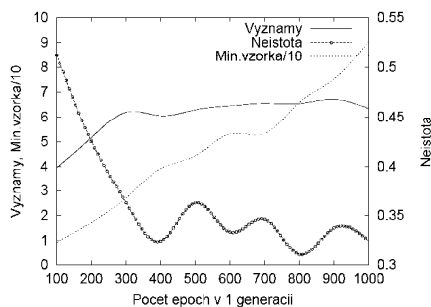
Aby sme vedeli kvalitatívne odhadnúť medzigeneračné posuny vo významoch, skúmali sme u agentov aj konkrétnu podobu interných reprezentácií významov. Zistili sme, že slovesné kategórie *narástol* a *zmenšil sa* reprezentované znamienkovými kritériami ostali rovnaké počas všetkých generácií. Takisto sa v nezmenenej podobe zachovali významy kategórií *trojuholník* (povinné 3 vrcholy, ostatné atribúty irelevantné) a *štvorec* (povinné 4 vrcholy a rovnaké dĺžky strán opísaného obdĺžnika). U kategórie *štíhly* sa presná hodnota pomeru výšky a šírky počas generácií menila, ale vlastnosť malej šírky v porovnaní s výškou bola korektné zachytená a zachovaná. Kategórie *veľký* a *malý*, v našom modeli založené na konkrétnych intervaloch navzájom nekorelujúcich atribútov, sa ukázali nestabilnými. Buď sa ich význam v priebehu generácií natoľko zúžil, že neprežili, alebo boli neprimerane rozšírené (pozri obr. 3).



**Obr. 2.** Kvalita lexikónu v experimente s iterovaným učením. Výmena generácií prebehla každých 500 krokov. Výsledky sú priemerom z 10 náhodných simulačných behov. (a) Podobnosť opisov a použiteľnosť sa rýchlo ustálili na hodnote blízkej 1 za cenu vyššej neistoty spôsobenej prílišným zovšeobecnením (rozšírením) niektorých významov. (b) Z pôvodného počtu 7 významov v prvej generácii sa zachovalo 6. Priemerne 1 význam neprežil kvôli svojej prílišnej špecializácii (zúženiu).



**Obr. 3.** Nestabilita niektorých kategórií. Zobrazená je projekcia receptívneho poľa kategórie do roviny  $sX, sY$ , projekčná rovina pretína ostatné dimenzie v priemerných hodnotách zameraných v kritériu. (a) Prílišná špecializácia: veľkosť receptívneho poľa s pribúdajúcimi generáciami konverguje k nule. (b) Prílišné zovšeobecnenie: náhodná korelácia niektorých atribútov spôsobí zanedbanie iných atribútov.



**Obr. 4.** Vplyv počtu učiacich epoch v jednej generácii na stabilitu významov. Zobrazené dáta sú priemerom z 10 náhodných simulačných behov. Miera „Významy“ udáva celkový počet žiakových významov spriemerovaný cez všetky učiace epochy vo všetkých generáciách. „Min. vzorka/10“ je veľkosť vzorky žiakovho kritéria s najmenšou vzorkou (spriemerovaná rovnakým spôsobom a zobrazená v škále 1/10). Miera „Neistota“ (definovaná v časti 4.2) vyjadruje (spriemerovaný) referenčnú nejednoznačnosť učiteľovej výpovede.

Kritickým parametrom pre stabilitu významov sa ukázala veľkosť množiny príkladov, z ktorej boli odvodené. Na obr. 4 sú zobrazené výsledky experimentovania s rôznymi hodnotami parametra počet

učiacich epoch v jednej generácii. Menší počet epoch spôsobil menšie množiny príkladov. Ak boli množiny príkladov príliš malé, významy sa stávali nestabilnými a u niektorých nastávala prílišná špecializácia (čo sa na grafe odráža ako zmenšenie počtu významov), u iných prílišné zovšeobecnenie (na grafe viditeľné ako nárast neistoty). Na grafe vidíme, že toto je prípad simulácií s 300 a menej krokovými generáciami, čomu zodpovedá menej ako 25 príkladov v množine významu s najmenším počtom príkladov. V simuláciách s viac ako 300 krokovými generáciami sa počet významov udrží medzi 6 a 7 a neistota je okolo 30–35 %. Tieto výsledky analyzujeme v nasledujúcej časti.

## 6 Diskusia

### 6.1 Vplyv zúženého profilu na významy

Keďže sa objekty a ich zmeny generujú náhodne, veľkosť žiakovej vzorky pre nejakú kategóriu v rámci fixného počtu učiacich epoch závisí od pravdepodobnosti, že vygenerovaný objekt bude dobrým príkladom učiteľovej kategórie. Táto pravdepodobnosť je vyššia pre všeobecnejšie kategórie ako pre špecifické. Toto vytvára zúžený profil pre prenos významov. V modeli emergencie gramatiky iterovaným učením zúžený profil vedie k objaveniu sa kompozičných štruktúr, keďže tie sa ľahšie prenášajú cez zúžený profil [14, 33].

Zachytávanie a amplifikácia náhodných regularít je inherentnou vlastnosťou iterovaného učenia. Zatiaľ čo táto vlastnosť je výhodná pre emergenciu gramatiky, v našom modeli môže viesť k skresľovaniu významov. Čím je vzorka menšia, tým je pravdepodobnejšie, že bude obsahovať nejakú náhodnú koreláciu, ktorá nebola súčasťou pôvodnej kategórie. Ak takúto koreláciu skonštruované rozlišovacie kritérium zachytí, učiteľ nasledujúcej generácie bude za inštancie tejto kategórie považovať už iba príklady obsahujúce túto koreláciu, čím sa pravdepodobnosť, že vzorka bude malá, ešte zvýši. Tým sa môžu kategórie stať príliš špecifickými (obr. 3a). Na tlmenie prílišnej špecializácie slúži zovšeobecňovanie pomocou SVD-rozkladu [30], ktorý zachytáva najinvariantnejšie vlastnosti a od ostatných abstrahuje, to však môže za istých okolností viesť k nesprávnemu zanedbaniu niektorého podstatného atribútu a k prílišnému rozšíreniu (zovšeobecneniu) významu (obr. 3b).

Prenos významov môžeme chápať ako evolučný proces, v ktorom sú významy replikátormi súťažiacimi o prežitie. Selektívnym tlakom je zúžený profil prenosu významov. Významy prejdú cez zúžený profil, ak sú relevantné prostrediu (teda ich inštancie sa v ňom vyskytujú dostatočne často). Veľmi špecifické významy

opisujúce zriedkavé situácie majú menšiu šancu na prežitie ako často aplikovateľné všeobecné významy.

Deti interagujú so svojim okolím, avšak zároveň sú vystavené lingvistickému vstupu. Pomenovávanie vplýva na formovanie nových konceptov [36]. Ukázali sme, ako si učiaci sa agent na základe verbálnej inštrukcie spolu s neverbálnou referenciou skonštruujú významy, ktoré sú pre plnenie pragmatických cieľov dostatočne podobné s význammi učiteľa. Vysoká podobnosť bola dosiahnutá rýchlo, čo je v súlade s pozorovaným fenoménom rýchleho učenia (*fast mapping*) u detí [4].

Ak necháme proces akvizície významov iterovať, významy neostanú nezmenené. Aj keď v každej generácii ostane vysoká podobnosť medzi význammi učiteľa a žiaka, medzigeneračne sa významy budú posúvať a vyvíjať. Tieto výsledky zodpovedajú tomu, že jazyky podliehajú historickým premenám bez toho, aby stratili svoju dorozumievaciu funkciu.

Významy konštituované jednoduchými štruktúrnymi vzťahmi sa ukázali medzigeneračne omnoho stabilnejšie ako významy založené na konkrétnych intervalových hodnotách nekorelovaných atribútov. V našom modeli boli významy *veľký* a *malý* založené na ordinálnom vzťahu k nejakým konkrétnym konštantám, čo nie je veľmi realistické. Sémantika väčšiny adjektív je závislá od podstatného mena, ktoré modifikujú [35], napr. adjektívum *veľký* odkazuje vo fráze „veľký zajac“ na inú veľkosť ako vo fráze „veľký slon“. Význam takýchto adjektív je konštituovaný štruktúrnymi vzťahmi, ktoré sa zobrazujú do domény generovanej príslušným podstatným menom. Naš výskum napovedá, že sémantická závislosť adjektív môže byť dôsledkom dynamiky selekčného procesu v rámci iterovaného prenosu, v ktorom sú významy založené na štruktúrných vzťahoch omnoho stabilnejšie.

## 6.2 Obmedzenia modelu a čo ďalej

Model v súčasnej podobe nezvláda prítomnosť šumu (chybných pomenovaní) vo vstupnej vzorke, ani homonymá (rovnaké pomenovanie pre rôzne významy). Takisto nezachytáva hierarchické a taxonomické vzťahy medzi konceptami. V budúcnosti plánujeme vybaviť agenta akciami, cieľmi a plánovaním, adekvátne rozšíriť sémantiku a študovať vzájomnú interakciu individuálnej pragmatickej konštrukcie významov s konštrukciou na základe sociálneho učenia. Takisto plánujeme rozšíriť model o jazyk s gramatikou.

## 7 Zhrnutie

V tomto článku sme opísali implementovaný model konštrukcie významov za pomoci sociálneho učenia a skúmali dynamiku významov v iterovanom

medzigeneračnom prenose. Treba zdôrazniť, že významy nie sú dané vopred ani nie sú interpretované externým pozorovateľom, ale sú konštruované (naučené) samotnými agentmi, pre ktoré majú inherentný význam. Učenie v modeli je inkrementálne a permanentné. Učiaci mechanizmus je citlivý na korelácie atribútov inštancií s príslušnosťou ku kategóriám aj na vzájomné korelácie medzi atribútmi. Tým sa stáva zaujímavým nielen pre kognitívne modelovanie, ale aj pre praktické aplikácie.

## Podakovanie

Tento výskum bol čiastočne podporený grantom APVV-20P04805 Agentúry na podporu výskumu a vývoja a grantom VEGA 1/3105/06.

## Literatúra

- [1] Barsalou, L.W.: Perceptual symbols systems. *Behavioral and Brain Sciences* 22 (1999) 577–660.
- [2] Cangelosi, A.: Approaches to grounding symbols in perceptual and sensorimotor categories. In: *Handbook of Categorization in Cognitive Science* (Cohen, H., Lefebvre, C., eds.), Elsevier, 2005: 719–737.
- [3] Cangelosi, A.: The grounding and sharing of symbols. *Pragmatics and Cognition* 14/2 (2006) 275–285.
- [4] Carey, S., Bartlett, E.: Acquiring a single new word. *Papers and Reports on Child Language Development* 15 (1978) 17–29.
- [5] Clark, E.: The principle of contrast: A constraint on language acquisition. In: *Mechanisms of language acquisition* (MacWhinney, B., ed.), Lawrence Erlbaum Assoc., Hillsdale, NJ, 1987: 1–33.
- [6] Feldman, J.: *From Molecule to Metaphor: A Neural Theory of Language*. MIT Press, Cambridge, MA, 2006.
- [7] Gärdenfors, P.: Cued and detached representations in animal cognition. *Behavioral Processes* 35 (1996) 263–273.
- [8] Gärdenfors, P.: Language and the evolution of cognition. In: *Penser l'esprit: Des sciences de la cognition à une philosophie cognitive* (Rialle, V., Fiset, D., eds.), Presses Universitaires de Grenoble, Grenoble, 1996: 151–172.

- [9] Gärdenfors, P.: *Conceptual Spaces*. MIT Press, Cambridge, MA, 2000.
- [10] Gärdenfors, P.: Cooperation and the evolution of symbolic communication. In: *The Evolution of Communication Systems* (Oller, K., Griebel, U., eds.), MIT Press, Cambridge, MA, 2004: 237–256.
- [11] Gibbs, R.: *Embodiment and Cognitive Science*. Cambridge University Press, Cambridge, 2006.
- [12] Harnad, S.: The symbol grounding problem. *Physica D* 42 (1990) 335–346.
- [13] Kelemen, J.: *Strojovia a agenty*. Archa, Bratislava, 1994.
- [14] Kirby, S., Hurford, J. The emergence of linguistic structure: an overview of the iterated learning model. In: *Computational Approaches to the Evolution of Language and Communication* (Parisi, D., Cangelosi, A., eds.), Springer-Verlag, Berlin, 2001: 121–148.
- [15] Kováč, L.: Fundamental principles of cognitive biology. *Evolution and Cognition* 6 (2000) 51–69.
- [16] Lakoff, G.: *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind*. University of Chicago Press, Chicago, 1987.
- [17] Maturana, H. R., Varela, F. J.: *The Tree of Knowledge: The Biological Roots of Human Understanding*. Shambhala, Boston, MA, 1987.
- [18] Nehaniv, C.: The making of meaning in societies: Semiotic and information-theoretic background to the evolution of communication. In: *Proceedings of the AISB 2000 Symposium: Starting from Society – the Application of Social Analogies to Computational Systems*. AISB (Edmonds, B., Dautenhahn, K., eds.), 2000: 73–84.
- [19] Pecher, D., Zwaan, R. A. (Eds.): *Grounding Cognition: The Role of Perception and Action in Memory, Language, and Thinking*. Cambridge University Press, Cambridge, U. K., 2005.
- [20] Piaget, J.: *The Child's Construction of Reality*. Routledge and Kegan Paul, London, 1955.
- [21] Piaget, J., Inhelder, B.: *La Psychologie de L'enfant*. PUF, Paris, 1966.
- [22] Pulvermüller, F.: Words in the brain's language. *Behavioral and Brain Sciences* 22/2 (1999) 253–279.
- [23] Quine, W.: *Word and Object*. MIT Press, Cambridge, MA, 1960.
- [24] Rizzolatti, G., et al.: Premotor cortex and the recognition of motor actions. *Cognitive Brain Research* 3 (1996) 131–141.
- [25] Spelke, E. S.: Principles of object perception. *Cognitive Science* 14 (1990) 29–56.
- [26] Steels, L.: Language as a complex adaptive system. In: *Proceedings of PPSN-VI*. (Schoenauer, M. ed.), Springer-Verlag, Berlin, 2000: 17–26.
- [27] Steels, L., Kaplan, F.: Bootstrapping grounded word semantics. In: *Linguistic Evolution Through Language Acquisition: Formal and Computational Models* (Briscoe, T., ed.), Cambridge University Press, Cambridge, UK., 2001: 53–73.
- [28] Steels, L., Kaplan, F., McIntyre, A., Looveren, J. V.: Crucial factors in the origins of word-meaning. In: *The Transition to Language* (Wray, A., ed.), Oxford University Press, Oxford, 2002: 252–271.
- [29] Takáč, M.: Categorization by sensory-motor interaction in artificial agents. In: *Proceedings of the 7th International Conference on Cognitive Modeling* (Fum, D., Del Missier, F., Stocco, A., eds.), Edizioni Goliardiche, Trieste, Italy, 2006: 310–315.
- [30] Takáč, M.: Kognitívna sémantika komplexných kategórií založená na rozlišovacích kritériách. In: *Mysel, inteligencia a život* (Kvasnička, V., Trebatický, P., Pospíchal, J., Kelemen, J., eds.), Vydavateľstvo STU, Bratislava, 2007.
- [31] Tomasello, M., Farrar, J.: Joint attention and early language. *Child Development* 57 (1986) 1454–1463.
- [32] Vogt, P.: The physical symbol grounding problem. *Cognitive Systems Research* 3/3 (2002) 429–457.
- [33] Vogt, P.: The emergence of compositional structures in perceptually grounded language games. *Artificial Intelligence* 167/1-2 (2005) 206–242.

- [34] Vygotsky, L. S.: *Mind in Society. The Development of Higher Psychological Processes*. Harvard University Press, Cambridge, MA, 1978.
- [35] Warren, B.: Ambiguity and vagueness in adjectives. *Studia Linguistica* 42/2 (1988) 122–171.
- [36] Waxman, S. R., Braun, I. E.: Consistent (but not variable) names as invitations to form object categories: new evidence from 12-month-old infants. *Cognition* 95 (2005) B59–B68.
- [37] Whorf, B. L.: *Language, Thought and Reality: Selected Writings of Benjamin Lee Whorf*. MIT Press, Cambridge, MA, 1956.