

Holografická redukovaná reprezentácia v umelej inteligencii a kognitívnej vede¹

Vladimír Kvasnička a Jiří Pospíchal²

Fakulta informatiky a informačných technológií

a

Fakulta chemickej a potravinárskej technológie
Slovenská technická univerzita v Bratislave

Abstrakt. Holografická redukovaná reprezentácia využíva konceptuálne n -rozmerné vektory, ktorých elementy vyhovujú normálnej distribúcii $N(0,1/n)$. Existujúce aplikácie tohto prístupu sa zaoberajú štúdiom rôznych modelov asociatívnej pamäti, ktoré využívajú jednoduchú algebraickú operáciu skalárneho produktu konceptuálnych vektorov, pomocou ktorej sa kvantifikuje podobnosť dátových štruktúr reprezentovanými týmito vektormi. V práci je popísaná metóda k modelovaniu podobnosti medzi rôznymi koncepciami a procesy usudzovania, ktoré využívajú deduktívne pravidlá *modus ponens* a *modus tollens*.

1. Úvodné poznámky

Moderný pohľad na vzťah medzi mozgom a mysl'ou vychádza z neurovednej paradigmy [7], podľa ktorej architektúra mozgu je určená spojmi medzi neurónmi, ich inhibičným, alebo excitačným charakterom a taktiež aj ich intenzitou. Ľudský mozog vykazuje neobyčajnú plasticitu, v priebehu učenia neustále vznikajú (ale taktiež aj zanikajú) synaptické spoje. Možno konštatovať, že *schopnosť mozgu vykonávať nielen kognitívne aktivity, ale byť aj pamäťou a aj riadiacim centrom pre našu motoriku, je plne zakódovaná jeho architektúrou*. Môžeme konštatovať, že predstava o ľudskom mozgu ako o počítači sa musí formulovať tak, že mozog je *paralelný distribuovaný počítač* (obsahujúci mnoho miliárd neurónov, elementárnych procesorov, ktoré sú medzi sebou poprepájané do zložitej neurónovej siete). Program v tomto paralelnom počítači je priamo zabudovaný do architektúry neurónovej siete, t. j. ľudský mozog je jednoúčelový paralelný počítač reprezentovaný neurónovou sieťou, ktorý nie je možné preprogramovať bez zmeny jeho architektúry.

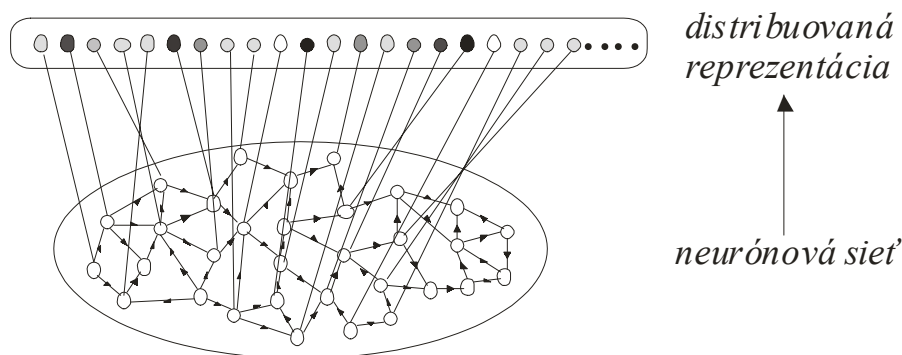
Z vyššie uvedených všeobecných úvah vyplýva, že myseľ s mozgom tvoria jeden integrálny celok, ktorý je charakterizovaný *komplementárnym dualizmom*. Mysleň je v tomto prístupe chápaná ako program vykonávaný mozgom, pričom tento program je špecifikovaný architektúrou distribuovanej neurónovej siete reprezentujúcej mozog. Mozog a myseľ tvoria dva rôzne pohľady na ten istý objekt:

- (1) Keď hovoríme o mozgu, myslíme tým „hardwarovú“ štruktúru, biologicky realizovanú neurónmi a ich synaptickými spojmi (formálne reprezentovanú neurónovou sieťou), v opačnom prípade,
- (2) keď hovoríme o mysli, myslíme tým kognitívne a iné podobné aktivity mozgu, ktoré sú vykonávané na symbolickej úrovni, kde prebieha transformácia symbolickej informácie na základe (jednoduchých) pravidiel.

¹ Táto prednáška je skrátenou verziou práce [6].

² Email: {kvasnicka, pospichal}@fiit.stuba.sk.

Komplementárny dualizmus medzi mozgom a myslou spôsobuje určité ťažkosti pri interpretácii kognitívnych aktivít mysle. Čisto neurálny prístup k interpretácii kognitívnych aktivít mysle sa upriamuje na hľadanie neurálnych korelátov medzi aktivitami neurónov a kognitívnymi aktivitami (konekcionizmus). Použitie neurálnej paradigmy k interpretácii symbolických kognitívnych aktivít má „vedľajší“ efekt v tom, že tieto sa nám „rozpúšťajú“ v ich mikroskopickom popise, symboly sa nám akoby strácajú v detailnom popise aktivít neurónov, intenzít synaptických spojov a pod. V opačnom prípade, absolutizovanie symbolickej paradigmy pri interpretácii kognitívnych aktivít mysli (kognitivizmus), ignorovanie skutočnosti, že myseľ je pevne ukotvená v mozgu, vedie ku koncepcnej sterilite, k snahe vysvetľovať kognitívne procesy ľudskej mysle len pomocou fenomenologických termínov odvodených od koncepcie symbolu. Navrhujú sa umelé symbolické konštrukty (metódy, algoritmy a pod.), pre ktoré obvykle neexistuje žiadna experimentálna podpora v neurovede. Cieľom tejto práce je poukázať na alternatívny prístup, ktorý prekonáva „priepasť“ medzi konekcionistickým a kognitivistickým prístupom k popisu a interpretácii kognitívnych aktivít ľudského mozgu [15-17]. Ukážeme, že použitie distribuovanej reprezentácie umožňuje integrovať konekcionizmus a kognitivizmus, kde mentálne reprezentácie (symboly) sú špecifikované pomocou distribuovaných obrazcov neurálnych aktivít, pričom nad týmito distribuovanými obrazcami môžeme zaviesť formálne algebraické operácie, ktoré umožňujú matematicky modelovať nielen kognitívne operácie, ale aj simulovať procesy prebiehajúci pri ukladaní a vyberaní informácie uloženej v pamäti.



Obrázok 1. Znázornenie prechodu od neurónovej siete k distribuovanej reprezentácii. Stav neurónovej siete v čase t je zadaný aktivitami jednotlivých neurónov, ktoré sú určené aktivitami v predchádzajúcom čase $t-1$ a váhovými koeficientmi jednotlivých orientovaných spojov. V určitej abstrakcie, môžeme tieto aktivity usporiadať do veľkého jednorozmerného poľa (vektora) útvaru, ktorého komponenty sú reálne čísla (ich veľkosť je určená stupňom šede príslušnej zložky – neurónu). V distribuovanej reprezentácii je ignorovaná architektúra neurónovej siete, t.j. dve distribuované reprezentácie musia byť chápané ako úplne nezávislé bez vzájomných súvislostí, ich kauzálna súvislosť pochádzajúca z neurónovej siete je teraz úplne ignorovaná. V distribuovanej reprezentácii zavádzame nové unárne a binárne operácie, pomocou ktorých vznikajú z pôvodných nové distribuované reprezentácie.

Našu pozornosť usmerníme na netradičný štýl uskutočňovania výpočtov pomocou distribuovaných obrazcov, ktorý sa diametrálne líši od klasických numerických a symbolických výpočtov a je vhodným modelovým prostriedkom pre pochopenie globálnych vlastností neurónových sietí. Ukážeme, že takéto „neuropočítanie“ je založené na rozsiahlych náhodne vytvorených obrazcoch (reprezentovaných mnohorozmernými vektormi s náhodnými zložkami), pozri obr. 1. Tento prístup, ktorého základné prístupy boli formulované už koncom 60-tich rokov [2,8,9,14,19], bol završený sériou prác Tony Plateom [10-12] o „holografickej redukovanej reprezentácii“ (HRR). Ukážeme, aké typy výpočtov sú týmto prístupom uskutočniteľné a či nám umožnia pochopiť procesy prebiehajúce v mozgu pri kognitívnych aktivitách. Naš príspevkom k rozvoju HRR spočíva v jeho aplikácia

k modelovaniu kognitívnych procesov usudzovania pomocou pravidiel modus ponens a modus tollens. Kanerva [3-5] v polovici 90-tich rokov navrhol určitú alternatívu k HRR, ktorá je založená na náhodne generovaných binárnych vektoroch.

2. Matematická formulácia holografickej reprezentácie

Cieľom tejto kapitoly je prezentácia základných vlastností holografickej reprezentácie, ktorá bola vypracovaná A. Plateom [10-12]. Základný pojem je *konceptuálny vektor*, ktorý je reprezentovaný n -rozmerným vektorom

$$\mathbf{a} \in R^n \Rightarrow \mathbf{a} = (a_0, a_1, \dots, a_{n-1}) \quad (1)$$

pričom jeho komponenty sú náhodné čísla so štandardným normálnym rozdelením

$$a_i = N(0, 1/n) \quad \forall i \in \{0, 1, \dots, n-1\} \quad (2)$$

kde $N(0, 1/n)$ je náhodné číslo so stredom v 0 a so strednou odchýlkou $1/n$.

Nad konceptuálnymi vektormi definujeme binárnu operáciu „konvolúciu“, ktorá dvojici vektorov priradí tretí vektor, $\otimes : R^n \times R^n \rightarrow R^n$, alebo

$$\mathbf{c} = \mathbf{a} \otimes \mathbf{b} \quad (3)$$

Zložky výsledného vektora $\mathbf{c} = (c_0, c_1, \dots, c_{n-1})$ sú určené vzťahom

$$c_i = \sum_{j=0}^{n-1} a_j b_{[i-j]} \quad (i = 0, 1, \dots, n-1) \quad (4)$$

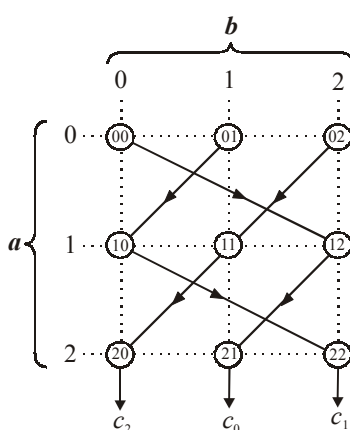
kde index v hranatej zátvorke, $[k]$, pomocou operácie *modulo n* takto³

$$k' = k \bmod n \quad (5a)$$

$$[k] = \begin{cases} k' & (\text{if } k' \geq 0) \\ n + k' & (\text{if } k' < 0) \end{cases} \quad (5b)$$

Konvolúcia dvoch vektorov \mathbf{a} a \mathbf{b} pre $n=3$ má tento tvar (pozri obrázok 2)

$$\begin{aligned} c_0 &= a_0 b_0 + a_1 b_2 + a_2 b_1 \\ c_1 &= a_0 b_1 + a_1 b_0 + a_2 b_2 \\ c_2 &= a_0 b_2 + a_1 b_1 + a_2 b_0 \end{aligned} \quad (6)$$



Obrázok 2. Znáozornenie konvolúcie dvoch vektorov \mathbf{a} a \mathbf{b} pre $n=3$ (pozri (6)).

³ Štandardná definícia aritmetickej operácie *k modulo n* je určená ako zvyšok po celočíselnom delení číslom n . Je potrebné poznamenať, že použitá definícia operácie *k modulo n* sa líši od tejto štandardnej definície pre záporné čísla k , ktorá poskytuje záporný výsledok. V našom prípade, ak je tento výsledok záporný, potom je transformovaný tak, že sa k nemu pripočíta n .

Konvolúcia vyhovuje týmto vlastnostiam:

- (1) komutatívnosť, $\mathbf{a} \otimes \mathbf{b} = \mathbf{b} \otimes \mathbf{a}$
- (2) asociatívnosť, $(\mathbf{a} \otimes \mathbf{b}) \otimes \mathbf{c} = \mathbf{a} \otimes (\mathbf{b} \otimes \mathbf{c})$
- (3) distributívnosť, $\mathbf{a} \otimes (\alpha \mathbf{b} + \beta \mathbf{c}) = \alpha (\mathbf{a} \otimes \mathbf{b}) + \beta (\mathbf{a} \otimes \mathbf{c})$
- (4) existencia jednotkového vektora, $\mathbf{1} \otimes \mathbf{a} = \mathbf{a}$ ($\mathbf{1} = (1, 0, \dots, 0)$)

Konvolúcia môže byť taktiež vyjadrená pomocou *cirkulantnej matice*[1]

$$\left. \begin{aligned} c_0 &= a_0 b_0 + a_1 b_2 + a_2 b_1 \\ c_1 &= a_0 b_1 + a_1 b_0 + a_2 b_2 \\ c_2 &= a_0 b_2 + a_1 b_1 + a_2 b_0 \end{aligned} \right\} \Rightarrow \begin{pmatrix} c_0 \\ c_1 \\ c_2 \end{pmatrix} = \underbrace{\begin{pmatrix} a_0 & a_2 & a_1 \\ a_1 & a_0 & a_2 \\ a_2 & a_1 & a_0 \end{pmatrix}}_{\text{circ}(a)} \begin{pmatrix} b_0 \\ b_1 \\ b_2 \end{pmatrix} \Rightarrow \mathbf{a} \otimes \mathbf{b} = \text{circ}(\mathbf{a})\mathbf{b} \quad (7)$$

Tento konkrétny príklad je zovšeobecnený pre ľubovoľnú dimenziu n takto

$$\mathbf{c} = \mathbf{a} \otimes \mathbf{b} \Leftrightarrow \mathbf{c} = \text{circ}(\mathbf{a})\mathbf{b} \Leftrightarrow \text{circ}(\mathbf{a}) = \begin{pmatrix} a_0 & a_{n-1} & \dots & a_2 & a_1 \\ a_1 & a_0 & \dots & a_3 & a_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{n-2} & a_{n-3} & \dots & a_0 & a_{n-1} \\ a_{n-1} & a_{n-2} & \dots & a_1 & a_0 \end{pmatrix} \quad (8)$$

kde všeobecná cirkulantná matica má maticové elementy

$$(\text{circ}(\mathbf{a}))_{ij} = a_{[i-j]} \quad (9)$$

Cirkulantná matica má tieto vlastnosti

$$\text{circ}(\mathbf{a} \otimes \mathbf{b}) = \text{circ}(\mathbf{a})\text{circ}(\mathbf{b}) \quad (10)$$

pretože konvolúcia je komutatívna operácia, potom cirkulantné matice navzájom komutujú

$$\text{circ}(\mathbf{a})\text{circ}(\mathbf{b}) = \text{circ}(\mathbf{b})\text{circ}(\mathbf{a}) \quad (11)$$

Nech X je inverzná matica k cirkulantnej matici $\text{circ}(\mathbf{a})$

$$X\text{circ}(\mathbf{a}) = \text{circ}(\mathbf{a})X = E \quad (12)$$

Jej alternatívny tvar je

$$X = \text{circ}^{-1}(\mathbf{a}) \quad (13)$$

Nech \mathbf{a}^{-1} je *inverzný vektor* vzhľadom k vektoru \mathbf{a} , $\mathbf{a}^{-1} \otimes \mathbf{a} = \mathbf{1} = (1, 0, \dots, 0, 0)$, potom za predpokladu, že cirkulantná matica je regulárna, $|\text{circ}(\mathbf{a})| \neq 0$, dostaneme

$$\text{circ}^{-1}(\mathbf{a}) = \text{circ}(\mathbf{a}^{-1}) \quad (14)$$

Definujme unárnu operáciu *involúcie* (pozri obrázok 3)

$$(\)^* : R^n \rightarrow R^n \quad (15)$$

vzťahom

$$\mathbf{b} = \mathbf{a}^* = (a_{[0]}, a_{[-1]}, \dots, a_{[-n+2]}, a_{[-n+1]}) \quad (16)$$

$$(a_0, \overset{\curvearrowright}{a_1}, \overset{\curvearrowright}{a_2}, \dots, \overset{\curvearrowright}{a_{n-2}}, \overset{\curvearrowright}{a_{n-1}})^* = (a_0, a_{n-1}, a_{n-2}, \dots, a_2, a_1)$$

Obrázok 3. Znázornenie unárnej operácie involúcie.

Operácia involúcie spĺňa tieto vzťahy

$$(\mathbf{a} + \mathbf{b})^* = \mathbf{a}^* + \mathbf{b}^* \quad (17a)$$

$$(\mathbf{a} \otimes \mathbf{b})^* = \mathbf{a}^* \otimes \mathbf{b}^* \quad (17b)$$

$$(\mathbf{a} \otimes \mathbf{b}^*) \cdot \mathbf{c} = \mathbf{a} \cdot (\mathbf{b} \otimes \mathbf{c}) \quad (17c)$$

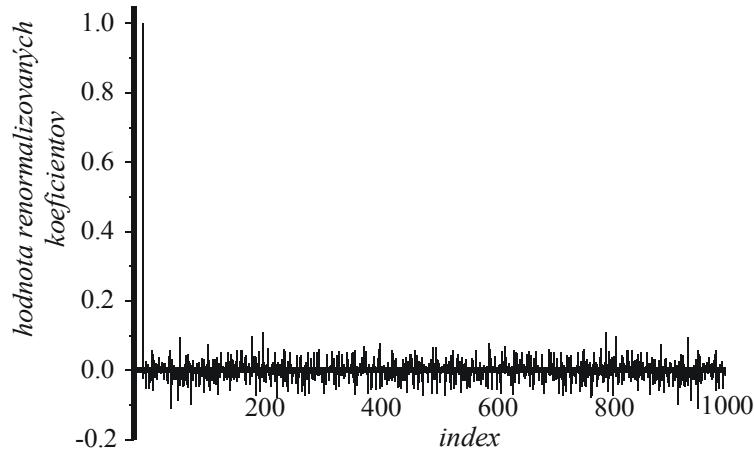
$$\mathbf{a}^{**} = \mathbf{a} \quad (17d)$$

$$\text{circ}(\mathbf{a}^*) = \text{circ}^T(\mathbf{a}) \quad (17e)$$

Dokážeme, že involúcia \mathbf{c}^* je približne rovná inverznému vektoru \mathbf{c}^{-1} , $\mathbf{c}^* \otimes \mathbf{c} \approx \mathbf{1}$. Uvažujme i -tú zložku konvolúcie $\mathbf{c}^* \otimes \mathbf{c}$

$$\begin{aligned} (\mathbf{c}^* \otimes \mathbf{c})_i &= \sum_{k=0}^{n-1} c_k^* c_{[i-k]} = \sum_{k=0}^{n-1} c_{[-k]} c_{[i-k]} \\ &= \begin{cases} \mathbf{c} \cdot \mathbf{c} & (\text{pre } i = 0) \\ \sum_{k=0}^{n-1} c_{[-k]} c_{[i-k]} & (\text{pre } i > 0) \end{cases} \end{aligned} \quad (18)$$

Nultá zložka konvolúcie $(\mathbf{c}^* \otimes \mathbf{c})_0$ je rovná skalárnemu súčinu $\mathbf{c} \cdot \mathbf{c}$, ktorý je vyjadrený ako suma kladných “diagonálnych” členov c_i^2 , pričom iné zložky $(\mathbf{c}^* \otimes \mathbf{c})_i$, pre $i \geq 1$, sú určené ako sumy “nediagonálnych” členov $c_i c_j$ s náhodným znamienkom. Táto skutočnosť nám implikuje, že absolútna hodnota nultého člena $(\mathbf{c}^* \otimes \mathbf{c})_0$ je o mnoho väčšia ako absolútne hodnoty zostávajúcich členov $(\mathbf{c}^* \otimes \mathbf{c})_i$, pre $i \geq 1$, čím sme dokázali $\mathbf{c}^* \otimes \mathbf{c} \approx \mathbf{1}$ (pozri obr. 2).



Obrázok 2. Histogram jednotlivých zložiek konvolúcie $\mathbf{c}^* \otimes \mathbf{c}$, pričom \mathbf{c} je náhodne vygenerovaný konceptuálny vektor dimenzie $n=1000$. Z obrázku vyplýva, že kladná hodnota nulej zložky $(\mathbf{c}^* \otimes \mathbf{c})_0$ je asi o dva rády väčšia, ako absolútna hodnota ostatných zložiek $(\mathbf{c}^* \otimes \mathbf{c})_i$, pre $i \geq 1$. To znamená, že vhodne normalizovaný súčin $\mathbf{c}^* \otimes \mathbf{c}$ sa približne rovná jednotkovému vektoru, $(\mathbf{c} \cdot \mathbf{c})^{-1} (\mathbf{c}^* \otimes \mathbf{c}) \doteq \mathbf{1} = (1, 0, \dots, 0, 0)$.

Jeden zo základných aspektov holografickej reprezentácie je možnosť rekonštrukcie jej pôvodných zložiek. Táto možnosť je veľmi dôležitá, pretože umožňuje dekódovať

pôvodnú informáciu obsiahnutú v komplexných konceptuálnych vektoroch. Rekonštrukcia \mathbf{x} z $\mathbf{c} \otimes \mathbf{x}$ je založená na vyššie dokázanej formuly $\mathbf{c}^* \otimes \mathbf{c} \approx \mathbf{1}$

$$\tilde{\mathbf{x}} = \mathbf{c}^* \otimes (\mathbf{c} \otimes \mathbf{x}) = (\mathbf{c}^* \otimes \mathbf{c}) \otimes \mathbf{x} \doteq \frac{1}{\mathbf{c} \cdot \mathbf{c}} \mathbf{1} \otimes \mathbf{x} = \frac{1}{\mathbf{c} \cdot \mathbf{c}} \mathbf{x} \quad (19)$$

Týmto sme dokázali, že konvolúcia \mathbf{c}^* s vektorom $\mathbf{c} \otimes \mathbf{x}$ produkuje vektor $\tilde{\mathbf{x}}$, ktorý je veľmi blízky pôvodnému vektoru \mathbf{x} , $\tilde{\mathbf{x}} \approx \mathbf{x}$. Tento dôležitý výsledok môžeme byť preformulovaný takto

$$\frac{1}{(\mathbf{c} \cdot \mathbf{c})} \tilde{\mathbf{x}} = \begin{pmatrix} x_0 \\ x_1 + \eta_1 \\ \dots \\ x_{n-1} + \eta_{n-1} \end{pmatrix} = \mathbf{x} + \boldsymbol{\eta} \quad (20)$$

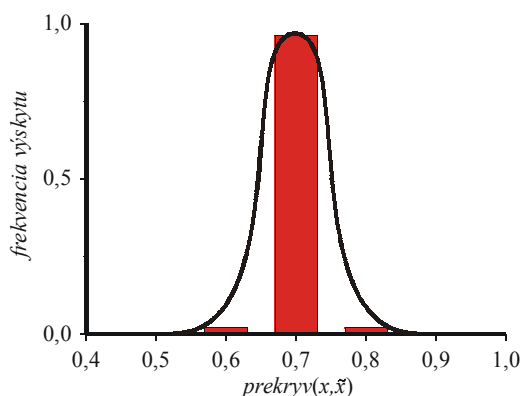
kde vektor $\boldsymbol{\eta}$ sa interpretuje ako šum s normálnou distribúciou so stredom v nule ale so štandardnou odchýlkou o mnoho menšou ako je štandardná odchýlka použitá pri generácii vektora \mathbf{x} .

Prekryv rezultujúceho vektora $\tilde{\mathbf{x}}$ s pôvodným vektorom \mathbf{x} je určený pomocou skalárneho súčinu takto

$$-1 \leq \text{prekryv}(\mathbf{x}, \tilde{\mathbf{x}}) = \frac{\mathbf{x} \cdot \tilde{\mathbf{x}}}{|\mathbf{x}| |\tilde{\mathbf{x}}|} \leq 1 \quad (21)$$

kde nerovnosti vyplývajú priamo zo Schwartzovej nerovnosti z lineárnej algebry. Čím je táto hodnota bližšie k maximálnej hodnote 1, tým sú si vektory $\tilde{\mathbf{x}}$ a \mathbf{x} podobnejšie⁴.

Na obrázku 4 je znázornený histogram prekryvov pre súčin $\mathbf{c} \otimes \mathbf{x}$, obsahujúci dvojicu náhodne vygenerovaných rôznych konceptuálnych vektorov \mathbf{c} a \mathbf{x} dimenzie $n=1000$. Z obrázku vyplýva, že najčastejší prekryv medzi $\tilde{\mathbf{x}} = \mathbf{c}^* \otimes \mathbf{c} \otimes \mathbf{x}$ a \mathbf{x} je okolo 0.7, z čoho vyplýva, že vektory $\tilde{\mathbf{x}}$ a \mathbf{x} sú si podobné, $\tilde{\mathbf{x}} \approx \mathbf{x}$.



Obrázok 4. Histogram prekryvov medzi vektormi $\tilde{\mathbf{x}}$ a \mathbf{x} (dimenzie $n=1000$) je s najväčšou frekvenciou okolo 0.7, z čoho vyplýva, že vektory $\tilde{\mathbf{x}}$ a \mathbf{x} sú si podobné.

Obráťme pozornosť na druhú možnosť verifikácie formuly (20b) pomocou prístupu nazývaného „aditívna pamäť“. Majme množinu obsahujúcu $p+q$ náhodne vygenerovaných konceptuálnych vektorov, $X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p, \mathbf{x}_{p+1}, \dots, \mathbf{x}_{p+q}\}$, pričom $p < q$. Pomocou prvých p vektorov z X definujme pamäťový vektor \mathbf{t} ako ich súčet

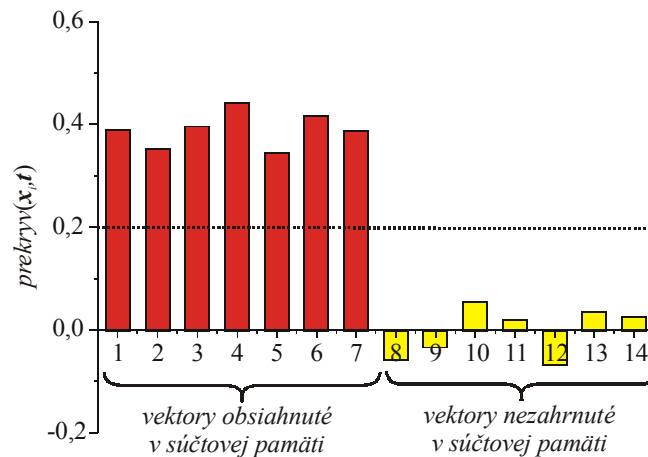
⁴ V prípade, že prekryv sa blíži k hodnote -1, potom vektory $\tilde{\mathbf{x}}$ a \mathbf{x} sú si taktiež podobné, aj keď sú opačne orientované (antikolineárne).

$$\mathbf{t} = \sum_{i=1}^p \mathbf{x}_i \quad (22)$$

Vektor \mathbf{t} reprezentuje aditívnu pamäť, ktorá jednoduchým aditívnym spôsobom obsahuje vektory z množiny X . Rozhodnutie o tom, či nejaký vektor $\mathbf{x} \in X$ je obsiahnutý v \mathbf{t} musí byť založené na hodnote prekryvu (21)

$$\text{prekryv}(\mathbf{x}, \mathbf{t}) = \frac{\mathbf{x} \cdot \mathbf{t}}{\|\mathbf{x}\| \|\mathbf{t}\|} \quad (23)$$

Ak je táto hodnota väčšia ako vopred zvolená prahová hodnota, $\text{prekryv}(\mathbf{x}, \mathbf{t}) \geq \vartheta$, potom vektor \mathbf{x} je zahrnutý v aditívnej pamäti \mathbf{t} , v opačnom prípade, ak $\text{prekryv}(\mathbf{x}, \mathbf{t}) < \vartheta$, potom vektor \mathbf{x} nie je zahrnutý v \mathbf{t} (pozri obr. 5).



Obrázok 5. Znázornenie súčtovej pamäti pre prvých 7 vektorov množiny X , ktorá obsahuje 14 náhodne vygenerovaných konceptuálnych vektorov dimenzie $n=1000$. Prahovú hodnotu ϑ v tomto prípade môžeme zvoliť 0.2.

V oboch predchádzajúcich príkladoch (pozri obr. 3 a 4) bol použitý rovnaký spôsob určenia konceptuálnych vektorov, ktoré sa „vyskytujú“ v nejakom inom komplexnom konceptuálnom vektore (ktorý môže byť výsledkom zložitých predchádzajúcich výpočtov – transformácií). Použitý spôsob sa nazýva „čistenie“ (angl. *clean-up*), ktorý je formálne špecifikovaný takto: Majme množinu vektorov $X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ a nejaký vektor \mathbf{t} . Stojíme pred problémom, rozhodnúť, či pamäťový vektor \mathbf{t} obsahuje aditívnu komponentu, ktorá je podobná (alebo nie je podobná) nejakému vektoru z množiny X . Tento problém riešime pomocou výpočtu tzv. prekryvu (23), formálne

$$\mathbf{x} \approx \mathbf{t} = \begin{cases} \text{áno} & (\text{prekryv}(\mathbf{x}, \mathbf{t}) \geq \vartheta) \\ \text{nie} & (\text{prekryv}(\mathbf{x}, \mathbf{t}) < \vartheta) \end{cases} \quad (24)$$

kde ϑ je zvolená *prahová hodnota* akceptovania veľkosti prekryvu ako pozitívnej odpovedi. Výsledkom tohto procesu čistenia je podmnožina vektorov

$$X(\mathbf{t}) = \{\mathbf{x} \in X; \mathbf{x} \approx \mathbf{t}\} \subseteq X \quad (25)$$

Môžeme si položiť otázku aj v trochu inej podobe, a to, či pamäťový vektor \mathbf{t} je podobný nejakému vektoru z množiny X ? Odpoveď na túto všeobecnejšiu otázku rozhodneme podľa maximálnej hodnoty prekryvu

$$\text{prekryv}(\mathbf{t}, X) = \max_{\mathbf{x} \in X} \text{prekryv}(\mathbf{t}, \mathbf{x}) \quad (26)$$

Potom môžeme prepísať (24) do tvaru

$$\mathbf{x} \approx X = \begin{cases} \text{áno} & (\text{prekryv}(\mathbf{x}, X) \geq \vartheta) \\ \text{nie} & (\text{prekryv}(\mathbf{x}, X) < \vartheta) \end{cases} \quad (27)$$

3. Asociačná pamäť

Konstruktia asociačnej pamäti patrí medzi hlavné výsledky holografickej redukovanej reprezentácie, ktorá môže byť ďalej zovšeobecnená pomocou tzv. agregácie (angl. chunking). Majme množinu konceptuálnych vektorov $X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ a tréningovú množinu $A_{train} = \{\mathbf{c}_i/\mathbf{x}_i; i=1,2,\dots,m\}$, ktorá obsahuje $m < n$ asociačných dvojíc konceptuálnych vektorov $\mathbf{c}_i/\mathbf{x}_i$, kde \mathbf{c}_i je vstup do asociačnej pamäti (angl. cue) a \mathbf{x}_i je výstup z pamäti. Zostrojíme pamäťový vektor \mathbf{t} , ktorý reprezentuje *asociačnú pamäť* nad tréningovou množinou A_{train}

$$\mathbf{t} = \mathbf{c}_1 \otimes \mathbf{x}_1 + \dots + \mathbf{c}_m \otimes \mathbf{x}_m = \sum_{i=1}^m \mathbf{c}_i \otimes \mathbf{x}_i \quad (28)$$

Predpokladajme, že vopred poznáme len vstupy asociačnej pamäti \mathbf{c}_i , nepoznáme možné výstupy z množiny $X_{train} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$. Odozva asociačnej pamäti na vstup – narážku \mathbf{c}_i je určená pomocou procesu „čistenia“ reprezentovaného formulou (27). V prvom kroku spočítame vektor $\tilde{\mathbf{x}}_i = \mathbf{c}_i^* \otimes \mathbf{t}$, potom pomocou procesu založenom na maximálnej hodnote prekryvu zistíme či $\tilde{\mathbf{x}}_i \approx \mathbf{x}_i \in X$

$$\text{prekryv}(\tilde{\mathbf{x}}_i, X) = \max_{\mathbf{x} \in X_{train}} \text{prekryv}(\tilde{\mathbf{x}}_i, \mathbf{x}) \quad (29)$$

Asociačnú pamäť bude ilustrovaná nasledujúcimi dvoma príkladmi.

Ilustračný príklad

Tento príklad využíva len tréningovú množinu $A_{train} = \{\mathbf{c}_i/\mathbf{x}_i; i=1,2,\dots,m\}$, ktorá je náhodne vygenerovaná pre $m=8$, pričom dimenzia konceptuálnych vektorov je $n=1000$. Pre každú asociačnú dvojicu $\mathbf{c}_i/\mathbf{x}_i$ spočítame $\mathbf{t}_i = \mathbf{c}_i \otimes \mathbf{x}_i$. Hodnoty $\text{prekryv}(\mathbf{c}_i^* \otimes \mathbf{t}_i, \mathbf{x}_j)$ sú uvedené v tabuľke

	1	2	3	4	5	6	7	8
1	0.71703	-0.01820	0.01452	0.02776	-0.01488	-0.01922	-0.02442	0.01358
2	-0.03998	0.73804	0.01510	0.01430	0.00276	0.02346	-0.00545	-0.01626
3	-0.02757	-0.01736	0.64667	0.00474	-0.11580	-0.00812	0.01476	0.00379
4	0.00785	0.00374	-0.01899	0.68728	-0.15340	0.00005	-0.00561	0.00136
5	-0.00466	0.00426	-0.01831	-0.00827	0.70767	0.04175	-0.03384	-0.00668
6	-0.01467	0.02522	-0.01403	-0.01316	-0.03000	0.71444	0.00078	-0.00526
7	0.02966	0.00892	-0.00301	-0.00358	0.01285	0.00971	0.70790	0.01816
8	-0.00344	-0.01080	0.00843	-0.01871	0.00324	-0.02629	0.00851	0.58957

Z tejto tabuľky jasne vyplýva, že prekryvy sú dostatočne veľké len pre diagonálne hodnoty, zatiaľ čo nediagonálne prekryvy sú rádovo menšie. Môžeme teda jednoznačne rozhodnúť pomocou prekryvu, či $\mathbf{c}_i^* \otimes \mathbf{t}_i \approx \mathbf{x}_i$ je asociantom s „narážkou“ \mathbf{c}_i .

Kombinácia súčtovej a asociačnej pamäti

Ukážeme, že aj kombinácia súčtovej a asociačnej pamäti je spoľahlivo rekognoskovateľná, čo sa ukáže v našich ďalších aplikáciách ako výhodná črta holografickej reprezentácie. Predpokladajme, že máme 10 konceptuálnych vektorov a, b, c, d, \dots, i, j , pomocou prvých štyroch zostrojíme kombináciu súčtovej a asociatívnej pamäti takto

$$t = a + b + a \otimes c + b \otimes d \quad (30)$$

Pomocou procedúry „čistenia“ získame, že vektor t obsahuje ako zložky vektory a a b , ktoré v následnom kroku odstránim z vektora t

$$t' = t - a - b = a \otimes c + b \otimes d \quad (31)$$

Z odstránenej aditívnej časti použitím techniky rekognoskácia asociatívnej pamäti zistíme, že obsahuje dve dvojice a/c a b/d , pozri obr. 6.

4. Postupnosť symbolov

Konstruktia asociačnej pamäti neumožňuje uchovať štruktúrovanú informácia, cieľom tejto kapitoly je ukázať, že holografická distribuovaná reprezentácia je schopná spracovať lineárny reťazec znakov, ktoré sú reprezentované postupnosťou konceptuálnych vektorov.

Pre konkrétnosť našich úvah, študujme postupnosť 6 konceptuálnych vektorov dimenzie $n=1000$

$$\text{postupnosť} = \{a \rightarrow b \rightarrow c \rightarrow d \rightarrow e \rightarrow f\} \quad (32)$$

Pre tieto vektory zostrojíme pamäťový vektor

$$t_0 = a + a \otimes b + a \otimes b \otimes c + a \otimes b \otimes c \otimes d + a \otimes b \otimes c \otimes d \otimes e + a \otimes b \otimes c \otimes d \otimes e \otimes f \quad (33)$$

Vieme, že tento vektor obsahuje postupnosť vektorov zakódovanú vektorom (33), avšak nevieme ktorých vektorov a v akom poradí. Ukážeme, že pomocou procedúry „čistenia“ z vektora t_0 môžeme postupne rekonštruovať pôvodnú postupnosť (32) nasledujúcou procedúrou (pozri obr. 7):

1. krok: $a = \text{clean_up}(t_0)$, $t_1 := t_0 - a$,

$$\tilde{t}_1 := a^* \otimes t_1,$$

2. krok: $b = \text{clean_up}(\tilde{t}_1)$, $t_2 := t_1 - a \otimes b$,

$$\tilde{t}_2 := (a \otimes b)^* \otimes t_2,$$

3. krok: $c = \text{clean_up}(\tilde{t}_2)$, $t_3 := t_2 - a \otimes b \otimes c$,

$$\tilde{t}_3 := (y_1 \otimes y_2 \otimes y_3)^* \otimes t_3,$$

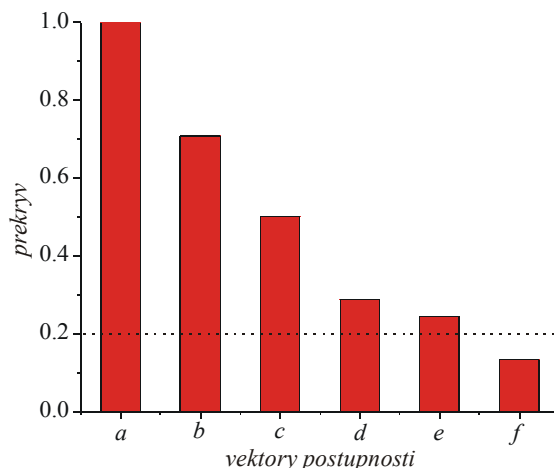
4. krok: $d = \text{clean_up}(\tilde{t}_3)$, $t_4 := t_3 - a \otimes b \otimes c \otimes d$,

$$\tilde{t}_4 := (a \otimes b \otimes c \otimes d)^* \otimes t_4,$$

5. krok: $e = \text{clean_up}(\tilde{t}_4)$, $t_5 := t_4 - a \otimes b \otimes c \otimes d \otimes e$,

$$\tilde{t}_5 := (a \otimes b \otimes c \otimes d \otimes e)^* \otimes t_5,$$

6. krok: $f = \text{clean_up}(\tilde{t}_5)$.



Obrázok 6. Prekryv pre jednotlivé vektory s postupnosti z (30), ktoré boli získané rekonštrukciou z vektora t_0 . Z obrázku vyplýva dôležitý poznatok, že nastáva pomerne rýchla degradácia rekonštrukcie, už šiesty vektor f je rekonštruovaný s pravdepodobnosťou menšou ako 0.20.

Funkcia $clean_up(\cdot)$ vykonáva „čistenie“ pre daný vektor t vzhľadom k množine vektorov $X = \{a, b, \dots, f, g, h, \dots\}$. Jednotlivé kroky rekonštrukcie postupnosti konceptuálnych vektorov – znakov sú znázornené na obr. 6, z ktorého vyplýva, že proces rekonštrukcie postupnosti symbolov pomerne rýchlo degraduje, už pri šiestom vektore je prekryv menší ako 0.2.

Postupnosť symbolov môže byť kódovaná taktiež pomocou asociačnej pamäti, kde vektor vstupu c_i špecifikuje i -tú pozíciu daného symbola. Vyššie uvedený ilustračný príklad je reprezentovaný pamäťovým vektorom

$$t = c_1 \otimes a + c_2 \otimes b + c_3 \otimes c + c_4 \otimes d + c_5 \otimes e + c_6 \otimes f \quad (34)$$

Potom rekognoskácia tejto postupnosti spočíva v hľadaní asocianta k vstupnému vektoru c_i , pomocou procesu čistenia zostrojíme „tréningovú množinu“

$$A_{train} = \{c_1/a, c_2/b, c_3/c, c_4/d, c_5/e, c_6/f\} \quad (35)$$

ktorá jednoznačne špecifikuje postupnosť vektorov. Výhoda tohto postupu je v tom, že presnosť rekognoskácia nedegraduje tak rýchlo ako pri pôvodnom postupe založenom na pamäťovom vektore (33).

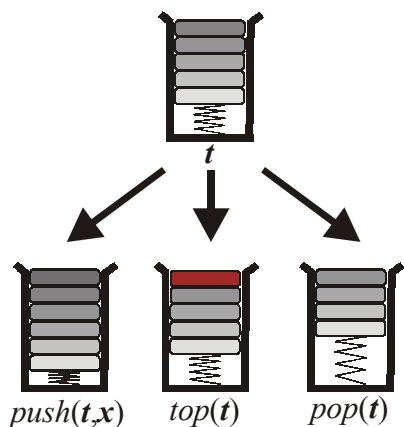
Asociačný prístup k implementácii pamäti pre postupnosť symbolov (reprezentovných vektormi a, b, \dots) je možné jednoducho zmodifikovať do tvaru tzv. „zásobníkovej pamäti“. Asociačné vstupy c_i sú určené takto

$$c_i = p^i \quad (i=1,2,\dots) \quad (34)$$

kde p^i je i -tá (konvolučná) mocnina náhodne generovaného konceptuálneho vektora p , $p^i = p^{i-1} \otimes p$. Potom pamäťový vektor (32) má tvar

$$t = p \otimes x_1 + p^2 \otimes x_2 + p^3 \otimes x_3 + \dots + p^n \otimes x_n \quad (35)$$

kde x_i sú jednotlivé položky pamäti $\{a, b, \dots\}$. Takto interpretovaná asociatívna pamäť pre postupnosť symbolov sa nazýva „zásobníková pamäť“, pomocou mocninových vstupov p^i môžeme jednoducho meniť jej obsah, pozri obr. 7. Nad touto pamäťou definujeme tri rôzne operácie, pomocou ktorých môžeme meniť jej obsah:



Obrázok 7. Tri možné operácie nad zásobníkovou pamäťou reprezentovanou vektorom.

- (1) $push(t, x) = p \otimes x + p \otimes t$, nová položka x je zavedená na vrchol zásobníka.
- (2) $top(t) = clean_up(p^* \otimes t)$, rekognoskuje vrchnú položku v zásobníku.
- (3) $pop(t) = p^{-1} \otimes t - top(t)$, zo zásobníka je odstránená vrchná položka.

Najproblematickejšia je tretia operácia, pomocou ktorého je zo zásobníka odstránená vrchná položka. Korektná implementácia požaduje použitie presného inverzného vektora p^{-1} , aproximácia tohto inverzného vektora pomocou involúcie, $p^{-1} \doteq p^*$, vedie k rýchlej degradácii zásobníkovej pamäti.

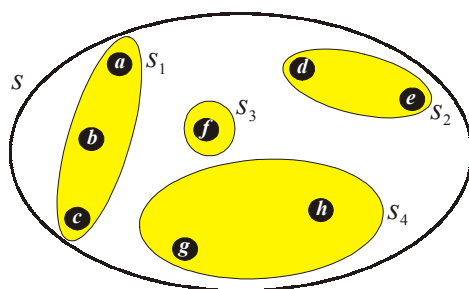
5. Agregovaná pamäť

Agregovaná pamäť pomáha prekonávať problémy s degradáciou pamäti pre postupnosť symbolov (pozri kap. 4). Majme množinu konceptuálnych vektorov $S = \{a, b, \dots, k, l, \dots\}$, túto množinu rozložíme na disjunktné podmnožiny - agregáty

$$S = S_1 \cup S_2 \cup S_3 \cup S_4 \cup \dots \quad (S_i \cup S_j = \emptyset, \text{ pre } i \neq j) \quad (36)$$

Študujme množinu $S = \{a, b, c, d, e, f, g, h\}$, jej rozklad na agregáty vypadá takto (pozri obr. 8)

$$S_1 = \{a, b, c\}, S_2 = \{d, e\}, S_3 = \{f\}, \text{ and } S_4 = \{g, h\} \quad (37)$$



Obrázok 8. Ilustrácia rozkladu množiny S na disjunktné agregáty, pozri (37).

Agregovaná pamäť je reprezentovaná pamäťovým vektorom, ktorý reprezentuje postupnosť agregátov $\{s_1 \rightarrow s_2 \rightarrow s_3 \rightarrow s_3\}$

$$t = s_1 + s_1 \otimes s_2 + s_1 \otimes s_2 \otimes s_3 + s_1 \otimes s_2 \otimes s_3 \otimes s_4 \quad (38)$$

pričom jednotlivé agregáty sú definované takto pomocou príslušných postupnosti vektorov (pozri obr. 9)

$$s_1 = a + a \otimes b + a \otimes b \otimes c \quad (39c)$$

$$s_2 = d + d \otimes e \quad (39d)$$

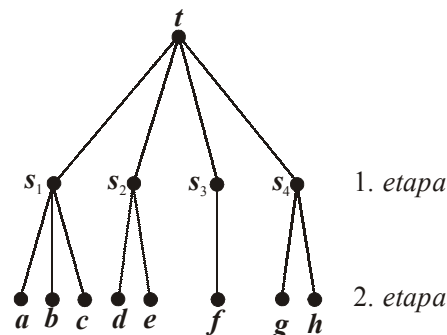
$$s_3 = f \quad (39e)$$

$$s_4 = g + g \otimes h \quad (39f)$$

Proces spracovania agregovanej pamäti obsahuje dve etapy:

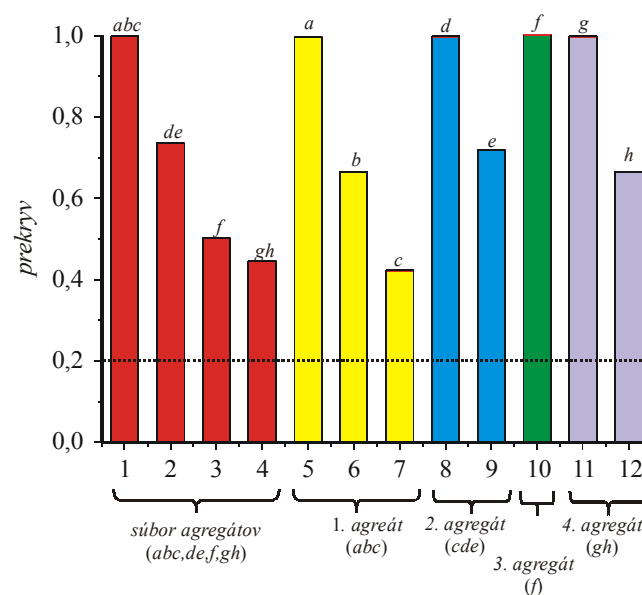
1. etapa – pomocou procesu „čistenia“ identifikujeme agregáty vyskytujúce sa v t (predpokladáme, že „upratovanie“ má množinu $X = \{x_1, x_2, \dots, x_n\}$ z konca 2. kapitoly, kde bol tento proces špecifikovaný, rozšírenú aj o agregáty s_1, s_2, s_3, s_4 , t.j. v našom ilustračnom príklade $X = \{a, b, c, d, e, f, g, h, s_1, s_2, s_3, s_4\}$).

2. etapa – identifikované agregáty sú ďalej analyzované pomocou procesu upratovania.



Obrázok 9. Agregácia 8 konceptuálnych vektorov na 4 agregáty. V 1. etape proces „upratovania“ identifikuje agregáty, ktoré sú potom v 2. etape ďalej identifikované na elementárne konceptuálne vektory.

Výsledky dvoj-etapového procesu „čistenia“ sú znázornené na obr. 10. Z tohto obrázku vyplýva, že v prípade dlhej sekvencie konceptuálnych vektorov rýchla degradácia procesu „čistenia“ môže čiastočne prekonaná pomocou agregácia konceptov na agregáty, ktoré sú na najvyššej úrovni samostatne kódované.



Obrázok 10. Znázornenie dvoj-etapového procesu „čistenia“, pričom v prvej etape sú identifikované agregáty, zatiaľ čo v druhej etape sú identifikované elementárne konceptuálne vektory z už identifikovaných agregátov.

6. Kódovanie relácií

Holografická redukovaná reprezentácia je vhodným prostriedkom aj na kódovanie relácií (predikátov). Študujme binárnu reláciu $P(x,y)$, použijúc Pascalovský zápis, táto relácia je formálne špecifikovaná touto hlavičkou

$$\text{function } P(x : \text{type}_1; y : \text{type}_2) : \text{type}_3 \quad (40)$$

Jednotlivé argumenty relácie sú špecifikované typmi type_1 a type_2 , ktoré špecifikujú doménu nad ktorou sú tieto premenné určené; podobne aj samotná relácia P je chápaná ako funkcia, ktorej doména hodnôt je špecifikovaná typom type_3 . V mnohých prípadoch doména premenných a aj samotnej relácie je jedna, preto ich špecifikácie môžu byť vynechané, čo podstatne zjednodušuje holografickú reprezentáciu relácií. Redukovaný tvar relácie (40) vypadá takto

$$\text{function } P(x; y) \quad (41)$$

kde už vopred vieme akého typu sú premenné x , y , a taktiež aj samotná relácia P . Holografická reprezentácia relácie (40) môže mať tento tvar

$$t = P + \text{variable}_1 + \text{variable}_2 + \\ P \otimes (\text{type}_3 + \text{variable}_1 \otimes (x + \text{type}_1) + \text{variable}_2 \otimes (y + \text{type}_2)) \quad (42)$$

Jej dekódovanie prebieha postupne. V prvej etape pomocou procedúry „čistenia“ rekognoskujeme názov (identifikátor) relácie P a taktiež aj názvy (identifikátory) jej jednotlivých premenných x a y . V druhom kroku rekognoskujeme typ type_3 relácie P , v záverečnom treťom kroku pomocou predchádzajúcich výsledkov rekognoskujeme premenné x , y a taktiež aj ich typy type_1 a type_2 . V mnohých prípadoch reprezentácia predikátora $P(x,y)$ je postačujúca v tomto zjednodušenom tvare (pozri (41))

$$t = P + \text{variable}_1 \otimes x + \text{variable}_2 \otimes y \quad (43)$$

Zvolená metóda holografickej reprezentácie relácie je ľahko zovšeobecniteľná aj zložitejšie relácie, kde premenné sú taktiež predikáty, napr. $P(x, Q(y, z))$, kde „vnútorný“ predikát Q je charakterizovaný

$$\text{function } Q(y : \text{type}_3; z : \text{type}_4) : \text{type}_5 \quad (44)$$

K tomu, aby sme mohli vytvoriť zloženú reláciu $P(x, Q(y, z))$, musíme predpokladať typovú kompatibilitu druhej premennej relácie P a typu relácie Q , t.j. $\text{type}_2 = \text{type}_4$. V zjednodušenom prístupe, kde všetky typy sú rovnaké, nie je potrebné rozlišovať typy jednotlivých premenných a samotných relácií. Zjednodušená holografická reprezentácia relácie (44) má tento tvar

$$t' = Q + \text{variable}_1 \otimes y + \text{variable}_2 \otimes z \quad (45)$$

Dosadením reprezentácie (45) za premennú y v reprezentácii (43) dostaneme túto výslednú reprezentáciu zloženej relácie $P(x, Q(y, z))$

$$t = P + \text{variable}_1 \otimes x + \text{variable}_2 \otimes (Q + \text{variable}_3 \otimes y + \text{variable}_4 \otimes z) \\ = P + \text{variable}_1 \otimes x + \text{variable}_2 \otimes Q + \\ \text{variable}_2 \otimes \text{variable}_3 \otimes y + \text{variable}_2 \otimes \text{variable}_4 \otimes z \quad (46)$$

Ilustračný príklad – podobnosť medzi geometrickými obrázkami

Na obrázku 11 je znázornených 48=6×8 geometrických útvarov, ktoré obsahujú v horizontálnom alebo vertikálnom usporiadaní dva objekty, ktoré navyše môžu byť ešte malé alebo veľké. Označme holografické reprezentácie príslušných atomárnych koncepcií takto:

Obrázce: *tr* (triangle), *sq* (square), *ci* (circle), *st* (star)

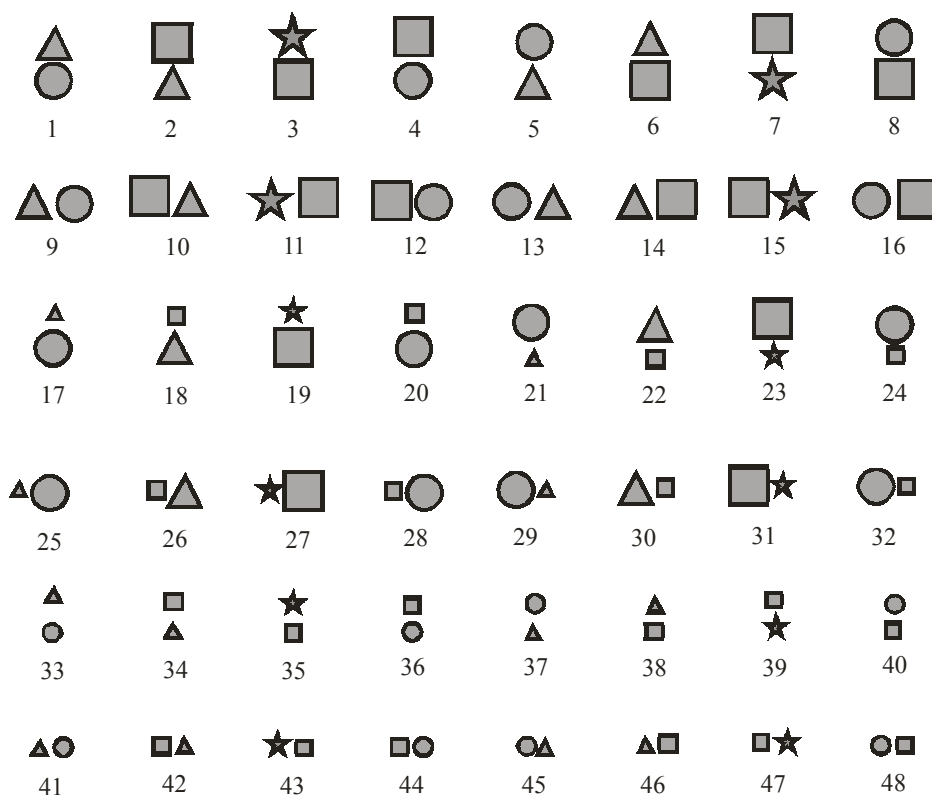
Unárne relácie: *sm* (small), *lg* (large)

Binárne relácie: *hor* (horizontal), *ver* (vertical)

Premenné: *ver_var₁* (1st variable for binary relation *ver*), *ver_var₂* (2nd variable for binary relation *ver*), *hor_var₁* (1st variable for binary relation *hor*), *hor_var₂* (2nd variable for binary relation *hor*)

Jednotlivé obrázky z obr. 11 sú charakterizované reláciami uvedenými v nasledujúcej tabuľke.

riadok	špecifikácia
1	<i>ver(lg(x),lg(y))</i>
2	<i>hor(lg(x),lg(y))</i>
3	<i>hor(sm(x),lg(y))</i> a <i>hor(lg(x),sm(y))</i>
4	<i>ver(sm(x),lg(y))</i> a <i>ver(lg(x),sm(y))</i>
5	<i>ver(sm(x),sm(y))</i>
6	<i>hor(sm(x),sm(y))</i>



Obrázok 11. Množina 48 podobných obrázkov, ktoré obsahujú dva objekty, umiestnené buď vedľa seba alebo nad sebou a objekty sú buď malé alebo veľké. Každý stĺpec obsahuje dvojicu podobných objektov, ktoré sa líšia len svojím umiestnením alebo veľkosťou.

Holografické reprezentácie jednotlivých prípadov z tejto tabuľky majú tento tvar (porovnaj s rovnicou (43)).

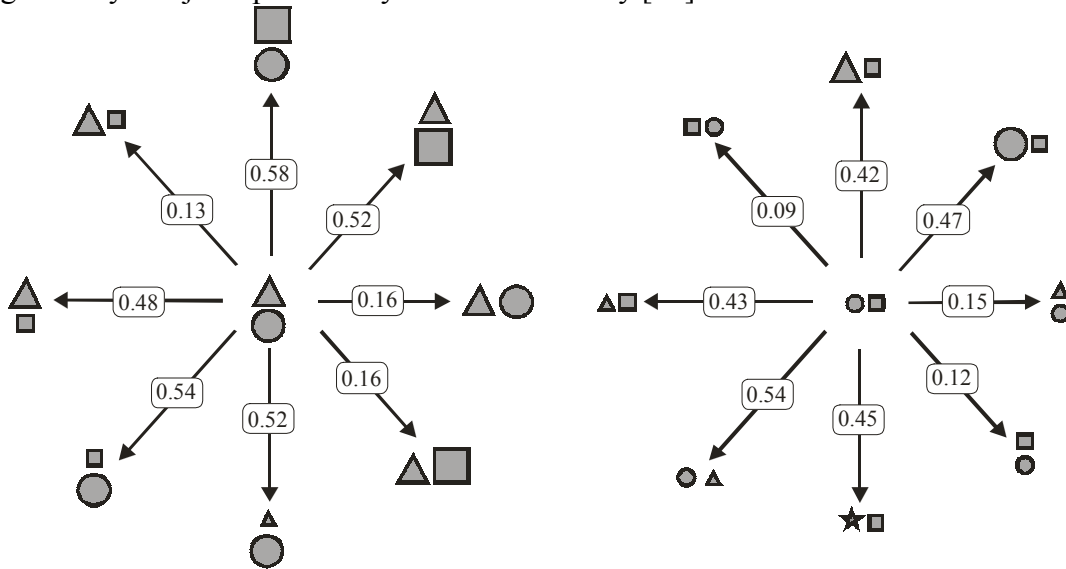
$$\begin{aligned}
t_{1,x,y} &= \mathit{ver} + \langle \mathit{ver_var}_1 \otimes \mathit{lg} \otimes x + \mathit{ver_var}_2 \otimes \mathit{lg} \otimes y \rangle \\
t_{2,x,y} &= \mathit{hor} + \langle \mathit{hor_var}_1 \otimes \mathit{lg} \otimes x + \mathit{hor_var}_2 \otimes \mathit{lg} \otimes y \rangle \\
t_{3,x,y} &= \begin{cases} \mathit{ver} + \langle \mathit{ver_var}_1 \otimes \mathit{lg} \otimes x + \mathit{ver_var}_2 \otimes \mathit{sm} \otimes y \rangle \\ \mathit{ver} + \langle \mathit{ver_var}_1 \otimes \mathit{sm} \otimes x + \mathit{ver_var}_2 \otimes \mathit{lg} \otimes y \rangle \end{cases} \\
t_{4,x,y} &= \begin{cases} \mathit{hor} + \langle \mathit{hor_var}_1 \otimes \mathit{lg} \otimes x + \mathit{hor_var}_2 \otimes \mathit{sm} \otimes y \rangle \\ \mathit{hor} + \langle \mathit{hor_var}_1 \otimes \mathit{sm} \otimes x + \mathit{hor_var}_2 \otimes \mathit{lg} \otimes y \rangle \end{cases} \\
t_{5,x,y} &= \mathit{ver} + \langle \mathit{ver_var}_1 \otimes \mathit{sm} \otimes x + \mathit{ver_var}_2 \otimes \mathit{sm} \otimes y \rangle \\
t_{6,x,y} &= \mathit{hor} + \langle \mathit{hor_var}_1 \otimes \mathit{sm} \otimes x + \mathit{hor_var}_2 \otimes \mathit{sm} \otimes y \rangle
\end{aligned} \tag{47}$$

kde x a y sú holografické reprezentácie jednotlivých obrázkov (tr , sq , ci , st) a zátvorka $\langle u \rangle$ indikuje, že vektor u je normalizovaný. Podobnosť medzi jednotlivými obrázkami je určená prekryvom ich holografickými reprezentáciami

$$\mathit{podobnost}(X, X') = \mathit{overlap}(t, t') \tag{48}$$

Získané výsledky sú znázornené na obr. 12. Dominantným determinantom podobnosti je ich vertikálne alebo horizontálne usporiadanie. Prekryv (t.j. aj podobnosť) medzi dvoma obrázkami, ktoré majú rôzne usporiadanie je obvykle menší ako 0.1.

Vo všeobecnosti, holografická redukovaná reprezentácia umožňuje pomerne jednoduché stanovenie podobnosti medzi objektmi, ktoré sú špecifikované predikátovou štruktúrou (40) alebo jej zovšeobecnením pomocou ďalších vložených (nested) predikátov (pozri (46)). Táto možnosť otvára úplne nové možnosti ako hľadať podobné objekty alebo riešiť problém analógie, ktorý je štandardne považovaný za veľmi obtiažny problém umelej inteligencie vyžadujúci špeciálne symbolické techniky [20].



Obrázok 12. Znázornenie podobných obrázkov pre dva vybrané obrázky 1 a 48 (pozri obr. 16). Jednotlivé šípky obsahujú prekryv medzi obrázkami spočítaný vzťahom (48).

7. Uvažovanie pomocou modus ponens a modus tollens

Simulácia procesov uvažovania (inferencie) patrí medzi základné problémy, ktoré sa znovu a znovu riešia v umelej inteligencii a v kognitívnej vede [13,16,18]. Fodorova kritika konekcionizmu [16] bola práve založená na unáhlenom závere, že umelé neurónové siete nie sú schopné simulovať vyššie kognitívne aktivity, ktoré sú vraj výhradnou doménou

klasického symbolického prístupu. Tento Fodorov názor sa ukázal ako nesprávny, ďalší rozvoj teórie neurónových sietí ukázal, že konekcionizmus je univerzálnym výpočtovým prostriedkom, ktorý nemá hranice aplikovateľnosti, neexistujú preň domény neaplikovateľnosti, ktoré by boli preň zakázané, samozrejme môže sa ukázať, že v niektorých doménach jeho použitie je neobyčajne ťažkopádne a zbytočne komplikované, že existujú iné prístupy, ktoré v danej doméne poskytujú podstatne jednoduchšie a priamočiarejšie riešenie, než ako umožňujú neurónové siete.

V tejto kapitole ukážeme na možnosť reprezentácie dvoch základných módov deduktívneho uvažovania výrokovej logiky,

$$\frac{p \Rightarrow q}{q} \quad \text{a} \quad \frac{p \Rightarrow q}{\bar{p}} \quad (49)$$

ktoré sa nazývajú *modus ponens* resp. *modus tollens*. Tieto módy uvažovania sú ekvivalentné týmto tautológiám výrokovej logiky

$$((p \Rightarrow q) \wedge p) \Rightarrow q \quad (50)$$

$$((p \Rightarrow q) \wedge \bar{q}) \Rightarrow \bar{p} \quad (51)$$

Implikáciu ' \Rightarrow ' môžeme chápať ako binárnu reláciu, ktorú je možno holografickým distribuovaným spôsobom reprezentovať takto (pozri vzťah (45))

$$\mathbf{t}_{p \Rightarrow q} = \mathbf{op} \otimes \mathbf{impl} + \mathbf{var}_1 \otimes \mathbf{p} + \mathbf{var}_2 \otimes \mathbf{q} \quad (52)$$

ktorá obsahuje špecifikuje typ relácie (implikáciu), druhá a tretia časť špecifikuje prvú (antedecedent) resp. druhú (konsekvent) premennú relácie implikácie. Tento konceptuálny vektor reprezentujúci reláciu implikácie môže byť transformovaný takto

$$\tilde{\mathbf{t}}_{p \Rightarrow q} = \mathbf{t}_{p \Rightarrow q} \otimes \mathbf{T} \quad (53)$$

kde

$$\mathbf{T} = \mathbf{var}_1^* \otimes \mathbf{p}^* \otimes \mathbf{p}^* \otimes \mathbf{q} + \mathbf{var}_2^* \otimes \mathbf{q}^* \otimes \bar{\mathbf{q}}^* \otimes \bar{\mathbf{p}} \quad (54)$$

Transformovaná reprezentácia implikácie je reprezentuje súčet dvoch asociačných dvojíc

$$\tilde{\mathbf{t}}_{p \Rightarrow q} \approx \mathbf{p}^* \otimes \mathbf{q} + \bar{\mathbf{q}}^* \otimes \bar{\mathbf{p}} \quad (55)$$

ktorá poskytuje holografickú reprezentáciu pravidiel *modus ponens* a *modus tollens*

$$\mathbf{p} \otimes \tilde{\mathbf{t}}_{p \Rightarrow q} \approx \mathbf{q} \quad (56a)$$

$$\bar{\mathbf{q}} \otimes \tilde{\mathbf{t}}_{p \Rightarrow q} \approx \bar{\mathbf{p}} \quad (56b)$$

Prvý vzťah (56a) môžeme chápať ako holografickú reprezentáciu *modus ponens* (pozri (51) a (52a)), zatiaľ čo druhý vzťah je holografická reprezentácia *modus tollens* (pozri (51) a (52b)).

Podobný výsledok môže byť získaný aj iným alternatívnym spôsobom, ktorý je založený na disjunktnej tvare implikácie

$$(p \Rightarrow q) \equiv (\bar{p} \vee q) \quad (57)$$

Distribuovaná reprezentácia implikácie v tomto alternatívnom tvare ma túto podobu

$$\mathbf{t}_{\bar{p} \vee q} = \mathbf{op} \otimes \mathbf{disj} + \mathbf{var}_1 \otimes \bar{\mathbf{p}} + \mathbf{var}_2 \otimes \mathbf{q} \quad (58)$$

Transformáciou tejto reprezentácie dostaneme (porovnaj (53))

$$\tilde{\mathbf{t}}_{\bar{p} \vee q} = \mathbf{t}_{\bar{p} \vee q} \otimes \mathbf{T} \approx \bar{\mathbf{p}} \otimes \bar{\mathbf{q}}^* + \mathbf{q} \otimes \mathbf{p}^* \quad (59a)$$

kde

$$\mathbf{T} = \mathbf{var}_1^* \bar{\mathbf{q}}^* + \mathbf{var}_2^* \mathbf{p}^* \quad (59a)$$

Táto transformácia je podstatne jednoduchšia než ako v predchádzajúcom prípade (54b). Pravidlá *modus ponens* a *modus tollens* sú teraz realizované vzťahmi podobnými s (56a-b). Navyše dostávame ešte tieto dve „pravidlá“

$$q^* \otimes \tilde{t}_{\bar{p} \vee q} \approx p^* \quad (60a)$$

$$\bar{p}^* \otimes \tilde{t}_{\bar{p} \vee q} \approx \bar{q}^* \quad (60b)$$

ktoré pripomínajú dobre známe pravidlá chybného uvažovania

$$\begin{array}{c} \cancel{p \Rightarrow q} \\ \cancel{q} \\ \cancel{p} \end{array} \quad \text{a} \quad \begin{array}{c} \cancel{p \Rightarrow q} \\ \cancel{\bar{p}} \\ \cancel{\bar{q}} \end{array} \quad (61)$$

ktoré sa nazývajú „*potvrdenie dôsledku*“ resp. „*poprenie predpokladu*“. Táto skutočnosť je spôsobená tým, že transformované reprezentácie implikácie $\tilde{t}_{p \Rightarrow q}$ a $\tilde{t}_{\bar{p} \vee q}$ nie sú identické, reprezentácia $\tilde{t}_{\bar{p} \vee q}$ vedie práve k neočakávaným výsledkom (61), ktoré reprezentujú nesprávne módy usudzovania (ktoré sú však často, aj napriek tomu, často používané ľuďmi nepoznajúcimi zásady logiky).

8. Predikátová logika

Budeme sa zaoberať jednoduchou formou predikátovej logiky, ktorá je založená na unárnych predikátoch, $P(x)$, ktorých distribuovaná reprezentácia má tvar (pozri kapitolu 6)

$$t_{P(x)} = \mathbf{pred} \otimes P + \mathbf{pred_var} \otimes x \quad (62)$$

Spojenie tohoto predikátu s univerzálnym kvantifikátorom, $(\forall x)P(x)$, môže byť reprezentované nasledujúcim spôsobom

$$t_{(\forall x)} = \mathbf{uni_quant} \otimes \mathbf{uni} + \mathbf{uni_quant_var} \otimes x \quad (63a)$$

$$t_{(\forall x)P(x)} = t_{(\forall x)} + t_{(\forall x)} \otimes (\mathbf{pred} \otimes P + \mathbf{pred_var} \otimes x) \quad (63b)$$

Oba konceptuálne vektory $t_{P(x)}$ a $t_{(\forall x)P(x)}$ môžu byť rekognoskované pomocou „čistiacej“ procedúry.

Pre naše ďalšie účely štúdia procesov uvažovania v rámci predikátovej logiky a ich distribuovanej reprezentácie je tento prístup zbytočne zložitý, používanie konceptuálneho vektora $t_{(\forall x)}$ pre reprezentáciu symbolu $(\forall x)$ v podstate len zbytočne komplikuje proces rekognoskácie zložených konceptuálnych vektorov obsahujúcich ako zložku $t_{(\forall x)}$. Preto opustíme od explicitného používania symbolu $(\forall x)$, jeho význam bude substituovaný používaním „*univerzálnej premennej*“ x , t.j. predikát $P(x)$ obsahujúci univerzálnu premennú x sa interpretuje ako $(\forall x)P(x)$, môžeme teda s určitou obozretnosťou používať „formulu“ $(\forall x)P(x) \equiv P(x)$.

V predikátovej logike platí pravidlo univerzálneho odlúčenia, pomocou ktorého sa konkretizuje predikát s univerzálnym kvantifikátorom na predikát s konkrétnou premennou a , $(\forall x)P(x) \Rightarrow P(a)$, ktorý je dôsledkom jednoduchej tautológie výrokovej logiky $((p \wedge q) \wedge p) \Rightarrow q$. Použitím konvencie univerzálnej premennej x túto konkretizáciu prepíšeme do jednoduchšieho tvaru

$$P(x) \Rightarrow P(a) \quad (64)$$

Zostrojíme distribuovanú reprezentáciu tejto univerzálnej eliminácie jednoduchého unárneho replikátora pomocou transformačného vektora T , pozri formuly (54-56). Distribuovaná reprezentácia (64) má tvar

$$t_{P(a)} \approx t_{P(x)} \otimes T \quad (65)$$

kde $T = x^* \otimes a$, potom

$$t_{P(a)} \approx t_{P(x)} \otimes x^* \otimes a \quad (66)$$

Sumarizujúc naše úvahy, distribuovaná reprezentácia pravidla univerzálnej eliminácie (konkretizácie) sa uskutočňuje pomocou transformačného vektora T , pomocou ktorého sa univerzálna premenná x substituuje „konkrétnou“ premennou a .

Pomocou tohto zjednodušenej reprezentácie kvantifikovaných predikátov budeme študovať tzv. *zovšeobecnený modus ponens* a *zovšeobecnený modus tollens*

$$\frac{(\forall x)(P(x) \Rightarrow Q(x)) \quad P(a)}{Q(a)} \quad \text{a} \quad \frac{(\forall x)(P(x) \Rightarrow Q(x)) \quad \bar{Q}(a)}{\bar{P}(a)} \quad (67a)$$

alebo v zjednodušenej podobe pomocou univerzálnej premennej x

$$\frac{P(x) \Rightarrow Q(x) \quad P(a)}{Q(a)} \quad \text{a} \quad \frac{P(x) \Rightarrow Q(x) \quad \bar{Q}(a)}{\bar{P}(a)} \quad (67b)$$

Tieto zovšeobecnené schémy deduktívneho uvažovania vyplývajú priamo z ich štandardnej výrokovej formy (51) a konkretizácie (61). Distribuovaná reprezentácia hlavnej (vrchnej) premisy týchto pravidiel má tvar

$$t_{P(x) \Rightarrow Q(x)} = op \otimes impl + var_1 \otimes (pred \otimes P + pred_var \otimes x) + var_2 \otimes (pred \otimes Q + pred_var \otimes y) \quad (68)$$

Konkretizácia implikácie $P(x) \Rightarrow Q(x)$ na $P(a) \Rightarrow Q(b)$, je formálne vyjadriteľná pomocou implikácie (pozri (64))

$$(P(x) \Rightarrow Q(x)) \Rightarrow (P(a) \Rightarrow Q(a)) \quad (69)$$

kde práva strana má nasledujúcu distribuovanú reprezentáciu

$$t_{P(a) \Rightarrow Q(a)} = op \otimes impl + var_1 \otimes (P + a) + var_2 \otimes (Q + b) \quad (70)$$

Podobne ako v úvodnom ilustračnom príklade (pozri (65)), tento prechod vyjadrený implikáciou (69), môže byť v distribuovanej reprezentácii vyjadrený pomocou transformácie

$$\tilde{t}_{P(a) \Rightarrow Q(a)} = t_{P(a) \Rightarrow Q(a)} \otimes T \approx (P + a)^* \otimes (Q + b) + (\bar{Q} + b)^* \otimes (\bar{P} + a) \quad (71)$$

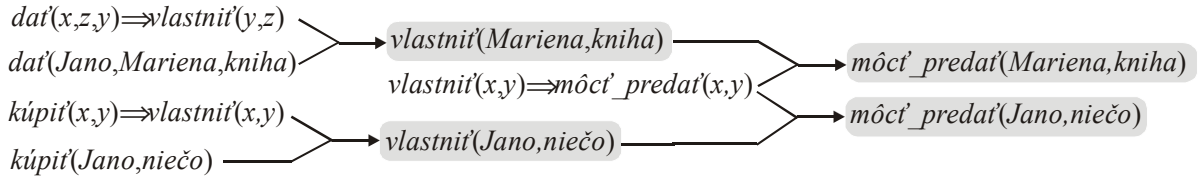
kde nová transformovaná distribuovaná reprezentácia $\tilde{t}_{P(a) \Rightarrow Q(a)}$ vyhovuje vzťahom, ktoré reprezentujú pravidlá (67) modus ponens resp. modus tollens

$$(P + a) \otimes \tilde{t}_{P(a) \Rightarrow Q(a)} \approx (Q + b) \quad (72a)$$

$$(\bar{Q} + b) \otimes \tilde{t}_{P(a) \Rightarrow Q(a)} \approx (\bar{P} + a) \quad (72b)$$

Ilustračný príklad – modelovanie reflexívneho uvažovania

V tomto ilustračnom príklade ukážeme, že holografická distribuovaná reprezentácia poskytuje formálne prostriedky, pomocou ktorého môžeme simulovať proces usudzovania založeného na zovšeobecnenom modus ponens (67). Tento proces bol široko študovaný Shastrim a Ajjanagaddom [13] pomocou konekcionistického systému nazývaného SHRUTI, ktorý bol schopný simulovať reflexívne uvažovanie založenom na predikátovej logike. Podobné výsledky budú dosiahnuté aj pomocou formalizmu holografickej distribuovanej reprezentácie.



Obrázok 13. Znázornenie použitia zovšeobecného pravidla modus ponens (64) pre vyvodenie poznatkov (označených tmavým podkladom a taktiež aj výstupnými šípkami) z implikácií (1-3) a zo vstupných pozorovaní (a-c), ktoré sú označené vstupnými šípkami.

Majme formálny systém obsahujúcou tieto tri všeobecné pravidlá (pozri obr.(13))

$$(1) \text{ dat}'(x, y, z) \Rightarrow \text{vlastniť}'(y, z), \text{ typ } x : \text{ darca}; \text{ typ } y : \text{ príjemca}; \text{ typ } z : \text{ vec},$$

$$(2) \text{ kúpiť}'(y, z) \Rightarrow \text{vlastniť}'(y, z), \text{ type } y : \text{ kupujúci}; \text{ typ } z : \text{ vec},$$

$$(3) \text{ vlastniť}'(y, z) \Rightarrow \text{môcť}'_ \text{predať}'(y, z), \text{ typ } y : \text{ vlastník}; \text{ typ } z : \text{ vec},$$

a tri pozorovania (skutočnosti – fakty)

$$(a) \text{ dat}'(Jano, Mariena, kniha),$$

$$(b) \text{ vlastniť}'(Mariena, kniha)$$

$$(c) \text{ kúpiť}'(Jano, niečo).$$

Aké sú deduktívne závery z tohto systému? Výsledky sú naznačené na obrázku 15, teraz pristúpime k ich odvodu pomocou distribuovanej reprezentácie založenej na konceptuálnych vektoroch a operáciách nad nimi.

Sledujme prvý zovšeobecný modus ponens z obr. 13

$$\text{dat}'(x, y, z) \Rightarrow \text{vlastniť}'(y, z)$$

$$\underline{\text{dat}'(Jano, Mariena, kniha)}$$

$$\text{vlastniť}'(Mariena, kniha)$$

(73)

Pomocou postupu (68-72) môžeme realizovať túto schému uvažovania pomocou reprezentácie konceptuálnych vektorov, jej jednotlivé zložky (idúc zhora nadol) sú reprezentované takto

$$\mathbf{t}_1^{(1)} = \mathbf{op} \otimes \mathbf{impl} + \mathbf{var}_1 \otimes \mathbf{t}_{\text{dat}'(x,y,x)} + \mathbf{var}_2 \otimes \mathbf{t}_{\text{vlastniť}'(y,x)} \quad (74a)$$

$$\mathbf{t}_2^{(1)} = \mathbf{t}_{\text{dat}'(Jano,Mariena,kniha)} = \mathbf{dat}' + \mathbf{dat}'_ \mathbf{var}_1 \otimes \mathbf{Jano} + \quad (74b)$$

$$\mathbf{dat}'_ \mathbf{var}_2 \otimes \mathbf{Mariena} + \mathbf{dat}'_ \mathbf{var}_3 \otimes \mathbf{kniha}$$

$$\mathbf{t}_3^{(1)} = \mathbf{t}_{\text{vlastniť}'(Mariena,kniha)} = \mathbf{vlastniť}' + \mathbf{vlastniť}'_ \mathbf{var}_1 \otimes \mathbf{Mariena} + \quad (74c)$$

$$\mathbf{vlastniť}'_ \mathbf{var}_2 \otimes \mathbf{kniha}$$

kde konceptuálne vektory $\mathbf{t}_{\text{dat}'(x,y,x)}$ a $\mathbf{t}_{\text{vlastniť}'(y,x)}$ sú zostrojené analogickým spôsobom ako v (43). V prvom kroku musíme vykonať konkrétnizáciu implikácie $\text{dat}'(x, y, z) \Rightarrow \text{vlastniť}'(y, z)$, tak že všeobecné premenné x, y, z sú substituované konkrétnymi premennými $Jano, Mariena, kniha$. Táto konkrétnizácia sa vykoná pomocou transformácie \mathbf{T} , ktorá je špecifikovaná formulami (65-66)

$$\hat{\mathbf{t}}_1^{(1)} \approx \mathbf{t}_1^{(1)} \otimes \mathbf{T} \quad (75)$$

kde

$$\hat{\mathbf{t}}_1^{(1)} = \mathbf{op} \otimes \mathbf{impl} + \mathbf{var}_1 \otimes \mathbf{t}_{\text{dat}'(Jano,Mariena,kniha)} + \mathbf{var}_2 \otimes \mathbf{t}_{\text{vlastniť}'(Mariena,kniha)} \quad (76a)$$

$$\mathbf{T} = \mathbf{t}_1^{*(1)} \otimes \hat{\mathbf{t}}_1^{(1)} \quad (76b)$$

Takto konkretizovaná reprezentácia $dat'(Jano, Mariena, kniha) \Rightarrow vlastniť'(Mariena, kniha)$ v následnom kroku je použiteľná pre modus ponens uskutočnenej pomocou formúl (71-72)

$$\tilde{\mathbf{t}}_1^{(1)} \approx \hat{\mathbf{t}}_1^{(1)} \otimes \mathbf{T}' \quad (77a)$$

kde rezultujúci konceptuálny vektor $\tilde{\mathbf{t}}_1^{(1)}$ už reprezentuje modus ponens, t.j. platí

$$\mathbf{t}_{dat'(Jano, Mariena, kniha)} \otimes \tilde{\mathbf{t}}_1^{(1)} \approx \mathbf{t}_{vlastniť'(Mariena, kniha)} \quad (78)$$

Podobným spôsobom môžu byť uskutočniteľné aj ostatné tri zovšeobecnené modus ponens z obr. 13.

9. Závery

Holografická redukovaná reprezentácia ponúka nové nekonvenčné riešenie jedného zo základných problémov umelej inteligencie a kognitívnej vedy, ktorý spočíva v distribuovanom kódovaní štruktúrovanej informácie (postupnosť znakov, vnorené relačné štruktúry, a pod.). Použitá distribuovaná reprezentácia je založená na dvoch binárnych operáciách: unárnej operácii „involúcie“ a binárnej operácii „konvolúcia“ nad doménou n -rozmerých náhodne generovaných konceptuálnych vektorov, ktorých elementy spĺňajú normálnu distribúciu $N(0,1/n)$. Pomocou tejto distribuovanej reprezentácie môžeme modelovať rôzne typy asociatívnej pamäti, ktoré sú reprezentované konceptuálnym vektorom a taktiež dekódovať pamäťový vektor, t.j. určiť konceptuálne vektory, ktoré ho vytvorili. Táto rekognoskácia pamäťového vektora sa vykonáva pomocou procedúry čistenia (clean-up), ktorá nám určí pomocou vzájomného prekryvu vektor, ktorý je mu najviac podobný. Popísali sme taktiež proces agregácie vektorov, ktorý umožňuje prekonať ťažkosti rýchlej degradácie rekognoskácie pamäťového vektora tak, niekoľko vektorov sa združí do agregátu a tento sa zapamätáva prostredníctvom pamäťového vektora. Holografická redukovaná reprezentácia umožňuje hľadať podobnosť medzi dvoma štruktúrovanými konceptami pomocou jednoduchej algebraickej operácie skalárneho súčinu medzi ich distribuovanými reprezentáciami. Táto skutočnosť môže byť veľmi užitočnou pre modelovanie procesov prehľadávania pamäti na základe podobnosti (analógie) jej jednotlivých položiek. V záverečnej časti práce sme ukázali, že holografická redukovaná reprezentácia môže byť taktiež použitá k modelovaniu procesov usudzovania založených na pravidlách modus ponens a modus tollens, taktiež sme ukázali plodnosť tohto prístupu pri modelovaní procesu zovšeobecnenia pomocou indukcie.

Literatúra

- [1] Davis, P. J.: *Circulant Matrices*. Chelsea Publishing, New York, 1999.
- [2] Gabor, D. Holographic model for temporal recall. *Nature*, **217** (1968), 1288–1289.
- [3] Kanerva, P.: Analogy as a Basis of Computation, In Uesaka, Y., Kanerva, P., Asoh, H.: *Foundations of real-world intelligence*. CSLI Publications, Stanford, CA, 2001.
- [4] Kanerva, P.: Binary Spatter-Coding of Ordered K-tuples. In von der Malsburg, C., von Seelen, W., Vorbruggen, J. C., Sendhoff, B. (eds.): *ICANN'96, Proceedings of the International Conference on Artificial Neural Networks*. Springer Verlag, Berlin, 1996, pp. 869-873.

- [5] Kanerva, P.: The Spatter Code for Encoding Concepts at Many Levels. In Marinaro, M., Morasso, P. G. (eds.): *ICANN '94, Proceedings of the International Conference on Artificial Neural Networks*. Springer Verlag, London, 1994, pp. 226–229.
- [6] Kvasnička, V., Pospíchal, J.: Holographic reduced representation in artificial intelligence and cognitive science. *Neurocomputing* (sent for publication, March 2004)
- [7] McClelland, J. L. M., Rumelhart, D. E., and the PDP research group (eds.): *Parallel distributed processing: Explorations in the microstructure of cognition*, volumes 1 and 2. The MIT Press, Cambridge, MA. (1986).
- [8] Metcalfe Eich, J.: Levels of processing, encoding specificity, elaboration, and charm *Psychological Review*, **92** (1985), 1–38.
- [9] Murdock, B. B.: A theory for the storage and retrieval of item and associative information. *Psychological Review*, **89** (1982), 316–338.
- [10] Plate, T. A.: Distributed Representations and Nested Compositional Structure. Department of Computer Science, University of Toronto. Ph.D. Thesis, 1994.
- [11] Plate, T. A.: *Holographic Reduced Representation: Distributed Representation for Cognitive Structures*, CSLI Publications, Stanford, CA, 2003.
- [12] Plate, T.: Holographic Distributed Representations. *IEEE Transaction on Neural Networks* **6**(1995), 623-641.
- [13] Shastri, L., Ajjanagadde, V.: From simple associations to systematic reasoning: connectionist representation of rules, variables, and dynamic bindings using temporal synchrony. *Behavioral and Brain Sciences*, **16** (1993), 417-494.
- [14] Slack, J. N.: The role of distributed memory in natural language processing. In O’Shea, T. (ed.): *Advances in Artificial Intelligence: Proceedings of the Sixth European Conference on Artificial Intelligence, ECAI-84*. Elsevier Science Publishers, New York, 1984.
- [15] Smolensky, P., Legendre, G., Miyata, Y.: Principles for an Integrated Connectionist/Symbolic Theory of Higher Cognition. Technical Report CU-CS-600-92, Department of Computer Science and 92-8, Institute of Cognitive Science. University of Colorado at Boulder, 1992, 75 pages.
- [16] Smolensky, P.: On the proper treatment of connectionism. *The Behavioral and Brain Sciences*, **11** (1988), 1-74.
- [17] Smolensky, P.: Tensor product variable binding and the representation of symbolic structures in connectionist networks. *Artificial Intelligence*, **46** (1990)., 159-216.
- [18] Smolensky, P.: Neural and conceptual interpretations of parallel distributed processing models. In J. L. McClelland, D. E. Rumelhart, & the PDP Research Group, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Volume 2: Psychological and Biological Models*. MIT Press - Bradford Books, Cambridge, MA, 1986, pp.. 390-431.
- [19] Willshaw, D. J., Buneman, O. P., and Longuet-Higgins, H. C.: Non-holographic associative memory. *Nature*, **222** (1969), 960–962.
- [20] Winston, P. H.: Learning new Principles From precedents and Exercises. *Artificial Intelligence* **19** (1982), 321-350.