

Návrh kognitívnej architektúry pre jazykové experimenty

Martin Takáč

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

Abstrakt. Článok prezentuje návrh kognitívne relevantnej architektúry pre multiagentové modely vývoja jazyka. Agenty autonómne konajúce a vnímajúce v dynamickom prostredí si na základe svojej individuálnej senzomotorickej skúsenosti vytvárajú reprezentácie nielen objektov, ale aj prebiehajúcich procesov a používajú ich ako významy v komunikácii.

1 Úvod

Popri empirických poznatkoch a teoretických modeloch sa v kognitívnych vedách ako významný nástroj overovania hypotéz čoraz viac uplatňujú výpočtové modely a simulácie. V oblasti skúmania rôznych aspektov koordinovanej komunikácie, vzniku, vývoja a osvojovania si jazyka zdôrazňujú publikované výpočtové modely obvykle buď vplyv vrodených faktorov a modelujú ich genetickú evolúciu [38], alebo zdôrazňujú vplyv učenia a prostredia a modelujú kultúrny prenos jazyka [13],[27]. Niektoré (koevolučné) modely sú kombináciou oboch prístupov [29].

V tomto článku sa budeme zaoberať modelovaním mechanizmov, ktoré ovplyvňujú jazyk na ontogenetickej škále (osvojovanie si jazyka), na glosogenetickej škále (historické premeny jazyka, kultúrny prenos), a ich vzájomnou interakciou.¹

Z existujúcich modelov vývoja jazyka kultúrnym prenosom považujeme za paradigmatické modely iterovaného učenia (prehľad v [13]) a modely založené na samoorganizácii (prehľad v [27]). Obe paradigmy majú spoločné to, že simulujú spoločenstvo používateľov jazyka pomocou multi-agentového systému, v ktorom navzájom interagujú jednotlivé autonómne softvérové entity – agenty podľa lokálne definovaných pravidiel.

V paradigme iterovaného učenia jazyk existuje v dvoch formách – ako interná reprezentácia jazykových znalostí hovoriaceho a ako jeho externe pozorovateľné jazykové výpovede. Samotný jazyk sa vyvíja striedaním procesov akvizície (osvojovania) a produkcie jazyka: dieťa si vytvára svoju vlastnú internú znalosť o jazyku na základe externého lingvistického vstupu od dospelých, neskôr sa samo stane dospelým a produkuje externé výpovede, ktoré sú vstupom akvizície pre ďalšiu generáciu atď.

Publikované modely založené na paradigme iterovaného učenia úspešne simulujú vznik niektorých gramatických črt jazyka (kompozičnosť), avšak syntaktická štruktúra výpovedí len odráža preddefinovanú štruktúru významov. Od súvisiacich sémantických problémov (najmä problém ukotvenosti symbolov [11]) sa abstrahuje – významy sú umelé, vopred dané a rovnaké pre všetky agenty. Navyše sa v komunikácii používa „telepatia“, pretože počúvajúcí dostane na vstupe osvojovacieho algoritmu

¹ Tým nijako nepopierame dôležitosť vrodených faktorov – biologických a kognitívnych predispozícií (najmä v podobe samotných osvojovacích algoritmov).

nielen externe pozorovateľnú výpoveď, ale spolu s ňou aj priamo hovorcovu internú reprezentáciu významu výpovede.

V modeloch samoorganizácie sa jazyk považuje za dynamický systém, ktorý emerguje ako globálny produkt decentralizovaných lokálnych interakcií jednotlivých používateľov jazyka. Globálna koherencia nastáva v dôsledku existencie pozitívnej spätnej väzby medzi výberom jazykového prostriedku a jeho úspešnosťou: Ak má hovoriaci viac spôsobov, ako vyjadriť nejaký význam, zvolí si ten, ktorý v minulosti viedol k úspešnému porozumeniu. A naopak, ak je jazykový výraz často používaný, rozšíri sa v populácii a stane sa úspešným dorozumievacím prostriedkom.

Agenty sú simulované alebo stelesnené v reálnych robotoch [25] a vnímajú objekty na scéne (napr. farebné geometrické tvary na tabuli). Počas tzv. *diskriminačných hier* si každý agent buduje vlastnú vnútornú reprezentáciu objektov (nepriístupnú priamo iným agentom). Význam symbolov je ukotvený v týchto reprezentáciách. Agenty hrajú rôzne typy *jazykových hier*, scenáre interakcií sú však vopred pevne dané a témami komunikácie sú vždy len statické objekty aktuálne prítomné na scéne. Komunikácia je situačne viazaná (*vyvolaná* [5]) a lexikón pozostáva iba zo slov zodpovedajúcich podstatným a prídavným menám; slovesá a komplexnejšie kategórie úplne absentujú. Modely teda úspešne simulujú vývoj jednoduchých komunikačných systémov približne na úrovni zvieracích systémov varovných volaní.

Z empirických výskumov vieme, že prvé detské slová neoznačujú len veci, ale aj vzťahy, akcie a vnútorné stavy [34] a prvé komplexnejšie lexikálne konštrukcie dieťaťa sú organizované okolo sloviess (hypotéza *verb islands* [32]). Osvojenie si sloviess je ťažšie ako podstatných mien [8] a ich korektný význam u dieťaťa kryštalizuje dlho a postupne [20].

Dieťa pri osvojovaní si jazyka využíva kognitívne schopnosti, ktorými v čase akvizície už disponuje: schopnosť vnímať a rozlišovať jednotlivé objekty [30], segmentovať udalosti v bodoch maximálnej zmeny vo vstupných dátach [39], vnímať akcie iných ako cieľovo-orientované a intencionálne [2]. Významnú úlohu hrá zdieľaná pozornosť s matkou [33].

V tomto príspevku vychádzame z centrálného predpokladu, že pri štúdiu vývoja jazyka nemožno vynechať sémantické otázky, teda pôvod, mechanizmy vzniku a vývoja reprezentácií, ktoré tvoria významy pre komunikáciu. Našou ambíciou je navrhnúť realistickejší model, ktorý by sa vyhol vyššie uvedeným zjednodušeniam a bol v súlade s uvedenými empirickými poznatkami.

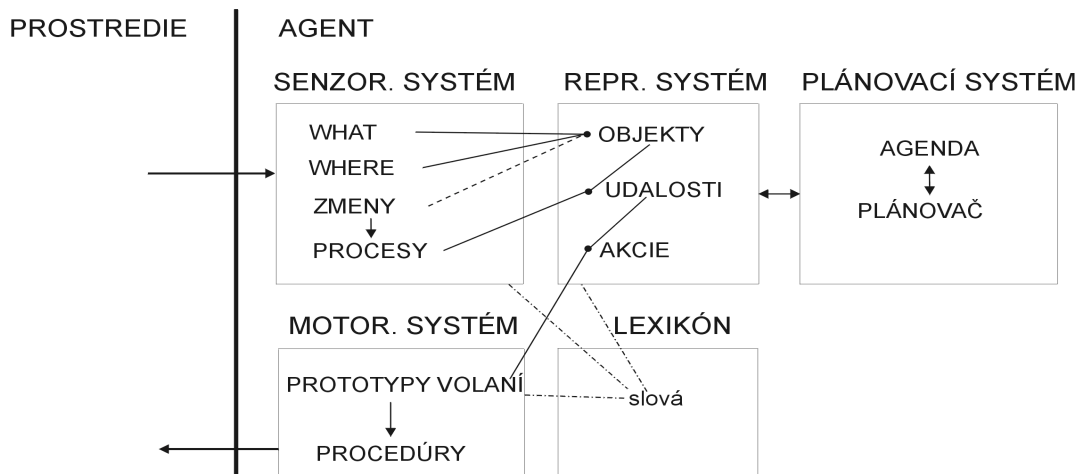
V ďalšej časti opíšeme základné kognitívne podsystemy navrhovanej architektúry potrebné pre vybudovanie predverbálnych reprezentácií prostredia agenta. Časť 3 opisuje jazykový aparát a vytvára prepojenie s modelmi vývoja jazyka. Posledná časť je diskusiou.

2 Architektúra pre predverbálnu kogníciu

2.1 Prostredie

Prostredie je dvojrozmerná mriežka, na ktorej sú rozmiestnené objekty – agenty alebo predmety. Každý objekt má vlastnosti, ktorých hodnota sa môže v čase meniť (napr. poloha). Čas sa mení v diskretných krokoch. Aktivita každého agenta je motivovaná jeho systémom potrieb, v ktorom je urgentnosť každej potreby vyjadrená číslom rastúcim s časom (agent ho môže znížiť vykonaním vhodnej akcie, napr. hlad sa môže znížiť zjedením nejakého jedlého objektu). Dynamika prostredia závisí od konkrétneho

nastavenia experimentu – rozmiestnenia a typov predmetov, potrieb agentov a vykonateľných akcií. Agent vníma prostredie, vytvára si jeho vnútornú reprezentáciu, plánuje a vykonáva akcie s cieľom uspokojiť svoje potreby. Tieto úlohy zabezpečujú kognitívne podsystémy (obrázok 1), ktoré si teraz postupne opíšeme.



Obrázok 1. Kognitívne subsystémy agenta.

2.2 Senzorický systém

Agentov aktuálny vnem tvorí zoznam objektov, ktoré sa momentálne nachádzajú v jeho *receptívnom poli* – výseku mriežky do určitej vzdialenosti od neho. Agent vníma každý objekt (vrátane seba) ako rámec – množinu trojíc $\langle \text{atribut} : \text{hodnota}, \Delta \rangle$, kde *atribut* je názov vlastnosti, *hodnota* je aktuálna hodnota vlastnosti v čase t a Δ je rozdiel medzi aktuálnou hodnotou a predchádzajúcou hodnotou v čase $t-1$. Tento zjednodušujúci predpoklad má abstrahovať od pedspracovania vstupu rutinami nižšej úrovne – segmentácia scény na objekty, identifikácia atribútov a prebiehajúcich zmien, atď.² Potreby sú implementované ako atribúty agentovho rámca viditeľné iba pre neho samého.

Agentov senzorický systém pozostáva zo subsystémov pre vnímanie typov objektov (WHAT), pozícií (WHERE) a dynamických zmien (ZMENY/PROCESY). Takúto dekompozícia má kognitívnu motiváciu – u ľudí boli dokázané samostatné neurálne cesty pre identifikáciu vnemu – WHAT systém, a priestorovú lokalizáciu vnemu – WHERE systém [36], funkčný systém pre rozlišovanie pohybu [14] a oblasť s tzv. zrkadliacimi neurónmi pre vnímanie a reprezentáciu akcií [21]. Dekompozícia má význam aj z funkčného hľadiska – subsystémy spracúvajú rôzne časti vstupnej informácie rôznym spôsobom: WHAT systém spracúva relatívne stabilné vlastnosti objektov pre triedenie do kategórií, WHERE systém dynamicky sa meniacu polohu a ZMENY/PROCESY zachytáva prebiehajúce procesy.

² Schopnosť vnímať jednotlivé objekty má dieťa veľmi skoro po narodení [30]. Pre umelý systém to vôbec nie je jednoduchý problém, rieši ho napr. [16].

2.3 Kategorizácia

Hlavným účelom sensorického systému je kategorizovať vstupy, t.j. zatriediť nové vnemy vzhľadom na už vnímané a zapamätané. Agent si zapamätáva/reprezentuje vnímané v podobe prototypov – typických reprezentantov kategórií [22]. Prototypy sú body resp. vektory v mnohorozmernom konceptuálnom priestore [7], ktorého dimenzie zodpovedajú atribútom objektov. Prototypy zabezpečujú úspornosť reprezentácie – spojitý priestor podelia na niekoľko diskretných kategórií, ktoré môžu byť neskôr v lingvistickej fáze experimentu označené slovami.

Agent kategorizuje vnem tak, že nájde k nemu najbližší prototyp. Ak sú dimenzie konceptuálneho priestoru separabilné (hodnoty atribútov sú nezávislé), Gärdenfors [7] odporúča použiť pre vzdialenosť blokovú (Manhattan) metriku, pre priestory s integrálnymi dimenziami (hodnoty atribútov navzájom korelujú) Euklidovskú metriku.³ Pre určenie vzdialenosti vnemu x od prototypu p sú:

$$d_{sep}(p, x) = \sum_{i=1}^N |p_i - x_i| \quad d_{int}(p, x) = \sqrt{\sum_{i=1}^N (p_i - x_i)^2} \quad (1)$$

Pri počítaní vzdialenosti nemusia byť rozdiely na všetkých dimenziách rovnako dôležité. Škály dimenzií nie sú nijako normalizované,⁴ posudzovanie rozdielu závisí od kontextu⁵ a ľudia majú tendenciu prisudzovať väčšiu váhu dimenziám, na ktorých sú hodnoty porovnávaných objektov podobné [18].⁶ Preto v našom modeli používame váhovanú verziu metrik:

$$d_{sep}(p, x) = \sum_{i=1}^N w_i^{(p)} |p_i - x_i| \quad d_{int}(p, x) = \sqrt{\sum_{i=1}^N w_i^{(p)} (p_i - x_i)^2} \quad (2)$$

kde váha $w_i^{(p)}$ dimenzie prislúchajúcej i -temu atribútu prototypu p je nepriamo úmerná rozptylu S_i^2 hodnôt atribútu i doteraz videných inštancií prototypu p :

$$w_i^{(p)} = \frac{1}{S_i^2 + 1} \quad (3)$$

Každý prototyp má teda svoje vlastné váhy.⁷

Podobnosť je nepriamo úmerná vzdialenosti a zodpovedá empirickému faktu, že pri blízkych objektoch klesá so vzdialenosťou rýchlo, zatiaľ čo u relatívne vzdialených objektov klesá omnoho pomalšie (c je konštanta citlivosti rozlišovania) [7]:

$$Sim(p, x) = e^{-c \cdot d(p, x)} \quad (4)$$

Takto zvolená podobnosť má ďalšie želané a kognitívne plauzibilné vlastnosti:

³ Delenie na integrálne a separabilné dimenzie je neostré. Keďže v našom modeli nemáme informácie o koreláciách dimenzií, predpokladáme, že dimenzie sú separabilné a používame blokovú metriku.

⁴ Niektoré klasterizačné algoritmy, napr. DCA [12] vyžadujú, aby boli škály všetkých dimenzií normalizované na interval $\langle 0,1 \rangle$.

⁵ Dynamicky sa meniace váhy v závislosti od kontextu a zamerania pozornosti navrhujú v [7] a [9], v našom modeli ich nepoužívame.

⁶ Väčšia váha spôsobí „rozťahnutie dimenzie“ a umožní jemnejšie rozlišovanie. Pre dimenzie s väčším rozptylom stačí hrubšie rozlišovanie (menšia váha).

⁷ S prototypom sú uložené priebežne iteratívne aktualizované rozptyly pre jednotlivé dimenzie, váhy sa z nich vypočítajú, keď sa určuje vzdialenosť. Pri použití blokovej metriky sa vzťah pre výpočet rozptylu upraví príslušným spôsobom.

- Atribúty, ktorých hodnoty u inštancií kategórie veľmi varujú, sa stanú pre určenie príslušnosti do tejto kategórie nedôležitými. To umožňuje vytvárať aj prototypy sémanticky zodpovedajúce adjektívam, teda s veľmi malými váhami pre všetky dimenzie okrem jednej (napr. pojem „byť vysoký“).
- Váhy implicitne reprezentujú „dosah“ alebo polomer kategórie: kategória, ktorej všetky doteraz vnímané inštanície mali nulový alebo len veľmi malý rozptyl (a teda váhy veľké, blízke 1), sémanticky zodpovedá skôr individuálnemu objektu. Čím väčší rozptyl, tým všeobecnejší pojem prototyp reprezentuje. Funkcia podobnosti to reflektuje – pri rovnakých rozdieloch hodnôt vracia väčšiu podobnosť pre všeobecnejšie pojmy (ak by sa prototyp „jablka“ líšil od prototypu „Golden Delicious“⁸ iba väčším rozptylom, objekt „Jonatan“ sa bude viac podobat’ na „jablko“ ako na „Golden Delicious“).
- Tversky [35] v množstve psychologických experimentov ukázal, že ľuďmi vnímaná podobnosť je asymetrická (napr. 99 sa viac podobá na 100 ako sa 100 podobá na 99). Rovnako sa správa aj nami definovaná funkcia podobnosti: ak porovnávam dva prototypy a p_1 je „typickejší“ prototyp (reprezentant rozmanitejšej množiny inštancií, kategória s väčším rozptylom) ako p_2 , bude $Sim(p_1, p_2) > Sim(p_2, p_1)$.

Pri vyššie uvedených vzorcoch pre výpočet vzdialenosti sme predpokladali, že dimenzie majú intervalové škály, teda, že možno určiť rozdiel $|p_i - x_i|$ pre ľubovoľné dve hodnoty p_i, x_i na i -tej dimenzii. Ak je škála dimenzie nominálna (hodnoty sa dajú porovnávať len na rovnosť), položíme

$$|p_i - x_i| = \begin{cases} 0 & \text{ak } p_i = x_i \\ 1 & \text{ak } p_i \neq x_i \end{cases} \quad (5)$$

Nie všetky atribúty sú relevantné pre všetky objekty (napr. chuť auta; ťažko tiež vyjadriť mieru podobnosti medzi jablkom a autom). Dva objekty (reprezentované rámcami s atribútmi) nazveme *kompatibilnými*, ak majú rovnakú množinu atribútov. Namiesto jedného všeobecného konceptuálneho priestoru budeme používať implicitné podpriestory/domény tak, že v každej doméne budú reprezentácie iba navzájom kompatibilných objektov.⁹ To zabezpečí funkcia podobnosti modifikovaná tak, že pre prototyp a vnem, ktoré sú navzájom nekompatibilné, vráti 0.

Algoritmus, ktorý kategorizuje vnemy, zároveň ukladá a modifikuje prototypy:¹⁰

1. Pre vstupný vnem x nájdi najpodobnejší prototyp p .
2. Ak neexistuje žiaden, alebo $Sim(p, x) < MinSim(p)$, vytvor prototyp z vnemu a ulož ho (vznikla nová kategória).
3. Inak je x inštanciou kategórie reprezentovanej prototypom p . Posuň prototyp p bližšie k vnemu x a aktualizuj váhy $w^{(p)}$ podľa rozptylov zahŕňajúcich hodnoty atribútov x .

⁸ Golden Delicious a Jonatan sú dve rôzne sorty jablka.

⁹ Teda napr. autá, jablká a hrušky nebudú reprezentované v priestore {farba, cena, kubatúra, chuť, sorta}, ale autá v priestore {farba, cena, kubatúra} a jablká a hrušky v priestore {farba, chuť, sorta}.

¹⁰ Algoritmus je inšpirovaný klasterizačným algoritmom DCA [12], líši sa však od neho vo viacerých dôležitých detailoch.

MinSim je funkcia, ktorá vracia prah minimálnej podobnosti požadovanej pre to, aby x bol považovaný za inštanciu kategórie reprezentovanej prototypom p .¹¹

Pre posúvanie prototypu možno zvoliť viacero stratégií:

- a) *Rigidné prototypy*: prototyp sa neposúva vôbec.
- b) *Iteratívny priemer*: prototyp je ťažiskom vnemov, hodnoty jeho atribútov sú priemerom hodnôt atribútov doteraz videných inštancií kategórie.¹² Počíta sa iteratívnym¹³ predpisom

$$\bar{p}_1 = x_1, \quad \bar{p}_N = \left(1 - \frac{1}{N}\right)\bar{p}_{N-1} + \frac{1}{N}x_N \quad (6)$$

- c) *Sledovanie vnemu*: hodnoty atribútov prototypu sa nahradia hodnotami atribútov vnemu. Rozptyl sa nepočíta, váhy sú konštantne 1.

Algoritmus možno obohatiť o ďalšie operácie, napr. zlúčenie veľmi podobných prototypov a vymazanie málo používaných prototypov.

V opísanej podobe algoritmus klasterizuje bez učiteľa (unsupervised). Krok 3 algoritmu by sa dal využiť aj na učenie kategórií s učiteľom – ak pre konkrétnu kategóriu dodávame jej inštancie/príklady, prototyp a váhy sa upravujú tak, že ich budú reflektovať.

2.4 Subsystémy WHAT a WHERE

Senzorické subsystémy využívajú opísaný kategorizačný algoritmus. Keďže každý subsystém spracúva iba časť vstupnej informácie, rámec každého objektu v receptívnom poli je pred vstupom do kategorizácie príslušným subsystémom filtrovaný.

Filter WHAT systému vynechá z rámcov informáciu o zmenách (Δ) a atribúty polohy. Naopak WHERE systém ponechá iba atribúty polohy a pridá počítaný atribút vzdialenosť vnímaného objektu od vnímajúceho¹⁴ (tiež bez Δ). Stratégiou posúvania prototypu pri kategorizácii vo WHAT systéme je iteratívny priemer.

Prototypy vo WHAT systéme reprezentujú typy resp. kategórie objektov, zatiaľ čo jednotlivé objekty sú reprezentované v reprezentačnom subsystéme OBJEKTY ako asociácie medzi kategóriou a polohou, napr. päť rôznych jabĺk v prostredí môže byť reprezentovaných piatimi asociáciami medzi prototypom kategórie jablko vo WHAT systéme a piatimi rôznymi prototypmi polohy vo WHERE systéme.

Stratégiou posúvania prototypu pre WHERE systém je sledovanie vnemu. Malá kontinuálna zmena polohy nejakého objektu bude sledovaná jej prototypom¹⁵ (až kým

¹¹ V našom modeli závisí od počtu nominálnych a intervalových atribútov. Tiež asymptoticky rastie s počtom doteraz videných inštancií prototypu: pre „mladý“ prototyp sa vyžaduje menšia podobnosť, s tým, že prototyp sa „opraví“ v kroku 3 posunom v smere vnemu, zatiaľ čo na opravovanie „osvedčeného“ prototypu treba vyššiu podobnosť, inak sa pre vnem vytvorí nový prototyp.

¹² Pre nominálne atribúty, kde sa nedá počítať priemer, sa s prototypom udržuje zoznam doteraz videných hodnôt atribútu a prototyp nadobúda hodnotu s maximálnou početnosťou (modus).

¹³ Rozptyl sa tiež počíta iteratívne. Vo výpočte rozptylu pre nominálne atribúty sa využije rovnica (5), čím rozptyl vyjadruje vlastne pomer počtu hodnôt rôznych od modu ku celkovému počtu (doteraz videných) hodnôt atribútu.

¹⁴ Tento atribút zabezpečí prepojenie objektívnej polohy so subjektívnym vnímaním (napr. ak sa objekt nehýbe a agent sa k nemu blíži). Agent ho môže využiť napr. na formuláciu čiastkového cieľa (ak chcem zjesť jablko, musím k nemu najprv prísť).

¹⁵ Ak mám viac rôznych objektov asociovaných s tou istou polohou, posúvanie prototypu vďaka zmene polohy jedného z nich by spôsobilo „odcestovanie“ aj ostatných objektov. V takom prípade sa prototyp rozdelí na dva a posúva sa len ten, ktorý je asociovaný s práve sa pohybujúcim objektom.

sa objekt nestratí z receptívneho poľa), zatiaľ čo ak sa objekt náhle objaví na inom¹⁶ mieste, vytvorí sa nový prototyp.¹⁷ Takto rôzne prototypy polohy asociované s rôznymi objektmi tvoria kognitívnu mapu miest, kde boli naposledy videné (vo veľmi dynamickom svete, kde sa objekty často hýbu, má zmysel implementovať zabúdanie dlho nepoužitých prototypov).

2.5 Vnímanie zmien

WHAT a WHERE systémy ignorujú Δ informáciu vstupných rámcov. Táto informácia sa spracováva v systéme ZMENY/PROCESY. Po sebe nasledujúce elementárne zmeny s rovnakým znamienkom sa kumulujú a tvoria proces trvajúci viac časových krokov. Agenty detekujú hranice medzi procesmi v bodoch, kde Δ zmení znamienko, napr. ak $\Delta(\text{Obj.položa}X, t-1) = 0$ a $\Delta(\text{Obj.položa}X, t) > 0$, objekt *Obj* sa začal hýbať doprava. Tento proces sa skončí v najbližšom okamihu t' , v ktorom je $\Delta(\text{Obj.položa}X, t') \leq 0$.

Senzorický subsystém ZMENY slúži ako dočasný krátkodobý pamäťový sklad pre kumulovanie prebiehajúcich zmien. Kumulatívne zmeny vstupného objektu *Obj* sa ukladajú ako trojice $\langle \text{atribut} : \Delta_{kum}, \text{trvanie} \rangle$ a sú asociované s reprezentáciou objektu *Obj* v systéme OBJEKTY. V každom časovom kroku t sa $\Delta(\text{Obj.atribut}, t)$ porovná s príslušným $\langle \text{atribut} : \Delta_{kum}, \text{trvanie} \rangle$ v systéme ZMENY. Ak Δ_{kum} a Δ majú rovnaké znamienko, Δ_{kum} sa nahradí $\Delta_{kum} + \Delta$ a *trvanie* sa zvýši o jedna (pokračujúci proces), inak sa Δ_{kum} nahradí Δ a *trvanie* sa nastaví na 1 (začal sa nový proces).

Agent má tiež dlhodobý pamäťový sklad pre ukladanie, rozpoznávanie (a neskôr aj pomenovávanie) procesov: subsystém PROCESY je podobný na ZMENY, ale procesy sa v ňom ukladajú v podobe prototypov. V každom časovom kroku sa kumulovaná prebiehajúca zmena $\langle \text{atribut} : \Delta_{kum}, \text{trvanie} \rangle$ kategorizuje v konceptuálnom podpriestore pre atribút *atribut* s dvoma dimenziami pre Δ_{kum} a *trvanie*. Reprezentácia procesov môže byť hrubá, kvalitatívna [31], alebo detailnejšia v závislosti od nastavenia prahovej funkcie *MinSim* v kategorizačnom algoritme subsystému PROCESY.

2.6 Reprezenačný systém

Reprezenačný systém integruje prototypové reprezentácie jednotlivých sensorických a motorických podsystémov a pozostáva z podsystémov OBJEKTY, UDALOSTI a AKCIE. Ako sme uviedli vyššie, jednotlivé objekty sú reprezentované ako asociácie medzi prototypom kategórie objektu vo WHAT systéme a polohy vo WHERE systéme (asociácie vytvára priamo sensorický systém pri dekompozícii vstupného vnemu).

Konkrétna udalosť je reprezentovaná v subsystéme UDALOSTI ako asociácia medzi reprezentáciou objektu v subsystéme OBJEKTY a prototypom v sensorickom subsystéme PROCESY. S každou udalosťou sa uchováva informácia o jej stave (*prebiehajúca, skončená*). Ak je udalosť spôsobená nejakou akciou agenta, bude s ňou

¹⁶ Teda ak je v druhom kroku kategorizačného algoritmu podobnosť miesta k najbližšiemu prototypu nižšia ako požadovaný prah podobnosti.

¹⁷ Sledované kotúľajúce sa jablko je vnímané ako stále to isté jablko, jablko objavené na úplne inom mieste, je iné jablko. Agent, rovnako ako človek, sa môže myliť, jablko mohol niekto premiestniť, kým sa pozeral inde. Dieťa v predoperačnom štádiu nevie rozlíšiť, či slimáky, ktoré stretá na prechádzke, sú jeden a ten istý slimák, alebo viaceré slimáky [24].

v reprezentačnom subsysteme AKCIE asociovaný príslušný prototyp volania motorickej procedúry (pozri v nasledujúcej časti).

2.7 Motorický systém

Agenty interagujú s prostredím vykonávaním akcií. Každý agent má sadu procedúr, ktoré za určitých podmienok menia atribúty objektov, napr. procedúra *chod'* môže meniť agentovu polohu, pokiaľ nemá pred sebou okraj ihriska alebo iného agenta. Procedúry modelujú motorické zručnosti, sú to čisto procedurálne znalosti a agent im nevidí „dovnútra“. Agentova deklaratívna znalosť o akciách a ich dôsledkoch vzniká až jeho interakciou so svetom. Agent si pamätá, s akými parametrami zavolał procedúru, a toto volanie si môže asociovať s jej dôsledkami – pozorovanými zmenami v prostredí¹⁸ (jedna akcia môže spôsobiť zmenu viacerých atribútov viacerých objektov, asociácie preto môžu byť viacnásobné).

Volania si ukladá v podobe motorických prototypov v konceptuálnom priestore, ktorého dimenzie tvoria parametre procedúr. (Pokiaľ by parametrom procedúry nebolo číslo alebo nominálna hodnota, ale zložitejšia štruktúra, napr. celý objekt, priestor by bol hierarchický, t.j. rozdiel hodnôt na dimenzii parametra objekt by sa určil na základe podobnosti objektov.)

2.8 Reprezentácia akcií

Systém AKCIE stelesňuje agentove (limitované) znalosti o kauzalite. Využíva ho plánovací systém na plánovanie akcií, ktoré by mali spôsobiť želané zmeny a splnenie cieľa. Preto sa s akciou ukladá aj jej kontext – situácia v čase jej vykonania, v podobe objektov ovplyvnených akciou. Keďže senzorické prototypy sa kontinuálne aktualizujú, motorické prototypy sa neasociujú priamo s relevantnými udalosťami v systéme UDALOSTI, ale s ich lokálnymi kópiami až na úroveň senzorických prototypov vytvorenými v čase vykonania procedúry.¹⁹

2.9 Plánovací systém

Úlohou plánovacieho systému je riadenie činnosti agenta, teda plánovanie akcií a vyhodnocovanie plnenia cieľov. V ľudskom mozgu sekvencovanie akcií zabezpečujú funkčné systémy prefrontálnej kôry [14], s. 284.

Plánovací systém je tvorený agendou cieľov a plánovačom. Každý cieľ v agende je boolovskou konštrukciou ordinálnych relácií nad reprezentáciami, napr. „*Self.hlad = 0* alebo *Obj1.polohaX > Obj2.polohaX*“ a má dynamicky sa meniacu prioritu. Trvalými cieľmi v agende sú ciele reprezentujúce splnené potreby a majú priority úmerné momentálnym urgenciám potrieb. Plánovač vyberie z agendy cieľ s najvyššou prioritou a testuje, či je splnený v aktuálnej situácii (tvorenej najaktívnejšími reprezentáciami, pozri časť Pozornosť). Ak áno, vyberie ďalší (ak už nemá nesplnený cieľ, nerobí nič).

¹⁸ Po vykonaní akcie agent môže pozorovať viacero zmien v prostredí vrátane tých, ktoré s akciou vôbec nesúvisia. Nájdenie zmien relevantných akcií je ťažké – má s ním problémy aj dieťa v senzomotorickom štádiu, keď za príčinu všetkých zmien považuje svoju činnosť (magicko-fenomenistická kauzalita [19], s. 24). Môžeme experimentovať s viacerými stratégiami od explicitného dodania relevantných zmien simulátorom prostredia cez heuristický odhad až po asociovanie všetkých práve prebiehajúcich zmien.

¹⁹ Namiesto „živých“ pohybujúcich sa prototypov teda akási fotka alebo momentka. Ak agent neskôr vykoná znova toto volanie procedúry a dostane inú „fotografiu“, môže spustiť abduktívny proces porovnávania oboch reprezentácií a nájdenia rozdielov, resp. spoločných čít.

Ak cieľ nie je splnený, nájde rozdiel medzi aktuálnou a cieľovou situáciou a hľadá akciu, ktorá by ho mohla eliminovať. (Např. ak cieľ je *Jablko.vzdialenosť* = 0, ale v momentálnej situácii je *Jablko.vzdialenosť* = 10, plánovač hľadá akciu, ktorá znižuje vzdialenosť od jablka.) Ak ju nenájde, hľadá sekvenciu elementárnych akcií a pridáva príslušné čiastkové podciele do agendy. Ak sa to nepodarí, agent vykoná náhodnú akciu.

Tento jednoduchý algoritmus nezaručuje rozumné správanie v ľubovoľnom prostredí. Deti si osvojujú jazyk oveľa skôr, ako disponujú usudzovaním na dospelú úroveň [19]. V experimentoch simulujúcich podmienky detí osvojujúcich si jazyk budú agenty situované vo veľmi jednoduchých a obmedzených prostrediach.

2.10 Pozornosť

Pozornosť je implementovaná ako dynamicky sa meniaci úroveň aktivity jednotlivých reprezentácií. Jej zaradenie do modelu je motivované tým, že človek vníma svoje okolie selektívne – ako figúru a pozadie [23], pričom figúru tvoria náhle zmeny v receptívnom poli alebo to, čo súvisí s jeho práve aktuálnym cieľom, všetko ostatné je pozadie [1],[10].

Každý prototyp v senzoryckom systéme má svoju úroveň aktivity – číslo, ktoré s časom klesá k nule, ak nie je restimulovaný. Restimulácia znamená, že úroveň aktivity reprezentácií pre každý objekt na vstupe sa zvýši priamo úmerne Δ jeho atribútov. Např. ak sa nejaký objekt pohol, zvýši sa aktivita prototypu jeho polohy vo WHERE systéme, ak sa začal alebo skončil nejaký dej, zvýši sa príslušná aktivita v systéme ZMENY atď. Ak sa v receptívnom poli objaví nový neznámy objekt, pridajú sa nové prototypy s vysokou úrovňou aktivity. Úroveň aktivity majú aj reprezentácie v reprezentačnom systéme, ich zmena aktivity sa počíta na základe zmien aktivít prototypov, ktoré asociujú. Reprezentácie môžu byť stimulované aj zhora, např. ak agent zvolí na dosiahnutie cieľa nejakú akciu.

Pozornosť umožňuje oddelené používanie reprezentácií (v zmysle [5],[6]): okrem oddeleného výberu akcie cieľom, nie aktuálnym vnemom, agent môže mať aktívne prototypy objektu, ktorý práve nie je v receptívnom poli. Pozornosť má dôležitú úlohu v jazykovej fáze experimentu, kde pomáha určiť spoločnú tému komunikácie.

3 Jazyková fáza

Doteraz opísaná architektúra modeluje predverbálnu kogníciu. Agent s potrebami a motorickými akciami situovaný v prostredí si vytvára interné reprezentácie konceptov relevantných jeho skúsenosti. Prototypy a asociácie uložené v senzoryckých a reprezentačných podsystemoch tvoria rôzne sémantické kategórie zodpovedajúce druhom objektov (WHAT), miestam (WHERE), jednotlivým objektom (OBJEKTY), procesom (PROCESY), udalostiam (UDALOSTI), zručnostiam (motorický systém) a akciám (AKCIE). Všetky tieto koncepty budú tvoriť významy v jazykovej fáze experimentu a budú označené slovami.

Lexikón agenta pozostáva z obojsmerných asociácií významov so slovami (reťazcami znakov). Jeden význam môže byť asociovaný s viacerými slovami (synonymia) a jedno slovo s viacerými význammi (homonymia). Každá asociácia má dynamicky sa meniacu silu. Keď agent počuje nejaké slovo a má aktívne niektoré reprezentácie (teda „venuje pozornosť nejakým významom“), zvýši silu všetkých

asociácií medzi počutým slovom a aktívnymi význammi. Lexikón vzniká samoorganizáciou, teda ak agent nemá asociované slovo na vyjadrenie nejakého významu, použije náhodný reťazec znakov (a asociuje si ho s významom), neskôr preferuje slovo s najsilnejšou asociáciou k danému významu.

3.1 Modelovanie vývoja jazyka

Prezentovanú architektúru možno použiť v modeloch vývoja jazyka samoorganizáciou aj iterovaným učením. V prvom prípade je populácia agentov rozmiestnená na mriežke, každý autonómne vykonáva akcie na uspokojenie svojich potrieb (v najjednoduchšom prípade experimentu sú na mriežke len agenty a jedlé objekty a agent má v repertoári dve akcie chod' a jedz). Každý agent si vytvorí vlastné významy, ktoré nie sú priamo prístupné iným agentom (nepripúšťame telepatiu). Pre zjednodušenie môže byť verbalizovanie konceptov bez zámerného rozhodnutia agenta²⁰ – každý agent automaticky v každom kroku verbalizuje svoju najaktívnejšiu reprezentáciu (význam, ktorému „venuje najväčšiu pozornosť“). Takisto si každé slovo produkované inými agentmi, ktoré sú aktuálne v jeho receptívnom poli, asociuje so svojimi aktívnymi význammi (silu asociácie zvýši úmerne aktivite významu). Možno skúmať rôzne verbalizačné stratégie, napr. preferovanie slov pre jednotlivé objekty (vlastné mená), zložené opisy kategóriou a polohou objektu, alebo „stručne a všeobecne“ (*succinct-and-general* [4]).

Tento prístup predpokladá významnú rolu zdieľanej pozornosti pri určovaní referenta výpovede. Na zabezpečenie toho, že pozornosť agentov je naozaj zdieľaná (upretá na veľmi podobné veci), bude pravdepodobne potrebné pridať do modelu mechanizmus analogický sledovaniu pohľadu (neverbálnu náповedu zužujúcu kontext možných referentov) alebo obmedzenia významu známe z výskumov akvizície jazyka deťmi, napr. vzájomnú výlučnosť [17].

Uvedené stratégie verbalizovania a asociovania možno použiť aj v rámci iterovaného učenia. V tom prípade sú na mriežke len dva agenty, dospelý a dieťa. Učenie je asymetrické – dieťa sa učí verbalizácie dospelého, ale nie naopak. Po nejakom čase je dospelý agent odstránený z plochy, dieťa sa stane dospelým a na plochu sa vloží nový agent s prázdny lexikónom (dieťa). Tento proces sa nechá iterovať mnoho generácie, počas ktorých sa sledujú zmeny v reprezentáciách a lexikóne.

4 Diskusia

4.1 Súvisiace práce

Cenným príspevkom k štúdiu vytvárania reprezentácii na základe senzomotorickej interakcie s prostredím sú práce v rámci paradigmy Spatial Semantic Hierarchy (SSH) [15], avšak tento výskum nebol aplikovaný na štúdium jazyka.

Iterované učenie s ukotvením významov diskriminačnými hrami možno nájsť vo Vogtovom modeli [37]. Agenty komunikujú o farebných geometrických tvaroch, ukladajú si ich reprezentácie v konceptuálnych priestoroch a vytvárajú zdieľanú pozornosť neverbálnymi prostriedkami (ukazovaním). Problémami komunikácie bez

²⁰ V tomto prípade komunikácia neslúži agentom pre žiadny účel a je analogická zvieračím škrekom, ktoré zvierať nemôžu potlačiť. Dômyselnejšia stratégia je umožniť agentom vybrať verbalizáciu ako jednu z akcií na dosiahnutie cieľa, to by si však vyžadovalo teóriu mysle a ďalšie netriviálne predpoklady.

explicitného prenosu významov sa zaoberá aj A. Smith [28]. V jeho modeli agenty komunikujú o objektoch simulovaných vektormi náhodných numerických hodnôt. Každý agent si vytvára vlastné významy hraním diskriminačných hier. Ani jeden z týchto modelov neumožňuje vnímanie a reprezentáciu dynamických udalostí potrebných pre modelovanie emergencie slovíes.

V modeli Steelsa [26] sú významy prepojené s cieľmi, plánovaním a akciami a sú dostatočne bohaté pre emergenciu komplexných gramatických štruktúr. Sémantická reprezentácia slovíes a mnohých ďalších komplexných jazykových štruktúr je rozpracovaná v rámci paradigmy Embodied Construction Grammar (ECG) [3], je však používaná len na modelovanie produkcie a porozumenia jazyka a ontogenetického osvojovania si jazyka, nie na glosogenetické premeny jazyka.

4.2 Čo ďalej

Cieľom tejto práce bolo navrhnúť novú architektúru pre jazykové experimenty s dôrazom na kognitívne reprezentácie a procesy. Agenty si vytvárajú reprezentácie prostredia na základe svojej individuálnej senzomotorickej skúsenosti a používajú ich ako významy v komunikácii. Architektúra sa vyhýba niektorým zjednodušeniam doteraz publikovaných modelov a môže byť použitá v rámci oboch paradigmatických prístupov k modelovaniu vývoja jazyka kultúrnym prenosom – samoorganizácie a iterovaného učenia. Možno ju aplikovať na modelovanie predverbálnej kognície, osvojovania si jazyka, emergencie komplexného lexikónu zahŕňajúceho slovesá a premien jazyka.

Funkčný systém autonómnych agentov dosahujúcich ciele v dynamickom prostredí, vytvárajúcich jeho reprezentácie a komunikujúcich o nich môže mať aj cenné praktické aplikácie. Bohaté sémantické kategórie spolu s verbalizačnými stratégiami a iterovaným učením sú dobrým prostriedkom na modelovanie emergencie gramatiky, čo je témou pre budúci výskum. V súčasnosti navrhnutú architektúru implementujeme, potom bude nasledovať testovanie navrhnutých stratégií a hypotéz.

Aj keď je nami navrhnutý model zložitejší ako paradigmatické modely vývoja jazyka, stále má veľmi ďaleko od modelovania jazykovej kompetencie na úrovni ľudí. Modelované reprezentácie zodpovedajú konkrétnym pozorovateľným kategóriám základnej úrovne [22], vytváranie abstraktnejších reprezentácií v modeli zatiaľ úplne chýba. Hoci je náš model v mnohom zjednodušený, principiálne umožňuje rozšírenia na modelovanie mnohých zaujímavých problémov kognitívnej sémantiky.

PodĎakovanie

Ďakujem všetkým kolegom zo štvrtkových seminárov za trpezlivé počúvanie, plodné diskusie a komentovanie prvých verzií tohto článku. Šaňovi Moravčíkovi patrí vďaka za nápad uchovávať s každým prototypom jeho vlastné váhy. Projekt bol podporený grantom VEGA 1/0172/03.

Literatúra

- [1] Assad, J. A.: Neural coding of behavioral relevance in parietal cortex. *Curr. Opin. Neurobiol.* **2** (2003) 194-7.

- [2] Baldwin, D. A., Baird, J. A.: Action analysis: A gateway to intentional inference. P. Rochat (ed.), *Early social cognition*, Hillsdale, NJ: Lawrence Erlbaum Associates 1999, 215-240.
- [3] Bergen, B., Chang, N.: Embodied construction grammar in simulation-based language understanding. In J. O. Ostman, M. Fried (eds.), *Construction Grammar(s): Cognitive and Cross-Language Dimensions*. Johns Benjamins 2003.
- [4] Bodík, P., Takáč, M.: Formation of a Common Spatial Lexicon and its Change in a Community of Moving Agents. *Frontiers in AI: Proceedings of SCAI '03*, IOS Press 2003.
- [5] Gärdenfors, P.: Cued and Detached Representations in Animal Cognition. *Behavioral Processes*, **35** (1996) 263-273.
- [6] Gärdenfors, P.: Language and the Evolution of Cognition. V. Rialle, D. Fissette (eds.): *Penser l'esprit: Des sciences de la cognition? une philosophie cognitive*, Grenoble: Presses Universitaires de Grenoble 1996.
- [7] Gärdenfors, P.: *Conceptual Spaces*. Cambridge, MA: MIT Press 2000.
- [8] Gentner, D.: Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. S. Kuczaj (ed.), *Language Development: Language, cognition, and culture*. Hillsdale, NJ: Lawrence Erlbaum 1982.
- [9] Geuder, W., Weisgerber, M.: Domains, Concepts, and Conceptual Spaces. F. Schmalhofer, R. M. Young, G. Katz (eds): *Proceedings of EuroCogSci 03, the European Cognitive Science Conference*, Mahwah/New Jersey/London: Erlbaum 2003, 388.
- [10] Gottlieb, J. P., Kusunoki, M., Goldberg, M. E.: The representation of visual salience in monkey parietal cortex (Letter to Nature) *Nature* (1998) 391-481.
- [11] Harnad, S.: The Symbol Grounding Problem. *Physica D* **42** (1990), 335-346.
- [12] Hulth, N., Grenholm, P.: A Distributed Clustering Algorithm. *Lund University Cognitive Studies* **74** (1998).
- [13] Kirby, S., Hurford, J.: The Emergence of Linguistic Structure: an Overview of the Iterated Learning Model. Parisi, D., Cangelosi, A. (eds.): *Computational Approaches to the Evolution of Language and Communication*. Berlin: Springer-Verlag 2001.
- [14] Koukolík, F.: *Lidský mozek: funkční systémy, norma a poruchy*. Praha: Portál 2000, 54.
- [15] Kuipers, B.: The spatial semantic hierarchy. *Artificial Intelligence*, **119** (2000, 1-2) 191-233.
- [16] Kuipers, B., Modayil, J.: Towards Bootstrap Learning for Object Discovery. *AAAI-2004 Workshop on Anchoring Symbols to Sensor Data* 2004.
- [17] Markman, E.: Constraints on word learning: Speculations about their origins and domain specificity. Gunnar, M. R., Maratsos, M. (eds.): *Modularity and constraints in language and cognition*. Hillsdale: Lawrence Erlbaum 1992.

- [18] Nosofsky, R. M.: Similarity scaling and cognitive process models. *Annual Review of Psychology*, **43** (1992) 25-53.
- [19] Piaget, J., Inhelder, B.: *Psychológia dieťaťa*. Bratislava: Vydavateľstvo SOFA 1997. (Originál: *La Psychologie de L'enfant*, Paris: PUF 1966.)
- [20] Pye C., Loeb D., Redmond S., Richardson L.: When Do Children Acquire Verbs? E. V. Clark (ed.) *The Proceedings of the Twenty-sixth Annual Child Language Research Forum*, Stanford Center for the Study of Language and Information 1995, 60-70.
- [21] Rizzolatti, G., Fadiga, L., Gallese, V., Fogassi, L.: Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, **3** (1996) 131-141.
- [22] Rosch, E.: Principles of categorization. E. Rosch, B. Lloyd (eds.), *Cognition and Categorization*. Hillsdale, NJ: Lawrence Erlbaum 1978.
- [23] Rubin, E.: Figure and ground. Beardslee, D., Wertheimer, M. (eds., trans.): *Readings in perception*. Princeton, NJ: Van Nostrand 1958, 35-101. (Originál: Synoplevde Figurer (1915), Kopenhagen, Gyldendalske.)
- [24] Rybár, J.: *Úvod do epistemológie Jeana Piageta*. Bratislava: IRIS 1997, 84.
- [25] Steels, L., Kaplan, F.: Situated grounded word semantics. T. Dean (ed.), *Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence*. San Francisco: Morgan Kauffmann 1999.
- [26] Steels, L.: The emergence of grammar in communicating autonomous robotic agents. W. Horn (ed.), *Proceedings of ECAI 2000*. Amsterdam: IOS Press, 2000.
- [27] Steels, L.: Language as a complex adaptive system. M. Schoenauer (ed.), *Proceedings of PPSN-VI*. Berlin: Springer-Verlag 2000.
- [28] Smith, A.: *Evolving Communication through the Inference of Meaning*. PhD práca, Theoretical and Applied Linguistics, The University of Edinburgh 2003.
- [29] Smith, K.: *The Transmission of Language: models of biological and cultural evolution*. PhD práca, Theoretical and Applied Linguistics, The University of Edinburgh 2003.
- [30] Spelke, E. S.: Principles of object perception. *Cognitive Science*, **14** (1990) 29-56.
- [31] Takáč, M.: *Kvalitatívne modelovanie a simulácia*. Bratislava: Vydavateľstvo Univerzity Komenského 2003.
- [32] Tomasello, M.: *First verbs: A case study of early grammatical development*. Cambridge: CUP 1992.
- [33] Tomasello, M., Farrar, J.: Joint attention and early language. *Child Development* **57** (1986) 1454-1463.
- [34] Tomasello, M., Merriman, W. E. (eds.): *Beyond Names for Things: Young Children's Acquisition of Verbs*. Mahwah, NJ: Lawrence Erlbaum Associates 1995.
- [35] Tversky, A.: Features of similarity. *Psychological Review*, **84** (1977) 327-352.

- [36] Ungerleider, L.G., Mishkin, M.: Two Cortical Visual Systems. Ingle, D. J., Goodale, M. A., Mansfield, R. J. W. (eds.), *Analysis of Visual Behavior*, The MIT Press 1982.
- [37] Vogt, P.: Iterated learning and grounding: from holistic to compositional languages. *Proceedings of Language evolution and computation workshop ESSLLI-03*, TU Wien 2003, 76–86.
- [38] Werner, G., Dyer, M.: Evolution of communication in artificial organisms. C. Langton, C. Taylor, J. Farmer, S. Rasmussen (eds.) *Artificial Life II*. Addison-Wesley: Redwood City, CA 1992, 659-687.
- [39] Zacks, J. M., Tversky, B.: Event structure in perception and conception. *Psychological Bulletin*, **127** (2001) 3-21.