

Konekcionistické modelovanie jazyka *

Igor Farkaš
Katedra aplikovanej informatiky
FMFI UK v Bratislave

1. Úvod

Prirodzený jazyk je jednou z najfascinujúcejších a najkomplexnejších kognitívnych schopností človeka. Pri bežnej komunikácii, a vzhľadom na ľahkosť, s akou si dieťa jazyk osvojuje, považujeme jazyk a reč za niečo samozrejmé. Avšak keď začneme pátrať po procesoch, ktoré sa za touto schopnosťou skrývajú, ich zložitosť sa nám zdá ohromujúca.

Úplná teória ľudského jazyka musí zahŕňať znalosti o neurálnych mechanizmoch, pomocou ktorých ľudská myseľ spracováva jazyk, vrátane povahy neurálnych reprezentácií, vrátane toho, ako sa učíme a aké potenciálne vrodené znalosti alebo štruktúry sú na to potrebné atď. Tým, že jazyk úzko súvisí s kogníciou, môžeme profitovať aj zo štúdií iných kognitívnych fenoménov, vrátane pamäti alebo videnia, ktoré zdieľame aj s nižšími živočíchmi. Pred desiatkami rokov boli hlavným zdrojom poznávania kognitívnych procesov rôzne invazívne metódy. Avšak pri poznávaní jazykovo-špecifických procesov vlastných človeku narážame na obmedzenia a z etických dôvodov musíme hľadať prevažne neinvasívne metódy skúmania. Rôzne poruchy jazyka u pacientov sú síce tiež zdrojom informácie, ale ťažiskom empirického výskumu ostávajú hlavne psycholingvistické experimenty, ktoré umožňujú odhaľovať mentálne procesy prebiehajúce pri rôznych jazykových úlohách. Od 90. rokov minulého storočia sa k nim pridali rôzne zobrazovacie metódy (napr. na báze pozitronovej emisnej tomografie alebo funkčnej magnetickej rezonancie), ktoré do istej miery umožňujú nachádzať neurálne koreláty týchto procesov (pozri kapitolu *Nervový systém a jazyk*).

Avšak aby boli empirické dáta užitočné, je nutné overiť, či vyhovujú nejakej teórii jazyka, nejakým procesom alebo funkciám. V porovnaní s teoretickými modelmi opísanými slovne sú implementované počítačové modely oveľa cennejšie, pretože sú explicitné a exaktné, a tým aj jednoznačne interpretovateľné. Tým, že sa pri modelovaní pozeráme na jazyk ako na výpočtový proces, môžeme zisťovať, ako jazyk funguje. Môžeme generovať a testovať hypotézy a porovnávaním predikovaných výsledkov s experimentálnymi pozorovaniami z psycholingvistických štúdií môžeme overovať, či je náš model psychologicky alebo kognitívne prijateľný. Motivácia však platí aj naopak: predikcie modelu môžu byť inšpiráciou pre ďalšie experimenty, ktoré by mohli predikcie potvrdiť alebo vyvrátiť.

Spomedzi počítačových modelov jazyka boli donedávna dominantnými symbolové modely, založené na predstave mysle ako stroja spracúvajúceho symboly. Ako sa však ukázalo, symbolové modely nedokážu vysvetliť niektoré kognitívne javy vznikajúce pri používaní jazyka. Preto – v súvislosti s požadovanou kognitívnou prijateľnosťou – sa popri klasických, symbolových modeloch začali v 80. rokoch minulého storočia používať konekcionistické modely jazyka, ale aj kognície všeobecne (McCLELLAND, RUMELHART, PDP RESEARCH GROUP 1986; QUINLAN 1991; McLEOD, PLUNKETT, ROLLS 1998). Prirodzený jazyk pritom vzbudil najviac záujmu, ale aj

* Táto práca bola publikovaná ako kapitola v knihe Rybár, Kvasnička, Farkaš (zost.): *Jazyk a kognícia*. Kalligram Bratislava 2005, str. 262-305. Práca bola podporená grantmi VEGA 1/9046/02 a 2/4026/04. Autor ďakuje Prof. Vladimírovi Kvasničkovi za cenné pripomienky k tejto kapitole.

kontroverzie (PINKER, MEHLER 1988). Podľa niektorých odborníkov sa totižto (minimálne) niektoré aspekty jazyka vôbec nedajú vysvetliť konekcionisticky, ale len pomocou symbolových modelov. Jadro kontroverzie súvisí s otázkou, aká je povaha mentálnych procesov prebiehajúcich v mozgu pri vyšších kognitívnych procesoch (vrátane jazyka). Ak sú tieto mentálne procesy naozaj symbolové, konekcionistické modely potom môžu byť len ich implementáciou na subsymbolovej úrovni (**implementačný konekcionizmus**; PINKER, PRINCE 1988). Avšak ak tieto mentálne procesy možno zredukovať na subsymbolovú úroveň, potom konekcionizmus bude znamenať alternatívnu paradigmu k vysvetleniu povahy jazyka (**elimináčny konekcionizmus**). Odpoveď na túto otázku je pri súčasnej úrovni poznania zložitá a zrejme bude závisieť aj od toho, ktorý aspekt jazyka máme na mysli. Jazyk ako komplexný fenomén zahŕňa celú škálu procesov, od fonológie až po sémantiku viet a pragmatiku, a práve tie „nižšie“ procesy sa zdajú ľahšie vysvetliteľné subsymbolovo. V ďalšom texte budeme orientovať pozornosť na konekcionistický prístup a pokúsime sa prezentovať úspešné modely zamerané nielen na „nižšie“ jazykové procesy prebiehajúce na lexikálnej úrovni, ale aj na tie „vyššie“, týkajúce sa spracovania viet. Najprv si však bližšie opíšme, čím sa symbolové a konekcionistické modely navzájom líšia.

1.1. Vlastnosti symbolových systémov

Symbolový pohľad na jazyk je veľmi podobný tomu, ako pracujú počítače, ktoré sa objavili zhruba v polovici minulého storočia. Podobnosť sa týka hlavne používania presne definovaných algoritmov a pravidiel, ktoré sú uložené v pamäti a vykonávané procesorom. Symbolový pohľad na jazyk vychádza z hypotézy, že znalosť jazyka zodpovedá znalosti gramatických pravidiel (CHOMSKY 1956).

Hlavnou črtou symbolových systémov je teda to, že pracujú s abstraktnými premennými – **symbolmi**, ktoré sa transformujú pomocou **explicitných pravidiel**. Takýto prístup je veľmi *transparentný* a veľmi vhodný na opísanie gramatík jazyka, čo je vlastne systém prepisovacích pravidiel. Napríklad $V \rightarrow (NF) VF$ pre slovenčinu znamená, že vetu možno vytvoriť zretazením (vynechateľnej¹) nominálnej frázy (NF) a slovesnej frázy (SF). Bežná NF sa skladá zo substantíva rozvitého nejakým adjektívom, čiže $NF \rightarrow (Adj) Sub$, a SF obsahuje sloveso a potenciálne nejakú NF a/alebo predložkovú frázu: $SF \rightarrow S1 (NF) (PF)$. Ako vidíme, NF sa môže vyskytnúť opakovane, z čoho vidieť, že takýto formalizmus veľmi dobre umožňuje opísať ďalšie podstatné charakteristiky jazyka – **rekurziu** a s ňou súvisiacu **generatívnosť** (generovanie teoreticky nespočetného množstva viet podľa danej gramatiky).

Na druhej strane **sémantika** je oblasť, kde sa symbolové modely dostávajú do ťažkostí. Na rozdiel od syntaktických kategórií sa sémantické koncepty nedajú tak ľahko klasifikovať a manipulovať pomocou abstraktných premenných. Je napríklad ťažké presne opísať význam alebo charakteristiky podobných konceptov v rámci nejakej kategórie (napr. ako sa líšia významy slov *krásny*, *nádherný*, *malebný*, *očarujúci*, *roztomilý*, *okúzľujúci* či *skvostný*). Podobné problémy vznikajú v sémantike na úrovni viet.

Ďalšou ťažkosťou symbolových modelov je súčasné **zohľadnenie gradovaných ohraničení** (graded constraints satisfaction), ktoré človek berie do úvahy pri analýze vety, a ktoré mu umožňujú porozumieť aj nejednoznačným alebo negramatickým vetám (negramatickým v klasickom, dichotomickom zmysle).² Problém spočíva v tom, že pravidlá sú rigidné a je ťažké zahrnúť do nich čiastočné a nediskrétné (nebinárne)

¹ Slovenská gramatika (na rozdiel od anglickej) umožňuje zmlčanie podmetu (napr. *Prší.*).

² Gramatickosť ako gradovaný fenomén rozoberáme i neskôr v časti 3.2. o indukcii gramatík. Aj Chomsky (1975) uznáva, že „adekvátna teória jazyka musí zahŕňať stupne gramatickosti“ (str. 131). Jej implementácia v symbolových systémoch je však dosť náročná.

informácie, ktoré nemajú rigidný syntaktický charakter (lexikálnu sémantiku, frekvencie výskytu slov, kontext atď.).

Vzhľadom na rigidnosť pravidiel symbolovým modelom **chýba robustnosť**. Minimálna porucha alebo nepresná či neúplná informácia môže mať za následok krach celého systému. Človek však porozumie vetám, aj keď sú negramatické alebo obsahujú skomolené slová. Človek takisto môže robiť chyby pri bežnom prejave.

Ďalším problémom symbolových systémov je **učenie**. Znalosť sa do systému obyčajne vkladá explicitne programátorom. Ak je k dispozícii dobrý externý učiteľ, symbolové systémy sú schopné naučiť sa nové pravidlá. Avšak dieťa pri akvizícii jazyka takéhoto dobrého explicitného korektora nemá k dispozícii. Dieťaťu sa obyčajne nehovorí po každej vete, akú urobilo chybu, a ak aj áno, dieťa si to nie vždy uvedomí (PINKER 1984; MORGAN, BONAMO, TRAVIS 1995). Napriek tomu sa jazyk naučí. Rôzne vysvetlenia tohto fenoménu sú príčinou pretrvávajúcej debaty o povahe vrodenej dispozície a o povahe akvizície jazyka, ktoré stručne spomenieme v časti 1.3.

Záverom tejto časti budeme konštatovať, že schopnosti človeka a správanie symbolových systémov sa javia z pohľadu spomínaných aspektov jazyka do značnej miery **komplementárne**. S tým, v čom sú symbolové systémy dobré, má človek problémy (napr. porozumenie viet so zložitejšími syntaktickými štruktúrami alebo viacnásobnými vnorenými rekurziami), a naopak, čo človek zvláda poľahky (napr. vnímanie jemných rozdielov medzi sémantickými konceptmi alebo využívanie informácií pravdepodobnostného charakteru), s tým má symbolový systém ťažkosti.

Ku komplementárnosti dospejeme i pri porovnávaní dvoch formálnych systémov na automatické spracovanie dát – symbolového a subsymbolového (HAMMER 2003). Symbolové systémy pracujú s dátami na vyššej úrovni (napr. vyjadrenými pomocou logických formúl či hierarchických stromov) a ich správanie je preto dosť ľahko pochopiteľné. Učenie symbolových systémov je však dosť ťažké, neefektívne a veľmi citlivé na šum. Subsymbologové systémy pracujú s dátami na nižšej úrovni (vyjadrenými pomocou vektorov čísel), čím sa ich správanie stáva človeku menej zrozumiteľné. Ich učenie je však dosť efektívne a odolné voči šumu. V tejto kapitole sa budeme snažiť v doméne jazyka ukázať, že konekcionistické systémy sú do značnej miery schopné spracovávať aj symbolovo reprezentované informácie, čím budeme v podstate argumentovať v prospech eliminačného konekcionizmu. Teraz si bližšie opíšme konekcionistické systémy a ich vlastnosti.

1.2. Vlastnosti konekcionistických systémov

Vďaka svojim kvalitatívnym odlišnostiam sú konekcionistické modely principiálnou alternatívou symbolových modelov kognície a jazyka. A to alternatívou kognitívne i biologicky prijateľnejšou, pretože nielen že sa viac podobajú architektúre mozgu, ale v mnohých prípadoch dokážu aj lepšie vysvetliť empirické dáta.

Konekcionistické modely (nazývané aj umelé neurónové siete) reprezentujú úsilie o výpočtový **prístup ku kognícii zdola** (KVASNIČKA A KOL. 1997; HAYKIN 1994). Konekcionistický model predstavuje zjednodušený model mozgu, ktorý pozostáva z množiny jednoduchých prvkov – uzlov (analógia k biologickým neurónom), navzájom komunikujúcich cez váhované spojenia, ktoré kvantifikujú silu prepojení (analógia k synaptickým spojeniam). Sila neurónovej siete nie je v samotných neurónoch, ale v ich vzájomnom prepojení a interakcii. Spojenia môžu byť kladné alebo záporné. V prvom prípade vstupný neurón excituje neurón, na ktorý sa pripája, v druhom prípade ho inhibuje. Každý neurón má svoj vnútorný potenciál, ktorý je súčtom všetkých pozitívnych a negatívnych príspevkov a na základe hodnoty potenciálu potom generuje svoju výstupnú aktivitu. Považa komunikácie medzi neurónmi má však skôr **numerický** než symbolový charakter. Biologické neuróny vysielajú impulzy s rôznou frekvenciou, v umelých

neurónových sieťach je táto frekvencia zvyčajne aproximovaná reálnym číslom na výstupe neurónu. Z matematického pohľadu sa na umelý neurón možno pozerat' ako na výpočtový element, ktorý transformuje vektor čísel (jeho rozmer je daný počtom vstupov neurónu) na jeden numerický výstup. Tento opis je síce tiež istou abstrakciou od dejov v biologických neurónoch, ale je veľmi vhodným nástrojom na modelovanie kognitívnych procesov.

Z pohľadu symbolovo-subsymbolovej dichotómie môžeme rozlíšiť dva typy konekcionistických modelov. *Lokalistické*³ konekcionistické modely sú podobné symbolovým modelom v tom zmysle, že každý neurón reprezentuje nejakú entitu (napr. tvár mojej starej mamy). Avšak v *distribovaných* konekcionistických modeloch aktivácia konkrétneho neurónu nemusí mať jednoznačnú interpretáciu, ale entity sú reprezentované množinou neurónov (population coding). Takéto distribuované reprezentácie, na ktoré sa možno z matematického hľadiska pozerat' ako na vysokorozmerné vektory čísel, majú niekoľko výhod:⁴ umožňujú **dobre reprezentovať sémantické koncepty** a podobnosti medzi nimi a vnášajú do systému **robustnosť**. Vďaka robustnosti má systém vlastnosť *postupnej degradácie* (graceful degradation), pretože i v prípade šumu či lokálnych porúch si zachová istú funkčnosť, ktorá bude závisieť od miery poškodenia.

Potenciál neurónovej siete závisí predovšetkým od jej architektúry. Neuróny sú obyčajne usporiadané vo vrstvách, medzi ktorými sa šíri informácia. Dopredné neurónové siete umožňujú tok informácie len jedným smerom, od vstupu na výstup, čo je postačujúce pre simuláciu problémov, ktoré možno formálne opísať nejakou jedno-jednoznačnou vstupno-výstupnou transformáciou. Výpočtovo silnejšie rekurentné neurónové siete, ktoré obsahujú **spätne väzby**, už umožňujú zahrnúť časovú dimenziu do transformácie, čo je nutné napríklad na generovanie viet jazyka. Viacvrstvové neurónové siete (t. j. s minimálne jednou skrytou vrstvou⁵) nie sú len jednoduchým asociátorom (čo je častý argument symbolikov), pretože na ich skrytých neurónoch môžu emergovať *interné reprezentácie*, ktoré charakterizujú abstraktné vlastnosti vyextrahované zo štruktúry vstupných dát (prostredia).

Neurónová sieť má jedinečnú **schopnosť učiť sa**, ktorá spočíva v zmene jej parametrov – váh spojení. To znamená, že parametre siete nemusíme explicitne nastaviť, len jej poskytnúť učiaci algoritmus. Podobne ako človek, sieť sa učí na základe tréningových príkladov. Učiaci algoritmus spôsobuje zmenu váh tak, aby sa po naučení dosiahlo požadované správanie siete, čo formálne zodpovedá minimalizácii nejakej chybovej funkcie. V závislosti od toho, s akou mierou explicitnosti sa sieť podáva informácia o tom, akú robí chybu, rozlišujeme tri skupiny učiacich algoritmov. Pri *učení s učiteľom* (supervised learning) každý neurón dostáva takúto informáciu, a ten mení potom svoje vstupné váhy tak, aby sa jeho chyba znižovala (najčastejšie používanou je metóda spätného šírenia chýb, back-propagation). Pri *učení odmenou a trestom* (reinforcement learning) len sieť ako celok dostáva informáciu o aktuálnej chybe v podobe skalárnej veličiny. Pri *učení bez učiteľa* (unsupervised learning) sieť nemá k dispozícii nijakú explicitnú informáciu o chybe, preto sa v tejto súvislosti hovorí o samoorganizácii siete.⁶ Existujú teórie, že všetky tri skupiny učiacich algoritmov by mohli byť relevantné pre rôzne

³ Page (2000) poukazuje na nepresnosti, pokiaľ ide o „nálepkovanie“ modelov. Argumentuje, že niektoré konekcionistické modely, označované za lokalistické, používajú aj distribuované reprezentácie.

⁴ Pre objektivnosť spomeňme, že distribuovaná reprezentácia má aj nevýhody, napr. ako súčasne reprezentovať viacero konceptov tými istými neurónmi (kvôli prekryvu). Tieto nevýhody sa však dajú značne zmierniť priblížením k lokalistickým reprezentáciám (sparse coding) alebo úplne odstrániť pomocou časovej synchronizácie v biologicky realističnejších modeloch s impulznými neurónmi.

⁵ Tieto neuróny sa nazývajú skryté, pretože z pohľadu vstup - výstup sú neviditeľné (neurónová sieť ako čierna skrinka).

⁶ Napriek neexistencii explicitnej chybovej informácie aj samoorganizujúca sa neurónová sieť realizuje nejakú optimalizáciu. Spravidla ide o nejakú extrakciu štruktúrálnych vzťahov zo vstupných dát (podľa učiaceho algoritmu).

časti mozgu, ale pre mozgovú kôru ako sídlo vyšších kognitívnych funkcií je najpriateľnejšou hypotézou samoorganizácia (DOYA 1999).

Učenie siete priamo súvisí s jej schopnosťou **zovšeobecňovať** (generalizovať) svoje znalosti a správne reagovať aj v nových situáciách. Neurónová sieť hľadá štrukturálne vzťahy počas učenia na príkladoch (trénovacia množina) a túto znalosť potom uplatňuje pri generalizácii (testovacie dáta). Ani dieťa nepotrebuje počuť všetky pádové koncovky všetkých podstatných mien, aby ich vedelo správne používať. Podobne ako dieťa i sieť robí spočiatku chyby pri generalizácii, to je však vo veľkom súlade s psycholingvistickou teóriou akvizície jazyka.

Aby sme to teda zhrnuli, konekcionistické modely (hlavne tie s distribuovanými reprezentáciami) poskytujú veľmi zaujímavú paradigmu na vysvetlenie jazyka, a to hlavne pre ich väčšiu biologickú i psychologickú prijateľnosť. Jazykový prejav človeka sa ukazuje vo väčšom súlade s prednosťami a ťažkosťami konekcionistických modelov než so symbolovými systémami.

1.3. Vrodené dispozície a akvizícia jazyka

Otázka povahy vrodenej dispozície jazyka (nature-nurture debate) je v súčasnosti veľmi horlivo diskutovanou témou aj vďaka nástupu konekcionistických modelov. Konekcionistický pohľad možno dobre konfrontovať so stále veľmi vplyvnou teóriou niektorých lingvistov a psycholingvistov (na čele s Chomským), ktorí tvrdia, že pre nedostatočnosť jazykových vstupov (poverty of stimulus) musí byť značná časť jazykových schopností vrodenej (MORGAN, TRAVIS 1989; MARCUS 1993; pozri aj kapitolu *Akvizícia jazyka u dieťaťa*). Konkrétne Chomsky postuloval existenciu univerzálnej gramatiky, ktorá vymedzuje priestor naučiteľných gramatík a existenciu jazykového modulu v mozgu (language acquisition device), ktorý umožňuje dieťaťu „nastaviť parametre“ osvojovaného jazyka. Teoretici okolo Chomského zároveň tvrdia, že ich teóriu o vrodenej jazykových dispozícií podporuje i logický dôkaz (GOLD 1967) o nenaučiteľnosti generatívnych jazykov len na základe konečného počtu pozitívnych príkladov. Chomsky však nevysvetľuje, ako by sa takéto jazykové dispozície mohli u človeka vyvinúť. Tu sú veľmi nápomocné simulačné modely, ktoré ponúkajú hypotézy evolúcie rôznych aspektov jazyka ako syntax a významy slov (PERFORS 2002), i samej univerzálnej gramatiky (pozri kapitolu *O nevyhnutnosti univerzálnej gramatiky* a referencie v nej). Alternatívne vysvetlenie povahy vrodenej dispozícií ponúkajú konekcionistické modely. Napríklad Kirby a Christiansen (2003) hypotetizujú na základe simulácií učenia rôznych typov jazykov a na základe psychologických experimentov, že jazykové univerzálne nemusia byť dôsledkom vrodenej biologickej adaptácie mozgu smerom k univerzálnej gramatike, ale že sa dajú vysvetliť pomocou adaptácie (vlastností a ohraničení) všeobecných mechanizmov, ktoré determinujú schopnosť človeka učiť sa a spracovávať sekvenčné dáta.

S otázkou vrodenej dispozícií je úzko prepojená otázka akvizície jazyka. Chomsky prikladá minoritný význam učeniu, tvrdiac, že prostredie je len spúšťačom (trigger) vrodenej mechanizmov nastavovania parametrov jazyka. Alternatívnu teóriu však ponúkajú konekcionistické prístupy a štatistický pohľad na jazyk: vďaka štruktúre si dieťa môže všimnúť distribučné charakteristiky slov, slovných spojení a viet v jazyku (SAFFRAN, NEWPORT, ASLIN 1996; SEIDENBERG 1997; ELMAN et al. 1997). Takáto štatistická informácia (neexistencia niektorých vetných foriem) môže predstavovať negatívne príklady v *implicitnom* tvare a podporovať tak učenie, aspoň teoreticky.⁷ Okrem toho Baker a McCarthy (1981) argumentujú, že pri štatistickom pohľade ani Goldove tvrdenia

⁷ Angluin (1988) dokázal, že stochastické bezkontextové jazyky sú naučiteľné len z pozitívnych príkladov.

nepredstavujú logický problém pre akvizíciu jazyka, pretože Goldove predpoklady a závery sú v takom prípade irelevantné.

Ako vidieť, výpočtové modely (vrátane konekcionistických) ponúkajú hypotézy, ktoré kontrastujú s klasickými teóriami učenia jazyka. V ďalšom texte sa zameriame na proces akvizície jazyka a systematicky uvedieme prehľad konekcionistických modelov, ktoré sa zameriavajú na rôzne aspekty jazykovej schopnosti na lexikálnej úrovni i na úrovni viet.

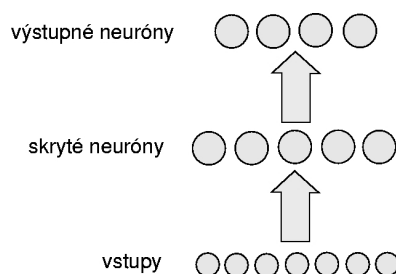
2. Lexikálne spracovanie jazyka

V konekcionistických modeloch orientovaných na lexikálnu úroveň jazyka pracujeme s reprezentáciami slov. Každé slovo je v podstate symbol, ktorý má svoju ortografickú (vizuálnu) alebo phonologickú (akustickú) formu, a ktoré má nejaký význam (sémantický obsah). Formy i význam sú v konekcionistických modeloch reprezentované oddelene a rôznymi spôsobmi, najčastejšie distribuovane pomocou vektorov čísel. Tieto reprezentácie sa potrebujú pri modelovaní rôznych lingvistických procesov.

Väčšina lingvistických procesov je **kváziregulárna** v tom zmysle, že medzi vstupmi a výstupmi existuje značná systematickosť, ale existujú aj viaceré výnimky. Na vysvetlenie takéhoto procesu sa pri symbolovom prístupe predpokladá, že systematickosť je reprezentovaná pomocou explicitných pravidiel, zatiaľ čo výnimky sa zabezpečujú osobitným asociačným mechanizmom (PINKER 1999). Konekcionistický prístup ponúka alternatívny pohľad, podľa ktorého existuje len jeden systém a jednotlivé vstupno-výstupné páry sa navzájom líšia iba *stupňom svojej konzistentnosti* z pohľadu vstupno-výstupnej transformácie. Tento alternatívny pohľad sa v literatúre objavil predovšetkým v morfológii a čítaní.

2.1. Morfológia

Spomedzi morfológických procesov bol stredobodom pozornosti problém tvorenia minulého času sloves v angličtine. Väčšina anglických sloves príberá v minulom čase koncovku /-ed/ (*walk/walked*), avšak pri nepravidelných slovesách jednoznačná systematickosť neexistuje (*go/went, take/took, make/made* atď.), existuje nanajvýš niekoľko podskupín (napr. *keep/kept, sleep/slept*). Napriek tomu sa neurónová sieť dokáže naučiť tvoriť minulý čas všetkých sloves a zovšeobecňovať svoju znalosť pri predložení nových tvarov. Okrem toho neurónová sieť dokáže **simulovať proces akvizície** minulého času sloves u dieťaťa, ktorý má tri typické fázy. Zjednodušene povedané, najprv dieťa správne tvorí minulé tvary u pravidelných i tých niekoľkých nepravidelných sloves, ktoré už pozná, potom však začne nesprávne aplikovať pravidlo /-ed/ aj na nové nepravidelné slovesá (*overregularization*), no napokon sa minulé tvary všetkých nepravidelných sloves naučí správne. Keďže tento proces sprevádza dočasný pokles korektnosti siete v druhej fáze, hovorí sa o učení v tvare U (*U-shaped learning*).⁸



⁸ Symbolové vysvetlenie tohto procesu by sa dalo formulovať takto: v prvej fáze dieťa len memoruje minulé tvary, v druhej fáze začne aplikovať pravidlo /-ed/ a až v tretej fáze nájde rovnováhu medzi oboma cestami.

Obrázok 1. Dvojvrstvová⁹ dopredná neurónová sieť (viacvrstvový perceptrón). Šípky označujú úplné prepojenie medzi susednými vrstvami, t. j. od každého neurónu nižšej vrstvy vedie váhované spojenie ku každému neurónu v nasledujúcej vrstve. Počty vstupov a výstupných neurónov sú vo všeobecnosti rôzne, závisia od reprezentácie použitých dát. Počtom skrytých neurónov určujeme zložitosť modelu. Takáto sieť je vhodná pre úlohy, ktoré sa dajú formálne opísať ako hľadanie statického (nelineárneho) vzťahu medzi vstupmi a výstupmi na základe predložených párov (x,y) , kde x je vstup a y je požadovaný výstup. Perceptróny sa najčastejšie učia pomocou metódy spätného šírenia chyby, ktorá podľa gradientu chyby minimalizuje rozdiely medzi skutočnými a požadovanými výstupmi.

Ako prví modelovali tento proces Rumelhart a McClelland (1986). Trénovali jednovrstvovú doprednú sieť (klasický perceptrón), ktorá mapovala korene slovies na ich minulé tvary, avšak dosť nerealistickým spôsobom. Po kritike tohto modelu (PINKER, PRINCE 1988) navrhli Plunkett a Marchman (1993) kognitívne prijateľnejší model pomocou dvojvrstvovej doprednej siete (obrázok 1), použijúc inkrementálny režim pri trénovaní (postupný nárast z 20 na 500 slovies). Tento model zároveň predikoval (podobne ako aj ďalší model, PLUNKETT, JUOLA 1999), že sieti je potrebné predložiť *kritické množstvo* slovies, aby mohlo dôjsť k prechodu od memorizácie k uplatneniu pravidla /-ed/.

Ďalšiu vlnu záujmu o problém podnietil kritický článok symbolikov (PRASADA, PINKER 1993), ktorého hlavným argumentom bolo, že neurónové siete zovšeobecňujú len na základe frekvencie výskytu slov a na základe fonologickej podobnosti medzi slovami. Pravidelných anglických slovies je veľa (high type frequency), ale jednotlivé slovesá sa vyskytujú zriedkavejšie (low token frequency), čo podľa Prasadu a Pinkera umožňuje sieti vytvoriť si dosť heterogénnu primárnu (default) kategóriu, na ktorú sa vzťahuje pravidlo /-ed/. Na druhej strane nepravidelných anglických slovies je relatívne málo, ale častejšie sa vyskytujú v jazyku, čo umožňuje memorizáciu ich tvarov minulého času podľa jednotlivých úzkych fonologických podskupín (napr. alternáciu i-a pre *sing/sang*, *ring/rang*, o-e pre *grow/grew*, *blow/blew* atď.). Prasada a Pinker zároveň predikovali, že nijaký konekcionistický model nedokáže simulovať správne zovšeobecňovanie pre primárnu kategóriu slovies, ak je táto málo početná a navyše sa nevyskytuje často.¹⁰ Takýto nedostatok multi-lingvistickej validity by bol vážnym problémom konekcionistických modelov morfológie.

Avšak na tento problém primárnej menšiny promptne reagovali Hare, Elman a Daugherty (1995) a ukázali, že ich viacvrstvová dopredná sieť dokázala správne aplikovať primárne pravidlo /-ed/ i v prípade, keď trieda pravidelných slovies predstavovala menšinu v trénovacej množine. Z toho možno usúdiť, že zovšeobecnenie v neurónovej sieti nemusí striktné závisieť od podobností reprezentácií nových slov (v tomto prípade fonologických) so známymi slovami v danej kategórii. Výsledky tejto práce ukazujú, že ak ostatné kategórie slov (v tomto prípade podskupiny nepravidelných slovies) majú výraznú vnútornú štruktúru, tá môže napomôcť správne zovšeobecneniu v rámci primárnej skupiny, aj keď v nej takáto štruktúra absentuje (pravidelné slovesá v angličtine sú navzájom odlišné či už v písanej alebo hovorenej forme).

Máme tu teda dve možné vysvetlenia, ako môže fungovať proces tvorby minulého času u človeka. Spomínané modely sú typickým príkladom eliminačného konekcionistického prístupu, ktorého silnou stránkou je to, že je v zhode s viacerými empirickými dátami. Dokonca i niektorí oponenti konekcionizmu pripúšťajú, že mechanizmy neurónových sietí vnášajú svetlo do komplexných vývojových procesov pozorovaných počas akvizície minulého času slovies.

⁹ Vstupnú vrstvu nepovažujeme za vrstvu neurónov, lebo nemajú aktivačnú funkciu.

¹⁰ To je napríklad prípad nemeckých podstatných mien, ktoré priberajú v množnom čísle až 60 rôznych koncoviek. Napriek tomu ľudia považujú /-s/ za primárnu, štandardnú koncovku a používajú ju pri novotvaroch, aj keď táto koncovka sa často v nemčine nevyskytuje (MARCUS A KOL., 1995).

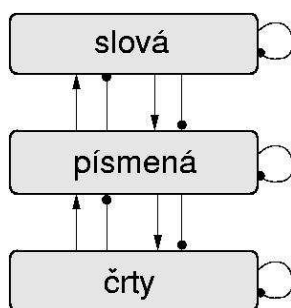
2.2. Čítanie

Psychické procesy, ktoré sú aktívne pri čítaní (vo všeobecnosti v kontexte), sú veľmi komplexné, počnúc rozpoznávaním ortografických črt jednotlivých grafém (písmen) cez syntaktickú, sémantickú a pragmatickú analýzu a končiac integráciou s kognitívnymi znalosťami všeobecne. Konekcionistické modely boli doteraz aplikované len na jednoduchšie aspekty čítania, ako sú vizuálne rozpoznávanie slov a čítanie slov.

Vizuálne rozpoznávanie slov

Jeden z prvých konekcionistických modelov v tejto oblasti bol interaktívno-aktivačný (IA) model vizuálneho rozpoznávania slov (McCLELLAND, RUMELHART 1981; RUMELHART, McCLELLAND 1982). Vstupné neuróny reprezentovali vizuálne črty jednotlivých grafém (čiary s rôznou orientáciou a na rôznych pozíciách v rámci slova), neuróny druhej vrstvy odpovedali jednotlivým písmenám a neuróny tretej jednotlivým slovám (obrázok 2). Sieť mala excitačné i inhibičné spojenia v rámci vrstiev i medzi susednými vrstvami, smerom zdola nahor i zhora nadol. Spojenia boli ručne nastavené tak (čiže sieť sa neučila), aby sieť správne fungovala. Proces rozpoznania slova prebiehal nasledovne: Prezentované slovo aktivovalo v prvej vrstve neuróny, ktoré reprezentovali existujúce črty slova. Tieto neuróny potom cez spojenia poslali svoje aktivity do druhej vrstvy, kde sa zaktivizovali neuróny reprezentujúce písmená v aktuálnom slove, a tie potom zaktivizovali neuróny v tretej vrstve. Výstupné neuróny medzi sebou súťažili (vdďaka inhibičným spojeniam), až ostal aktívny len jeden víťaz. Tento IA model dokázal vysvetliť viaceré psycholingvistické javy, hlavne javy týkajúce sa kontextových efektov pri rozpoznávaní písmen: napr. správne rozpoznanie zašumených písmen na základe kontextu a/alebo znalosti slov, alebo rýchlejšie rozpoznanie písmen v slovách než v nezmyselných reťazcoch (word superiority effect).

Napriek svojej úspešnosti mal tento model aj značné obmedzenia. Spomeňme aspoň dve. Fungoval iba pri slovách so štyrmi písmenami a nebol pozične invariantný (neuróny prvej a druhej vrstvy sa museli vyskytovať štyrikrát). Cieľom ďalších modelov bolo tieto nedostatky odstrániť (napr. McCLELLAND 1986; MOZER 1991).



Obrázok 2. Interaktívno-aktivačný model vizuálneho rozpoznávania slov (McCLELLAND, RUMELHART 1981; RUMELHART, McCLELLAND 1982). Susedné vrstvy neurónov sú poprepájané excitačnými spojeniami (označené šípkami) a inhibičnými spojeniami (krúžky). V rámci každej vrstvy existujú len inhibičné spojenia.

Čítanie slov

Podobne ako v morfológii i proces čítania slov predstavuje kváziregulárnu doménu, hoci z pohľadu systematickosti ortograficko-fonologickej transformácie sa jednotlivé jazyky

značne líšia.¹¹ Psychologické štúdie predikujú, že kváziregulárnosť by mohla byť zabezpečená *dvoma cestami* (dual route), ktoré človek pri čítaní využíva. Jedna z ciest je priamočiara, fonologická, a zabezpečuje aplikáciu pravidiel výslovnosti (sublexikálna cesta). Druhá cesta zabezpečuje správnu výslovnosť výnimiek (lexikálna cesta) a môže viesť cez sémantiku. Táto hypotéza pramení zo štúdií normálnych subjektov, ale aj pacientov s poruchami čítania (COLTHEART 1987).

Dominantná väčšina doterajších prác bola zameraná na angličtinu.¹² Ako prví sa venovali problému čítania Sejnowski a Rosenberg (1987) svojím modelom NETtalk (dvojvrstvová dopredná sieť). Slová boli sieti predkladané pomocou kľzavého okna, ktoré v každom kroku obsahovalo trojicu po sebe idúcich grafém (počítali sa i medzery medzi slovami). V každom kroku mala sieť vysloviť fonému asociovanú s grafémou v strede aktuálneho okna. Keďže v angličtine neexistuje jedno-jednoznačné priradenie medzi grafémami a fonémami (a grafém je viac), niektoré grafémy sa neasociovali so žiadnou fonémou. Napriek celkom dobrej úspešnosti čítania, modelu NETtalk chýbala nadväznosť na psychologické dáta.

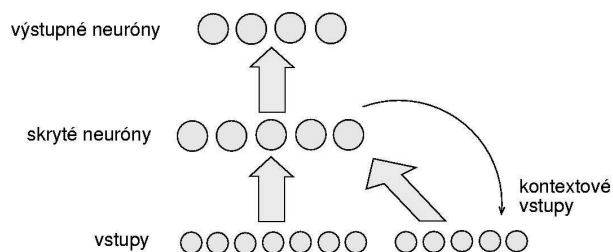
Prvý detailný psychologický model čítania prezentovali Seidenberg a McClelland (1989). Použili rovnakú sieť ako v NETtalk a brali do úvahy množinu (takmer 3000) izolovaných slov, ktoré boli sieti predkladané ako celok: v ortografickej forme na vstupe, vo fonologickej forme na výstupe. Na reprezentáciu použili ručne navrhnuté príznaky (Wickelfeatures), ktoré boli potom v literatúre dosť spochybňované. Seidenberg a McClelland ukázali, že ich model ako jeden systém dokáže správne čítať slová s nepravidelnou výslovnosťou, ale aj nezmyselné, hoci vysloviteľné reťazce (nonwords). Tým spochybnili nutnosť existencie duálnej cesty pri čítaní.

Kritizované nedostatky tohto modelu (COLTHEART et al. 1993), napr. jeho nízka úspešnosť pri čítaní nezmyselných reťazcov v porovnaní s človekom, a nedodržanie skutočných frekvencií výskytu slov v jazyku počas tréningu (frekvenčné efekty sú u človeka dôležité) viedli k návrhu vylepšeného modelu Plauta a kol. (1996), ktorí trénovali sieť na tých istých slovách, ale podľa ich skutočných frekvencií. Tento model sa odlišoval od svojho predchodcu v učiacom algoritme (metóda spätného šírenia chýb v priebehu času), použitých reprezentáciách (vylepšené distribuované reprezentácie) i architektúre: výstupné neuróny mali aj modifikovateľné spätné spojenia na skryté neuróny a tiež slučky samy na seba. Tým vznikol na výstupe dynamický systém, ktorého vektor aktivity po pár cykloch mohol skonvergovať do ustáleného stavu – atraktora (tento model sa preto nazýva aj atraktorová neurónová sieť). Takéto ustáľovanie odpovede neurónovej siete (dané počtom potrebných cyklov) bolo interpretované ako analógia reakčných časov pri čítaní, čo je tiež dôležitá psychologická charakteristika čítania. Týmto modelom sa Plaut a kol. viac priblížili k experimentálnym výsledkom, vrátane simulácie rôznych porúch pri čítaní (dyslexií). Na základe simulácií argumentovali, že človek, ako sme spomínali v úvode tejto časti, využíva dve cesty pri normálnom čítaní (pri rôznych typoch dyslexie býva jedna alebo druhá cesta porušená) a že pri akvizícii čítania vzniká proces del'by práce medzi oboma cestami.

Vo svojom doteraz poslednom modeli Plaut (1999) pridal ďalšie vylepšenie: použitím jednoduchej rekurentnej siete (obrázok 3) sa dalo simulovať čítanie ako sekvenčný proces (fonémy na výstupe sa generujú ako sekvencia). Tento pokus možno považovať za prvý krok smerom vpred k vysvetleniu temporálnych aspektov čítania.

¹¹ Niektoré jazyky majú veľmi pravidelnú výslovnosť (slovenčina, taliančina), iné zase veľmi nepravidelnú (hebrejčina, japonské kanji). Angličtina je niekde v strede, takže nájdeme v nej dosť pravidiel výslovnosti, ale aj výnimiek (napr. *have, women*).

¹² Konekcionistické modelovanie čítania slov v slovenčine je len v počiatočnom štádiu. Prvou lastovičkou je diplomová práca (LACHOVÁ 2004), kde autorka na základe simulácií poukazuje na to, že v slovenčine zrejme postačuje sublexikálna cesta vďaka oveľa väčšej pravidelnosti vo výslovnosti, než je to v prípade angličtiny.



Obrázok 3. Jednoduchá rekurentná sieť (simple recurrent network, SRN; ELMAN 1990). V porovnaní s viacvrstvovým perceptrónom je obohatená o spätnú väzbu od skrytých neurónov. Široké šípky označujú úplné prepojenie medzi susednými vrstvami neurónov pomocou adaptovateľných váh. Tenká šípka označuje kopírovanie aktivít skrytých neurónov na kontextové vstupy, čím sa aktivity skrytých neurónov v aktuálnom kroku stanú kontextovými vstupmi v nasledujúcom kroku. Túto sieť možno trénovať aj pomocou štandardnej metódy spätného šírenia chýb, ale aj pomocou efektívnejších algoritmov. SRN je vhodná na niektoré časopriestorové úlohy, kde kontextové vstupy slúžia ako pamäť (s exponenciálnym zabúdaním do minulosti) a umožňujú sieti generovať výstup nielen podľa aktuálneho vstupu, ale aj podľa ľavého kontextu.

2.3. Spracovanie reči

Spracovanie fonologických reprezentácií je proces analogický vizuálnemu rozpoznávaniu slov. Najznámejším modelom v tejto kategórii je TRACE (McCLELLAND, ELMAN 1986), ktorého architektúra i princípy sú analogické IA modelu vizuálneho rozpoznávania slov (obrázok 2): v modeli TRACE neuróny prvej vrstvy zodpovedajú fonetickým črtám foném, v druhej vrstve samotným fonémam a v tretej vrstve slovám. Vstupné slovo sa prezentuje sieti sekvenčne zľava doprava po fonémach. Rozpoznanie slova nastane v okamihu, keď jeden z kandidátov v tretej vrstve výrazne inhibuje všetky ostatné neuróny v tej vrstve. Rozpoznanie však môže nastať i predtým, ako sa sieti prezentuje celý vstup. TRACE dokáže simulovať rôzne empiricky pozorované javy, napr. kategorickú percepciu foném alebo vplyv lexikálneho kontextu pri identifikácii foném. Tieto efekty model vysvetľuje na základe povoleného obojsmerného toku informácie: zdola nahor i zhora nadol. Avšak otázka, či človek pri percepcii reči využíva lexikálnu znalosť (informácia zhora nadol), je otvorená (MASSARO 1989).

Iné konekcionistické modely tiež vysvetľujú tieto percepčné efekty, ale bez využitia lexikálnej informácie. Napríklad Norris (1993) trénoval SRN (obrázok 3), ktorá sa naučila rozpoznávať slová bez použitia lexikálnej znalosti. Norris si však všimol jeden nedostatok svojho modelu (ktorý v modeli TRACE nefiguroval, lebo tam sa mohli kandidáti inhibovať) – nemožnosť revidovať svoje rozhodnutia.¹³ Tento nedostatok bol odstránený v kombinovanom modeli, ktorý zahrnul aj lexikálnu kompetíciu medzi výstupnými neurónmi (NORRIS 1994). Niektoré ďalšie konekcionistické modely však tiež ukázali, že modelovanie lexikálnych efektov sa dá dosiahnuť bez toku informácie zhora nadol (napr. GASKELL, HARE, MARSLEN-WILSON 1995). Debata okolo interaktívneho prístupu verzus prístup zdola nahor je ilustráciou toho, ako použitie konekcionistických modelov podnietilo hlbší empirický výskum pri hľadaní odpovede na túto otázku.

2.4. Artikulácia slov

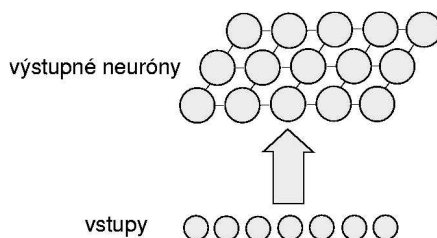
Prvé produkčné modely boli zamerané na lexikálnu úroveň a zaoberali sa lexikálnym prístupom, produkciou foném a morfológickými aspektami produkcie. Jedným z prvých

¹³ Napríklad pri slove *catalog* by sa najprv aktivoval víťaz pre *cat*, potom pre *log*. Sieť by nemala možnosť finálne odpovedať víťazom pre *catalog*.

modelov bola IA neurónová sieť (DELL 1986), podobná modelu TRACE. Model mal tri vrstvy neurónov, ktoré zodpovedali syntaktickým (lexikálne neuróny), morfológickým a fonológickým vlastnostiam slov. Na rozdiel od modelu TRACE, Dellov model začínal spracovávať informáciu zvrchu, aktiváciou lexikálnych neurónov, smerom dolu k morfológickej a fonológickej úrovni. Tento model bol tiež statický (s ručne nastavenými váhami). Dokázal vysvetliť rôzne chyby vznikajúce pri hovorení, napr. substitúcie, chybné vynechanie či pridanie fonémy. Zároveň poskytol kvantitatívne predikcie týkajúce sa vyvolania fonológických foriem z pamäti počas hovorenia. Tieto predikcie boli neskôr potvrdené experimentálne (DELL 1988).

Dell a kol. (1997) rozšírili tento model na simulovanie rôznych druhov afázie – jazykových porúch prejavujúcich sa i pri pomenovaní obrázkov. Chyby môžu byť sémantické (namiesto *cat* pacient povie *dog*), fonológické (*fat*), zmiešané (*rat*), bez nejakej súvislosti s cieľovým slovom (*pen*) alebo vyúsťujúce v nezmyselný výraz (*lat*). Trojvrstvový model mal jednu vrstvu určenú pre sémantické príznaky konceptov, druhú pre samotné koncepty (slová) a tretiu na reprezentovanie foném. Susedné vrstvy neurónov boli plne prepojené váhami v oboch smeroch. Vďaka obojsmernosti spojení medzi neurónmi mohli nastávať rôzne chybové efekty. Predložený vzor sémantických príznakov evokoval aktivity na konceptoch, ktoré tieto príznaky zdieľali (tým sa pri *cat* aktivoval aj koncept *dog*). Neskôr, po aktivovaní fonológických príznakov v prvej vrstve, sa spätným šírením informácie ku konceptom mohli aktivovať aj fonológicky blízke koncepty (*mat*).

Hlavnou nevýhodou Dellovho modelu bola jeho neschopnosť učiť sa. Dell, Juliano a Govindjee (1993) použili SRN (obohatenú o spätnú väzbu od výstupných neurónov), ktorá sa naučila zobrazovať slová na sekvencie fonológických segmentov. Tento model takisto dokázal vysvetliť rôzne fonológické chyby vznikajúce pri reči, ale niektoré nie (napr. vzájomná zámena dvoch začiatočných foném v *left hemisphere*, ktorá by vyúsťila v *heft lemisphere*). Efektívnejší fonológický model s podobným prístupom navrhli Plaut a Kello (1999).



Obrázok 4. Samoorganizujúca sa mapa (SOM; KOHONEN 2001). Vstupy sú úplne prepojené adaptovateľnými váhami so všetkými neurónmi v mape, usporiadanými do pravidelnej 2D mriežky. SOM zobrazí distribuovane reprezentované vstupy na lokalistické odpovede v mape (zobrazenie redukuje dimenziu). Požadovanou vlastnosťou tohto zobrazenia je vlastnosť zachovania topológie: blízke vstupy sa zobrazia na blízke neuróny v mape. Takéto zobrazenie možno učením dosiahnuť práve vďaka existencii neurálnej štruktúry v mape (každý neurón „vie“, akých má susedov), čo sa využije pri mechanizmoch súťaženia a kooperácie medzi neurónmi. Učenie má charakter samoorganizácie, pretože neurónom v mape neposkytujeme informáciu o požadovanom výstupe; SOM sa organizuje podľa štatistických vlastností vstupov.

2.5. Modelovanie lexikálnych máp

Dopredné a rekurentné neurónové siete neboli jedinými konekcionistickými modelmi, ktoré sa v literatúre objavili v súvislosti s lexikálnym modelovaním jazyka. Ďalšia kategória modelov využíva samoorganizujúcu sa mapu (SOM; KOHONEN 2001; FARKAŠ 1997), ktorá tiež umožňuje modelovať rôzne lexikálne efekty a navyše je biologicky zaujímavejšia, pretože ide o model samoorganizácie (a nie učenia s učiteľom). SOM (obrázok 4) realizuje zobrazenie, ktoré zachováva topologické vzťahy v dátach (topology-

preserving property), čo je vlastnosť systematicky figurujúca v mozgovej kôre, kde máme napr. rôzne sensorické mapy (napr. v doméne vizuálnej alebo akustickej), ktoré si mozog interne vytvára. Predikuje sa, že v mozgu môžu existovať i mapy vyšších, abstraktnejších entít, napr. konceptov (RITTER, KOHONEN 1989), čo by mohlo vysvetliť rôzne kognitívne poruchy pozorované u pacientov, napr. neschopnosť pomenovať predmety patriace do nejakej kategórie. Vysvetlenie spočíva práve v predpoklade o topologickom (kategorigickom) usporiadaní neurónov v mape: simulácia lokálnej poruchy s veľkou pravdepodobnosťou poškodí len neuróny v jednej časti mapy, zodpovedajúce nejakej kategórii.

Niektoré modely založené na SOM majú komplexnejšiu architektúru, čo rozširuje ich použiteľnosť. Napríklad Dislex (MIIKKULAINEN 1997) predstavuje systém troch prepojených máp (fonologická a ortografická mapa sú prepojené so sémantickou mapou), čo umožňuje ich vzájomné aktivovanie. Tým sa dajú modelovať jednotlivé jazykové procesy, napr. čítanie (ortograficko-sémantická linka) alebo artikulácia (sémanticko-fonologická linka), ale aj rôzne poruchy ako dyslexia alebo špecifické afázie (napr. poruchy pomenovania vecí), a to pomocou simulovaných lokálnych porúch v modeli. Na báze tohto modelu navrhli Li, Farkaš a McWhinney (2004) model DevLex, pozostávajúci z *rastúcich* máp, ktorý bol navrhnutý ako psychologicky a kognitívne prijateľný model vznikajúceho lexikónu u malých detí (pozri aj FARKAŠ 2003). Ďalším spôsobom spájania máp je ich hierarchické usporiadanie ako v Miikkulainenovej práci (1990), kde pyramidovito usporiadané vrstvy boli použité na reprezentáciu slov, viet i celých skriptov. Takýto systém pracuje ako epizodická pamäť skriptov, ktoré sa dajú vyvolať pomocou čiastočnej alebo zašumenej informácie. SOM môže taktiež slúžiť ako súčasť nejakého heterogénneho systému (napr. syntaktického analyzátora, MAYBERRY, MIIKKULAINEN 1998; pozri aj časť 3.4).

Na záver tejto časti zhrňme, že rôzne lexikálne procesy, o ktorých sa zdá, že vyžadujú explicitnú formuláciu a rôzne mechanizmy spracovania, možno úspešne modelovať integrovane pomocou distribuovaných konekcionistických modelov učiacich sa vstupno-výstupné vzťahy v kváziregulárnej doméne. V rámci modelovania jazyka ako komplexného javu je pororumenie týchto procesov predpokladom a stavebným blokom pre vyššiu, vetnú úroveň analýzy jazyka.

3. Spracovanie viet

Pri vetnej analýze už prichádza do hry viacero procesov, ktoré v konečnom dôsledku umožňujú človeku porozumieť vete (alebo vysloviť vetu). Doterajšie konekcionistické modely boli navrhnuté tak, aby riešili jeden z niekoľkých podproblémov: predikcia nasledujúcich slov, indukcia gramatiky jazyka, reprezentovanie štruktúrovaných dát, syntaktická analýza, komprehenzia viet a produkcia viet. Ďalej si uvedieme stručný prehľad modelov podľa tohto členenia.

3.1. Predikcia slov

Pri počúvaní hovoreného slova človek zvyčajne dokáže s istou presnosťou predikovať, aké slová môžu nasledovať. Je to vďaka syntaktickým, sémantickým a pragmatickým obmedzeniam, ktoré determinujú štruktúru jazyka. Aj keď predikcia nasledujúceho slova nie je finálnym produktom pri spracovaní jazyka na úrovni vety (cieľom je porozumenie vete, t. j. vybudovanie jej sémantickej reprezentácie v mysli), predikcia bude zrejme jedným z užitočných komponentov modelu jazyka. Preto je táto úloha relevantnou a vzbudila pozornosť i v konekcionistickej komunite.

Ako prvý modeloval predikciu nasledujúcich slov Elman (1990), ktorý na tento účel navrhol svoj model jednoduchej rekurentnej siete (SRN, obr. 3). Elman aplikoval SRN

najprv na predikciu nasledujúcich grafém v súvislom texte. Siet' teda v každom kroku prečítala aktuálnu grafému a jej ľavý kontext, a mala predikovať nasledujúcu grafému (vrátane medzier medzi slovami). Elman ukázal, že predikcie začiatku nového slova boli vždy najhoršie. Tento prístup implikuje, že schopnosť predikcie nasledujúcej fonémy môže byť primárnym mechanizmom, ktorý novorodenci používajú pri segmentácii slov (SAFFRAN, NEWPORT, ASLIN 1996). Elman takisto vyskúšal SRN na predikciu nasledujúcich slov (tiež lokalisticky kódovaných) v korpuse jednoduchých viet (*John runs, Mary loves John, John broke glass* atď.). SRN sa naučila s istou mierou presnosti predikovať vhodných kandidátov v každom kroku, a to vďaka emergencii interných reprezentácií, ktoré si SRN vybuodovala vo svojej skrytej vrstve. Zaujímavým zistením bolo, že tieto distribuované reprezentácie slov boli štruktúrované, čím ich klasterizácia viedla k detekcii jednotlivých syntaktických i sémantických kategórií. Z toho vyplýva, že i zo samotných viet sa dá do istej miery extrahovať implicitná znalosť gramatiky, podľa ktorej boli vety generované. Podobné výsledky bolo možné pozorovať i v zložitejších vetách (ELMAN 1991).

Druhým významným výsledkom podľa Elmana bolo zistenie, že SRN sa naučí správne predikovať nasledujúce slová v komplexných vetách len vtedy, ak sa jej najprv predkladajú jednoduchšie vety. Podobne Elman (1993) argumentoval, že učenie komplexnejších viet sieti uľahčíme, ak jej budeme spočiatku narúšať rozsah pamäti, čo jej spočiatku umožní predikovať len slová v rámci jednoduchších viet. Tieto zistenia sa považovali za veľmi dôležité, pretože boli aj v súlade s vplyvnou hypotézou **menej-je-viac** (NEWPORT 1990), podľa ktorej tým, že kognitívny aparát dieťaťa je spočiatku limitovaný, zvládne dieťa jazyk (i cudzí) lepšie než dospelý človek, ktorého jazykové štruktúry sú už úplne vyvinuté.

I keď Elmanove zistenia boli veľmi vplyvné pre konekcionistické teórie akvizície jazyka, boli neskôr spochybnené. Rohde a Plaut (1999) zistili, že manipulácia štruktúry vstupných viet môže sieti, naopak, aj škodiť a že manipulácia rozsahu pamäti siete nemá takmer nijaký vplyv na učenie. Ich vysvetlením bolo, že učiaca sa sieť prirodzene začína so slabšou pamäťou (pre náhodne inicializované váhy), ktorú si musí postupne vybudovať. Preto aj keď je SRN vystavená od začiatku zložitým vetám, jej spočiatku neustálené váhy jej spočiatku umožňujú správne predikovať len slová v jednoduchších vetách, a až neskôr v komplexnejších vetách (napr. predikcia *said* vo vete *The boy who broke the vase said...*). Ako Rohde a Plaut argumentujú, manipulovaním rozsahu pamäti Elman len simuloval jej prirodzený vývoj, a preto táto manipulácia nemá na učenie nijaký efekt.¹⁴

Christiansen (1994) rozšíril gramatiku použitú Elmanom a zahrnul do nej predložkové frázy, genitívne nominálne frázy (*John's friend*), konjunkcie nominálnych fráz (*John and Mary walk*) a vetné komplementy (*I thought Peter said he left*). Jedna verzia gramatiky produkovala súvetia s vnorenými vetami (*the boy who came loves Mary*), druhá verzia zase krížne závislosti (cross dependencies; *John and Mary walked and hopped, respectively*). V oboch prípadoch sa SRN celkom dobre naučila gramatiky týchto jazykov a jej správanie sa dosť približovalo experimentálnym výsledkom na podobných vetách s rovnakou zložitosťou. Christiansen a Chater (1999b) ďalej rozšírili tieto výsledky a detailnejšie porovnali správanie neurónovej siete so správaním človeka. Zaujímavým výsledkom v kontexte správania sa rekurentných neurónových sietí je ich tzv. architekturný bias (CHRISTIANSEN, CHATER 1999b). Ešte skôr Kolen (1994) ukázal, že už v rekurentnej sieti s náhodne inicializovanými malými váhami vznikajú v jej stavovom priestore klastre aktivít, ktoré priestorovo kódujú históriu vstupov (kontext), na základe čoho potom sieť predikuje nasledujúce symboly. Tento stav bol teda pozorovaný už *pred učením*, a nevznikol len ako produkt učenia, na čo poukazoval Elman. Tiño, Čerňanský a Beňušková (2004) analyzovali tento jav a ukázali, že klastre aktivít zodpovedajú markovovským

¹⁴ Elmanovi sa však nepodarilo zopakovať výsledky Rohdeho a Plauta a ich vzájomná komunikácia potom stroskotala, takže problém ostáva otvorený (Elman, osobná komunikácia).

predikčným kontextom, na základe ktorých možno skonštruovať predikčný model s markovovskými vlastnosťami. Poznatok o existencii architekturného biasu poukazuje na to, že už nenaučená sieť má nejaké predikčné schopnosti. Z toho vyplýva, že na to, aby sme porozumeli, nakoľko prispieva učenie pri indukovaní nejakého požadovaného predikčného správania siete, treba porovnať naučenú sieť s nenaučenou.

3.2. Indukcia gramatík

Trénovanie siete ako prediktora nasledujúceho slova môžeme zaradiť do širšej kategórie konekcionistických modelov, ktoré sa učia gramatiku nejakého jazyka podobného prirodzenému jazyku. Učia sa ju **implicitne**, trénovaním na vetách generovaných danou gramatikou. Konekcionistické modely indukcie gramatík (pozri prehľad vo FORCADA 2002) sú tiež typickým príkladom spracovania symbolových dát subsymbolovým spôsobom. Modely boli testované predovšetkým na formálnych gramatikách rôznej zložitosti podľa Chomského (1956) klasifikácie (pozri kapitolu *O nevyhnutnosti univerzálnej gramatiky*). Chomsky (1965) sformuloval tézu, že bezkontextové gramatiky sú nedostatočným a kontextové zase príliš silným aparátom na zachytenie syntaxe prirodzeného jazyka. Preto prirodzené jazyky svojou zložitou spadajú medzi tieto dve formálne triedy.

Bolo ukázané, že rekurentné neurónové siete sa vedú naučiť jednoduché regulárne jazyky. Jedným zo spôsobov, ako učiť sieť gramatiku, je trénovať ju ako vyššie spomínaný prediktor nasledujúceho slova (CLEEREMANS, SERVAN-SCHREIBER, McCLELLAND 1989; ELMAN 1990). Iným spôsobom je trénovať sieť, aby fungovala ako detektor gramatickosti testovanej vety: na konci vety musí sieť svojím výstupom povedať explicitne, či bola veta gramatická, alebo nie (GILES ET AL. 1992). V oboch prípadoch sa rekurentná sieť naučí imitovať konečný automat, ktorý regulárnu gramatiku generuje. Klasterizáciou aktivít skrytých neurónov siete dostaneme množinu bodov v stavovom priestore siete, ktoré predstavujú jednotlivé stavy konečného automatu. Boli navrhnuté procedúry, ako sa dá automat zo siete formálne vyextrahovať.

Rekurentné neurónové siete boli použité i na učenie bezkontextových jazykov z príkladov, a to s použitím externého zásobníka (ZENG, GOODMAN, SMYTH 1994; MOZER, DAS 1993; GILES ET AL. 1990), alebo dokonca i bez neho (RODRIGUEZ, WILES, ELMAN 1999).¹⁵ Rodriguez a kol. ukázali, že teória formálnych automatov síce predikuje, že na učenie bezkontextového jazyka je nutný zásobník, avšak ich SRN sa bez neho zaobišla tým, že si vybudovala interné počítadlo vo svojom stavovom priestore. Ich sieť sa naučila počítat' (smerom nahor i nadol) pri slovách typu $a^n b^n$ (napr. *aabb*), čo je potrebné pri vetách typu *ak... potom...* alebo súvetiach s vnorenými frázami (napr. *Chlapec, ktorého sledujem, má zmrzlinu*). Rodriguez a kol. úspešne overili sieť aj na dlhších reťazcoch (väčšie n), hoci ich sieť počas trénovania nevidela. Rodriguez (2001) dokonca rozšíril uplatnenie „počítajúcej“ SRN aj pri niektorých kontextových jazykoch.

Ako sme spomínali v úvode kapitoly, na indukciu gramatiky sa však možno pozerat' aj inou optikou než ako na formálny diskretný proces, kde znalosť jazyka znamená znalosť pravidiel gramatiky. Pri alternatívnom, štatistickom pohľade chápeme gramatickosť vety ako **gradovaný fenomén** (graded grammaticality) a znalosť jazyka možno reprezentovať neurónovou sieťou a jej schopnosťou rozpoznávat' i produkovať vety z nejakého jazyka, či už gramatické, alebo nie (ALLEN, SEIDENBERG 1999). Podobne aj o človeku môžeme povedať, že vie jazyk (aspoň dosť dobre), ak rozumieme jeho vetám, aj keď niektoré z nich by boli negramatické. Gramatickosť vety nemusí mať z tohto pohľadu diskretný charakter (napr. veta *Deti malé do školy idú* sa nám zdá menej negramatická ako *Idú malé do školy*

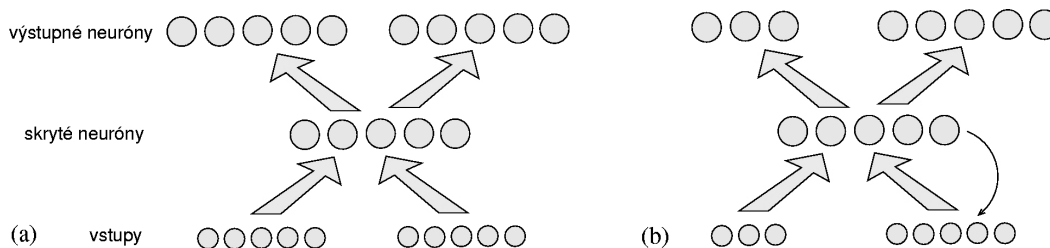
¹⁵ Prvý prístup možno chápať ako implementáciu symbolového prístupu, zatiaľ čo druhý ako príklad eliminačného konekcionizmu.

deti). Allen a Seidenberg argumentujú, že dôležitú úlohu pri porozumení hrá sémantika a že schopnosť človeka posúdiť gramatickosť vety je založená na použití viacerých kritérií, hlavne tých, ktoré sa opierajú o lokálnu štatistickú informáciu sekvencií slov. Parfitt, Tiño a Dorffner (2000) tiež poukazujú na štatistické jazykové modely (natrénované na Brownovom korpuse), ktorých správanie je konzistentné s hypotézou, že gramatická štruktúra vety (wellformedness) je gradovaná, nie absolútna.

3.3. Reprezentovanie štruktúrovaných dát

Schopnosť človeka porozumieť štruktúrnym vzťahom medzi jednotlivými prvkami (konštituentmi) je neoddeliteľnou súčasťou jazyka a kognície všeobecne. Konekcionistické modely boli vo vplyvnom článku Fodora a Pylyshyna (1988) kritizované, že nedokážu reprezentovať a spracovať štruktúrované dáta, ako sú zoznamy, grafy a stromy. A ak dokážu, tak iba v implementácii symbolových prístupov. Táto kritika podnietila návrh niekoľkých konekcionisticky orientovaných výpočtových modelov, ktoré sa snažia riešiť problém reprezentovania (rekurzívnych) štruktúr a tzv. viazania (binding).¹⁶ Možno ich rozdeliť do dvoch tried: jedna trieda modelov je založená na *statických* konjunktívnych kódach (POLLACK 1990; SMOLENSKY 1990; PLATE 1995), druhá na *dynamickom* viazaní reprezentácií (SHASTRI, AJJANAGADDE 1993; HUMMEL, HOLYOAK, 1997). Najprv sa bližšie venujme prvej triede.

V súvislosti s reprezentovaním štruktúr pomocou konjunktívnych kódov sú dôležitými vlastnosťami **systematickosť** a **kompozičnosť** týchto reprezentácií.¹⁷ Systematickosť a kompozičnosť reprezentácie sú predpokladom na možnosť realizácie **holistických** operácií. Koncept holistikosti vyvolal pre svoju nejasnú definíciu trochu zmätok v literatúre. Hammerton (1998) navrhuje nasledujúcu definíciu: Holistická operácia je akýkoľvek výpočet, ktorý sa môže aplikovať na všetky konštituenty objektu súčasne, bez nutnosti lokalizovať, extrahovať či modifikovať jeho konštituenty (ako je to pri konvenčných symbolových prístupoch).



Obrázok 5. (a) Binárna rekurzívna autoasociatívna pamäť (RAAM). Autoasociáciou vzniká na skrytých neurónoch komprimovaná reprezentácia. Komprimácia sa dosahuje pomocou spodnej časti siete, dekomprimácia pomocou hornej časti. Binárna RAAM sa používa na zakódovanie binárnych stromov (pre stromy valencie n treba použiť n skupín neurónov na vstupe i výstupe). Trénuje sa pomocou štandardnej metódy spätného šírenia chýb. (b) Sekvenčná RAAM (SRAAM) sa podobá na SRN, ale funguje ako autoasociátor, takže spätne predikuje aj kontext. Používa sa na kódovanie (a dekódovanie) acyklických symbolových štruktúr.

¹⁶ Napr. vo vete *Janko ľúbi Aničku* treba dať do súvislosti (zviazať) Jana so subjektom vety a Aničku s objektom.

¹⁷ Reprezentácie sú štruktúrne systematické, ak sú podobné entity reprezentované nejakým spôsobom podobne (napr. *John loves Mary* a *Mary loves John* by mali mať podobné reprezentácie). Kompozičnosť (compositionality) reprezentácie predpokladá možnosť syntakticky preusporiadať konštituenty reprezentácie. Distribúované konekcionistické reprezentácie majú vlastnosť tzv. *funkčnej* kompozičnosti (VAN GELDER 1990) bez nutnosti *konkatenatívnej* kompozičnosti (ktorá je typická pre symbolové systémy a vzniká zret'azením reprezentácií jednotlivých konštituentov).

Pollackova (1990) rekurzívna autoasociatívna pamäť (recursive autoassociative memory, RAAM) je jedným z najznámejších a najčastejšie používaných modelov (obrázok 5a). Ako prvý simuloval holistické operácie Chalmers (1990), a to pri transformácii (pomocou doprednej siete) medzi činným a trpným rodom jednoduchých (trojslovných) anglických viet, ktorých komprimované reprezentácie boli vopred skonštruované pomocou separátnych ternárnych RAAM. Chrisnan (1991) simuloval holistické operácie pomocou RAAM, ktorá prekladala anglické vety do španielčiny a späť. Vstupom binárnej RAAM bola veta a jej preklad, a sieť – autoasociátor sa naučila zakódovať význam vety (spoločný pre oba jazyky) na svojej skrytej vrstve. RAAM bola použitá vo viacerých prácach ako prostriedok na vybudovanie reprezentácií syntaktických stromov viet (HO, CHAN 1995; SHARKEY, SHARKEY 1992), ktoré sa potom použili pri syntaktickej analýze viet (časť 3.4). Alternatívou RAAM na tento účel je sekvenčná RAAM (SRAAM), ktorá je vhodná skôr na reprezentovanie stromov s ľubovoľnou valenciou než n-árnych stromov (obrázok 5b). Kwasny a Kalman (1995) argumentujú, že ich SRAAM umožňuje efektívnejšie tréningovanie (pomocou modifikovaného algoritmu spätného šírenia chýb) a lepšie zovšeobecnenie v porovnaní s RAAM. Navrhli techniku linearizácie stromových štruktúr (t. j. jednoznačné zobrazenie stromu na sekvenciu), ktoré potom použili na tréningovanie SRAAM. Napriek zlepšeniam má však SRAAM značné obmedzenia týkajúce sa zovšeobecnenia, ktoré pravdepodobne pramenia z vlastností algoritmu učenia s učiteľom. Zaujímavé vylepšenie ponúka Voegtlin (2002), ktorého model PCA-RAAM s *lineárnymi* neurónmi sa učí v procese *samoorganizácie* pomocou algoritmu rekurzívnej PCA (principal component analysis). Ako autor ukazuje, PCA-RAAM dokáže reprezentovať štruktúrované dáta oveľa vyššej zložitosti, pričom vďaka lineárnej povahe modelu sa zložitejšie štruktúry dajú vytvárať kombináciou jednoduchších štruktúr. Tým sa dosahuje aj pozoruhodne vyššia schopnosť zovšeobecnenia na nové dáta, čo sa týka štruktúry i obsahu.

Alternatívny prístup ku konštruovaniu štruktúrovaných reprezentácií predstavujú holografické redukované reprezentácie (HRR; PLATE 1995). Táto efektívna metóda založená na kruhovej konvolúcii (circular convolution) taktiež podporuje holistické operácie. S modelom RAAM má spoločné to, že dimenzia výslednej reprezentácie štruktúry má fixnú, nemennú dĺžku.¹⁸ Hlavným rozdielom oproti RAAM je to, že holistické reprezentácie vznikajú v HRR *bez učenia*; iba vďaka vlastnostiam použitej matematickej operácie (konvolúcie). Zaujímavou vlastnosťou modelu HRR je aj jeho škálovanie (scaling property): pre jednoduché štruktúry (obsahujúce zopár konštituentov) potrebujeme rádovo tisícrozmerný vektor, avšak na oveľa zložitejšie štruktúry stačí len niekoľko tisícrozmerný vektor (teda nepotrebujeme rádovo viac dimenzií).

Druhá trieda modelov, založená na dynamickom viazaní, využíva časovú synchronizáciu pálenia neurónov. Tieto modely umožňujú reprezentovať a spracovávať štruktúrované dáta a súčasne simulovať s tým súvisiace „vysokúrovňové“ jazykové úlohy ako usudzovanie (reasoning) s využitím inferencií a spracovanie analógií. Model SHRUTI (SHASTRI, AJJANAGADDE 1993) využíva lokalistické reprezentácie konceptov i objektov, zatiaľ čo model LISA (HUMMEL, HOLYOAK 1997) využíva distribuované kódy na reprezentáciu konceptov a lokalistické kódy na reprezentáciu štrukturálnych vzťahov. S tým súvisia výhody i nedostatky oboch prístupov, pričom spoločným nedostatkom oboch modelov je neexistencia efektívnych adaptačných stratégií (učenia). Napriek tomu oba modely dokážu simulovať spomínané kognitívne prejavy človeka. Zaujímavý model založený na dynamickom viazaní, ktorý je schopný adaptácie, je model SSN (LANE, HENDERSON 2001). Ten tiež využíva štruktúrované reprezentácie, a to pri syntaktickej analýze viet.

¹⁸ na rozdiel od Smolenského modelu založenom na tenzorových produktoch, kde dimenzia rastie s komplexitou štruktúry

3.4. Syntaktická analýza viet

Schopnosť výpočtového systému spracovávať štruktúrované dáta úzko súvisí so syntaktickou analýzou viet (sentence parsing), čo je proces vytvorenia syntaktického stromu vety (parse tree) z jej povrchovej formy. Tento strom vyjadruje vzťahy formálnej súvislosti medzi frázami vo vete. Napríklad syntax vety *Jano sedí na veľkom strome* by sme LISP-ovsky zapísali ako štruktúru (Sub (SI (Predl Adj Sub))), kde vidieť podštruktúry ako PF → Predl Adj Sub a SF → SI PF. Svojou explicitnosťou vyvolala syntaktická analýza viet veľký záujem hlavne vo výpočtovej lingvistiky. Napriek značnému pokroku symbolové analyzátory nedokážu obísť niektoré problémy vyplývajúce z ich podstaty, najmä pre ich rigidnosť a chýbajúcu robustnosť. Návrh konekcionalistických analyzátorov mal za cieľ odstrániť práve tieto nedostatky. Prvé konekcionalistické analyzátory využívali lokalistické reprezentácie a boli v podstate implementáciou symbolových algoritmov, i keď pravidlá už boli menej rigidné vďaka použitiu neurónových sietí. Tieto modely boli robustnejšie, hoci ich schopnosť korekcie chýb v negramatických vetách bola nedostačujúca (napr. POLLACK, WALTZ 1985; MIIKKULAINEN 1996; HO, WONG 1998).

Požiadavkou pre úspešné konekcionalistické analyzátory je ich schopnosť realizovať holistickú analýzu, t. j. holisticky zobrazit' reprezentáciu vety na reprezentáciu jej zodpovedajúceho syntaktického stromu. Na zakódovanie vety i syntaktického stromu sa najčastejšie používajú SRN alebo SRAAM (HO, CHAN 1999). Keď už máme vety i ich stromy zakódované, môžeme naučiť doprednú sieť, aby holisticky asociovala jednu reprezentáciu s druhou. Druhým spôsobom realizácie holistickej transformácie je vopred zakódovať vetu alebo jej strom, a takto získanú reprezentáciu potom použiť ako požadovaný výstup siete (SRN), trénovanej na zakódovanie druhej reprezentácie (čím obe reprezentácie budú identické, i keď boli oddelene kódované). Po tretie, obe reprezentácie sa môžu vyvíjať súčasne a spoločne (konfluentná inferencia) ako v prípade Chrismanovej translačnej úlohy.

Konekcionalistické analyzátory sa snažia vysvetliť schopnosť človeka manipulovať so štruktúrovanými dátami subsymbolovým spôsobom. Medzi ich nedostatky patrí to, že sú založené číro syntakticky. Nie sú preto napríklad citlivé na sémantické obmedzenia, o ktoré sa človek pri analýze vety určite opiera. Okrem toho využívajú znalosti (napr. kódované syntaktické stromy), ktoré človek pri učení má sotva k dispozícii.

3.5. Komprehenzia viet

Komprehenzné modely sú tie, ktoré sa snažia riešiť hlavný cieľ spracovania vety – vytvorenie (hlbkovej) sémantickej reprezentácie vety z jej povrchovej formy. Dve vety môžu mať veľmi odlišné povrchové formy (a aj syntaktické stromy), avšak môžu znamenať v podstate to isté.¹⁹ Reprezentácia a spracovanie sémantickej informácie je ťažká úloha (pozri aj kapitolu *Významy neexistujú*), zrejme i preto nebolo doteraz navrhnutých veľa komprehenzných modelov. Väčšina modelov sa opiera o distribuované reprezentácie, ktoré dokážu oveľa lepšie podchytiť jemné sémantické rozdiely medzi jednotlivými konceptmi (na to už dávnejšie poukázal HINTON 1981).

Jedným z prvých modelov je práca McClellanda a Kawamota (1986), ktorých jednoduchý perceptrón bol trénovaný tak, aby priradil sémantické roly (agens, paciens, inštrument, modifikátor) jednotlivým konštituentom vety (podmet, prísudok, predmet, PF „s“). Sémantické reprezentácie viet teda neboli holistické (ale konkatenatívne kompozičné). Priradenie rôl sa však považuje za dôležitý krok pri komprehenzii, čo bolo motiváciou tejto

¹⁹ Napr. vety *Karol zaplatil kartou* a *Pomocou karty sa Karolovi podarilo realizovať platbu* sú syntakticky veľmi odlišné, sémanticky však nie. Je to preto, lebo neexistuje jedno-jednoznačné priradenie medzi syntaktickými (podmet, predmet, prívlastok...) a sémantickými rolami (agens, paciens, inštrument...).

práce. Reprezentácie slov boli navrhnuté ručne. Model bol trévaný na vetách generovaných len podľa niekoľkých šablón (templates) s požadovanými sémantickými rolami (napr. „human *ate* food utensil“ umožňuje vytvoriť vetu *Man ate pasta (with) spoon* → agens, paciens, inštrument). Napriek tomu úloha nie je triviálna, lebo priradenie rôl závisí od kontextu (vo vetách *Ball hit girl* a *Ball moved* je ten istý podmet raz inštrumentom a raz paciensom), taktiež závisí od sémantiky slovesa (vo vetách *Man ate pasta (with) spoon* a *Man ate pasta (with) cheese* v prvom prípade PF *with* modifikuje paciensa, zatiaľ čo v druhom prípade je inštrumentom). Napokon v niektorých prípadoch nejednoznačnosť ostáva nerozlúštená (napr. vo vete *Boy hit girl with ball* nevieme, či *ball* je inštrumentom alebo modifikátorom paciensa *girl*). Model dokázal tieto nejednoznačnosti zohľadniť a správne zovšeobecňovať. Ako obmedzenia modelu autori spomenuli pevne definované šablóny vymedzujúce typy viet, ktoré vie model spracovať, alebo formuláciu problému ako statického zobrazenia, so statickými reprezentáciami slov.

St. John a McClelland (1992) použili SRN a rozšírili predchádzajúci model v tom zmysle, že ich sieť čítala vstup po slovách, čím proces vytvárania sémantickej reprezentácie vety získal dynamický charakter. Výstupná časť SRN bola trévaná tak, aby pri testovaní mohla odpovedať na podnety (po predložení konkrétneho slova musela predikovať jeho sémantickú rolu, alebo naopak). Tento model úspešne preukázal kognitívne zaujímavé správanie, napr. zohľadnenie syntaktických i sémantických črt pri budovaní sémantickej reprezentácie vety, aktualizované interpretácie vety počas jej sekvenčného spracovania, inferencie chýbajúcich konštituentov a zovšeobecnenie na nové vety. Pretrvávajúcim nedostatkom bola nemožnosť spracovávať súvetia.

Miikkulainen a Dyer (1989) navrhli zaujímavý učiaci mechanizmus vytvárania sémantických reprezentácií slov v úlohe priradovania sémantických rôl jednotlivým konštituentom vety (použili tie isté vety ako McClelland a Kawamoto). Vstupy i výstupná vrstva neurónov doprednej dvojvrstvej siete boli rozdelené na rovnaké bloky, z ktorých každý slúžil pre jedno slovo (syntaktický konštituent na vstupe, sémantická rola na výstupe). Náhodne inicializované reprezentácie slov uložené v slovníku sa adaptovali (súčasne s váhami) pomocou modifikovaného algoritmu spätného šírenia chýb dovtedy, kým vstupno-výstupné zobrazenie (t. j. priradenie rôl) nedosiahlo najväčšiu konzistentnosť. Správnosť sémantických reprezentácií autori overili pomocou dendrogramu, kde bolo vidieť, že sémanticky blízke koncepty nadobudli podobné reprezentácie. Táto metóda bola rozšírená aj pre SRN (MIKKULAINEN, DYER 1991), ktorá spracovávala tie isté vety, ale sekvenčne, čo eliminovalo nutnosť vopred explicitne špecifikovať vetnú šablónu. Miikkulainen a Dyer (1991) navrhli integrovaný kognitívny systém – parafrázovač krátkych príbehov (skriptov; pozri časť 3.7). Prv než sa dostaneme k integrovaným modelom detailnejšie, ešte stručne spomeňme čisto produkčné modely.

3.6. Produkcia viet

Produkcia vety je transformácia významu vety na sekvenciu slov alebo zvukov. Produkčné modely v porovnaní s komprehenznými modelmi alebo modelmi syntactickej analýzy upútali oveľa menej pozornosti možno práve preto, lebo ich vstupom sú sémantické reprezentácie, ktorých vytvorenie je samo osebe ťažkou úlohou. Prvé modely produkcie boli lokalistické modely, bez schopnosti učenia. Kalita a Shastri (1994) sa zamerali na problém produkcie a usporiadania slov do vety na základe znalosti o ich sémantických rolách a indikácii o požadovanom čase (napr. prítomný) a rode (činný/trpný). Ich model bol schopný produkovať jednoduché vety typu SVO (subject-verb-object) v požadovanom rode a v niekoľkých časoch. Gasser (1988) opísal model, ktorý produkuje vety podľa jednotlivých schém a ktorý využíva IA princípy. Aktivácia slov v správnom poradí je v modeli zabezpečená pomocou špeciálnych neurónov (sequencing units), ktoré inhibujú ostatné neuróny po vykonaní svojej činnosti. Viazanie slov so syntaktickými rolami je

založené na synchronizovanom pálení neurónov. Správne poradie slov na výstupe sa dosahuje pomocou mechanizmu štartovacích a ukončujúcich neurónov pre každú produkovanú frázu, čo je v istom ohľade podobný mechanizmus, aký neskôr použili Kalita a Shastri (1994). Gasserov model bol navrhnutý s cieľom vysvetliť rôzne javy ako chyby reči, robustné správanie pri neúplnom vstupe, flexibilita pri usporadúvaní slov a prenos znalosti do iného jazyka. Model bol však schopný produkovať len jednoduché vety. Tretí lokalistický model (WARD 1991) už dokázal produkovať širšie spektrum viet, avšak ani tento neobsahoval učiaci mechanizmus. Dell, Chang a Griffin (1999) opisujú aktuálne konekcionistické modely produkcie a načrtávajú problémy, ktoré treba riešiť.

3.7. Integrované modely

Miikkulainen (1993) navrhol subsymbolový model DISCERN (DIstributed SCRipt processing and Episodic memoRY Network), ktorý spracováva jednoduché skripty a rozumie im.²⁰ Skripty sú organizované v hierarchicky organizovanej epizodickej pamäti (implementovanej pomocou SOM máp). Okrem epizodickej pamäti a SOM lexikónu (pozri časť 2.5) systém obsahuje procesné časti (SRN), konkrétne modul syntaktickej analýzy, produkčný modul a modul generujúci odpovede na otázky. Systém DISCERN bol testovaný takmer na 100 skriptoch, ktoré dokázal správne parafrázovať. Vie odpovedať aj na otázky (cues), týkajúce sa skriptov. Počas učenia systém extrahuje inferencie automaticky podľa štatistických korelácií vo vstupných dátach. Týmito vlastnosťami sa DISCERN kvalitatívne líši od symbolových systémov, kde inferencie bývajú založené na explicitne vložených pravidlách a kde pamäťové štruktúry bývajú explicitne naprogramované (KOLODNER 1984). Vďaka redundancii v distribuovaných reprezentáciách je systém DISCERN robustný. Je aj prijateľným modelom na neurálnej úrovni, lebo dokáže vysvetliť rôzne chyby pri sémantickej analýze (priradenie sémantických rôl), v epizodickej pamäti, v lexikálnom prístupe, ako aj rôzne lézie v dôsledku lokálneho poškodenia.

Druhým významným integrovaným konekcionistickým modelom spracovania jazyka je CSCP model (Connectionist model of Sentence Comprehension and Production), ktorého cieľom je rozpoznávanie a produkcia viet (ROHDE 2002). Vety (vrátane súvetí) sú generované podľa gramatiky zjednodušeného anglického jazyka (proto-english alebo penglish), ktorý má však pomerne veľkú komplexnosť.²¹ Význam vety pozostáva z množiny propozícií, ktoré sú reprezentované distribuovane. Jedna časť systému – sémantický systém – sa učí skomprimovať sekvenciu týchto propozícií do jednej, statickej reprezentácie významu celej vety. CSCP využíva moduly SRN, ktoré sa naučia zobrazit' vstupnú vetu (sekvencia slov) na jej celkový význam (komprehenzia) a opačne (produkcia). CSCP je schopný simulovať rôzne aspekty spracovania viet u človeka, vrátane citlivosti na frekvenciu výskytu slov i slovných fráz, sémantickú plauzibilitu, inflektívnu morfológiu a lokálne efekty. Model poukazuje na dôležitosť prepojenia medzi komprehenziou a produkciou počas učenia jazyka. Oba spomínané modely predstavujú významný krok smerom k realistickým modelom jazyka.

²⁰ Skripty (SCHANK, ABELSON 1977) sú mentálne schémy častých, stereotypických udalostí (napr. nakupovanie alebo návšteva divadla), ktoré človek používa pri spracovaní informácie (napr. na doplnenie chýbajúcej informácie). Príkladom skriptu použitého v DISCERNe je tento: „*John went to the airport. John checked in for flight to JFK. The plane landed at JFK.*” Schéma nám umožní napríklad doplniť, že *John boarded the plane.*

²¹ Tento jazyk zahŕňa slovesný čas, množné číslo, prídavné mená a príslovky, predložkové frázy, vzťažné vety, vetné komplementy a má slovnú zásobu okolo 300 slov.

4. Záver

Cieľom tejto kapitoly bolo ukázať, akým vývojom za posledných zhruba 25 rokov prešlo konekcionistické modelovanie prirodzeného jazyka, ktoré vyústilo do vzniku špecializovaného výskumného smeru označovaného ako **konekcionistická psycholingvistika**. Otázka, či konekcionistické modely jazyka plne nahradia symbolové systémy, je asi predčasná, pretože i v konekcionistickom tábore sa názory rôznia (SEIDENBERG, MCDONALD 1999; SMOLENSKY 1999; STEEDMAN 1999). Napriek tomu použitie týchto modelov má veľký význam (CHRISTIANSEN, CHATER 1999a). Po prvé, konekcionistické modely podnietili teoretický výskum vo viacerých oblastiach a ponúkli alternatívne vysvetlenia (učenie v kváziregulárnej doméne). Po druhé, učiace mechanizmy v konekcionistických modeloch oživilo záujem o výpočtové modely akvizície jazyka (BATES, ELMAN 1993). Zatiaľ čo Chomsky (1986) argumentuje v prospech vrodenných univerzálnych aspektov jazyka, zástancovia konekcionizmu tvrdia, že veľkú časť jazykovo-špecifických znalostí musí dieťa nadobudnúť učením. Konekcionistické modely ponúkajú vysvetlenie, ako sa to deje. Po tretie, závislosť konekcionistických modelov od štatistických vlastností vstupných dát podnietil výskum štatistických faktorov pri učení a spracovaní jazyka (REDINGTON, CHATER, FINCH 1998).

Aj keď správanie konekcionistických modelov lepšie vysvetľuje jazykové prejavy človeka, súčasné modely majú aj nedostatky. Väčšina konekcionistických prác modelovania jazyka sa orientovala na malé podproblémy, ktoré sa dajú modelovať jednou neurónovou sieťou. V snahe o budovanie integrovaných konekcionistických systémov vystupuje problém návrhu efektívnej komunikácie (riadenia) medzi jednotlivými modulmi (Miikkulainen, 1994). Alternatívou sú hybridné systémy, kde sa takáto komunikácia dá predprogramovať na symbolovej báze (WERMTER, SUN 2000). Ďalšou otázkou ostáva použitie vhodných reprezentácií a učiaceho algoritmu, čo podmieňuje úspešnosť modelovania. Vhodnosť reprezentácií a učenia treba zrejme posudzovať separátne. Napríklad reprezentácie môžu podporovať realizovateľnosť nejakej úlohy (ak sa podarí váhy ručne nastaviť), ale učiaci algoritmus zlyháva (SHARKEY, SHARKEY, JACKSON 2000). Problémom je takisto škálovanie na väčšie úlohy.

Tieto problémy by mali slúžiť ako výzva na hľadanie nových paradigiem reprezentácie jazykových entít a učenia. Zatiaľ to vyzerá tak, že konekcionistická psycholingvistika si úspešne razí cestu zdola pri vysvetľovaní stále komplexnejších aspektov jazykového správania a iba ďalší výskum ukáže, či tento prístup narazí na neprekonateľné bariéry pri snahe subsymbolovo vysvetliť zložitejšie mentálne procesy prebiehajúce v mozgu pri používaní jazyka.

Literatúra

- [1] ALLEN, J., SEIDENBERG, M. S. (1999). The emergence of grammaticality in connectionist networks. In B. MacWhinney (ed.), *Mechanisms of Language Acquisition*. Hillsdale, NJ: Erlbaum, 115-151.
- [2] ANGLUIN, D. (1988). Identifying languages from stochastic examples. *Technical report YALEU/DCS/RR-614*, Yale University, New Haven, CT.
- [3] BAKER, C. L., MCCARTHY, J. J. (eds.) (1981). *The Logical Problem of Language Acquisition*. Cambridge, MA: MIT Press.
- [4] BATES, L., ELMAN, J. (1993). Connectionism and the study of change. In: M. Johnson (ed.), *Brain development and cognition*. Blackwell: Oxford, 623-642.
- [5] CHALMERS, D. (1990). Syntactic transformations on distributed representations. *Connection Science*, 2, 53-62.
- [6] CHARNIAK, E. (1993). *Statistical language learning*. Cambridge, MA: MIT Press.

- [7] CHOMSKY, N. (1956). Three models for the description of language. *IRE Transactions on Information Theory*, 2(3), 113-124.
- [8] CHOMSKY, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- [9] CHOMSKY, N. (1975). *The Logical Structure of Linguistic Theory*. New York, NY: Plenum Press.
- [10] CHOMSKY, N. (1986). *Knowledge of Language*. New York: Prager.
- [11] CHRISMAN, L. (1991). Learning recursive distributed representations for holistic computation. *Connection Science*, 3(4), 345-366.
- [12] CHRISTIANSEN, M.H. (1994). Infinite Languages, Finite Minds: Connectionism, Learning and Linguistic Structure. *PhD thesis*, University of Edinburgh, Scotland.
- [13] CHRISTIANSEN, M. H., CHATER, N. (1999a). Connectionist natural language processing: The state of the art. *Cognitive Science: a special issue on connectionist models of human language processing: progress and prospects*, 23(4), 417-437.
- [14] CHRISTIANSEN, M. H., CHATER, N. (1999b). Toward a connectionist model of recursion in human linguistic performance. *Cognitive Science*, 23(2), 157-205.
- [15] CLEEREMANS, A., SERVAN-SCHREIBER, D., McCLELLAND, J. L. (1989). Finite state automata and simple recurrent networks. *Neural Computation*, 1(3):372-381.
- [16] COLTHEART, M. (1987). Cognitive neuropsychology and the study of reading. In M. Posner and O.S. Marin (eds.), *Attention and Performance XI*. Hillsdale, NJ: Erlbaum, 3-37.
- [17] COLTHEART, M., CURTIS, B., ATKINS, P., HALLER, M. (1993). Models of reading aloud: Dual route and parallel-distributed processing approaches. *Psychological Review*, 100, 589-608.
- [18] DELL, G. S. (1986). A spreading activation theory of retrieval in sentence production. *Psychological Review*, 93(3), 283-381.
- [19] DELL, G. S. (1988). The retrieval of phonological forms in production: Tests of predictions from a connectionist model. *Journal of Memory and Language*, 104, 123-147.
- [20] DELL, G. S., CHANG, F., GRIFFIN, Z. M. (1999). Connectionist models of language production: Lexical access and grammatical encoding. *Cognitive Science*, 23(4), 517-542.
- [21] DELL, G. S., JULIANO, C., GOVINDJEE, A. (1993). Structure and content in language production: A theory of frame constraints in phonological speech errors. *Cognitive Science*, 17, 149-195.
- [22] DELL, G. S., SCHWARTZ, M. F., MARTIN, N., SAFFRAN, E. M., GAGNON, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review*, 104(4), 801-838.
- [23] DOYA, K. (1999). What are the computations of the cerebellum, the basal ganglia, and the cerebral cortex? *Neural Networks*, 12, 961-974.
- [24] ELMAN, J. L. (1990). Finding structure in time. *Cognitive Science*, 14, 179-211.
- [25] ELMAN, J. L. (1991). Distributed representations, simple recurrent networks, and grammatical structure. *Machine Learning*, 7, 195-224
- [26] ELMAN, J. L. (1993). Learning and development in neural networks: The importance of starting small. *Cognition*, 48, 71-99.
- [27] ELMAN, J. L., BATES, E. A., JOHNSON, M. H., KARMILOFF-SMITH, A., PARISI, D., PLUNKETT, K. (1997). *Rethinking Innateness: A Connectionist Perspective on Development*. Bradford Books.
- [28] FARKAŠ, I. (1997). Samoorganizujúce sa mapy. Kapitola v knihe V. Kvasnička a kol., *Úvod do teórie neurónových sietí*, IRIS, Bratislava, 142-189.
- [29] FARKAŠ, I. (2003). Lexical acquisition and developing semantic map. *Neural Network World*, 13(3), 235-245.
- [30] FODOR, J. A., PYLYSHYN, Z. W. (1988). Connectionism and cognitive architecture: A critical architecture. *Cognition*, 28, 3-71.
- [31] FORCADA, M.L. (2002). *Neural networks: Automata and formal models of computation*. Elektronická kniha <http://www.dlsi.ua.es/~mlf/nnafmc/pbook/index.html>.

- [32] GASKELL, M. G., HARE, M., MARSLEN-WILSON, W. D. (1995). A connectionist model of phonological representation in speech perception. *Cognitive Science*, 19, 407-439.
- [33] GELDER VAN, T. (1990). A compositionality: a connectionist variation on a classical theme. *Cognitive Science*, 14, 355-384.
- [34] GILES, C. L., SUN, G., CHEN, H., LEE, Y., CHEN, D. (1990). Higher order recurrent networks and grammatical inference. In Touretzky, D. (ed.), *Advances in Neural Information Processing Systems 2*, 380-387, San Mateo, CA: Morgan Kaufmann.
- [35] GILES, C. L., MILLER, C. B., CHEN, D., CHEN, H. H., SUN, G. Z., LEE, Y. C. (1992). Learning and extracted finite state automata with second-order recurrent neural networks. *Neural Computation*, 4(3), 393-405.
- [36] HAMMER, B. (2003). Perspectives on learning symbolic data with connectionistic systems. In R.Kühn et al. (eds.), *Adaptivity and Learning*, 141-160, Springer.
- [37] HAMMERTON, J. A. (1998). Holistic computation: Reconstructing a muddled concept. *Connection Science*, 10(1), 3-19.
- [38] HARE, M., ELMAN, J., DAUGHERTY, K. G. (1995). Default generalization in connectionist networks. *Language and Cognitive Processes*, 10(6), 601-630.
- [39] HAYKIN, S. (1994). *Neural Networks: A Comprehensive Foundation*. Prentice Hall.
- [40] HINTON, G.E. (1981). Implementing semantic networks in parallel hardware. In G. E. Hinton, J. A. Anderson (eds.), *Parallel Models of Associative Memory*, Hillsdale, NJ: Erlbaum, 161-187.
- [41] HO, E. K. S., WONG, K. F. (1998). A neural network-based LR parser. In *Proceedings of the 1st Int. Symposium on Intelligent Data Engineering and Learning, Hong Kong*, 289-294.
- [42] HUMMEL, J. E., HOLYOAK, K. L. (1997). Distributed representation of structure: a theory of analogical access and mapping. *Psychological Review*, 104, 427-466.
- [43] KALITA, J., SHASTRI, L. (1994). A connectionist approach to generation of simple sentences and word choice. In G. Adriaens, U. Hahn (eds.), *Parallel Natural Language Processing*. Norwood, NJ: Ablex, 395-400.
- [44] KIRBY, S., CHRISTIANSEN, M. H. (2003). From language learning to language evolution. In S. Kirby, M. H. Christiansen, (eds.), *Language Evolution*, Oxford University Press, 272-294.
- [45] KOHONEN, T. (2001). *Self-Organizing Maps* (3. ed.), Springer: Berlin.
- [46] KOLEN, J. F. (1994). The origin of clusters in recurrent neural network state space. In *Proc. 16th Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum Associates, 508-513.
- [47] KOLODNER, J. L. (1984). *Retrieval and Organizational Strategies in Conceptual Memory: A Computer Model*. Hillsdale, NJ: Erlbaum.
- [48] KVASNIČKA V., BEŇUŠKOVÁ, Ľ., POSPÍCHAL J., FARKAŠ I., TIŇO P., KRÁL A. (1997). *Úvod do teórie neurónových sietí*. IRIS, Bratislava.
- [49] KWASNY, S. C., KALMAN, B. L. (1995). Tail-recursive distributed representations and simple recurrent networks. *Connection Science*, 7(1), 61-80.
- [50] LACHOVÁ, J. (2004). Model procesu transformácie grafém na fonémy pri čítaní slov v slovenčine. *Diplomová práca*, Ústav informatiky FMFI UK, Bratislava.
- [51] LANE, P., HENDERSON, J. (2001). Incremental syntactic parsing of natural language corpora with simple synchrony networks. *IEEE Transactions on Knowledge and Data Engineering*, 13(2), 219-231.
- [52] LI, P., FARKAŠ, I., MACWHINNEY B. (2004). Early lexical development in a self-organizing neural network. *Neural Networks*, v tlači.
- [53] MARCUS, G. (1993). Negative evidence in language acquisition. *Cognition*, 46(1), 53-85.
- [54] MARCUS, G. F., BRINKMANN, U., CLAHSSEN, H., WIESE, R., PINKER, S. (1995). German inflection: The exception that proves the rule. *Cognitive Psychology*, 29, 189-256.
- [55] MASSARO, D. (1989). Testing between TRACE and the fuzzy logical model of speech perception. *Cognitive Psychology*, 21, 398-421.

- [56] MAYBERRY, M. R., MIIKKULAINEN, R. (1998). SARDSRN: A neural network shift-reduce parser. *Technical report AI98-275*, University of Texas at Austin, Texas.
- [57] McCLELLAND, J. L. (1986). The programmable blackboard model of reading. In J. L. McClelland, D. E. Rumelhart (eds.) *Parallel distributed processing, Vol. 2*. Cambridge, MA: MIT Press.
- [58] McCLELLAND, J. L., ELMAN, J. L. (1986). Interactive processes in speech perception: The TRACE model. In *Parallel distributed processing, Vol. 2*. Cambridge, MA: MIT Press, 58-121.
- [59] McCLELLAND, J. L., KAWAMOTO, A. H. (1986). Mechanisms of sentence processing: Assigning roles to constituents of sentences. In *Parallel Distributed Processing, Vol. 2*. MIT Press, Cambridge, MA, 318-362.
- [60] McCLELLAND, J. L., RUMELHART, D. E. (1981). An interactive activation model of the effect of context in perception: Part 1. An account of basic findings. *Psychological Review*, 88, 375-405.
- [61] McCLELLAND, J.L., RUMELHART, D.E., THE PDP RESEARCH GROUP (eds.) (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Vol. 2: Psychological and Biological Models*. Cambridge, MA: MIT Press.
- [62] McLEOD, P., PLUNKETT, K., ROLLS, E. T. (1998). *Introduction to Connectionist Modelling of cognitive processes*. Oxford, UK: Oxford University Press.
- [63] MIIKKULAINEN, R. (1990). Script recognition with hierarchical feature maps. *Connection Science*, 2(1 & 2), 83-101.
- [64] MIIKKULAINEN, R. (1994). Integrated connectionist models: Building AI systems on subsymbolic foundations. In *Artificial Intelligence and Neural Networks: Steps Toward Principled Integration*. Honavar, V. Uhr, L. (eds.), San Diego, CA: Academic Press, 483-508.
- [65] MIIKKULAINEN, R. (1996). Subsymbolic case-role analysis of sentences with embedded clauses. *Cognitive Science*, 20 (1), 47-73.
- [66] MIIKKULAINEN, R. (1997). Dyslexic and category-specific aphasic impairments in a self-organizing feature map model of the lexicon. *Brain and Language*, 59, 334-366.
- [67] MIIKKULAINEN, R., DYER, M. G. (1989). Encoding input/output representations in connectionist cognitive systems. In: Touretzky, D., Hinton, G., Sejnowski, T. (eds.). *Proceedings of the 1988 Connectionist Models Summer School*. Morgan Kaufmann, 347-356.
- [68] MIIKKULAINEN, R., DYER, M. G. (1991). Natural Language Processing with Modular PDP Networks and Distributed Lexicon. *Cognitive Science*, 15(3), 343-399.
- [69] MORGAN, J. L., BONAMO, K. M., TRAVIS, L. L. (1995). Negative evidence on negative evidence. *Developmental Psychology*, 31, 180-197.
- [70] MORGAN, J. L., TRAVIS, L. L. (1989). Limits on negative information in language input. *Journal of Child Language*, 16, 531-552.
- [71] MOZER, M. C. (1991). *The Perception of Multiple Objects: A Connectionist Approach*. Cambridge, MA: MIT Press.
- [72] MOZER, M. C., DAS, S. (1993). A connectionist chunker that induces the structure of context-free languages. In *Advances in Neural Information Processing Systems 5*, San Mateo, CA: Morgan Kaufmann.
- [73] NEWELL, A., SIMON, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice Hall.
- [74] NEWPORT, E. (1990). Maturation constraints on language learning. *Cognitive Science*, 14, 11-28.
- [75] NORRIS, D. (1993). Bottom-up connectionist models of interaction. In R. Shillock, D. Altmann (eds.), *Cognitive models of speech processing*. Hove: Erlbaum, 211-234.
- [76] NORRIS, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, 52, 189-234.
- [77] PAGE, M. (2000). Connectionist modelling in psychology: A localist manifesto. *Behavioral and Brain Sciences* 23(4): 443-512.

- [78] PARFITT, S., TIÑO, P., DORFFNER, G. (2000): Graded grammaticality in prediction fractal machines. In *Advances in Neural Information Processing Systems 12*, MIT Press, 52-58.
- [79] PALMER-BROWN, D., TEPPER, J. A., POWELL, H. M. (2002). Connectionist natural language parsing. *Trends in Cognitive Sciences*, 6(10), 437-442.
- [80] PERFORIS, A. (2002). Simulated evolution of language: a review of the field. *Journal of Artificial Societies and Social Simulation*, 5(2), <http://jasss.soc.surrey.ac.uk/5/2/4.html>.
- [81] PINKER, S. (1984). *Language Learnability and Language Development*. Cambridge, MA: Harvard University Press.
- [82] PINKER, S. (1999). *Words and rules: The ingredients of language*. New York: Basic Books.
- [83] PINKER, S., MEHLER, J. (eds.) (1988). *Connections and symbols*. Cambridge, MA: MIT Press.
- [84] PINKER, S., PRINCE, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition*, 28, 73-193.
- [85] PLATE, T. (1995). Holographic reduced representations. *IEEE Transactions on Neural Networks*, 6(3), 623-641.
- [86] PLAUT, D. C. (1999). A connectionist approach to word reading and acquired dyslexia: Extension to sequential processing. *Cognitive Science*, 23, 543-568.
- [87] PLAUT, D. C., MCCLELLAND, J. L., SEIDENBERG, M. S., PATTERSON, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, 103, 56-115.
- [88] PLAUT, D. C., KELLO, C. T. (1999). The emergence of phonology from the interplay of speech comprehension and production: A distributed connectionist approach. In B. MacWhinney (ed.), *The Emergence of Language*, Mahwah, NJ: Erlbaum, 381-415.
- [89] PLUNKETT, K., MARCHMAN, V. A. (1993). From rote learning to system building: acquiring verb morphology in children and connectionist nets. *Cognition*, 48(1), 21-69.
- [90] PLUNKETT, K., JUOLA, P. (1999). A connectionist model of English past tense and plural morphology. *Cognitive Science*, 23, 463-490.
- [91] POLLACK, J. (1990). Recursive distributed representations. *Artificial Intelligence*, 46, 77-105.
- [92] POLLACK, J. B., WALTZ, D. (1985). Massively parallel parsing: A strongly interactive model of natural language interpretation. *Cognitive Science*, 9, 51-74.
- [93] PRASADA, S., PINKER, S. (1993). Similarity-based and rule-based generalizations in inflectional morphology. *Language and Cognitive Processes*, 8, 1-56.
- [94] QUINLAN, P. (1991). *Connectionism and Psychology: A Psychological Perspective on New Connectionist Research*. Chicago, IL: University of Chicago Press.
- [95] REDINGTON, M., CHATER, N., FINCH, S. (1998). Distributional information: A powerful cue for acquiring syntactic categories. *Cognitive Science*, 22, 425-469.
- [96] RITTER, H., KOHONEN, T. (1989). Self-organizing semantic maps. *Biological Cybernetics*, 61, 241-254.
- [97] RODRIGUEZ, P. (2001). Simple recurrent networks learn context-free and context-sensitive languages by counting. *Neural Computation*, 13(9), 2093-2118.
- [98] RODRIGUEZ, P., WILES J., ELMAN, J. L. (1999). A recurrent neural network that learns to count. *Connection Science*, 11(1), 5-40.
- [99] ROHDE, D. L. T. (2002). A connectionist model of sentence comprehension and production. *Doctoral thesis*, Carnegie Mellon University, Pittsburg, PA.
- [100] ROHDE, D. L. T., PLAUT, D. C. (1999). Language acquisition in the absence of explicit negative evidence: How important is starting small? *Cognition*, 72, 67-109.
- [101] ROHDE, D. L. T., PLAUT, D. C. (2003). Connectionist models of language processing. *Cognitive Studies*, 10, 10-28.
- [102] RUMELHART, D. E., MCCLELLAND, J. L. (1982). An interactive activation model of context effects in letter perception: Part 2. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review*, 89, 60-94.

- [103] RUMELHART, D. E., MCCLELLAND, J. L. (1986). On learning the past tenses of English verbs. In *Parallel Distributed Processing, Vol. 2*, Cambridge, MA: MIT Press, 216-271.
- [104] SAFFRAN, J. R., NEWPORT, E. L., ASLIN, R. N. (1996). Word segmentation: the role of distributional cues. *Journal of Memory and Language*, 35, 606-621.
- [105] SCHANK, R., ABELSON, R. (1977). *Scripts, plans, goals and understanding*. Hillside, NJ: Erlbaum.
- [106] SEIDENBERG, M. (1997). Language acquisition and use: Learning and applying probabilistic constraints. *Science*, 275, 1599-1603.
- [107] SEIDENBERG, M.S., MCCLELLAND, J.L. (1989). A distributed developmental model of word recognition and naming. *Psychological Review*, 96, 523-568.
- [108] SEIDENBERG, M.S., MACDONALD, M.C. (1999). A probabilistic constraints approach to language acquisition and processing. *Cognitive Science*, 23, 569-588.
- [109] SHARKEY, N., SHARKEY, A., JACKSON, S. (2000). Are SRNs sufficient for modelling language acquisition? In P. Broeder, J. Murre (eds.), *Models of Language Acquisition: Inductive and Deductive Approaches*. Oxford University Press, 33-54.
- [110] SHASTRI, L., AJJANAGADDE, V. (1993). From simple associations to systematic reasoning. *Behavioral and Brain Sciences*, 16(3), 417-494.
- [111] SMOLENSKY, P. (1990). Tensor product variable binding and the representation of symbolic structures in connectionist systems. *Artificial Intelligence*, 46(1-2), 159-216.
- [112] SMOLENSKY, P. (1999). Grammar-based connectionist approaches to language. *Cognitive Science*, 23(4), 589-613.
- [113] ST. JOHN, M., MCCLELLAND, J. (1992). Parallel constraint satisfaction as a comprehension mechanism. In Reilly, R., Sharkey, N. (eds.), *Connectionist Approaches to Natural Language Processing*, 97-136. Hove, UK: Erlbaum.
- [114] STEEDMAN, M. (1999). Connectionist sentence processing in perspective. *Cognitive Science*, 23(4), 615-634.
- [115] TIŇO, P., ČERNÁNSKÝ M., BEŇUŠKOVÁ, I (2004): Markovian architectural bias of recurrent neural networks. *IEEE Transactions on Neural Networks*, 15(1), 6-15.
- [116] VOEGTLIN, T. (2002). Neural networks and self-reference. *Doctoral thesis*, Université Lyon 2.
- [117] WARD, N. (1991). A flexible, parallel model of natural language generation. *Doctoral thesis*, University of California at Berkeley.
- [118] WERMTER, S., SUN, R. (eds.) (2000). *Hybrid Neural Systems*. Springer: Heidelberg.
- [119] ZENG, Z., GOODMAN, R. M., SMYTH, P. (1994). Discrete recurrent neural networks for grammatical inference. *IEEE Transactions on Neural Networks*, 5(2), 320-330.