

Reprezentace významu pomocí multimodálních reprezentací

Michal Vavrečka

Katedra psychologie, Fakulta Sociálních studií Masarykovy Univerzita
Joštova 10, Brno
Katedra kybernetiky, Fakulta elektrotechniky ČVUT
Karlovo náměstí 13, Praha
vavrecka@fss.muni.cz

Abstrakt

Príspevek je shrnutím teoretických předpokladů z oblasti reprezentace významu, které jsem aplikoval v modelu ukotvení symbolů realizovaném v disertační práci. Teoretická analýza předcházející jeho realizaci prokázala, že v oblasti umělé inteligence bývají slučovaný a často zaměňovány konceptualistická a formální teorie významu. Proto je nutné provést jejich přesnější rozlišení. Cílem příspěvku je identifikovat rozdíly ve dvou základních přístupech k významu a pomocí realizovaného modelu ukázat, jakým způsobem a za jakých podmínek lze přístupy sjednotit do společného rámce. Jednotný model je založen na samoorganizujících se neuronových sítích a tvorbě multimodálních reprezentací, sloužících k identifikaci symbolické a konceptuální úrovně znaku. Hlavním přínosem článku je detailní analýza procesu ukotvení symbolů a zdůraznění problematických míst v současných přístupech.

1. Teoretická východiska

Návrh modelu je založený na teoretických principech kognitivní sémantiky, a proto bych chtěl v úvodu shrnout, základní východiska, které daný přístup zastává. Kognitivní sémantika je brána jako oblast kognitivní lingvistiky. Reprezenčním mechanismem je tvorba konceptů, které tvoří centrální oblast tohoto přístupu. Základní východiska se dají shrnout do tří základních bodů:

- 1. Gramatika je konceptualizace*
- 2. Konceptuální struktura je vtělená a závislá na motivovaném užití.*
- 3. Schopnost používat jazyk je založena na obecných kognitivních schopnostech a nikoliv na specializovaném jazykovém modulu.*

Kognitivní sémantika (podobně jako kognitivní lingvistika) odmítá dělení analýzy jazyka na syntax, sémantiku a pragmatiku. Nahlíží tradiční rozdělení syntax – sémantika jako kontinuum a nepředpokládá samostatnou lexikální složku. Sémantika v kognitivně

lingvistickém podání zahrnuje také vše, co se tradičně popisuje jako pragmatika.

Všechny lingvistické vrstvy fungují společně v jednom systému a nelze je interpretovat samostatně. Oblast sémantiky se dělí na konstrukci významu a reprezentaci znalostí. Na rozdíl od analytických přístupů, je kognitivní sémantika založená na konceptuálním přístupu k významu. Teorie založená na pravdivostních hodnotách má totiž omezenou explanační sílu a nedokáže nám zajistit plné pochopení vět. Tento způsob reprezentace významu není v kognitivní sémantice odmítán, ale není brán jako absolutní a jediný. Cílem je totiž nalézt reprezentaci významu, který nám umožní zachytit veškeré nuance jazyka. Základním rozdílem je vnímání významu jako dynamického a tvořeného na základě konstrukce a konvence. Významy jsou tedy budovány jako koncepty procesem učení a může docházet k jejich rekonstruování.

Kognitivní sémantika jde ve svých základech hlouběji než pojem dostatečných a nutných podmínek, se kterým se setkáme v souvislosti s teorií pravdivosti. Význam slov je zde chápán v kontextu naší tělesné a kulturní zkušenosti. Způsob jakým jsou zachyceny asociační vazby mezi těmito koncepty nemůže být zachycen analýzou nutných a dostačujících podmínek. Pravdivostní hodnoty navíc nedokáží zachytit asymetrie ve způsobu užití slov. Pokud se podíváme na významy slov chlapec a dívka, přestože je můžeme definovat na základě pohlaví a věku, nalezneme v oblasti jazykového užití značné rozpory ve vymezení užití těchto slov, které jsou založeny na mnoha aspektech, které nedokážeme pomocí nutných a dostatečných podmínek zachytit. Základní principy kognitivní sémantiky shrnuje kompendium Leonarda Talmyho (2003) prezentující obor jako celistvý přístup k jazyku z hlediska pojmových struktur. Talmyho práci lze jejím významem považovat za paralelní Langackerově kognitivní gramatice (1987), přístupující se stejnou motivací k problému vztahu významu a formy, s důrazem na reprezentaci významu. Cílem zmíněného oboru je nalezení jednotného organizačního principu, který by spojoval aspekty sémantické s formálními. Pro lepší rozlišení uvádím

jednotlivé přístupy (formální, psychologický a lingvistický a počítačový), abychom přesněji určili oblasti, které je nutné sjednotit.

Formální přístupy se soustřeďují na strukturální vzorce lingvistických forem, které jsou abstraktní a považovány jako nezávislé na asociovaných významech. Základem je studium syntaxe a morfologie. Klasickým příkladem může být generativní gramatika. Sémantika takových systémů je odlišná od kognitivní lingvistiky či psychologického pojetí významu.

Psychologické přístupy zkoumají jazyk z hlediska obecných kognitivních systémů jako je vnímání, paměť, pozornost a uvažování. Psychologie zkoumá jazyk a tvorbu významů z hlediska konceptuálního. Centrem zájmu je sémantická paměť, asociace mezi koncepty, struktura kategorií, tvorba inference apod. Je však nutné zdůraznit, že v psychologii je dosud nedostatečně zpracován integrovaný strukturální systém, pomocí kterého jazyk organizuje konceptuální obsah, což jsou vlastnosti, které jsou centrální pro předchozí přístup. Možným řešením je užití kognitivního modelování, které nám umožní lépe napodobit některé procesy související s reprezentací významu. Příkladem je prezentovaný model.

V kognitivní lingvistice je pozornost zaměřena na strukturu a procesy, kterým je konceptuální obsah organizován v jazyce. Jedná se o lingvistickou strukturaci základních konceptuálních kategorií (čas a prostor, scéna a událost, objekty a procesy, pohyb a umístění, dynamika a kauzace apod.). Zajímá se o sémantickou i syntaktickou strukturu. Jedná se tedy o integraci předchozích přístupů do společného rámce.

Z hlediska počítačového přístupu k reprezentaci významu jsme se v oblasti umělé inteligence setkávali se způsobem, který je založen na čistě formálních principech. Jak bylo zmíněno v souvislosti se sémantikou formálních systémů, je takový způsob reprezentace redukcionistický. Čistě symbolické přístupy bývají používány v reprezentačních systémech založených na seznamech, rámcích, schématech či sémantických sítích. Pomocí takového způsobu reprezentace dokáží vytvořit plně funkční symbolický systém s kombinatorickou syntaxí a sémantikou, který podporuje vyšší kognitivní funkce, jako jazyk, paměť, znalosti. Zastánci daného přístupu jsou např. Newell a Simon (1972), Fodor (1975), Pylyshyn (1984) apod. Soustřeďují svůj zájem pouze na formální sémantiku pravdivostních hodnot a zcela pomíjejí způsob propojení symbolické roviny s rovinou empirickou.

Používání čistě symbolického systému jako jediného způsobu reprezentace znalostí se setkává s neschopností zachycení významu. Narážíme na potřebu existence referenčního systému pro dostatečnou sémantickou reprezentaci.

1.1 Multimodální reprezentace

Možným přístupem, který vychází z principu kognitivní sémantiky je perceptuální symbolický systém (Barsalou, 1999). Patří do oblasti konceptuálního přístupu ke kognici, která narozdíl od čistě symbolického systému, zdůrazňuje funkci perceptuálně získaných, a subsymbolických tvořených konceptů, které jsou použity, společně se symbolickou vrstvou, k reprezentaci významu.

Centrálním pojmem, vyskytujícím se v souvislosti s touto teorií jsou multimodální reprezentace. Jak již napovídá samotný název, je základním požadavkem pro jejich tvorbu reprezentace významu napříč sensorickými modalitami. Barsalou hovoří o specifických oblastech, integrujících více modalit dohromady. Reprezentace jsou multimodální, protože vznikají na základě např. sensorických modalit, propriorecepce a introrecepce. Integrace modalit se objevuje již u Damasia (1989), který hovoří o zónách konvergence, které integrují informace ze senzomotorických map, čímž dochází k reprezentaci map. Konvergenční zóny pomocí hierarchických množin asociálních oblastí integrují specifické modalitní do společného rámce. Jedná se o tvorbu mentálních představ napříč modalitami, přičemž mentální představy jsou nazírány jako vědomé a proces jejich tvorby jako nevědomý. (Barsalou, 2003)

Damasio (1989) také tvrdí, že není rozdíl mezi perceptuální a kognitivní pamětí. Místo toho hovoří o gradientu, z posteriorní do anteriorních asociálních oblastí, jež vyjadřuje míru komplexity a specifity paměťových obsahů, vzhledem k síle aktivace senzomotorických oblastí. Damasio říká, že paměťové oblasti jak uvnitř či vně senzomotorického systému se podílí na tvorbě kognitivních reprezentací. (Barsalou, 1999)

2 Východiska pro návrh modelu

Při návrhu modelu vycházím z principů kognitivní sémantiky, přičemž základní inspiraci nacházím v Peircově teorii znaku (1931-35). Jedná se o tvorbu znaku, který referuje k externímu objektu pomocí konceptu (interpretant) a arbitrární symbolické vrstvy (representamen). Konceptuální úroveň je vytvářena na základě perceptuálních vstupů, což je v souladu s perceptuální teorií kognice. Tím se přístup liší od Saussureovy teorie znaku (1965) a také logiky v rámci analytické filosofie (Tarski, 1944). Oproti formálnímu symbolickému systému a formální sémantice je tento systém odlišný užitím perceptuálně získané konceptuální úrovně. Model je v teoretické úrovni řešením problematiky ukotvení symbolů (Harnad, 1990), která se

zabývá způsobem převodu perceptu do konceptuální úrovně a její následné propojení s arbitrárním symbolickým označením. Jeden z možných způsobů převodu, který je inspirován teorií perceptuálního symbolu je tvorba multimodálních reprezentací, tzn. integrace informací z různých sensorických modalit do společného rámce. Podobnou inspiraci nacházíme také v teorii duálního kódování (Paivio, 1986). Systém nejprve primárně zareprezentuje auditivní a vizuální vstupy a poté je integruje do společné multimodální vrstvy. Následně nás zajímá otázka, který vstup se na strukturaci multimodální vrstvy podílí.

V oblasti psychologie a lingvistiky se totiž setkáváme v souvislosti se propojením symbolické a konceptuální roviny s otázkami, která z těchto úrovní je hierarchicky výše, tzn. privilegovanější. Jedná se o spory ohledně privilegovanosti lingvistické či nelingvistické reprezentace (Landau, Hofmann, 2005), zdůrazňující nadřazenost konceptuální, popřípadě symbolické úrovně. Použitím multimodální vrstvy, která tyto dvě úrovně integruje, máme zajištěny podmínky pro analýzu způsobu jejich činnosti.

V realizovaném modelu vnímání prostoru (Vavrečka, 2007) je pro strukturaci důležitá auditivní (lingvistická) informace, která je jednoznačná (kategorie se nepřekrývají), narozdíl od vizuální, jejíž kategorie se překrývají a jsou zatíženy mírou neurčitosti. V případě rozpoznávání objektů tomu může být opačně, jelikož pomocí jednoznačného tvaru lze odlišit slova, která jsou v symbolické úrovni identická, přestože nabývají odlišných významů (polysémie). Proto není možné tuto otázku zjednodušovat pouze na konkrétní oblast, ve které provádíme testování.

2.1 Vnímání prostoru

Jelikož model slouží k reprezentaci prostoru, je potřeba zmínit detaily, které s touto oblastí souvisí. V rovině teorie vycházím z Coventryho FGR (Functional Geometric Framework) teorie (2004), která odvozuje prostorové vztahy na základě geometrických a extra-geometrických vztahů, a také znalosti objektu. Jedná se rozlišení tří základních kategorií, spolupodílejících se na identifikaci prostorového označení mezi objekty. Geometrické vztahy souvisí s prostorovou polohou objektu. Extra-geometrické doplňují informace o gravitaci a silách působících mezi objekty. Subsystém znalost objektu dodává informace o vlivech kontextu a dalších aspektů, které také determinují označení prostorového vztahu. Dohromady tyto subsystému tvoří jednotný funkční geometrický rámec (FGR), pomocí kterého můžeme reprezentovat většinu prostorových vztahů.

V oblasti vnímání prostoru je také nutné vzít v potaz referenční rámce. Na jejich základě určujeme počátek a natočení souřadného systému (centrum deixie), sloužící k následnému posouzení prostorových vztahů objektů. V základním dělení hovoříme o vnitřním (centrem deixie je objekt zájmu), relativním (centrum deixie je pozorovatel) a absolutním (deiktické osy tvořené na základě okolního prostředí) referenčním rámci.

V současné fázi realizovaný model řeší pouze geometrické prostorové vztahy dvou objektů v relativním referenčním rámci, bez přesné identifikace tvaru objektů. V další fázi bude model rozšiřován v souladu s FGR teorií, pomocí přidávání modulů, které jsou schopny reprezentovat extra-geometrické vztahy, znalost objektu a dokáží pracovat také v absolutním a vnitřním referenčním rámci.

3 Realizace modelu

Z hlediska integrace modalit jsou v popředí zájmu realizovaného modelu auditivní a vizuální vstupy. Jelikož auditivní informace má vlastnosti symbolické úrovně (je arbitrární a jednotlivé slova jsou jednoznačně identifikovatelná) a vizuální informace naopak obsahuje informaci o externím prostředí a je velmi neurčitá (fuzzy), vznikne jejich integrací multimodální reprezentace, ve které jsou spojeny vlastnosti obou. Model tedy dokáže vytvořit primární reprezentace jednotlivých sensorických vstupů a následně vytvořit společnou reprezentaci. Z hlediska teorie zpracování informace se jedná o výpočetně náročné procesy, přičemž je obtížné určit algoritmičnou posloupnost zpracování. Proto je použito metody modelování pomocí učicích se sítí, které umožní umělému systému tvorbu reprezentací na základě interakce s prostředím. Při adekvátním nastavení počtu neuronů a metod učení můžeme docílit reprezentační úrovně blízké lidským schopnostem (Vavrečka, 2007).

3.1 Architektura systému

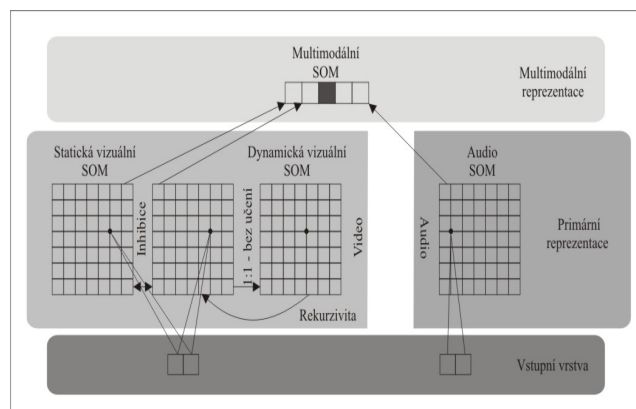
Architekturou pro realizaci modelu byly zvoleny neuronové sítě a způsobem *učení bez učitele*. Důraz je kladen na *učitelský přístup* (Ziemke, 1999), jehož cílem je eliminovat vliv designéra, tvořícího fixní apriorní systémy, které nelze během vývoje modelu modifikovat. Právě proto jsou neuronové sítě vhodnou alternativou. Také užití *učení bez učitele* zaručuje, že systém je vytvářen, a jeho reprezentace strukturovány, na základě informací z prostředí, což opět omezuje fixnost architektury a reprezentačního systému. Poté, co bude model rozšířen o motorickou část a systém efektorů, které se naučí mapovat své reprezentace do multimodální

vrstvy, stane se systém plnohodnotným zástupcem enaktivního přístupu (Varela, 1991). Protože v oblasti kognitivních věd neexistuje jednota přístupů ke konstrukci umělých inteligentních systémů, a jednotlivé proudy používají rozličná teoretická východiska, jsou v následující tabulce shrnuty použité principy.

<i>Teoretický rámec</i>	Symbolický přístup (analytický)	Konceptuální přístup (empirický)
<i>Forma reprezentací</i>	Symbolická	Subsymbolická
<i>Způsob realizace</i>	Kognitivismus	Konekcionismus
<i>Způsob tvorby systému</i>	Designérský	Učitel'ský

Tab.1 Tabulky použitých přístupu (zvýrazněné)

Architektura využívá jako základní prvek pro vizuálního i auditivního reprezentační vrstvu vlastnosti samoorganizujících se map (Kohonen, 1989). Pro vizuální systém použil zmíněný typ sítě již Riga (Riga, Cangelosi a Greco 2004). V jeho případě však mapy uchovávaly invariantní reprezentace objektu a barev, nikoliv prostorových vztahů. V navrhovaném modelu jsou navíc samoorganizující se mapy (SOM) použity jak pro vizuální informace, tak pro zvukové vstupy. Právě reprezentace zvukového vstupu, který má strukturu jazykové promluvy a lze ji nazírat jako symbolickou, je slabinou předchozích modelů. Jak Riga (2004), tak Regier (1996) tuto část systému vytvářejí redukcionisticky. Regierův model auditivní vrstvu vůbec neobsahuje. Prostorová (jazyková) označení jsou výstupní vrstvou (výsledným produktem činnosti neuronové sítě), což můžeme interpretovat následovně. Model buď symbolickou vrstvu apriorně obsahuje, nebo dochází k auditivnímu předzpracování a slova nemají žádné auditivní vlastnosti. Tím ale pomíjíme jejich potenciál a výhody (podrobněji níže). Rigaův model auditivní vstup obsahuje, ale je reprezentován jako arbitrární řetězec, asociovaný na perceptuální kategorie pomocí *učení s učitelem*. Takový způsob ukotvení opět plně nevyužívá potenciál auditivního vstupu. Proto je základní návrh vlastního modelu odlišný od těchto řešení (viz. Obr.1). Z hlediska použité architektury se blíží Dorffnerovu modelu (1996), vycházející z principu *radikálního konekcionismu*, tedy použití neuronových sítí v celém systému.



Obr.1 Základní schéma modelu. Audio a video vstup je reprezentován odděleně pomocí SOM sítí. Následně jsou výstupy integrovány v multimodální vrstvě. Dynamické scény jsou reprezentovány ve vizuální dynamické SOM síti. Inhibice vizuálních vrstev není dosud realizována.

3.1.1 Audio vstup

Podívejme se nejdříve podrobněji na možnosti, jak zarepresentovat auditivní informaci. U předchozích badatelů (Riga, 2004; Cangelosi, 2006) se jednalo o sekvence arbitrárních symbolů, které nahrazovaly slova. V realizovaném modelu jsem se od takového přístupu odklonil a uvažoval o reprezentaci auditivní informace jako n-dimenzionálního vektoru fonetických vlastností. Pro stanovení optimálního počtu dimenzí jsem vycházel z následujících úvah.

V případě jednodimenzionálního vstupu zdůrazňujeme arbitrárnost symbolů a blížíme se výše zmíněným přístupům. Abstrahujeme od schopností slov být asociovány (vytvářet strukturu na základě podobnosti). Tím mám na mysli schopnost vybavit si slova s podobnou strukturou (pes, les, ves) bez velké počítačové a časové náročnosti. Pouze jednodimenzionální reprezentace neumožňuje vznik asociace na základě zvukové (tvarové) podobnosti ani na základě kauzální blízkosti (výskyt slov vedle sebe ve větě). Taková reprezentace není psychologicky plausibilní a v dalších úrovních zpracování zvyšuje výpočetní náročnost. V případě dvoudimenzionálních vstupů, které byly během realizace modelu použity, už můžeme reprezentovat rozdílné vlastnosti a vytvářet asociativní vazby. U vícedimenzionálních vektorů lze zachytit další akustické, ale i lingvistické vlastnosti, např. společný výskyt slov v promluvě. Přesto nejsou lingvistické vlastnosti v základní verzi využity. Důvodem je větší přehlednost a snadnější interpretovatelnost modelu z hlediska procesu ukotvení symbolů.

V navrhovaném modelu se snažím pracovat se zvukovým vstupem podle následujících pravidel. Podobně jako je vizuální vstup kategorizován podle topologických vlastností, je audio vstup také nejdříve reprezentován

topologickou mapou. Vycházím z předpokladu, že ze zvukového vstupu, který je strukturován jako jazyková promluva, lze bez apriorní sémantické či syntaktické znalosti získat mnoho informací, které nám mohou pomoci při tvorbě primární reprezentace. Pokud vycházím z předpokladu, že malé dítě dokáže v prvních fázích analyzovat promluvu na úrovni slov, nabízejí se následující možnosti. Bez předchozí znalosti lze členit slova podle zvukové podobnosti, podle délky, či podle frekvenční křivky, neboť informace je v promluvě bezprostředně obsažena. Přestože se může jevit, že vlastnosti neusnadňují proces integrace zvukových reprezentací s vizuálními, není tomu tak. Jestliže přistupujeme k otázce ukotvení symbolů s požadavkem minimálního zásahu designéra (tzn. žádné externí interpretace vstupu), je nutné využít vlastností vstupů při tvorbě primárních reprezentací v co největší míře. Reprezentujeme-li auditivní vstup samostatně v primární auditivní reprezentaci, bylo by krátkozraké redukovat veškeré vlastnosti, které jej charakterizují, jelikož bychom při použití principu samoorganizace ztratili informace, sloužící k tvorbě jednoznačných kategorií. I když se nabízí možnost reprezentovat jazykovou promluvu jako arbitrární shluk symbolů, pokusíme se uvažovat odlišně. Přestože samotná promluva není sémantická (k ničemu neodkazuje), ve způsobu svého kódování obsahuje informace (jak bylo zmíněno výše), které nám pomohou při její reprezentaci, ale i v následných transformacích. V našem zjednodušeném modelu, vyjádříme jednotlivé slova 2D vektorem vyjadřujícím frekvenční křivku a časovou délku daného označení .

3.1.2 Vizuální vstup

Vizuální informace jsou reprezentovány také pomocí SOM sítě. Takto vytvořená vrstva je schopna pomocí laterálně propojených neuronů vytvořit prostorově uspořádaná prototypická centra, citlivá na vstupy určitých hodnot. Vstupem je dvourozměrný vektor, reprezentující prostorovou polohu předmětu.

V modelu abstrahuji od prvotního zpracování informace, které lze snadno vytvořit pomocí umělé retin. Jedná se ale pouze o architektonické zjednodušení. Systém by mohl mít na vstupu umělou sítnici, na kterou by přicházely obrazy z prostředí, ale pro simulaci je postačující vektorový vstup, jelikož nedochází k redukci informace o signálu z prostředí.

Ve realizované verzi je také oproti Regierovu modelu (1996) odlišně řešeno zpracování sekvencí. Na rozdíl od identifikace dynamické scény ve „vrchních“ vrstvách sítě, je v modelu vytvořen duální systém zpracování již v prvotní úrovni. V případě statické scény je informace zpracována samoorganizující se mapou citlivou na polohu objektu. V případě dynamického scény jí zpracovává RecSOM mapa (Voegtlin, 2002), obsahující

neurony citlivé na pohyb. Ty pomocí rekurzivity dokáží klasifikovat dynamické prostorové vztahy (nad, pod, ven, skrz, okolo).

3.1.3 Multimodální vrstva

Jádrem systému je multimodální vrstva, která zajišťuje proces identifikace, tedy tvorbu jednoznačné kategorie a zároveň její reprezentace. V případě modelu se jedná o nalezení vhodného propojení mezi primárním auditivním a vizuálním výstupem, který je multimodální vrstvě prezentován v jeden okamžik. Cílem učení na základě těchto dvou informací je nalézt kategorie, které reprezentují společné vlastnosti těchto vstupů. Rozdíl oproti klasickému Harnadovu přístupu je podrobně rozebrán v následující kapitole. V teoretické rovině je výsledkem analýzy modelu odhalení principu *privilegovanosti jednoznačného vstupu*, což je nový přístup k řešení otázky privilegovanosti lingvistické a nelingvistické reprezentace, podle kterého není privilegovaná reprezentace apriorně daná, ale závisí na aktuálních vstupech z prostředí. V obecné rovině se jedná o příspěvek do diskuse ohledně platnosti Sapir-Whorfovy hypotézy (Whorf, 1956).

4 Analýza modelu

V této části se pokusím na realizovaný model podívat z pohledu teoretických východisek, která byla zmíněna v souvislosti s kognitivní sémantikou, tedy konkrétně s perceptuální teorií kognice a konceptuálními přístupy k reprezentaci znalostí.

Již ze základního popisu modelu je jasné, že je navržen v souladu s požadavky na tvorbu konceptuální úrovně na základě perceptuálních vstupů. Také užití multimodálních reprezentací znamená příklon k teoriím podobným perceptuálnímu symbolickému systému (Barsalou, 1999). Jelikož jsou perceptuální symbol a multimodální reprezentace definovány pouze teoreticky a dosud se v oblasti modelování neobjevil pokus o jejich formalizaci a převod do úrovně simulace, pokusím se jednotlivé části svého systému interpretovat v kontextu těchto teorií, ale s cílem přesněji identifikovat jejich funkci. Zároveň se budu snažit interpretovat model v termínech souvisejících s ukotvením symbolů či Peircovou teorií znaku. Přestože tyto teorie používají odlišné pojmy, je možné nalézt sjednocující prvky.

Z hlediska teorie perceptuálního symbolu (Barsalou, 1999) můžeme v modelu identifikovat pouze multimodální reprezentaci, která společně reprezentuje senzory vstupů. Mentální simulátory nejsou v systému realizovány, jelikož schopnost produkovat „neomezené“ simulace možných stavů na základě perceptuálních vlastností leží za hranicemi současného vývoje v

modelování. Simulátory nejsou ve výchozím článku (Barsalou, 1999) definovány způsobem, umožňujícím jejich formalizaci, takže je obtížné jejich modelování. Proto je třeba nalézt jiné metody, jak s reprezentacemi operovat, a jak je využít v procesech vyvozování či myšlení obecně.

Podívejme se nyní na prezentovaný model z hlediska Peircovy teorie znaku. Jestliže Peirce dělí jednotlivé úrovně znaku na objekt, representamen a interpretant, pokusme se najít jejich koreláty v prezentovaném modelu. V případě externího objektu je nám k dispozici pouze informace o jeho poloze v podobě vektoru dvou souřadnic. Tato vstupní informace je přesto dostatečným zástupcem objektu v externím prostředí, neboť potvrzuje existenci objektu a také jeho umístění ve světě, se kterým systém interaguje.

Obtížnější je v našem případě rozlišení representamen a interpretantu. Podle Peircovy teorie, je významem právě interpretant, sdružující vlastnosti, které jsou pro externí objekt signifikantní. V konceptuálních teoriích odpovídá danému označení koncept, který reprezentuje vlastnosti objektu. Representamen je symbol, arbitrárně přiřazený interpretantu, společně vytvářející znak, který je následně užít v procesech vyvozování.

Podle teorie ukotvení symbolů vzniká koncept procesy diskriminace perceptuálních informací z prostředí a symbolická informace je arbitrární označení, který je procesem identifikace propojen s konceptem.

4.1 Diskriminace v symbolické úrovni

V případě hybridního systému jsou procesy diskriminace subsymbolické a identifikace přiřadí subsymbolickému konceptu symbolickou „nálepkou“. V takto koncipovaných modelech (Regier, 1996; Harnad, 1990) vzniká apriorní předpoklad, že symbolická úroveň je přiřazována konceptu přímo „uvnitř“ systému, tzn. že pomocí ní designér „označí“ vytvořený konceptu. Tyto přístupy vycházejí z předpokladu, že uvnitř systému je s reprezentací pracováno na základech symbolického systému a proto je klíčovou právě symbolická vrstva. Faktem zůstává, že při tvorbě obdobných systému autoři zjednodušují proces akvizice symbolické vrstvy, kterou na základě designérského přístupu vkládají přímo do systému. Pro účely základních metod ukotvení je tato metoda dostačující, ale unikají nám některé důležité aspekty akvizice symbolické roviny pomocí auditivního vstupu. Takový přístup je zároveň v rozporu s *učitelským přístupem*, jelikož symbolickou úroveň do systému apriorně vkládáme, místo aby vznikla na základě učení. Odklon od podobného způsobu tvorby symbolické úrovně můžeme nejlépe demonstrovat na realizovaném modelu. V prvé řadě se jedná o způsob akvizice symbolické roviny a její reprezentace. Jestliže v předchozích modelech dochází k jejímu přiřazení, je v

popsaném modelu tento způsob zajištěn pomocí učení ze vstupní vrstvy. Podobný přístup nacházíme u modelu *přenosu ukotvení* (Riga, Cangelosi a Greco 2004). V jejich případě byla vizuální vstup učen na základě SOM sítě, pomocí které byly vytvořeny subsymbolické reprezentace objektů, tzn. koncept. Následně jsou tyto subsymbolické reprezentace asociovány se symbolickým auditivním vstupem pomocí MLP (Multi level perceptron). Druhá fáze tedy probíhá na základě *učení s učitelem*. V našem případě je první část architektury shodná, ale auditivní vrstva je trénována na základě *učení bez učitele*. Následné multimodální reprezentace jsou oproti Rigovu modelu opět tvořeny na principu samoorganizace (SOM). Realizovaný model je tedy posunem k *učitelskému přístupu*, jelikož je celý založen samoorganizujícími se sítěmi. Důležité změny jsou také v chápání konceptuální a symbolické úrovně.

V realizovaném modelu je konceptuální úroveň, která souvisí s reprezentací významu (interpretant), uložena v primární vizuální síti. Pomocí ní jsme schopni detekovat polohu objektu v prostoru, tzn. uchovávat konstantní vlastnosti polohy. Jedná se tedy o subsymbolický koncept, který vznikl na základě diskriminace objektů v prostoru.

V případě auditivní vrstvy je tomu narozdíl od výše zmíněných přístupů podobně. Přestože auditivní vstup slouží k akvizici symbolické úrovně (jazyková označení), je způsob jeho tvorby stejný jako u vizuálního vstupu. V tom je základní rozdíl oproti klasickým způsobům ukotvení.

Připomeňme si, že základní vlastností symbolů je arbitrárnost, tzn. nemají žádnou podobnou vlastnost s objektem, který reprezentují (tedy i s vlastnostmi, které jsou v souvislosti s objektem zareprezentovány v konceptuální úrovni). Arbitrárnost ovšem neznamená, že mezi symbolickým označením jednotlivých slov, neexistují žádné vzájemné vazby, popřípadě že je nelze reprezentovat v samostatné vrstvě, která dokáže zachytit vztahy podobnosti (v prezentovaném modelu na úrovni zvukové podobnosti, ale je možné zachytit vlastnosti související s reprezentací kontextových informací). Jestliže symbolickou vrstvu do systému nevkládá designér, je potřeba vytvořit způsob její reprezentace. Pro tyto účely slouží v realizovaném modelu primární auditivní vrstva. Přestože jsou symboly arbitrární vzhledem ke konceptům, ke kterým referují, mohou být navzájem propojeny např. podle fonetické blízkosti. V našem případě se jedná o podobnost na úrovni celých slov.

Takový požadavek je založen na psychologických poznatcích. Člověk dokáže volně produkovat slova, která jsou zvukově podobná, aniž by měla jakoukoliv podobnost z hlediska významu (koncepty s podobnými vlastnostmi). Příkladem může být produkce slov „banka,

branka, Hanka“, které jsou významově zcela odlišné, ale člověk je dokáže produkovat velmi rychle, což svědčí o způsobu reprezentace, který nevyžaduje komplexní sémantické zpracování. Proto je primární auditivní reprezentace vytvořeny obdobným způsobem, tzn. blízké jsou slova, které jsou foneticky podobné. Pokud bychom se chtěli přiblížit psychologické plausibilitě ještě více, museli bychom symbolické označení reprezentovat na úrovni slabik. Člověk dokáže kromě produkce foneticky podobných slov, vytvářet také zcela nové označení, která jsou kombinací slabik, ale nereferují k žádnému konceptu ani známému symbolickému označení a lze je považovat za nesmyslná. I tato produkce je natolik rychlá a automatická, že svědčí o samostatné reprezentaci symbolické úrovně, které není propojena s konceptuální úrovní, jelikož její kombinace nereferují k žádnému konceptu. Příkladem je schopnost produkovat slova jako „jala, bala, bola“, aniž bychom je někdy předtím slyšeli. Přesto je pro účely modelu dostačující reprezentace v úrovni slov, jelikož se jedná o zjednodušenou demonstraci primární reprezentace symbolické úrovně a základní analýzu procesu ukotvení symbolů. Hlavním důvodem zvoleného způsobu je schopnost reprezentovat auditivní úroveň samostatně a vytvářen vztahy mezi symbolickými výrazy na základě jejich vzájemné tvarové (zvukové) podobnosti. Dalším důležitým aspektem primární auditivní reprezentace je vzájemná oddělenost jednotlivých slovních označení, tzn. jsou reprezentována tak, abychom je mohli rozdělit do ohraničených a nepřekrývajících se kategorií. Tato vlastnost je velmi důležitá, jelikož zajistí jednoznačné odlišení objektů okolního světa a vytvoření jednoznačného znaku (reprezentaci obsahující konceptuální a symbolickou úroveň, v našem případě integrovanou v multimodální reprezentaci). Během testování modelu pro statické prostorové vztahy (Vavrečka, 2007) se objevují pokusy měnit velikost shluků (tedy fonetických rozdílů) auditivního vstupu, až do úrovně, kdy se odlišné označení překrývají. Taková situace znamená, že bychom na vstupu přijímali slova, která jsou natolik nejednoznačná, že nelze určit, do které kategorie spadají. V případě prostorových vztahů by se jednalo o slova jako „vlelu“ nebo „dovo“, vytvořené průnikem kategorie vlevo a dolů, kdy auditivní kategorie začínají být „fuzzy“. Pokud bychom během učení přijímali vstupy, které vytvářejí překrývající se shluky v auditivní i vizuální vrstvě, není možné naučit systém správně kategorizovat ve vrstvě multimodální. Tento požadavek je nemožný již z teoretického hlediska na základě prosté úvahy. Jestliže máme dva nejednoznačné vstupy, které nelze kategorizovat do oddělených tříd, nelze z nich bez doplňujících informací odvodit jednoznačné kategorie.

Z těchto úvah lze vyvodit několik závěrů, týkajících se vztahu mezi jazykem a obrazem (lingvistickou a nelingvistickou reprezentací popřípadě symbolickou a konceptuální vrstvou) a způsobu, jakým se navzájem ovlivňují a jaká je jejich vzájemná hierarchie. Jedná se o ustavení vztahu mezi koncepty (nelingvistické) a symboly (lingvistické). Již psychologické studie (Landau, Hofmann, 2005) týkající se Williamssova syndromu (genetická vada ovlivňující vnímání prostoru při zachování jazykové produkce) prokázaly, že jednotlivé subsystémy se mohou během činnosti vzájemně doplňovat a nahrazovat deficit jednoho z nich. Z výsledků tedy vyplývalo, že jeden systém není nadřazen druhému, ale že fungují společně. Pomocí zmíněného experimentu (Landau, Hofmann, 2005) ovšem nelze přesněji určit jakým způsobem se systémy vzájemně strukturují. Zkusme se nejprve podívat na problematiku z lingvistického hlediska. V nejobecnější rovině se o vztahu lingvistických a nelingvistických reprezentací vyjadřuje Sapir-Whorfova hypotéza. Jedná se o lingvisticko-antropologická tezi (Whorf, 1956), podle níž je pojetí reálného světa vystavěno na jazykových zvyklostech konkrétní dané komunity, jež pak předurčují určitý výběr interpretace reality. To bývá v oblasti reprezentace znalostí a významu interpretováno, jako nadřazenost lingvistických reprezentací nad nelingvistickými. Z toho plyne hierarchie jednotlivých úrovní reprezentace. Příkladem v oblasti modelování může být mnohokrát zmiňovaný Harnadův model (1990), ve kterých jsou pomocí symbolické vrstvy vytvářeny procesy myšlení, jako logické operace v rámci formálního symbolického systému, tzn. strukturace probíhá pomocí symbolické vrstvy. Odpůrci Sapir-Whorfové hypotézy argumentují (Gumperz, 1996), že myšlení není zcela redukovatelné pouze na jazyk a řada významů může být vyjádřena neverbálními či mimojazykovými prostředky, které se mohou podílet při realizaci myšlenkových operací. Jelikož z těchto poznatků opět nedokážeme určit vzájemný vztah těchto reprezentačních vrstev, pokusím se nyní interpretovat realizovaný model a jeho přínos pro danou oblast v kontextu těchto teorií. V první fázi docházíme k závěru shodně se Sapir-Whorfovou hypotézou, jelikož lingvistické označení strukturuje nelingvistické koncepty. Systém pracuje na základě ohraničených jazykových vstupů (slovo je přesně identifikovatelné) a překrývajících se prostorových kategorií. Z takových výsledků konkrétního modelu můžeme dojít k závěru, že lingvistický systém strukturuje nelingvistický právě pro svou jednoznačnost a ohraničenost. Přesto nelze podobný závěr přijmout a zobecnit pro celou oblast ukotvení symbolů.

Některá jazyková označení totiž bývají polysémická, tzn. jsou propojená s rozdílnými koncepty (např. klíč jako označení rozdílných předmětů). V takovém případě ovšem jazykové označení přesně nevymezuje koncept, ale stává se nejednoznačným. Přestává platit vztah nadřazenosti lingvistické reprezentace nad nelingvistickou. V dané situaci se systém řídí konceptuální nelingvistickou informací, pomocí které strukturuje způsob propojení se symbolickou rovinou. Konceptuální úroveň je totiž pro polysémická slova jednoznačná (koncepty jsou na rozdíl od identického označení rozdílné a lépe odlišitelné) a proto mají během učení větší vliv na kategorizaci. Jelikož neexistují polysémická slova, která by referovala k překrývajícím se konceptům, což je ověřeno v Modelu 1, který nedokáže kategorizovat dva „fuzzy“ vstupy, docházíme k závěru, který popírá privilegovanost jednoho ze systému. Můžeme tedy shrnout.

V oblasti reprezentace významu se podílí na způsobu kategorizace na základě symbolického a konceptuálního vstupu ten, jehož vstupy jsou jednoznačně rozlišitelné, jelikož jednoznačné kategorie determinují způsob tvorby multimodálních reprezentací. Neexistuje předem daná privilegovanost.

Dané tvrzení tvoří jádro navrženého přístupu k reprezentaci významu založeném na multimodálních reprezentacích pomocí systému dvou primárních reprezentací. Apriorní privilegovanost jednoho systému je v rozporu psychologickými poznatky o způsobu reprezentace znalostí u člověka. Naopak pomocí systému, který kategorizuje na základě jednoznačných vstupů dokážeme vysvětlit mnoho teoretických poznatků z oblasti experimentální psychologie.

Jestliže panují rozpory o privilegovanosti systémů, na základě rozdílných výsledků experimentů s odlišně definovanými úlohami (Landau, Hoffman, 2005), lze pro jednotnou interpretaci užít výše popsaného principu privilegovanosti jednoznačného vstupu, jelikož dokáže eliminovat vliv typu úlohy a experimentálních podmínek na výsledky výzkumu. Správnost navržené principu je navíc ověřena konkrétní realizací formou modelu, jehož efektivita fungování verifikuje princip v empirické rovině. Právě provedení a odzkoušení modelu na konkrétní úloze jej činí argumentačně silnější oproti rozporným teoretickým předpokladům. Na druhou stranu je nutné říci, že model byl zkoušen pouze v konkrétní oblasti vnímání prostoru, při které se neobjevují jednoznačné vstupy konceptuální reprezentace, jelikož podobný typ úlohy se v této doméně nevyskytuje. Přesto je jasné, že z hlediska (obecné matematické) teorie platí princip jednoznačného vstupu také v nerealizované variantě (konceptuální vstup jednoznačný a symbolický

nejednoznačný), jelikož mají obě primární reprezentativní vrstvy totožnou architekturu a také způsob propojení s multimodální vrstvou. Argumenty byly uvedeny výše. Daný způsob propojení navíc činí architekturu univerzálně použitelnou. Další výhodou je možnost rozšířit počet modalit, takže je možné společně reprezentovat vizuální, auditivní, motorickou a introceptivní informace ve společném multimodálním systému. Pokud zachováme princip privilegovanosti jednoznačného vstupu, bude zajištěna tvorba přesně identifikovatelných kategorií.

Můžeme tedy shrnout, že způsob primární reprezentace je odlišný od dosud realizovaných modelů. Harnad (1990), Regier (1996), Siskind (2001) aj. řeší problematiku symbolických vstupů jako apriorně dané, či jednoznačně vymezující kategorie a omezují již v počátku možnost systému fungovat podle principu privilegovanosti jednoznačného vstupu, jelikož jednoznačný vstup je fixně dán. Takový systém nedokáže řešit úlohy, ve kterém nelze pomocí symbolického vstupu určit jednoznačnou kategorii a bude tedy selhávat v obecných úlohách. Systém realizovaný v této práci je svou strukturou blízký Dorffnerovu řešení (1996), založeného na plně konekcionistickém modelu. Již v jeho práci se objevuje odlišné pojmání procesu ukotvení v oblasti symbolické roviny, která je zpracovávána na základě diskriminace. Podobně je tomu se způsobem propojení symbolické a konceptuální úrovně do jedné společné vrstvy. Jelikož svůj model realizoval ještě před teorií perceptuálního symbolu (Barsalou, 1999), není společná vrstva nazývána multimodální reprezentace.

Dorffner používá diskriminaci pro tvorbu auditivních (symbolických) kategorií, které mapuje do jednorozměrné sítě SOM, což je ojedinělý přístup k ukotvení. Ve shrnutí činnosti modelu však interpretuje model pouze v kontextu základního principu ukotvení a nepřichází s návrhem *privilegovanosti jednoznačného vstupu*, přestože by ho architektura měla být schopná. Důvod můžeme hledat v omezeném zadání úlohy, které podobný princip pro řešení úlohy nevyžaduje, popřípadě jej teoreticky nedopracoval. Přestože jeho řešení obsahuje schopnost vytvářet kategorie na základě jednoznačného vstupu, potenciál daného přístupu je prezentován až v této práci.

Z hlediska procesu kategorizace tedy můžeme shrnout, že klasické modely ukotvení (Harnad, 1990; Siskind, 2001; Regier, 1996) se jeví ve světle psychologických teorií jako nedostačující, jelikož pomocí nich nedokážeme řešit široké spektrum úloh, kterému je lidský kognitivní aparát během svého života vystaven, a které dokáže velmi dobře zvládnout. Proces prvotní diskriminace musí být proveden pro všechny úrovně znaku (symbolickou i konceptuální), abychom mohli pro jejich následnou

integraci využít vlastnosti, které jim umožní efektivně vytvářet společné propojení.

4.2 Identifikace

Jestliže jsme podrobněji analyzovali prvotní reprezentaci symbolů a konceptů, můžeme přikročit ke druhé části procesu ukotvení. Zaměříme pozornost na oblast identifikace symbolů a konceptů, která je pro realizovaný model výpočetně nejnáročnější. Pro dobrou orientaci bude interpretace probíhat v kontextu teorie znaku a také teorie ukotvení symbolů.

Primární vizuální reprezentace je možné nazírat jako interpretant (koncept či konceptuální úroveň), která vznikla na základě procesu diskriminace, tzn. první část procesu ukotvení. Primární auditivní reprezentace je totožná s representamen (symbol či symbolická úroveň) a je také vytvořena na základě procesů diskriminace. Připomeňme si v této chvíli Harnadovo (1990) dělení diskriminace a identifikace. Diskriminace je posouzení, zda jsou dva vstupy stejné nebo rozdílné, popřípadě jak jsou rozdílné. Jedná se o relativní posouzení tzn. hledání míry shody. Identifikace je přiřazení arbitrární odpovědi - "jména" - klasifikující vstupy, které následně tvoří kategorii podle shodného prvku. Identifikace je absolutní posouzení, které rozhodne, zda je vstup členem kategorie či nikoli (Harnad, 1990).

Pokud se na definici podíváme optikou realizovaného modelu, potřebují pojmy přesnější vymezení. Přestože je diskriminace posouzení míry shody, vytváří ohraničené kategorie již v primární auditivní (symbolické) vrstvě. I když se podle definice diskriminace jedná o relativní posouzení, dostáváme v tomto případě absolutní kategorie. Realizovaný model se začíná odlišovat od Harnadova vymezení. Rozpor můžeme hledat ve způsobu definice pojmů. Ve spojitosti s procesem identifikace je nutné zdůraznit, že pro její provedení je nutné použít dvou odlišných zdrojů (reprezentačních úrovní) referující k totožné entitě či procesu v externím prostředí, přičemž identifikace znamená jejich vzájemné namapování.

Rozdíl mezi diskriminací a identifikací je tedy hlavně ve způsobu aplikace. Diskriminace probíhá v rámci jednoho zdroje vstupu, který je kategorizován pouze v rámci těchto vstupů. Pro identifikaci jsou nutné dva vstupní zdroje, u kterých hledáme vzájemné propojení.

V realizovaném modelu tedy probíhá diskriminace samostatně pro vizuální a auditivní vstup a identifikace tvoří propojení mezi jejich primárními reprezentacemi. V Harnadově návrhu se zdůraznění tohoto faktu neobjevuje, přesto je z hlediska realizace klíčový. Pouhé dělení na relativní a absolutní posouzení neimplikuje počet reprezentačních vrstev.

Samotný proces identifikace probíhá v realizovaném modelu dle principu *privilegovanosti jednoznačného vstupu*. V oblasti reprezentace prostoru je vždy

jednoznačný auditivní symbolický vstup, přesto je pro obecnou interpretaci nutné připomenout variantu funkce systému, při které je jednoznačná vizuální kategorie a nejednoznačná auditivní (symbolická) úroveň, tzn.

multimodální reprezentace je strukturována na základě vizuální vstupu (konceptuální reprezentace).

Můžeme tedy říci, že proces identifikace nemusí probíhat pouze na základě jednoznačné symbolické informace, ale že je možné provádět identifikaci na základě konceptuální informace, přičemž rozhodující je *privilegovanost jednoznačného vstupu*. Oproti Harnadovu způsobu identifikace podle symbolického vstupu, je v naší teorii identifikace závislá na tom, z kterého vstupu přichází jednoznačné informace. Proces identifikace není fixně dán, ale je závislý na povaze vstupů z různých modalit. Obecně lze říci, že nezávisí na počtu modalit, jelikož na způsobu identifikace a struktura multimodální vrstvy se podílí vždy pouze jednoznačné vstupy.

Z hlediska způsobu zpracování informace je proces identifikace jádrem celého systému a nejobtížnější a výpočetně nejnáročnější úlohou. Jedná se o nalezení způsobu mapování mezi konceptem a symbolem. Jestliže je v Harnadově případě tento proces specifikován, jako jednoznačné označení kategorie, v realizovaném modelu přistupují k problematice na základě předpokladů zmíněných výše. Harnad provede symbolické přiřazení jako asociaci řetězce znaků s konceptem reprezentovaným neuronovou sítí (hybridní model). V našem případě vzniká propojení na základě učení a jeho výsledkem je multimodální vrstva. V ní se propojují informace ze primárních systémů a jsou na základě učení identifikovány jako jednoznačné kategorie. Rozdíl je v užití metody učení v celém systému. Jedná se o zdůraznění *učitelského přístupu*, který v našem případě přináší výhody v podobě principu *privilegovanosti jednoznačného vstupu*, který nelze v klasických modelech (Harnad, 1990; Regier, 1996) realizovat právě díky jejich designérskému přístupu a fixnosti a apriornosti reprezentace symbolického vstupu. Právě samoorganizující se sítě umožní systému univerzálně kategorizovat dle rozličných vstupů. Navíc lze systém rozšířit o další modalitu, v čemž spatřuji výhodu a přínos realizovaného modelu v oblasti ukotvení symbolů.

4.3 Kritika navrženého přístupu

Pro komplexní posouzení interpretuji realizovaný model také z hlediska možné kritiky daného přístupu. První námitka by mohla směřovat právě na způsob reprezentace symbolické roviny pomocí samoorganizující se sítě, která reprezentuje označení v úrovni slov.

V první fázi je možno napadnout reprezentaci celých slov pro jejich neschopnost zachytit zvukovou podobu na úrovni fonému. Odpovědí je fakt, že změna způsobu reprezentace z úrovně jednotlivých slov do úrovně

fonémů je pouze technická záležitost související s rozšířením architektury modelu a povahou vstupů. V budoucí verzi je s tímto způsobem reprezentace počítáno, jelikož může odhalit mnoho zajímavého o způsobu tvorby slov. Důvodem redukce na reprezentaci celých slov během současného řešení byla výpočetní náročnost modelu, založeného na reprezentaci fonémů, který je v oblasti ukotvení symbolů nadbytečný, jelikož nás v první fázi zajímají podrobnosti týkající se procesu diskriminace a identifikace.

Přechod na reprezentaci v úrovni fonémů by neměl být obtížný, jelikož se vyskytují práce, které se danou problematikou zabývají. Způsobem reprezentace symbolických označení na úrovni jednotlivých písmen či fonémů se zabývá ve svých pracích Farkaš (2003,2007) a dosahuje užitím principu samoorganizace výsledky, které lze užít a propojit s navrženým modelem.

Další námitkou týkající se reprezentace slov pomocí samoorganizujících se map je zbytečná reprezentační složitost a nadbytečnost auditivního (symbolického) vstupu. Důvodů se nabízí několik. Jestliže se jedná o arbitrární symbolický vstup, stačí jej reprezentovat pomocí sekvence symbolů, která jej bude jednoznačně identifikovat.

Použití samoorganizace může být považováno za neefektivní, jelikož nám nepřináší žádné další výhody pro následné zpracování. Taková námitka je oprávněná, jelikož pro úlohu reprezentace prostorových vztahů, ve které se vyskytují pouze jednoznačné symbolické vstupy, není potřeba jejich reprezentace pomocí tvorby 2D mřížky, organizující prostorová označení dle podobných vlastností a vytvářející „asociační mapu“. Model však slouží pro obecnější přístup k procesu ukotvení a proto vytvořen tak, aby byl schopen reprezentovat v multimodální vrstvě také nejednoznačné symbolické vstupy, tzn. byl schopen kategorizovat na základě konceptuálního vstupu. Pro lepší orientaci uvádím rozdíly mezi odlišnými typy vstupů.

Harnadovo základní řešení a od něj odvozené způsoby ukotvení vycházejí z předpokladu, že symbolické označení je vždy jednoznačné a dokáže přesně identifikovat koncept, tzn. přiřadit kategorii či prototypu jednoznačnou symbolickou nálepkou, se kterou je následně operováno v symbolickém systému.

Nejčastějšími vstupy takové formy ukotvení mohou být jednotlivé objekty externího prostředí, jejichž koncepty vytvořené na základě perceptuálních vstupů a za pomoci procesu diskriminace, jsou považovány za překrývající se prototypy (není jasná hranice prototypu kočky a rysa, popřípadě jiného zvířete). Naopak symboly jsou vždy jednoznačné označení (označení kočka,rys,pes apod.), které vždy odliší třídu či koncept (nejednoznačný).

Harnadův model je vhodný pro označení jednotlivých objektů, tříd objektů, popřípadě jednoznačných procesů.

Pokud bychom chtěli percentuálně vyjádřit zastoupení takových objektů a procesů během normálního způsobu interakce s okolím, těžko bychom našli v současných studiích odpovídající statistiku. Přesto můžeme říci, že se s danou formou tvorby znaků (jednoznačný symbol, nejednoznačný koncept) setkáváme ve většině případů. To nás ale neopravňuje říci, že je Harnadův model obecným řešením procesu ukotvení. Lidský kognitivní aparát je totiž schopen reprezentovat také objekty a situace, mimo rámec možností navržené Harnadem. V prvé řadě se jedná o polysémická slova. Jestliže je jedna sekvence symbolů identifikována s různými koncepty, nelze o takovém symbolickém označení říci, že je jednoznačné. Oproti předchozí situaci, kdy byla konceptuální vrstva nejednoznačná, nastává podobná situace v symbolické vrstvě. Rozdíl je v také v míře nejednoznačnosti. Jestliže u konceptuální úrovně se kategorie překrývají, u symbolické úrovně můžeme hovořit (v případě reprezentace pomocí fonetických vlastností) o identických kategoriích, jelikož polysémická slova (např. jazyk, klíč, dát apod.) nevykazují v symbolické úrovni žádné rozdíly. Taková vstupy nedokáže Harnadův model zpracovat díky své architektuře a hierarchii úrovní. Harnad postuluje symbolickou rovinu hierarchicky nejvýše, jelikož následně slouží pro operace v rámci formálního symbolického systému. Model realizovaný v této práci je zásadně odlišuje způsobem procesu ukotvení. Hierarchicky nejvýše stojí multimodální vrstva, která vytváří propojení mezi konceptuální a symbolickou úrovní. Právě ona dokáže absolutně odlišit jednotlivé znaky (tvořené propojením konceptu a symbolu). Při porovnání zmíněných přístupů docházíme k překvapivým závěrům. Pokud bychom interpretovali multimodální vrstvu z hlediska Harnadovy teorie, jednalo by se o proces identifikace, vedoucí k ukotvení symbolické vrstvy, se kterou je dále manipulováno. V našem modelu je multimodální vrstva procesem identifikace a zároveň vrchní vrstvou celého systému, která tvoří podklad pro následnou manipulaci (např. pomocí formálního symbolického systému). To znamená, že navržený model ukotvení obsahuje samostatnou reprezentační úroveň, která identifikuje symboly s koncepty. Identifikace je procesem a zároveň produktem této činnosti ve formě multimodální reprezentace. Je schopná vytvářet vždy jednoznačná označení, na rozdíl od symbolické úrovně, která selhává v případě polysémie. Pomocí přidání této třetí (multimodální vrstvy) vytváříme systém, který je univerzálně použitelný v procesu ukotvení. Změnám v nazírání reprezentace významu užitím navrženého modelu se budeme věnovat v samotném závěru.

Podívejme se zatím podrobněji na další aspekty polysémie. Kritika může namítat, že polysémická slova

tvoří pouze malé procento výrazu z celkového slovníku a že je možné zanedbat tyto výrazy. Tímto krokem však zavíráme cestu k napodobení lidského kognitivního aparátu, jelikož Harnadův systém není schopen podobné úlohy zvládnout. Na tento argument může kritika namítat, že je možné polysémická slova rozlišit na základě analýzy kontextu dané promluvy, ze které zjistíme, který koncept je potřeba symbolickému označení přiřadit. Takové argumentaci lze oponovat v několika úrovních. V elementární úrovni můžeme použít následující příklad. Jestliže jsou nám během fáze učení prezentována pouze jednoslovné symbolické označení objektu (např. jazyk), nelze kontextuální analýzu použít, jelikož symbolická (auditivní) informace obsahuje pouze tuto promluvu. Systém nedokáže provést jednoznačnou identifikaci konceptu se symbolem. Pokud použijeme architekturu realizovanou v této disertaci a modifikovanou pro oblast reprezentace objektů, lze pomocí konceptuální úrovně identifikovat koncept se symbolem na základě odlišnosti konceptuálních kategorií (např. při promluvě je mi vizuálně prezentován klíč od bytu nebo klíč na utahování), jejichž perceptuální vlastnosti a následně tvořené koncepty jsou natolik odlišné, že jsou klíčové pro proces identifikace v multimodální vrstvě. Identifikace závisí na jednoznačnějším, nikoli na apriorně privilegovaném vstupu a proto jsou oba vstupy v multimodální vrstvě zarepresentovány odlišně. Příklad, ve kterém nelze provést rozlišení na základě analýzy kontextu, je ale příliš elementární a můžeme se s ním setkat pouze během raného vývoje dítěte. Proto se pokusíme podrobněji analyzovat vliv kontextu, který nám umožní komplexnější pohled na fungování modelů ukotvení. Nejprve si načrtneme způsob zpracování (polysémických) slov v kontextu promluvy. Před jejím začátkem je potřeba specifikovat, jak definujeme kontext. Z našeho pohledu můžeme brát kontext jako vliv vzájemného působení časově blízce prezentovaných označení, na význam, který v promluvě mají. Laicky řečeno, společný výskyt slov tvoří kontext, pomocí kterého lze specifikovat význam jednotlivých označení. Jedná se o elementární formu kontextu, při které nebereme v potaz subjektivitu, intencionální postoje mluvčího a významy slov v úrovni obecnějších lingvistických kategorií (téma, žánr, styl). Jedná se nám o analýzu kontextu v úrovni jednoduchých promluv, pomocí které je možné upřesnit významy (polysémických) slov. Pokusme se nalézt rozdíly v analýze kontextu s použitím Harnadova a vlastního modelu. V případě prvního zmíněného máme reprezentovány polysémické slova pomocí symbolického označení, identifikující více rozdílných konceptů, které slouží k reprezentaci nesymbolických vlastností a odvozeně tedy významu. Jestliže potřebujeme určit, ke kterému konceptu

referujeme, můžeme analyzovat slova, související s daným symbolickým označením, které se vyskytují společně v promluvě. V symbolické rovině provedeme logickou analýzu, určující platnost výroků v dané promluvě, či soustavě promluv (výroků) a přiřadí jim pravdivostní hodnoty. V Harnadově modelu se nejedná o aplikaci symbolického systému dle pravidel formální sémantiky (pravdivost je dána na základě nerozpornosti a kompletnosti), ale o aplikaci sémantiky, která bere v potaz konceptuální vlastnosti, spojené s daným symbolem. Kontext nám slouží jako způsob testování pravdivosti pro různé koncepty, se kterými je polysémické slovo spojeno. V symbolickém systému dochází k ověřování výroků na základě odlišných konceptuálních vlastností, které nám odhalí, který koncept se pro dané symbolické označení hodí nejlépe. Pro tyto účely již nestačí užití pravdivostních hodnot pravda/nepřavda, ale na základě odlišných kombinací vlastností konceptů je potřeba vyjádřit míru konzistence výroku na zvolené škále (nejčastěji 0-1). Jedná se o posun směrem k fuzzy logice.

V případě že užijeme Harnadův model, můžeme se setkat v oblasti zpracování kontextu s obtížemi, které jsme analyzovali již dříve. Jestliže jsou symboly definovány jako fixní a apriorní označení, znamená analýza kontextu velmi výpočetně náročný proces, jelikož je potřeba analyzovat vliv kontextu na význam slov vždy pomocí „testování“, které určuje koncepty, jež jsou pro daná symbolická označení nositeli významu, pomocí míry pravděpodobnosti. To je dáno způsobem reprezentace symbolické roviny, a užitím pouze dvou reprezentačních úrovní, ze kterých je jedna brána jako fixní a absolutní z hlediska jednoznačnosti. V případě že bychom použili Harnadem navrženou symbolickou vrstvu jako vrchní úroveň, která slouží díky své jednoznačnosti j manipulaci s reprezentacemi, znamenalo by každé polysémické slovo pro systém extrémní výpočetní zátěž, jelikož by nebylo možné tyto výjimky systém naučit. Nedostatek vyplývá z fixní symbolické roviny a způsobu jejího propojení s konceptuální úrovní a je v rozporu s poznatky z oblasti psychologie, jelikož člověk dokáže zpracovávat odlišné významy slov velmi rychle a automaticky. Způsob zpracování se daleko více blíží přístupu, navrženému v realizovaném modelu. Hlavním přínosem je užití multimodální vrstvy, která se dokáže naučit odlišnosti slov na základě jak vizuálního, tak auditivního vstupu. Právě přidáním vrstvy, která slouží k jednoznačné kategorizaci je jádrem systému, schopného univerzálně reprezentovat okolní prostředí. Výhody systému se objevují i v případě potřeby rozlišit polysémické slovo na základě analýzy kontextu promluvy. Opět využijeme schopnosti systému vytvářet reprezentace pomocí metod učení, které nám umožní reprezentovat společný výskyt slov a následně určit význam jednotlivých slov v

promluvě. Oproti fixní symbolické rovině u Harnada můžeme pomoci primární auditivní (symbolické) reprezentace, zachytit nejen fonetické vlastnosti slov, ale také vytvářet asociační mapy na základě jejich vzájemného výskytu. Jedná se pouze o přidání dalších vlastností, které během učení reprezentujeme do primární samoorganizující se sítě. Stačí použít další dimenze, které slouží ke sdružování podobných slov dle blízkosti v promluvě. Tyto vlastnosti pak slouží v procesu identifikace s konceptuální rovinou, jelikož dokáží strukturovat tvorbu multimodální vrstvy a odlišit polysémická slova na základě kontextu. Rozdíl oproti Harnadovu modelu je ve způsobu činnosti modelu. Není třeba provádět pokaždé analýzu kontextu na základě testování vhodnosti možných konceptů, ke které slovo v promluvě referuje, ale je možné reprezentovat časté společné výskyty slov (kontext) v primární vizuální vrstvě pomocí metod učení, a následně použít tyto informace ke tvorbě jednoznačných kategorií v multimodální vrstvě, odlišující polysémická slova. Hlavním rozdílem je opět užití metod učení oproti fixním mechanismům.

4.4 Sémantika modelu ukotvení

Skrze ukázkou odlišnosti modelů ukotvení se dostáváme k nejobtížnější pasáži, týkající se samotné reprezentace významu. Abychom dokázali spolehlivě určit, co pojmem reprezentace významu znamená, bude třeba opět provést odlišení na základě identifikace rozdílů mezi Harnadovým přístupem a realizovaným řešením. Podívejme se opět nejdříve na Harnadovo řešení. V jeho pojetí dochází ke tvorbě Peircovského znaku, který je tvořen na základě konceptu a symbolu. Způsobem propojení je proces identifikace jednoho s druhým. V jeho pojetí tedy vzniká systém jakýchsi uspořádaných dvojic, přičemž jedna z nich zachycuje subsymbolické konstantní vlastnosti objektu v externím prostředí a druhá je tvořena arbitrárním shlukem symbolů, které jednoznačně určují daný objekt. Takto vybudovaný systém je následně připraven vykonávat základní myšlenkové operace na základě manipulace s ukotvenými reprezentacemi. Význam je zde reprezentován v konceptuální úrovni (interpretant) a zajišťuje korespondenci interního systému reprezentací s externím prostředím, se kterým systém interaguje. Pokud přistoupíme k následné tvorbě myšlenkových operací, nenacházíme u Harnadova přístupu (1990) příliš mnoho informací o funkci takového modelu. Ve svém hybridním systému navrhuje realizovat myšlenkové procesy pomocí pravidel inference, se kterými jsme se setkali v oblasti formálního symbolického systému. Základem jsou pravidla vyvozování, které na základě vstupních výroků vytvoří závěr. Harnad blíže neurčuje jakým způsobem proces probíhá. Ve svém návrhu

zdůrazňuje nutnost formálního symbolického systému, jako jediného způsobu manipulace s interními reprezentacemi.

„Sémantická interpretovatelnost, aby byla symbolická, musí být propojena s explicitní reprezentací, syntaktickou manipulovatelností a systematicitou. Žádné z těchto kritérií není arbitrární, a pokud je zlehčíme, ztratíme možnost pracovat s pojmy jako přírodní kategorie a také možnost operovat na základě formální teorie výpočtu, čímž se stává ze slova "symbolický" nevysvětlitelná metafora (která se liší interindividually). Proto budeme přemýšlet pouze o klasickém formálním systému (tak jak je definován), chceme-li ukotvovat symboly.“ (Harnad, 1990)

Tuto pasáž lze považovat za klíčovou. Hovoří o tom, že neexistuje alternativa k formálnímu systému, kterou bychom mohli použít jako podklad pro myšlenkové operace. Problematické je určit, jaký druh sémantiky měl Harnad na mysli v souvislosti se svým hybridním modelem. Jestliže se jedná o pravdivostní hodnoty, které podporují konzistentnost a úplnost systému výroků, nepotřebuje systém konceptuální úroveň, která zajišťuje korespondenci s externím světem. Takový systém se nijak neliší od čistě symbolických systémů. Harnadův model ovšem konceptuální úroveň obsahuje. Rozdílnost spočívá ve způsobu, jakým přiřadíme pravdivostní hodnotu výroků, na jejichž základě chceme provést inferenci. Můžeme použít konceptuální úroveň, tudíž nepotřebujeme systému interpretovat význam a pravdivost výroků externě. Systém vyvozuje na základě pravdivosti získané z perceptuálních dat a reprezentovaných konceptuálně. Použitím obdobného systému se ale dostáváme do obtíží. Jestliže má být systém konzistentní a úplný, je potřeba aby tuto vlastnost měly i koncepty, které přiřazují pravdivost výroků. Jelikož ale víme, že perceptuálně získané koncepty obsahují velké množství nepřesností, redukci a víceznačnosti, je tato konzistence nerealizovatelná. Přestože Harnad tvrdí, že symbolická rovina je jednoznačná a přesně identifikuje kategorie, je nutné si uvědomit, že význam je uložen v nejednoznačné konceptuální rovině. Symbolický systém pouze manipuluje s významy konceptuální roviny, ale není na ní nezávislý. Naopak se jedná o roviny, které jsou ve vzájemné korespondenci. Proto není možné nekonzistentní a nepřesnost konceptuální roviny pomocí symbolického systému redukovat. Jestliže jsou kategorie nejednoznačné, překrývající se a obsahují chybějící informace, nelze předpokládat, že v symbolickém systému lze tyto vlastnosti redukovat a dospět k logicky jednoznačnému závěru *pravda/nepravda*. Přestože Harnad tvrdí, že procesem identifikace získáme jednoznačné kategorie, zjistili jsme, že tato jednoznačnost neplatí pro všechny případy (polysémie, kontext). V

případě překrývajících se konceptuálních kategorií je možné, že systém identifikuje nejednoznačné vstupy s nesprávným symbolickým označením a nemůže být úplný a konzistentní. Máme zde tedy tři faktory (polysémie, kontext, nejednoznačnost konceptů), které způsobují, že použití klasické sémantiky pravdivostních hodnot *pravda/nepravda* je pro Harnadův model nerealizovatelné.

Alternativou je možnost užití *fuzzy logiky* v procesech vyvozování. V takových systémech je konzistentnost systému posunuta do roviny pravděpodobnosti, a výroky mohou nabývat pravdivosti na škále hodnot (nejčastěji 0 až 1) a systém je schopen vytvářet inference, pro které nebude problematická nekonzistentnost konceptuální úrovně. Logická konsistence je zde nahrazena pravděpodobností. Díky fixnímu způsobu reprezentace symbolické roviny, si však podobný systém nedokáže poradit s polysémií či vlivy kontextu.

Shrněme tedy, že v Harnadově přístupu je význam reprezentován v konceptuální vrstvě, která slouží jako podklad pro následnou pravdivostní hodnotu symbolických označení. Pomocí pravidel vyplývání jsou vytvářeny myšlenkové operace a jejich závěry. Dosud jsme se však nezmínili (ani Harnad tuto informaci neuvádí), jakým způsobem je reprezentován význam tohoto produktu myšlenkových operací. Abychom zajistili zpětnou korespondenci s konceptuální úrovní, je nutné závěr inference do ní opět namapovat. V případě, že odpovídající koncept neexistuje, je nutné jej vytvořit na základě konceptů stávajících. V případě že koncept existuje, je nutné závěr v symbolické rovině porovnat s vlastnostmi v rovině konceptuální. Tím ale narážíme na problém zpětného převodu symbolické roviny do konceptuální. Jestliže je výrok pravdivý na základě pravidel vyplývání, jaké to má implikace pro konceptuální úroveň?

Přestože celá disertační práce byla věnována procesu propojení konceptuální úrovně se symbolickou, neobjevily se zde žádné informace o navazující část, tzn. zajištění způsobu převodu logických operací do úrovně konceptů. Jelikož se jedná o oblast, která není přesněji popsána ani v úrovni teorie, byly by další vývody v úrovni spekulace. Přesto této oblasti přikládám prioritní význam a budu se jí věnovat v souvislosti s rozvojem modelu reprezentace prostorových vztahů.

Vraťme se nyní k původnímu tématu této kapitoly, kterým je porovnání Harnadova a vlastního modelu z hlediska reprezentace významu. První zmíněny jsme již popsali, takže můžeme přikročit k definici rozdílů. I v případě vlastního modelu je naplněna podstata Peircovy teorie znaku, jelikož systém obsahuje symbolickou a konceptuální úroveň, která je založena na informaci z externího prostředí. Kromě zmíněných dvou však systém obsahuje také multimodální vrstvu, která

nekoresponduje s klasickou teorií znaku. Multimodální vrstva slouží k propojení konceptuální a symbolické úrovně. Takové propojení není u Peirce přesněji specifikováno, ani neobsahuje žádný název. U Harnada můžeme hledat podobnost multimodální vrstvy s procesem identifikace. Narážíme tedy na rozdíl, jelikož v realizovaném modelu se nejedná o pouhý proces, ale o další reprezentační vrstvu, která tento proces provádí a samostatně reprezentuje. Jestliže Harnad vidí identifikaci jako mechanickou proceduru, jejíž výsledky jsou uloženy ve způsobu propojení a vzájemného mapování symbolů a konceptů, v našem případě jsme vytvořili novou reprezentační úroveň, která ukládá nejen způsob propojení konceptů a symbolů, ale také vytváří jednoznačně identifikovatelné kategorie, které jsou nositeli společných vlastností. Jestliže u Harnada jsou jednoznačné kategorie pouze v symbolické rovině, v realizovaném modelu jsou jednoznačné kategorie v rovině multimodální. Proces identifikace se stává samostatnou částí systému, se schopností reprezentovat a jednoznačně (absolutně) kategorizovat. Opodstatněnost nového prvku v modelu ukotvení byla prokázána na základě předchozí argumentace. Harnadův model není schopen vytvářet jednoznačné kategorie v případě vlivu polysémie či kontextu, jelikož obsahuje fixní symbolickou rovinu, která způsobí neefektivní způsob analýzy kontextu. Užití třetí reprezentační vrstvy se tedy nabízí jako logický krok směrem k univerzálnímu reprezentačnímu systému. Proces identifikace je prováděn na základě učení, což zajišťuje jeho schopnost přizpůsobit se vstupním informacím. Jelikož je identifikace realizována pomocí samostatné vrstvy, umožní nám reprezentovat kategorie na základě *privilegovanosti jednoznačného vstupu*, čímž eliminuje vliv polysémie. V případě rozšíření vlastností primární symbolické vrstvy a způsobu učení v multimodální vrstvě, můžeme zajistit reprezentaci kontextuálních informací v multimodální vrstvě, což sníží výpočetní náročnost spojenou s nutností testování možných variant, jako v případě Harnadova modelu. Další výhodou je možnost napojení ostatních modalit na tuto vrstvu, přičemž identifikace bude vždy probíhat na základě jednoznačného vstupu, což zajistí tvorbu jednoznačných kategorií.

Podobně jako u Harnadova modelu nás zajímá způsob jakým můžeme s takto vybudovaným systémem provádět myšlenkové operace. Opět budeme hovořit pouze o elementární formě, která je spojena s logickými principy vyplývání na základě výroků. Již v začátku můžeme odmítnout přístup, založený na formální sémantice a také přístup založený na konzistenci a úplnosti výroků, které získaly pravdivost na základě konceptuálních úrovně. Jedná se o přístupy, používající pravdivost v úrovni *pravda/nepravda*, která je ve svém základu postavena na

jednoznačnosti výroků do systému vstupujících. Jak bylo zmíněno výše, je význam v systému reprezentován pomocí konceptuální roviny, která obsahuje velkou míru nejasnosti a nekonzistence, takže podobný typ posuzování pravdivosti výroků není možný.

V této chvíli je nutné zdůraznit, že právě způsob reprezentace významu pomocí konceptuální úrovně neumožňuje systému naplnit Harnadův předpoklad, tzn. fungování dle pravidel formálního systému a jeho sémantiky. Přestože jsou jevy v externím prostředí zcela deterministické na úrovni elementárních vztahů (jednoznačné fungování fyzikálních zákonů, pokud nebereme v potaz názory kvantové fyziky), není schopen perceptuální aparát podobný lidskému zachytit a reprezentovat tyto fakta a jevy jednoznačným a konzistentním způsobem. Pokud chceme s obdobnými reprezentacemi následně pracovat v rovině abstraktních a jednoznačných kategorií, projeví se nekonzistence konceptuálního pokladu (reprezentujícího význam) v rozporných výsledcích inferencí v úrovni symbolické. Je proto nutné zdůraznit, že nejenom formální sémantika není vhodným nástrojem pro reprezentaci významu v oblasti umělé inteligence, ale také sémantika postavená na konceptuálních základech, která operuje s pravdivostí v úrovni *pravda/nepravda* je příliš hrubým nástrojem reprezentace významu.

Tento fakt zní poněkud neintuitivně v kontextu systému, který je v disertaci vytvořen. Jestliže multimodální vrstva vytváří vždy jednoznačné kategorie, na základě *privilegovanosti jednoznačného vstupu*, jedná se o totožnou reprezentační vrstvu, jakou je symbolická v případě Harnadova modelu, s tím rozdílem že je odolná proti vlivům polysémie.

Přesto se můžou během procesu akvizice symbolických a konceptuálních reprezentací objevit šumy a nekonzistence, promítající se do reprezentací v multimodální vrstvě, a v rámci následných operací na základě inference bude docházet k rozporným výsledkům. Takové chyby nemůžeme na základě architektury, kterou používáme pro reprezentaci významu eliminovat a proto je potřeba hledat alternativní způsob práce v úrovni inferencí. Jak již bylo zmíněno výše, alternativu může představovat fuzzy logika, u které je konsistence vyjádřena mírou pravděpodobnosti. Takový systém je daleko bližší lidskému způsobu operací s reprezentacemi a nabízí možnost eliminovat nekonzistentnost konceptuální úrovně. Nalezení vhodného způsobu implementace tohoto logického aparátu do stávajícího systému je dalším krokem jeho vývoje, který potvrdí správnost směřování a univerzálnost navrhovaného řešení.

5 Poděkování

Chtěl bych poděkovat doc. Igoru Farkašovi z Katedry aplikované informatiky a matematiky Komenského univerzity v Bratislavě za pomoc při realizaci modelu a RNDr. Martinu Takáčovi za inspirativní připomínky.

4 Literatura

- [1] L. W. Barsalou: Perceptual symbol systems. *Behavioral and Brain Sciences* 22 (1999) 577-609
- [2] L. W. Barsalou, A. K. Kyle, C. D. Wilson: Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences*, 7(2) (2003) 84-91.
- [3] A. Cangelosi, A. Greco and S. Harnad: From Robotic Toil to Symbolic Theft: Grounding Transfer from Entry-Level to Higher-Level Categories. *Connection Science* 2 (2000) 143-162
- [4] K. Coventry: *Saying, Seeing and Acting: The Psychological Semantics of Spatial Prepositions (Essays in Cognitive Psychology)* (2004) Psychology Press.
- [5] G. Dorffner, M. Hentze, G. Thurner: A Connectionist Model of Categorization and Grounded Word Learning, In: *Proceedings of the Groningen Assembly on Language Acquisition: (Koster C., Wijnen F. eds.)*, (GALA '95)
- [6] A.R. Damasio: Time-locked multiregional retroactivation: A systems level proposal for the neural substrates of recall and recognition. *Cognition* 33: (1989) 25-62.
- [7] I. Farkas, M. Crocker: Systematicity in sentence processing with a recursive self-organizing neural network. *Proceedings of the 15th European Symposium on Artificial Neural Networks (ESANN'07)*, Bruges, Belgium, (Ed.) M. Verleysen, pp. 49-54, 2007.
- [8] I. Farkas: Lexical acquisition and developing semantic map. *Neural Network World*, 13(3), 235-245, 2003.
- [9] J. A. Fodor: *The Language of Thought*. New York: Crowell. (1975)
- [10] J. Gumperz: *Rethinking Linguistic Relativity*. Edited by John Gumperz & Stephen Levinson. Cambridge University Press. 1996
- [11] S. Harnad: The Symbol Grounding Problem. *Physica D* (1990) 335-346.

- [12] W. G. Hayward, M. J. Tarr: Spatial language and spatial representation. *Cognition*, 55(1) (1995) 39-84.
- [13] J. Huttenlocher, L. Hedges, S. Duncan: Categories and particulars: Prototype effects in estimating spatial location. *Psychological Review*, 98, (1991) 352-376.
- [14] T. Kohonen: *Self-organization and associative memory*, Springer, New York 1989
- [15] B. Landau, J.E. Hoffman: Parallels between spatial cognition and spatial language: Evidence from Williams syndrome. *Journal of Memory and Language*. 53 (2) (2005) 163-185.
- [16] R. W. Langacker: *Foundations of cognitive grammar: Theoretical Prerequisites*. Stanford, CA: Stanford University Press.(1987)
- [17] A. Newell, H. Simon: *Human Problem Solving*. Prentice Hall. 1972.
- [18] A. Paivio: *Mental representation: A dual coding approach*. New York: Oxford, 1986
- [19] C.S. Peirce: *Collected Papers of Charles Sanders Peirce*, vols. 1-6, Charles Hartshorne and Paul Weiss (eds.), vols. 7-8, Arthur W. Burks (ed.), Harvard University Press, Cambridge, MA, 1931-1935, 1958.
- [20] Z.W. Pylyshyn: *Computation and Cognition: Towards a Foundation for Cognitive Science*. MIT Press, (1984).
- [21] T. Regier: *The Human Semantic Potential: Spatial Language and Constrained Connectionism* Cambridge, MA: MIT Press, 1996
- [22] F. de Saussure: *Course In General Linguistics*. McGraw-Hill Humanities/Social Sciences/Languages. 1965
- [23] J. M. Siskind: Grounding the Lexical Semantics of Verbs in Visual Perception using Force Dynamics and Event Logic. *J. Artif. Intell. Res. (JAIR)* 15, (2001), 31-90
- [24] L. Talmy: *Toward a Cognitive Semantics, Volume 1: Concept Structuring Systems; Volume 2: Typology and Process in Concept Structuring*, MIT Press, Language, Speech, and Communication series. (2003)
- [25] A. Tarski: The Semantic Conception of Truth: and the Foundations of Semantics'. *Philosophy and Phenomenological Research* 4(3): (1944) 341-376.
- [26] F. Varela, E. Thompson, E. Rosch: *The Embodied Mind - Cognitive Science and Human Experience*. Cambridge, MA: MIT Press / Bradford Books, 1991
- [27] M. Vavrečka: Ukotvení prostorových vztahů, In J. Kelemen, V.Kvasnička, P.Trebatický: *Kognicia a umelý život VII*. Bratislava, (2007)
- [28] T. Voegtlin. "Recursive self-organizing maps". *Neur.Netw.*, 15(8-9):979-991, 2002.
- [29] B. Whorf: *Language, Thought, and Reality: Selected Writings of Benjamin Lee Whorf*. MIT Press (1956).
- [30] T. Ziemke: Rethinking Grounding. In: *Understanding Representation in the Cognitive Sciences*, (A. Riegler, M. Peschl and A. von Stein, Eds.), New York: Plenum Press, 1999: 177-190.