

# KOGNITIVNÍ VĚDA A UMĚLÝ ŽIVOT

*sestavili*

*Jozef Kelemen*

*Ján Rybár*

*Igor Farkaš*

*Martin Takáč*

**Kognitivní věda a umělý život**  
(editovaná kniha)

Tato kniha je spolufinancována Evropským sociálním fondem a státním rozpočtem České republiky v rámci projektu CZ.1.07/2.2.00/28.0014 Interdisciplinární vzdělávání v ICT s jazykovou kompetencí.

Recenzenti příspěvků:

Doc. Ing. Igor Farkaš, PhD.  
Doc. PhDr. Ján Rybár, PhD.  
RNDr. Martin Takáč, PhD.  
Doc. RNDr. Gabriela Andrejková, CSc.  
RNDr. Zuzana Černeková, PhD.  
Ing. František Gyárfáš, PhD.  
RNDr. Martin Homola, PhD.  
PhDr. Dezider Kamhal, PhD.  
RNDr. Andrej Lúčny, PhD.  
RNDr. Marek Nagy, PhD.  
RNDr. Jozef Šiška, PhD.  
Mgr. Martin Čajági  
Mgr. Ľudovít Malinovský  
Mgr. Bogumiła Suwara, PhD.

Vydavatel: Slezská univerzita v Opavě  
Filozoficko-přírodovědecká fakulta v Opavě  
Bezručovo náměstí 13  
746 01 Opava

Sestavili © Prof. RNDr. Jozef Kelemen, DrSc., 2013  
Doc. PhDr. Ján Rybár, CSc., 2013  
Doc. Ing. Igor Farkaš, Ph.D., 2013  
RNDr. Martin Takáč, Ph.D., 2013

Autorská práva © autoři příspěvků, 2013

Cover Design © Róbert Kelemen, 2013

Tisk: Optys, spol. s. r. o., Opava

Vydáno v Opavě v květnu 2013

**ISBN 978-80-7248-863-6**

## Předslov

Kniha *Kognitivní věda a umělý život* slouží primárně jako vzdělávací materiál pro cílovou skupinu projektu. Doporučujeme ji však také odborné veřejnosti zabývající se výzkumem kognitivnosti a fenoménu života z rozličných hledisek, představuje část výzkumů, které se v těchto oblastech uskutečňují na Slovensku a v České republice. Takové výzkumy se v obou zemích dělají již roky a přinášejí výsledky, reflektované i nezanedbatelným počtem našich kolegů z jiných zemí. Domníváme se tedy, že předkládaná kniha zaujme také v našich jazykových kontextech a podnítky k dalšímu zkoumání, které přivedou k novým poznatkům o myšlení, poznávání, vědomí a životě.

Kniha je svým obsahem těsně spojená s v pořadí už 13. česko-slovenským seminářem *Kognícia a umelý život*. Jedná se o vědeckou interdisciplinární akci s mnohaletou tradicí, která vytváří prostor na prezentaci výsledků práce odborníků z Čech a Slovenska zabývajících se o různé aspekty a metody zkoumání kognice a života, vedené však především z pozic informatiky v živých, anebo umělých systémech. Hlavním cílem této každoroční konference je podnítit odborné diskuse, hledat styčné body a oblasti spolupráce v této dynamicky se rozvíjející interdisciplinární oblasti poznávání. Rádi konstatujeme, že právě interdisciplinarita je jedna z charakteristických vlastností této akce. Svědčí o tom i pestrost profesionálního zaměření autorů příspěvků této knihy – od informatiky přes neurovědu, psychologii, filosofii, až po literární vědu.

Už 13. ročník konference se v roce 2013 konal ve Vysokých Tatrách v Staré Lesné ve dnech 27. - 30. května. Za to, že se mohl konat v malebném prostředí Vysokých Tater v hotelu Academia, děkujeme řediteli Filozofického ústavu SAV prof. Tiborovi Pichlerovi, CSc. Také děkujeme zástupci ředitele a vědeckému tajemníkovi tohoto ústavu PhDr. Karolovi Kollárovi, CSc., že nám byl nápomocen při vyřizování administrativy související s touto akcí. Poděkování si taktéž zaslouží všichni členové Centra pre kognitívnu vedu na Fakulte matematiky, fyziky a informatiky Univerzity Komenského, kteří vyvinuli patřičné úsilí při přípravě a realizaci konference. Osobně chceme poděkovat RNDr. Kristíně Rebrové, autorce a administrátorce konferenční stránky. Děkujeme také firmě Nuabi, s.r.o., která poskytla finanční příspěvek na pokrytí konferenčních nákladů třem studentům vybraným programovým výbo-rem.

Děkujeme všem autorům a recenzentům za jejich práci s příspěvkem. Bez ní by tato kniha neexistovala.

Mgr. Viliamovi Dillingerovi děkujeme za technickou pomoc při editování knihy. Za účinnou spolupráci při jejím vydávání děkujeme slečně Haně Černínové z Ústavu informatiky Filozoficko-přírodovědecké fakulty Slezské univerzity v Opavě.

Tato kniha je spolufinancována Evropským sociálním fondem a státním rozpočtem České republiky v rámci projektu CZ.1.07/2.2.00/28.0014 Interdisciplinární vzdělávání v ICT s jazykovou kompetencí.

V Bratislavě a Opavě, květen 2013

Jozef Kelemen  
Ján Rybár  
Igor Farkaš  
Martin Takáč



# Obsah

Kombinácia SOM modelu neurónovej siete a vrstvy pre tréning s učiteľom <i>Gabriela Andrejková, Jozef Oravec</i> . . . . .	5
Fenomenologie volní akce: motorické intence a pocit záměru <i>Ondřej Bečev</i> . . . . .	17
Detekce patologií v B-MODE ultrazvukových obrazech s automatickou kontrolou kvality obrazu <i>Jiří Blahuta, Petr Čermák</i> . . . . .	27
Elektronická literatúra v paradigme interakcií človek-stroj <i>Petra Blšťáková</i> . . . . .	37
Multiagentní model monopsonu na trhu práce <i>Jan Burian</i> . . . . .	43
Preferencia k deliberácii: (ne)istota úspechu v kognitívnych úlohách <i>Vladimíra Čavojová, Eva Ballová Mikušková, Róbert Hanák</i> . . . . .	53
O komunitách v sociálnych sieťach <i>David Chalupa</i> . . . . .	59
P kolonie ve 2D <i>Luděk Cienciala, Lucie Ciencialová</i> . . . . .	65
Principy zdravého rozumu <i>Martina Číhalová, Jan Štěpán</i> . . . . .	73
Emócie a kognícia: ich interakcie a poruchy (vybrané aspekty) <i>Barbora Cimrová, Fedor Jagla, Igor Riečanský</i> . . . . .	79
Analýza efektu krátkodobej a dlhodobej motivácie pomocou fuzzy prístupu <i>Ladislav Clementis</i> . . . . .	87

Hierarchické učenie posilňovaním <i>Viliam Dillinger</i> . . . . .	93
Vyhrávajú stereotypy nad štatistikou? Normatívne a stereotypné uvažovanie vo vzťahu k úrovni kognitívnej reflexie <i>Katarína Dudeková, Lenka Kostovičová</i> . . . . .	101
Je reprezentačný pluralizmus v kognitívnej vede nevyhnutný? <i>Igor Farkaš</i> . . . . .	107
Multi-agentové systémy pro modelování biologických fenoménů <i>Martina Husáková</i> . . . . .	115
Primární mentální schopnosti jako měřítko inteligence v internetovém pro- storu <i>Anna Janikova, Kamil Matula, Štěpánka Tůmová</i> . . . . .	123
Tak mohli by stroje myslieť alebo nie? <i>Dezider Kamhal</i> . . . . .	129
Čo ešte ukáže čas vo výskume efektu spätného pohľadu? <i>Matúš Konečný</i> . . . . .	133
Podporou kognitívnej reflexie k tolerancii vyšších strát? Vplyv akontability na averziu voči strate je čiastočne mediovaný mierou analytického myslenia. <i>Lenka Kostovičová, Katarína Dudeková, Miroslav Sirota, Viera Bačová</i> . . .	141
Odhad atribútov písma na základe požiadaviek človeka <i>Miron Kuzma</i> . . . . .	147
Učenie reprezentácie objektov pomocou prekrývajúcich sa vývinových štruktúr <i>Andrej Lúčny, Michal Vician</i> . . . . .	153
Výpočtový model zavádzania ako spoločnej akcie <i>Ľudovít Malinovský</i> . . . . .	157
ARCH/GARCH modely a ich neurónová reprezentácia <i>Dušan Marček, Marek Horváth</i> . . . . .	167
Komplexné siete a modelovanie javov reálneho sveta <i>Mária Markošová</i> . . . . .	173

Datamining nad systémem eLogika <i>Marek Menšík, Jakub Gerlich</i> . . . . .	181
Možnosti rozsáhlých aplikací ambientní inteligence <i>Peter Mikulecký</i> . . . . .	189
Hybridní neuronové systémy pro návrh architektur autonomních agentů v oblasti umělého života <i>Pavel Nahodil, Jaroslav Vítků</i> . . . . .	197
Využití metod vicekriteriálního rozhodování pro analýzu e-mailových zpráv <i>Libor Olajec</i> . . . . .	205
Konekcionalistický model uchopovania <i>Matej Pecháč</i> . . . . .	211
Ještě jednou: Turing v Searlově čínské místnosti <i>Karel Pstružina</i> . . . . .	217
Robotický model systému zrkadliacich neurónov: experimentálna analýza <i>Kristína Rebrová, Igor Farkaš</i> . . . . .	223
Etapy a mechanizmy raného jazykového vývinu <i>Ján Rybár</i> . . . . .	231
Experimentální zkoumání kognitivních procesů pomocí původního webového nástroje MuTEP: zkoumání práce na mapách v kontextu osobnosti uživatele <i>Čeněk Šašínska, Zdeněk Stachoně</i> . . . . .	235
Kognitívny priestor <i>Ján Šefránek</i> . . . . .	243
Dopravní chování jako úkol pro agentní simulaci <i>Michal Šimeček</i> . . . . .	249
Prostor ve fotografii a v obrázku <i>Michal Šimeček</i> . . . . .	255
Socio-technická interakce: virtuální prostředí jako důsledek informatizace společnosti <i>Zdeněk Smutný</i> . . . . .	259

Konekcionalistický model epizodické pracovní paměti <i>Martin Takáč, Alistair Knott</i> . . . . .	265
Mobilní monitorovací systém pro seniory <i>Petr Tučník, Petr Šuba</i> . . . . .	273
Vliv kontextu na orientaci v prostoru <i>Michal Vavrečka, Jakub Kužilek, Lenka Lhotská</i> . . . . .	279
Brána vědění otevřena: nový pohled na výpočty <i>Jiří Wiedermann</i> . . . . .	285
Register autorů . . . . .	293



# Kombinácia SOM modelu neurónovej siete a vrstvy pre tréovanie s učiteľom

Gabriela Andrejková, Jozef Oravec

Ústav informatiky, Prírodovedecká fakulta UPJŠ v Košiciach

Jesenná 5, 041 54 Košice

Email: gabriela.andrejкова@upjs.sk, jozi.oravec@gmail.com

## Abstrakt

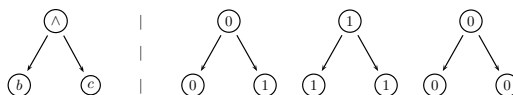
Použitie samoorganizujúcich sa modelov (SOM) neurónových sietí pri tréovaní dátových štruktúr je možné nájsť napríklad v [4] a v ďalších prácach. V prípade, že je potrebné okrem natréovanej štruktúry vypočítať aj hodnotu (charakterizujúcu túto štruktúru), napríklad u strojov pre aritmetické výrazy, zapamätanie štruktúry nepostačuje. V tomto príspevku je uvedené rozšírenie samoorganizujúcich sa neurónových sietí s laterálnymi váhami pre dátové štruktúry o možnosť tréovania ďalších hodnôt s dozorom. Navrhnutý model neurónovej siete bol otestovaný na acyklických dátových štruktúrach študijných programov a jeho vyhodnotenie pomocou viacerých kritérií ukazuje prijateľnú úspešnosť a zároveň poukazuje na vznikajúce problémy.

## 1 Úvod

V súvislosti s hľadaním modelu neurónovej siete, ktorú by bolo možné naučiť dátové štruktúry ako štruktúry zachovávajúce svoju podstatu, nás zaujal tento problém doplnený o možnosti prípadných výpočtov realizovaných pomocou týchto štruktúr. Ilustráciu urobíme na nasledujúcich dvoch príkladoch.

Na obrázku 1 je príklad, ako by mohli vyzeráť tréovacie vzory. Na ľavej strane je vstupná štruktúra  $D$  odpovedajúca výrazu  $b \wedge c$  a na pravej sú tri príklady výstupných štruktúr  $Y_1$ ,  $Y_2$  a  $Y_3$ , ktoré odpovedajú vyhodnoteniam pre konkrétne hodnoty premenných  $b$  a  $c$ . Z týchto štruktúr môžeme vytvoriť tréovacie dvojice, konkrétne v tomto prípade budú tri. Dvojice z tréovacej množiny majú kostry totožné. Predpokladaná štruktúra reprezentuje výraz, ktorého premenné sú v listoch a výstupná štruktúra reprezentuje postupný výpočet smerom ku koreňu. Cieľom je natréovať tieto dvojice a po natréovaní sa pýtať siete:

Ak na vstup siete vložíme číselné ohodnotenia odpovedajúce listom výstupnej štruktúry, aká bude vypočítaná hodnota výrazu odpovedajúceho podštruktúre ľubovoľného vrcholu vstupnej štruktúry, ktorý nie je list? Napríklad, na

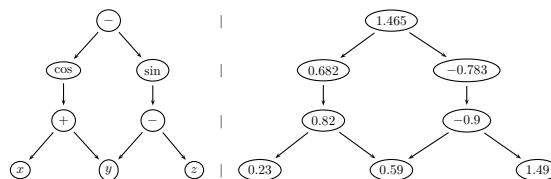


Obr. 1: Príklad tréovacej vzorky pre učenie štruktúry a nastavovaných hodnôt.

obrázku 1 nás môže zaujímať aká bude hodnota pre vrchol  $\wedge$ , ak hodnoty v listoch budú 1, 1.

Ďalším príkladom je výraz  $\cos(x + y) - \sin(x + z)$ , ktorý je reprezentovaný Orientovaným Acyklickým Grafom  $OAG$  zobrazeným na obrázku 2 v ľavej časti. Aj v tomto prípade na základe príkladov je cieľom natréovať výraz tak, aby neurónová sieť poskytla štruktúru a výpočet, pokiaľ jej dáme na vstup hodnoty v listoch. Očakávame, že nám vráti vypočítanú hodnotu pre ľubovoľný vrchol štruktúry (výrazu), ktorý nie je listom. Napríklad v tomto prípade, nastavíme hodnoty listov  $x, y, z$  a očakávame realizovanie výpočtu pomocou neurónovej siete pre výraz, ktorého supervrchol je "cos", "sin", "+", alebo "-".

Takúto sieť, ktorá je natréovaná na štruktúry spolu s výpočtami nimi realizovanými, nazveme sieť s laterálnymi váhami a učením s učiteľom (SLVT).



Obr. 2: Príklad tréovacej vzorky, kde vstupná štruktúra odpovedá výrazu  $\cos(x + y) - \sin(x + z)$  a ohodnotenia vrcholov výstupnej štruktúry odpovedajú hodnotám premenných alebo očakávaným vypočítaným hodnotám.

Modely neurónových sietí na tréovanie dátových štruktúr je možné nájsť napríklad v prácach [4], [5], [6], [7], [8], [11] a [12]. V článku nadväzujeme na práce [1], [2] a [3]. Všeobecné poznatky o neurónových sieťach boli čerpané z [9] a [10].

Štruktúra článku: V druhej kapitole je uvedený modi-

fikovaný model SOM siete, ktorý upravuje laterálne váhy spôsobom podporujúcim pamäť na grafových dátových štruktúrach, a zároveň podporuje učenie s dozorom. Tretia kapitola uvádza kritériá na vyhodnotenie tohto modelu siete. Štvrtá kapitola obsahuje popis použitia navrhnutého modelu na reálnych údajoch acyklických údajových štruktúr predstavujúcich prerekvizity študijných programov. V piatej kapitole je uvedené vyhodnotenie výsledkov pre danú aplikáciu. V závere je uvedená rekapitulácia dosiahnutých výsledkov a možnosti ďalšej práce.

## 2 Model modifikovanej neurónovej siete SOM a učenie s učiteľom (SLVT)

Na základe vyššie uvedených motivačných príkladov je treba si uvedomiť, že tréningová množina bude daná predpisom (1):

$$T^T = \{(\mathbf{D}, \mathbf{Y}); \mathbf{D} \in \mathcal{U}^\#, \mathbf{Y} \in \mathcal{Y}^\#\}, \quad (1)$$

kde  $\mathcal{U}^\#$  je vstupná doména štruktúr nad doménou  $\mathcal{U} \subseteq \mathcal{R}^n$  ( $\mathcal{R}$  je množina reálnych čísel),  $\mathcal{Y}^\#$  je výstupná doména štruktúr nad doménou  $\mathcal{Y} \subseteq \mathcal{R}^m$  a navyše pre každú dvojicu  $(\mathbf{D}, \mathbf{Y})$  z tréningovej množiny  $T^T$  platí, že  $skel(\mathbf{D}) = skel(\mathbf{Y})$ , t. j. kostry oboch štruktúr sú rovnaké. Budeme pracovať s dvojicami štruktúr, kde vstupná štruktúra predstavuje výraz zaznamenaný v podobe OAG grafu.

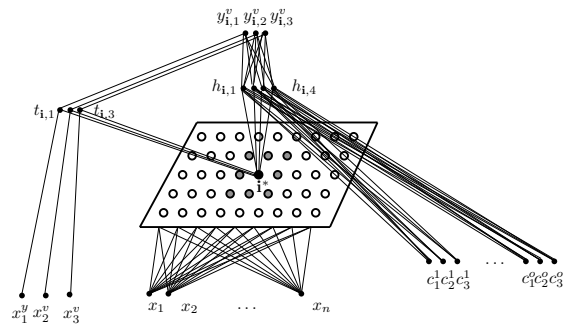
Množinu všetkých vrcholov v štruktúre  $\mathbf{D}$  budeme označovať  $V(\mathbf{D})$ , možinu listov budeme označovať  $L(\mathbf{D})$ ,  $ch_{\mathbf{D}}[v]$  je množina všetkých priamych potomkov a  $ch_{\mathbf{D},k}[v]$  je  $k$ -tý priamy potomok vrchola  $v$ . Podgraf (podštruktúru) grafu (štruktúru)  $\mathbf{G}$  indukovaný vrcholmi všetkých potomkov vrchola  $v$  označíme  ${}_v\mathbf{G}$ . Ohodnotenia listov budú odpovedať identifikátorom premenných a ohodnotenia ostatných vrcholov budú odpovedať identifikátorom  $|ch_{\mathbf{D}}[v]|$ -árnym operátorom, kde  $v \in V(\mathbf{D}) - L(\mathbf{D})$ . Listy výstupnej štruktúry budú obsahovať hodnoty z oboru hodnôt premenných, ktoré odpovedajú listom vstupnej štruktúry a ostatné vrcholy budú obsahovať hodnoty, ktoré sú výsledkom aplikácie  $|ch_{\mathbf{D}}[v]|$ -árneho operátora na predchodcov. Konkrétne, pre tréningovú vzorku  $(\mathbf{D}, \mathbf{Y})$  očakávame, že ak ohodnotením vrchola  $v$  štruktúry  $\mathbf{D}$  je  $|ch_{\mathbf{D}}[v]|$ -árny operátor  $f(., \dots, .)$ , tak pre ohodnotenie  $\mathbf{Y}_v$  vrchola  $v$  výstupnej štruktúry  $\mathbf{Y}$  bude platiť  $\mathbf{Y}_v = f(\mathbf{Y}_{ch_{\mathbf{D},1}[v]}, \dots, \mathbf{Y}_{ch_{\mathbf{D},k}[v]})$ , kde  $k = |ch_{\mathbf{D}}[v]|$ .

V texte sa bude vyskytovať rovnosť  $\mathbf{D}_v = \mathbf{Y}_v$  alebo iná rovnosť týkajúca sa "rovnakých vrcholov" oboch štruktúr  $\mathbf{D}$  a  $\mathbf{Y}$ . Slovné spojenie "rovnakým vrcholom" bude odpovedať izomorfnému zobrazeniu  $g : V(\mathbf{U}) \rightarrow V(\mathbf{Y})$  a v podstate pre každý vrchol  $v \in V(\mathbf{D})$  bude rovnosť  $\mathbf{D}_v = \mathbf{Y}_v$  odpovedať rovnosti  $\mathbf{D}_v = \mathbf{Y}_{g(v)}$ . Podobne to bude platiť pre iné typy nerovností.

Model neurónovej siete vychádza z modelov sietí študovaných v [1] a [3]. Základom je sieť, ktorá je tréningovaná na dátové štruktúry (reprezentované AOG), ktorá vznikla modifikáciou SOM siete s kontextovými vrstvami. Ku každému neurónu mriežky je pripojená skrytá vrstva, na ktorú sú pripojené výstupné neuróny. Ku skrytej vrstve je ešte pripojená kontextová vrstva, ktorá odpovedá aktivite výstupných neurónov odpovedajúcich potomkom spracúvaného vrcholu. Kontextová vrstva odpovedá čiastočným medzivýsledkom pri spracúvaní výpočtu podľa v sieti zapamätanej (interpretovanej) štruktúry získanej rekurzívnu aktiváciou. Sieť ešte obsahuje vrstvu vstupu hodnôt priradených listom štruktúry a transportnú vrstvu počítačnych hodnôt do výstupnej vrstvy.

### 2.1 Štruktúra siete SLVT

Sieť bude pozostávať z  $n$  vstupných neurónov  $x_i$ ,  $i = 1, \dots, n$ ,  $N$  neurónov usporiadaných v mriežke, ktorých stavy označíme  $y_i$ ,  $i = 1, \dots, N$ . Medzi vstupnou vrstvou a mriežkou sú váhy  $w_{ji}$ ,  $j = 1, \dots, n$ ,  $i = 1, \dots, N$  a medzi jednotlivými neurónmi mriežky sú laterálne váhy  $w_{ji}$ ,  $j = n + l$ ,  $i \neq l$ ,  $l = 1, \dots, N$ . Aby mohla neurónová sieť vypočítať výstup pre nejaký vrchol, ktorý nie je listom, pripojíme ku každému neurónu  $i$  v mriežke dve vrstvy - skrytú a výstupnú. Skrytá vrstva obsahuje  $k$  neurónov, ktorých výstupy označíme  $h_{i,j}$  a sú spojené s neurónom  $i$  mriežky váhami  $w_{i,j}^h$ , kde  $j = 1, \dots, k$ . Výstupná vrstva, obsahujúca  $m$  neurónov, ktorých každý výstupný neurón  $y_{i,l}^v$  je spojený váhami  $w_{i,j,l}^v$ ,  $j = 1, \dots, k$ ,  $l = 1, \dots, m$  s neurónmi skrytej vrstvy  $h_{i,j}$ . Výstupná vrstva bude slúžiť ako výstup vypočítaných hodnôt.



Obr. 3: Príklad siete SLVT.

Na obrázku 2 je príklad tréningovej vzorky. Vľavo je štruktúra vstupu odpovedajúca výrazu a vpravo je odpovedajúca štruktúra, ktorej ohodnotenia vrcholov nesú správne hodnoty medzivýsledkov pri spracovaní výrazu, resp. hodnoty premenných výrazu. Pri tréningu sa budú štruktúry spracovávať rekurzívne, t. j. keď sa bude spracovávať su-

pervrchol (vrchol, z ktorého nevedú žiadne hrany) ľavej štruktúry v tréningovom režime, je nurné mať vypočítané hodnoty pre vrcholy  $\cos$  a  $\sin$ . Tieto vypočítané hodnoty z predchádzajúceho kroku, podobne ako súradnice víťazných neurónov odpovedajúcich potomkom v SOM, budú vložené do vrstvy, ktorú nazveme kontextová vrstva hodnôt. Tá odpovedá výstupným hodnotám, ktoré sú ohodnotením potomkov spracovávaného vrcholu. V tejto vrstve bude  $o * m$  neurónov označených  $c_j^l$ , kde  $j = 1, \dots, m$  a  $l = 1, \dots, o$ . Neuróny kontextovej vrstvy budú spojené váhami  $w_{i,j}^{h,l}$  s neurónmi skrytej vrstvy, ktoré sú pripojené k ľubovoľnému neurónu  $i$  mriežky, kde  $h = 1, \dots, k$ ,  $l = 1, \dots, o$  a  $j = 1, \dots, m$ .

Počas tréningovania aj počítania je nutné nastaviť hodnoty, ktoré sú priradené listom štruktúr. Tieto hodnoty postupne budeme vkládať cez vrstvu, ktorá obsahuje neuróny  $x_1^y, \dots, x_m^y$  a je spojená s vrstvou hradíel  $t_{i,1}, \dots, t_{i,m}$  s váhami  $w_{ji}^t = 1$ , kde  $j = 1, \dots, m$ . Úlohou vrstvy hradíel je prideliť hodnoty listov do výstupnej vrstvy pre  $i$ -tý neurón mriežky, ktorý odpovedá premennej výrazu odpovedajúcemu vstupnej štruktúre. Vrstva hradíel bude spojená s výstupnými neurónmi  $y_{i,1}, \dots, y_{i,m}$  váhami  $w_{ij}^y = 1$ , kde  $j = 1, \dots, m$ .

Na obrázku 3 je príklad SOM siete s laterálnymi váhami, ktorá má kontext hodnôt a je tréningovaná s učiteľom SLVT. Pre názornosť je zobrazená skrytá vrstva pre víťazný neurón  $i^*$  obsahujúca neuróny  $h_{i^*,1}, \dots, h_{i^*,4}$  a výstupnú vrstvu obsahujúcu neuróny  $y_{i^*,1}^v, y_{i^*,2}^v$  a  $y_{i^*,3}^v$ . Kontextová vrstva hodnôt má 30 neurónov. Hodnoty v listoch budú vkladané do siete pomocou neurónov  $x_1^y, x_2^y$  a  $x_3^y$ . Táto sieť môže byť tréningovaná pomocou tréningovej množiny vzoriek uvedených na obrázkoch 1 a 2.

## 2.2 Práca siete SLVT

Jednou z úloh v sieťach SLVT je vypočítať hodnoty výrazov, na ktoré sú tréningované (ktoré sú uvedené na ľavej strane prezentovaných príkladov). Teda, vstupná štruktúra je výraz a výstupná štruktúra s ohodnoteniami vrcholov obsahuje postupný výpočet výrazu, ktorý je už získaný neurónovou sieťou. Na vstup siete pri tréningovaní dávame štruktúru, ktorá odpovedá nejakému tréningovému výrazu alebo podvýrazu a hodnoty, ktoré odpovedajú listom vstupnej štruktúry, teda premenným výrazu. Po natréningovaní očakávame, že sieť nebude počítat hodnotu výrazu, ktorý je na vstupe, ale natréningovaného výrazu, pomocou ktorého je vykonaný výpočet rekurzívne. Tento výpočet je realizovateľný pomocou nasledujúcej transdukcie

$$\mathcal{M}_T : \mathcal{U}^\# \times (\mathcal{U} \times \mathcal{Y})^p \rightarrow \mathcal{Y}^\#, \quad (2)$$

kde  $\mathcal{U}$  je doména, nad ktorou je vstupná doména štruktúr  $\mathcal{U}^\#$  a  $\mathcal{Y}$  sú hodnoty listov z domény, nad ktorou je výstupná

doména štruktúr. Vstupom bude postupne priestor označení vrcholov vstupnej štruktúry a  $p$  dvojíc identifikátora vrchola listu a hodnoty odpovedajúcej listu. Bude sa vykonávať výpočet

$$\mathcal{M}_T (\mathbf{D}, \mathbf{D}_{i_1}, \mathbf{x}_{i_1}^v, \dots, \mathbf{D}_{i_p}, \mathbf{x}_{i_p}^v) = \mathbf{Y}, \quad (3)$$

kde  $\mathbf{D}$  vstupná štruktúra,  $\mathbf{D}_{i_j}$  je ohodnotenie vrchola odpovedajúce listu  $i_j$  a  $\mathbf{x}_{i_j}$  je hodnota, ktorá odpovedá listu v našom prípade premennej výrazu. Výsledná hodnota bude v ohodnotení supervrcholu výstupnej štruktúry  $\mathbf{Y}$ .

Výpočet rozdelíme do nasledovných krokov: (1) nastavenie vstupných hodnôt neurónom, ktoré odpovedajú vrcholom listov vstupnej štruktúry, (2) interpretácia štruktúry časťou neurónovej siete pre tréningovanie štruktúr, pomocou ktorej sa bude vykonávať rekurzívne výpočet a (3) rekurzívny výpočet výstupných hodnôt, čím získame výstupnú štruktúru spolu s čiastočnými výpočtami v ohodnoteniach vrcholov a v ohodnotení supervrchola hodnotu výrazu podľa neurónovej siete.

V prvom kroku je určený víťazný neurón vstupnej štruktúry na mriežke pomocou zobrazenia

$$\mathcal{M}_{TV}^\# : \mathcal{U} \rightarrow \mathcal{A}, \quad (4)$$

kde  $\mathcal{A}$  je priestor odpovedajúci súradniciam neurónov umiestnených v mriežke. Výpočet je vykonaný nasledovne

$$\mathcal{M}_{TV}^\# (\mathbf{D}_v) = \arg \min_{j=1, \dots, N} (\|\mathbf{D}_v - \mathbf{w}_j^I(t)\|) = \mathbf{i}_{\mathbf{D}_v}^*. \quad (5)$$

Takto je získaný víťazný neurón odpovedajúci ohodnoteniu vrchola  $v$  vstupnej štruktúry  $\mathbf{D}$ . V prípade, že vrchol na vstupe je list, je potrebné nastaviť stav neurónov vo výstupnej vrstve odpovedajúci listu, a to nasledujúcim zobrazením

$$\mathcal{M}_{TL}^\# : \mathcal{U} \times \mathcal{Y} \rightarrow \mathcal{Y}. \quad (6)$$

Výpočet  $\mathcal{M}_{TL}^\#$  je urobený nasledovne

$$\mathcal{M}_{TL}^\# (\mathbf{D}_{i_l}, \mathbf{x}_{i_l}) = (\mathbf{x}_{i_l}^v)^T \mathbf{t} = \mathbf{y}_{i_l}^v, \quad (7)$$

kde  $i_l^* = \mathcal{M}_{TV} (\mathbf{D}_{i_l})$ ,  $\mathbf{D}_{i_j}$  je ohodnotenie listu  $i_j$  vstupnej štruktúry  $\mathbf{D}$  a  $\mathbf{x}_{i_j}$  je hodnota listu. Takto pre všetky listy  $i_j, j = 1, \dots, p$  sú nastavené aktivity výstupných neurónov odpovedajúcich víťazným neurónom  $i_j^*$ .

Predtým, než sa začne vykonávať samotný rekurzívny výpočet výstupných hodnôt, je potrebné získať štruktúru  $\mathbf{Z}$  pomocou transdukcie nasledovne:  $\mathcal{M}_L (\mathbf{D}) = \mathbf{Z}$ . štruktúra  $\mathbf{Z} \in \mathcal{I}^\#$ , kde  $\mathcal{I} = \mathcal{R} \times \mathcal{A}$ ,  $\mathcal{R}$  je množina reálnych čísel, nesie informáciu v ohodnoteniach vrcholov aj o súradnici odpovedajúcej neurónu vrchola štruktúry  $\mathbf{Z}$  a táto informácia nám posluží na aktiváciu stavov neurónov kontextovej vrstvy hodnôt. Konkrétne pre vrchol  $v \in V(\mathbf{Z})$  bude jeho ohodnotenie  $\mathbf{Z}_v = (u_v, \mathbf{i}_v)$ .

Samotný výpočet je rekurzívny výpočet hodnôt medzi výsledkov pri spracúvaní výrazu zapamätaného ako štruktúra  $\mathbf{Z}$ . V (9) je uvedený rekurzívny výpočet postupujúci podľa štruktúry  $\mathbf{Z}$ , ktorý vypočíta výstupnú hodnotu odpovedajúcu supervrcholu  $s_{\mathbf{Y}}$  štruktúry  $\mathbf{Y}$  nasledovne

$$\mathcal{M}_{node}^{\#YR} : \mathcal{I}^{\#} \rightarrow \mathcal{Y}, \quad (8)$$

kde  $\mathcal{I}^{\#}$  je doména štruktúr získaných rekurzívnou aktiváciou neurónov a  $\mathcal{Y}$  doména, nad ktorou budú ohodnotenia vrcholov výstupnej štruktúry. Výpočet ohodnotenia vrcholov výstupnej štruktúry získame rekurzívnym prechádzaním štruktúry  $\mathbf{Z}$  nasledovne:

$$\mathcal{M}_{node}^{\#YR}(\mathbf{Z}) = \begin{cases} \mathbf{0}, & \mathbf{Z} = \xi, \\ \mathbf{y}_{s_{\mathbf{Z}}}, & |V(\mathbf{Z})| = 1, \\ \mathcal{M}_{node}^{\#YR}, & \text{inak.} \end{cases} \quad (9)$$

kde

$$\mathcal{M}_{node}^{\#YR} = \left( s_{\mathbf{Z}}, \mathcal{M}_{node}^{\#YR}(ch_{1|s_{\mathbf{Z}}}(\mathbf{Z})), \dots, \mathcal{M}_{node}^{\#YR}(ch_{o|s_{\mathbf{Z}}}(\mathbf{Z})) \right)$$

Výsledkom výpočtu je ohodnotenie supervrchola  $\mathbf{Y}$ , t.j.  $\mathbf{Y}_{s_{\mathbf{Y}}} = \mathcal{M}_{node}^{\#YR}(\mathbf{Z})$ . Samotný výpočet funkcie (9) je rozdelený do troch podmienok. Prvou je, ak dáme na vstup prázdnu štruktúru, dostaneme nulový vektor. Tejto situácii odpovedajú chýbajúci potomkovia. Druhá možnosť nastáva, ak štruktúra má len jeden vrchol a tomu odpovedá situácia, keď je spracovávaný list štruktúry. Posledný prípad odpovedá rekurzívnemu spracovaniu výpočtu štruktúry  $\mathbf{Z}$ , ktoré je dané zobrazením

$$\mathcal{M}_{node}^{\#YR} : \mathcal{I} \times \underbrace{\mathcal{Y} \times \dots \times \mathcal{Y}}_{o \text{ krát}} \rightarrow \mathcal{Y} \quad (10)$$

kde  $\mathcal{I}$  je doména, nad ktorou je doména obsahujúca štruktúry získané rekurzívnou aktiváciou a  $\mathcal{Y} \times \dots \times \mathcal{Y}$  sú výstupné hodnoty získané pre podštruktúry vrchola budovanej výstupnej štruktúry. Výstupom je aktivita výstupných neurónov k spracovávanému neurónu, ktorý odpovedá vrcholu zapamätanej štruktúry. Výpočet je daný v (11)

$$\mathcal{M}_{node}^{\#YR}(u, \mathbf{i}, \mathbf{y}_{i_1}^v, \dots, \mathbf{y}_{i_o}^v) = \mathbf{y}_i = (y_{i_1}, \dots, y_{i_m}), \quad (11)$$

a je vykonaný v nasledujúcich dvoch krokoch. Prvým krokom je výpočet aktivity skrytých neurónov odpovedajúcich i-tému spracovávanému vrcholu nasledovne

$$h_{i,d} = \sigma(\xi_{i,d}) = \sigma\left(w_{i,d}^h y_i + \sum_{l=1}^o \sum_{j=1}^m w_{i,j}^{d,l} c_j^l\right), \quad (12)$$

kde  $d = 1, \dots, k$  a  $c_j^d$  je aktivita neurónu v kontextovej vrstve danej nasledovne  $c_j^d = y_{i_l}^v$ . Výstup pre neurón mriežky  $\mathbf{i}$  je vypočítaný pomocou (13).

$$y_{i,l} = \sigma(\xi_{i,l}^v) = \sigma\left(\sum_{j=1}^k w_{i,j}^v h_{i,j}\right), \quad (13)$$

kde  $l = 1, \dots, m$ . V oboch predchádzajúcich prípadoch je  $\sigma$  aktivačná funkcia a bola použitá sigmoidálna funkcia  $\sigma(x) = \frac{1}{1+e^{-x}}$ . Treba zdôrazniť, že tento výpočet sa vykonáva rekurzívne, a teda musíme vrcholy z množiny  $V(\mathbf{Z})$  inverzne usporiadať a až potom začať realizovať výpočet.

Pomocou funkcie (9) získame výstupnú hodnotu odpovedajúcu supervrcholu výstupnej štruktúry. Konkrétne,  $\mathcal{M}_{node}^{\#YR}(\mathbf{Z}) = \mathbf{Y}_{s_{\mathbf{Y}}}$ . Pre ostatné vrcholy  $v \in V(\mathbf{Z}) - \{s_{\mathbf{Z}}\}$  bude ohodnotenie vrcholov  $\mathbf{Y}_v = \mathcal{M}_{node}^{\#YR}(v, \mathbf{Z})$ . Takto dostaneme ohodnotenie vrcholov výstupnej štruktúry  $\mathbf{Y}$ . Štruktúry  $\mathbf{Z}$  a  $\mathbf{Y}$  majú rovnaké kostry. Formálny postup výpočtu v aktívnom režime je zapísaný v algoritme 1, štruktúra  $\mathbf{Y}$  je výstupnou z transdukcie  $\mathcal{M}_T$  a ohodnotenie supervrchola je výsledok výpočtu výrazu v neurónovej sieti.

### Algoritmus 1: Trénovanie siete SLVT

**Vstup:** Vstupná štruktúra  $\mathbf{D}$ ,  $p = |L(\mathbf{D})|$ , dvojice  $L = \left\{ (\mathbf{D}_{i_1}, \mathbf{x}_{i_1}^v), \dots, (\mathbf{D}_{i_p}, \mathbf{x}_{i_p}^v) \right\}$ , kde  $l = 1, \dots, p$ .

**Výstup:** Štruktúra  $\mathbf{Y}$ .

**Metóda:**

**Inicializácia:** Stav všetkých výstupných neurónov  $y_{i,l}, i = 1, \dots, N, l = 1, \dots, m$  nastaviť na 0.

**K1:** Pre každú dvojicu  $(\mathbf{D}_{i_l}, \mathbf{x}_{i_l}^v) \in L$  nastaviť výstupné neuróny  $y_{i_l,k}^v, k = 1, \dots, m$  podľa (7).

**K2:** Pomocou algoritmu pre trénovanie štruktúr [2] získať štruktúru  $\mathbf{Z}$  pomocou rekurzívnej aktivácie neurónov, pričom vstupom do tohto algoritmu je štruktúra  $\mathbf{D}$ .

**K3:** Pre každý vrchol  $v \in V_{inv}(\mathbf{Z})$ , ktorého ohodnotenie je  $\mathbf{Z}_v = (u_v, \mathbf{i}_v)$ , vytvoriť štruktúru  $\mathbf{Y}$  bez ohodnotenia a vykonať postupný výpočet nasledovne:

K3.1: Stav neurónov v kontextovej vrstve hodnôt nastaviť na 0.

K3.2: Do kontextovej vrstvy hodnôt vložiť hodnoty výstupnej vrstvy odpovedajúcej potomkom. Konkrétne, pre  $u = ch_{\mathbf{Z},k}[v]$ , kde  $k$  je najvyššia, bude  $c_j^k = y_{i_u,j}^v$ .

K3.3: Podľa (9) vypočítať ohodnotenia vrcholu  $\mathbf{Y}_v$ .

K3.4: Ak  $v = s_{\mathbf{X}}$ , skončiť, inak pokračovať na K3.1.

Algoritmus 1 pre trénovanie siete SLVT, ktorý používa trénovanie štruktúr, nemusí interpretovať štruktúru s presnou topológiou ako sme očakávali, a tu vzniká

problém. Pri rekurzívnom výpočte výstupných hodnôt je používaná štruktúra  $\mathbf{Z}$ , ktorá je interpretáciou pôvodného výrazu a nevieme zaručiť, že počet potomkov nejakého vrchola takto získanej štruktúry bude väčší než najväčší počet potomkov vrchola trénovacej množiny. Je to otvorený problém. V algoritme 1 sme sa obmedzili na počet potomkov tak, aby neprekročil hodnotu  $o$ .

**Trénovaním** upravíme konfiguráciu siete tak, aby sa hodnota chybovej funkcie siete znižovala. Vzniká otázka, ako definovať chybovú funkciu, pretože transdukcija (2) nie je  $IO$ -izomorfná. Musíme zvoliť počas trénovania alternatívnu cestu. Najprv natrénujeme mriežku odpovedajúcu sieti pre trénovanie štruktúr [2] a až na základe ňou interpretovaných štruktúr budeme adaptovať váhy medzi kontextovou vrstvou hodnôt a skrytou vrstvou, resp. medzi skrytou a výstupnou vrstvou. Pri trénovaní sme zo štruktúry  $\mathbf{D}$  získali najprv interpretovanú štruktúru  $\mathbf{Z}$  a z nej sme získali štruktúru  $\mathbf{Y}$ . Poznamenajme, že toto zobrazenie odpovedá  $\mathcal{M}_{node}^{\#YR}$   $IO$ -izomorfnéj transdukcii a môžeme na ňu aplikovať chybovú funkciu (14).

$$E = \sum_{(\mathbf{U}, \mathbf{Y}) \in \mathcal{T}} E(\mathbf{U}, \mathbf{Y}) = \sum_{(\mathbf{U}, \mathbf{Y}) \in \mathcal{T}} \sum_{v \in V(\mathbf{U})} \|\tau(\mathbf{U})_v - \mathbf{Y}_v\|^2. \quad (14)$$

Analizujme detailnejšie druhú fázu, v ktorej sa bude vykonávať adaptácia váh medzi kontextovou vrstvou hodnôt, skrytou a výstupnou vrstvou. Vychádzajúc z trénovacej množiny, nemôžeme vykonať priame priradenie  $\mathbf{U}_v = \mathbf{Y}_{\mathcal{M}^T(v)}$ , kde  $\forall v \in V(\mathbf{U})$ , môžeme vykonať iba priradenie  $\mathbf{Z}_v = \mathbf{Y}_{\mathcal{M}_{node}^{\#YR}(v)}$ , kde  $\forall v \in V(\mathbf{Z})$ , lebo štruktúry  $\mathbf{U}$  a  $\mathbf{Y}$  nemajú rovnaké kostry. S touto skutočnosťou sa vysporiadame tak, že využijeme to, že supervrchol vstupnej štruktúry  $\mathbf{U}$ , štruktúry  $\mathbf{Z}$  interpretovanej časťou siete pre trénovanie štruktúr a výstupnej štruktúry  $\mathbf{Y}$  je možné priradiť cez transdukciju  $\mathcal{M}_L$  a  $\mathcal{M}_{node}^{\#YR}$ . O ostatných vrcholoch týchto štruktúr nemôžeme toto tvrdiť, lebo  $\mathcal{M}_L$  nie je  $IO$ -izomorfná transdukcija. Nasledujúci výraz

$$\begin{aligned} \tau\left((v\mathbf{U})_{s_{v\mathbf{U}}}\right) &= \mathcal{M}_{node}^{\#YR}\left(\mathcal{M}_T\left((v\mathbf{U})_{s_{v\mathbf{U}}}\right)\right) = \\ &= \mathcal{M}_{node}^{\#YR} = (v\mathbf{Z})_{s_{v\mathbf{Z}}} = \mathbf{Y}_{s_{v\mathbf{Z}}} \end{aligned} \quad (15)$$

je zhrnutím vyššie popísanej skutočnosti za predpokladu, že pracuje so supervrcholmi týchto troch štruktúr. Inými slovami vo vzťahu (15) pracujeme so supervrcholmi odpovedajúcimi podštruktúram  $v\mathbf{U}$  štruktúry  $\mathbf{U}$  a túto výhodu využijeme pri budovaní modifikovanej trénovacej množiny. Na začiatku výrazu je zobrazenie  $\tau : \mathbf{R}^n \rightarrow \mathbf{R}^m$ , ktoré je uvedené pre zjednodušenie odvodenia adaptačnej dynamiky, ktorú zhrnieme do nasledujúceho postupu:

Z množiny trénovacích vzorov  $\mathcal{T}^T$  vytvoríme modifikovanú trénovaciu množinu

$$\mathcal{T}_U^T = \{(\mathbf{U}, \mathbf{Y}); \forall (\mathbf{A}, \mathbf{B}) \in \mathcal{T}^T, \forall v \in V(\mathbf{A}) - L(\mathbf{A}),$$

$$\mathbf{U} = {}_v \mathbf{A}, \mathbf{Y} = {}_v \mathbf{B}\}. \quad (16)$$

Množina  $\mathcal{T}_U^T$  bude obsahovať dvojice štruktúr, ktorých supervrcholy sú jednotlivé vrcholy vstupných a výstupných štruktúr okrem ich listov. Takto zabezpečíme postupnú adaptáciu pre všetky vrcholy štruktúr v trénovacej množine  $\mathcal{T}^T$  a nie len pre ich supervrcholy.

Vychádzajúc z vyššie uvedených skutočností definujeme chybovú funkciu siete nasledovne

$$E = \sum_{(\mathbf{U}, \mathbf{V}) \in \mathcal{T}_U^T} \|\tau(\mathbf{U}_{s_{\mathbf{U}}}) - \mathbf{V}_{s_{\mathbf{V}}}\|^2 \quad (17)$$

$$= \frac{1}{2} \sum_{(\mathbf{U}, \mathbf{V}) \in \mathcal{T}_U^T} \sum_{l=1}^m \left(y_{i,l}^v - (\mathbf{V}_{s_{\mathbf{V}}})_l\right)^2, \quad (18)$$

kde v (17) pre jednoduchosť ďalších výpočtov označme  $E_s = \sum_{l=1}^m \left(y_{i,l}^v - (\mathbf{V}_{s_{\mathbf{V}}})_l\right)^2$ . Pre výpočet zníženia chyby potrebujeme určiť parciálne derivácie podľa jednotlivých váh  $w_{i,j,l}^v$ ,  $w_{i,j}^{d,l}$  a  $w_{i,j}^d$ , lebo tie sa budú meniť v druhej fáze adaptačného režimu.

Výpočet aktivity skrytých neurónov odpovedajúcich i-tému spracovanému vrcholu je nasledovný

$$h_{i,d} = \sigma(\xi_{i,d}) = \sigma\left(w_{i,d}^h y_i + \sum_{l=1}^o \sum_{j=1}^m w_{i,j}^{d,l} c_j^l\right) \quad (19)$$

kde  $d = 1, \dots, k$  a  $c_j^d$  je aktivita neurónu v kontextovej vrstve danej nasledovne  $c_j^d = y_{i,j}^v$ .

Po odvodení potrebných vzťahov celý proces adaptácie zhrnieme v nasledujúcom algoritme 2.

**Algoritmus 2:** Adaptácia siete SLVT.

**Vstup:** Trénovacia množina  $\mathcal{T}^T$  definovaná v (1).

**Výstup:** Konfigurácia siete SLVT.

**Metóda:**

K1: Vytvorí množinu  $\mathcal{T}_U^T$  danú predpisom (16).

K2: Inicializovať a natrénovať časť siete pomocou algoritmu 1. Trénovacia množina bude pozostávať zo vstupných štruktúr množiny  $\mathcal{T}^T$ .

K3: Inicializovať náhodne váhy  $w_{i,j,l}^v$ ,  $w_{i,j}^{d,l}$  a  $w_{i,j}^d$ , kde  $i = 1, \dots, N$ ,  $j = 1, \dots, o$ ,  $l = 1, \dots, m$  a  $d = 1, \dots, k$ . Inicializovať váhy  $w_{ji}^t$  a  $w_{ij}^v$  na hodnotu 1, kde  $i = 1, \dots, N$  a  $j = 1, \dots, m$ .

K4: Opakovať pre každú dvojicu štruktúr  $(\mathbf{U}, \mathbf{V}) \in \mathcal{T}_U^T$ , pokiaľ sieť neskonverguje do stabilného stavu nasledujúce kroky:

K4.1: Vypočítať výstupné ohodnotenia výstupnej štruktúry  $\mathbf{Y}$  k vstupnej štruktúre  $\mathbf{U}$  s hodnotami listov zo štruktúry  $\mathbf{Y}$  podľa algoritmu 1.

K4.2: Najší víťaza  $i$  pre vrchol  $s_U$  podľa (5).

K4.3: Výstupnými hodnotami z neurónov odpovedajúcim potomkom supervrchola  $s_U$  aktivovať neuróny kontextovej vrstvy hodnôt, resp. ak chýba potomok, vložiť vektor  $\mathbf{0}$ .

K4.4: Vypočítať aktivitu neurónov v skrytej vrstve pomocou (19) a výstupnej vrstve pomocou (13) odpovedajúcej neurónu  $i$ .

K4.5: Pomocou gradientov určených v odvodených vzťahoch vypočítať zmenu váh.

Treba poznamenať, že pri tréovaní nás zaujíma len výpočet pre supervrchol vstupnej, a teda aj výstupnej štruktúry. Tým, že sme vytvorili modifikovanú tréovaciu množinu  $T_U^T$ , dostali sme dvojicu, v ktorých supervrcholy odpovedajú všetkým "nelistom" z pôvodnej tréovacej množiny  $T^T$ , a tak zabezpečíme natréovanie postupne celých výrazov.

**Priestorová zložitosť** samoorganizujúcej sa siete s laterálnymi váhami pre tréovanie s učiteľom je

$$lUS = O(n + N + (n + N)N + 2Nk + nN + kmN + mo + mokN + 2m + 2mN), \quad (20)$$

kde priestorová zložitosť časti siete pre tréovanie štruktúr je  $O(n + N + (n + N)N)$ , počet skrytých neurónov mriežky je  $Nk$  a počet váh medzi mriežkou a skrytými neurónmi je  $Nk$ . Počet výstupných neurónov je  $mN$  a počet váh medzi skrytou a výstupnou vrstvou je  $kmN$ . Počet neurónov v kontextovej vrstve hodnôt je  $mo$ , počet váh medzi skrytou a kontextovou vrstvou hodnôt je  $mokN$ , počet neurónov vrstvy rozdeľujúcej hodnoty listov je  $m$ , počet neurónov vstupnej vrstvy pre hodnoty listov je  $m$  a počet váh medzi vrstvou rozdeľujúcej hodnoty, vstupnou vrstvou pre hodnoty listov a výstupnou vrstvou je  $2mN$ .

**Časová zložitosť** adaptácie jednej tréovacej vzorky za predpokladu, že časť siete pre tréovanie štruktúr, je

$$T = O((om + k(1 + om) + mk) |V(\mathbf{Z}) - L(\mathbf{Z})| + nN + mo + k(1 + om + m) + km(1 + o)) \quad (21)$$

kde je treba  $(om + k(1 + om) + mk) |V(\mathbf{Z}) - L(\mathbf{Z})|$  operácií na vytvorenie výstupnej štruktúry  $\mathbf{Y}$  zo štruktúry  $\mathbf{Z}$ , pre nájdenie víťazného neurónu pre supervrchol je potrebných najviac  $nN$  operácií, pre nastavenie aktivity kontextových neurónov je potrebných  $mo$  operácií, pre vypočítanie aktivity neurónov v skrytej a vo výstupnej vrstve zaberie  $k(1 + om + m)$  operácií a na zmenu váh bude potrebných  $km(1 + o)$  operácií.

### 3 Kritéria na vyhodnotenie výsledkov

Na vyhodnotenie siete SLVT ola pripravená množina overovacích vzorov  $()$ .

$$T^O = \{(\mathbf{U}, \mathbf{Y}) ; \mathbf{U} \in \mathcal{U}^\#, \mathbf{Y} \in \mathcal{Y}^\#\}, \quad (22)$$

Štruktúra  $\mathbf{U}$  je vstupná a  $\mathbf{Y}$ . Množinu  $()$  nazveme **overovacia**.

#### 3.1 Diferenciácia víťazov

Vkladaním ohodnotení vrcholov štruktúr z overovacej množiny na vstup sieťam a následným výpočtom siete získame víťazný neurón. Postupným vložením všetkých ohodnotení všetkých vrcholov z overovacej množiny získame množinu víťazných neurónov, ktoré odpovedajú vrcholom štruktúry. Pomer medzi počtom víťazných neurónov a počtom vrcholov v štruktúrach overovacej množiny sa bude nazývať **diferenciácia víťazov (winner differentiation WD)** [12] a bude určená vzťahom

$$WD = \frac{|\{j; j = i^*(\cdot), v \in V(\mathbf{U}), \mathbf{U} \in T\}|}{N}, \quad (23)$$

kde  $i^*(\cdot)$  je funkcia, ktorá určí víťazný neurón pre ohodnotenie vrchola  $\mathbf{U}_v$ . Zrejme, ak  $WD < 1$ , víťazi nie sú priradení jednoznačne pre jednotlivé vstupy - ohodnotenia vrcholov. Očakáva sa, že daný parameter je čo najväčší.

#### 3.2 Vyhodnotenie kvality množiny potomkov

Pod **interpretáciou vstupnej štruktúry pomocou NS** (vstupných dát) rozumieme štruktúru, ktorú natréovaná sieť vypočíta (výstupná štruktúra) k daným vstupným údajom.

Dôležité je vyhodnotenie kvality natréovania štruktúr z pohľadu vymenovania potomkov vrchola natréovanej štruktúry. Bude nás zaujímať správnosť aktivácie neurónov počas rekurzívnej aktivácie. Pred zadaním ohodnotenia vrcholu na vstup, je stav neurónov na mriežke, resp. výstupnej vrstve nastavený na 0, t.j. neaktívny. Po vykonaní výpočtu budú stavy niektorých neurónov nastavené na 1-aktívny, t.j. neurón, ktorý je aktívny odpovedá vrcholu na vstupe alebo jednému z jeho všetkých potomkov. Vymenované vrcholy sú určené sieťou a nemusia korešpondovať so skutočnosťou. Pod **správnou aktiváciou neurónu** budeme rozumieť, že po výpočte je stav neurónu 1 a očakáva sa, že stav má byť 1 a pod **nesprávnou aktiváciou neurónu** budeme rozumieť, že jeho stav je iný než sme očakávali.

Z pohľadu interpretácie štruktúry budeme rozlišovať dva prístupy. Prvý je vymenovanie všetkých potomkov

vrchola štruktúry vrátane jeho samého a neočakávame, že sieť reprezentuje topológiu podštruktúry. Druhým spôsobom bude interpretácia podštruktúry neurónovou sieťou. V tomto prípade získame nielen vymenovanie všetkých potomkov, ale aj topologické usporiadanie podštruktúry danou sieťou.

Budeme vyhodnocovať dôveryhodnosť interpretácie štruktúr t.j pomer medzi správne aktivovanými neurónmi a počtom vrcholov štruktúry.

V overovacej množine (3) sú uvedené vstupné štruktúry, ktorých ohodnotenia vrcholov sú na vstupe. Toto ohodnotenie tvorí identifikátor vrcholov. Dôveryhodnosť sa určí porovnaním pomeru počtu správne aktivovaných neurónov a počtu očakávaných vrcholov. Počet očakávaných potomkov vrcholu  $v$  je  $|de_{\mathbf{U}}[v] \cup \{v\}|$ , resp.  $|de_{l,\mathbf{U}}[v] \cup \{v\}|$  v úrovni  $l$ , ktorého ohodnotenie  $\mathbf{U}_v$  je na vstupe a  $\mathbf{U}$  je dátová štruktúra z množiny štruktúr  $\mathcal{T}_B^O$  alebo  $\mathcal{T}^O$ . Potom **dôveryhodnosť**  $SA$  bude daná všeobecným vzťahom:

$$SA = \frac{p^{SA}}{p^S}, \quad (24)$$

kde  $p^S$  ako súčet počtu všetkých správne vymenovaných všetkých potomkov vrcholov z overovacej množiny na vstupe vrátane daného vrcholu a  $p^{SA}$  je súčet počtu všetkých správne aktivovaných neurónov v mriežke počas rekurzívnej aktivácie siete pri postupnom vložení všetkých ohodnotení vrcholov vstupných štruktúr z overovacej množiny na vstup siete.

### 3.3 Vyhodnotenie kvality výpočtu pri tréovaní siete s laterálnymi váhami a učením s učiteľom

V tréovacej množine (1) pre sieť SLVT sú vstupné a odpovedajúce očakávané výstupné štruktúry. Vychádzajúc z tejto tréovacej množiny, vytvoríme overovaciu množinu, v ktorej z každej dvojice kostra vstupnej štruktúry a výstupnej štruktúry bude rovnaká. Overovacia množina  $\mathcal{T}^O$ , daná predpisom (3), bude vytváraná tak, že vstupná štruktúra bude odpovedať výrazu, ktorý reprezentuje nejaká vstupná štruktúra z tréovacej množiny a výstupná štruktúra bude vždy iná než z tréovacej množiny, pričom ohodnotenia listov bude mať vypočítané podľa výrazu vstupnej štruktúry a priradených hodnôt premenných odpovedajúcich listom.

V prípade sietí sa budeme vyhodnocovať rozdiel medzi skutočným ohodnotením vrchola výstupnej štruktúry a vypočítaným ohodnotením neurónovou sieťou. Pre potreby porovnávania zavedieme tri typy chýb, ktoré sú bežne používané:

**M.E.** - Mean Error - stredná chyba odhadu

$$M.E. = \frac{\sum_{\mathbf{U} \in \mathcal{T}^O} \left( \sum_{v \in V(\mathbf{U})-L(\mathbf{U})} (\mathbf{Y}_v - \mathbf{U}_v) \right)}{N^{NL}}, \quad (25)$$

**M.A.E.** - Mean Absolute Error - stredná absolútna chyba odhadu

$$M.A.E. = \frac{\sum_{\mathbf{U} \in \mathcal{T}^O} \left( \sum_{v \in V(\mathbf{U})-L(\mathbf{U})} |\mathbf{Y}_v - \mathbf{U}_v| \right)}{N^{NL}}. \quad (26)$$

**M.S.E.** - Mean Squared Error - stredná kvadratická chyba odhadu

$$M.S.E. = \frac{\sum_{\mathbf{U} \in \mathcal{T}^O} \left( \sum_{v \in V(\mathbf{U})-L(\mathbf{U})} (\mathbf{Y}_v - \mathbf{U}_v)^2 \right)}{N^{NL}}, \quad (27)$$

Poznamenajme, že rozdiel  $\mathbf{Y}_v - \mathbf{U}_v$  je písaný v zmysle intuitívneho významu "rovnakých vrcholov" uvedeného v odseku 2.

Využívajúc vyššie uvedené chyby definujeme chyby postupného spracovávania štruktúr od listov k supervrcholu štruktúr. Bude nás zaujímať, ako dobre vypočíta výstup neurónová sieť pri spracovaní vrcholov podľa vzdialenosti od najvzdialenejšieho listu. Pre každú štruktúru definujeme vyššie spomenuté chyby podľa úrovni  $l = 0, \dots, h-1$  postupne od  $h-1$  úrovne pri spracovaní štruktúry obsahujúca aspoň jeden vrchol, ktorý nie je listom až po úroveň  $l = 0$ , ktorá obsahuje supervrchol u štruktúry, ktorej najväčšia vzdialenosť medzi dvoma vrcholmi je  $h$ .

Chyby v úrovniach  $l = 0, \dots, h-1$  uvádzané nižšie postupne vychádzajú z (25), (26) a (27):

**M.E.<sup>l</sup>** - Mean Error - stredná chyba odhadu

$$M.E.^l = \frac{\sum_{\mathbf{U} \in \mathcal{T}^O} \sum_{v \in V(\mathbf{U})} \left( \sum_{u \in de_{l,\mathbf{U}}[v]-L(\mathbf{U})} (\mathbf{Y}_u - \mathbf{U}_v) \right)}{N_l^{RNL}}, \quad (28)$$

**M.A.E.<sup>l</sup>** - Mean Absolute Error - stredná absolútna chyba odhadu

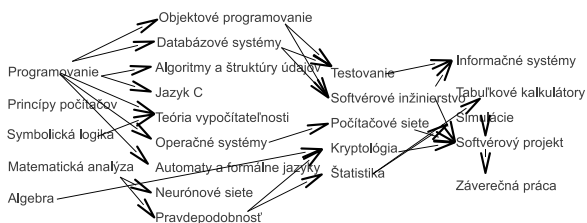
$$M.A.E.^l = \frac{\sum_{\mathbf{U} \in \mathcal{T}^O} \sum_{v \in V(\mathbf{U})} \left( \sum_{u \in de_{l,\mathbf{U}}[v]-L(\mathbf{U})} |\mathbf{Y}_u - \mathbf{U}_v| \right)}{N_l^{RNL}}. \quad (29)$$

**M.S.E.**<sup>1</sup> -] Mean Squared Error - stredná kvadratická chyba odhadu

$$M.S.E.^1 = \frac{\sum_{U \in T^O} \sum_{v \in V(U)} \left( \sum_{u \in del_U, U[v]-L(U)} (Y_u - U_v)^2 \right)}{N_l^{RNL}}, \quad (30)$$

#### 4 Vyhodnotenie dosiahnutých výsledkov

Sieť je schopná natrénovať sa dvojice štruktúr, v ktorých vstup odpovedá výrazu a výstup odpovedá priebežne vypočítaným hodnotám. Vstupné štruktúry boli vytvorené k acyklickým grafom študijných programov. Príklad štruktúry je na obrázku 4.



**Obr. 4:** Príklad použitej štruktúry.

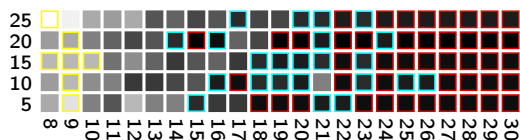
Boli vybrané typické tvary booleovských výrazov vyjadrujúcich prerekvizity v študijných programoch. K ním bola vytvorená množina  $V_V$ . Na základe týchto vstupných štruktúr (výrazov) vytvoríme tréningovú a testovaciu množinu.

Aby sme mohli pripraviť tréningovú a overovaciu množinu, je potrebné uviesť, ako sme vytvárali dvojice  $(U; Y)$ . V množine  $V_V$  sa nachádza 8 vrcholov odpovedajúcich premenným. Keďže trénujeme booleovské výrazy, hodnoty listov môžu nadobúdať hodnoty  $\{0; 1\}$  a teda ôsmim premenným odpovedá 28 možných priradení hodnôt. Ku každej štruktúre  $D \in V_V$  vytvoríme výstupné štruktúry  $\{Y^1, \dots, Y^{256}\}$  podľa pravidiel určených pre dvojice štruktúr  $(D; Y)$ . Takto získame 2 304 dvojíc, ktorých 25% náhodne vyberieme do množiny  $T^T$  a 4% náhodne vyberiem do množiny  $T^O$ .

Na tréningovanie sme použili siete s rozmerom mriežky 8, 9, 10 až 30 a počtom neurónov v skrytej vrstve je 5, 10, 15, 20 a 25. Parameter  $\beta = -2 0281$ . Počet iterácií prvej fázy tréningovania bol 3 000 a druhej fázy bol 5 000. Učiaci koeficient bol 0,01.

#### 4.1 Diferenciácia víťazov

Prvou skúmanou kvalitatívnou charakteristikou je diferenciácia víťazov WD, hodnoty ktorej sú uvedené na obrázku 5 pomocou grafu s nameranými hodnotami. Zistené hodnoty sú reprezentované odtieňom šedivej farby štvorcika. Biela farba odpovedá najnižšej hodnote a čierna farba odpovedá najvyššej hodnote z rozsahu hodnôt určených pre graf. Keďže identifikátorov vrcholov je podstatne menej ako v prípade predmetov študijných programov, siete dosahujú veľmi dobré výsledky rozdelenia vrcholov medzi neuróny mriežky. Dokonca väčšina sietí má hodnotu WD v rozsahu najlepších hodnôt. Namerané hodnoty diferenciácie víťazov boli v rozsahu intervalu  $\langle 0.511; 0.9305 \rangle$ . Už aj sieť s rozmerom mriežky 15 dosiahla maximálnu hodnotu WD.



**Obr. 5:** Diferenciácia víťazov pre siete SLVT.

#### 4.2 Topologické vlastnosti vstupných štruktúr

Sieť SLVT obsahuje časť pozostávajúcu z tréningovania štruktúr. Keďže sme pre jeden rozmer trénovali hneď 5 sietí (počty skrytých neurónov 5, 10, 15, 20, 25), tak je dobré aspoň v krátkosti sa pozrieť na schopnosť natrénovať sa štruktúry. V tabuľke 1 sú uvedené niektoré hodnoty dôveryhodnosti pre sieť SLVT cez všetky úrovne vnorenia pri rekurzívnom spracovaní štruktúr. V prvom riadku je uvádzaný rozmer štvorcovej mriežky. Namerané hodnoty sú z rozsahu intervalu  $\langle 0.328; 0.831 \rangle$ . Opäť sú tu uvedené hodnoty podľa očakávaného usporiadania. Zistené hodnoty kvality d overyhodnosti sú rozdelené do troch pozorovaní, a to z pohľadu usporiadania daného sieťou (US), očakávaného relatívneho usporiadania (RU) a očakávaného úplného usporiadania (UU).

#### 4.3 Chyby M.E., M.A.E. a M.S.E

Pri vyhodnocovaní chýb nás zaujímali chyby medzi očakávanými a vypočítanými hodnotami. Pri vyhodnocovaní interpretácie štruktúr nás zaujímali topologické vlastnosti, a to schopnosť interpretácie štruktúr do hĺbky. V prípade siete SLVT nás zaujíma kvalita schopnosti interpretovať výpočty výrazov v jednotlivých úrovniach spracovania od listov/ premenných smerom vyššie v zmysle rekurzívneho spracovania výrazov až k supervrcholu,



Skr.v.	Usp.	8	9	18	19	26	27
5	US:	0,432	0,432	0,682	0,649	0,672	0,705
	RU:	0,432	0,432	0,682	0,649	0,672	0,705
	UU:	0,383	0,406	0,682	0,571	0,666	0,705
10	US:	0,503	0,464	0,669	0,500	0,786	0,821
	RU:	0,490	0,464	0,669	0,494	0,734	0,821
	UU:	0,497	0,393	0,487	0,432	0,753	0,821
15	US:	0,432	0,412	0,692	0,656	0,705	0,731
	RU:	0,432	0,409	0,688	0,649	0,705	0,727
	UU:	0,383	0,367	0,649	0,549	0,627	0,701
20	US:	0,474	0,448	0,753	0,718	0,812	0,802
	RU:	0,474	0,442	0,753	0,718	0,812	0,802
	UU:	0,419	0,396	0,646	0,627	0,802	0,776
25	US:	0,406	0,422	0,636	0,604	0,643	0,701
	RU:	0,406	0,422	0,636	0,604	0,643	0,701
	UU:	0,367	0,351	0,620	0,487	0,545	0,662

**Tabuľka 1:** Namerané hodnoty dôveryhodnosti cez všetky štruktúry sietí SLVT.

ktorému odpovedá vypočítaná hodnota výrazu. Hodnoty dôveryhodností sú zobrazené pre úrovne od potomkov supervrchola k listom a hodnoty nameraných chýb pre úrovne od supervrchola k predchodcom listov.

Posledné pozorovanie bolo zamerané na hĺbku jednotlivých štruktúr. Z množín  $T^T$  a  $T^O$  boli vytvorené množiny  $T_x^T$  a  $T_x^O$ , ktoré obsahujú dvojice štruktúr s najväčšou hĺbkou pri rekurzívnom spracovaní  $x$ , kde  $x = 1, 2, 3, 4$ . Na týchto množinách sme postupne vykonali výpočty chýb  $M.E.^l$ ,  $MA.E.^l$  a  $M.S.E.^l$ .

Poznamenajme, že pre prehľadnosť budeme uvádzať najlepšie a najhoršie namerané chyby pri argumentoch

**arg=hodnota; rozmer mriežky; počet neurónov**

v nasledujúcom zápise pre najmenšiu chybu z trénovacej množiny  $Tr_{min}(arg)$ , najväčšiu chybu z trénovacej množiny  $Tr_{max}(arg)$ , pre najmenšiu chybu z overovacej množiny  $Tst_{min}(arg)$  a najväčšiu chybu z overovacej množiny  $Tst_{max}(arg)$  pre jednotlivé pozorovania. Poznamenajme, že pri vyhodnocovaní dôveryhodnosti a hodnôt nameraných chýb bolo použité usporiadanie danej siete, lebo pri spracovaní výrazov nezáleží na poradí spracovania potomkov.

V tabuľke 2 sú uvedené hraničné hodnoty zistených chýb  $M.E.$  podľa úrovni. Zo zistených hodnôt môžeme pozorovať, že sieť zväčša dosiahla zlé výsledky pri menšom rozmere mriežky a počte neurónov v skrytej vrstve. S rastúcim rozmerom sa hodnota chýb blížila k nule. Ak uvažujeme vyhodnotenie chyby v smere spracovania výrazov od úrovne 3, tak pre menšie rozmery dosahujú hodnoty chýb výrazné výkyvy v úrovni 3. Túto skutočnosť môžeme vysvetliť slabšou schopnosťou siete interpretovať štruktúry vo väčšej hĺbke, čo potvrdzuje aj nameraná dôveryhodnosť podľa úrovni pre vstupné štruktúry.

Poslednou skupinou sú chyby s označením "Tst  $M.E.$  ur.  $x - y$ " a "Tr.  $M.E.$  ur.  $x - y$ ", kde  $x = 1, 2, 3, 4$  a  $y = 0, 1, 2, 3$  a  $y < x$ . Tieto chyby sú vypočítané pre testovacie množiny  $T_x^O$  a trénovacie množiny  $T_x^T$ , ktoré sú rozdelením štruktúr podľa najhlbšieho vnorenia  $x$  pri spracovaní štruktúry.

Úr.	min	max
0	$Tst(-0, 4312; 9; 10)$ $Tr(-0, 4396; 12; 5)$	$Tst(0, 1224; 21; 5)$ $Tr(0, 0982; 29; 15)$
1	$Tst(-0, 3791; 12; 5)$ $Tr(-0, 3391; 12; 10)$	$Tst(-0, 0096; 13; 20)$ $Tr(-0, 0053; 13; 20)$
2	$Tst(-0, 4881; 12; 5)$ $Tr(-0, 4842; 12; 5)$	$Tst(0, 0314; 30; 15)$ $Tr(-0, 0044; 30; 15)$
3	$Tst(-0, 26; 8, 10, 11; 15)$ $Tr(-0, 2290; 20; 25)$	$Tst(0, 4108; 9; 15)$ $Tr(0, 4443; 9; 15)$

**Tabuľka 2:** Hraničné chyby  $M.E.$  podľa úrovni.

vacie množiny  $T_x^O$  a trénovacie množiny  $T_x^T$ , ktoré sú rozdelením štruktúr podľa najhlbšieho vnorenia  $x$  pri spracovaní štruktúry.

Úr.	min	max
1 - 0	$Tst_{nim}(-0, 5571; 10; 5)$ $Tr_{nim}(-0, 4609; 8; 5)$	$Tst_{max}(0, 1501; 10; 5)$ $Tr_{max}(0, 1232; 15; 10)$
2 - 0	$Tst_{nim}(-0, 7667; 13; 25)$ $Tr_{nim}(-0, 6902; 21; 25)$	$Tst_{max}(0, 26689; 14; 5)$ $Tr_{max}(0, 0091; 20; 5)$
2 - 1	$Tst_{nim}(-0, 3769; 13; 15)$ $Tr_{nim}(-0, 2905; 13; 15)$	$Tst_{max}(0, 0921; 12; 20)$ $Tr_{max}(0, 0524; 8; 5)$
3 - 0	$Tst_{nim}(-0, 4759; 11; 5)$ $Tr_{nim}(-0, 4921; 18; 20)$	$Tst_{max}(0, 3218; 21; 5)$ $Tr_{max}(0, 2180; 25; 15)$
3 - 1	$Tst_{nim}(-0, 5585; 8; 5)$ $Tr_{nim}(-0, 5273; 8; 25)$	$Tst_{max}(0, 0253; 13; 15)$ $Tr_{max}(-0, 0067; 11; 5)$
3 - 2	$Tst_{nim}(-0, 4642; 12; 5)$ $Tr_{nim}(-0, 4596; 12; 5)$	$Tst_{max}(0, 0415; 26; 15)$ $Tr_{max}(0, 0031; 26; 15)$
4 - 0	$Tst_{nim}(-0, 1; 26; 10)$ $Tr_{nim}(-0, 1219; 13; 5)$	$Tst_{max}(0, 8461; 21; 25)$ $Tr_{max}(0, 7831; 21; 25)$
4 - 1	$Tst_{nim}(-0, 2143; 27; 5)$ $Tr_{nim}(-0, 2260; 27; 5)$	$Tst_{max}(0, 3935; 18; 20)$ $Tr_{max}(0, 3656; 18; 20)$
4 - 2	$Tst_{nim}(-0, 5939; 10; 5)$ $Tr_{nim}(-0, 5514; 12; 5)$	$Tst_{max}(0, 0910; 24; 10)$ $Tr_{max}(0, 0488; 29; 20)$
4 - 3	$Tst_{nim}(-0, 5939; 10; 5)$ $Tr_{nim}(-0, 2260; 9; 15)$	$Tst_{max}(0, 0910; 24; 10)$ $Tr_{max}(-0, 4443; 9; 20)$

**Tabuľka 3:** Hraničné chyby  $M.E.$  podľa  $x - y$ .

Najprv budeme analyzovať štruktúry typu 1 - 0. Tu boli dosiahnuté vcelku dobre výsledky, pre siete od mriežky 24 oscilovali okolo nuly. Zo znižujúcim sa rozmerom mriežky sa namerané hodnoty chyby viac rozptyľujú. Pre spracúvané štruktúry 2 - y sa chyby pre vrcholy v úrovni 1 postupne stabilizujú so zväčšujúcim sa rozmerom mriežky, ale pre úroveň 0, ktorá odpovedá supervrcholu, a teda aj vo výsledku výrazu je chyba väčšia a viac "rozhádzaná". Tiež sa vylepšuje s rozširujúcim sa rozmerom mriežky, ale nie až tak dobre ako pre úroveň 1. Podobné pozorovanie môžeme spraviť pre štruktúry typu 3 - y a 4 - y. V najnižšej úrovni, kde sa merala chyba, boli dosiahnuté výsledky najlepšie a so zväčšujúcim sa rozmerom chyba začala oscilovať okolo nuly. Postupným spracovaním výrazu smerom k supervrcholu sa aj chyba znižovala. Pre siete typu 3 - y bližšie k supervrcholu nebola až taká veľká ako pre štruktúry typu 4 - y. Ťažko vysvetliteľným je prípad 4 - 0, kde pre siete s väčším rozmerom mriežky sú dosahované veľmi veľké hodnoty

chýb. V tabuľke 3 sú zobrazené zistené hraničné hodnoty nameraných chýb  $M.E.$  pre štruktúry typu  $x - y$  v jednotlivých úrovniach.

Chyby  $M.A.E.$  a  $M.S.E.$  vyhodnotíme spolu. Základnou charakteristikou týchto chýb je to, že majú kladné hodnoty a teda hodnoty blízke nule budú najlepšie. Prvá chyba je súčet absolútnych hodnôt a druhá je súčet štvorcov. Keďže vypočítané aj očakávané hodnoty výrazov a medzivýpočtov sú v rozsahu intervalu  $(0, 1)$ , tak chyba  $M.A.E.$  bude "špicatejšia". Najprv vyhodnotíme chyby spolu cez všetky úrovne a pre množiny  $T^O$  a  $T^T$ . Hraničné hodnoty chýb  $M.A.E.$  pre  $T^O$  boli  $Tst_{nim}(0, 01388; 12; 5)$ ,  $Tst_{max}(0, 4483; 30; 20)$  a pre tréningovú množinu  $T^T$  boli  $Tr_{min}(0, 1370; 12; 5)$ ,  $Tr_{max}(0, 4385; 27; 20)$ .

Hraničné hodnoty chýb: chyba  $M.S.E.$  pre  $T^O$  boli  $Tst_{nim}(0, 0635; 30; 15)$ ,  $Tst_{max}(0, 3598; 12; 5)$  a hodnoty  $Tr_{min}(0, 0720; 30; 15)$ ,  $Tr_{max}(0, 3492; 12; 5)$  boli namerané pre tréningovú množinu  $T^T$ . Aj v prípade týchto chýb môžeme pozorovať, že so zvyšujúcim sa rozmerom mriežky chyby klesajú, čo môžeme vysvetliť lepšou hodnotou dôveryhodnosti pre siete s väčším rozmerom mriežky. Pre tento prípad je možné pozorovať, že sieť s 5 skrytými neurónmi v mriežke mala väčšiu chybu pri menšom rozmere mriežky než ostatné siete.

Úr.	min	max
0	$Tst_{nim}(0, 0819; 8; 15)$ $Tr_{nim}(0, 0957; 8; 15)$	$Tst_{max}(0, 5193; 29; 5)$ $Tr_{max}(0, 5017; 8; 15)$
1	$Tst_{nim}(0, 1569; 26; 20)$ $Tr_{nim}(0, 1856; 26; 20)$	$Tst_{max}(0, 4338; 23; 10)$ $Tr_{max}(0, 4037; 8; 25)$
2	$Tst_{nim}(0, 0919; 12; 5)$ $Tr_{nim}(0, 0841; 27; 20)$	$Tst_{max}(0, 5572; 27; 20)$ $Tr_{max}(0, 5201; 12; 5)$
3	$Tst_{nim}(0, 0182; 10; 5)$ $Tr_{nim}(0, 0603; 28; 10)$	$Tst_{max}(0, 5012; 9; 15)$ $Tr_{max}(0, 5170; 9; 15)$

**Tabuľka 4:** Hraničné chyby  $M.A.E.$  podľa úrovni.

Aj tu sme pozorovanie kvality natrénovania výrazov rozdelili do úrovni. V tabuľkách 4 a 5 sú uvedené namerané hraničné hodnoty chýb. Aj tu sú chyby zväčša menšie pre väčší rozmer mriežok. Namerané hodnoty v úrovni 2 a 3 sú klesajúce s rozmerom mriežky. Zaujímavé je, že hodnoty chýb pre úroveň 1 oscilujú zväčša medzi hodnotami 0, 2 až 0, 4 pre všetky rozmery mriežky. I v tomto meraní podľa úrovni svoju úlohu zohráva dôveryhodnosť aj napriek tomu, že jej vplyv je menej viditeľný než v prípade chýb  $M.E.$

V závere vyhodnotenia chýb sa ešte zmienime o chybách  $M.A.E.$  a  $M.S.E.$  pre štruktúry typu  $x - y$ .

Zistené chyby pre štruktúry typu 1 - 0 potvrdzujú, že čím väčšia mriežka, tým lepšia schopnosť siete vypočítať

Úr.	min	max
0	$Tst(0, 0242; 29; 5)$ $Tr_{nim}(0, 0345; 29; 25)$	$Tst(0, 4437; 8; 15)$ $Tr_{max}(0, 4214; 8; 15)$
1	$Tst_{nim}(0, 0932; 21; 20)$ $Tr_{nim}(0, 1123; 21; 20)$	$Tst_{max}(0, 3637; 23; 10)$ $Tr_{max}(0, 3129; 15; 15)$
2	$Tst_{nim}(0, 0255; 30; 25)$ $Tr_{nim}(0, 0255; 30; 25)$	$Tst_{max}(0, 4599; 12; 5)$ $Tr_{max}(0, 4291; 12; 5)$
3	$Tst_{nim}(0, 022; 19; 5)$ $Tr_{nim}(0, 0040; 28; 10)$	$Tst_{max}(0, 2701; 9; 15)$ $Tr_{max}(0, 28123; 9; 15)$

**Tabuľka 5:** Hraničné chyby  $M.S.E.$  podľa úrovni.

Úr.	min	max
1 - 0	$Tst_{nim}(0, 0160; 26; 20)$ $Tr_{nim}(0, 0119; 21; 25)$	$Tst_{max}(0, 5571; 10; 25)$ $Tr_{max}(0, 4609; 8; 5)$
2 - 0	$Tst_{nim}(0, 0536; 28; 10)$ $Tr_{nim}(0, 0756; 30; 10)$	$Tst_{max}(0, 7667; 21; 25)$ $Tr_{max}(0, 6902; 13; 25)$
2 - 1	$Tst_{nim}(0, 0308; 29; 5)$ $Tr_{nim}(0, 0362; 29; 5)$	$Tst_{max}(0, 3914; 9; 25)$ $Tr_{max}(0, 3338; 9; 25)$
3 - 0	$Tst_{nim}(0, 1224; 22; 5)$ $Tr_{nim}(0, 1378; 22; 25)$	$Tst_{max}(0, 6088; 22; 5)$ $Tr_{max}(0, 6337; 16; 25)$
3 - 1	$Tst_{nim}(0, 1986; 16; 15)$ $Tr_{nim}(0, 2517; 21; 20)$	$Tst_{max}(0, 5950; 10; 5)$ $Tr_{max}(0, 5547; 16; 10)$
3 - 2	$Tst_{nim}(0, 0926; 28; 20)$ $Tr_{nim}(0, 0866; 26; 20)$	$Tst_{max}(0, 5413; 12; 5)$ $Tr_{max}(0, 4994; 12; 5)$
4 - 0	$Tst_{nim}(0, 30; 20)$ $Tr_{nim}(0, 000; 30; 20)$	$Tst_{max}(0, 8461; 21; 25)$ $Tr_{max}(0, 7831; 21; 25)$
4 - 1	$Tst_{nim}(0, 0177; 24; 10)$ $Tr_{nim}(0, 0177; 24; 10)$	$Tst_{max}(0, 4615; 18; 20)$ $Tr_{max}(0, 5185; 16; 25)$
4 - 2	$Tst_{nim}(0, 0573; 21; 15)$ $Tr_{nim}(0, 0586; 27; 25)$	$Tst_{max}(0, 6217; 12; 5)$ $Tr_{max}(0, 5768; 12; 5)$
4 - 3	$Tst_{nim}(0, 0182; 10; 5)$ $Tr_{nim}(0, 0603; 28; 10)$	$Tst_{max}(0, 5012; 9; 15)$ $Tr_{max}(0, 5170; 9; 15)$

**Tabuľka 6:** Hraničné chyby  $M.A.E.$  podľa  $x - y$ .

Úr.	min	max
1 - 0	$Tst_{nim}(0, 0010; 26; 20)$ $Tr_{nim}(0, 0008; 24; 20)$	$Tst_{max}(0, 47671; 10; 25)$ $Tr_{max}(0, 3709; 8; 5)$
2 - 0	$Tst_{nim}(0, 0033; 28; 10)$ $Tr_{nim}(0, 0094; 30; 10)$	$Tst_{max}(0, 6767; 16; 25)$ $Tr_{max}(0, 6002; 13; 25)$
2 - 1	$Tst_{nim}(0, 0015; 29; 5)$ $Tr_{nim}(0, 0020; 29; 5)$	$Tst_{max}(0, 2869; 13; 15)$ $Tr_{max}(0, 2178; 9; 25)$
3 - 0	$Tst_{nim}(0, 0368; 22; 5)$ $Tr_{nim}(0, 0505; 16; 25)$	$Tst_{max}(0, 5277; 8; 15)$ $Tr_{max}(0, 5384; 29; 25)$
3 - 1	$Tst_{nim}(0, 1221; 16; 15)$ $Tr_{nim}(0, 1471; 21; 20)$	$Tst_{max}(0, 5077; 25; 10)$ $Tr_{max}(0, 4700; 16; 10)$
3 - 2	$Tst_{nim}(0, 0299; 30; 25)$ $Tr_{nim}(0, 0299; 26; 20)$	$Tst_{max}(0, 4396; 12; 5)$ $Tr_{max}(0, 4045; 12; 5)$
4 - 0	$Tst_{nim}(0, 000; 9; 20)$ $Tr_{nim}(0, 0000; 9; 20)$	$Tst_{max}(0, 7189; 21; 25)$ $Tr_{max}(0, 6412; 21; 25)$
4 - 1	$Tst_{nim}(0, 0016; 24; 10)$ $Tr_{nim}(0, 0058; 16; 10)$	$Tst_{max}(0, 3977; 28; 20)$ $Tr_{max}(0, 4086; 16; 25)$
4 - 2	$Tst_{nim}(0, 0043; 21; 15)$ $Tr_{nim}(0, 0069; 30; 25)$	$Tst_{max}(0, 5420; 12; 5)$ $Tr_{max}(0, 4965; 12; 5)$
4 - 3	$Tst_{nim}(0, 0022; 19; 5)$ $Tr_{nim}(0, 0040; 28; 10)$	$Tst_{max}(0, 2701; 9; 15)$ $Tr_{max}(0, 2823; 9; 15)$

**Tabuľka 7:** Hraničné chyby  $M.S.E.$  podľa  $x - y$ .

výraz. Ale toto tvrdiť nemôžeme pre štruktúry typu 2 - y. Namerané chyby pre úroveň  $y = 1$  len veľmi pozvoľna klesajú so zväčšujúcim sa rozmerom mriežky a to isté môžeme pozorovať aj pre úroveň 2-0. Štruktúry typu 3-y dosahujú v úrovni 3 - 2 očakávaný vývoj chyby, t.j. jej znižovanie so

zväčšujúcim sa rozmerom mriežky. Ale pre úrovne  $3 - 1$  a  $3 - 0$  sú namerané hodnoty nezávislé od rozmeru mriežky. To isté potvrdzuje aj pozorovanie pre štruktúry typu  $4 - y$ . Pre úroveň  $4 - 3$  je ešte pozorovateľná závislosť medzi znižujúcou sa chybou a zväčšujúcim sa rozmerom mriežky. V tabuľkách 6 a 7 sú uvedené hraničné hodnoty nameraných chýb *M.A.E.* a *M.S.E.*, ktoré nie sú zreteľne zadené v zmysle očakávaného pravidla “Väčšia chyba, menší rozmer mriežky”.

## 5 Záver

Na základe testovacích výsledkov je možné konštatovať, že úspešnosť výpočtu siete SLVT závisí hlavne od schopnosti interpretácie štruktúry odpovedajúcej zadávanému výrazu. S väčším počtom vnorení pri ich spracovaní sa prejavuje nižšia schopnosť správne vypočítať výraz neurónovou sieťou, čo je podmienené tým, že neurónová sieť sa nemusí dobre naučiť asociácie do hĺbky. V ďalšej práci by sme sa chceli zamerať na zjednodušenie modelu navrhutej siete, ktorá bola vytvorená zo siete pre tréning štruktúr.

## PodĎakovanie

Tento príspevok vznikol za podpory grantovej agentúry VEGA v rámci grantového projektu VEGA 1/0492/12.

## Literatúra

- [1] G. Andrejková, J. Oravec: Processing Acyclic Data Structures using modified Self-Organizing Maps. In: LNCS 6692: Advances in Computational Intelligence, part II., Springer, 2011, pp. 145-152.
- [2] G. Andrejková, J. Oravec: Application of Modified SOM Neural Networks on Acyclic Data Structures. Acta Electrotechnica et Informatica, Vol. 12, No. 2, 2012, p. 3-8.
- [3] G. Andrejková, J. Oravec: Spracovanie acyklických dátových štruktúr pomocou SOM neurónových sietí s modifikovanými laterálnymi váhami. In: Kognice a umelý život XII : 22.-25. máj 2012, Pruhonice u Prahy. Agentura Action M, 2012, S. 10-15.
- [4] P. M. Frasconi, M. Gori, and A. Sperduti: A general framework of adaptive processing of data structures. IEEE-NN, 9(5), September 1998, pp. 768-786.
- [5] M. Hagenbuchner, A. Sperduti, A. C. Tsoi : A self-organizing map for adaptive processing of structured data. IEEE Transactions on Neural Networks, 14 (3), 2003, pp. 491-505.
- [6] M. Hagenbuchner, A. C. Tsoi: A supervised self-organizing map for structures, Proceedings IEEE International Joint Conference on Neural Networks, 25-29 July 2004, vol 3, pp. 1923-1928.
- [7] M. Hagenbuchner, A. C. Tsoi, A. Sperduti: A supervised self-organising map for structured data. In: Proc. WSOM 2001: Advances in Self-Organizing Maps. Springer, Berlin, 2001, pp. 21-28.
- [8] B. Hammer, A. Micheli, A. Sperduti, M. Strickert: A general framework for unsupervised processing of structured data, Neurocomputing, No. 57, Elsevier, 2004, pp. 3-35.
- [9] S. Haykin: Neural networks: A comprehensive foundation (2nd ed.), Prentice-Hall, New Jersey (1999)
- [10] P. Sinčák, G. Andrejková: Neural Networks I., II., ELFA Košice, 1996 (in Slovak).
- [11] A. Sperduti: Tutorial on neurocomputing of structures. In: Knowledge-Based Neurocomputing, I. Cloete and J. M. Zurada, Eds. Cambridge, MA: MIT Press, 2000, pp. 117-152
- [12] P. Vančo, I. Farkáš: Experimental comparison of recursive self-organizing maps for processing tree-structured data. Neurocomputing 73 (7-9), 2010, pp. 1362-1375.



# Fenomenologie volní akce: motorické intence a pocit záměru

Ondřej Bečev

Lékařská fakulta, Masarykova Univerzita,  
Kamenice 5, 625 00 Brno, Česká Republika

Email: [ondrej@becev.cz](mailto:ondrej@becev.cz)

Fakulta informatiky a statistiky; Vysoká škola ekonomická, Praha  
nám W. Churchilla 4, 130 67 Praha 3, Česká Republika

Email: [ondrej.becev@vse.cz](mailto:ondrej.becev@vse.cz)

## Abstrakt

Přehledový článek poskytuje seznámení se současnou podobou diskuse o fenomenálním aspektu jednání, zejména s důrazem na pojem motorických intencí zkoumaných v návaznosti na Libetovy experimenty. Je zde poskytnut stručný úvod do základních pojmů, jakými jsou Sense of Agency, Awareness of Intention a další. Těžištěm článku je snaha o uchopení *jednání* (agence) z fenomenálního pohledu a poskytnutí plastičtějšího pohledu na toto zdánlivě monolitické téma. Popsáno je též studium časových horizontů vědomých intencí a jejich prožitkových forem a souvislost fenomenologie volní akce a různých úrovní vědomí. Představeny jsou dále také významné perspektivní, retrospektivní a integrativní modely vzniku *prožitku jednání*.

## 1 Úvod

### 1.1 Empirický výzkum svobodné vůle

Empirický výzkum svobodné vůle, rozvíjený zejména v posledních dvaceti letech, vyvolal otázky po kauzální potenci vědomí, konkrétně po kauzální účinnosti vědomých intencí. Počínaje experimenty BENJAMINA LIBETA s *potenciálem připravenosti* (Bereitschaftspotential, readiness potential) [16] se v centru pozornosti ocitla témata *motorického jednání* (motor agency), *korespondujících neurálních procesů* a *vědomé intence* (conscious intention). Pozdější výzkum zpřesnil náš vzhled a zdokumentoval dynamiku mozkové aktivity předcházející prožitku volního rozhodnutí. Výrazně k tomu přispěl například výzkum *lateralizovaného potenciálu připravenosti* (lateralized readiness potential) [10], či objev kódování specifického rozhodnutí v prefrontálním kortexu, předcházející okamžik vědomého rozhodnutí až o 8 sekund [21, 2]. Výtečný přehled současných neuropsychologických poznatků poskytuje F. KOUKOLÍK [13]. Přesné časování, vnitřní dynamika prožitku a související neurální koreláty

*fenomenálního prožitku rozhodnutí* však zůstávají nejasné.

### 1.2 Fenomenální prožitek jednání

Vzhledem k výrazným úspěchům neurovědy *volního jednání* a jeho centrální role v každodenním životě je překvapující, že bylo toto téma až do relativně nedávné doby ve *filosofii mysli* a *filosofii vědomí* přehlíženo. K důsledkům tohoto opomíjení bychom mohli přiřadit i panující pojmové zmatení. My své úvahy o *vědomém prožitku motorických záměrů* (conscious experience of motor intentions) postavíme kolem dvou zásadních pojmů: *Sense of Agency* a *Awareness of Intention*<sup>1</sup>, kterým se věnuje kapitola 3. Nejprve se však stručně seznámíme s širší pojmové plurality, která filosofům v této disciplíně dělá těžkou hlavu (sekce 2).

Někdo by se mohl ptát, proč se při své intelektuální výpravě za fenomenálními aspekty *záměru* vůbec pouštím do rozsáhlých vod *Sense of Agency* a *Experience of Agency*. Vždyť jednání by, alespoň podle obecného teoretického nahlížení, mělo být až **důsledkem intence**. *Prožitek jednání* by nám tedy neměl být co platný při zkoumání *prožitku intence*. Jak se však ukáže, panují dobré důvody se domnívat, že alespoň určité aspekty *prožitku jednání* ve skutečnosti ještě nejsou ve vědomí přítomny v okamžiku, kdy později reportujeme, že jsme se rozhodli. Jak ukazují současné filosofické a neuropsychologické výzkumy, vědomí je jako pohyblivé písky. To, co je v něm v jednom okamžiku zcela zřetelné, tam za zlomek vteřiny nejen být nemusí, ale nadto bychom svatosvatě odpřisáhli, že to tam nikdy nebylo. Výzkum vědomí je proto vzhledem k jeho mimořádně fluidní a retrospektivní povaze nadmíru obtížný.

<sup>1</sup>V textu dále využívám jak anglických pojmů *sense of agency* a *awareness of intention*, tak jejich českých obdob *prožitek jednání* a *prožitek intence*, která však chápu jako pojmy širší, zastřešující také několik dalších.

## 2 Problematická terminologie

U problému *fenomenálního prožitku jednání* se můžeme často setkat množstvím různých jiných, široce rozšířených termínů, přičemž není jasné, zda označují stejný jev. I krátký pohled na experimentální literaturu odhalí pojmy jako *conscious intention, the judgment of will, intention awareness, intention perception, feeling of will, sense of agency, ownership of agency, conscious awareness, awareness of the intention to move, feeling of „will“, intention judgment, Intention experience, Attention to intention*. Stejně věci si všímá ELISABETH PACHERIE [18], která předkládá seznam pojmů a rozlišení, se kterými se můžeme setkat v literatuře. Ty zahrnují *awareness of a goal, awareness of an intention to act, awareness of initiation of action, awareness of movements, sense of activity, sense of mental effort, sense of physical effort, sense of control, experience of authorship, experience of intentionality, experience of purposiveness, experience of freedom, a experience of mental causation*. Všimá si, že kromě zbytečného terminologického zmnožení zůstává nejasné, jak spolu tyto různé aspekty fenomenologie akce souvisí, do jaké míry jsou disociovatelné, a zda jsou některé z nich základnější než ostatní.

V následujících kapitolách se proto pokusím některé z těchto pojmů vysvětlit a uvést do vzájemného vztahu.

## 3 Prožitek jednání

### 3.1 Sense of Agency

*Sense of Agency* (SoA), je pocit, že já jsem tím, kdo způsobuje, nebo generuje, danou akci. Např. že jsem to já, kdo způsobuje, že se něco pohybuje, nebo že jsem to já, kdo způsobuje nějakou myšlenku v mém proudu vědomí [7]. PATRICK HAGGARD SoA vykládá jako *prožitek ovládnutí jak vlastních akcí, tak skrze ně i událostí ve vnějším světě* [12]. Tento pocit je zásadním způsobem přítomen v jakékoli naší interakci se světem.

Pojem SoA se rychle stal generickým termínem pro celou rodinu témat a chápání se liší vědec od vědce. To vyvolává potřebu se na věc podívat poněkud blíže. Základní rozvrh současné terminologické diskuse byl načrtnut SHAUNEM GALLAGHEREM [7], který na příkladu schizofrenie ukázal nutnost oddělit od *Sense of Agency* složku *Sense of Ownership*. **Sense of Ownership** (SoO) je prožitek, že *jsem to já, kdo zakouší vnímané*. Například pocit, že se hýbe mé tělo, bez ohledu na to, zda se jedná o pohyb volní nebo mimovolní [7]. Gallagher opírá tuto svou distinkci o dichotomii *minimálního a narativního já*.

**Minimální já** (minimal self): Fenomenologicky, tedy ve smyslu prožívání, je *minimální já* vědomí sebe

sama jako bezprostředního subjektu zkušenosti, nerozpostraněného v čase. Minimální já takřka s jistotou závisí na mozkových procesech a na těle zasazeném do prostředí, ačkoliv si toho prožívající nemusí být vědom a uvědomování si této skutečnosti není předpokladem k tomu mít zkušenost, kterou bychom mohli považovat za *zkušenost sebe*. Minimální já je prereflexivní, čistě fenomenální.

**Narativní já** (narrative self): Víceméně koherentní *já* (nebo *obraz sebe*) konstituovaný v průběhu času skrze „příběhy“ a narativní vzpomínky, které (o) sobě vyprávíme a které nám o nás říkají druzí. Neuropsychologické teorie by identifikovali funkci narativního já s epizodickou pamětí a jejím narušením.

*Narativní já*, odpovídající odkazování k sobě jako objektu, v tuto chvíli není tolik zajímavé, jako *minimální já*. To totiž odpovídá referování k sobě jako *subjektu* a vyznačuje se některými zvláštnostmi, plynoucími z *principu imunity*, což rozebírají následující tři oddíly.

### 3.1.1 Imunita vůči omylu z chybné identifikace

Wittgensteinovo rozlišení dvou způsobů reference k sobě pomocí zájmena *já*, na odkazování „jako subjekt“ a „jako objekt“, dalo vzniknout teorii o *imunitě vůči omylům z chybné identifikace* (immunity to error through misidentification) vztahující se k zájmenu první osoby. SYDNEY SHOEMAKER ukázal [20], že **princip imunity** se uplatňuje pouze u použití *já* „jako subjekt“. V současné době však probíhají diskuse, že prožitky pacientů s diagnózou schizofrenie mohou být námitkou k principu imunity. Pacient se schizofrenií, trpící intrapsychickými halucinacemi, jako je vkládání myšlenek, může například tvrdit, že on není tím, kdo myslí nějakou konkrétní myšlenku, ačkoliv ve skutečnosti je tím, kdo ji myslí. Tento druh specifického myšlenkového procesu schizofrenních ilustruje následující příklad výpovědi: „*Do mé mysli jsou vkládány myšlenky jako třeba „Zabij Boha“*“. Je to, jako kdyby běžely v mé mysli, ale není to tak. Pochází od tamtoho chlápka, Chrise. Jsou to jeho myšlenky“. Citace dle [7]. Na případu schizofrenie se tak ukazuje potřeba rozlišovat SoA a SoO. Fenomeny, jako *blud kontroly*, sluchové halucinace a vkládání myšlenek se zdají být spojené spíše s problémem *sense of agency*, než se *sense of ownership*.

### 3.1.2 Vysvětlení vkládání myšlenek pomocí distinkce SoA a SoO

Právě situace schizofrenie a následující krátký příklad s *motorickou akcí* nám pomůže ujasnit dva výše představené a příbuzné aspekty *minimálního sebe-vědomí*. Při normální zkušenosti volní nebo chtěné akce se SoA a SoO shodují a

jsou neodližitelné. Když se natáhnu pro hrnek, vím, že to je *můj* úkon. Tento obvyklý společný výskyt a shoda nás může vést ke (zavádějícímu) chápání *vlastnictví akce* (ownership of action) jakožto *jednání* samotného: tedy, že vlastník akce je tou osobou, která je (specifickým způsobem) kauzálně zapojená do vytváření akce. I přes obvyklý spoluvýskyt se však jedná o dva různé mechanismy.

Již v případě *mimovolní akce* je vcelku dobře možné rozlišovat mezi SoA a SoO. Mohu sice potvrdit *vlastnictví pohybu* (ownership of a movement) – to jest mám pocit, že já jsem ten, kdo se pohybuje nebo je pohybován – a mohu jej připsat sám sobě jakožto *svůj* pohyb, ale nemusím mít pocit zapříčinění nebo kontroly tohoto pohybu. Tehdy bychom řekli, že mám *sense of ownership*, ale nemám *sense of agency*. A přesně taková situace dle Gallaghera dává vzniknout *pocitu vkládání myšlenek* u schizofrenního onemocnění.

### 3.1.3 Je naše jednání vyvoláno záměrem, nebo naopak?

Ukazuje se, že *pocit vlastnictví motorické akce* (sense of ownership for motor action) může být vysvětlen pomocí sebe-uvědomění skrze jednání v prostředí, které je zabudované již do naší percepce a mechanismu řízení pohybu. Experimentální výzkum na zdravých subjektech na druhou stranu naznačuje, že *prožitek jednání* u motorických akcí je založen na mechanismu, který předchází akci a překládá intenci v akci [7].

### 3.1.4 Distinkce SoA/OoA a Comparator model

Gallagher nachází oporu pro svou teorii v klasickém *komparátorovém modelu* (comparator model, **Obrázek 1** na následující straně) C. FRITHA [5], když ukazuje, že poškození SoA, zodpovědné za intrapsychické halucinace, koresponduje s dysfunkcí konkrétních komponent tohoto modelu. Schizofrenní pacienti, kteří trpí vkládáním myšlenek a bludem kontroly mají také problém s feed-forward monitoringem pohybu, přecházejícím akci. Nemají však poškozený mechanismus motorické kontroly založený na porovnávání zamýšleného pohybu a senzoričké zpětné vazby.

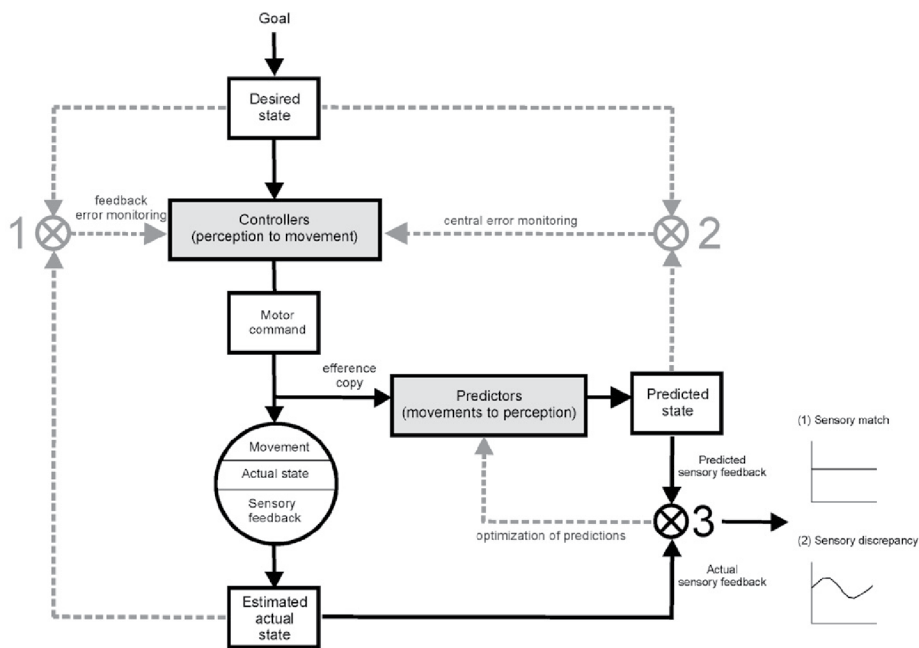
Komparátorový proces předvídá senzoričkou zpětnou vazbu z pohybu a podtrhuje online prožitek jednání, který doplňuje eko-logický (tj. z prostředí) *pocit příslušnosti sobě* (sense of self-ownership), který je založen na reálné smyslové zpětné vazbě [7]. Pokud selže *forward model*, nebo není eferentní kopie správně vytvořena, senzoričká zpětná vazba může stále ještě vytvořit SoO, např. „jsem v pohybu“, ale SoA bude nesprávný („tento pohyb nezpůsobuji já“). A to i přes to, že uskutečňovaný pohyb

odpovídá pohybu zamýšlenému.

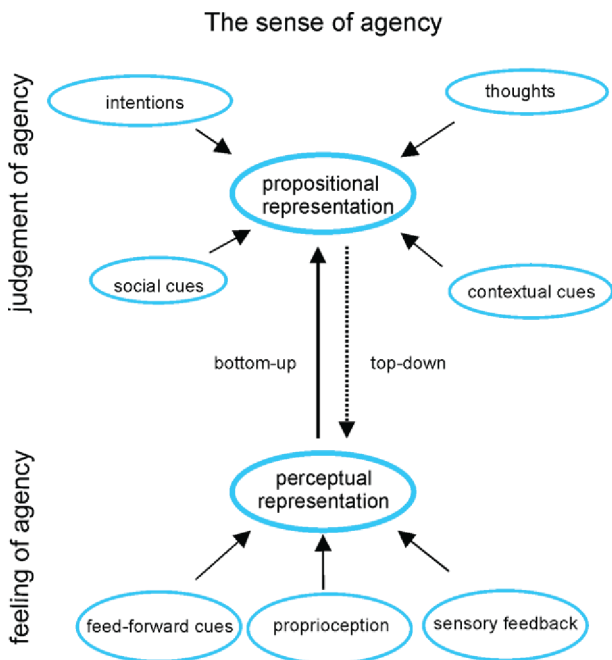
Comparator model se ovšem v současné době střetává s rozsáhlou kritikou, např. [22, 23], která napadá jeho vhodnost v roli *neurokognitivního modelu vlastního jednání*, což souvisí mj. právě se schopností vysvětlit výše zmíněné jevy při schizofrenii.

## 3.2 Experience of Agency

M. Synofzik a G. Vosgerau dávají namísto *sense of agency* postupně přednost pojmu **experience of agency**, definovaném velice podobně, jako *registraci, že já jsem iniciátorem svých akcí* [23]. Shodně s autory považují pojem SoA za příliš generický a variabilní mezi autory [22]. Vysoce heterogenní je však nejenom chápání pojmu jako takového, ale i jeho operacionalizace a experimentální testování. Zde se setkáváme s ukázkovým příkladem interdisciplinárního výzkumu nezbytně závislého na spolupráci experimentálních vědců a filosofů. Bez ujasnění, kdo čím co míní, se experimentální výzkum vědomí neposune za hranice neporozumění a vzájemné neporovnatelnosti výsledků. Synofzik a Vosgerau proto na základě analýzy implicitního chápání SoA ve stávajících pracích navrhuji jemnější dělení *pocitu jednání* na *Feeling of Agency* a *Judgment of Agency* (viz **Obrázek 2**). Tyto dva pojmy představují dva různé aspekty SoA. **Feeling of agency** (FoA) je předpojmový, nízkourovňový pocit bytí agentem dotyčné akce. Subjektivní pocit kontroly akce. **Judgment of agency** (JoA) je *soud o jednání*, tedy explicitní, konceptuální, interpretativní soud o tom, že jsem agentem. Běžně je sice FoA dostatečným pro následnou přítomnost JoA, ale explicitní JoA vyžaduje dodatečnou kognitivní funkci monitorování výsledku akce. Jako takový je tedy citlivý vůči post-hoc faktorům, které mohou „zpětně“ ovlivňovat pocit prožitého. Debata o stalinistickém a orwellovském vědomí, kognitivních biasech (např. hindsight bias), přepisování vzpomínek či klamech vědomí (flash-lag efekt, color phi, intentional binding atd.) by však dalekosáhle přesáhl rámec tohoto výkladu. Experimentální výzkum ale ukázal, že vnímané časové umístění prožitku *awareness of intention* (viz dále) se zdá být plně ustaveno až po vykonání odpovídající akce a může být manipulováno pomocí TMS pulzu aplikovaného i několik set milisekund *po* vykonání samotné akce [15].



**Obrázek 1:** Představa neurokognitivního komparátorového modelu, jakožto mechanismu pocitu jednání. Zpětnovazební kontrolní smyčka počítá motorickou odchylku prostřednictvím porovnání požadovaného stavu a aktuálního odhadovaného stavu. Výsledek vstupuje zpět do systému a slouží ke zlepšení jeho funkce (komparátor 1). Na základě daných motorických pokynů systém predikuje výstup agentova chování („predikovaný stav“). Takové predikce mohou být použity pro feed-forward motorickou kontrolu („central error monitoring“) pohybu, která probíhá ještě před jakýmkoliv sensorickým feedbackem (komparátor 2). Navíc mohou být predikce použity k vyrušení nebo oslabení sensorické zpětné vazby vytvořené systémem samotným (komparátor 3)



**Obrázek 2:** Synofzik a Vosgerau argumentují proti klasickému komparátorovému modelu jednání (comparator model of agency). Dvojstupňový model jednání, který navrhují [22], popisuje jak explicitní, tak implicitní faktory SoA, reprezentované Judgment of Agency (JoA), respektive Feeling of Agency (FoA). Obrázek převzat ze [22]

#### 4 Prožitek záměru

V předchozí sekci jsme se zabývali širokou a pojmově neustálenou rodinou *prožitek jednání*. Literatuře o sense of agency je velmi blízké téma *vědomého záměru* (conscious intention), spřízněné také s fenoménem *uvědomění si intence* (awareness of intention, conscious awareness of intention). SoA zkoumá obecně *prožitek jednání*, ať již chápaném jako mentální pokyn k akci, přítomný před nebo během provedení akce, nebo naopak retrospektivně vytvořený „příběh o nás jako původci jednání“, držící pohromadě naše narativní já. *Prožitek* či *uvědomění si záměru* si všímá aspektu agence souvisejícího s intencí. *Prožitek záměru* je zpravidla chápán v kontextu *motorických intencí*, tj. záměrů uplatňovaných v motorickém jednání, které jsou experimentálně testovatelné. Tento prožitek bychom ale měli chápat jako obecnější, nevztahující se výlučně k motorickému typu jednání. *Prožitek jednání* u *mentální agence* je aktuálně kontroverzní téma.

Zůstává otázkou, zda vůbec *prožitek záměru* jako fenomenální úkaz existuje, nebo má spíše podobu *retrospektivního soudu*. Odpověď bude záviset jak na vymezení



prožitku jednání vůči prožitku záměru (Sekce 3.1.4), tak na pohledu na časovou souslednost záměru a aktu (Sekce 6).

#### 4.1 Awareness of Intention

Nejlépe popsáním druhem prožitku záměru je *Awareness of Intention* (AoI). V obecné rovině mu můžeme rozumět jako jevu *vědomé zkušenosti vlastního záměru* (conscious experience of having an intention)<sup>2</sup>. Ve specifickém kontextu *motorického jednání* (motor agency) bychom AoI měli chápat jako *vědomý prožitek (bytí si vědom) vlastního záměru vykonat motorickou akci*. S pojmem jsem se setkal poprvé u Patricka Haggarda [9], aby byl dále použit množstvím autorů, včetně zejména JOHNEM-DYLANEM HAYNESEM [21] a HAKWANEM LAU. Poslední zmíněný definuje AoI jako *vědomé uchopení akce, kterou se agent chystá provést* [14]. Klasicky je *uvědomění* (awareness) považováno za substanciální a podstatnou kvalitu *intence*, ačkoliv tento postoj je v současné době předmětem sporu v literatuře o nevědomých intencích. AoI se dle mého názoru vzhledem ke svým metodologickým předpokladům řadí do prediktivních paradigmat. **Obrázek 3** ukazuje nejvýznamnější časové události při volním motorickém úkonu, relevantní vůči AoI.

##### 4.1.1 Neurovědné koreláty AoI

V souvislosti s pracemi P. Haggarda panuje názor, že na půdě neurověd odpovídá AoI *okamžiku W* z úlohy na *lateralizovaný readiness potential* (LRP). Haggard tvrdí, že jeho experimenty na LRP ukazují, že *W* se odehrává bezpochyby až když je připraven *kompletní motorický plán*. To ve svém důsledku vylučuje Libetovo „*veto*“ ve prospěch Haggardovy „*specificity*“, která předpokládá mnohem menší rozsah volního zásahu do běžícího plánu [11].

## 5 Vztah mezi Awareness of Intention a Sense of Agency

Pokud se mi podařilo zjistit, vztah mezi *sense of agency* a *awareness of intention* nebyl doposud v literatuře explicitně rozpracován a je těžké nalézt nějaký jednotně uznávaný implicitní princip. Tento problém jistě souvisí s celkově neujednoceným pojmoslovím v této doméně, jak je naznačeno výše v **Sekci 2**. Jádrem vymezení je to, že SoA se týká *jednání* (agence), kdežto AoI *prožitku záměru* (intence). Předběžná definice, kterou zde předkládám, pouka-

<sup>2</sup>Zůstává zde filosofická otázka, nakolik je možno ztotožňovat *vědomý prožitek* (conscious experience) a *bytí si vědom* (awareness). Zatímco první se zdá být případem vědomí prvního řádu, awareness bude patrně typicky definována jako reflexivní, tedy náležící do vědomí druhého řádu, meta-vědomí.

zuje k některým odlišným fenomenálním charakteristikám výše zmíněných jevů.

**Sense of Agency** chápou jako generický pojem reprezentující spíše širokou rodinu rozličných forem *zkušenosti jednání*, spojených spíše s motorickým jednáním. **Awareness of intention** se zabývá specifitěji tím, jak povstává *uvědomění intence, tedy situace, kdy si člověk uvědomuje svůj vlastní záměr, je si vědom svého záměru* (conscious awareness of an intention). Vztahuje se obecně k *intencím*, které neodkazují vždy nutně k *motorickým akcím*. Ačkoliv v některých případech oba tyto koncepty odkazují ke stejné zkušenosti (např. u Libetových experimentů s mačkáním tlačítka), v jiných situacích to tak není (např. rozhodnutí přestat kouřit). Z hlediska fenomenálního aspektu se AoI překrývá s *pocitem jednání* (feeling of agency) ale na rozdíl od něj se nevztahuje jen na *motorické intence*. Tento text však uvažuje zejména motorické intence.

Vztahem mezi *prožitkem jednání* a *prožitkem intence* u motorického jednání se zabývá také Haggard [8]. Podle něj může být *prožitek jednání* (experience of agency) vyvolán jak *prožitkem záměru* (experience of intention), tak *volní akcí* samotnou. To pak pocítujeme, jakoby naše volní jednání způsobovalo specifické události ve vnějším světě.

*Prožitek jednání* tedy nutně zahrnuje prožitek externích smyslových důsledků (tj. vnímanou změnu působení na sensorické orgány, přisuzovanou vlastní akci), kdežto *záměr* se poutá blížeji s *přípravou* a *úsilím* (soustředěním). Tento Haggardův postřeh považuji za velmi koncentrovaný vhled do tohoto zapeklitého problému.

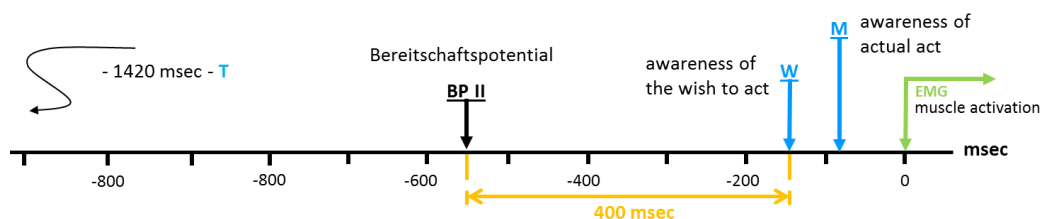
### 5.1 Neurovědné souvislosti

Neurální základ prožitku jednání zůstává nejasný. Explicitní přiřazování sensorických efektů, jako výsledku vlastní akce, či akce cizího agenta, zahrnuje parietální kortex. Avšak normální, implicitní *pocit vlastního jednání* může zahrnovat predikce vytvořené na základě intencí, které jsou vytvářeny ve *frontálních motorických oblastech*, a ještě dříve v *prefrontálních* (LPFC, MPFC) a také *parietálních oblastech* (Precuneus/PCC) [21].

### 5.2 Časové horizonty vědomých intencí a jejich prožitkové formy

Nejen podle Haggarda je volní jednání spojeno s dvěma (či více) poměrně odlišnými formami vědomé zkušenosti: myšlenky „*dlouhého dosahu*“, týkající se budoucích akcí, a okamžité pocity „*chystání se něco udělat*“ [8].

**Intence dlouhého dosahu** připomínají mentální plánování a vědomé rozvažování a jsou spojeny s perspektivní pamětí. Postrádají také konkrétní motorickou



**Obrázek 3:** Časová osa motorického jednání. BP II, W a M, zjištěné v Libetových experimentech [16], značí okamžik detekovatelnosti specifického EEG potenciálu BP II, okamžik nejčasnějšího *uvědomění intence* pohnout prstem (odpovídá awareness of intention) a okamžik uvědomění si již uskutečňovaného pohybu. Mozková aktivace BP II je detekovatelná až 400 ms před okamžikem W (AoI). T je okamžik *uvědomění intence* (AoI) při použití jiné metody snímání mentálního časování událostí (probe-testu) [17]. Pověšněte si, že T výrazně předchází ostatní časové body a leží mimo zobrazenou část časové osy, což by zpochybňovalo hypotézu o nevědomých iniciátorech našeho jednání. Otázkou však zůstává, zda obě metody testují stejnou úroveň (typ) vědomí. Upraveno dle schématu z přednášek I. M. Havla

složku. Zahrnují spíše mentální „cestování časem“ a proces časového plánování, podobný tomu, když plánujeme co budeme nyní dělat. Současný výzkum *intencí dlouhého dosahu* se zaměřuje na rozpoznání mechanismů v mozku, zodpovědných za prospektivní paměť.

**Intence krátkého dosahu** jsou blížeji k tématu *volního jednání a volní akce*. V literatuře se o nich často dočteme jako o nutkáních („urges“), přestože někteří vědci považují jako vhodnější pojem „*vědomé intence*“. Tento prožitek *vědomých intencí* v posledních letech přilákal výrazný zájem výzkumníků. Přesto však na sebe nechaly první pokusy o fenomenologii těchto jevů čekat dlouho a objevili až relativně nedávno [19].

### 5.2.1 Fenomenální aspekty prožitku intence

Fenomenální obsah vědomých intencí má nejméně tři aspekty: *časování*, *specifičnost efektoru* (tj. prostředku jednání) a *intenzity* [8]. Vědomé intence se zdají odehrávat během poměrně úzkého časového okna cca 1 s před zahájením motorického vykonání pohybu [16]. Zahrnují konkrétní detaily o zapojené části těla a o způsobu chystaného pohybu.

Prožitek intence je také jevem jasně stupňovitým: stěží si můžeme být vědomi přesného procesu *dělání dalšího kroku při chůzi*, ale *mačkání kohoutku pistole* si budeme vědomi velice intenzivně. Tento *rozměr intenzity* prožívání intence může být však jen stěží experimentálně zkoumán [8].

### 5.2.2 Distální, proximální a motorické intence

Propracovaný systém časových aspektů a horizontů vědomých intencí má ELISBATEH PACHERIE [19]. Konceptuální rámec, který navrhuje, pracuje s trojstupňovou

distinkcí intence, dle jejich časového začlenění, rozvinutosti a fenomenálních atributů. Rozlišuje tak *distální intence*, *proximální intence* a *motorické intence* (D-intence, P-intence, M-intence). Zjednodušeně lze říci, že rozumová kontrola se nalézá na úrovni D-intencí, situační kontrola na úrovni P-intencí a motorická kontrola na úrovni M-intencí. Tomu odpovídají i tři úrovně specifikace prostředků realizace. Na úrovni D-intencí jsou prostředky typicky reprezentovány podcíli a podakcemi. Na úrovni P-intencí jako pohyby určitého typu a konečně na úrovni M-intencí se jedná o plně specifikované pohyby.

Aktuální neurovědecký výzkum [21] se pochopitelně pokouší vysvětlit způsob, jak se *intence dlouhého dosahu* (které bychom přiřadili asi k D-, nebo P-intencím) mění v *intention-in-action* (ztotožnitelné s M-intencemi).

### 5.3 Je AoI jevem vědomí, nebo metavědomí?

Zabýváme-li se otázkou vědomých prožitků, jejich vnitřní dynamiky a struktury, nevyhneme se alespoň dílčím pohledem na vědomí jako takové. I experimentální vědci v oboru svobodné vůle a motorických intencí poznávají, že další výzkum je stěží proveditelný bez začlenění filosofie. Pozoruhodný je právě pohled na různé úrovně vědomí, kde dochází ke spontánní konvergenci moderních filosofických teorií [1] a aktuálních neurovědných objevů [17]. Filosof Ned Block tvrdí, že vědomí je nepřesný název pro konglomerát dvou fenoménů. Navrhl proto rozdělit pojem *vědomí* na *P-consciousness* a *A-consciousness*. **P-vědomí** (phenomenal consciousness) je podle Blocka syrovým prožitkem. Je to fenomenální složka vědomí - zvuky, pohyb, barevné formy, emoce a pocity, s naším tělem jako centrálním prvkem. Tato rovina - rovina kvalit - je prekonceptuální. **A-vědomí** (access consciousness) je naopak rovina intencionální, kdy jsou informace v naší mysli

k dispozici (přístupné) pro verbální reporty, usuzování a řízení chování. Jestliže vnímáme, tak informace o tom *co* vnímáme je součástí *přístupového vědomí*. A-vědomí má tedy charakter *meta-vědomí*, kdežto P-vědomí je *vědomím prvního řádu*. Zdá se, že Chalmersův *těžký problém* se proto vztahuje právě na P-vědomí.

Experimentální výzkum prožitku jednání a svobodné vůle nyní stojí před stejnou výzvou. Tým vědců MASUO MATSUHASHIHO a MARKA HALLETTA totiž před několika roky vytvořili novou metodu testování časového určení vědomé intence. Takovou, která nebude trpět mnohými neduhy, za které byl Libetův design využívající Wundtových hodin těžce kritizován, zejména od filosofů a metodologů. Přesto se stal v disciplíně jakýmsi „industry-standardem“. Cílem Matsushahiho a Halletta nebylo testovat prožitek intence *po* vykonání akce, ale *během* ní, aby se zamezilo možným retrospektivním cenzurním zásahům vědomí, o kterých mluví např. Dennett. Tito dva vědci použili stop-signal paradigmatu a na základě této odlišné metody měření stejného fenoménu dospěli ke zcela odlišným výsledkům. Okamžik prožitku vědomého rozhodnutí se nejenom neopozdoval o cca 400 ms za EEG komponentou zvanou *Bereitschaftpotential*, ale naopak ji cca o 700 ms předcházela [17]. Autoři samotní výsledek neinterpretovali příliš radikálně, zdá se však, že (snad bezděky) testovali jiný druh (úroveň) vědomí. Zatímco většina experimentů používá retrospektivní dotazování, tedy vlastně měří Judgment of Agency, kdy se ptáme na *přístupové vědomí*, oni se snažili zprostředkovaně testovat v reálném čase přímo Feling of Agency nebo Awareness of Intention - tedy jevy *fenomenálního vědomí*.

## 6 Dynamika prožitku jednání a prožitku záměru

Zatímco některé přístupy zdůrazňují prediktivní procesy jako stěžejní mechanismus (např. *comparator model* C. Fritha), jiné zdůrazňují význam retrospektivní inferencí (např. postdiktivní model Daniela Wegnera, založený na post-hoc inferenci).

### 6.1 Prožitek jednání jako retrospektivní konstrukt

Někteří autoři zastávají radikální náhled, že vědomá intence vlastně vůbec není mentálním stavem, ale spíše inferencí, která je retrospektivně a antedatovaně vsunuta do proudu vědomí, jako hypotetická příčina fyzického pohybu našeho těla. Naše akce jsou tedy iniciovány nevědomě na neuronální úrovni a k pocitu awareness of intention je odkazováno retrospektivně, až po tom, co byla provedena odpovídající akce [24, 25, 3]. Vědomá intence se pohnout je podle nich součástí psychologického narativu, kteří si my

všichni vytváříme, abychom vysvětlili, proč se naše ruka znenadání dala do pohybu.

Podobný efekt, kdy události proběhlé po akci také přispívají a ovlivňují prožitek intence, byl experimentálně prokázán. Podpora přichází ze studií psychotických onemocnění, kdy bývá prožitek intence asociován s velmi neobvyklými kauzálními vysvětleními o propojenosti mezi událostmi (blud kontroly, paranoia). Efekt se však projevuje i u zdravých osob, při specifických experimentálních úlohách, kdy se ukazuje, že důsledek akce dokáže silně ovlivnit prožitek akce samotné (dle [8]). Tento efekt je obzvláště silný v případech chyb při akci, kdy zpětná vazba přináší informaci o neočekávaných důsledcích akce. Další ze studií, které naznačují, že alespoň některé složky AoI jsou tvořeny post-hoc způsobem, tedy po dokončení motorické akce, je TMS experiment Hakwana Lau [15].

Jak si ale všímá Haggard [8], nejlepší důkaz proti „konfabulacionismu“ retrospektivních interpretací pochází ze studie ITHZAKA FRIEDA, kdy byly motorické oblasti stimulovány sadou elektrod [4]. Když Fried a kolegové stimulovali oblast *preSMA*, pacienti reportovali „nutkání“ (‘an urge’) pohnout specifickou částí těla kontralaterálně. Vzhledem k tomu, že se pacienti ve skutečnosti fyzicky nepohnuli, *prožitek intence* (nebo nutkání) nemůže být (pouze) retrospektivním konstruktem. Na druhou stranu ale pro *prožitek jednání* není nutná předchozí přítomnost intence nebo uvědomění záměru [19, :203].

### 6.2 Prožitek jednání jako prospektivní proces

Hlavním zástupcem prospektivních modelů je *comparator model*, rozvíjený zejména Frithem a Blakemoorovou [6]. Charakteristika modelu výše v 3.1.4.

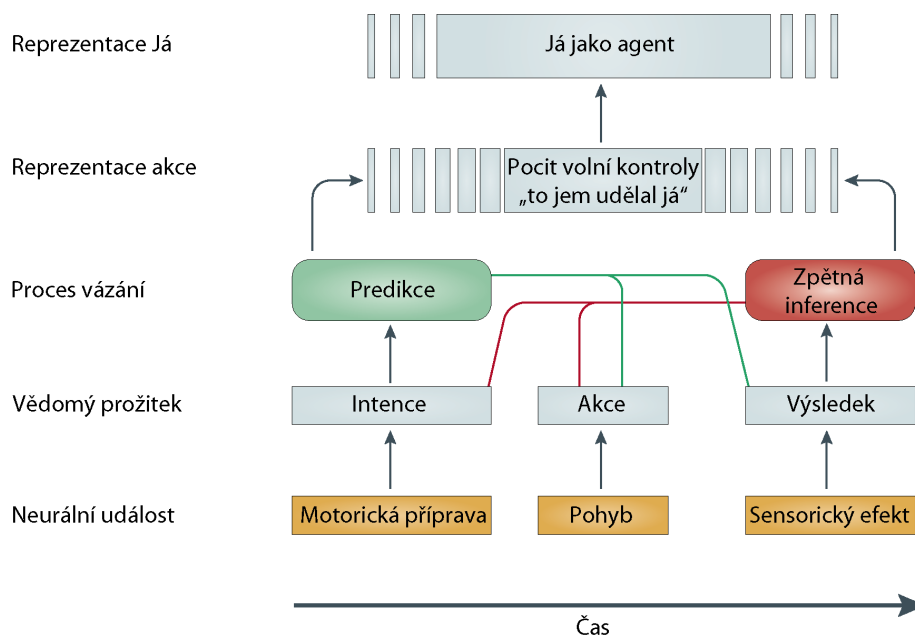
### 6.3 Prožitek jednání jako integrativní proces

#### 6.3.1 Framework pro uvědomování si akce

Se zajímavým integrativním modelem, sjednocujícím prospektivní a retrospektivní vlivy, přichází PATRICK HAGGARD. Jeho pětiúrovňový *framework pro uvědomování si akce* (Framework For Action Awareness, **Obrázek 4** na následující straně) popisuje intenci a akt jednání v čase a to současně na různých explanačních rovinách. Setkáváme se tedy s pokusem sjednotit do jednoho systému různé aspekty aktu - od nejnižšího neurálního, až po ten nejvyšší, na úrovni plně konceptuálního, narativního Já.

**První úroveň - neurální procesy:** Nervové signály, podléající se na prožitku volního jednání, zahrnují pokročilou přípravu akce, reaférentní somatosenzorickou zpětnou vazbu a senzorickou informaci o důsledcích akce.

**Druhá úroveň - vědomý prožitek:** Oddělené prožitky, zapříčiněné těmito signály, jsou zde synteti-



**Obrázek 4:** Haggardův model usiluje o vysvětlení prožitku jednání skrze integraci kognitivních procesů různých úrovní. Upraveno dle [8]

zovány, aby zformovaly celkový dojem akce jako jedné události.

**Třetí úroveň - vazební procesy:** K této syntéze přispívají dva specifické kognitivní procesy. Za prvé intence predikují akce a cílový efekt, ke kterému se vztahují (zelené linky). Za druhé smyslový prožitek akce a jejích důsledků spouští zpětnou re-konstrukci *záměru jednat* (červená linka). Na třetí úrovni se tedy uplatňují jak prediktivní, tak postdiktivní faktory.

**Čtvrtá úroveň - reprezentace akce:** Procesy na této úrovni k sobě vážou rozličné složky zkušenosti, aby vytvořili prožitek a uvědomění si akce jako celku, stlačeného v čase vůči podkladným neurálním procesům.

**Pátá úroveň - reprezentace Já:** Toto dává dále vzniknout reprezentaci sebe jako agenta, schopného plynulé volní kontroly [8].

Ke korelaci mezi rovinami potom dochází proto, že jak *vědomý záměr jednat*, tak *provedení akce* samotné, jsou hnány stejnou příčinou, a to neurální přípravou akce. Podle Haggarda vstupuje vědomí do „hry“ příliš pozdě na skutečnou kontrolu akce. Jeho vysvětlení fenomenální složky jednání je takové, že důvodem přítomnosti doprovodného vědomého prožitku je vytvoření silných paměťových stop „pro příště“. Pokud bych příště pocíval takový fenomenální prožitek, který minule „vedl“ (kon-

strukce kauzality a vlastního vlivu) k následnému negativnímu prožitku, budu si toho velice dobře vědom a akci zastavím, dokud je to ještě možné. Tento pokus „vysvětlit“ význam fenomenální roviny patrně nebude bezproblémový, přesto bych však byl dalek podezírání Haggarda ze „sprostého“ eliminativního materialismu.

## 7 Závěr

V této analýze *fenomenálních aspektů volního jednání*, postavené na aktuálních teoriích autorů S. Gallaghery, E. Pacherie, P. Haggarda, M. Synofzika a D. Vosgerau jsem se pokusil nastínit kontury současného studia *prožitku intence* a souvisejícího *prožitku jednání*. Zůstává však řada témat spřízněných této diskusi, kterých jsme se nedotkli. Namátkou jsou to sense of error, pocit plynulosti při jednání, či sense of initiation. Slibné je také téma mentální agence, kterému se plánuji věnovat v jiném článku. Tématem metodologie časového určení vědomých prožitků, kterého jsem se zde také dotkl, se zabývám v paralelně připravovaném článku, který vyjde v nejbližších měsících ve filosofickém žurnálu E-Logos.

## Reference

- [1] Block, N.: Consciousness, accessibility, and the mesh between psychology and neuroscience. *The Behavioral and Brain Sciences* 30(5-6), 481–99; discussion 499–548 (Dec 2007) 5.3
- [2] Bode, S., He, A.H., Soon, C.S., Trampel, R., Turner, R., Haynes, J.D.: Tracking the unconscious generation of free decisions using ultra-high field fMRI. *PloS one* 6(6), e21612 (Jan 2011) 1.1
- [3] Dennett, D., Kinsbourne, M.: Time and the Observer: the Where and When of Consciousness in the Brain. *Behavioral and Brain Sciences* XV, 183–247 (1992) 6.1
- [4] Fried, I., Katz, A., McCarthy, G., Sass, K., Williamson, P., Spencer, S.: Functional Organization of Human Supplementary Studied by Electrical Stimulation Motor Cortex. *Journal of Neuroscience* 11(11), 3656–3666 (1991) 6.1
- [5] Frith, C.D., Blakemore, S., Wolpert, D.M.: Explaining the symptoms of schizophrenia: abnormalities in the awareness of action. *Brain research. Brain research reviews* 31(2-3), 357–63 (Mar 2000) 3.1.4
- [6] Frith, C.D., Blakemore, S.J., Wolpert, D.M.: Abnormalities in the awareness and control of action. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 355(1404), 1771–88 (Dec 2000) 6.2
- [7] Gallagher, S.: Philosophical conceptions of the self: implications for cognitive science. *Trends in Cognitive Sciences* 4(1), 14–21 (2000) 3.1, 3.1.1, 3.1.3, 3.1.4
- [8] Haggard, P.: Human volition: towards a neuroscience of will. *Nature reviews. Neuroscience* 9(12), 934–46 (Dec 2008) 5, 5.2, 5.2.1, 6.1, 4, 6.3.1
- [9] Haggard, P., Clark, S., Kalogeras, J.: Voluntary action and conscious awareness. *Nature Neuroscience* 5(4), 382–5 (Apr 2002) 4.1
- [10] Haggard, P., Eimer, M.: On the relation between brain potentials and the awareness of voluntary movements. *Experimental Brain Research* 126(1), 128–133 (1999) 1.1
- [11] Haggard, P., Libet, B.: Conscious Intention and Brain Activity. *Journal of Consciousness Studies* 8(11), 47–63 (2001) 4.1.1
- [12] Haggard, P., Tsakiris, M.: The Experience of Agency. *Current Directions in Psychological Science* 18(4), 242–246 (2009) 3.1
- [13] Koukolík, F.: Základy kognitivní, afektivní a sociální neurovědy XVII. Svobodná vůle. *Praktický lékař* 92(4), 191–197 (Jan 2012) 1.1
- [14] Lau, H., Maniscalco, B., Shaver, E.C., Graves, T.: Awareness of Intention (2012) 4.1
- [15] Lau, H.C., Rogers, R.D., Passingham, R.E.: Manipulating the experienced onset of intention after action execution. *Journal of Cognitive Neuroscience* 19(1), 81–90 (2007) 3.2, 6.1
- [16] Libet, B., Gleason, C.A., Wright, E.W., Pearl, D.K.: Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). The unconscious initiation of a freely voluntary act. *Brain* 106(3), 623–642 (1983) 1.1, 3, 5.2.1
- [17] Matsushashi, M., Hallett, M.: The timing of the conscious intention to move. *European Journal of Neuroscience* 28(11), 2344–51 (Dec 2008) 3, 5.3
- [18] Pacherie, E.: Towards a dynamic theory of intentions. In: Pockett, S., W.P.Banks, Gallagher, S. (eds.) *Does Consciousness Cause Behavior? An Investigation of the Nature of Volition*, pp. 145–167. Cambridge, MA (2006) 2
- [19] Pacherie, E.: The phenomenology of action: a conceptual framework. *Cognition* 107(1), 179–217 (Apr 2008) 5.2, 5.2.2, 6.1
- [20] Shoemaker, S.: Self-Reference and Self-Awareness. *The Journal of Philosophy* 65(19), 555–567 (1968) 3.1.1
- [21] Soon, C.S., Brass, M., Heinze, H.J., Haynes, J.D.: Unconscious determinants of free decisions in the human brain. *Nature Neuroscience* 11(5), 543–5 (May 2008) 1.1, 4.1, 5.1, 5.2.2
- [22] Synofzik, M., Vosgerau, G., Newen, A.: Beyond the comparator model: a multifactorial two-step account of agency. *Consciousness and cognition* 17(1), 219–39 (Mar 2008) 3.1.4, 3.2, 2
- [23] Synofzik, M., Vosgerau, G., Voss, M.: The experience of agency: an interplay between prediction and postdiction. *Frontiers in Psychology* 4(March), 1–8 (2013) 3.1.4, 3.2
- [24] Wegner, D.M.: The mind's best trick: how we experience conscious will. *Trends in Cognitive Sciences* 7(2), 65–69 (2003) 6.1
- [25] Wegner, D.M.: Self is Magic. In: Baer, J., Kaufmann, J., Baumeister, R.F. (eds.) *Psychology and Free Will*. Oxford University Press, New York (2006) 6.1



# Detekce patologií v B-MODE ultrazvukových obrazech s automatickou kontrolou kvality obrazu

Jiří Blahuta a Petr Čermák

Slezská univerzita v Opavě  
Filozoficko-přírodovědecká fakulta  
Bezručovo nám. 13, 74601 Opava  
jiri.blahuta@fpf.slu.cz, petr.cermak@fpf.slu.cz

## Abstrakt

Vyvinuli jsme aplikaci v prostředí MATLAB pro detekce patologií v ultrazvukových B-MODE obrazech založenou na ROI-based zpracování. Aplikace je díky povaze B-MODE obrazu univerzálně použitelná na obrazy ve škále šedi. Jednou ze součástí aplikace je automatické hodnocení kvality vstupního minimálního jasu. To slouží k tomu, aby ve výsledku nebyly vyhodnoceny patologie tam, kde ve skutečnosti nejsou. Hlavní zpracování je založeno na výpočtu obsahu v definovaném ROI pomocí binárního prahování pro všechny přípustné hodnoty prahu a výsledkem je klesající křivka obsahu, ze které je hodnocena patologie či fyziologie. Aplikace jsme otestovali na intrakraniálních sonografických obrazech mozkového kmene z transtemporálního okna k detekci parkinsonismu dle hodnocení echogenity v substantia nigra.

## 1 Význam detekce B-MODE obrazů

Zpracování obrazových medicínských dat z různých modalit, např. CT, RTG či ultrazvuk, je stále aktuální a potřebné téma. V této práci budeme prezentovat zpracování B-MODE<sup>1</sup> ultrazvukových obrazů, které jsou nativně ve škále šedi a dle intenzity šedi se hodnotí echogenita, která rozhoduje o patologii či fyziologii. B-MODE je dvojrozměrné zobrazení. Jednotlivým výchylkám odpovídají různě jasné body („B“ podle „Brightness“ = jas), sonda přitom vysílá ultrazvukové vlny v 1 rovině. Detaily k B-obrazu jsou dostupné v [3], [5], [6] a [9]. Tyto obrazy jsou základními obrazy v sonografii, existují i další typy, pomocí kterých můžeme například sledovat průtokovou cévu. My se budeme věnovat intrakraniálním B-MODE obrazům, konkrétně detekci příznaku parkinsonismu na základě hyperechogenity (vysoký jas) v oblasti substantia nigra (dále jen SN). Nejen v neurologii má ultrazvuk

nezastupitelné místo díky efektivitě vyšetření, diagnostický ultrazvuk<sup>2</sup> může být opakován bez následků radiační zátěže a je vhodnou modalitou pro měkké tkáně, jako je právě mozek. Více v [3] a [4].

Ve spolupráce s erudovaným neurosonologem jsme společně vyvinuli aplikaci k detekci patologie v B-MODE obrazech na základě měření obsahu ve zvolené malé oblasti ROI pomocí binárního prahování obrazu pro všechny prahy  $T \in (0; 255)$  a výsledkem je klesající křivka obsahu ROI dle prahu  $T$ . Aplikace byla řádně testována pro hodnocení parkinsonických příznaků hyperechogenity – byla provedena korelační a ROC analýza ke zjištění míry korelace a shody mezi 3 lidmi a mezi obrazy ze 2 různých ultrazvukových přístrojů.

Nedílnou součástí aplikace je automatické hodnocení kvality vstupního obrazu. Jedná se o fázi předzpracování obrazu. Ze vstupního obrazu vybereme (či je automaticky vybrána) oblast 50×50 mm dle nativní osy snímku. Kvalita obrazu je hodnocena pomocí agentového systému, který je samostatnou součástí aplikace a nemá nijak vliv na další zpracování. To slouží k tomu, že pokud je obraz příliš přejasněn vlivem nastavení sondy, aplikace vydá varování, že výsledné hodnocení echogenity nemusí být objektivní. Tuto situaci může také přivodit užití funkce *Gray Scale Map*<sup>3</sup>, která umožní přednastavit určitý rozsah intenzit šedi obrazu dle vyšetřované tkáně pro získání subjektivně optimálního rozlišení či zobrazení echogenity.

## 2 Vstupní obrazy pro aplikaci

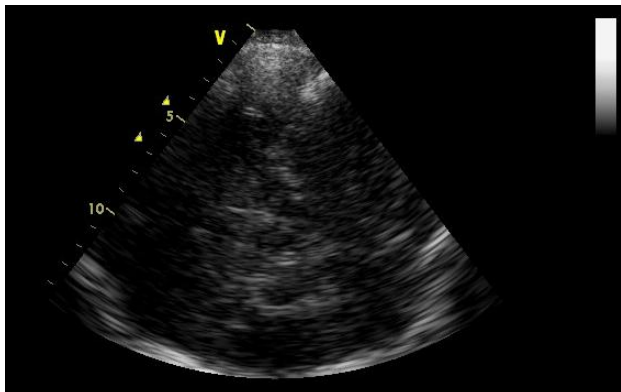
Jako vstupní obrazy, na kterých jsme aplikaci testovali, používáme obrazy mozkového kmene získané transtemporálním přístupem [3] sondy se zobrazením SN

<sup>1</sup> <http://medical-dictionary.thefreedictionary.com/B-mode>

<sup>2</sup> [http://www.solen.sk/index.php?page=pdf\\_view&pdf\\_id=2642&magazine\\_id=3](http://www.solen.sk/index.php?page=pdf_view&pdf_id=2642&magazine_id=3)

<sup>3</sup> <http://www.ncbi.nlm.nih.gov/pubmed/8521072>

k detekci parkinsonických příznaků hyperechogenity v SN. Aplikaci jsme také testovali na oblasti nuclei raphe.

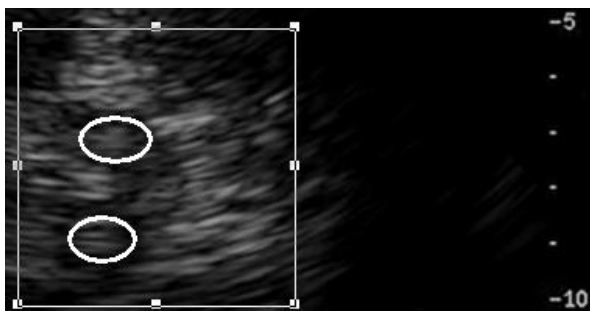


Obr. 1. Vstupní B-MODE obraz mozkového kmene se SN

Používáme obrazy v nativním formátu DICOM<sup>4</sup>, který je dnes informačně-technologickým standardem v oblasti zobrazování a zpracování RDG dat, či obrazy převedené do JPEG. DICOM soubor není pouze obrazem, ale nese spoustu metadat o modalitě, obrazu, jeho rozlišení, o pacientovi, atd. Taková metadata v bitmapových obrazech nemáme a do formátu DICOM dnes ukládají všechny přístroje. Z tohoto obrazu je automaticky či manuálně vyřezáno okno 50×50 mm dle nativní osy (žlutá vlevo) pomocí výpočtu eukleidovské vzdálenosti mezi 2 body  $A[a_1, a_2]$  a  $B[b_1, b_2]$  ve 2-D rovině.

$$d = \sqrt{(a_1 - b_1)^2}. \quad (1)$$

Následující obrázek ukazuje vyřezané okno (jiný obraz) s vyznačenou echogenní SN (elipsy). Podobná ROI budeme dále používat.



Obr. 2. Okno 50×50 mm se SN (patologické)

Výběr může být automatickým se zacílením okna dle použitého obrazu (obrazy z různých přístrojů) či manuálně výběrem bodu na ose a posunem okna. Následně je obraz (okno) převeden do škály šedi, vstup může být ve 24-bit, přestože platí  $R = G = B$ . Pak se převede do hodnot intenzity

$$I = 0,299 \times R + 0,587 \times G + 0,114 \times B, \quad (2)$$

kde intenzita  $I$  nabývá hodnot z intervalu  $\langle 0; 255 \rangle$ . Aplikace ověřuje, zda je vstup ve škále šedi či nikoli. Pro další zpracování tedy máme každý pixel vyjádřen jako 1 hodnotu šedi, nikoli vektorem RGB.

Následně již může být aplikována kontrola jasu vstupu. Více o samotné fázi předzpracování lze nalézt v [5] a [6].

### 3 Substantia nigra a její echogenita v B-obrazu k detekci Parkinsonovy choroby

Parkinsonova choroba (Morbus Parkinson; [3]) je degenerativní chronické onemocnění, jde o poruchu bazálních ganglií. SN hraje stěžejní roli v CNS a ovlivňuje koordinaci pohybů a produkuje důležitý dopamin. Porušení SN, tedy porucha produkce dopaminu, vede Parkinsonově chorobě (dále pouze PD; Parkinson's Disease). Získala název po lékaři Jamesi Parkinsonovi, který ji na začátku 19. století popsal. Na ultrazvukovém B-MODE obraze se detekuje pomocí míry echogenity (jasu) SN [3]. Ve stejné publikaci je uvedeno 5 tříd echogenity k detekci PD v B-obrazu.

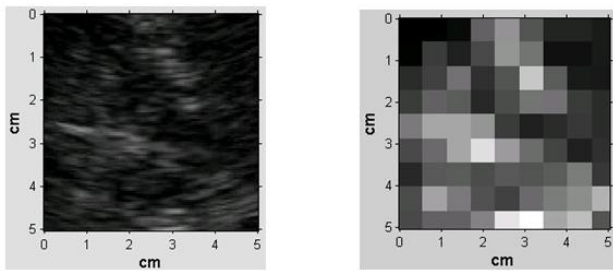
K detekci jsme vyvinuli algoritmus k spočtení hyperechogenního obsahu v SN [3], [4] a tím klasifikujeme patologii či fyziologii dle obsahu echogenních oblastí v potřebném ROI SN, viz Obr. 2.

### 4 Subsystem k posouzení kvality vstupního obrazu s použitím softwarových agentů

Po výřezu okna je aplikován subsystem, který na základě minimálního jasu vyhodnotí vstupní kvalitu obrazu. Jelikož je nesmyslné hodnotit intenzitu  $I$  (2) každého pixelu, tam se mohou vyskytnout ojedinělé skoky, rozdělíme obraz na bloky 8×8 mm, což lépe vypoví o celkové charakteristice jasu.

<sup>4</sup> poslání standardu DICOM na webu <http://medical.nema.org/dicom/geninfo/Strategy.pdf>





Obr. 3. Rozdělení okna na bloky 8×8 mm

Každé políčko v obraze představuje zástupnou hodnotu jasu  $I(2)$  získanou jako průměrnou hodnotu z pixelů v políčku.

#### 4.1 Inicializace agentového systému

Chápejme tuto část zpracování jako jednoduchý softwarový multiagentový systém, který je zcela autonomní, není nijak vázán na předchozí kroky – jedinou podmínkou je převedení obraz do šedi (1). Tento systém lze chápat jako multiagentový systém se 3 agenty, jak definujeme MAS jako množinu agentů ve svém prostředí:

$$MAS = \{Ag_1, Ag_2, \dots, Ag_n\}, \quad (3)$$

kde  $A_i$  je  $i$ -tý agent. V našem případě se jedná o softwarové agenty (softboty) [1] vymezené HW a operační platformou bez HW senzorů a aktuátorů [1]. Prostředí agenta je v tomto případě 2D obraz čili bitová mapa. Definice softwarového agenta může být různá, avšak stále z ní musí být patrný základ autonomie. Například definice v časopise *Automa*<sup>5</sup> uvádí následující definici: Softwarový agent je programová entita, jejíž chování je do jisté míry autonomní<sup>6</sup>. Každého agenta lze popsat agentovou funkcí<sup>6</sup>:

$$FA: V \rightarrow A, \quad (4)$$

Kde  $V$  je konečná množina vjemů a  $A$  je konečnou množinou akcí, které agent na základě vjemu vykoná. V našem případě detekce jasu můžeme funkci popsat takto pro každého agenta:

$Ag_1$  má vjemy  $V_1 = \{no\_echo, echo\}$  a akce  $A_1 = \{continue, warn\}$  a tedy funkce

$$okno(no\_echo) \rightarrow continue, \text{ či } okno(echo) \rightarrow warn. \quad (5)$$

A podobně pro  $Ag_2$  pouze v SN:

$Ag_1$  má vjemy  $V_2 = \{no\_echo, echo\}$  a akce  $A_2 = \{continue, warn\}$  a tedy funkce

$$SN(no\_echo) \rightarrow continue, \text{ resp. } SN(echo) \rightarrow warn. \quad (6)$$

Třetí agent pak srovnává shodu nálezu v levé a pravé SN. Důležitá je nezávislost tohoto subsystému na dalším i předchozím zpracování. Taktéž multiagentový systém můžeme definovat různými způsoby a existují různé formy MAS dle centralizace, hierarchie, atd., detaily ohledně zavedení pojmu agent a MAS lze nalézt v [1] či [2]. Komplexnější pohled na problematiku umělé inteligence ve spojitosti s autonomitou lze nalézt např. v [1], [9] či [10]. Pro náš případ uvažujme softwarové agenty jako autonomní výsek kódu. Máme 3 agenty:

- agent hodnotící jas globálně pro celé okno
- agent hodnotící jas v levé a pravé SN
- agent srovnávající hodnoty od druhého agenta

Mohl by tedy takto stejně pracovat i pro celý obraz. Samozřejmě ve vyřezaném okně bude rychlejší. Okolní obraz nás ani nezajímá, proto jsme na začátku vyřezali toto okno (Obr. 2.). Prostředím námi užitých softwarových agentů je 2D obraz.

Agent je implicitně inicializován do počátku souřadného systému obrazu, tedy levý horní roh uvažujeme jako počátek o souřadnicích [0, 0]. Cílem této detekce jasu je:

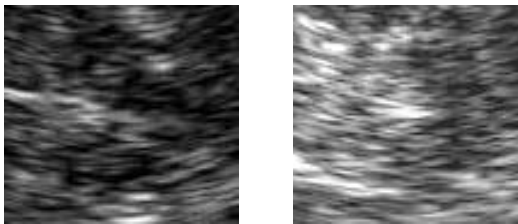
1. Zjistit hodnoty intenzity  $I$  ve všech blocích
2. Zjistit minimální hodnotu z těchto intenzit
3. Zjistit, zda toto minimum je  $> 25$

Tento poměrně jednoduchý autonomní algoritmus nám umožní detekovat, zda je obraz příliš přejasněný a tudíž může být nekvalitní, resp. nevhodný pro další zpracování. V normálně nastavených obrazech by se měla vyskytovat minimální hodnota prakticky černá, tj. 0 a blízké kladné okolí. U obrazů, kde je vysoký jas a černá je prakticky šedá, bude toto minimum vysoké. Následující obrázek ukazuje normálně nastavený obraz a přejasněný obraz nevhodný ke zpracování.

<sup>5</sup>

[http://www.odbornecasopisy.cz/index.php?id\\_document=38998](http://www.odbornecasopisy.cz/index.php?id_document=38998)

<sup>6</sup> <http://kti.mff.cuni.cz/~bartak/ui/lectures/lecture02.pdf>



Obr. 4. Normálně nastavený obraz (vlevo) a přejasněný (vpravo)

Popišme postup formálně. Rozdělili jsme obraz okna na bloky  $8 \times 8$ , získali jsme tedy matici se 64 prvky, kdy každý prvek matice představuje číselnou hodnotu  $I$  pro daný blok. Ta je dána průměrem z okolních pixelů, které náleží do bloku. Samozřejmě obrazová matice popisuje i celý obraz, ten je daný rozlišením. Matici s bloky je tedy

$$B = \begin{bmatrix} I(1) \\ \vdots \\ I(64) \end{bmatrix} \quad (7)$$

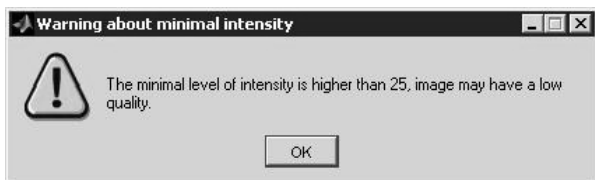
Krok 1. Je tedy zjištění daných intenzit bloků. Druhým krokem je zjištění minima z těchto intenzit, tedy

$$\min[I(1), I(2), \dots, I(64)]. \quad (8)$$

A posledním krokem je srovnání toho minima s prahovou hodnotou 25. Poté se aplikuje jednoduché pravidlo o rozhodnutí kvality obrazu

IF min > 25 THEN obraz\_nevhodny ELSE vhodny. (9)

Pravidlo říká, že pokud je minimální jas > 25, obraz není vhodný k dalšímu zpracování a výsledek měření nemusí být objektivní, mohou být detekovány defekty, tam kde ve skutečnosti nejsou a jsou rozpoznány pouze vlivem přejasnění obrazu, způsobeným např. aplikací funkce Gray Scale Map (viz kap. 1) či jiným faktorem. Tato prahová hodnota byla stanovena po konzultaci a praktickém měření se zkušeným neurosonologem vzhledem k vyšetřované tkáni SN a obecně povaze B-MODE obrazu. V případě minima > 25 program vypíše následující hlášku, která varuje před neobjektivním zpracováním.



Obr. 5. Je-li minimální jas vyšší než 25, jsme varováni

Praktický vliv přejasněného obrazu bude ukázán v následující kapitole, kde ukážeme hodnocení stejného obrazu při normálním nastavením jasu a přejasnění. Díky tomuto subsystému jsme schopni předcházet chybné detekci patologie vlivem nesprávné konfigurace jasu. Pokud bychom chtěli takové obrazy použít, museli bychom nastavit ruční konfigurace funkce Gray Scale Map a otestovat na každé examinované oblasti korektnost detekce.

## 4.2 Lokální detekce jasu v oblasti SN

Doposud jsme se zabývali globálním hodnocením jasu v celém výřezu. Máme však dalšího agenta, který detekuje jas pouze v oblasti SN. Cílení na oblast SN je automatické. Oblast je menší než celý výřez a proto u toho agenta namísto dělení bloků  $8 \times 8$  užíváme jemnější dělení na  $32 \times 32$  bloků, dohromady tedy 1024. Tato část je prvním zpracováním hodnocení echogenity před vlastním měřením. Tentýž agent bude mít 2 úlohy – provádí kontrolu jasu v levé a pravé SN (v obraze horní a dolní) a uloží si zjištěné hodnoty. Výpočet polohy SN v obraze odvodíme poměrně snadno z faktu, že pokud máme správně zacílené okno, nachází se SN prakticky ve stejných místech. Úkoly agenta k lokální detekci jasu:

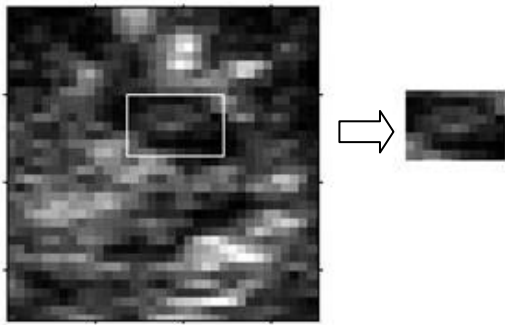
1. určit střed obrazu  $C[c_x, c_y]$
2. vypočítat polohu pro levou SN
3. vypočítat polohu pro pravou SN
4. v obou spočítat minimální jas
5. uložit hodnoty

Souřadnice středu okna  $C = [1/2x, 1/2y]$ . V praxi jsme zjistili nejvhodnější polohu levé i pravé SN. V MATLABu je tato situace zapsána jako ořez obrazu:

```
hsn = imcrop(relg2, [xc-40 yc-50 50
45]); % horni
dsn = imcrop(relg2, [xc-40 yc+60 50
45]); % spodni
```

V tomto kódu představuje *relg2* okno s  $32 \times 32$  bloky a jednotlivé souřadnice a velikost okna, viz Obr. 6.

Lokální agent má stejnou úlohu jako popsany agent pro globální detekci, avšak dostane od prvního agenta zprávu, pokud ten vyhodnotil obraz jako nepoužitelný. V tomto případě lokální detekce nepokračuje.



Obr. 6. Rozdělení bloků 32×32 pro lokální detekci průměrného jasu

Rozdělili jsme výstup toho agenta na 3 možné případy:

- Průměrná hodnota jasu < 20 – pacient je pravděpodobně zdravý
- Průměrná hodnota jasu 20 – 25 – pacient může být nemocný
- Průměrná hodnota jasu > 30 – pacient je pravděpodobně nemocný

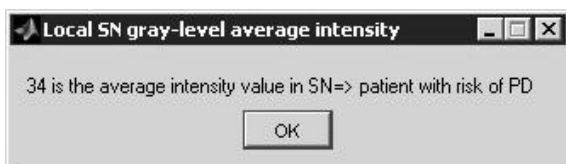
Implementace v MATLABu následovně:

```

if ag2 < 20
msgbox(int2str(ag2) 'is the average
intensity value in SN, No lesions with
high gray-level intensity, there is no
risk of PD')
elseif 20 < ag2 < 25
    msgbox(int2str(ag2) 'is the
average intensity value in SN =>
potential risk of PD')
else > 30
    msgbox('is the average intensity
value in SN => patient with risk of
PD')
end

```

Tento postup hodnocení echogenity je prvním krokem k odhalení potenciální patologie před vlastním měřením. Program upozorní v jednotlivých případech:

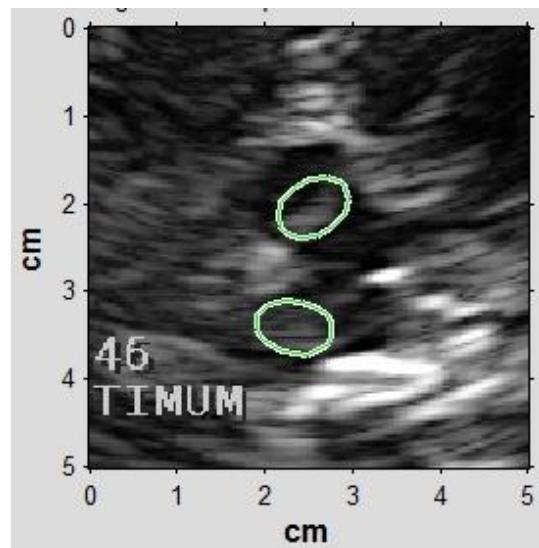


Obr. 7. Hodnocení echogenity SN agentem

#### 4.2.1 Srovnání lokálního jasu v SN

Po lokální detekci minimálního jasu v oblasti SN máme třetího agenta, který bude výsledek vyhodnocovat na základě předchozích detekcí. Zaujímá nás, zda se liší jas v levé a pravé SN (Obr. 8). Globální jas v tuto chvíli nehraje roli, protože pokud první agent zjistil překročení hranice jasu 25, druhý již nepokračuje a rovnou se obraz vyhodnotí jako nepoužitelný ke zpracování (Obr. 5). Tento třetí agent vyhodnocuje výsledky takto:

- je-li přibližně stejná úroveň jasu v obou SN, pak se jedná pravděpodobně o možnou patologii v SN
- je-li jas zcela rozdílný, není možné spolehlivě určit výsledek



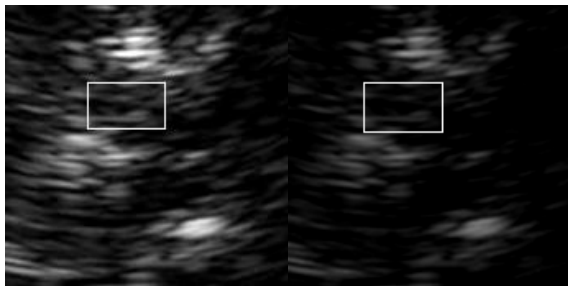
Obr. 8. Levá a pravá SN

Připustíme shodu  $\pm 10$ , v tomto případě považujeme nález za shodný. Přesnější interpretaci ukáže až samotné měření obsahu. I tak je však možné odhalit ve snímcích potenciální patologie či nesrovnalosti. Agent jas hodnotí stejně jako v předchozím případě dle vzorce (5) k detekci minimálního jasu.

#### 4.2 Experimentální problém ekvalizace histogramu

Nejen zvýšení úrovně jasu může mít fatální následek na hodnocení, ale také egalizace histogramu. V případě obrazů ve škále šedi, histogram vyjadřuje četnost intenzity  $I$  (2) v obraze pro každou intenzitu  $I \in \langle 0; 255 \rangle$ . Pro každou intenzitu je dána četnost výskytu pixelů s danou intenzitou, jejich suma je pak celý histogram.

Častou úpravou předzpracování obrazu je jeho ekvalizace neboli vyrovnání. Ideálně vyrovnaný histogram má stejný počet pixelů všech četností. Ekvalizace má často pozitivní vliv na kvalitu obrazu, srovná úrovně. Musíme si však uvědomit, že pro určitou třídu obrazů může být na škodu. Mezi tyto případy řadíme také ultrazvuková data, kde manuální úprava histogramu za účelem zviditelnění určitých struktur, může mít fatální následky na hodnocení. A proto i tento agentový systém je k tomu určen.



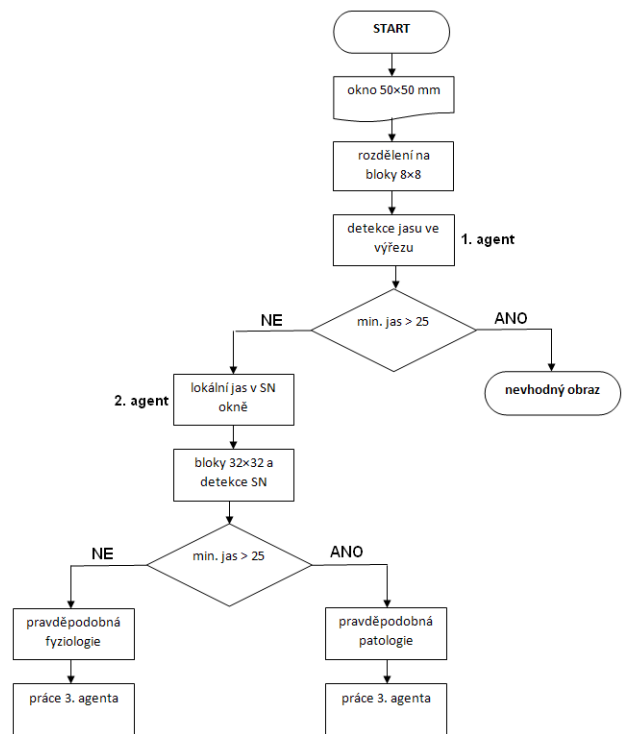
Obr. 9. Vliv ekvalizace histogramu na obraz

Je patrné, že v případě neupraveného obrazu (vlevo) bude SN vyhodnocena jako nevýznamně echogenní, zatímco v případě ekvalizace (vpravo) může dojít k falešné detekci patologie.

Konkrétní nastavení však musí srovnat erudovaný neurosonolog. Pokud bychom chtěli ekvalizace použít, musela by být součástí předzpracování. Mohlo to mít však také negativní následek – každý obraz může mít jinou úroveň původního jasu.

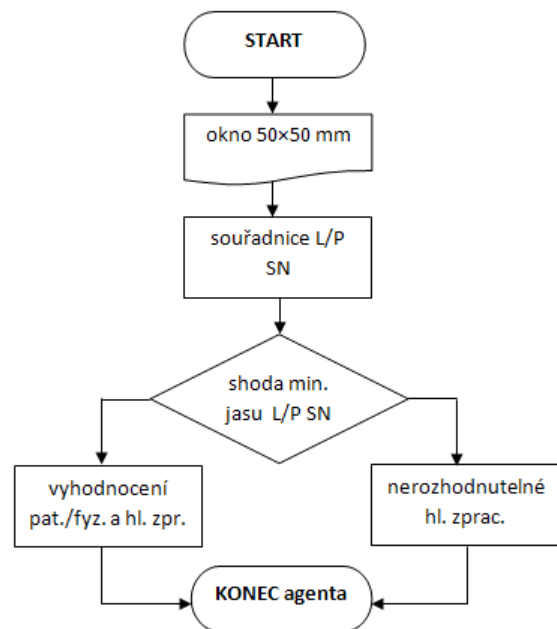
#### 4.3 Shrnutí přístupu agentového systému k detekci jasu vstupního obrazu

Následující diagram shrnuje postup užití agentového systému k hodnocení kvality vstupního obrazu. Předpokládáme již vyřezané okno 50×50 mm ze vstupního B-MODE obrazu. První 2 agenty řeší globální a lokální detekce minimálního zprůměrovaného jasu v blocích.



Obr. 10. Role agentů ke globální a lokální detekci jasu

A třetí agent hodnotí popsanou shodu mezi levou a pravou SN oblastí pro patologii i fyziologii.

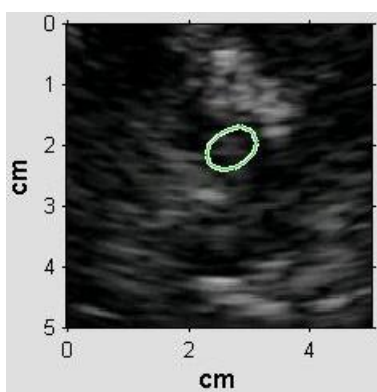


Obr. 11. Role třetího agenta k detekci shody jasu pro levou a pravou SN

## 5 Měření echogenity v ROI oblasti

Po detekci kvality obrazu přecházíme k hlavní fázi zpracování pomocí ROI-based processingu. Jedná se o hlavní fázi zpracování po zjištění kvality obrazu. Algoritmus je založen na principu binárního prahování<sup>7</sup> ve vyznačené oblasti ROI. Na Obr. 2 vidíme vyznačené eliptické ROI oblasti pro SN, podobně budeme používat v praxi.

Budeme používat elipsu s obsahem  $A = 50 \text{ mm}^2$  a sklonem  $60^\circ$  od  $x$ -ové osy. Dle (2) máme obraz ve škále šedi a to potřebujeme pro aplikaci principu binárního prahování. Zajímá nás dále pouze oblast uvnitř elipsy, okolí nikoli.



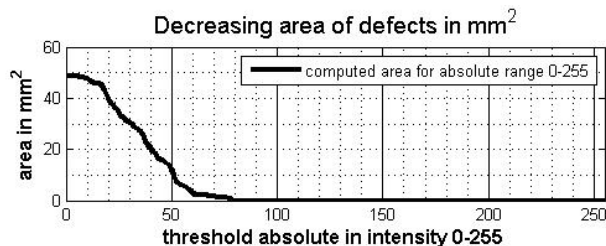
Obr. 11. Eliptické ROI pro měření echogenity v SN

Výstupem binárního prahování je binární obraz. Princip algoritmu spočívá v porovnání intenzity pixelu  $I \in \langle 0; 255 \rangle$  v obraze s prahem  $T \in \langle 0; 255 \rangle$  následovně:

$$\text{IF } I \geq T \text{ THEN } B = 1 \text{ ELSE } B = 0, \quad (10)$$

kde  $B$  je výstupní pixel s logickou boolean hodnotou 0/1. Tato operace srovnání je provedena pro všechny pixely obrazu. Pro náš případ je prahování provedeno pro všechny možné hodnoty  $T \in \langle 0; 255 \rangle$  a sledujeme počet bílých pixelů uvnitř ROI, které nám reprezentují obsah, tedy pro ty pixely, pro něž platí  $I \geq T$ . Výsledkem je klesající křivka obsahu ROI, tedy bílých pixelů  $B = 1$  v závislosti na zvyšujícím se  $T$ , čímž sledujeme echogenitu v daném ROI.

<sup>7</sup> <http://inperc.com/wiki/index.php?title=Thresholding>

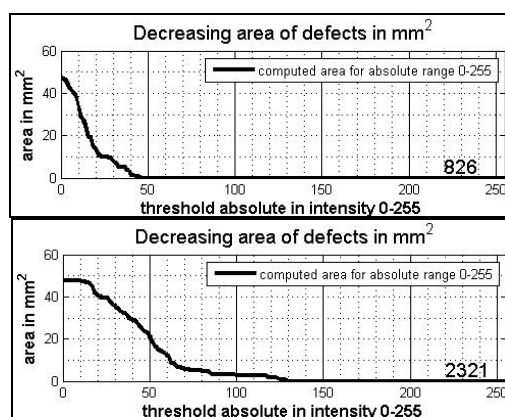
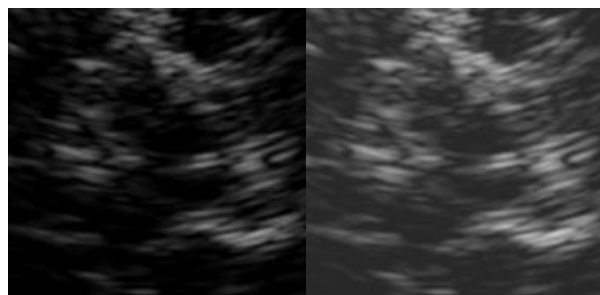


Obr. 11. Křivka s klesajícím obsahem v ROI v závislosti na prahu  $T$

V závislosti na klesající tendenci křivky, resp. rychlosti poklesu obsahu, stanovujeme patologii či fyziologii vyšetření. Tuto křivku budeme dále uvažovat jako vektor o 256 diskretních hodnotách.

### 5.1 Vliv přejasnění na výsledné hodnocení křivky

V souvislosti s popsáním algoritmu kontroly jasu v předchozí kapitole si ukažme příklad stejného obrazu s různým nastavením jasu, a jak se mění křivka obsahu v ROI, která je rozhodující k detekci patologie. Podobně jako na Obr. 4 uvažujme nyní stejný obraz s různým nastavením jasu a jeho výstupní křivku.



Obr. 12. Vliv přejasnění na výslednou křivku obsahu v ROI

Spočtené číslo vpravo dole v grafu představuje sumu spočtených obsahů pro všechny  $T \in \langle 0; 255 \rangle$  (Obr. 11).

$$\sum ap_i, \quad (11)$$

kde  $ap_i$  představuje vypočtený obsah pro  $i$ -tou úroveň prahu  $T$ .

Na tuto sumu pak navazuje statistická analýza s výpočtem referenčního percentilu k hodnocení fyziologie a patologie. Poté uvažujeme hodnoty jako vektor, viz kap. 5.2.

K zabránění této situaci slouží detekce globálního jasu ve výřezu. Tentýž obraz, který je normálně fyziologickým nálezem, by mohl být nesprávně vyhodnocen jako patologický případ, což je nežádoucí a aplikace by pak neměla žádný význam.

## 5.2 Statistická analýza výsledků testování aplikace

Pro posouzení objektivního hodnocení byl program testován:

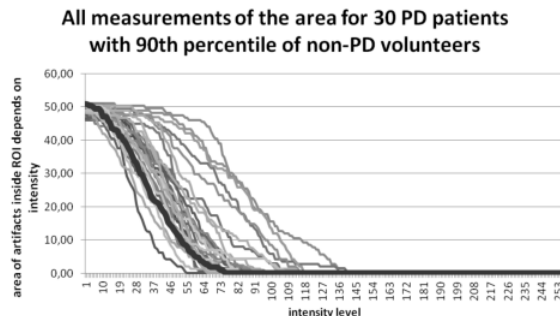
- 3 pozorovatelé, z toho 1 neurosonologický expert
- na reálných datech ze 2 různých přístrojů a korelace mezi výsledky
- třemi průchody měření u 1 pozorovatele

Toto testování bylo provedeno pro oblast SN a podrobné výsledky jsou dostupné v [5] a [6]. Dosáhli jsme vysokých korelačních koeficientů a kappa-koeficientů:

$$diff = \sum_{i=0}^{255} (ap_i - ar_i), \quad (12)$$

Pro účely testování byla vytvořena referenční křivka 90. percentilu vypočtená z průměrných měření zdravých dobrovolníků. Křivku obsahu (Obr. 8) chápeme jako vektor 256 hodnot a referenční percentilovou křivku taktéž. K hodnocení patologie či fyziologie byla použita metodika hodnocení na základě spočtení difference mezi vektorem naměřených hodnot a 90. percentilem:

Byli testováni také zdraví dobrovolníci, na nichž se také ukázala správná funkčnost měření, a výsledek křivky pouze u 5 % případů neodpovídal fyziologii. Shoda mezi měřeními stejných obrazů ze 2 přístrojů byla dosažena  $r > 0,88$ . Následující obrázek ukazuje výsledek 30 patologických pacientů s referenční křivkou pro fyziologii.



Obr. 13. Testování 30 parkinsoniků s hyperechogenní SN

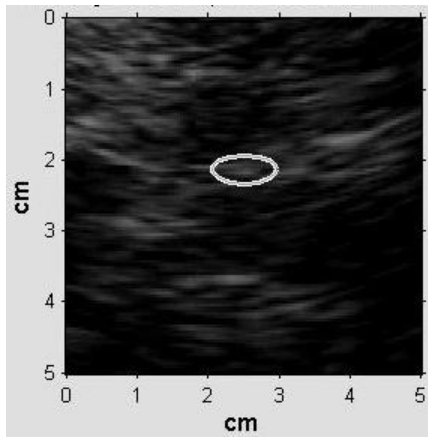
Několik vybraných výsledků statistické analýzy:

- průměrný korelační koeficient mezi obrazy z levého transtemporálního okna pro ESAOTE  $r = 0.895$
- průměrný korelační koeficient mezi obrazy z levého transtemporálního okna pro GE  $r = 0.88$
- pro GE přístroj průměrný kappa koeficient ze všech obrazů levého i pravého okna  $\kappa = 0.854$
- pro obrazy levého okna z přístrojů GE a ESAOTE je průměrný kappa koeficient  $\kappa = 0.943$

Analýza korelace a kappa koeficientů nám pomohla ke sledování spolehlivosti algoritmu v rámci měření mezi různými snímky, přístroji a lidmi. Lékař potvrdil, že program je reprodukovatelný a tato skutečnost byla publikována ve (WSEAS).

## 5.3 Měření echogenity nuclei raphe

V současnosti aplikaci testujeme pro jinou oblast, pro oblast NR, která má vliv na multisystémovou atrofii a deprese. Také zde je velmi důležité posouzení celkového jasu obrazu. Pro NR však platí opačné hodnocení – zcela černé ROI je zde patologické a opačně. Pro NR používáme menší ROI s obsahem  $A = 30 \text{ mm}^2$ . Dosažené výsledky se zatím zdají být dobré, včetně korelace opět ze 2 přístrojů. Správnost měření bylo potvrzeno neurosonologickým expertem. Pro NR používáme jiné ROI:



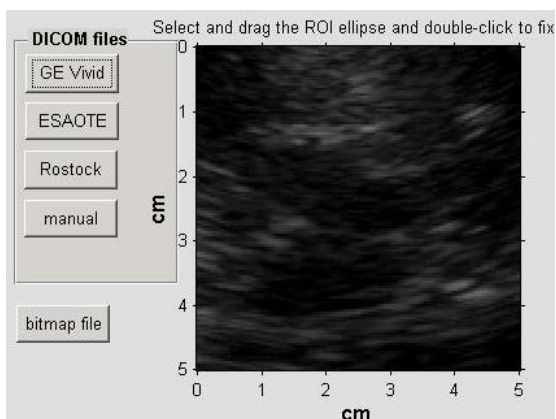
Obr. 14. ROI pro NR s obsahem  $30 \text{ mm}^2$

Samozřejmě i pro oblast NR je velmi důležité hodnocení vstupního jasu obrazu podle výše popsaného agentového systému, tentokrát však s opačným efektem oproti SN – vliv přejasnění by mohl vyhodnotit fyziologický nález místo patologického.

## 5 Automatická verze programu s cílením okna a ROI dle přístroje

Používáme-li stejné sady obrazů ze stejného přístroje, je dobré zefektivnit cílení okna a následně ROI. K našim účelům používáme obrazy z různých přístrojů, převážně GE a ESAOTE. Pro tyto přístroje jsme vyvinuli automatickou verzi, která cílí okno a ROI dle přístroje.

Obraz z každého přístrojů má jiné rozlišení a trochu jinou polohu okna  $50 \times 50 \text{ mm}$ , které nás zajímá. To nás vedlo k vytvoření verze, ve které máme několik tlačítek pro načítání obrazu, a dle předem definovaných souřadnic se okno správně zacílí. V rámci jednoho přístroje je umístění stále stejné, proto je možné jej automatizovat.



Obr. 15. Automatické cílení okna dle přístroje

Tento systém cílení je užitečný a zefektivní práci stále se stejnými snímky, kdy nemusíme ručně umístit okno a zároveň máme vysokou pravděpodobnost správného umístění okna. Při testování na 100 snímcích z jednoho přístroje se pouze v 5 případech zacílilo okno mimo SN, což vykazuje 5%-ní chybu.

Díky použitému formátu DICOM (viz kap. 2) máme v obrazech k dispozici také metadata, ze kterých můžeme snadno extrahovat rozměry obrazu a tedy nastavit pro daný přístroj souřadnic okna, čímž odpadá nutnost manuální cílení okna.

Máme k dispozici 4 tlačítka pro modely přístrojů. Program po stisku příslušného tlačítka ověří z DICOM metadat, zda se skutečně jedná o daný přístroj. Pokud ne, program upozorní na skutečnost, že nesouhlasí výrobce přístroje se stisknutým tlačítkem. To je užitečné k prevenci omylu. V MATLABu použijeme funkci `dicominfo.Manufacturer(filename)`.

Attribute	Value
SpecificCharacterSet	ISO_IR 100
ImageType	DERIVEDSECONDARYTRANSCRANIAL0001
SOPClassUID	1.2.840.10008.5.1.4.1.1.6.1
SOPInstanceUID	1.3.76.2.3.2.90.341770937.2197.3112012.14317921
StudyDate	20120131
SeriesDate	20120131
ContentDate	20120131
StudyTime	1502
SeriesTime	1502
ContentTime	150317
AccessionNumber	-
Modality	US
Manufacturer	ESAOTE
ReferringPhysicianName	[1x1 struct]
StudyDescription	ADULT CEPHALIC
SeriesDescription	Transcranial
AdmittingDiagnosesDescription	-
ManufacturerModelName	6100
DerivationDescription	JPEG lossy compression - compression ratio 9:1
PatientName	[1x1 struct]
PatientID	25

Obr. 16. DICOM metadata k detekci přístroje

Následně se také automaticky umístí ROI do oblasti SN opět dle a priori nastavených souřadnic v okně podle rozlišení. Samozřejmě nechybí ruční možnost zacílení, pokud dostaneme obraz, který nespadá do žádné kategorie.

## 6 Závěr a budoucí práce

Aplikace má být prakticky využitelná pro medicínskou praxi pro hodnocení patologií v ultrazvukových B-obrazech. Nedílnou a důležitou součástí je předzpracování obrazu cílené na hodnocení kvality vstupního obrazu. Tato část je v programu řešena autonomně pomocí softwarového multiagentového systému k detekci průměrného jasu ve výřezu obrazu a

v ROI. Jen tak může být celé zpracování objektivní. Nevhodné nastavení obrazu či například ekvalizace histogramu může vést k falešným detekcím.

Jádro hlavního zpracování je založeno na výpočtu obsahu v potřebném ROI po binárním prahování, které provádíme pro každý práh  $T \in \langle 0; 255 \rangle$ . Získali jsme tak křivku charakterizující klesající obsah v ROI vůči prahu. K objektivnímu hodnocení fyziologie či patologie byla stanovena na základě 100 vyšetřených fyziologických pacientů referenční křivka percentilu z těchto hodnot. Poté je možno srovnávat průběh poklesu obsahu v ROI s touto křivkou. Provedená statistická analýza pro SN ukázala možnost reproducibility aplikace, dále se testuje NR a opět srovnání s výsledky SN. Doposud bylo testován vzorek 100 pacientů k detekci poškození NR, z toho soubor 30 zdravých.

Cílem další práce na programu je otestování výpočtů na větším vzorku pacientů a také z jiných oblastí než SN, aktuálně řešena oblast NR. Možné otestování i jiných struktur v TCS B-obrazech. Současně je vyvíjena verze aplikace s automatickým cílením okna a eliptického ROI dle přístroje ze kterého byly obrazy nahrány. Další práce bude zaměřena na vývoj této verze a další statistickou analýzu v examinovaných strukturách.

## Poděkování

Tento příspěvek vznikl za podpory grantové soutěže SGS 6/2011.

## Literatura

- [1] A. Kubík: *Inteligentní agenti: Tvorba aplikačního software na bázi multiagentových systémů*, Computer Press, Praha, 2004.
- [2] D. Floreano, F. Mondada: Automatic Creation of an Autonomous Agent: Genetic Evolution of a Neural Network Driven Robot. In: *From Animals to Animates III: Proceedings of the Third International Conference on Simulation of Adaptive Behaviour* (D. Cliff, P. Husbands, J. Meyer, and S. Wilson, eds.), MIT Press-Bradford Books, Cambridge, MA, 1994: 402-410.
- [3] D. Školoudík a kol.: *Neurosonologie*, Galén, Praha, 2003.
- [4] D. Školoudík: Reproducibility of sonographic measurement of the substantia nigra, In: *Ultrasound in Medicine & Biology* (9), 2007, 1347-1352.
- [5] J. Blahuta, T. Soukup, P. Čermák, M. Večerek, M. Jakel, D. Novák : ROC and reproducibility analysis of designed algorithm to potential diagnosis of Parkinson's disease in ultrasound images. In: *Mathematical Models and Methods in Modern Science*, WSEAS Press, Puerto de la Cruz, ISBN:978-1-61804-046-6, 2011, 88-93.
- [6] J. Blahuta, T. Soukup, P. Čermák, D. Novák: The ROI defect statistical analysis of substantia nigra to reproducibility of designed experimental algorithm for potential PD diagnosis. In: *Mathematical Methods and Techniques in Engineering and Environmental Sciences*, WSEAS Press, Catania, 2011, ISBN: 978-1-61804-046-6, 261-266.
- [7] J. Blahuta, T. Soukup, P. Čermák, M. Večerek: Ultrasound medical image recognition with artificial intelligence for Parkinson's Disease classification. In: *2012 Proceedings of the 35th International Convention, IEEE, Mipro 2012, Opatija, 2012*, ISBN: 978-953233072-4, 958-962.
- [8] J. Kelemen, M. Ftáčnik, I. Kalaš, P. Mikulecký: *Základy umelej inteligencie*. Alfa Bratislava, Alfa, 1992, ISBN: 80-05-00957-7.
- [9] P. Hoskins, K. Martin, Abigail Thrush: *Diagnostic Ultrasound: Physics and Equipment*. Second Edition, Cambridge University Press, Cambridge, 2010.
- [10] R. Sun, V. Honavar, G. Oden: *Cognitive Systems Research*, Vol. 5 (1-4 issues), Science Direct, Netherlands, 2004, ISSN: 1389-0417.



## Elektronická literatúra v paradigme interakcií človek–stroj

Petra Blšťáková

Ústav svetovej literatúry SAV  
Konventná 13, 813 64, Bratislava  
petra.blstakova@gmail.com

### Abstrakt

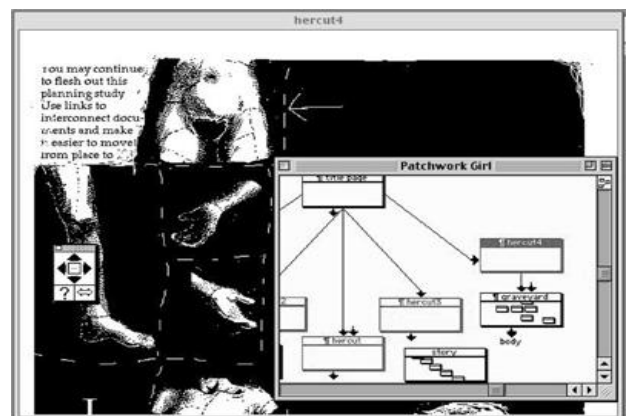
Elektronická literatúra ako hypermediálny artefakt (podľa B. Suwary [16], s. 180) preformúva zážitok čítania a povahu interakcie v literárnej komunikácii. Príspevok potvrdzuje potrebu pestovania a popisovania novej gramotnosti, nových čitateľských kompetencií a upozorňuje na zmeny percepcie (a zmeny čitateľských kompetencií) pri takomto čítaní. Zatiaľ čo pri printovej literatúre prebieha najmä vizuálne spracovanie textu, pri elektronickej je užívateľ v interakcii aj cez dotyk či sluch, a zároveň je nútený spracúvať rôznorodé vizuálne podnety (nie iba text, ale aj obraz, video). Takáto súčinnosť vytvára situácie, ktoré podmieňujú u každého individuálneho čitateľa diferencované (estetické) objekty. Ďalej sa príspevok sústreďuje na úlohy človeka ako súčasť „elektronicko-literárneho“ okruhu, na mieru a spôsob jeho zapojenia. Skúma ju v dvoch pozíciách formujúcich otázku kyborga – ako čitateľa a ako autora. Túto problematiku demonštrujeme na príklade diel elektronickej literatúry využívajúcich virtuálnu realitu (perspektíva užívateľa) a generatívne texty (perspektíva autor[stvá]).

### 1 Úvod do problematiky

V ostatných dvoch desaťročiach tvorcovia literatúry získali výnimočnú, neštandardnú ponuku – možnosť prepojenia literatúry s technikou a novými médiami. Nepôjde tu však o staré podmaňovanie umenia ideológiou revolúcie, resp. technickej revolúcie. Pôjde o nepochybniteľnú spätosť a nerozlúčnosť procesu vznikania literatúry s informačnou technológiou.

Zrod a vývoj elektronickej literatúry je nutné sledovať spolu s rozvojom počítačov. N. Katherine Hayles [5] ju definuje ako literatúru, ktorá je „digital bom“, teda zrodená v digitálnom prostredí a takisto určená na čítanie v ňom (s. 3). Za prvé dielo elektronickej (digitálnej) literatúry sa považuje program *Mark I* od Christophera Stracheyho z roku 1952, ktorý generoval ľubostné listy (základom bol Turingov náhodný číselný generátor) [18]. Z empirického pozorovania rokmi pribúdajúcej produkcie elektronickej literatúry sú zrejme paralely a korelácie s posunmi výpočtovej techniky.

Rozvoj hardvéru, neskôr však najmä softvéru (aj toho primárne určeného pre iné účely), otváral autorom elektronickej literatúry nové možnosti práce a tvorby a čitateľom nové výzvy pri kontakte s takýmito dielami. Výraznou črtou elektronickej literatúry je, okrem digitálnosti, jej multimedialná povaha. Väčšina jej diel prepája najmenej dve médiá, z ktorých tvorí – viac-menej – jednoliaty, zviazaný obsahový celok – hypermediálny artefakt. Najlepšie je to ilustrovať na niekoľkých príkladoch. Dnes už kultovou je *Patchwork Girl* (1995) od Shelley Jackson (Obr. 1), vytvorená pomocou softvéru Storyspace od Eastgate pre potreby tvorcov hypertextových diel.



Obr. 1

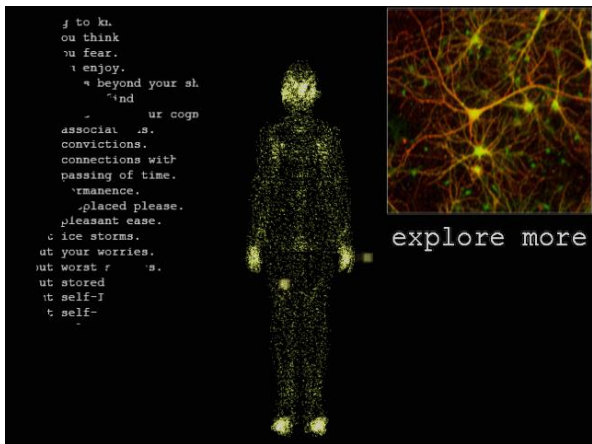
Navigovaním cez časti ženského tela, ktorého obrázok je štartovacím bodom, sa čitateľ preklikáva k jednotlivým lexiám (textovým celkom, [10], s. 25). Tie v sebe nesú niekoľko ďalších hypertextových odkazov a ponúkajú tak alternatívne cesty čítania príbehu, rozkúskovaného presne ako aluzívna (Mary Shelley – Frankenstein) „networka“ v hlavnej úlohe. Neskoršia inštalácia *Text Rain* (1999) autorskej dvojice Camille Utterback a Romy Achitov (Obr. 2) už nepracuje len s textom a obrazom, ale pripája aj real-time video projekciu.

Užívateľ je cez video čierno-bielo projektovaný na plátno, po ktorom padajú na prvý pohľad náhodné písmenká. Tmavými časťami svojho tela (alebo predmetmi, ktoré drží) dokáže písmenká zachytávať a sledovať, aké slová tvoria.



Obr. 2

V skutočnosti totiž nejde o náhodný dážď písmen, ale báseň, ktorá odkazuje k problematike média ako nosiča (tzv. materialita média). Podobným, no ešte o čosi mediálne rozšírenejším a komplexnejším, je dielo slovenských autorov Zuzany **Husárovej** a **Ľubomíra Panáka** *Enter:in' Wodies* (2011) (Obr. 3). Je to (najprv fragmentárna) báseň, ktorú užívateľ odкрýva pohybom snímaným cez Kinect, sprevádzaná ilustratívnym zvukom<sup>1</sup>.



Obr. 3

Už z týchto zopár nesystematicky vybraných príkladov je zrejme, že elektronická literatúra vzniká ako výsledok prepojenia viacerých (digitálnych) médií, ktorými autor sleduje ponúknutie inovovaného literárneho zážitku „čitateľovi“. Jeho vnímanie je značne odlišné od percepcie tradičných (printových) diel. Bližšie túto zmenu ilustrujem v nasledujúcej časti príspevku.

<sup>1</sup> Ukážka: <http://projects.delezu.net/enterin-wodies/>

## 2 Percepcie elektronickej literatúry

Báseň *The Sweet Old Etcetera* (2006) od Alison Clifford (Obr. 4) je inšpirovaná tvorbou e.e. cummingsa, známeho svojimi krátkymi básňami pripomínajúcimi fotografie-momentky. Práve silná obraznosť (cez farby) jeho básní tvorí podnet pre grafické spracovanie, ktoré je možno dosiahnuť v digitálnom prostredí. Báseň Alison Clifford začína ako malý puk (vyzerá ako prázdna zátvorka ( ) ), ktorý čaká, kým naň užívateľ klikne, aby rozkvitol do stromu z písmen. Ak chce užívateľ ďalej čítať, musí hľadať správne miesta, kam kliknúť. Úspešný čitateľ v závere na obrazovke „číta“ vyobrazenú krajinku, sfarbenú odtieňom zapadajúceho slnka a vytvorenú z písmen básne – teda čosi, čo môžeme porovnať k určitej vizualizácii, a všeobecnejšie, k odkazovaniu do poľa problematiky mentálnych konštruktov.



Obr. 4

To, ako sa čitateľ k tomuto záveru dostal a čo a ako pritom vnímal, je však odlišné od toho, ako by pracoval s tlačenu podobou niektorej z básní e.e. cummingsa (alebo iného autora). Pri čítaní tlačenej básne (alebo akéhokoľvek iného printového literárneho diela) vnímame text najmä očami. Málokedy čítame nahlas, takže auditívny dojem slov a zvukomalebná stránka básne sa „vokalizujú“ len akýmsi vnútorným hlasom, prípadne sa nevokalizujú vôbec. Jediný dotyk s dielom je prevrátenie knižného listu a držanie väzby, ktoré nemá nič spoločné s obsahom či interpretáciou básne. Ak v našej mysli, našich jazykových predstavách, báseň vyvoláva isté obrazy, tie sú veľmi prchavé a vytvárajú v nás emočnú stopu, skôr než by stimulovali konkrétny obraz, „grafický odtlačok“ v pamäti. Pri elektronickej literatúre naša interakcia s dielom funguje podstatne odlišne a ako zdôrazňuje Roberto Simanowski [14], nečítame už len text, ale aj to, čo sa s textom deje (s. 13). A preto aj keby Alison Clifford iba remediovala pôvodnú báseň e.e. cummingsa (a nevytvorila na jej motívy

obsahovo novú), tak by sme ju v digitálnom prostredí vnímali a interpretovali odlišne ako v tlačenej forme. Ako hovorí N. K. Hayles [6], materialita diela ovplyvňuje jeho význam, lebo keď sa zmení materiálny substrát, zmení sa aj dielo, a tým, logicky, aj jeho vnímanie (s. 38). Azda najväčšou zmenou oproti tradičnému čítaniu je zapojenie dotyku a taktilného vnímania. Bez toho, aby sme v *The Sweet Old Etcetera* klikali myšou, sa nedopracujeme ani len k prvému veršu básne. V prípade iných elektronických diel je dotyková zložka ešte integrálnou a dominantnejšou črtou – napr. v *Text Rain* a *Enter:in' Wodies* používa užívateľ celé telo a pohybmi absolútne netypickými pre tradičné čítanie sa zapája, naviguje dielom. Okrem toho musí „čítať“ rôzne mediálne obsahy – text je dynamický, mieša sa s videom, často sa pridáva aj zvuk, ktorý je jednou z podstatných zložiek významu. Čitateľ/užívateľ je teda atakovaný cez niekoľko zmyslových kanálov naraz – cez zrak, sluch, dotyk, pohyb – ide o multisenzorické čítanie.

V procese percepcie týchto javov, ktoré bádatelia skúmajú ako problematiku „intermediality“, evidentne dochádza k mixovaniu médií do jedného celku, v ktorom rozlišovanie pôsobenia jednotlivých médií stráca opodstatnenosť – preto niektorí bádatelia hovoria o tom, že hranice jednotlivých médií sa mažia a tie pôsobia ako integrálny celok ([7], s. 92). Mohli by sme predsa hovoriť o akejkoľvek synergii, pri ktorej čiastočky tvorené jednotlivými médiami spoločne vytvárajú omnoho silnejší a komplexnejší podnet ako keď stoja samostatne. Je potrebné uvedomiť si a skúmať kognitívne procesy, ktoré pri percepcii takéhoto diela prebiehajú. Intermedialita diel a ich multisenzorické čítanie menia to, ako ich vnímame, prežívame. Napriek existencii štúdií mapujúcich vnímanie diel elektronickej literatúry (z množstva napr. Mial, Dobson 2001, Brinkerhoff, Klein, Koroghlanian 2001, White 2007), problematika ešte stále nie je celkom preskúmaná. Dôvod zrejme treba hľadať v potrebe spojenia viacerých vedných disciplín a odborníkov – literárnych vedcov, neurológov, kognitivistov, psychológov, ako aj v náročnosti a ťažkej uchopiteľnosti/merateľnosti niektorých kognitívnych procesov. Zdá sa, že bránou k reflektovaniu procesov prebiehajúcich pri čítaní elektronickej literatúry je pozornosť – jej kolísanie, zameranie, priestorová a časová distribúcia, selektívnosť (či je niektorý zmyslový kanál dominantný, či je prítomný „filter“ uľtujúci ostatné), jej determinanty (vonkajšie, vnútorné) [8] (s. 111-159). Na aspekt pozornosti poukazuje totiž už priekopníčka problematiky tvorby a interpretovania techno-textov N. K. Hayles, keď sa zamýšľa nad využitím a striedaním „deep attention“ a „hyper attention“ pri čítaní elektronickej literatúry ([5], s. 121-130). Okrem toho však bádatelka kladie dôraz aj na distribúciu agentnosti medzi telo a stroj, medzi ktorými vzniká „feedback loop“ [5]. Oba tieto aspekty sú predmetom záujmu aj u Janeza

Strehovca, ktorý si cez koncept „ozvlášťňovania“<sup>2</sup> literárneho diela všima instantnosť textov elektronickej literatúry, vyžadujúcu text s „efektom jazdy výt'ahom“, čiže jeho rýchlu premenlivosť (zážitkovosť). Tá od užívateľa požaduje koncentráciu pozornosti v procese „nielen- čítania“ ([15], s. 31), no zároveň naplňa dnes aktuálnu požiadavku rýchleho zaujatia. J. Strehovec upozorňuje, že dnešný čitateľ (resp. dnešný človek) žije obklopený obrazovkami, je v interakcii so svetom cez nimi tvorený interfejs, a žije tak v „nomádskom kokpíte“, čiže zliatine reálneho sveta s tým na/v obrazovkách, cez ktorý sa naviguje tlačidlami (s. 25).

Zo všetkých načrtnutých súvislostí je zřejmé, že reflektovanie elektronickej literatúry je okrem vyslovenia literárnovedných súvislostí odkazané aj na kontexty komunikácie (či kooperácie) medzi človekom a strojom. Čiže ide o dôverne známu predstavu kyborga.

### 3 Kyborgizácia cez elektronicnú literatúru

V prvom je na mieste pripomenúť si starú otázku: čo je to kyborg? Autori odbornej literatúry nie sú v definovaní zhodní. Definovanie závisí od miery či stupňovania dôvodov, pre ktoré je možné niekoho (niečo) považovať za kyborga. Škála postupuje od nositeľa kontaktných šošoviek cez človeka s kardiosťimulátorom, prípadne robotickou rukou, až po terminátora z rovnomenného sci-fi filmu. Z tejto diferencovanosti vyplývajú viaceré problematické body. Je kyborgom človek technicky vylepšený s účelom korekcie (napr. robotická ruka človeka s amputovanou hornou končatinou), alebo až ten, ktorý vďaka technickému vylepšeniu (enhancement) presahuje rámec ľudských možností? Ďalej – je kyborgom ten, kto má techniku „implantovanú v tele“, alebo pripevnenú zvonku (tzv. „wearable technology“, prípadne celý „exoskeleton“ – teda kybernetický zjav ale ľudské/organické vnútro), alebo môže byť reč o tzv. „low-tech“ kyborgoch, medzi ktorých patrí každý užívateľ mobilného telefónu, auta a i.<sup>3</sup>? Kodifikátormi pojmu kyborg sú Manfred E. Clynes a Nathan S. Kline, ktorí ho v roku 1960 použili v článku *Cyborgs and Space* (Astronautics, september 1960)<sup>4</sup> zameranom na teóriu vytvárania prístrojov implantovaných človeku tak, aby bol schopný prežiť vo vesmíre. Namiesto prenášania celého doterajšieho prostredia navrhovali prispôsobenie sa tomu novému – vesmíru. Nejde tu teda o korektúru v snahe obstať pri porovnaní s priemerom, ale o vylepšenie s cieľom dosiahnutia nových schopností (o evolúciu?) – a to tak, aby bola zachovaná homeostáza, aby bol

<sup>2</sup> termín V. Šklovského – „ostranenie“

<sup>3</sup> [http://cyborganthropology.com/What\\_is\\_a\\_Cyborg%3F](http://cyborganthropology.com/What_is_a_Cyborg%3F)

<sup>4</sup> <http://web.mit.edu/digitalapollo/Documents/Chapter1/cyborgs.pdf>

mechanizmus integrovaný a aby jeho ovládanie nemuselo byť vedomé<sup>5</sup>. Účelom tohto vylepšenia malo byť: „...leaving the man free to explore, to create, to think, and to feel.“<sup>6</sup> Z iného uhla pohľadu kyborga konceptualizuje Donna Haraway, ktorá naň nazerá ako na sociálnu konštrukciu, ktorá umožňuje dekonštruovať základné kategórie (najmä rodové). Podľa nej „cyborg is (...) a creature of social reality as well as a creature of fiction“ ([4], s. 149), ktorý vznikol zrútením troch hraníc: medzi človekom a zvierat'om, živým a neživým, hmotným a nehmotným. Aby boli východiská kompletné, je treba spomenúť jedného zo zakladateľov kybernetiky, Norberta Wienera. Vo svojej publikácii *The Human Use of Human Beings – Cybernetics and Society* tvrdí, že spoločnosť je možné pochopiť iba štúdiom komunikácie, pričom (správne) predpovedá, že v budúcnosti nepôjde len o medziľudskú komunikáciu, ale aj komunikáciu na osiach človek–stroj a stroj–stroj ([20], s. 16). Teraz sa pokúsím načrtnúť, kde v tejto sieti pohľadov a definícií leží elektronická literatúra. Wienerova kybernetická teória je založená na komunikácii a kontrole, teda na slučke (loop) spätnej väzby a odpovede. Literárna komunikácia je založená na podobnej slučke výmeny informácií medzi autorom/dielom a čitateľom, pričom elektronická literatúra obzvlášť kladie dôraz aj na opačný smer komunikácie, teda späť od čitateľa k dielu/autorovi. Preto je možné Wienerovu teóriu aplikovať aj pri analýze elektronickej literatúry. Interakcia v komunikácii (slučke) elektronickej literatúry nebude iba v podobe kladného alebo záporného hodnotenia diela (ako to navrhuje T. Tenney, napr. pri divadelnej hre so živým publikom, v časovej a priestorovej jednote<sup>7</sup>), ale aj výmenou a reagovaním na informácie priamo od čitateľa, ktorý do diela aktívne zasahuje, naviguje ním a doslova s ním komunikuje (napr. zadávaním vlastného textu). Technicky zapojenie čitateľa najbežnejšie prebieha cez používanie klávesnice, myši, prípadne Kinectu. Môžeme tak hovoriť prinajmenšom o „low-tech“ kyborgovi. Existujú však aj diela elektronickej literatúry využívajúce virtuálnu realitu (VR). Sú tvorené pre CAVE (cave automatic virtual environment), ako je to aj v prípade diela *Screen* (kolektív autorov, 2003) (Obr. 5), kde užívateľ cez dátovú rukavicu a displejom vybavený headset „vstupuje“ do miestnosti tvorenej stenami zo slov, ktoré vylievajú, padajú a nechávajú sa užívateľom manipulovať.



Obr. 5

Napriek tomu, že nejde o vnútorné implantované mechanizmy, je tu viditeľná súvislosť s prispôbením sa prostrediu v chápaní Clynesa a Klinea ([3], s. 26). Užívateľ sa vnára (immersion) do nového prostredia vďaka transparentnosti sveta tvoreného médiami, je v ňom („presence“ – Marie-Laure Ryan) ([13], s. 66). Táto prítomnosť vo svete (jej pocit, prežívanie) je kontrolovaná cez tri premenné: rozsah senzorických informácií, kontrola vzťahu senzorov k prostrediu, schopnosť modifikovať fyzické prostredie (s. 67). Inými slovami, ide o Wienerovu „feedback-and-response loop“: informácie vysielané simulátorom VR, reakcia na ne/ich prežívanie užívateľom, konanie užívateľa na základe podnetov, opätovné reagovanie na VR atď.

Pomocou technických prístrojov pripravených k človeku je možné vstúpiť do virtuálneho sveta reálne vnímateľného zmyslami. Na priesečníku čitateľ/užívateľ – stroj sa v dielach využívajúcich VR nachádza bod ostraňovania a difúzie hraníc medzi človekom a strojom, ako aj medzi hmotným a nehmotným. Tieto aspekty považuje Donna Haraway, ako bolo spomenuté vyššie, za príčiny (a podmienky) vzniku kyborga. Podobná difúzia cez „mäkké“ hranice nastáva aj na začiatku procesu literárnej komunikácie, čiže medzi autorom a strojom. Výborným príkladom tejto hranice je generatívna poézia – teda básne tvorené počítačom, generovaním na základe algoritmu. Autor síce vkladá do takéhoto diela konceptuálnu myšlienku a programuje samotné generovanie, no len čo dielo sprístupní užívateľovi, program už koná sám, prípadne na základe podnetov čitateľa. Je otázne, do akej miery potom môžeme danú prácu považovať za výtvor autora (viacerí bádatelia spochybňujú estetické hodnoty takýchto básní). Avšak, kde sa v tejto spleti ocitá autor? Je iba prvotným impulzom vzniku diela, zdieľa spoluautorstvo so strojom, alebo časť autorského podielu odkazuje čitateľovi? Espen Aarseth v tejto súvislosti hovorí o „kyborgovskej aktivite písania“, ktorá podľa neho poukazuje na to, ako sa význam snaží sám seba vyprodukovať ([1], s. 57). Tým naznačuje, že literárne dielo vznikajúce v

<sup>5</sup> ibid, s.27

<sup>6</sup> ibid

<sup>7</sup> <http://inc.ongruo.us/2010/12/29/cybernetics-in-art-and-the-myth-of-the-cyborg-artist>

digitálnom prostredí je do vysokej miery „samoproduktívne“. Kyborgom by tu potom bola autorská entita počítača, či jeho programov (alebo hardvéru a softvéru spoločne). Zároveň Aarseth dodáva, že podiel kontroly autora/textu/čitateľa nad dielom treba posudzovať individuálne, vzhľadom na konkrétne dielo [1].

Pri celom spektre literárnej komunikácie (od autora, cez text a stroj, po čitateľa), pri využití VR, generatívnej poézii a ďalších dielach elektronickej literatúry dochádza k interakcii, integrácii, ba úzkemu spojeniu medzi človekom a strojom. Autorská dvojica Elena Matějková a Peter Sýkora kyborgizáciu ľudského tela charakterizujú ako „čo najtesnejšie prepojenie biologického hardvéru (tela) s hardvérom kybernetickým“ ([11], 185). K analogickým záverom privádzajú uvedené príklady elektronickej literatúry. Je z nich zrejmé, že ide o veľmi úzke spojenie človeka so strojom, keďže ich kontakt prebieha najmä dotyk, ovládaním počítača či nosením prístrojov, pričom sa z ľudského tela stáva komunikačný interfejs, cez ktorý počítačový softvér modifikuje percepciu „obsahov“ a napokon aj literárnu recepciu diela. V týchto všetkých súvislostiach je prinajmenšom opodstatnené hovoriť o „low-tech“ variante a postupne narastajúcej kyborgizácii.

#### 4 Záver

Podľa Jozefa Kelemenena môžeme kyborgizáciu vnímať ako globálny vývoj na všetkých úrovniach existencie postmoderného človeka, za predpoklad pokračovania evolúcie ([9], s. 322). Tento predpoklad podporuje jednak aktuálny technický pokrok a aj niektoré umelecké trendy – generácia kyborgických umelcov ako Stelarc, Orlan, či Scott Draves. Na zahraničnom trhu stále pribúda kníh odkazujúcich k augmentovanej realite, lebo táto čoraz invazívnejšie preniká do bežného života (BrainGate technológia BCI, Muse, Google Glas multifunkčné okuliare, aplikácia Layar, Yelp – Monocle pre iPhone a i.) a stále intenzívnejším spôsobom rozširuje naše dimenzie vnímania skutočnosti. Čím sme závislejší od techniky, tým väčšmi kyborgizujeme (GPS, instantný prístup na internet, zálohovanie dát, plánovanie, komunikácia...). Táto tendencia – ako tomu nasvedčujú uvedené príklady – je pre autorov elektronickej literatúry zdrojom plodných možností a inovatívnych prostriedkov. Teda tak, ako sa technikou a kyborgizáciou menia podmienky nášho vnímania sveta, mení sa podoba nášho sveta, menia sa aj podmienky percepcie literatúry – primáme tej, ktorú autori koncipujú ako „netriviálnu“ prácu s elektronickými dielami ([1], s. 125). Nezvratným dôkazom tohto vývoja je generácia malých detí, ktoré od dvoch rokov bez problémov navigujú tablety, no na skúškach školskej zrelosti demonštrujú nedostatočné

grafo-motorické zručnosti. Stojí za tým neuroplasticita, ktorá pod vplyvom techniky pomáha formovať generáciu (nových) „mozgov“ (vnímateľov), ktoré sú ešte o krok ďalej v množnom procese kyborgizácie.

Uvedené príklady z oblasti elektronickej literatúry sú zrejmou šancou pre participovanie na reflektovaní týchto javov – tak aspektov literárnej tvorby ako jej recepcie.

#### PodĎakovanie

Príspevok vznikol v rámci grantu VEGA 2/0182/10, Text na internete ako jav (r)evolúcie kultúry.

#### Literatúra

- [1] Aarseth, E. 1997: *Cybertext: Perspectives on Ergodic Literature*. The Johns Hopkins University Press, Baltimore.
- [2] Brinkerhoff, J. D., Klein, J. D., Koroghlanian, C. M. 2001: Effects of Overviews and Computer Experience on Learning from Hypertext. *Educational Computing Research* 25(4), [cit. 2.4.2013] [http://mkoehler.educ.msu.edu/OtherPages/Courses/CEP\\_909\\_FA02/Readings/Brinkerhoff.pdf](http://mkoehler.educ.msu.edu/OtherPages/Courses/CEP_909_FA02/Readings/Brinkerhoff.pdf)
- [3] Clynes, M. E., Kline, N. S. 1960: *Cyborgs and Space*. In: *Astronautics*, September, s. 26–27 a 74–75; reedícia cez Gray, Mentor, and Figueroa-Sarriera, ed., *The Cyborg Handbook*, New York: Routledge, 1995, s. 29–34. <http://web.mit.edu/digitalapollo/Documents/Chapter1/cyborgs.pdf>
- [4] Haraway, D. 1991: *A Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late Twentieth Century*. In: *Simians, Cyborgs, and Women: The Reinvention of Nature*. Routledge, New York.
- [5] Hayles, N. K. 2010: *Electronic Literature: New Horizons for the Literary*. University of Notre Dame, Notre Dame.
- [6] Hayles, N. K. 2002: *Writing Machines*. The MIT Press, Cambridge.
- [7] Husárová, Z. 2012: *Zážitok\* multisenzorického čítania elektronickej literatúry*. In: (Eds.) B. Suwara, Z. Husárová: *V sieti strednej Európy: nielen o elektronickej literatúre*. SAP a Ústav svetovej literatúry SAV, Bratislava.
- [8] Chalupa, B. 2011: *Studie z kognitivní psychologie*. Nakladatelství Littera, Brno.

- [9] Kelemen, J. 2010: *Myslenie a stroj*. Kalligram, Bratislava.
- [10] Landow, G. P. 1997: *Hypertext 2.0: The Convergence of Contemporary Critical Theory and Technology*. The Johns Hopkins University Press, Baltimore.
- [11] Maťejková, E., Sýkora, P. 2011: *Kognitívne vylepšovanie človeka, evolúcia a etika*. In: *Umelá inteligencia a kognitívna veda III*. STU, Bratislava.
- [12] Mial, D. S., Dobson, T. 2001: Reading Hypertext and the Experience of Literature. *Journal of Digital Information*. 2(1). [cit. 1.4.2013] <http://journals.tdl.org/jodi/index.php/jodi/article/view/35/37>
- [13] Ryan, M.-L. 2001: *Narrative as Virtual Reality*. The Johns Hopkins University Press, Baltimore.
- [14] Simanowski, R. 2009: *What is and Toward What End Do We Read Digital Literature?* In: (Ed.) F.J. Ricardo: *Literary Art in Digital Performance: Case Studies in New Media Art and Criticism*. Continuum International Publishing Group, New York.
- [15] Strehovec, J. 2012: *Digitálny literárny text v ríši nových médií: Text ako jazda*. In: (Eds.) B. Suwara, Z. Husárová: *V sieti strednej Európy: nielen o elektronickej literatúry. SAP a Ústav svetovej literatúry SAV*, Bratislava.
- [16] Suwara, B. 2012: *Na ceste k hypermediálnemu artefaktu*. In: (Eds.) B. Suwara, Z. Husárová: *V sieti strednej Európy: nielen o elektronickej literatúry. SAP a Ústav svetovej literatúry SAV*, Bratislava.
- [17] Tenney, T. 2010: *Cybernetics in Art and the Myth of the Cyborg Artist*. [cit. 4.4.2013] <http://inc.ongruo.us/2010/12/29/cybernetics-in-art-and-the-myth-of-the-cyborg-artist/>.
- [18] Wardrip-Fruin, N. 2005: Christopher Strachey: The first digital artist? [cit. 2.4.2013] <http://grandtextauto.org/2005/08/01/christopher-strachey-first-digital-artist/>.
- [19] White, A. 2007: Understanding hypertext cognition: Developing models to aid users' comprehension. *First Monday*, 12(1) [cit. 1.4.2013] <http://journals.uic.edu/ojs/index.php/fm/article/view/1425/1343>.
- [20] Wiener, N. 1989: *The Human Use of Human Beings – Cybernetics and Society*. Free Association Books, London.
- [21] *What is a Cyborg?* [online] [cit. 4.4. 2013] [http://cyborganthropology.com/What\\_is\\_a\\_Cyborg%3F](http://cyborganthropology.com/What_is_a_Cyborg%3F).

# Multiagentní model monopsonu na trhu práce

Jan Burian

Vysoká škola ekonomická

Nám. W Churchilla 3, Praha 3

burianj@vse.cz

## Abstrakt

Model monopsonu na trhu práce zkoumá situaci, kdy zaměstnavatel může díky své síle snižovat mzdy až k existenčnímu minimu. Pracovníci jsou v takovém případě "paradoxně" nuceni zvyšovat nabídku práce při snižujících se mzdách. Model dále zkoumá vliv spotřebních půjček na zisky zaměstnavatele, výši mezd, nezaměstnanost a blahobyt. Model ukazuje, že zvyšování pravděpodobnosti, že si pracovníci vezmou spotřební půjčku, zvyšuje zisky zaměstnavatelů, zvyšuje nezaměstnanost, snižuje blahobyt a v některých situacích umožňuje zaměstnavatelům snížit mzdy až k existenčnímu minimu pracovníků. Model také napovídá, že pro velmi nízké úrovně mezd povede snížení minimálních mezd "paradoxně" ke zvýšení nezaměstnanosti. V závěru vyvozují doporučení jak pro zaměstnavatele snažící se maximalizovat zisk, tak pro zaměstnance bránící se snižování mezd.

## 1 Úvod

Učebnicová mikroekonomická analýza individuální křivky nabídky práce je založená na křivkách indiference mezi dobou odpočinku a výší mzdy (viz např. [2, 3]). Tyto křivky jsou výslednicí protichůdného působení důchodového a substitučního efektu. Výsledek může být pro každé individuum velmi odlišný. Ekonomové předpokládají, že individuální nabídka práce je rostoucí – tj. roste se zvyšující se mzdou. Učebnice ale popisují i situaci, kdy zvyšování mezd nejprve vede ke zvyšování agregátní nabídky práce, ale po překročení určité úrovně mezd vydělávají pracovníci již tolik, že dojde k převaze důchodového efektu a nabídka práce se začne snižovat. Ovšem učebnice předpokládají, že agregátní nabídka práce zůstává rostoucí, protože do odvětví mohou vstupovat noví pracovníci přilákaní vysokými mzdami.

Téměř nikdo se nevěnuje opačnému postupu – co se bude dít, když budeme mzdy dlouhodobě snižovat. Lze předpokládat, že nejdříve se bude agregátní nabídka

práce se snižujícími mzdami také snižovat. Ale jak se budou mzdy blížit k existenčnímu minimu, tedy objemu prostředků, který je nutný k holému přežití, začne se agregátní nabídka práce „paradoxně“ zvyšovat. Pracovníci musí začít nabízet stále více práce, aby přežili. Dosáhnout této situace by samozřejmě bylo výhodné pro majitele podniků – získali by tím maximální množství práce za minimální ceny. V rozvinutých zemích existují dva limity takového vývoje. Jedním je minimální mzda a druhým je předpokládaná konkurence mezi zaměstnavateli. Vzájemná konkurence mezi zaměstnavateli předpokládá, že zaměstnavatelé neuzavírají nekalé dohody o výši mezd, a že pracovníci jsou dostatečně mobilní a mohou mezi zaměstnavateli volit. Minimální mzda může být v důsledku kombinace tlaku zaměstnavatelských lobby a inflace dotlačena téměř k existenčnímu minimu. Možnost pracovníků vybrat si mezi různými zaměstnavateli bývá omezená – je omezena jejich mobilita, náklady na rekvalifikaci jsou vysoké a konečně různé podniky mohou být vlastněny jednou korporací nebo kartelem, který může uplatňovat stejnou mzdovou politiku. Tyto faktory umožňují, aby se ze zaměstnavatele stal monopson, tedy jediný nakupující na trhu práce.

V této situaci není pro pracovníky možné odcházet a hledat si práci u jiných zaměstnavatelů, a tak snižovat agregátní nabídku práce (alespoň v krátkodobém či střednědobém horizontu). Díky tomu se u monopsonu může při snižování mezd blízkých existenčnímu minimu zvyšovat agregátní nabídka práce.

## 2 Popis modelu

### 2.1 Hlavní prvky modelu

Níže popsany multiagentní model je model monopsonu na trhu práce. Agenti představují jednotlivé pracovníky. Zaměstnavatel je specifický agent, kterého můžeme chápat jako prostředí, ve kterém pracovníci žijí. Pracovníků může být velké množství a parametry jejich

chování se mohou lišit. Zaměstnavatel je pouze jeden. Pracovníci a zaměstnavatel se střetávají na trhu práce, jehož klíčovými parametry jsou mzdy a nabídka práce.

Model navíc zkoumá na mikro i makro úrovni vliv dalších faktorů, jako je výše úspor a zadlužení pracovníků a vliv spotřebních půjček.

Tento článek představuje jen dílčí analýzu tohoto modelu. Úplnou analýzu modelu a podrobný popis metodologie, na jejímž základě byl model vytvořen, může čtenář nalézt v [1].

## 2.2 Pracovníci

### 2.2.1 Individuální nabídka práce

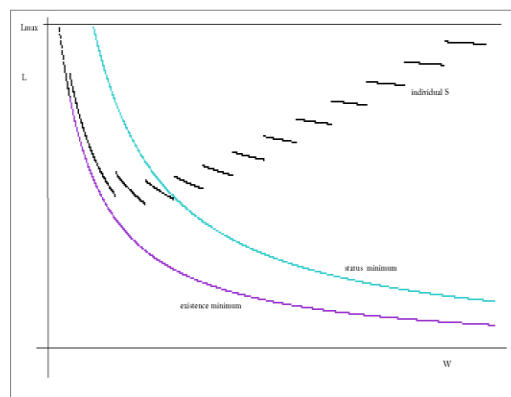
Individuální nabídka práce je v modelu chápána jako průměrná denní práce za časovou jednotku. Maximum je 14 hodin práce denně, minimum 0. Časová jednotka odpovídá jednomu kroku modelu, což by realitě odpovídala cca jednomu měsíci až jednomu roku. Tímto způsobem můžeme zachytit to, že pracovníci nabízejí práci v odlišené struktuře. Víme, že řada pracovníků nepracuje každý den stejný počet hodin a nepracuje ve všech dnech, nebo týdnech. Průměrná denní práce 6 hodin denně tak může odpovídat pracovníkovi, který pracuje každý den v časové jednotce 6 hodin nebo pracovníkovi, který pracoval jen polovinu dní v časové jednotce 12 hodin denně.

### 2.2.2 Existenční a statusové minimum

Model předpokládá, že individuální nabídka práce je ovlivňována existenčním minimem a statusovým minimem. Existenční minimum je částka, kterou pracovník potřebuje k tomu, aby zabezpečil své základní potřeby a reprodukoval svoji pracovní sílu. Existenční minimum je stejné pro všechny pracovníky. Spotřeba nad existenční minimum umožňuje pracovníkům dosahovat sociálního statusu, proto ji nazýváme statusovou spotřebou. Statusové minimum je částka, podle které se řídí užitek ze statusové spotřeby. Statusové minimum může být několikrát vyšší než je existenční minimum a pro různé pracovníky může být odlišné. V modelu je statusové minimum generováno pro každého pracovníka náhodně s normálním rozložením se středem rovným trojnásobku a směrodatnou odchylkou rovnou jednomu násobku existenčního minima (statusové minimum však nemůže být nižší než existenční).

Vztah mezi mzdou a počtem hodin práce nutným pro dosažení obou minim je daný hyperbolami:

práce nutná k dosažení minima =  
 existenční (statusové) minimum / hodinová mzda



Graf 1. - Závislost mezi hodinovou mzdou (osa x) a individuální nabídkou práce (osa y) potřebnou k dosažení existenčního minima (fialová hyperbola) a statusového minima (modrá hyperbola). Černá křivka znázorňuje individuální nabídku práce jednoho pracovníka. Tato křivka je v modelu nespojitá, protože zvýšení nabídky práce znamená vždy jednu celou hodinu práce navíc.

Individuální nabídka práce tedy musí minimálně dosahovat takové hodnoty, aby výsledný výdělek pokryl existenční minimum. Každá další jednotka výdělku přináší užitek ze statusové spotřeby - statusový užitek. Tento užitek odpovídá substitučnímu efektu. V modelu je statusový užitek dán mocninnou funkcí, jejímž základem je rozdíl mezi výdělkem a statusovým minimem. Exponent této funkce je nižší než jedna, a statusový užitek tedy roste rychleji pro mzdy nižší, než je statusové minimum a pomaleji pro mzdy vyšší. Důchodovému efektu naopak odpovídá negativní užitek z času obětovaného práci, který je také dán mocninnou funkcí, ale s exponentem vyšším než jedna, a tedy roste stále rychleji s větším množstvím práce. Pracovník bude nabízet takové množství práce, při kterém bude maximalizovat svůj celkový užitek, k čemuž dojde tehdy, když se mezní hodnoty statusového užitku a negativního užitku z času obětovaného práci budou navzájem rovnat<sup>1</sup>.

<sup>1</sup> Podrobný popis tohoto mechanismu nalezne zájemce ve zdrojovém kódu modelu.



### 2.2.3 Půjčky a dluhy, utrácení a spoření

Agenti - pracovníci jsou reaktivní, ale mají jakousi minimální paměť, která spočívá v tom, že si „pamatují“ jaké mají dluhy a úspory.

Každý pracovník, který není schopen v každém kroku modelu zaplatit existenční minimum, si musí na jeho zaplacení půjčit. Pracovníci se však mohou zadlužit i tím, že si vezmou spotřební půjčku. Tuto půjčku si v modelu mohou vzít jen pracovníci, kteří mají práci a nenulové úspory (tj. nejsou již v dlužích). Velikost půjčky je vždy desetinásobek existenčního minima. Pravděpodobnost, s jakou si pracovníci berou spotřební půjčky (dále jen zkráceně „pravděpodobnost půjček“), je jedním z klíčových parametrů modelu. Jedním z hlavních cílů experimentů s modelem je popsat vliv tohoto parametru na chování modelu.

Pokud si půjčku vezme pracovník, jehož úspory jsou větší než půjčka, tak ji okamžitě splatí a jeho úspory se sníží o velikost půjčky. Je to tedy to samé, jako by si půjčku nevzal, jen se rozhodl pro drahý nákup, na který má peníze. Pokud má pracovník nižší úspory než je výše půjčky, využije je celé na pokrytí části půjčky a zbytek se stane dluhem.

Dluh splácejí pracovníci v každém kole splátkou o velikosti existenčního minima. Zadlužený pracovník tak potřebuje nabízet při stejné mzdě dvakrát tolik práce, aby zaplatil existenční minimum i splátku a nezadlužil se ještě více. Zadlužení tak ovlivňuje nabídku práce. Může se stát, že při stejné mzdě může pracovník zvýšit nabídku práce, když se zadluží, a po splacení dluhu zase nabídku práce snížit. Splátky nejsou v modelu úročeny.<sup>2</sup>

Na začátku běhu modelu získají pracovníci náhodnou výši úspor o maximální velikosti desetinásobku existenčního minima. Pokud mají pracovníci práci, získávají v každém kole celkovou mzdu rovnou součinu hodinové mzdy a jimi nabízené práce. Tuto mzdu vždy nejprve přičte k úsporám (přijde mu na účet) a následně může z úspor utrácet.

Pokud je pracovník schopen splatit životní minimum a půjčky, může část mzdy dále spotřebovat a část uspořit. Při modelování spotřeby jsem se snažil vystihnout psychologii spotřebitele. Spotřeba se řídí velikostí mzdy, úspor, tím zda má pracovník zrovna práci a

---

<sup>2</sup> Ve skutečnosti je samozřejmě možné, aby se zadlužení dostali kvůli nesplaceným úrokům a penále do dluhové pastí, ze které se nemohou dostat nikdy. Tento problém by si jistě zasloužil modelovat, ale vyžadovalo by to jiný typ modelu, který by kladl větší důraz na modelování sociálních parametrů dlužníků a praktik lichvářských firem.

velikostí statusového minima. Rozhodně nepovažuji za vhodné modelovat utrácení jedním parametrem jako například sklon ke spotřebě (resp. k úsporám). Jelikož je model spotřeby relativně komplikovaný a přitom není pro chod modelu klíčový (úspory neovlivňují nabídku práce), umístil jsem podrobný popis modelu spotřeby do poznámky pod čarou<sup>3</sup>.

## 2.3 Zaměstnavatelé

### 2.3.1 Mzdy a zisk

Mzdy představují v modelu vždy hodinové mzdy. Mzdy mohou být buď pevně dané jako externí parametr nebo je určuje zaměstnavatel tak, aby maximalizoval zisk. Zaměstnavatel nemá žádné informace o celkovém tvaru funkce agregátní nabídky práce. Zaměstnavatel však může sledovat, jaká agregátní nabídka vzniká při jisté úrovni mezd a reagovat na ni. Mzdy jsou proto v modelu chápány jako vysvětlující proměnná a v grafech je představovány osou x. To je rozdíl oproti učebnicovým modelům, které ji chápou spíše jako proměnou závislou na agregátním množství nabízené práce. To je však problematické, protože, jak jsem naznačil v úvodu, může snadno nastat situace, kdy je nabízeno stejné množství práce za různé mzdy. Naopak nejsou časté situace, kdy za jinak stejných externích podmínek budou pro jednu úroveň mzdy nabízeny různé hodnoty práce. Výjimkou zde může být vliv zadlužení pracovníků.

---

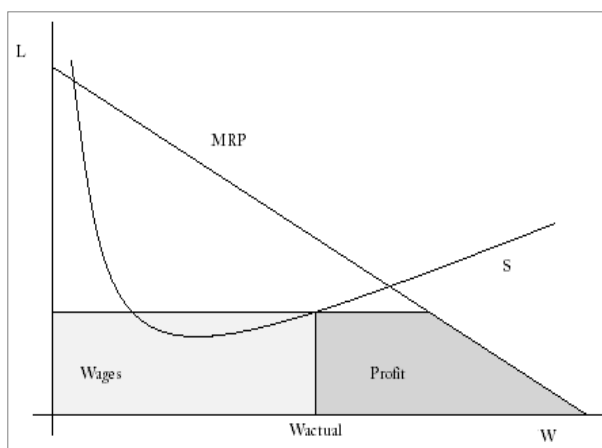
<sup>3</sup> Pracovník nejprve splatí existenční minimum a případnou splátku dluhu. Pokud je zbytek úspor menší než rozdíl mezi existenčním a statusovým minimem, spotřebuje zaměstnaný i nezaměstnaný agent polovinu svých úspor (kompromis mezi potřebou dosáhnout statusové spotřeby a snahou nechat si železnou rezervu).

Pokud pracovník není zaměstnaný, ale úspory jsou vyšší než statusové minimum, chce se pracovník přiblížit spotřebě na úrovni statusového minima. Spotřebuje nejprve tolik, aby dosáhl poloviny statusového minima. Kolik spotřebuje navíc, závisí na velikosti úspor. Pro úspory nízké, blízké se statusovému minimu, nespotřebuje navíc téměř nic, pro úspory zvyšující se limitně do nekonečna bude se spotřeba limitně blížit druhé polovině statusového minima (celková spotřeba se blíží statusovému minimu).

Pokud je pracovník zaměstnaný a jeho mzda i úspory jsou vyšší než statusové minimum, spotřebuje nejprve tolik, aby dosáhl statusového minima. Kolik spotřebuje ze zbylé mzdy, závisí na velikosti úspor. Pro úspory nízké, blízké se statusovému minimu, nespotřebuje ze zbytku mzdy téměř nic, pro úspory zvyšující se limitně do nekonečna se spotřeba bude limitně blížit celému zbytku mzdy.

Ještě podrobnější popis lze v případě zájmu vyčíst ze zdrojového kódu modelu.

Na výši mezd je závislá zaměstnavatelova poptávka po práci. Ve shodě s učebnicovými modely předpokládám, že zaměstnavatel maximalizuje zisk, pokud mzda je rovna příjmu z mezního produktu práce<sup>4</sup>. Proto můžeme příjem z mezního produktu ztotožnit s poptávkou. Důležitým předpokladem a jistým omezením modelu je konstantní velikost kapitálu, a tedy stabilní funkce příjmu z mezního produktu práce. To odpovídá učebnicové situaci v krátkém období. Předpokládáme, že příjem z mezního produktu práce je lineárně závislý na množství práce a je klesající. To znamená, že zaměstnavatelova poptávka po práci bude lineárně klesat s rostoucí mzdou.



Graf 2. – Graf znázorňuje objem mezd (Wages) a zisk zaměstnavatele (Profit) při dané mzdě (Wactual) a odpovídající agregátní nabídce práce (S). Poptávku po práci představuje příjem z mezního produktu MRP. Agregátní nabídka práce je pouze hypotetická, v modelu nemá stabilní podobu a mění se v závislosti na subjektivních preferencích a zadlužení pracovníků. Proto není možno maximalizovat zisk analyticky odvozením křivky mezních nákladů z křivky agregátní nabídky. Zaměstnavatel tedy musí neustále experimentovat se zvyšováním a snižováním mzdy, aby stanovil mzdu maximalizující zisk.

4 Příjem z mezního produktu práce znamená zvýšení příjmu z produktu získaného každou novou jednotkou práce. Jednotkou práce v tomto modelu nerozumíme jednoho pracovníka navíc, ale jednu hodinu práce navíc. Pokud bychom chtěli trvat na tom, že jednotka práce, která mění mezní produktivitu, je jeden pracovník, mohli bychom to snadno provést tak, že by funkce mezní produktivity nebyla spojitě lineární, ale měla by tvar schodiště – byla by konstantní pro počet hodin odpracovaný jedním pracovníkem a skokově by se měnila u dalšího pracovníka.

Jak jsem již zmínil, modelujeme dvě možné situace: jednak může být mzda po celou dobu běhu modelu fixní a zaměstnavatel jí jednoduše přizpůsobí poptávku po práci, anebo může zaměstnavatel nabídku práce měnit. První situace nám umožňuje určit, jak se pro různé stabilní hodnoty mezd budou vyvíjet parametry jako je zisk, poptávka po práci, nezaměstnanost, zadluženost pracovníků a jejich úspory.

V druhé situaci mění zaměstnavatel výšku mezd o relativně malou částku a sleduje, zda mu změna nabídky práce, kterou vyvolá změna mezd, přinese vyšší zisk či nikoliv. Pokud se zisk zvýší, pokračuje zaměstnavatel ve změně mezd stejným směrem (tj. stále je zvyšuje nebo snižuje) dokud se zisk zvyšuje. Pokud se zisk sníží, obrátí zaměstnavatel směr změny mezd. Situace, kdy zaměstnavatel mění mzdy, umožňuje zjistit, za jakých podmínek dotlačí zaměstnavatel mzdy až k existenčnímu minimu a kdy zůstanou mzdy na úrovni umožňující pracovníkům blahobytný život.

Blahobyt je obtížné definovat, protože se na něm podílejí jak nezaměstnanost, tak výše mezd a liší se pro různé počty pracovníků. Nakonec jsem se rozhodl definovat míru blahobytu jako podíl pracovníků, kteří za dobu běhu modelu naspořili desetinásobek statusového minima. V dalším textu budeme pojmem „blahobytný život“ rozumět situaci, která se blíží maximu blahobytu.

Mzdy umožňující pouze existenční minimum představují pro zaměstnavatele vždy globální maximum zisku, neboť pracovníci jsou nuceni nabízet maximum práce za minimální mzdy. Ovšem pokud je počáteční úroveň mezd nastavena dost vysoko, může zaměstnavatel při postupném hledání maxima zisku malými změnami mezd uváznout v lokálním maximu zisku. K tomu dojde tehdy, když nabídka práce bude nižší než je poptávka po práci a zvyšující se mzdy budou pracovníky díky substitučnímu efektu motivovat k výrazně vyšší nabídce práce. To odpovídá situaci, kterou předpokládá učebnicová ekonomie.

Zaměstnavatel však vždy uspokojí jen takovou nabídku práce, která je nižší než poptávka po práci. Pokud se nabídka práce bude zvyšovat nad poptávku, na vyšší zisku už to mít vliv nebude.

Omezením modelu je abstrahování od trhu s výrobky. Model nezachycuje, zda zaměstnavatel dokáže své výrobky prodat. Předpokládáme předem danou funkci příjmu z mezního produktu práce, která určuje poptávku po práci při daných mzdách. Model nepředpokládá existenci zpětné vazby, díky které by mzdy a poptávka po práci ovlivňovaly příjem z mezního produktu práce. K tomu by docházelo, pokud by například pracovníci nakupovali větší část zboží vyráběného

zaměstnavatelem. Pak by nízké mzdy, nebo nízká zaměstnanost mohl vést k poklesu poptávky po tomto zboží v důsledku snížení kupní síly pracovníků. V globalizovaném světě se však často zboží vyvážá a spotřebuje v zahraničí nebo na jiném místě, než je vyráběno. Zaměstnavatel tak může snížení kupní síly svých pracovníků kompenzovat zvětšením množství vyvážené zboží. Pro takového zaměstnavatele je předpoklad pevné funkce příjmu z mezního produkt ospravedlnitelný.

### 2.3.2 Výběr pracovníků

Pokud je agregátní nabídka práce nižší než poptávka zaměstnavatele, přijme zaměstnavatel všechny zaměstnance, až na ty, kteří nemohou být přijati kvůli strukturální nezaměstnanosti (v modelu je její výše 3%).

Pokud je nabídka práce vyšší než poptávka, přijímá zaměstnavatel do práce nejprve ty pracovníky, kteří už v minulém kole pracovali. S jistou nízkou pravděpodobností (v modelu 0.02) jsou však někteří pracovníci zaměstnaní v minulém kole propuštěni (přirozená fluktuace). Teprve poté jsou přijímáni další pracovníci, až dokud není poptávka uspokojena. Navíc jsou přednostně přijímáni ti pracovníci (ať už byli nebo nebyli v předchozím kole zaměstnaní), kteří nabízejí více práce (což je dáno buď jejich vyšším statusovým minimem nebo zadlužeností nebo obojím).

Jedním z největších zjednodušení modelu je to, že zaměstnavatel nerozlišuje kvalifikaci, schopnosti nebo různé pracovní pozice pracovníků a všichni dostávají stejnou hodinovou mzdu. To je samozřejmě nerealistické, pokud bychom modelovali celou škálu pracovních příležitostí. Ale pro užší profesní skupinu jako jsou třeba dělnické profese, státní správa nebo střední management platí, že jejich mzdy zůstávají v relativně kompaktních mzdových mantinelech, kde rozdíly mezi nejnižší a nejvyšší odměnou nedosahují řádových rozdílů. Stejně mzdy pro všechny pracovníky proto můžeme chápat jako hrubou aproximaci platů v nějaké profesní skupině.

### 2.3.3 Interakce mezi pracovníky probíhá nepřímo přes zaměstnavatele

Pracovníci spolu přímo neinteragují. To odpovídá situaci atomizovaných pracovníků, kteří nejsou schopni se sdružovat do odborových hnutí nebo se jinými politickými prostředky bránit snižování mezd.

Pracovníci však spolu interagují nepřímo skrze zaměstnavatele. Pokud je poptávka po práci nižší než nabídka, vzájemně se vytlačují z práce. Zaměstnaný

pracovník, který si nevezal půjčku a pobírá mzdu dostačující na zaplacení existenčního minima, nabízí méně práce než agent se stejným statusovým minimem, který si půjčku vzal. Pracovník, který nemá práci, rychle spotřebuje případné úspory, musí si vzít půjčku a začne nabízet více práce. Zaměstnavatel preferuje pracovníky nabízející více práce. Nezaměstnaní pracovníci s půjčkou tedy vytlačují zaměstnané pracovníky bez půjčky.

### 2.3.4 Nastavení modelu

Pro všechny běhy modelu byly fixně nastaveny následující parametry:

#### Na straně pracovníků:

- Existenční minimum: 1000
  - *Poznámka: Hodnoty ex. minima a mezd jsou v abstraktních měnových jednotkách.*
- Exponent užitku ze statusové spotřeby: 1.85
- Exponent negativního užitku z času obětovaného práci: 0.85
  - *Poznámka: Velikosti exponentů byly nastaveny tak, aby individuální nabídkové funkce odpovídaly předpokládané nabídce v realitě. K přesnější kalibraci by byl potřeba empirický výzkum.*

#### Na straně zaměstnavatele:

- Strukturální nezaměstnanost: 3%
- Pravděpodobnost propuštění: 0.02
- Příjem z mezního produktu, resp. poptávková funkce byla dána lineární rovnicí:  $L = 2000 - W$ 
  - *Poznámka: Při nulové mzdě je tedy zaměstnavatel ochoten poptávat za časovou jednotku průměrně 2000 člověkohodin práce denně. S rostoucí mzdou (odpovídající příjmu z mezního produktu) se poptávané množství práce lineárně snižuje se směrnici -1. Zaměstnavatel přestává poptávat práci při velikosti hodinové mzdy 2000.*

Při opakovaných spouštěních modelu pro tu samou

velikost populace byly hodnoty statusových potřeb v populaci pracovníků stejné<sup>5</sup>.

Proměnné parametry byly počáteční velikost mezd, počet pracovníků a pravděpodobnost, že si pracovník vezme spotřební půjčku.

Počáteční hodnoty hodinových mezd byly nastavovány pro různé běhy modelu v rozmezí 50 až 1500 s krokem 10. Hodnota mzdy 50 je ještě o něco nižší, než stačí na zajištění existenčního minima. V realitě by samozřejmě taková situace vedla ke společenským nepokojům, stávkám apod. Při maximální individuální nabídce práce 14hodin denně lze vydělat na existenční minimum ve výši 1000 až při hodinové mzdě cca 71.43.

Mzda byla nastavena na začátku a zaměstnavatel nemohl mzdu měnit. Počet pracovníků byl nastavován na hodnoty: {100, 200, 300}. Pravděpodobnost půjček byla nastavována na hodnoty: {0, 0.05, 0.1, 0.15, 0.2}. Každý běh modelu trval 100 kroků. Celkově bylo provedeno 2190 běhů modelu s tímto nastavením.

### 3 Výsledky experimentů

Výsledky experimentů v situaci s fixní mzdou jsou zachyceny na grafech na konci článku. Grafy 3.-23. jsou organizovány v tabulce. Řádky zachycují vývoj několika klíčových parametrů pro různé počáteční hodnoty mzdy (které zůstávají fixní po dobu běhu modelu). Barevně jsou odlišeny hodnoty pro různou pravděpodobnost půjček. V levém sloupci jsou hodnoty pro 100 pracovníků, v prostředním pro 200 pracovníků a v pravém pro 300 pracovníků. Každý bod v grafech v jednom řádku (tj. pro jeden parametr) odpovídá jednomu z konečných stavů 2190 běhů modelu. V prvním řádku, v pravém grafu je pomocná tabulka popisující barevné odlišení hodnot pro různé pravděpodobnosti půjček.

V prvním řádku sledujeme vývoj zisku zaměstnavatele. Vidíme, že zisk má globální maximum vždy pro minimální mzdu. Pro nižší počet pracovníků (tj. nižší agregátní nabídku práce) a nižší pravděpodobnost půjček existuje i lokální maximum zisku. Toto lokální maximum se objevuje při takových hodnotách mezd, kdy snížení/zvýšení mezd vede k relativně výraznému snížení/zvýšení nabídky práce. Pro větší míru zaměstnanosti a vyšší pravděpodobnost spotřebních

půjček toto maximum mizí a objevuje se pouze pomalejší růst zisku při snižování mezd.

V druhém řádku sledujeme vývoj uspokojení nabídky práce (měřené v průměrných hodinách práce za den). Vidíme, že při velmi nízkých mzdách jsou pracovníci nuceni nabízet maximum práce, se zvyšujícími se mzdou klesá nabízená práce až do okamžiku, kdy převládne substituční efekt. Pokles zisků zaměstnavatele souvisí právě s tímto poklesem nabídky práce. Při dalším zvyšování mezd může zaměstnavatel dosáhnout lokálního maxima zisků ve fázi vzestupu nabídky práce, ale ještě předtím než nabídka práce dosáhne poptávky po práci. V tomto bodě se vývoj uspokojené nabídky práce láme a se zvyšující mzdou kopíruje křivku poptávky po práci (zaměstnavatel poptává méně práce, než činí nabídka).

Ve třetím řádku sledujeme vývoj nezaměstnanosti. Nezaměstnanost vzniká, když je nabídka práce vyšší než poptávka. Pro nižší počet pracovníků není znát vliv pravděpodobnosti půjček. Teprve u vysokého počtu pracovníků je zřetelné, že vyšší pravděpodobnost půjček vede k nárůstu nabídky práce a následně k růstu nezaměstnanosti. Zadlužení pracovníci nabízejí více práce a navzájem se tak vytlačují ze zaměstnání.

Ve čtvrtém řádku sledujeme procento zadlužených pracovníků. Není zde rozlišeno, zda je jejich zadlužení malé či velké. Zadlužení může být jak důsledkem toho, že pracovník nemá na existenční minimum, tak toho, že si vezme spotřební půjčku. Připomeňme, že spotřební půjčku si pracovníci mohou vzít, jen pokud jsou aktuálně zaměstnání a mají nenulové úspory. Pro nižší pravděpodobnosti spotřebních půjček je zřetelná závislost procenta zadlužených na nezaměstnanosti – musí si brát půjčky, aby splatili existenční minimum. Pro vysoké pravděpodobnosti spotřebních půjček a nízké mzdy je tato závislost zřetelná méně, neboť je zadlužena většina pracovníků. I zaměstnaní pracovníci se nyní zadlužují, berou si spotřební půjčky, které jsou desetinásobně vyšší než existenční minimum a proto nestačí půjčky splácet. S rostoucími mzdami začne počet zadlužených pracovníků klesat, neboť je stále snazší půjčky splácet. Od jisté výše mezd se však začne projevovat převis nabídky práce nad poptávkou a počet zadlužených klientů začne kopírovat počet nezaměstnaných. V této fázi přestává být zřetelný vliv velikosti půjček na počet zadlužených pracovníků.

V pátém a šestém řádku sledujeme vývoj průměrných úspor a dluhu. Maximum úspor dosahují pracovníci v situaci, kdy je malá část pracovníků nabízejících nejvíce práce zaměstnána za extrémně vysoké mzdy, přičemž zbytek je nezaměstnaný. V této situaci je nabídka práce vyšší než poptávka po práci.

<sup>5</sup> Díky stejnému „seed“ generátoru náhodných čísel. U ostatních náhodných akcí (např. propouštění ze zaměstnání), byl nastaven náhodný seed.

Zaměstnavatelé nerealizují maximum zisků. Tato situace je tedy nepříznivá jak z hlediska velké části pracovníků, tak z hlediska zaměstnavatele.

Při extrémně nízkých mzdách je zadlužení vysoké pro všechny počty pracovníků, neboť ani zaměstnaní pracovníci nejsou schopni vydělat existenční minimum. Pro vyšší počty pracovníků se při velmi nízkých mzdách navíc projevuje vliv nezaměstnanosti. Při nízkých mzdách a plné zaměstnanosti je zadlužení nejprve relativně nízké. Zaměstnanci si sice vezmou spotřební půjčky, ale jejich mzda je celá spotřebována na splácení půjčky, takže nic nenaspoří a díky tomu si nemohou znovu půjčit. Když se ale mzdy zvýší, vydělávají zaměstnanci více, než vydají na splátkách a při vysokých pravděpodobnostech půjček si začínají brát další a další půjčky a celkový dluh začne růst. Při ještě vyšších mzdách jsou zaměstnanci schopni splácet půjčky ve stále kratším intervalu a dluh začne klesat. Nakonec vydělávají pracovníci tolik, že půjčku splatí z úspor okamžitě a vliv spotřebních půjček na dluh je eliminován. Při vysokých mzdách se však začne projevovat vliv vzrůstající nezaměstnanosti, která vede ke vzrůstu dluhu v důsledku půjček na splácení existenčního minima.

V sedmém řádku (na další straně nahoře) sledujeme vývoj blahobytu, který jsem definoval jako podíl pracovníků, jejichž úspory dosáhly na konci běhu modelu alespoň desetinásobku statusového minima potřeby. Nejvyšší míra blahobytu je dosahována při vyšším počtu pracovníků pro nižší mzdy, což je dáno zejména tím, že pro vyšší počty pracovníků začíná růst nezaměstnanost při nižších úrovních mezd. Můžeme si všimnout, že maxima blahobytu je dosaženo při o něco nižších mzdách, než je dosaženo maxima úspor – viz výše odstavec o úsporách. Vliv pravděpodobnosti spotřebních půjček na snižování blahobytu je zřetelný zejména při mzdách, které nejsou ani příliš nízké, ani příliš vysoké. Při nízkých mzdách je blahobyt nízký pro všechny pravděpodobnosti půjček. Při vysokých mzdách jsou všichni zaměstnaní pracovníci schopni pro všechny pravděpodobnosti půjček ušetřit desetinásobek statusového minima a velikost blahobytu bude pro danou mzdu záviset hlavně na míře nezaměstnanosti.

Grafy 24.-27. na konci článku zachycující závislost některých klíčových parametrů na pravděpodobnosti spotřebních půjček. Jednotlivé body v grafech zachycují průměr dané hodnoty pro různé mzdy, horní a dolní meze značí 90% interval spolehlivosti. Barevně rozlišeny jsou spojnice pro různé hodnoty počtu pracovníků.

Na grafu vlevo nahoře vidíme vývoj zisku. Zisk se vždy zvyšuje se zvyšující se pravděpodobností spotřebních

půjček. To je důsledek zvýšené nabídky práce zadlužených. Zvýšená nabídka práce vždy buď zisk zvýší, nebo ho nechá beze změny.

Na grafu vpravo nahoře vidíme vývoj nezaměstnanosti. S výjimkou nízkého počtu pracovníků, kdy je nabídka práce vždy uspokojena, roste nezaměstnanost s pravděpodobností spotřebních půjček. Toto je jeden z nejzajímavějších nalezených výsledků.

Na spodních grafech vidíme vývoj průměrného zadlužení a úspor. Dluhy rostou a úspory klesají se zvyšující se pravděpodobností spotřebních půjček.

## 4 Závěr

V rámci zde popsaného modelu je patrné, že pokud se v některých lokalitách nebo pro jisté profesní skupiny blíží trh práce situaci monopsonu, pak je zájmem zaměstnavatele maximalizujícího zisk snížit mzdy až na hranici existenčního minima. Při nízkém počtu pracovníků a nízké pravděpodobnosti půjček se však může stát, že zaměstnavatel, který mění mzdy jen po relativně malých krocích, uvázne v lokálním maximu zisku při velikosti mezd, která umožňuje pracovníkům blahobytný život.

Samozřejmě, kdyby zaměstnavatel mzdy snižoval dost dlouho (riskoval přitom krátkodobé snížení zisku), nebo kdyby mzdy snížil o velkou částku, mohl by se dostat do stejné situace, jaká nastává při nízkých počátečních mzdách. Takový manévr však model neumožňuje, protože předpokládáme, že tomu brání různá institucionální omezení. Je pravděpodobnější, že se zaměstnavatel bude spíše snažit snižovat mzdy postupně a minimalizovat svůj ušlý zisk podporou nabídky práce. Nabídku práce může podpořit nejen podpora spotřebních půjček, ale i další faktory zesilující obavy z nezaměstnanosti, které model nezachycuje, jako jen např. oslabování veřejných sociálních služeb<sup>6</sup>, dočasný import pracovní síly ze zahraničí apod.

Zaměstnavatel může též využívat kombinace obav ze ztráty zaměstnání a zadluženosti a nutit pracovníky dělat neplacené přesčasy. To přesně odpovídá situaci, ve které se snižují průměrné hodinové mzdy a zvyšuje se nabídka práce.

Zájmem pracovníků je naopak bránit se buď zavedením minimální mzdy, nebo alespoň nedovolit zaměstnavateli příliš velké snižování mezd, což může vést k tomu, že zaměstnavatel uvázne v lokálním maximu zisku. Pro

---

<sup>6</sup> Zejména podpory v nezaměstnanosti, podpory rodin s dětmi, zdravotnictví a školství.

pracovníky jako jednotlivce i skupinu je vysoce rizikové i při relativně vysokých mzdách a úsporách nejen brát si spotřební půjčky, ale i nechat se nadměrně motivovat statusovou spotřebou k vysoké nabídce práce. Vysoká nabídka práce představuje riziko, že jejich zaměstnavatelé budou moci v dlouhodobém horizontu stlačit mzdy na podstatně nižší hodnotu. Pracovníci nabízející hodně práce vytlačují ze zaměstnání ostatní pracovníky a ti se postupně dostávají do dluhů nikoliv kvůli spotřebním půjčkám, ale kvůli půjčkám nutným na splacení existenčního minima. Nezaměstnanost nevzniká jen v důsledku nechuti nezaměstnaných pracovat, ale i v důsledku toho, že je jejich „šťastnější“ kolegové vytlačili ze zaměstnání.

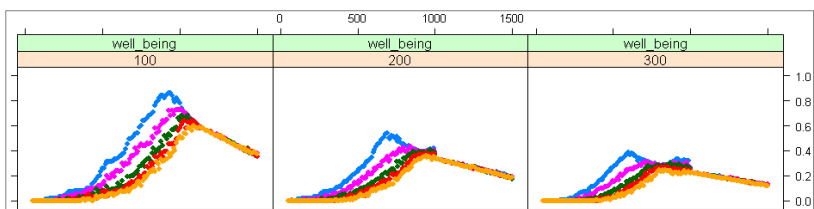
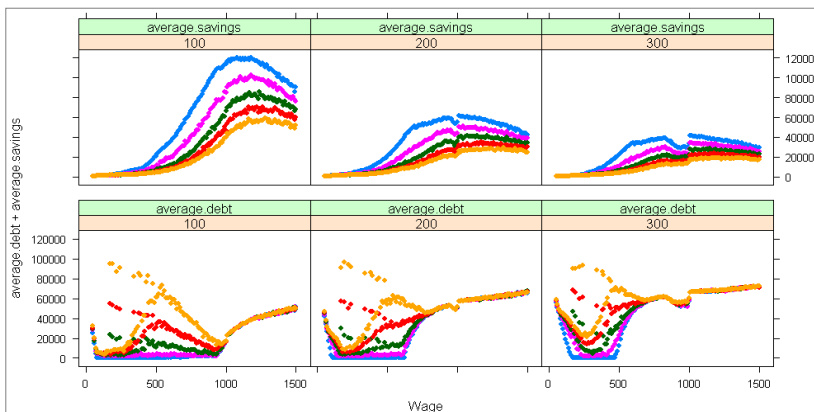
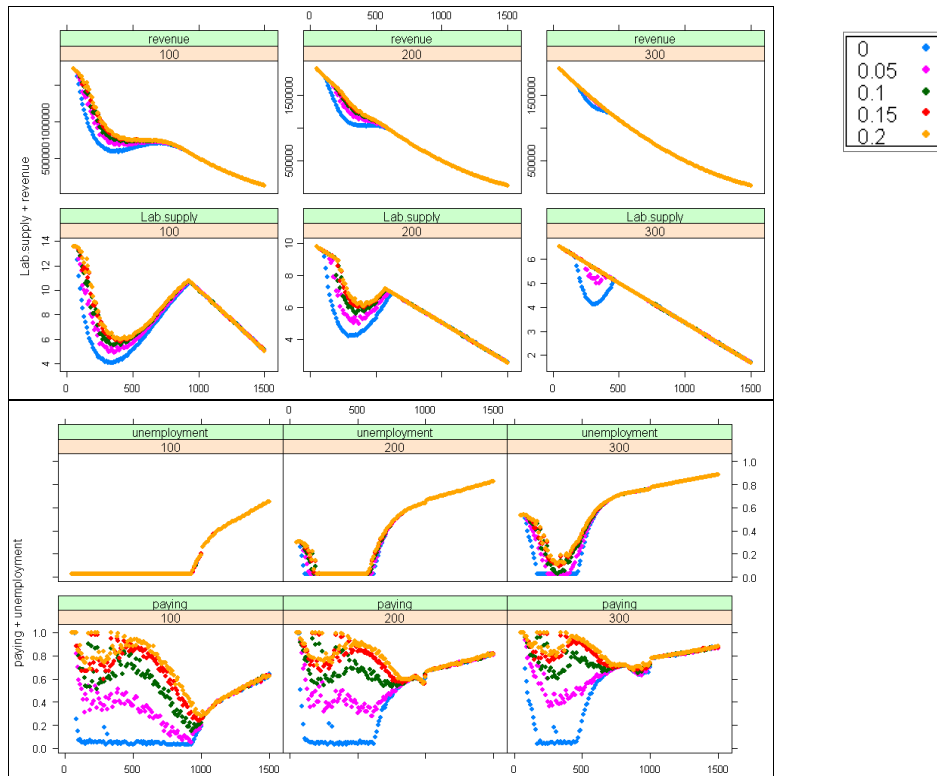
Model sice přímo neřeší otázku minimální mzdy, ale ukazuje následující skutečnosti: Pokud jsou mzdy tak vysoké, že agregátní nabídka práce roste se zvyšujícími se mzdami, pak můžeme souhlasit s učebnicovou ekonomickou teorií, že zvýšení minimálních mezd bude zvyšovat nezaměstnanost (ovšem jen pokud agregátní nabídka práce převyšuje poptávku). Model však ukazuje i to, že pokud jsou minimální mzdy tak nízké, že agregátní nabídka práce roste se snižujícími se mzdami,

pak povede ke zvýšení nezaměstnanosti naopak další snížení či dokonce zrušení minimálních mezd.

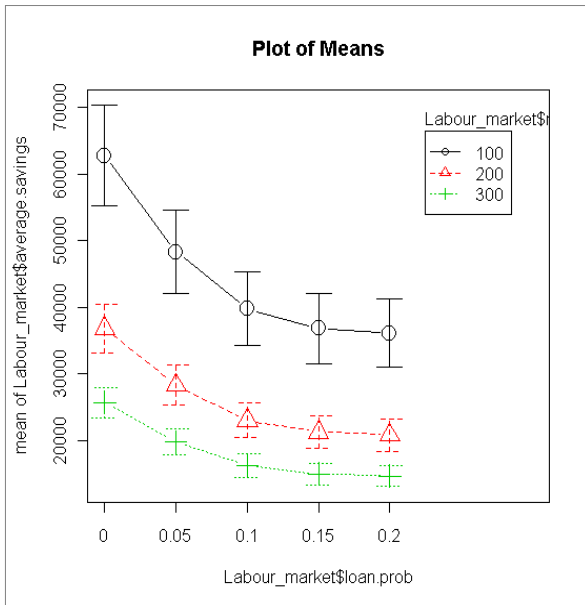
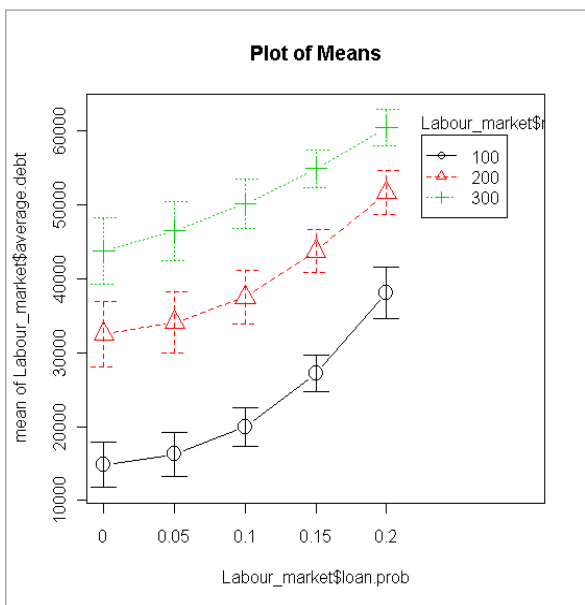
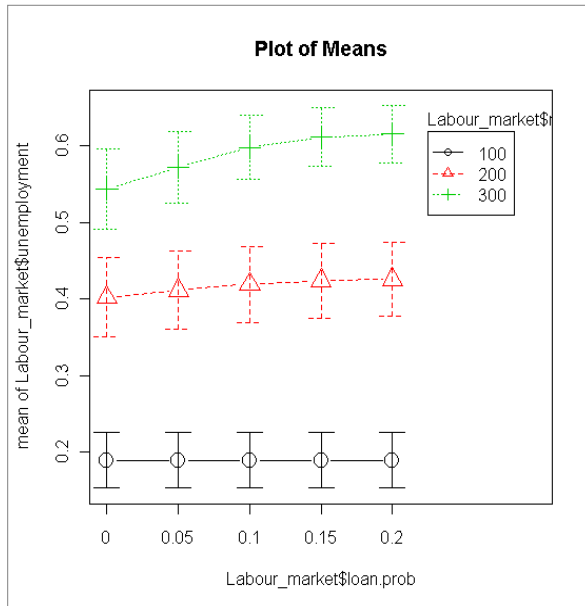
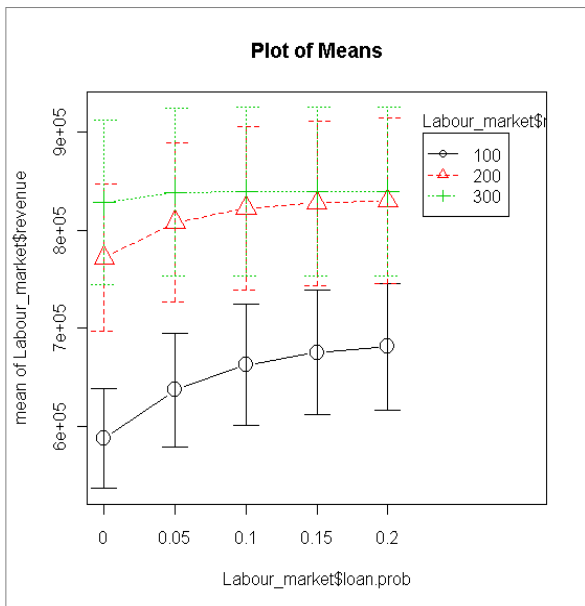
Všechny zde uvedené závěry jsou samozřejmě platné jen v rámci zjednodušujících předpokladů, které jsem stanovil při tvorbě modelu. Aplikace těchto závěrů v reálném světě by bylo nutné doplnit empirickým výzkumem.

## Literatura

- [1] Burian, J.: *Reaktivní multiagentní modely v ekonomii*. Doktorská práce. Praha: Vysoká škola ekonomická, 2010.
- [2] W. Samuelson, W. Nordhaus: *Ekonomie*. Praha: Svoboda, 1991.
- [3] J. Soukupová, B. Hořejší, L. Macáková, J. Soukup: *Mikroekonomie*. Management Press 2002.



Grafy 3.-23. – Hodnoty zisku, uspokojené nabídky práce, nezaměstnanosti, zadlužených pracovníků, průměrných úspor, průměrných dluhů a blahobytu na konci běhu pro různé hodnoty fixních mezd a kombinace počtu pracovníků a pravděpodobnosti půjčky.



Grafy 24.-27. – Průměrné hodnoty zisku nezaměstnanosti, dluhu a úspor při fixní mzdě pro různé kombinace počtu pracovníků a pravděpodobnosti půjčky.



## Preferencia k deliberácii: (ne)istota úspechu v kognitívnych úlohách

Vladimíra Čavojová<sup>1</sup>, Eva Ballová Mikušková<sup>2</sup>, Róbert Hanák<sup>3</sup>

Ústav experimentálnej psychológie, Slovenská akadémia vied  
Dúbravská cesta 9, Bratislava

<sup>1</sup>vladimira.cavojova@savba.sk, <sup>2</sup>eva.ballova-mikusko@savba.sk, <sup>3</sup>robohanak@gmail.com

### Abstrakt

Cieľom príspevku je identifikovať intuitívne a deliberatívne typy ľudí na základe ich výsledkov v dvoch dotazníkoch intuitívneho spracovania informácií (REI a PID) a porovnať ich výkon v kognitívnych úlohách (CRT). Identifikovali sme 4 typy ľudí na základe ich preferencie k deliberácii a intuícii. Výsledky ukázali, že preferencia k deliberácii nezaručuje, že sa participant nedá „nachytať“ na chybný výsledok, ani že intuitívni ľudia majú menší sklon rozmýšľať nad kognitívnymi úlohami. V príspevku tiež porovnávame oba dotazníky na intuícii z hľadiska ich efektívnosti správne identifikovať intuitívne typy ľudí a výsledky rozoberáme aj vo svetle sociálnej dezirability vyzerať „analyticky a uvážlivo“, najmä v kontexte testovej situácie.

### 1 Úvod

Pri skúmaní rozhodovania stála dlhé roky v centre pozornosti stratégia založená na dôkladnej a objektívnej analýze všetkých dostupných informácií, bolo ju možné matematicky vyjadriť, vzbudzovala dôveru a úctu a preto bola aj prezentovaná ako správny a istý spôsob správneho rozhodovania [1]. Väčšina ľudí sa však nerozhoduje striktne analyticky, racionálne, deliberatívne, preto sa do popredia vedeckého záujmu dostala intuícia, automatické a nevedomé procesy v rozhodovaní, hoci sa kvôli kognitívnym omylom intuitívneho rozhodovania [2] považoval tento spôsob uvažovania za menej racionálny.

#### 1.1 Racionálne verzus intuitívne uvažovanie a rozhodovanie

Deliberatívne a intuitívne uvažovanie a rozhodovanie je najčastejšie vnímané v rámci duálnych procesov: intuícia a deliberácia ako dva póly jednej dimenzie alebo ako samostatné dimenzie, ktoré sa dopĺňajú. V rámci teórie duálnych procesov [3] býva intuícia stotožňovaná so Systémom 1 (rýchle, automatické, paralelné, nenútené, asociatívne, implicitné a často emočne nabité uvažovanie a rozhodovanie) a deliberácia so Systémom 2 (pomalšie, postupné, namáhavé, často vedome monitorované a

úmyselne kontrolované, relatívne flexibilné, pravidlami sa riadiace uvažovanie a rozhodovanie).

Vo väčšine prípadov máme automatický systém (intuícii) nastavený ako predvolený: pri každom uvažovaní alebo riešení nejakého problému nevieme zabrániť automatickému systému pracovať. Aj napriek tomu, že výsledok uvažovania alebo rozhodovania je produktom úmyselného systému myslenia, zdá sa, že činnosť automatického systému je vždy prítomná [4], či už sa analytické a intuitívne výstupy zhodujú alebo nie. To znamená, že ľudia majú vždy nejakú automatickú (hoci minimálnu) reakciu na každý podnet. Výnimkou sú situácie, kedy je podnet prezentovaný vo forme evokujúcej analytické myslenie (hlavne ak človek vie, ako má postupovať).

Správanie ľudí je ovplyvňované oboma systémami, niektoré rozhodnutia sú ovplyvnené viac jedným systémom ako druhým, takže možno povedať, že v mnohých prípadoch je intuícia rovnako dôležitá ako racionálne uvažovanie [5]. Podľa niektorých výskumníkov zaoberajúcich sa duálnymi teóriami rozhodovania sú silné prvky racionality v oboch systémoch [6]. Ľudia sa však líšia v tom, nakoľko majú tendenciu spoliehať sa na intuitívny, zážitkový a na deliberatívny, racionálny systém a táto individuálna preferencia pre rozhodovaciu stratégiu je silným prediktorom výberu tejto stratégie [7]. U nás napr. Sirota [8] zistil, že participant preferujúci racionálnu stratégiu rozhodovania majú výhodu oproti participantom preferujúcim intuitívne stratégie v úlohách vyžadujúcich si explicitné zdôvodnenie rozhodnutia (merané adaptovaným Racionálno-experienciálnym inventárom od Paciniho a Epsteina, [9]).

#### 1.1.1 Rozdiely medzi intuitívnymi a deliberatívnymi v kognitívnych úlohách

Wittemanová s kolektívom [10] testovali dotazník REI v súvislosti riešením desiatich rôznych kognitívnych úloh na veľmi malej vzorke ( $n=28$ ); pričom jednou z nich bolo aj CRT od Fredericka ([11]; u nás CRT validizovali Sirota a Juanchich, [12]). U všetkých týchto úloh bola intuitívna (nesprávna) a racionálna (správna odpoveď).

Autori našli významné pozitívne vzájomné vzťahy ( $r = 0,4$ ,  $p = 0,034$ ) medzi REI-R, racionálnou škálou a výkonnosťou v úlohách a súčasne negatívne vzťahy medzi experienciálnou REI -E škálou a výkonnosťou v úlohách ( $r = -0,57$ ,  $p = 0,002$ ). Hoci tieto výsledky vyzerajú veľmi sľubne, je nevyhnutné poznamenať významné obmedzenie tohto výskumu. Prvou je malá vzorka a druhou je fakt, že boli opätovne použité úlohy, ktoré používal sám autor dotazníka REI v minulosti a preukázali už svoju významnú súvislosť.

Aj u nás napr. Sirota [8] zistil, že participanti, ktorí riešili usudzovacie problémy normatívne adekvátne, dosahovali vysokú mieru racionálnej schopnosti a postoja a zároveň nízke sýtenie zážitkovým postojom a schopnosťou (merané adaptovaným Racionálno-experienciálnym inventárom od Paciniho a Epsteinov [9]). To môže tiež naznačovať, že participanti preferujúci racionálnu stratégiu rozhodovania majú výhodu oproti participantom preferujúcim intuitívne stratégie v úlohách vyžadujúcich si explicitné zdôvodnenie rozhodnutia.

### 1.1.2 Meranie intuície a racionality

Pri snahe merať intuíciu narážame na niekoľko problémov. Po prvé, samotná definícia intuície – rôzni autori vnímajú a definujú intuíciu odlišne (prehľad definícií uvádzajú napríklad v Čavojová [1], Ballová Mikušková [13], Hanák [14]). Po druhé, existuje niekoľko typov intuície (afektívna, kumulatívna, asociatívna; napr. Glöckner a Wittemanová [15] a po tretie, na intuíciu možno nazerať z rôznych uhlov pohľadu, napr. v zmysle osobnostnej črty, kognitívneho štýlu, preferencie alebo aktuálne používanej stratégie. V súčasnosti sa pri skúmaní intuície používajú rôzne meracie nástroje, otázná je však adekvátnosť ich použitia vzhľadom na cieľ výskumu.

## 2 Metódy

### 2.1 Participanti

Výskumu sa zúčastnilo 169 (114 žien) študentov externej formy štúdia s ekonomickým zameraním vo veku 20–60 rokov ( $M=27,51$ ,  $SD=7,64$ ). 82% participantov bolo v čase zberu dát ekonomicky aktívnych.

### 2.2 Metódy

Participanti vyplňali slovenské verzie dvoch dotazníkov spôsobu spracovávania informácií, Betschovej [16] *Dotazník preferencie k intuícii a deliberácii* (PID) a Paciniho a Epsteinov [9] *Racionálno-zážitkový inventár* (REI). Participanti ďalej riešili tri kognitívne úlohy

*Cognitive Reflective Task* vytvorené Frederickom [11]. Všetky metódy boli participantom administrované online.

### 2.2.1 Dotazník preferencie k intuícii a deliberácii (PID)

Betschová [16] vyvinula PID (The Preference for Intuition and Deliberation) ako nástroj na mapovanie afektívnej intuície, pričom intuícia a deliberácia v PID predstavujú dve dimenzie, nie dva protipóly jednej dimenzie. Intuícia meraná subškálou PID-I je založená na implicitných vedomostiach a je definovaná ako spôsob rozhodovania, ktorý využíva priame afektívne reakcie ako rozhodovacie kritérium. Na druhej strane, deliberácia meraná subškálou PID-D je založená na explicitných vedomostiach a pravidlách rozhodovania sa [17].

PID pozostáva z 18 výrokov (9 pre každú subškálu), ktoré sú hodnotené na 5 bodovej škále (1 znamená „vôbec nesúhlasím“ a 5 znamená „úplne súhlasím“).

Vnútorňá konzistencia PID (meraná Cronbachovou alfou) bola 0,77 pre PID-I a 0,79 pre PID-D [16]. V slovenskej populácii bola nameraná vnútorňá konzistencia PID 0,73 pre PID-I a 0,82 pre PID-D.

### 2.2.2 Racionálno – zážitkový inventár (REI)

Epstein, Pacini, Denes-Raj a Heier [18] vytvorili REI (The Rational Experiential Inventory) na základe Kognitívno-zážitkovej seba-teórie: REI podľa nich meria, ktorý z dvoch paralelných spôsobov spracovania informácií ľudia používajú, racionálny alebo zážitkový. Racionálny systém operuje primárne s vedomou úrovňou a je úmyselný, analytický, verbálny. Zážitkový systém je automatický, nevedomý, holistický, asociatívny, neverbálny a je blízko spätý s emóciami.

Intuícia v REI mapuje skôr schopnosť ako osobnostnú črtu.

Používali sme verziu so 40 položkami [9], kde 20 výrokov je racionálnych a 20 zážitkových. Celkovo REI mapuje štyri dimenzie: racionálna schopnosť, racionálny postoj, zážitková schopnosť, zážitkový postoj [8].

### 2.2.3 Test kognitívnej reflexie (CRT)

CRT (Cognitive Reflection Test) je jednoduchý merací nástroj, ktorý bol navrhnutý Frederickom [11] tak, aby zachytil špecifické kognitívne schopnosti participanta, konkrétne hodnotí schopnosť potlačiť spontánnu, intuitívnu, avšak nesprávnu odpoveď (Systém 1) a nahradiť ju uvážanou, deliberatívnou a zároveň správnu odpoveďou (Systém 2).

CRT pozostáva z troch relatívne jednoduchých úloh s možnosťou voľných odpovedí, ktoré majú potenciál

evokovať rýchle, intuitívne avšak nesprávne odpovede. Všetky správne odpovede indikujú kognitívny typ (Systém 2), žiadna správna odpoveď indikuje intuitívny typ (Systém 1).

## 2.3 Postup

Pri identifikácii kognitívnych typov z CRT sme postupovali podľa Fredericka [11]: participanti so žiadnou správnou odpoveďou boli označení ako intuitívni („slabí“ vo Frederickovom ponímaní), participanti so všetkými správnymi odpoveďami boli označení ako analytici („silní“ vo Frederickovom ponímaní). Participanti, ktorí mali 1 alebo 2 správne odpovede, boli zaradení do typu strední. Analyzované boli iba extrémne skupiny, intuitívni a analytici.

Pri identifikácii participantov s intuitívnou verzus deliberatívnou preferenciou spracovávaní informácií prichádzali do úvahy 4 spôsoby spracovania skóre PID a REI: 1) jednoduché sčítanie skóre v jednotlivých subškálach, 2) metóda median split, 3) nútená hranica, 4) klastrová analýza. Na základe zistení z analýz sme skóre PID a REI spracovávali metódou median split, podobne ako Betschová [16].

V prvom kroku sme vypočítali medián pre jednotlivé sledované dimenzie (intuícia a deliberácia v PID a racionálna a zážitková v REI). Následne sme identifikovali 4 skupiny participantov podľa PID a rovnako 4 skupiny participantov podľa REI: (1) intuitívni – skórovali nad mediánom v intuícii/zážitkovej dimenzii a pod mediánom v deliberácii/racionálnej dimenzii, (2) deliberatívni – skórovali nad mediánom v deliberácii/racionálnej dimenzii a pod mediánom v intuícii/zážitkovej dimenzii, (3) zmiešaní – skórovali nad mediánom v oboch preferenciách, a (4) nevyhranení – skórovali pod mediánom v oboch preferenciách.

V ďalšom kroku sme porovnávali PID a REI z hľadiska ich efektívnosti správne identifikovať intuitívne verzus deliberatívne typy ľudí. Nakoniec sme porovnávali výkon jednotlivých typov participantov v kognitívnych úlohách.

## 3 Výsledky

### 3.1 Deskriptívna štatistika

V Tabuľke 1 uvádzame deskriptívnu štatistiku získaných skóre v PID, REI a CRT pre celý výskumný súbor (minimálne a maximálne hodnoty, medián, priemer, štandardná odchýlka).

**Tabuľka 1** Deskriptívna štatistika

	N	Min.	Max.	Mdn	M	SD
PID-D	169	9	44	36	35,75	4,84
PID-I	169	19	45	32	31,39	5,33
REI-R	169	49	104	78	77,66	10,10
REI-E	169	43	117	79	78,80	14,08
CRT	169	0	3	0	0,75	1,08

Na základe vypočítaných mediánov pre subškály v PID a REI sme metódou median split identifikovali typy participantov na základe ich preferencie k intuícii verzus deliberácii. V Tabuľke 2 uvádzame rozloženie a prieniky typov PID a REI.

**Tabuľka 2** Krížová tabuľka typov PID a REI

		PID-D	PID-I	PID-M	PID-N	spolu
REI-R	N	22	2	3	12	39
	%	13,0	1,2	1,8	7,1	23,1
REI-Z	N	3	10	15	5	33
	%	1,8	5,9	8,9	3,0	19,5
REI-M	N	7	10	30	5	52
	%	4,1	5,9	17,8	3,0	30,8
REI-N	N	8	10	6	21	45
	%	4,7	5,9	3,6	12,4	26,6
Spolu	N	40	32	54	43	169
	%	23,7	18,9	32,0	25,4	100,0

PID-D – deliberatívni (PID), PID-I – intuitívni (PID), PID-M – zmiešaní (PID), PID-N – nevyhranení (PID), REI-R – racionálni (REI), REI-Z – zážitkoví (REI), REI-M – zmiešaní (REI), REI-N – nevyhranení (REI)

### 3.2 Komparatívna štatistika

Porovnávali sme jednotlivé rozhodovacie typy (PID a REI) v ich výkone v kognitívnych úlohách CRT (one-way ANOVA). Nezistili sme žiadne významné rozdiely, ani pre typy PID, ani pre typy REI v kognitívnych úlohách.

## 4 Diskusia

Cieľom príspevku bolo na jednej strane porovnať rôzne metódy identifikovania intuitívnych a deliberatívnych ľudí, ako aj dva rôzne dotazníky merajúce intuíciiu, a na druhej strane zistiť, či preferencia k intuícii resp. deliberácii má vplyv na správnosť odpovedí v kognitívnych úlohách.

Ukázalo sa, že najspôľahlivejším spôsobom identifikácie intuitívnych ľudí je metóda *median split*

(ktorú používala aj Betschová, autorka PID), hoci si vyžaduje výpočet mediánu osobitne pre každý skúmaný súbor. Na druhej strane týmto spôsobom dokážeme zachytiť najväčší počet participantov a zároveň ich dobre diferencovať. Zistili sme, že je dôležité porovnávať nielen obe extrémne skupiny (intuitívni vs. deliberatívni), ale že je dôležité pozrieť sa aj na výhody toho, keď má človek vysokú úroveň obidvoch preferencií. Dá sa totiž predpokladať, že aj intuícii vedia využiť lepšie ľudia, ktorí majú zároveň sklony k deliberácii, a tým pádom vedia lepšie posúdiť, na aký typ problémov je vhodné použiť jednu či druhú stratégiu. Na rozdiel od klasického kritického prístupu k intuícii ako k skratkovitým, heuristickým riešeniam, ktoré sa odráža aj v úlohe CRT [11] a v programe heuristik a zaujatostí [2], novšie prístupy poukazujú na to, že za určitých okolností je lepšie spoľahnúť sa na intuitívne riešenia [19]. Podobne Dijksterhuisova a Nordgrenova teória nevedomého myslenia [20] postuluje, že na komplexné problémy je lepšie použiť nevedomé myslenie za predpokladu, že máme k dispozícii všetky relevantné informácie, ktoré boli získané vedome.

Okrem toho sme porovnávali dotazníky PID a REI z hľadiska ich schopnosti spoľahlivo identifikovať intuitívnych a deliberatívnych ľudí. Z výsledkov vyplýva, že oba dotazníky sa do určitej miery prekrývajú a dokážu spoľahlivo identifikovať tých istých participantov ako intuitívnych, resp. deliberatívnych. Zároveň však merajú trochu odlišné koncepcie intuície. Kým PID konceptualizuje intuícii ako tendenciu k intuícii, resp. deliberácii, navyše pravdepodobne založenú na motivácii, REI meria skôr preferenciu k zážitkovému alebo kognitívnemu spracovávaniu založenú na schopnosti.

Hlavným cieľom bolo však zistiť, či sklony k deliberácii budú mať pozitívny vplyv na správnosť riešenia predložených kognitívnych úloh, pričom sme vychádzali aj zo Sirotových [8] zistení, že racionálna škála REI súvisela s normatívne adekvátnejším riešením bayesiánskych úloh. Naše výsledky ukázali, že preferencia k deliberácii nezaručuje, že sa participant nedá „nachytať“ na chybný výsledok, ani že intuitívni ľudia majú menší sklon rozmýšľať nad kognitívnymi úlohami.

Jedným z možných vysvetlení tohto zistenia môže byť aj to, že kontext testovej situácie mohol spôsobiť, že aj participant, ktorí mali väčšiu preferenciu intuície, v dotazníkoch priradzovali vyššiu hodnotu odpovediam, ktoré vnímali ako „racionálnejšie“. Keďže išlo o študentov, ktorí vyplňali dotazníky v rámci účasti na kurze, mohli to vnímať ako „skúšku“ a poskytovať sociálne žiadúcejšie odpovede. Preto je predmetom skúmania v našom nasledujúcom výskume, nakoľko preferencia k deliberácii (PID) či racionálna škála (REI) súvisia so sociálnou dezirabilitou.

Iné vysvetlenie súvisí s prediktívnou validitou oboch dotazníkov, nakoľko dokážu predpovedať výkon v reálnych rozhodovacích úlohách či situáciách v bežnom živote. Na problém prediktívnej validity dotazníkov merajúcich intuícii poukázal aj Hanák [21], ktorý prezentuje prehľad výskumov mapujúcich konštruktívnu a prediktívnu validitu dotazníkov PID a REI. Napríklad vo výskume Wittemanovej a jej výskumného tímu [10] kognitívne úlohy (merané CRT) pozitívne súviseli s racionálnou škálou a negatívne so zážitkovou škálou dotazníka REI, no výsledky boli získané na veľmi malej vzorke [21]. Okrem toho, hoci je CRT kognitívna úloha, stále ide len o iný typ testovej úlohy a nie reálne (racionálne) správanie v rozhodovacej situácii.

Preferenciu k intuícii resp. deliberácii je preto nevyhnutné skúmať v kontexte reálnych rozhodovacích úloh v experimentálnom dizajne. Na záver však môžeme konštatovať, že nie na všetky úlohy platí zdravý rozum a že ani sklony k prílišnému rozvažovaniu nám nezaručia, že sa nedáme výskumníkom nachytať.

## Pod'akovanie

Tento príspevok vznikol za podpory grantovej agentúry VEGA v rámci grantovej úlohy 2/0064/13 s názvom Rozhodovanie expertov: využívanie intuície expertmi pri riešení strategických úloh.

## Literatúra

- [1] V. Čavojová, O racionalite intuície, in *Rozhodovanie a usudzovanie IV*, R. Hanák, Ed. Bratislava: Ústav experimentálnej psychológie SAV (v tlači) 2013.
- [2] A. Tversky and D. Kahneman, Judgment under Uncertainty: Heuristics and Biases, *Science*, 185(4157): 1124–1131, 1974.
- [3] K. E. Stanovich and R. F. West, Individual differences in reasoning: implications for the rationality debate?, *The Behavioral and brain sciences*, 23, 645–726, 2000.
- [4] R. M. Hogarth, *Educating Intuition*. Chicago: The University of Chicago Press, Ltd., 2001, p. 335.
- [5] E. Sadler-Smith and E. Shefy, The intuitive executive: Understanding and applying 'gut feel' in decision-making, *Academy of Management Executive*, 18(4):76–91, 2004.
- [6] P. Slovic, M. Finucane, E. Peters, and D. G. Macgregor, Rational actors or rational fools: implications of the affect heuristic for behavioral economics, *Journal of Socio-Economics*, 31, 329–342, 2002.

- [7] C. Betsch, *Intuition in judgment and decision making*. New York: Lawrence Erlbaum Associates, 2008.
- [8] M. Sirota, Individuálne rozdiely v racionálnom usudzovaní. Kognitívne parametre a komputačné stratégie v „bayesovskom“ usudzovaní v úlohách s odlišnými číselnými reprezentáciami, Univerzita Komenského, 2008.
- [9] R. Pacini and S. Epstein, The relation of rational and experiential information processing styles to personality, basic beliefs, and the ratio-bias phenomenon, *Journal of Personality and Social Psychology*, 76(6): 972–987, 1999.
- [10] C. Witteman, J. van den Bercken, L. Claes, and A. Godoy, Assessing Rational and Intuitive Thinking Styles, *European Journal of Psychological Assessment*, 25(1): 39–47, 2009.
- [11] S. Frederick, Cognitive Reflection and Decision Making, *Journal of Economic Perspectives*, 19(4):25–42, 2005.
- [12] M. Sirota and M. Juanchich, Role of numeracy and cognitive reflection in Bayesian reasoning with natural frequencies, *Studia Psychologica*, 53(2):151–161, 2011.
- [13] E. Ballová Mikušková, “Intuícia: dobrý sluha, zlý pán? Kritický pohľad na intuíciu: Kedy a pre koho je intuícia výhodná?,” in *Rozhodovanie a usudzovanie IV*, R. Hanák, Ed. Bratislava: Ústav experimentálnej psychológie SAV (in press), 2013.
- [14] R. Hanák, “Nástroje na meranie intuície,” in *Rozhodovanie a usudzovanie IV*, R. Hanák, Ed. Bratislava: Ústav experimentálnej psychológie SAV (in press), 2013.
- [15] A. Glöckner and C. Witteman, “Foundation for tracing intuition: models, findings, categorizations,” in *Foundation for tracing intuition: Challenges and Methods*, A. Glöckner and C. Witteman, Eds. Hove and New York: Psychology Press, 2010, 1–23.
- [16] C. Betsch, Präferenz für Intuition und Deliberation (PID): Inventar zur Erfassung von affekt- und kognitionsbasiertem Entscheiden, *Zeitschrift für Differentielle und Diagnostische Psychologie*, 25(4): 179–197, 2004.
- [17] C. Betsch and P. Iannello, Measuring individual differences in intuitive and deliberative decision-making styles, in *Foundation for tracing intuition: Challenges and Methods*, A. Glöckner and C. Witteman, Eds. Hove and New York: Psychology Press, 2010, 251–278.
- [18] S. Epstein, R. Pacini, V. Denes-Raj, and H. Heier, Individual differences in Intuitive-Experiential and Analytical-Rational Thinking Styles, *Journal of Personality and Social Psychology*, 71(2): 390–405, 1996.
- [19] G. Gigerenzer, *Gut Feelings: Short Cuts to Better Decision Making*. London: Penguin Books Ltd, 2008, p. 288.
- [20] A. Dijksterhuis and L. F. Nordgren, “A Theory of Unconscious Thought,” *Perspectives on Psychological Science*, 1(2):95–109, 2006.
- [21] R. Hanák, Intuícia ako schopnosť vybudovaná expertmi verus intuícia ako kognitívny štýl, in *Rozhodovanie a usudzovanie IV*, R. Hanák, Ed. Bratislava: Ústav experimentálnej psychológie SAV (in press), 2013.



## O komunitách v sociálnych sieťach

David Chalupa

Ústav aplikovanej informatiky  
Fakulta informatiky a informačných technológií  
Slovenská technická univerzita v Bratislave  
Ilkovičova, 842 16 Bratislava  
Email: chalupa@fiit.stuba.sk

### Abstrakt

Pojem komunity sa v informatických vedách objavuje čoraz častejšie, napriek tomu ho však nie je jednoduché definovať. Komunita v sociálnej sieti je vo všeobecnosti zoskupenie ľudí, ktorí majú čosi spoločné, a preto je hustota prepojení v rámci komunity väčšia, než mimo nej. Tento pojem je zovšeobecniteľný aj pre siete z iných domén, kde sa de facto ako synonymum pojmu komunity využíva pojem zhuk. Formulácia problému detekcie komunít je však otvoreným problémom, ktorý má dopad nielen na zmyslupnosť dosiahnutých výsledkov, ale aj výpočtovú náročnosť optimalizačného problému detekcie komunít. V tomto príspevku diskutujeme možnosti formulácie problému detekcie komunít v sociálnych sieťach na úrovni, v ktorej pracujeme len s informáciou o kontaktoch jednotlivých používateľov - vrcholov siete. Krátko sumarizujeme výsledky o výpočtovej náročnosti príslušných optimalizačných problémov. Diskutujeme zároveň ich vplyv na voľbu vhodných heuristických algoritmov detekcie komunít, nakoľko exaktné prístupy majú v mnohých variantoch problému detekcii komunít tendenciu si vyžadovať exponenciálny výpočtový čas.

### 1 Úvod

Sociálna sieť môže byť formalizovaná ako neorientovaný graf  $G = [V, E]$ , t.j. diskrétna štruktúra opísaná množinou vrcholov  $V$  siete (ľudí) a množinou hrán  $E$ , t.j. dvojíc vrcholov siete, ktoré sú prepojené. Hovoríme, že dva vrcholy sú susedné, keď majú medzi sebou hranu. Stupeň vrchola  $v$  je definovaný ako počet susedov vrchola  $v$ , značíme ho  $\deg(v)$ . Inými slovami, je to počet kontaktov človeka v sociálnej sieti. Hustota  $d(G)$  grafu  $G$  je definovaná ako pomer počtu hrán k počtu všetkých dvojíc vrcholov, t.j.  $d(G) = 2|E|/[|V|(|V| - 1)]$  [17].

Komunita v sociálnej sieti môže byť v tomto kontexte definovaná ako podgraf  $G(V') = [V', E']$  indukovaný podmnožinou vrcholov siete  $V' \subset V$ , t.j. obsahujúci

len vrcholy z množiny  $V'$  a hrany medzi nimi, t.j.  $E'$  obsahuje všetky hrany z  $E$  medzi vrcholmi z  $V'$  a žiadne iné. Otázkou však je, aké ohraničenia by mal spĺňať tento podgraf. Na túto otázku ani súčasný výskum v informatických vedách nenachádza jasnú odpoveď. Existuje široké spektrum možných formulácií, aké kritériá by mala komunita spĺňať, avšak je ťažké exaktne vyhodnotiť, ktorá z nich je lepšia, resp. horšia [23]. Dodávame, že pojem komunity môže byť zovšeobecnený aj pre siete z iných domén, kedy sa preferuje pojem zhuk, resp. klaster (z angl. cluster) [27]. Vtedy hovoríme o probléme zhukovania, ktorý možno nájsť vo veľkom množstve aplikácií a je možné ho riešiť širokou škálou najvzájom pomerne odlišných prístupov [6, 16, 34].

Medzi ďalšie typy sietí, v ktorých je zhukovanie prakticky zaujímavým problémom, patria napr. siete vedeckých citácií, počítačové siete [29] či génové regulačné siete [2] a siete interakcií proteínov [10], študované v bioinformatike. Podobne ako sociálne siete, aj tieto siete majú netriviálnu „asymetrickú“ štruktúru. Takéto siete nazývame vo všeobecnosti komplexnými sieťami [1, 11, 21, 22].

S masívnou popularizáciou webových sociálnych sietí prirodzene stúpol aj záujem širšej verejnosti o túto problematiku. Najmä v oblasti štúdií o webe sa otázky hľadania komunít a zhukovania vo všeobecnosti stávajú čoraz atraktívnejšími [19, 30]. Pravdou však je, že vo viacerých vedeckých doménach majú sociálne siete svoje miesto už dlho, napr. vo výskume štruktúry teroristických organizácií [23] či vo výskume vzorov šírenia pohlavne prenosných ochorení, kde citlivou otázkou je najmä vplyv štruktúry siete sexuálnych partnerov na šírenie ochorenia [24].

V tomto článku diskutujeme otázku formulácie problému detekcie komunít a jej dopad na zložitosť príslušného optimalizačného problému a heuristických algoritmov na jeho riešenie. V kapitole 2 rozoberáme pojem komunity z rôznych perspektív. Kapitola 3 nadväzuje rozborom otázky formulácie hľadania komunít v sociálnych

sieťach. V tejto kapitole zároveň spomíname aj príslušné problémy matematickej optimalizácie. V kapitole 4 sa venujeme otázke úlohy heuristických algoritmov v hľadaní komunit. V kapitole 5 následne predkladáme niektoré zaujímavé otázky pre budúci výskum a v kapitole 6 sumarizujeme naše zistenia.

## 2 Pojem komunity v rôznych perspektívach

Pojem komunity možno vnímať z pohľadu viacerých disciplín. Pre účely tohto príspevku predstavujeme najprv veľmi krátko niekoľko pohľadov na komunity v sociálnych vedách. Podrobnejšie sa následne venujeme problematike ukotvenia pojmu komunity v informatických vedách. Toto ukotvenie je kľúčové najmä pre stanovenie vlastností problémov matematickej optimalizácie, na ktoré je následne detekcia komunit redukovaná.

V sociálnych vedách je pojem komunity študovaný relatívne široko. Jedným z relevantných pohľadov je, že existujú tri základné kritériá, na základe ktorých možno stanoviť, čo pojem komunita vyjadruje: miesta, záujmy a spoločensvá [28]. Kritérium miesta predstavuje pohľad, v ktorom komunita vzniká na základe geografických faktorov [33]. Pri primárnom zohľadnení kritéria záujmov sa dívame skôr na spoločné charakteristiky osôb, než spoločnú lokáciu. V rámci tohto kritéria možno skúmať napr. komunity vznikajúce na základe vierovyznania, sexuálnej orientácie či etnickej príslušnosti [18]. Z charakteru komunit opísaných týmto kritériom je zrejme, že mnohé sú elektívne, t.j. človek si volí komunity, ktorých súčasťou bude. Iné však zrejme túto vlastnosť spĺňať nemusia. Kritérium spoločensiev je z uvedených zrejme najmenej intuitívne, predstavuje viazanosť človeka na určitú ideu, resp. zážitok [7].

Vo všeobecnosti môžeme skonštatovať, že prístup sociálnych vied ku skúmaniu komunit sa snaží najmä hľadať dôvody, prečo sa ľudia združujú. Na základe toho je možné vysloviť hypotézy o možnom ukotvení významu pojmu komunity a študovať vlastnosti komunit. Komunity však možno študovať aj opačne, kedy sa dívame na ich samotnú štruktúru a vlastnosti a na základe dodatočných informácií môžeme vysloviť hypotézu o ich pôvode. Tento prístup je už viac vlastný informatike a vyskytuje sa najmä v oblasti dolovania v dátach.

Pristúpme teraz k rôznym možnostiam ukotvenia významu pojmu komunity v informatickom prístupe. Najjednoduchším prípadom je tzv. kompletný podgraf, niekedy tiež nazývaný klika (pre podobnosť s anglickým ekvivalentom clique sa výraz používa aj v slovenčine, napriek tomu, že sa jedná o bohemizmus).

**Kompletný podgraf (klika).** Kompletný podgraf sociálnej siete je zoskupenie ľudí, v ktorom sa každý s každým pozná. Formálnejšie, pre graf  $G = [V, E]$  a množinu  $V' \subset V$  platí, že indukovaný podgraf  $G' = [V', E']$  nazývame kompletným podgrafom (klikou), keď pre jeho hustotu platí, že  $d(G') = 1$ .

**Podgraf s hustotou väčšou ako konštanta  $d_l$ .** Takýto podgraf je definovaný pomocou takej množiny  $V' \subset V$ , že ňou indukovaný podgraf  $G' = [V', E']$  má hustotu aspoň  $d(G')$ . Formálnejšie,  $d(G') \geq d_l$ , kde  $d_l \in [0, 1]$ . Tento pojem je evidentne zovšeobecnením predošlého pojmu, nakoľko pre  $d_l = 1$  dostávame jednoducho kompletný podgraf. Preto aj problémy zamerané na detekciu podgrafov s určitou hustotou sú len zovšeobecnením problémov zameraných na kliky.

**Podgraf s hustotou uzavretou v intervale  $[d_l, d_u]$ .** Definujme komplementárny graf  $\bar{G}$  tak, že  $\bar{G}$  obsahuje hrany medzi práve tými dvojicami vrcholov, v ktorých nie sú hrany v  $G$  a naopak. Potom, podgraf grafu  $G$  s hustotou najviac  $d_u$  je ekvivalentný s podgrafom s hustotou aspoň  $1 - d_u$  v grafe  $\bar{G}$ . Z tohto vyplýva, že tento pojem je definovaný opäť len prostredníctvom predošlého pojmu.

Všetky uvedené formulácie pojmu komunity sa spoliehajú na ohraničenie lokálnej hustoty. Vidíme pritom, že už takéto jednoduché videnie pojmu vytvára nejasnosť pri ukotvení jeho významu. Keď vezmeme do úvahy tretiu, t.j. najvšeobecnejšiu formuláciu, tak zásadnou otázkou je, ako nastaviť hodnoty  $d_l$  a  $d_u$ . Takisto, už jednoduchá dvojica vrcholov spojených hranou vytvára kompletný podgraf s hustotou 1. Preto je dôležité tiež hľadať na veľkosť príslušného podgrafu, pri ktorom je opäť otázne, aký podgraf označíme za dostatočne veľký.

## 3 Hľadanie komunit v sociálnych sieťach

Na základe predošlej analýzy môžeme skonštatovať, že hľadanie komunit je skôr množinou navzájom súvisiacich problémov, než exaktne definovaným konceptom. Z tohto vyplývajú aj niektoré ďalšie fakty, ktoré ovplyvňujú rozhodnutia pri návrhu algoritmov a aplikácií detekcie komunit v informatických vedách. Problémy hľadania komunit v sociálnych sieťach môžeme z teoretického hľadiska rozdeliť na dva základné typy: problém najväčšej komunity a problém minimálnej dekompozície siete do komunit. V nasledujúcich riadkoch tieto problémy formalizujeme.

**Problém najväčšej komunity.** Nech  $G = [V, E]$  je neorientovaný graf. Tento problém formálne definujeme pro-



stredníctvom nasledujúceho predpisu:

$$\max k = |V'|; V' \subset V, \quad (1)$$

za predpokladu, že sú splnené ohraničenia:

$$\text{val} \left( \bigwedge_{i=1}^m C_i(G(V')) \right) = 1, \quad (2)$$

kde  $C_i(G(V'))$  sú predikáty, ktoré vyjadrujú, či indukovaný podgraf  $G(V_i)$  spĺňa určitú vlastnosť. Potom, uvedená konjunkcia je pravdivá vtedy a len vtedy, keď indukovaný podgraf spĺňa všetky vlastnosti definované predikátmi  $C_i$ .

Jedným zo špeciálnych prípadov uvedeného problému je problém najväčšieho kompletného podgrafu, ktorý získame pre  $m = 1$  a predikát  $C_1(G(V'))$ , ktorý hovorí, že  $d(G(V')) = 1$ . Definícia nám však umožňuje zájsť aj ďalej a definovať napr. ohraničenie  $C_2$ , ktoré by určovalo, že fixný vrchol  $v \in V'$ . Potom by sme získali problém hľadania najväčšieho kompletného podgrafu, v ktorom sa nachádza určitý vrchol siete. Takáto možnosť je zaujímavá pre aplikácie, v ktorých potrebujeme nájsť čo najväčšiu komunitu, do ktorej patrí konkrétna osoba.

Všimnime si zároveň, že charakter problému je tu určený len predikátmi  $C_i$ , samotná minimalizácia je pre všetky „rozumné“ formulácie spoločná. Uvedená formulácia má tiež výhodou v tom, že problém môžeme stanoviť ako multikriteriálny, t.j. zaviesť viacero ohraňení, pomocou ktorých určíme, či nájdený podgraf môžeme nazvať komunitou, resp. nemôžeme. Samotné určenie vhodnosti príslušných kritérií sa však stále v značnej miere spolieha na subjektívny názor človeka.

**Problém minimálnej dekompozície siete do komunít.** Nech  $G = [V, E]$  je neorientovaný graf. Tento problém formálne definujeme prostredníctvom nasledujúceho predpisu:

$$\min k \leq |V|, \quad (3)$$

za predpokladu, že existujú také po dvojiciach disjunktné podmnožiny vrcholov  $V_1, V_2, \dots, V_k \subset V$  také, že:

- $\bigcup_{i=1}^k V_i = V$  a zároveň
- $\forall i = 1..k \text{ val} \left( \bigwedge_{i=1}^m C_i(G(V_i)) \right) = 1,$

kde  $C_i(G(V'))$  sú opäť rovnaké predikáty ako v predošlom probléme.

V probléme minimálnej dekompozície sa už nezameriavame na hľadanie jednej konkrétnej komunity, ale problém formulujeme ako rozklad vrcholov celej siete, resp. vzorky siete do skupín, ktoré spĺňajú analogické ohraňenia. Opäť máme výhodnú situáciu, v ktorej práve ohraňenia  $C_i$  určujú charakteristiky príslušného problému. Typickým príkladom takéhoto problému je pokrytie vrcholov kompletnými podgrafmi (v tomto prípade toto pokrytie možno vnímať ako dekompozíciu do kompletných podgrafov) [3, 4, 15]. Tento problém z hore uvedenej všeobecnej formulácie získame opäť tak, že položíme  $m = 1$  a predikát  $C_1(G(V_i))$  bude stanovený tak, že  $d(G(V_i)) = 1$ .

Mnohé problémy, ktoré je možné získať upresnením uvedených všeobecných formulácií, sú však výpočtovo veľmi náročné [31]. Problém najväčšieho kompletného podgrafu i problém pokrytia kompletnými podgrafmi patria medzi prvé problémy, o ktorých bolo dokázané, že sú NP-ťažké [15]. V súčasnosti sú preto známe len také exaktné algoritmy pre riešenie týchto problémov, ktoré ich dokážu pre ľubovoľný vstup vyriešiť v exponenciálnom čase. Dokonca aj keď vieme, že existuje kompletný podgraf siete s  $k$  vrcholmi, tak určiť, či existuje aj kompletný podgraf s  $k+1$  vrcholmi je tiež výpočtovo veľmi náročný problém. Podobná je situácia aj v prípade pokrytia kompletnými podgrafmi.

Dodávame, že nielen problém najväčšieho kompletného podgrafu a minimálneho pokrytia kompletnými podgrafmi sú náročné. Bolo ukázané, že aj hľadanie podgrafov s určitou minimálnou lokálnou hustotou je NP-ťažké [9]. Existujú pritom aj ďalšie metriky, ktoré sa dajú použiť na určenie, či podgraf môžeme nazvať komunitou, resp. zhľukom. Tieto metriky sú často založené na rezoch grafov a patria medzi ne relatívna hustota, či tzv. vodivosť (angl. conductance). O problémoch súvisiacich s optimalizáciou týchto metrick je tiež dokázané, že sú NP-ťažké [9].

## 4 Rola heuristických algoritmov v hľadaní komunít

Vzhľadom k výpočtovej náročnosti problémov hľadania komunít je veľmi dôležité sa zamerať na výber vhodných algoritmov. Len táto otázka je sama o sebe relatívne náročná. Exaktné algoritmy sú síce schopné vyriešiť ľubovoľný optimalizačný problém hľadania komunít optimálne, majú však zásadne exponenciálnu zložitosť, nakoľko aj priestor možných riešení je exponenciálne veľký. Preto sú použiteľné len na malé inštancie, rádovo sa jedná o siete s desiatkami vrcholov. Pre reálne použitie vo veľkých sieťach je preto na mieste uvažovať o využití prístupu umelej inteligencie [25], t.j. heuristických, resp. metaheuristických algoritmov, ktoré poskytujú kvalitné subop-

timálne riešenie problému [12]. Otáznou je tiež reprezentácia problému, nakoľko existuje viacero možností reprezentácie pre problémy hľadania najväčšej komunity i problémy minimálnej dekompozície do komunit.

**Reprezentácie problémov hľadania komunit.** V probléme hľadania najväčšej komunity je možné zvoliť viacero reprezentácií. Priamočiarou reprezentáciou je zrejme jednoduchý zoznam indexov vrcholov, ktoré zaraďujeme do komunity, t.j. priamo množina  $V' \subset V$ . Takýto zoznam je zároveň ekvivalentný s reprezentáciou vo forme bitového vektora z priestoru  $\{0,1\}^{|V|}$ , kde pre každý vrchol máme samostatný bit, ktorý vyjadruje, či vrchol patrí do komunity, resp. nepatrí. Takáto reprezentácia by bola vhodná pre klasické evolučné algoritmy, napr. pôvodný genetický algoritmus [13]. Jej úskalím je však to, že v takomto bitovom vektore nemáme garantované, že súvisiace bity (t.j. bity reprezentujúce susedné vrcholy) budú mať tendenciu byť vedľa seba. Preto by zrejme bolo klasické kríženie neúčinné a vhodné by bolo použiť algoritmus založený na mutácii.

Pre problém minimálnej dekompozície siete do komunit môžeme využiť reprezentáciu pomocou značenia vrcholov. Každý vrchol dostane jednu z  $k$  značiek, pričom vrcholy s rovnakými značkami budeme považovať za príslušné komunity. Nakoľko platí, že  $k \leq |V|$ , tak sa jedná o priestor  $\{1, 2, \dots, |V|\}^{|V|}$ . Priestor takýchto značení je možné prehľadávať rôznymi spôsobmi, pre veľké siete je však rozumné používať nepriamu reprezentáciu vo forme greedy (lačných) algoritmov, ktoré takéto značenie zostrojujú na základe vopred definovanej permutácie vrcholov. Medzi algoritmy tohto typu patria greedy pokrytie kompletnými podgrafmi [4] a greedy farbenie grafov [32].

**Heuristické algoritmy hľadania komunit.** Nie je jednoduché vybrať vhodný heuristický algoritmus pre hľadanie komunit. V probléme hľadania komunity pre konkrétny vrchol existujú metódy stochastického lokálneho hľadania, využívajú však rôzne funkcie ohodnotenia kvality [26].

V probléme pokrytia kompletnými podgrafmi je zaujímavé využitie iterovanej greedy heuristiky [4], ktorá využíva blokové vlastnosti problému. Momentálne identifikované kompletné podgrafy sú vložené do permutácie vrcholov ako bloky, následne sa poprehadzujú a greedy algoritmus je použitý znova. Uvedený prístup negarantuje kvalitu riešenia, ale ukázalo sa, že poskytuje veľmi dobré riešenia na reálnych sociálnych sieťach. Dodávame, že pôvodne sa táto idea už dávnejšie využívala vo farbení grafov [8]. Podobný prístup bol využitý aj pre multikriteriálny problém dekompozície do komunit [5].

## 5 Otvorené otázky vo výskume hľadania komunit

Predošlé úvahy v tejto časti sumarizujeme a hľadáme východiská a zaujímavé otvorené problémy pre ďalší výskum. Identifikujeme štyri zaujímavé oblasti, v ktorých zostávajú niektoré kľúčové otázky hľadania komunit a jeho využitia nezodpovedané, resp. iba čiastočne zodpovedané.

### Formulácie problému a zložitosť, ku ktorej vedú.

V predošlých úvahách sme spomenuli, že mnohé formulácie problému detekcie komunit vedú k výpočtovo náročným úlohám. Navyše, nie je priamočiaro jasné, ktoré formulácie sú vhodnejšie z pohľadu vierohodnosti výsledného riešenia, než iné formulácie. Je napr. zrejme, že chápanie komunity ako skupiny ľudí, v ktorej sa všetci poznajú (t.j. ako kompletného podgrafu), je relevantné pre mnohé aplikácie. Otvorenou však ostáva otázka vhodného zovšeobecnenia tohto prístupu. V tomto kontexte je možné vidieť viacero možných smerov výskumu. Z nich je zaujímavou multikriteriálna formulácia problémov detekcie komunit, ktorá sa javí ako vhodný prístup k hybridizácii rôznych pohľadov na pojem komunity. Druhou zaujímavou otázkou zostáva, či všeobecne náročné problémy, ktoré sme spomínali, nie sú v skutočnosti v reálnych sieťach výpočtovo jednoduchšie. Pre riešenie tejto otázky by bolo zaujímavé študovať napr. vlastnosti existujúcich heuristických algoritmov hľadania komunit v zjednodušených modeloch komplexných sietí. V tomto smere sa diskutované otázky začínajú dotýkať oblasti analýzy evolučných a iných stochastických algoritmov [20].

**Výber vhodných algoritmov.** Táto otázka je zrejme veľmi úzko spojená s predošlou. Existujú techniky, ako napr. multikriteriálne konštrukčné procedúry [5], ktoré sa javia ako všeobecne využiteľné pre rôzne varianty problému hľadania komunit. Vo všeobecnosti sú však metódy detekcie komunit natoľko ideovo diverzifikované [27], že všeobecné techniky (napr. metaheuristiky či evolučné algoritmy) majú evidentné limity z hľadiska pokrytia tých myšlienok, ktoré už pre detekciu komunit boli navrhnuté. V tomto smere by mohla pomôcť bližšia teoretická analýza problémov hľadania komunit, resp. využitie strojového učenia pre voľbu vhodného algoritmu detekcie komunit, resp. dokonca výber vhodnej formulácie problému detekcie komunit.

**Metriky kvality komunit a zhlukov.** Vyhodnotiť kvalitu zhlukov dát je netriviálny problém prakticky vo všetkých doménach, v ktorých sa zhlukovanie využíva, komunity nevytvárajúc. Vo výskume detekcie komunit, resp. zhlukovania v sieťach vo všeobecnosti je trendom algoritmy

testovať na zjednotených sadách dát (často sa využívajú extrakty webových sociálnych sietí, kolaboračné a citačné siete, podsiete Internetu a pod.). V takomto prípade je možné vyhodnotiť niektoré metriky kvality, napr. vodivosť či lokálnu hustotu [19, 31]. Často však poznáme aj správne zhlukovanie, a preto je v problémoch minimálnej dekompozície použiť metriky založené na podobnosti nájdených dekompozícií s referenčnou, napr. upravený Randov index (angl. adjusted Rand index) [14]. Problémom zostáva objektívne vyhodnotenie kvality v sieťach, v ktorých referenčné riešenie nepoznáme.

**Praktické aplikácie.** Aplikačných oblastí detekcie komunit a zhlukovania v sieťach vo všeobecnosti je veľké množstvo, prehľadový článok [27] tieto oblasti dobre zachytáva. Na tomto mieste spomeňme predovšetkým využitie detekcie komunit pri odporúčaní, napr. vhodných kandidátov na pracovné pozície (na základe zaradenia do komunity s iným vhodným kandidátom), či v predikcii virálneho šírenia informácie. Ďalšie aplikácie možno nájsť pri odporúčaní možných vedeckých kolaborácií alebo aj vo vhodne cielej kontextovej reklame. Mnohé z prístupov, ktoré výskum hľadania komunit v sociálnych sieťach zachytáva, sú však tiež aplikovateľné v bioinformatike [2, 10], spracovaní dát vo veľkých databázach [27] či v operačnom výskume [12].

## 6 Zhodnotenie

V príspevku sme diskutovali možnosti všeobecnej formulácie problému hľadania komunit v sociálnych sieťach, nakoľko doposiaľ v informatike nie je jasne stanovený pojem komunity. Krátko sme diskutovali o metrikách, ktoré možno využiť na určenie, či podgraf je komunitou, zamerali sme sa najmä na metriku lokálnej hustoty. Identifikovali sme dva základné typy týchto problémov: problém najväčšej komunity a problém minimálnej dekompozície do komunit. Mnohé z problémov, ktoré možno odvodiť z týchto dvoch základných formulácií, napr. problém maximálneho kompletneho podgrafu a problém minimálneho pokrytia kompletnými podgrafmi, sú NP-ťažké, a preto výpočtovo náročné problémy.

V otázke roly heuristických algoritmov v hľadaní komunit sme načrtli víziu všeobecného algoritmu, ktorý by bol vhodný na riešenie rôznych typov problémov detekcie komunit. Zdôraznili sme tiež rolu heuristických algoritmov v týchto problémoch, vzhľadom k tomu, že exaktné prístupy si vyžadujú exponenciálne množstvo času na riešenie mnohých z nich. Identifikovali sme štyri základné oblasti pre možné ďalšie smerovanie výskumu detekcie komunit, vo všeobecnom pohľade, ktorý opisuje tento príspevok.

**PodĎakovanie.** Tento príspevok bol podporený Vedeckou grantovou agentúrou Slovenskej republiky, v rámci grantov 1/0553/12 a 1/0458/13.

## Literatúra

- [1] Albert, R., Barabási, A.L.: Statistical mechanics of complex networks. *Reviews of Modern Physics*, 2002, vol. 74, no. 1, pp. 47–97.
- [2] Boyer, F., Morgat, A., Labarre, L., Pothier, J., Viari, A.: Syntons, metabolons and interactons: an exact graph-theoretical approach for exploring neighbourhood between genomic and functional data. *Bioinformatics*, 2005, vol. 21, no. 23, pp. 4209–4215.
- [3] Chalupa, D.: On the Ability of Graph Coloring Heuristics to Find Substructures in Social Networks. *Information Sciences and Technologies, Bulletin of ACM Slovakia*, 2011, vol. 3, no. 2, pp. 51–54.
- [4] Chalupa, D.: On the efficiency of an order-based representation in the clique covering problem. In: *Proceedings of the 14th annual conference on Genetic and evolutionary computation*. GECCO '12, New York, NY, USA, ACM, to appear, 2012.
- [5] Chalupa, D., Pospíchal, J.: Metaheuristically Optimized Multicriteria Clustering for Medium-Scale Networks. In: *Soft Computing Models in Industrial and Environmental Applications, 7th International Conference, SOCO'12*. Volume 188 of *Advances in Intelligent Systems and Computing*., Springer, 2013, pp. 337–346.
- [6] Comellas, F., Miralles, A.: A fast and efficient algorithm to identify clusters in networks. *Applied Mathematics and Computation*, 2010, vol. 217, no. 5, pp. 2007–2014.
- [7] Crow, G., Allan, G.: *Social Networks, Informal Care and Public Policy*. Hemel Hempstead: Harvester Wheatsheaf, 1994.
- [8] Culberson, J.C., Luo, F.: Exploring the k-colorable Landscape with Iterated Greedy. In: *Dimacs Series in Discrete Mathematics and Theoretical Computer Science*, American Mathematical Society, 1995, pp. 245–284.
- [9] Drineas, P., Frieze, A., Kannan, R., Vempala, S., Vinnay, V.: Clustering in large graphs and matrices. *Machine Learning*, 2004, vol. 56, pp. 9–33.
- [10] Gao, L., Sun, P., Song, J.: Clustering Algorithms for Detecting Functional Modules in protein Interaction Networks. *Journal of Bioinformatics and Computational Biology*, 2009, vol. 7, no. 1, pp. 217–242.
- [11] Girvan, M., Newman, M.E.J.: Community structure in social and biological networks. *Proceedings of the*

- National Academy of Sciences*, 2002, vol. 99, no. 12, pp. 7821–7826.
- [12] Glover, F., Kochenberger, G.A.: *Handbook of metaheuristics. International Series in Operations Research & Management Science*. Springer, 2003.
- [13] Holland, J.H.: Genetic algorithms and classifier systems: foundations and future directions. In Grefenstette, J.J., ed.: *Proceedings of the Second International Conference on Genetic algorithms and their application*, Hillsdale, NJ, USA, L. Erlbaum Associates Inc., 1987, pp. 82–89.
- [14] Hubert, L., Arabie, P.: Comparing partitions. *Journal of Classification*, 1985, vol. 2, no. 1, pp. 193–218.
- [15] Karp, R.M.: Reducibility among combinatorial problems. In Miller, R., Thatcher, J., eds.: *Complexity of Computer Computations*. Plenum Press, New York, NY, USA, 1972, pp. 85–103.
- [16] Kaufman, L., Rousseeuw, P.J.: *Finding groups in data: an introduction to cluster analysis*. Wiley, 1990.
- [17] Kvasnička, V., Pospíchal, J.: *Algebra and Discrete Mathematics (Algebra a diskrétna matematika, in Slovak)*. SUT Press, 2008.
- [18] Lee, D., Newby, H.: *Community Life. An introduction to local social relations*. Unwin Hyman, London, 1983.
- [19] Leskovec, J., Lang, K.J., Mahoney, M.W.: Empirical comparison of algorithms for network community detection. In Rappa, M., Jones, P., Freire, J., Chakrabarti, S., eds.: *Proceedings of the 19th International Conference on World Wide Web, WWW 2010*, ACM, 2010, pp. 631–640.
- [20] Neumann, F., Witt, C.: *Bioinspired Computation in Combinatorial Optimization - Algorithms and Their Computational Complexity*. Springer, 2010.
- [21] Newman, M.E.J.: Detecting community structure in networks. *The European Physical Journal B*, 2004, vol. 38, no. 2, pp. 321–330.
- [22] Newman, M.E.J., Girvan, M.: Mixing patterns and community structure in networks. In: *Statistical Mechanics of Complex Networks: Proceedings of the XVIII Sitges Conference on Statistical Mechanics, Lecture Notes in Physics 625*, Berlin, Heidelberg, Germany, Springer, 2003.
- [23] Pattillo, J., Youssef, N., Butenko, S.: Clique Relaxation Models in Social Network Analysis. In Thai, M.T., Pardalos, P.M., eds.: *Handbook of Optimization in Complex Networks*. Springer, 2012, pp. 143–162.
- [24] Rothenberg, R.B., Potterat, J.J., Woodhouse, D.E.: Personal Risk Taking and the Spread of Disease: Beyond Core Groups. *The Journal of Infectious Diseases, Supplement 2*, 1996, vol. 174, pp. S144–S149.
- [25] Russell, S.J., Norvig, P.: *Artificial Intelligence: A Modern Approach (2nd ed.)*. Upper Saddle River, New Jersey: Prentice Hall, 2003.
- [26] Schaeffer, S.E.: Stochastic Local Clustering for Massive Graphs. In Ho, T.B., Cheung, D., Liu, H., eds.: *Advances in Knowledge Discovery and Data Mining*, Springer, 2005, pp. 354–360.
- [27] Schaeffer, S.E.: Graph clustering. *Computer Science Review*, 2007, vol. 1, no. 1, pp. 27–64.
- [28] Smith, M.K.: Community. Published in the encyclopedia of informal education <http://www.infed.org/community/community.htm> [retrieved 08 April 2013], 2001.
- [29] Sun, J., Xie, Y., Zhang, H., Faloutsos, C.: Less is More: Sparse Graph Mining with Compact Matrix Decomposition. *Statistical Analysis and Data Mining*, 2008, vol. 1, no. 1, pp. 6–22.
- [30] Tang, J., Wang, T., Wang, J., Lu, Q., Li, W.: Using complex network features for fast clustering in the web. In Sadagopan, S., Ramamritham, K., Kumar, A., Ravindra, M.P., Bertino, E., Kumar, R., eds.: *Proceedings of the 20th international conference companion on World wide web. WWW '11*, New York, NY, USA, ACM, 2011, pp. 133–134.
- [31] Šíma, J., Schaeffer, S.E.: On the NP-completeness of some graph cluster measures. In Wiedermann, J., Tel, G., Pokorný, J., Bielíková, M., Štuller, J., eds.: *Proceedings of the 32nd International Conference on Current Trends in Theory and Practice of Computer Science*, Berlin, Heidelberg, Springer, 2006, pp. 530–537.
- [32] Welsh, D.J.A., Powell, M.B.: An upper bound for the chromatic number of a graph and its application to timetabling problems. *The Computer Journal*, 1967, vol. 10, no. 1, pp. 85–86.
- [33] Willmott, P.: *Social Networks, Informal Care and Public Policy*. Policy Studies Institute, London, 1986.
- [34] Xu, R., Wunsch, D.: Survey of Clustering Algorithms. *IEEE Transactions on Neural Networks*, 2005, vol. 16, no. 3, pp. 645–678.

## P kolonie ve 2D

Luděk Cienciala, Lucie Ciencialová

Ústav informatiky, Výzkumný ústav Centra excelence IT4Innovations, Slezská univerzita v Opavě  
Bezručovo nám.13, 746 01 Opava

Email: ludek.cienciala@fpf.slu.cz, lucie.ciencialova@fpf.slu.cz

### Abstrakt

V našem příspěvku prezentujeme výsledky získané při výzkumu P kolonií, které představují jeden z řady biologicky inspirovaných výpočetních modelů patřících do oblasti membránových systémů, které byly poprvé představeny Gheorgem Păunem v roce 1998. P kolonie jsou inspirovány strukturou a činností živých organismů žijících ve společném prostředí. Uvedený model je založený na nezávislých membránových agentech umístěných do prostředí. Naším předmětem zkoumání budou P kolonie, kde daný model rozšíříme o tzv. 2-rozměrné prostředí. V původním modelu pracujeme s „bezrozměrným“ prostředím. Agenty ve 2D P koloniích jsou omezeny, co se týče získávání informací o obsahu prostředí na „blízké“ okolí a mohou se pohybovat v prostředí ve čtyřech směrech. V práci definujeme 2D P kolonie, výpočet a na příkladech (Běžec, Conwayova hra života, Stigmergie) ukážeme princip, jak probíhá výpočet ve 2D P kolonii. Pro větší názornost jsme použili program – simulátor s grafickým uživatelským rozhraním.

## 1 Úvod

Sledování přirozených systémů a procesů nám dává možnost nového úhlu pohledu na výpočty. Můžeme vytvořit nekonvenční výpočetní modely založené na nových ideách. Mezi takto motivované výpočetní modely patří i modely inspirované biochemií buněčné struktury, nazývané membránové systémy nebo také P systémy. Tyto modely byly zavedeny v roce 1998 Gheorghem Păunem v [10]. Existují různé druhy P systémů v závislosti na způsobu reprezentace objektů, na základě typů pravidel, způsobu formalizace vývoje systému, popisu výpočtu a výsledky výpočtu. Membránové systémy patří do skupiny nedeterministických systémů, kde výpočet probíhá na bázi stanovených pravidel v diskretních krocích a paralelně pro všechny objekty.

P kolonie zavedené v článku [8] představují jeden z modelů P systémů. Jsou inspirované strukturou a činností živých organismů, žijících ve společném prostředí. Role živých organismů budou hrát membránové

agenty, které umístíme do prostředí. Prostředí v daném modelu P koloniích je statické a slouží pouze jako komunikační kanál pro organismy (agenty) v P kolonii. V práci Eco-P kolonie [2, 1] jsme rozšířili daný model P kolonií o možnost měnit prostředí nezávisle na daných organismech. Výpočetní model PCol automaty popsány v pracích [4], [3] vychází také z P kolonií, kde původní model jsme rozšířili o vstupní pásku, kde chování agentů bylo ovlivněno nejen prostředím, ale také symboly umístěnými na pásce.

Více informací o P systémech se můžete dozvědět v [10], [12] a [13].

V tomto článku se budeme zabývat P koloniemi, kde daný model rozšíříme o tzv. 2-rozměrné prostředí a budeme je označovat jako P kolonie ve 2D nebo také 2D P kolonie. Daná modifikace byla poprvé popsána v článku [5].

V P kolonii živé organismy představují agenty, které jsou umístěny do prostředí. Každý agent je tvořen jednou membránou obsahující dvojici objektů. Počet objektů uvnitř každého agenta je stejný a nemění se během celého výpočtu. Prostředí slouží jako komunikační kanál a agenty jej využívají jako úložiště objektů. V každém okamžiku agenty mají informaci o existenci všech objektů v prostředí a mají ke každému objektu bezprostřední přístup.

Každý agent má svou množinu programů určující činnost agenta. Programy jsou jednoduché a mají vliv na množinu objektů, která přísluší danému agentu, ale také i na multimnožinu objektů obsaženou v prostředí. To znamená, že každý agent může měnit obsah prostředí pomocí programů a tím ovlivnit chování ostatních agentů prostřednictvím prostředí. Tato komunikace agentů přes prostředí tvoří základní princip fungování P kolonie. V každém okamžiku každý objekt obsažený v agentu je ovlivněn právě aplikovaným programem.

V reálném světě (stejně tak i v kybernetickém světě) se liší koncentrace látek v závislosti na místě a živé organismy nemají možnost vědět, co je „za horizontem“. Tyto úvahy nás inspirovaly k zavedení nového modelu P kolonií, kde agenty jsou umístěny ve 2D kartézské mřížce. Agenty tedy umístíme do této mřížky a „pohled“ agenta omezíme na buňky bezprostředně ho obklopující. Na základě obsahu

těchto buněk je ovlivněna poloha agenta v mřížce v dalším kroku výpočtu.

Co se týče pravidel, budeme vycházet z původního modelu P kolonií. Program agenta může být tedy sestaven ze dvou typů pravidel. První typ pravidel nazýváme přepisovací - evoluční a druhý typ pravidel nazýváme komunikační. Komunikační pravidlo bude mít navíc vliv pouze na místo, kde se právě agent nachází.

Přidáme nový typ pravidel pro pohyb agenta ve 2D prostředí. Podmínkou pro pohyb agenta je nalezení konkrétních objektů v konkrétních místech prostředí. To je určeno pomocí matice základních objektů. Agent hledá nejvýše jeden objekt v každé buňce okolí. Pokud daný objekt najde, přemístí se o jednu buňku nahoru, dolů, dolva nebo doprava.

Programy jsou tvořeny pravidly. Protože každý agent obsahuje dva objekty, programy se skládají ze dvou pravidel. Program navíc obsahuje nejvýše jedno pravidlo pro pohyb agenta. Pro jednoduchost chování agenta stanovíme ještě jednu podmínku a to podmínku: pokud agent mění polohu, pak nemůže komunikovat s prostředím. Tzn., pokud program obsahuje pravidlo pro pohyb, tak druhé pravidlo které obsahuje musí být pravidlo přepisovací.

## 2 Definice

V této kapitole uvedeme základní definice a označení použité v článku. Předpokládáme, že čtenáři jsou známy základní definice a poznatky z oblasti teorie formálních jazyků a automatů.

$NRE$  budeme označovat množiny rekurzivně spočítatelných nezáporných čísel.  $\Sigma$  je označení pro abecedu. Nechť  $\Sigma^*$  je množina všech slov nad abecedou  $\Sigma$  (včetně prázdného slova  $\varepsilon$ ). Pro délku slova  $w \in \Sigma^*$  budeme používat označení  $|w|$  a pro počet výskytů symbolů  $a \in \Sigma$  ve slově  $w$  pak  $|w|_a$ .

Multimnožina objektů  $M$  je dvojice  $M = (V, f)$ , kde  $V$  je libovolná (ne nutně konečná) množina objektů a  $f$  je zobrazení  $f : V \rightarrow N$ . Zobrazení  $f$  přiřazuje každému objektu z množiny  $V$  jeho násobnost v multimnožině  $M$ .  $V^\circ$  označíme množinu všech konečných multimnožin nad množinou objektů  $V$ . Množinu  $V' \subseteq V$  nazveme support množiny  $M$  a budeme ji označovat  $supp(M)$ , pokud pro všechna  $x \in V'$   $f(x) \neq 0$ . Kardinalitu množiny  $M$  značíme  $|M|$  a je definována jako  $|M| = \sum_{a \in V} f(a)$ . Každá multimnožina objektů  $M$  s množinou objektů  $V' = \{a_1, \dots, a_n\}$  může být reprezentována jako řetězec  $w$  nad abecedou  $V'$ , kde  $|w|_{a_i} = f(a_i)$ ;  $1 \leq i \leq n$ . Samozřejmě všechna slova  $w$  permutací znaků jsou reprezentována stejnou multimnožinou  $M$ .  $\varepsilon$  představuje prázdnou multimnožinu.

### 2.1 P kolonie ve 2D

Stručně popíšeme 2D P kolonie. P kolonie jsou tvořeny agenty a prostředím. Agenty i prostředí obsahují objekty. Prostředí má rozměr  $m \times n$ ,  $m$  sloupců a  $n$  řádků buněk,  $m, n \in N$ . Každý agent má svou množinu programů.

První typ pravidla nazýváme přepisovací (evoluční) pravidlo a má tvar  $a \rightarrow b$ . Vykonáním pravidla objekt  $a$  v agentu je přepsán na objekt  $b$ . Druhý typ pravidla nazýváme komunikační a má tvar  $c \leftrightarrow d$ . Pokud je aplikováno komunikační pravidlo, potom objekt  $c$  uvnitř agenta a objekt  $d$  vně agenta si zamění pozice, tzn., že po vykonání pravidla se objekt  $d$  nachází uvnitř agenta a objekt  $c$  je obsažen v prostředí.

Třetí typ pravidla nazýváme pohybové pravidlo a má tvar matice  $3 \times 3 \rightarrow$  směr pohybu. Umístění agenta odpovídá středu matice Jestliže sousední buňky obsahují stejné objekty, jako jsou objekty uvnitř matice, potom se agent může pohybovat do některé ze sousedních buněk následujícím směrem pohybu - vlevo, vpravo, nahoru nebo dolů. Pokud je tedy pozice agenta  $[r, s]$ ,  $0 \leq r \leq m - 1$ ,  $0 \leq s \leq n - 1$ , po vykonání pravidla může agent změnit pozici na  $[r - 1, s]$ ,  $[r + 1, s]$ ,  $[r, s - 1]$  or  $[r, s + 1]$ .

Program může obsahovat pouze jedno pohybové pravidlo. Pokud program se skládá z pohybového pravidla, tak daný program už nemůže obsahovat komunikační pravidlo.

**Definice 1** 2D P kolonie je struktura

$$\Pi = (A, e, Env, B_1, \dots, B_k, f), k \geq 1, \text{ kde}$$

- $A$  je abeceda P kolonie, její prvky nazýváme objekty;
- $e$  je základní objekt automatu, který nazýváme environmentální,  $e \in A$ ;
- $Env$  je dvojice  $(m \times n, w_E)$ , kde  $m \times n$ ,  $m, n \in N$  je matice o velikosti  $m \times n$  multimnožin objektů nad množinou  $A - \{e\}$ ,  $m, n \in N$  je velikost prostředí a  $w_E$  je počáteční obsah prostředí.
- $B_i$ ,  $1 \leq i \leq k$  jsou agenty, každý agent je struktura  $B_i = (O_i, P_i, [r_i, s_i])$ , kde
  - $[r_i, s_i]$  je počáteční pozice agenta  $B_i$  ve 2D prostředí,  $0 \leq r_i \leq m - 1$ ,  $0 \leq s_i \leq n - 1$ ,  $1 \leq i \leq k$ ,
  - $O_i$  je multimnožina nad množinou  $A$ , určuje počáteční stav (obsah) agenta,  $|O_i| = 2$ ,
  - $P_i = \{p_{i,1}, \dots, p_{i,j}\}$ ,  $j \geq 1$ ,  $1 \leq i \leq k$  je konečná množina programů, kde každý program je sestaven ze dvou 2 pravidel, které mají jeden z následujících tvarů:

- \*  $a \rightarrow b$  pravidla v daném tvaru nazýváme evoluční pravidla,
- \*  $c \leftrightarrow d$  pravidla v daném tvaru nazýváme komunikační pravidla,
- \*  $[a_{u,v}] \rightarrow t, 0 \leq u, v \leq 2,$   
 $t \in \{\Leftarrow, \Rightarrow, \Uparrow, \Downarrow\}$  pravidla v daném tvaru nazýváme pohybová pravidla;

- $f \in A$  je finální objekt kolonie.

Výpočetní krok se skládá ze tří částí. První část spočívá ve stanovení aplikovatelné množiny programů v závislosti na aktuální konfiguraci P kolonie. Existují programy patřící do množin programů všech agentů. Ve druhé části je vybrán jeden program u každého agenta patřící do množiny aplikovatelných programů. Neexistuje kolize mezi komunikačními pravidly různých programů. Ve třetí části jsou vykonány vybrané programy.

Vykonáním programů se změní konfigurace kolonie. Změna konfigurace je založena na změně stavu prostředí, obsahu a umístění agentů.

Výpočet probíhá nedeterministicky a maximálně paralelně a končí zastavením, kdy žádný agent nemá aplikovatelný program.

Výsledek výpočtu je počet kopií finálního objektu umístěného v prostředí na konci výpočtu.

Další možný způsob, jak definovat výsledek výpočtu je vzít v úvahu nejen počet finálních objektů, ale také i jejich umístění. Výsledkem by pak mohl šedotónový obraz, řetězec znaků nebo číslo závislé na počtu a umístění objektů (například  $g = \sum_{j=0}^{n-1} \left( \sum_{i=0}^{m-1} f(i, j) \right) \cdot n^i$ , kde  $f(i, j)$  je počet kopií objektu  $f$  v  $[i, j]$ -buňce).

Motivací pro zavedení 2D P kolonií není studium jejich výpočetní síly, ale sledování jejich chování v průběhu výpočtu. Můžeme definovat několik pohledů jak posoudit dynamiku výpočtu:

- počet pohybů agentů,
- počet „navštívených“ buněk (nebo „nenavštívených“ buněk),
- počet kopií určitého objektu v „domovské“ buňce nebo v celém prostředí.

Navíc můžeme uvažovat, že daný počet sledujeme po jednotlivých krocích výpočtu nebo až na konci výpočtu.

### 3 Příklady

V této kapitole ukážeme příklady 2D P kolonií. První 2D P kolonii můžeme nazvat běžec po objektech  $b$ .

**Příklad 1** Necht'  $\Pi_1$  2D P kolonie je definována:  $\Pi_1 = (A, e, Env, B_1, f)$ , kde

- $A = \{e, f, a, b\}$ ,
- $e \in A$  je základní objekt v prostředí kolonie,
- $Env = (5 \times 5, w_E)$ ,

$$w_E = \begin{bmatrix} a & a & a & a & a \\ a & b & b & b & a \\ a & b & a & b & a \\ a & b & b & b & a \\ a & a & a & a & a \end{bmatrix},$$

- $B_1 = (aa, P_1, [1, 1])$ ,

$$P_1 = \left\{ \left\langle \begin{bmatrix} * & b & * \\ * & b & * \\ * & * & * \\ * & * & * \\ * & b & * \\ * & b & * \\ * & * & * \\ b & b & * \\ * & * & * \\ * & * & * \\ * & b & b \\ * & * & * \end{bmatrix} \rightarrow \Uparrow; a \rightarrow a \right\rangle; \left\langle \begin{bmatrix} * & * & * \\ * & b & * \\ * & b & * \\ * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \end{bmatrix} \rightarrow \Downarrow; a \rightarrow a \right\rangle; \left\langle \begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \end{bmatrix} \rightarrow \Leftarrow; a \rightarrow a \right\rangle; \left\langle \begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \end{bmatrix} \rightarrow \Rightarrow; a \rightarrow a \right\rangle \right\}$$

Hvězdičky v matici znamenají, že agentu „nezáleží“ na obsahu odpovídajících buněk.

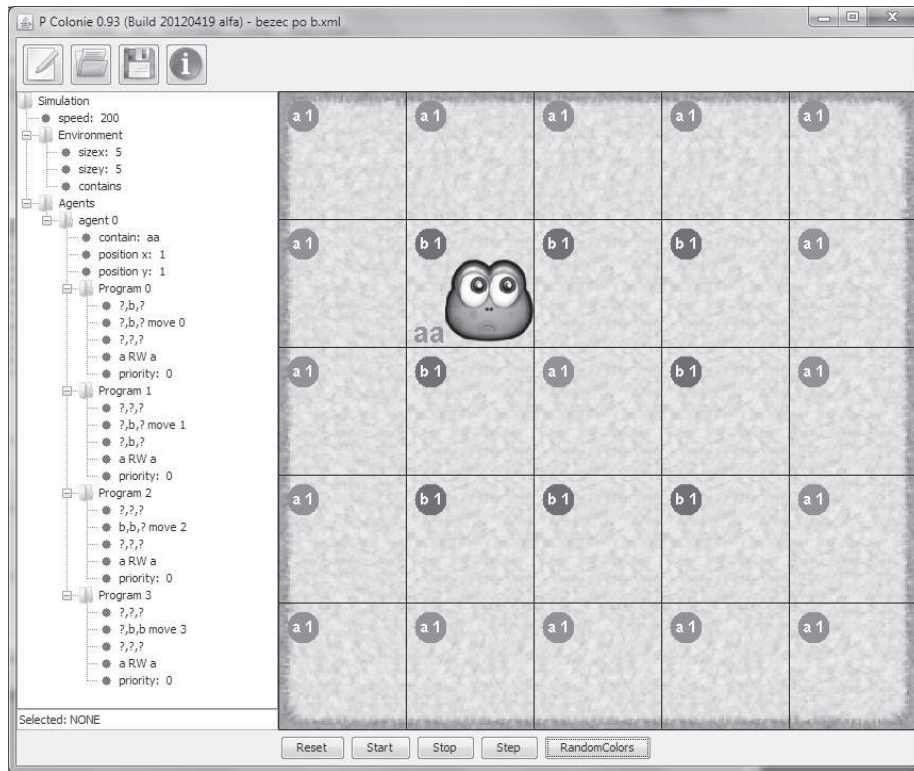
- $f \in A$  je konečný objekt kolonie.

Police agenta v prostředí je druhý řádek a druhý sloupec. Každé pohybové pravidlo má objekt  $b$  ve středu matice. Takže, všechny programy jsou aplikovatelné pouze v případě, kdy je agent umístěn v buňce obsahující alespoň jednu kopii objektu  $b$ . Na základě daných pohybových pravidel se agent přesune náhodně vybraným směrem do okolní buňky obsahující objekt  $b$ . Agent provede přesun v každém kroku výpočtu. Pro prostředí a obsah agenta se nemění. Počáteční konfigurace je na obrázku 1.

Druhý příklad 2D P kolonií je motivovaný Conwayovou Hrou života([7]). Britský matematik John Horton Conway v roce 1970 navrhl nejznámější příklad celulárního automatu. Uzavřený systém Hra života je nekonečná dvojrozměrná ortogonální síť čtvercových automatů, z nichž každý je v jednom ze dvou možných stavů, živý nebo mrtvý. Každý automat interaguje s osmi sousedy, které ho obklopují.

**Příklad 2** Necht'  $\Pi_2$  je 2D P kolonie definována tímto způsobem  $\Pi_2 = (A, e, Env, B_1, \dots, B_{16}, f)$ , kde

- $A = \{e, f, D, S, Z, M, O, L, N\}$ ,



Obrázek 1: Počáteční konfigurace  $\Pi_1$

- $e \in A$  je základní objekt prostředí kolonie,

- $Env = (6 \times 6, w_E)$ ,

$$\bullet w_E = \begin{bmatrix} D & D & D & D & D & D \\ D & S & S & D & D & D \\ D & S & S & D & D & D \\ D & D & D & S & S & D \\ D & D & D & S & S & D \\ D & D & D & D & D & D \end{bmatrix},$$

- $B_1 = (ee, P_1, [1, 1])$ ,  $B_2 = (ee, P_2, [1, 2]), \dots$ ,  
 $B_{16} = (ee, P_{16}, [4, 4])$ ,

- $f \in A$  je finální objekt kolonie.

Stavy automatu jsou uloženy uvnitř buněk ( $D$  - mrtvý automat,  $S$  - živý automat). Máme jen jeden druh agentů umístěných v matici  $4 \times 4$  vnitřních buněk (vidíme na obrázku 2). Množiny jejich programů jsou definovány v souladu s pravidly automatu:

- Živý automat s méně než dvěma živými sousedními automaty umírá kvůli malého výskytu živých automatů v okolí.

- Živý automat s více než třemi živými sousedními automaty umírá z důvodu velkého výskytu živých automatů v okolí.

- Živý automat se dvěma nebo třemi živými automaty v okolí nezmění svůj stav pro další generaci.

- Mrtvý automat právě se třemi živými automaty v okolí změní stav na živý.

První program  $\langle e \leftrightarrow e; e \rightarrow Z \rangle$  umožňuje inicializaci agenta;

Seřadíme programy podle počtu kopií objektu  $S$  v matici pohybového pravidla.

1. Pro situaci kdy všechny sousední automaty jsou ve stavu mrtvé, má agent jediný program nezávisle na stavu, ve kterém se nachází, tzn. zda je mrtvý či živý.

$$\left\langle \left[ \begin{array}{ccc} D & D & D \\ D & e & D \\ D & D & D \end{array} \right] \rightarrow \uparrow; Z \rightarrow M \right\rangle.$$

2. Pro situaci kdy je 1 ze sousedních automatů ve stavu živý, má agent ve stavu živý či mrtvý 8 možných programů



$$\left\langle \begin{bmatrix} S & D & D \\ D & e & D \\ D & D & D \end{bmatrix} \rightarrow \uparrow; Z \rightarrow M \right\rangle$$

a 7 dalších kombinací.

3. Pokud jsou 2 sousední automaty ve stavu živé, pak pro živé agenty je to 28 programů

$$\left\langle \begin{bmatrix} S & S & D \\ D & S & D \\ D & D & D \end{bmatrix} \rightarrow \uparrow; Z \rightarrow O \right\rangle$$

a 27 dalších kombinací.

4. Pokud jsou 2 sousední automaty ve stavu živé - 28 programů programů pro živé agenty

$$\left\langle \begin{bmatrix} S & S & D \\ D & D & D \\ D & D & D \end{bmatrix} \rightarrow \uparrow; Z \rightarrow M \right\rangle$$

a 27 dalších kombinací.

5. Pro případ 3 živých sousedních automatů - 56 programů pro živé a mrtvé agenty

$$\left\langle \begin{bmatrix} S & S & S \\ D & e & D \\ D & D & D \end{bmatrix} \rightarrow \uparrow; Z \rightarrow O \right\rangle$$

a 55 dalších kombinací.

6. Pro 4 živé sousední automaty - 70 možných programů pro mrtvé i živé agenty

$$\left\langle \begin{bmatrix} S & S & S \\ S & e & D \\ D & D & D \end{bmatrix} \rightarrow \uparrow; Z \rightarrow M \right\rangle$$

a dalších 69 kombinací.

7. Pro případ alespoň 5 živých sousedních automatů - 56 možných programů pro živé i mrtvé agenty

$$\left\langle \begin{bmatrix} S & S & S \\ S & e & S \\ * & * & * \end{bmatrix} \rightarrow \uparrow; Z \rightarrow M \right\rangle$$

a dalších 55 kombinací.

Po provedení jednoho z předchozích uvedených programů, všechny agenty se přesunou o jeden krok vpřed a přepíšou jeden ze svých objektů  $e$  na objekt  $M$  (automat bude mrtvý) nebo na objekt  $O$  (automat bude živý). Následující programy jsou určeny pro přesun agentů dolů a pro aktualizaci stavu automatů, tj. nahrazení objektu v buňce za objekt v agentu a tím dochází ke změně stavu automatu.

$$\left\langle \begin{bmatrix} * & * & * \\ * & e & * \\ * & * & * \end{bmatrix} \rightarrow \Downarrow; O \rightarrow S \right\rangle;$$

$$\left\langle \begin{bmatrix} * & * & * \\ * & e & * \\ * & * & * \end{bmatrix} \rightarrow \Downarrow; M \rightarrow D \right\rangle;$$

$$\langle e \rightarrow L; S \leftrightarrow S \rangle; \quad \langle e \rightarrow N; D \leftrightarrow S \rangle;$$

$$\langle S \rightarrow e; L \rightarrow e \rangle; \quad \langle S \rightarrow e; N \rightarrow e \rangle;$$

$$\langle e \rightarrow L; S \leftrightarrow D \rangle; \quad \langle e \rightarrow N; D \leftrightarrow D \rangle;$$

$$\langle D \rightarrow e; L \rightarrow e \rangle; \quad \langle D \rightarrow e; N \rightarrow e \rangle.$$

Je zřejmé, že můžeme tímto způsobem simulovat každý klasický celulární automat.

Ve třetím příkladě se budeme zabývat mravenci.

**Příklad 3** Cílem bude zkonstruovat 2D P kolonii simulující pohyb mravenců při hledání potravy. Agenty - mravenci - jsou umístěné do domovských buněk, ze kterých začínají hledat jídlo. Mravenci prohledávají nedeterministicky prostředí, dokud nenaleznou jídlo nebo stopu. Pokud najdou potravu, vezmou jeden „kus“ (jeden objekt) a vrátí se nejkratší cestou do domovské buňky. Tuto trasu si značí pomocí konkrétního objektu stopu. Pokud naleznou stopu, jdou po ní k jídlu.

Agenty v 2D P kolonii mají tolik programů, že uvedení seznamu a popsání daných programů by přesáhlo rámec tohoto článku.

Sestavili jsme agenty - mravence, kteří používají 57 programů. Agent prozkoumává prostředí. Jakmile nalezne potravu, vezme si jeden objekt potravy a vrací se do domovské buňky. Na zpáteční cestě umístí objekt-značku do každé buňky na této cestě. Jakmile přinese agent potravu do domácí buňky, vrací se po označované cestě zpátky ke zdroji potravy. Konfigurace se třemi agenty a se dvěma cestami je na obrázku 3. Po vyčerpání zdroje potravy se agent zastaví. Pokud vložíme do agenta program umožňující návrat agenta do domovské buňky, pak má agent možnost po nalezení dalšího zdroje tento zdroj vyčerpat nebo použít nové pravidlo, které ho bez potravy vrátí zpět do domovské buňky.

Daná situace může být řešena přidáním nového agenta 2D P kolonie. Tento nový agent nemá za úkol vyhledávat zdroj potravy, není tzv. průzkumníkem. Jeho programy umožňují odblokovat agenty, které se zastavily u vyčerpaného zdroje potravy. Alternativním řešením je přidání priorit mezi programy. Každému programu přiřadíme přirozené číslo představující důležitost programu. Během výpočtu je prováděn vždy aplikovatelný program s nejvyšším číslem, tzn. s nejvyšší prioritou.

#### 4 Implementace 2D Pkolonie - simulátor

Pro simulační prostředí použijeme jazyk Java. V simulátoru je možné načítat, ukládat a vytvářet simulace za pomoci značkovacího jazyka XML. Soubor simulace načítáme pomocí XML parseru a s využitím DOM a JAXP je vytvořena stromová struktura objektů. Tyto objekty představují parametry simulace, obsah prostředí a popis



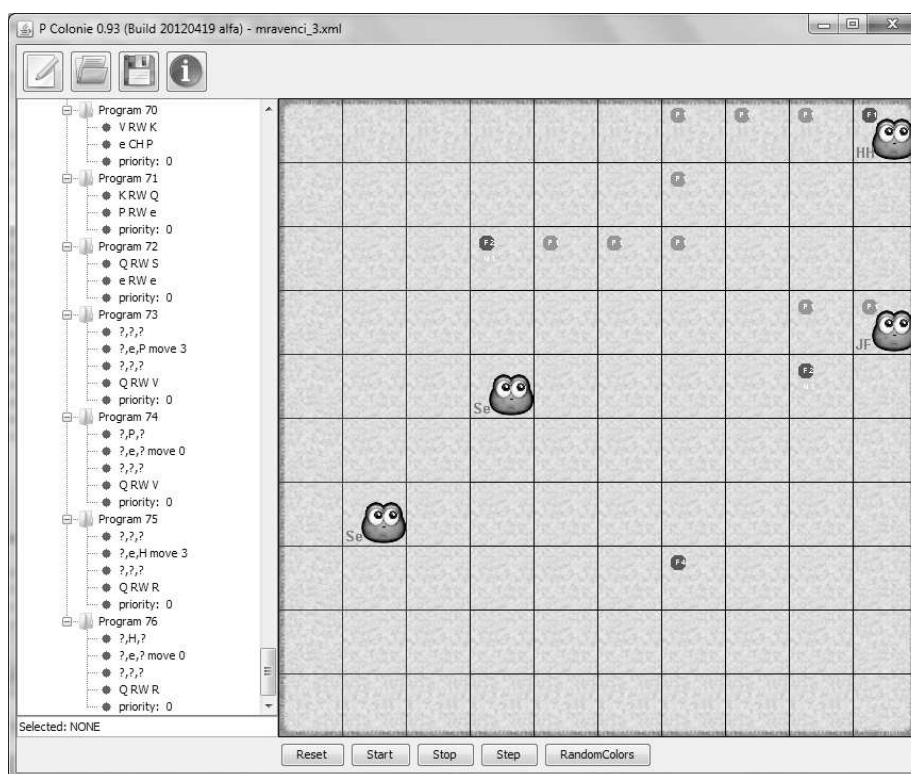
Obrázek 2: Počáteční konfigurace  $\Pi_2$

agentů nacházejících se v prostředí simulace. Informace popisující prostředí zahrnují několik parametrů: rychlost simulace, velikost prostředí a jeho obsah. Rychlost zde určuje časový interval mezi jednotlivými kroky simulace. Prostředí a jeho obsah reprezentujeme dvojrozměrným polem objektů, které je uživateli vyobrazeno jako 2D mřížka. Agent se nachází v této mřížce a má možnost se pohybovat prostředím a za pomoci prepisovacích pravidel ovlivňovat jeho obsah. Prostředí může obsahovat speciální objekt # reprezentující překážku nebo také pozici, na kterou se agenty nemohou dostat. V každém kroku simulace aktivuje každý agent v prostředí jeden ze svých aplikovatelných programů. Jednotlivým programům můžeme přiřadit priority a tímto způsobem ovlivnit výběr aplikovaného programu. V případě, že lze aktivovat několik programů se stejnou prioritou, využíváme pseudonáhodného výběru pouze jednoho z těchto použitelných programů. Simulační prostředí může obsahovat více agentů, které se mohou nacházet na různých nebo totožných pozicích. V jednom kroku simulace aplikuje každý agent v prostředí jeden ze svých aplikovatelných programů. Změny prostředí se ukládají do zásobníku, ze kterého jsou následně promítnuty do prostředí. Tímto způsobem jsme schopni se vyhnout situacím, kdy agent v rámci jednoho kroku simulace ovlivní okolí jiného agenta nebo objekty na

pozici, kde se nachází více agentů. V těchto případech by mohlo dojít k použití dříve nepoužitelných programů agenta. To však v rámci jednoho kroku simulace není přípustné. Pro grafické uživatelské rozhraní simulačního nástroje využíváme knihovny Swing. Vzhled prostředí, agentů a překážek je možné nahradit vlastní grafikou a přizpůsobit si tak simulační prostředí a vizualizaci simulace k dané problematice. S možností editace simulace přímo v simulačním prostředí nebo libovolném textovém editoru, se tak uživateli dostává do rukou nástroj pro provádění zajímavých simulací z oblasti membránových systémů.

## 5 Závěr

V tomto článku jsme představili nový typ P kolonií, který je vhodný pro simulaci reálných situací, například pro simulaci šíření vody při blezkových povodních v dané lokalitě v práci [6]. Vytvořili jsme 2-rozměrné prostředí, do kterého jsme umístili agenty. Agenty mají omezené informace o obsahu prostředí, což více odpovídá realitě. Pro simulaci v příkladu stigmergie jsme navíc použili mezi pravidly priority. Chování agentu je tedy více přirozené, protože v reálném světě mravenci dávají také přednost určitým ak-



**Obrázek 3:** Konfigurace  $\Pi_2$  se 4 mravenci a 2 cestami od zdroje potravy po domovskou buňku

tivitám před některými dalšími. Pro názornost a prezentaci jsme vytvořili simulační prostředí 2D P kolonií. Do budoucna plánujeme rozšíření simulátoru o statistické nástroje a o dynamické prostředí.

## Poděkování

Tento článek byl vypracován v rámci projektu Centrum excelence IT4Innovations, reg. č. CZ.1.05/1.1.00/02.0070 podporovaného Operačním programem Výzkum a vývoj pro inovace, financovaného ze strukturálních fondů EU a ze státního rozpočtu ČR, SGS/7/2011 a projektu OPVK č. CZ.1.07/2.2.00/28.0014.

## Reference

- [1] L. Cienciala, L. Ciencialová: Eco-P colonies. In: *Proceedings of the 10th International Workshop on Membrane Computing* (G. Păun, M. Pérez-Jiménez, A. Riscos-Núñez, eds.), Sevilla, 2009: 201 - 209.
- [2] L. Cienciala, L. Ciencialová: Eco-P kolonie. In: *Kognice a umělý život IX* (J. Kelemen, V. Kvasnička, J. Rybár, eds.), Opava, 2009: 47 - 56.
- [3] L. Cienciala, L. Ciencialová, E. Csuhanj-Varjú, G. Vaszil: P Col Automata: Recognizing Strings with P colonies. In: *Proceedings Eight Brainsorming Week on Membrane Computing* (M. A. Martínez del Amor, G. Păun, I. P. Hurtado de Mendoza, A. Riscos-Núñez, eds.), Fénix editora, Sevilla, 2010: 65 - 76.
- [4] L. Cienciala, L. Ciencialová, E. Csuhanj-Varjú, G. Vaszil: P kolonie řízena vstupem - PCol automat. In: *Kognice a umělý život X* (J. Kelemen, V. Kvasnička, eds.), Slezská univerzita Opava, 2010: 61 - 67.
- [5] L. Cienciala, L. Ciencialová, M. Perdek, M.: 2D P colonies. In: *Proceedings CMC 2012* (E. Csuhanj-Varjú et al. , eds.), Springer, LNCS 7762, 2013: 161-172.
- [6] L. Cienciala, L. Ciencialová, M. Langer: 2D P colonies and modelling of liquid flow over the earth's surface. V přípravě.
- [7] M. Gardner: Mathematical Games - The fantastic combinations of John Conway's new solitaire game "life". *Scientific American* 223. (1970) 120-123. Archived from the original on 2009-06-03. Retrieved 2011-06-26.

- [8] J. Kelemen, A. Kelemenová, Gh. Păun. Preview of P colonies: A biochemically inspired computing model. In: *Workshop and Tutorial Proceedings. Ninth International Conference on the Simulation and Synthesis of Living Systems (Alife IX)*. (M. Beaud et al. , eds.), Boston Mass., 2004: 82 - 86.
- [9] J. Kelemen, A. Kelemenová. On P colonies, a biochemically inspired model of computation. In: *Proc. of the 6<sup>th</sup> International Symposium of Hungarian Researchers on Computational Intelligence, Budapest TECH*, Hungary, 2005: 40 - 56.
- [10] Gh. Păun. Computing with membranes. *Journal of Computer and System Sciences* 61 (2000) 108-143.
- [11] Gh. Păun. *Membrane computing: An introduction*, Springer-Verlag, Berlin, 2002.
- [12] Gh. Păun, Gh. Rozenberg, A. Salomaa: *The Oxford Handbook of Membrane Computing*, Oxford University Press, 2009.
- [13] P systems web page: <http://ppage.psyste.ms.eu> [online 2013-3-22].

# Principy zdravého rozumu

Martina Číhalová, Jan Štěpán

Filozofická fakulta Univerzity Palackého v Olomouci, Křížkovského 10, 771 47 Olomouc  
m.tina.cihal@gmail.com, jan.stepan@upol.cz

## Abstrakt

Cílem příspěvku je představit problematiku zdravého rozumu a specifikovat jeho základní principy. Pojem zdravého rozumu je často směřován s tzv. rozumem selským, proti tomuto pojetí se však příspěvek vymezuje. Jsou uvedeny základní principy zdravého rozumu, jako jsou například předpoklad ontologického realismu, důvěra v přímou zkušenost, důvěra v přijatý kognitivní rámec a schopnost využívání inferenčních mechanismů, které jsou v příspěvku podrobněji představeny. Je rozlišována inference pojmová a propoziční, kdy inference pojmová se týká odvozování nových pojmů z pojmů již existujících a propoziční pak odvozování propozic z jiné skupiny propozic. Inference pojmová probíhá metodou konkretizace, abstrakce, celek-část a škálováním. Propoziční inference pak na základě metod indukce, dedukce a analogie. Inference propoziční operuje na přijatém systému přesvědčení, kdy přesvědčení je tvrzení, které je justifikováno, přičemž v příspěvku jsou uvedeny různé typy justifikace. Je ukázán vztah zdravého rozumu k vědeckému poznání, kdy je zdravý rozum považován za jeho předpoklad.

## 1 Úvod

Pojem zdravého rozumu (*common sense*) je velmi těžké přesně vymezit, neboť je často chápán spíše intuitivně. Tento pojem byl zaveden v 18. a počátkem 19. století filozofy skotské školy v čele s Thomasem Reidem, kteří ve zdravém rozumu viděli zdroj poznání evidentních pravd, předpoklad poznání a pramene jistoty o existenci Boha. Vycházeli z toho, že co se přiči zdravému rozumu, vede na scestí a k chybným závěrům. Toto pojetí však zdravý rozum v podstatě staví do protikladu s některými odvětvími věd, jako je kvantová mechanika nebo obecná teorie relativity. Pojem zdravého rozumu dále rozpracoval G. E. Moore v [1]. Byl přesvědčen o tom, že víry a přesvědčení o světě, které jsou obsaženy v běžném jazyce, tj. víry „zdravého rozumu“, jsou správné, tak jak jsou. Úkolem filosofie proto podle něj nemůže být bádání či diskuse o pravdivosti toho, co jazykem vyjadřujeme, ale spíše hledání běžnému jazyku přiměřeného způsobu analýzy

významů toho, co je jím řečeno, neboli analýzy odhalující právě tato přesvědčení, která jsou z běžného užívání jazyka rozpoznatelná. Hájíl tezi, že všichni dobře známe pravdivost mnohých propozic o nás samotných, našich tělech i ostatních lidech, přestože si nemusíme být jisti v případech jejich analýzy. Jak idealisté, tak skeptici naopak odmítají toto jednoduché, každodenní vědění.

Zdravý rozum je často spojován se *selským rozumem*, který je lidovou psychologíí. Toto spojení pak přináší problémy v tom smyslu, že zdravý rozum se často mylí, naše intuice jsou často nesprávné a podléháme sebeklamu a předpojatostem mnoha druhů. Pokud bychom zdravý rozum chápali tímto způsobem, pak právě různé druhy omylů mohou vést k předpokladu lidového relativismu. Právě proti takovému pojetí zdravého rozumu se chce tento příspěvek vymezit a ukázat, že zdravý rozum není v rozporu s vědeckým poznáním, naopak tvoří právě jeho základy.

V oddíle 2 budou představeny různé přístupy k problematice zdravého rozumu a bude vymezen jeho pojem tak, jak jej chápeme v tomto příspěvku. V následujícím oddíle 3 se budeme věnovat inferenci, jako jednomu z principů zdravého rozumu. V oddíle 4 budou specifikovány základní principy zdravého rozumu, aby pak v následujícím oddíle 5 byla vedena paralela mezi těmito principy a vědeckým poznáním.

## 2 Vymezení zdravého rozumu

Zdravý rozum je spojen s naší každodenní zkušeností a v podstatě umožňuje orientaci ve světě této každodenní zkušenosti.

*Zdravý rozum a jeho svět jsou spojovány s každodenním myšlením, s každodenním jednáním a prožíváním a každodenním světem ... [2, s. 93]*

Často je zdravý rozum ztotožňován s určitým specifickým systémem přesvědčení, jak je tomu například u filozofů vycházejících z G. E. Moora a T. Reida. T. Reid přesvědčení zdravého rozumu vysvětloval lidskou povahou, jež je stvořena Bohem, zárukou jejich pravdivosti mu byl tedy Bůh. G. E. Moore vycházel z přesvědčení, že stanoviska zdravého rozumu jsou správná a evidentní sama o sobě a nepotřebují dalšího důkazu nebo zjištění jejich pramene.

*And there are, of course, enormous numbers of other features in 'the Common Sense view of the world' which, if these are true, are quite certainly true too: e.g. that there have lived on the surface of the earth not only human beings, but also many different species of plants and animals, etc. [1, s. 45]*

Stephen Boulter při formulování konkrétních přesvědčení týkajících se zdravého rozumu vychází z G. E. Moora (*Some Main Problems of Philosophy*) [3] a T. Reida (*Essays on the Intellectual Powers, Essays on the Active Powers*) [4], [5] a shrnuje je do následujících 7 základních:

- *That there is a real world that exists independently of us, of our thoughts, language and representations*
- *That human beings have direct, non-projective perceptual access to this world via the senses*
- *That causation is a real relation among objects and events in the world, a relation whereby one phenomenon, the cause, brings about another, the effect*
- *That statements are either true or false in virtue of states of affairs in the world*
- *That human beings have beliefs, desires, hopes and fears and other mental states to which one can appeal to in order to explain and predict human actions*
- *That human beings are responsible for their actions in certain specifiable circumstances, and so are proper objects of approval, condemnation, praise, blame and punishment. [6, s. 30]*

Noath Lemos pak tento systém přesvědčení dává do souvislosti s všeobecnou znalostí (common knowledge).

*Moreover, the common sense tradition holds that some common sense propositions are known to be true. For example, it holds that we do know that there are other people, that they have bodies, that they think, and that they know various things about the world. Indeed, the tradition holds that these and many other common sense propositions are such that almost everyone knows them. In this respect, the tradition holds that some common sense propositions are matters of common knowledge. [7, s. 4]*

Alfred Schuetz vymezuje zdravý rozum poněkud šíře a řadí k jeho charakteristikám i organizaci myšlení prostřednictvím abstrakce, generalizace, formalizace a idealizace:

*All our knowledge of the world, in common-sense as well as in scientific thinking, involves constructs, namely, a set of abstractions, generalizations, formalizations, idealizations specific to the respective level of thought organization. [8, s. 2]*

V tomto příspěvku budeme zdravý rozum rovněž chápat poněkud šíře, než jako systém určitých přesvědčení. Zdravý rozum budeme považovat za vrozenou *dispozici, která umožňuje veškeré kognitivní schopnosti člověka*. A sem patří nejen nějaký systém základních přesvědčení, ale především také inferenční mechanismy, které umožňují rozšiřování tohoto systému, případně jeho revizi.

### 3 Inference

Inferenci se nejčastěji rozumí odvozování jedněch výroků z jiných. J. Křemen u kapitoly o inferenci uvádí, že

*... nesmírně složitá nativní a sémantická inference se ve službách exaktní vědy užívající formální jazyky ujala jen vtěch nejjednodušších rozpoznatelných složkách, které jsou ve schematizujícím zjednodušení dobře definovatelné. Jsou to indukce, dedukce a analogie. [9, s. 20]*

Budeme rozlišovat inferenci *pojmovou*, tj. odvozování nových pojmů z pojmů konceptuálního schématu a inferenci *propoziční*, tj. odvozování propozic z jiné skupiny propozic.

Jakmile je dáno určité konceptuální schéma pojmů a vztahů mezi nimi, člověk může vydělovat nové pojmy, na základě pojmů již existujících, které toto schéma dále obohacují. Pokud například pojmově rozlišil „ryby“ od „ptáků“, může na základě metody *konkretizace* odlišovat různé druhy ptáků jako je „tetřev“, „strnad“ apod. Naopak může pojmy „ryby“ a „ptáci“ shrnout pod společný obecnější pojem „živočicha“ metodou *abstrakce*.

Další metodou pro vydělování nových pojmů z pojmů již existujících je vztah *celek – část*. Jakmile již existuje nějaký pojem pro entity, které jsou strukturované, lze si všimnout částí tohoto celku a uchopit je prostřednictvím nových pojmů. Tak si uvědoměním částí konkrétní ryby vyděluje člověk pojmy jako „ploutev“, „zuby“ aj. Metodou *celek – část* je i proces opačný, kdy na základě pojmenované entity pojmenujeme nově i sdružení těchto entit, tj. například „společenství jednotlivých stromů“ může být souhrnně pojmenováno jako „les“.

Stejně tak *kontrární* pojmy, které vnímáme pomocí uvědomění si nějakého kontrastu mezi entitami, které popisují. Kontrární pojmy jsou například pojmy pro vztahy v prostoru jako „vpravo“, „vlevo“, „vepředu“,

„vzadu“; dále například pojmy označující jednotlivé barvy.

Rozlišení mezi pojmy protikladnými a kontrárními lze provést pomocí operátoru negace. Pokud znegujeme kontrární pojem, dostaneme jeho protějšek. Pokud však znegujeme pojem kontrastní, dostaneme doplněk, který tvoří celá skupina pojmů. Například když řekneme *toto není červené*, tak barva předmětu může být vyjádřena kterýmkoli pojmem pro barvy, který je fixovaný v daném konceptuálním schématu. Rozdíl mezi protikladnými a kontrárními pojmy poněkud splývá v případech fuzzy pojmů jako „mladý“, „rychlý“, „velký“, „studený“, „krásný“, apod. Fuzzy pojmy však můžeme pojmově defuzzifikovat prostřednictvím definování nějaké škály či stupnice, a zde se již jedná o *inferenci pojmovou*, nazvěme tuto metodu *škálování*.

Dalším typem pojmové inference je *metafora*, jde o vydělování nových pojmů prostřednictvím přenášení významu na základě vnější nebo vnitřní podobnosti.

Metody *pojmové inference* jsou tedy výše zmíněné metoda *abstrakce*, *konkretizace*, *metoda celek-část*, *škálování*, *metafora* a případně nějaké další, které nejsou v příspěvku reflektovány.

Inference se však týká také odvozování propozic v rámci báze poznatků. Zde pak patří známé formy inference jako *indukce*, *abdukce*, *dedukce* a *analogie*. V případě indukce, abdukce a analogie se neobejdeme bez teorie pravděpodobnosti. Z výše uvedených inferenčních mechanismů si pro nedostatek prostoru blíže specifikujeme pouze analogii.

*Analogie je úsudek založený na stanovení podobnosti dvou různých objektů na základě částečně společných znaků. Uvažované objekty se tedy musí lišit aspoň v jednom znaku, který při této úvaze považujeme za nepodstatný. Pravděpodobnostní úsudek z analogie pro individua a a b má tvar:*

$$P_1(a), \dots, P_n(a), P_1(b), \dots, P_{i-1}(b), P_{i+1}(b), \dots, P_n(b) \gg P_i(b),$$

*tj. ta individua se liší v jediném z (podstatných) znaků, který má (ověřeně) a, soudíme tedy, že jej má (pravděpodobně) i b. Spolehlivost úsudků z analogie roste s počtem společných znaků. Zřejmě nemá smysl pro  $n = 1$ .*

*Např. úsudek: Země je planeta. Země obíhá kolem Slunce. Země se otáčí kolem své osy. Merkur je planeta. Merkur obíhá kolem Slunce. Tudiž Merkur se otáčí kolem své osy.*

*Obdobně se postupuje v případě strukturální analogie. Mějme systémy*

$$S_1 = [A, \{R_i\}], \text{ kde } i = 1, 2, \dots, n, \text{ kde } n > 1, \text{ a}$$

$$S_2 = [B, \{P_j\}], \text{ kde } j = 1, 2, \dots, n-1.$$

*Nechť existuje homomorfismus nebo izomorfismus  $S_1$  na  $S_2$  mimo relaci  $R_n$ . Pak usuzujeme na existenci jisté relace  $P_n$  analogické relaci  $R_n$ , která vyhovuje příslušnému zobrazení mezi uvažovanými systémy. Závěr i zde bude tím spolehlivější, čím "větší" bude podobnost struktur obou systémů především kvantitativně, tj. čím vyšší bude  $n$ .*

*Analogie tohoto druhu např. zdůvodňuje tzv. pokusy se zvířaty. Zvířecí a lidský organismus jsou podobné. Když určitý lék pomáhá nemocnému zvířecímu organismu, tedy pomůže i lidskému. [10, s. 60]*

Existují ještě jiné typy inference než výše uvedené? David Danks [11, s. 451] hovoří o *kauzální inferenci*, kdy jde o asociativní učení, na základě něhož z nějaké opakující se vazby mezi dvěma jevy či událostmi usuzujeme, že se mezi nimi jedná o vztah příčinnosti. Zde však nelze o inferenci hovořit. Přisouzení kauzálního vztahu dvěma objektům je apriorní hypotézou, kdy tato hypotéza je podmíněná vrozenou dispozicí ke kladení této hypotézy. Na kauzalitu se z opakovaného spojení mezi dvěma jevy neusuzuje, ta se apriori předpokládá a dále se pak využívá jako předpoklad pro inferenci výše uvedené.

### 3 Principy zdravého rozumu

V této kapitole se pokusíme shrnout základní principy zdravého rozumu, které v podstatě zdravý rozum specifikují. Zdravý rozum se týká jednak našich přesvědčení a jednak inferenčních mechanismů aplikovaných na tato přesvědčení. Za principy zdravého rozumu budeme považovat následující:

Týkající se přesvědčení:

- Apriorní důvěra v bezprostřední zkušenost subjektu<sup>1</sup>
- Předpoklad ontologického realismu<sup>2</sup>
- Schopnost uvažovat negaci
- Schopnost revize přesvědčení

Týkající se inferenčních mechanismů:

- Schopnost pojmové inference
- Schopnost indukce, abdukce, dedukce
- Dispozice k všímání si podobných jevů a inference z analogie
- Dispozice k všímání si spojení mezi jevy

<sup>1</sup> Nikoliv zkušenost zprostředkovaná, ať už slovem mluveným, nebo projevem písemným.

<sup>2</sup> Jedná se o předpoklad, že svět má svoji vlastní ontologickou realitu nezávislou na naší percepci a kognitivních mechanismech.

- Schopnost uvažování na základě pravděpodobnosti
- Dispozice k vytváření modelů světa za použití výše uvedených principů

Dodržování principů zdravého rozumu samo o sobě nevede k poznání pravdivému. Dodržování principů zdravého rozumu však umožňuje možnost racionální diskuse ve společnosti o tom, co je pravdivé a co nikoli, o tom, které principy je potřeba dodržovat, o relevanci určitých příčin a důsledků.

#### 4 Zdravý rozum a vědecké myšlení

Tato kapitola se zabývá otázkou, jaký má zdravý rozum vztah k vědeckému poznání. Častým argumentem, který vede k oddělování vědeckého poznání a zdravého rozumu je to, že zdravý rozum se často mýlí. Fakt, že zdravý rozum se mýlí, neznamená však jeho zpochybnění či zavržení. Ve výše uvedených principech zdravého rozumu zahrnujeme potencialitu zdravého rozumu k revizi vlastních přesvědčení. Zdravý rozum je tak rozumem kritickým a je právě tou okolností, která nám umožňuje rozlišit pravdu od klamu a uvědomit si skutečnost, že jsme se mýlili. Naši omylnost si tak uvědomujeme právě prostřednictvím zdravého rozumu, jak o tom hovoří například Barry Smith.

*Thus common sense is not, in spite of this reputation, naive; it draws a systematic distinction between reality and appearance, or in other words between the way the world is and the way the world seems or appears via one or other of the sensory modalities and from the perspective of one or other perceiving subject in one other context. ... The putative information supplied by perception is always partial, and sometimes erroneous, but it can in every day case be supplemented and corrected by the gathering of further information about the sides of objects we cannot see, about the future behaviour of objects, and so on. [12 s. 645]*

Obdobný názor zastává i Howard Sankey, když hovoří o tom, že přijmout vědecké poznatky, které jsou v rozporu s naší perceptuální zkušeností, neznamená odvržení zdravého rozumu. Věda podává vysvětlení, proč je naše perceptuální zkušenost taková, jaká je, ale nezpochybňuje tuto zkušenost jako takovou [13, s. 20].<sup>3</sup>

<sup>3</sup> *Science discovers the truth about the independently existing world in which we find ourselves. It starts from common sense, which embodies a realist view of the objects of every day experience. Occasionally it conflicts with common sense. But science does not lead to the overthrow of common sense. It explains why commonsense objects appear as they do. It*

Empirická evidence dostupná zdravému rozumu pak hraje hlavní roli u vědeckého experimentu.

*Conformity with empirical evidence is a standard requirement for theory-acceptance in science. Because it is primarily observational, empirical evidence typically forms part of or is at least available to common sense. [13, s. 16].*

Schuetz dále hovoří o tom, že racionalita zdravého rozumu má s vědeckou racionalitou společné to, že buduje konstrukty pomocí metody abstrakce, formalizace, zobecňování a idealizace, že však konstrukty zdravého rozumu jsou nahrazeny konstrukty vědy.

*All our knowledge of the world, in common-sense as well as in scientific thinking, involves constructs, namely, a set of abstractions, generalizations, formalizations, idealizations specific to the respective level of thought organization. Strictly speaking, there are no such things as facts, pure and simple. All facts are from the outset facts selected from a universal context by the activities of our mind. They are, therefore, always interpreted facts, namely, either facts looked at as detached from their context by an artificial abstraction or facts considered in their particular setting. ... This does not mean that, in daily life or in science, we are unable to grasp the reality of the world. It just means that we grasp merely certain aspects of it, namely those which are relevant to us either for carrying on our business of living or from the point of view of a body of accepted rules of procedure of thinking called the method of science. If, according to this view, all scientific constructs are designed to supersede the constructs of common sense thought ... [8, s. 2-3]*

Je potřeba si na tomto místě uvědomit, že konstrukty vědy vykazují vyšší míru abstrakce, generalizace, formalizace a idealizace (vzpomeňme jen na fyzikální zákony, které se vztahují na idealizované prostředí). Nahrazení konstruktů zdravého rozumu konstrukty vědy však vychází právě z kompetencí zdravého rozumu, neboť zdravý rozum je kritický a je schopen reflexe svých vlastních omylů. Vědecká teorie tvořená pojmy a vztahy vykazující vyšší míru abstrakce umožňuje vyšší schopnost predikce jevů než různé folk teorie, a tedy má vyšší explanační sílu. Věda využívá určité obecné metody, které jsou jejím instrumentariem. A přestože se nedospělo jejich zobecňováním k žádné univerzální metodě poznání, je jejich efektivnost nesporná. Znalost obecných metod a dodržování všech pravidel v nich obsažených patří k profesionalitě vědecké práce. Jak je uvedeno v [14, s. 16]: jedná se o metody

*explains why in some cases the commonsense appearance of things is misleading. [13, s. 20]*



empirické a obecně teoretické: Mezi empirické metody patří pozorování, měření a experiment, mezi obecně teoretické jsou pak syntéza, dedukce, indukce, analogie, srovnání, specifikace.

Domníváme se, že nelze vést ostré oddělení zdravého rozumu a vědeckého myšlení, neboť zdravý rozum, jako vrozená racionální schopnost umožňující poznání, má ve své potenci precizování svých vlastních pojmů a jejich vztahů a také svých metod, které ho k poznání vedou.

Jiří Nosek spatřuje rozdíl mezi vědeckým myšlením a zdravým rozumem v tom, že teorie používané zdravým rozumem jsou užívány nereflektivně, neuvědoměle, zatímco vědecký rozum je používá reflektivně a uvědoměle.

*Rozdíl mezi metodou používanou zdravým rozumem a metodou používanou vědeckým rozumem by byl opět v tom, že zdravý rozum si neuvědomuje, že této metodě podléhá, že si tedy neuvědomuje, že jeho samozřejmé a běžné pravdy se stanou jednou omyly, a proto trvá na jejich samozřejmosti a běžnosti. [2 s. 148]*

Je pravdou, že až reflexe na principy a metody zdravého rozumu umožnila vznik vědecké racionality, ale tato reflexe musí být zároveň potencialitou zdravého rozumu. Zkušenost subjektu však ukazuje, že přesvědčení zdravého rozumu se jednoho dne mohou stát omyly a netrvá tak na jejich samozřejmosti, v rozporu s tím, co tvrdí Nosek. Jedním z principů zdravého rozumu pak je schopnost provádět revizi svých přesvědčení. Taková revize předpokládá, že zdravý rozum je schopen vidět situaci, která ho k omylu vedla, ve větší komplexnosti, což předpokládá právě schopnost reflektovat své principy. Zdravý rozum má k této reflexi dispozici.

Přesvědčení zdravého rozumu jsou tvrzení, která jsou justifikována. Typy justifikace pak mohou být různé. Nejčastěji se jedná o justifikaci pomocí přímé zkušenosti, může však být i pomocí zkušenosti zprostředkované (jako je zkušenost sdělená, vyčtená apod.).

Justifikační pravidla mohou být správná, nebo nesprávná. A. I. Goldmann rozlišuje několik kritérií správnosti systému justifikačních pravidel:

*(C1) R is a system of rules derivable from logic (and probability theory).*

*(C1\*) R is the system of rules that would be chosen by someone who believes all truths about logic (and probability theory), but is ignorant of all contingent facts.*

*(C2) R is the system of J-rules accepted by the players of one's language game (Wittgenstein)<sup>4</sup>*

<sup>4</sup> Pojem jazykové hry je diskutabilní, neboť Wittgenstein nikde ve svém díle [15] nepodává přesnou definici tohoto termínu.

*(C2\*) R is the system of J-rules accepted by members of one's disciplinary matrix (Kuhn).*

*(C2\*\*) R is the system of J-rules accepted by one's peers (Rorty).*

*(C3) Conformity with R would guarantee a coherent set of beliefs.*

*(C4) R permits doxastic attitudes proportioned to the strength of one's evidence.*

*(C5) Conformity with R would maximize the total number of true beliefs a cognizer would obtain.*

[16, s. 66]

Výše uvedené typy kritérií správnosti však nejsou neproblematické a mohou se vzájemně doplňovat. Pokud jde o kritéria správnosti, nabízí se meta otázka, které kritérium správnosti je správné. Až tehdy, pokud je přesvědčení justifikováno správným kritériem správnosti, můžeme hovořit o znalosti. V praxi však neexistuje žádné kritérium, které by nám definitivně zaručovalo správnost našich přesvědčení, a to ani ve vědeckém myšlení. Ve vědeckém myšlení místo o přesvědčeních hovoříme o hypotézách. Hypotéza musí být testovatelná a kompatibilní s dřívějšími hypotézami. Pokud je experimentálně potvrzená, je justifikována, což samozřejmě neznamená to, že je dokázána a jedná se tak o znalost. K justifikaci hypotéz ve vědeckém myšlení je využíváno empirické testování. Přesvědčení zdravého rozumu však vždy nepodléhají takto přísným justifikačním kritériím. Zdravý rozum je však schopen rozlišit míru přijatelnosti různých justifikačních kritérií na základě reflektování četnosti omylů, které jsou způsobeny mylnými přesvědčeními. Tak je spíše revidováno nějaké přesvědčení, které je založeno na sekundární obezřetnosti, než přesvědčení, které vychází z vlastní zkušenosti. Je v potenci zdravého rozumu precizovat justifikační kritéria tvrzení na základě četnosti omylů, k nimž vedly. Experimentální metoda je pak zdravým rozumem vyhodnocena jako nejspolehlivější justifikační kritérium.

## 6. Závěr

Zdravý rozum se netýká pouze systému našich přesvědčení, ale také inferenčních mechanismů, které operují na tomto systému. Principy zdravého rozumu jsou nám vrozené a reflexe na tyto principy a jejich precizace

---

Jazykovou hrou tak může být v podstatě i komunikace na základě rozkazů, viz jeho příklad s dělníky ze str. 14. V rámci takové jazykové hry nelze mluvit o justifikaci, neboť justifikaci podléhají pouze tvrzení, nikoli rozkazy.

umožnila vznik vědecké metody. Vědecké myšlení tak není v rozporu se zdravým rozumem, ale zdravý rozum je podmínkou možnosti vzniku vědeckého myšlení vůbec.

## Poděkování

Tento příspěvek byl podporován projektem „Podpora vytváření excelentních výzkumných týmů a intersektorální mobility na Univerzitě Palackého v Olomouci“, reg. č. CZ.1.07/2.3.00/30.0004 (POST-UP).

## Literatura

- [1] Moore, G. E.: A Defence of Common Sense, *Philosophical Papers*, London, George Allenand Undin, p. 45., 1959.
- [2] Nosek, J.: *Věda a zdravý rozum. Komparativní studie dvou kompetencí racionality*, Plzeň, Západočeská univerzita v Plzni, 2007, ISBN: 978-80-7043-639-4.
- [3] Moore, G. E.: Some Main Problems of Philosophy. *Philosophy* (ročník 31, číslo 119), s. 362-366, 1956.
- [4] Reid, T.: *The Works of Thomas Reid*. Bristol, Thoemmes Press, 1994.
- [5] Reid, T.: *The Works of Thomas Reid*. Boston, Adamant Media Corporation, 2005.
- [6] Boulter, S.: *The Rediscovery of Common Sense Philosophy*. New York, Palgrave Macmillan, 2007, ISBN-10: 0-2300-0246-3
- [7] Lemos, N.: *Common Sense. A Contemporary Defence*. New York, Cambridge University Press, 2004, ISBN-13 978-0-521-83784-2.
- [8] Schuetz, A.: Common-sense and scientific interpretation of human action. *Philosophy and Phenomenological Research* 14 (1), s. 1-38, 1953.
- [9] Křemen, J.: *Modely a systémy*, Praha, Academia, 2007, ISBN: 978-80-200-1477-1.
- [10] Štěpán J.: *Logika a právo*. 3. vydání. Praha, C. H. Beck, 2011
- [11] Danks, D. (2012): Psychology of causal perception. *The Oxford Handbook of Causation*, eds. H. Beebe, Ch. Hitchcock, P. Menzies, Oxford, Oxford University Press, s. 447-464, 2012, ISBN: 978-0-19-964258-8.
- [12] Smith, B.: *Formal Ontology, Common Sense and Cognitive Science*, 1995.  
<http://ontology.buffalo.edu/focscs.pdf>  
(stránka navštívena: 11. 4. 2013)
- [13] Sankey, H.: Science, Common Sense and Reality. In J. E. Ford (ed.), *The Ontic Return: The Current Epochal Shift from Meaning to Being*. Palgrave Macmillan, 2010.  
[http://philsci-archive.pitt.edu/5560/1/Science%2C\\_Common\\_Sense\\_and\\_Reality.pdf](http://philsci-archive.pitt.edu/5560/1/Science%2C_Common_Sense_and_Reality.pdf) (stránka navštívena: 11. 4. 2013)
- [14] Ochrana, F.: *Metodologie vědy. Úvod do problému*. Praha, Karolinum, 2009, ISBN 978-80246-1609-4.
- [15] Wittgenstein, L.: *Filosofická zkoumání*. Praha, Filosofický ústav AVČR, 1993, ISBN 80-7007-040-4.
- [16] Goldman, A. I.: *Epistemology and Cognition*. Cambridge, Massachusetts and London, Harvard University Press, 1988, ISBN: 978-0674258969

## Emócie a kognícia: ich interakcie a poruchy (vybrané aspekty)

Barbora Cimrová<sup>1,2,3</sup>, Fedor Jagla<sup>1</sup>, Igor Riečanský<sup>1,4</sup>

<sup>1</sup> Laboratórium kognitívnej neurovedy, Ústav normálnej a patologickej fyziológie, Slovenská akadémia vied, Sienkiewiczova 1, 813 71 Bratislava

<sup>2</sup> Centrum pre kognitívnu vedu, Fakulta matematiky, fyziky a informatiky, Univerzita Komenského v Bratislave, Mlynská dolina, 842 48 Bratislava

<sup>3</sup> Katedra psychológie, Filozofická fakulta, Trnavská univerzita v Trnave, Hornopotočná 2, 918 43 Trnava

<sup>4</sup> SCAN-Unit, Fakulta psychológie, Viedenská univerzita, Liebiggasse 5, A-1010 Viedeň, Rakúsko

{barbora.cimrova, fedor.jagla, igor.riecansky}@savba.sk

### Abstrakt

Kognícia a emócie sú dva významné aspekty ľudskej psychiky, ktoré sa tradične považujú za rozdielne, ba až protichodné. Vďaka rozvoju funkčných zobrazovacích techník je v súčasnosti do značnej miery možné priame skúmanie oboch týchto systémov, pričom výskumy posledných rokov potvrdzujú obrovskú mieru ich prepojenia a interakcie. V rozpore s tradičným pohľadom na protiklad rozumu a emócií sa zdá, že za určitých okolností môžu emočné procesy ovplyvniť uvažovanie a rozhodovanie výhodným spôsobom. Podobne, pri narušenom prežívaní emócií dochádza často aj k neadekvátnemu kognitívnemu spracovaniu. Pri viacerých psychických poruchách sú pozorované emočné aj kognitívne dysfunkcie, pričom smer ich interakcie nie je jednoznačný. V nedávnej štúdií sme skúmali vzťah medzi úzkosťou a explicitnou pamäťou. Naše výsledky poukazujú na rozdielny mechanizmus zabezpečujúci vybavovanie si emočných informácií z pamäte, v závislosti od miery úzkosti.

### 1 Kognícia a emócie – úvod

Kognícia, z latinského „cognoscere“ – „poznávať“ predstavuje schopnosť človeka získať poznanie. Zahŕňa procesy ako vnímanie, uvedomovanie si, pamäť, pozornosť, predstavivosť, porozumenie reči, jej tvorenie a používanie, schopnosť počítať, exekutívne funkcie ako sú plánovanie, riešenie problémov, rozhodovanie sa atď. Teda všetky mentálne procesy, ktoré umožňujú získať poznatky a vedomosti o veciach vnímaných, ako aj výsledky z informácií spracúvaných myslou.

Vzťah emócie a kognície je témou diskutovanou už od dávnych dôb. O jeho opísanie, či bližšie preskúmanie sa pokúšali už mnohí filozofi, učitelia, či špičkové vedecké tímy. Vďaka rozvoju funkčných zobrazovacích techník je v súčasnosti do značnej miery možné priame skúmanie oboch týchto systémov, pričom výskumy posledných rokov potvrdzujú obrovskú mieru ich prepojenia a interakcie. Tieto interakcie môžu prebiehať vo viacerých formách a na rozličných úrovniach. Niektoré formy týchto interakcií sú z hľadiska funkčnej organizácie mozgu skôr difúzne, alebo nešpecifické, iné sú detailnejšie a vyžadujú si účasť komplikovanejších prepojení v rámci funkčnej organizácie mozgu. Napriek mnohým poznatkom, ku ktorým sa v priebehu tohto času dospelo, zostáva vzťah medzi emóciami a kogníciou naďalej nedostatočne objasnený.

Kognícia a emócia sú dva významné aspekty ľudskej psychiky sú z tradičného hľadiska považované za rozdielne, či dokonca protichodné (Oxford dictionary: „cognition as opposed to emotion and volition“). A skutočne: „rácio“ býva často „zahmlené emóciami“, „triezvenu úsudku“ bránia „nerozumné city“, vnímanie reality môže byť v afekte skreslené, pamäť je v mnohých ohľadoch emóciami ovplyvnená. Takéto a ďalšie podobné príklady ľudového vyjadrenia vzťahu emócie a kognitívnych funkcií nájdeme v mnohých knihách, filmoch, ale aj v bežnej hovorovej reči. Ale je to naozaj tak? Je vplyv emócií na procesy kognície naozaj takto neziaduci? Na čo nám potom emócie slúžia? Nebolo by pre prežitie vhodnejšie a účelnejšie používať iba „chladnú hlavu“? Potrebujeme vôbec emócie? V skutočnosti sú podľa prevládajúceho vedeckého názoru v súčasnosti emócie naopak považované za funkčné, účelné a majúce vysokú evolučnú hodnotu pre sociálne cicavce, medzi ktorých patrí aj človek [1].

## 2 Emócie

*„Žiaden aspekt mentálneho sveta nie pre kvalitu a zmysel ľudského života tak dôležitý, ako emócie. Práve emócie robia život tým, že stojí za to žiť ho (a niekedy ukončiť ho).“*

R.de Sousa

Aby sme sa mohli pokúsiť odpovedať na otázky nastolené na konci predchádzajúcej state, je vhodné najskôr si zadať definovať a vymedziť základný pojem – emóciu.

### 2.1 Definícia emócie

Definovať niečo tak základné a intuitívne jasné, ako sú emócie, je paradoxne veľmi komplikované a pri rozličných prístupoch sa stretávame s mnohými rozdielnymi pokusmi o definovanie pojmu emócia. Podľa učebnice fyziológie predstavuje emócia silný nutkavý stav pudového cítenia, ktoré sa vzťahuje na určitú cieľovú činnosť a navonok sa prejavuje apetitívnym alebo averzívny správaním [2]. V psychológii je emócia považovaná za odpoveď na podnet, ktorá zahŕňa charakteristické fyziologické zmeny, akými sú zrýchlenie pulzu, nárast telesnej teploty, vyššia, alebo nižšia aktivita niektorých žliaz, zmena rýchlosti dýchania, a ktoré vo svojej podstate vedú k motivácii jedinca k ďalšiemu konaniu. Anglický Websterov slovník definuje emóciu ako afektívny stav vedomia, v priebehu ktorého sa prežívajú stavy ako sú radosť, smútok, strach, nenávisť a pod., čím sa odlišuje od kognitívnych alebo vôľových stavov vedomia [3].

### 2.2 Diskrétné emočné kategórie

Emócie sú tradične rozdeľované do niekoľkých základných kategórií (hnev, strach, smútok, radosť, láska). Jednotlivé kategórie sú vyvolávané rozdielnymi podnetmi, majú rozdielne prejavy a sú charakterizované rozdielnymi subjektívnymi pocitmi pri ich prežívaní. Takéto delenie prirodzene vyplýva aj z každodennej ľudovej slovnej zásoby. To však môže byť zavádzajúce, nakoľko v rozličných jazykoch existujú síce podobné, ale nie totožné výrazy pre jednotlivé emócie a takisto je rozdielny aj počet pomenovaní, ktorý sa od jazyka k jazyku líši [4]. Paul Ekman navrhol kategorizáciu emócií podľa výrazu tváre ako univerzálneho emočného komunikačného nástroja [5]. Neskôr však túto ideu sám opustil z dôvodu existencie emócií bez tvárových prejavov a existencie rozdielných emócií zdieľajúcich rovnaké prejavy [6]. Ďalšie spôsoby delenia emócií sú založené na rozdielnej fyziologickej reakcii na základe

rozdielnej aktivácie autonómneho nervového systému [7], na základe behaviorálnej odpovede [10], prípadne na základe zapojenia rozdielných mozgových štruktúr [11]. Avšak zdá sa, že iba niektoré mozgové štruktúry sú takto špecificky previazané s konkrétnou emóciou. Takouto štruktúrou je bezpochyby amygdala – kľúčová štruktúra pre emóciu strachu. Funkčné zobrazovacie techniky potvrdili selektívnu aktiváciu amygdaly počas prezentácie vystrašených tvárí [12], [13]. Amygdala je tiež nevyhnutná pri podmieňovaní strachu (z angl. „fear conditioning“) [14]. Ako ďalšia špecifická emočná štruktúra sa javí inzula v súvislosti so znechutením (hnusom, z angl. „disgust“). Vnímanie znechutenia vo výraze tváre je spojené s aktiváciou prednej inzulárnej kôry [15]. Elektrická stimulácia inzuly vyvoláva pocity nauzey a neprijemných chutí [16].

### 2.3 Dimenzionálny model emócií

Na rozdiel od diskkrétnej kategorizácie emócií, motivačná hypotéza navrhuje organizáciu emočných reakcií pozdĺž dvoch motivačných systémov – apetitívneho a averzívneho, ktoré sa evolučne vyvinuli na zabezpečenie adekvátnej reakcie v situáciách, ktoré sú buď prospešné pre fyzické prežitie alebo ho naopak ohrozujú [17]. V súlade s týmto modelom, sa charakterizácia jednotlivých emócií pohybuje v rámci kontinua dvoch základných parametrov. Jednu dimenziu predstavuje valencia alebo miera potešenia, ktoré daná emócia prináša. Druhý rozmer súvisí s mierou aktivácie alebo nabudenia jedinca. Oba parametre korelujú s reakciami autonómneho nervového systému. Zmeny elektrodermálnej aktivity (vodivosti resp. odporu kože) sú modulované nabudením – počas sledovania emočných podnetov je elektrodermálna aktivita vyššia (zvýši sa vodivosť) v porovnaní so sledovaním neutrálnych podnetov [18]. Valencia podnetu moduluje úľakový (angl. „startle reflex“) reflex – čím nepríjemnejší emočný stimul, tým je žmurkací reflex frekventnejší [19]. Zdá sa, že určité mozgové štruktúry majú špecifický vzťah k valencii podnetu. Počas negatívnych emócií nastáva zvýšená aktivácia amygdaly, inzuly, pravej ventro-laterálnej prefrontálnej kôry a ľavej orbito-frontálnej mozgovej kôry, počas pozitívnych emócií je aktívnejšia ľavá dorzo-laterálna a ventro-mediálna prefrontálna kôra a pravá orbito-frontálna kôra [20]. Canli navrhol teóriu lateralizácie emočnej valencie: pozitívne emócie sú spojené so zvýšenou aktivitou ľavej hemisféry, negatívne emócie s vyššou aktivitou pravej hemisféry [21]. Ďalšie mozgové oblasti súvisia s intenzitou emócií: amygdala, inzula, senzorická kôra (v závislosti od modality), talamus, dorzo-mediálna prefrontálna kôra [22].

## 2.4 Separácia emočných podjednotiek

Ďalší prístup k nazeraniu na emócie predstavuje pohľad Russella, ktorý navrhuje rozlišovať jednotlivé podzložky tradičného termínu emócia. Predstavuje termín „core affect“ ako základný subjektívny emočný stav prítomný bez ohľadu na vonkajší objekt (nálada) a tzv. prototypickú emočnú epizódu – komplexná emočná udalosť vyvolaná určitým objektom v určitom čase, pričom obe sa navzájom neustále ovplyvňujú [23]. Napriek tomu, že neposkytuje vyčerpávajúci výpočet emočných zložiek, Russellova práca poukazuje na vhodnosť rozlišovania emočných podkategórií.

## 2.5 Kognitívny prístup k emóciám

V ostatných pár desaťročiach sa dostáva do popredia tzv. kognitívna teória emócií, resp. teória kognitívneho zhodnotenia (z angl. „appraisal“, zhodnotenie, posúdenie, ohodnotenie). Podobné myšlienky však môžeme nájsť už u Aristotela či Spinozu. Základnou myšlienkou teórie zhodnotenia je, že každá jednotlivá emócia je založená na vedomom alebo nevedomom vyhodnotení situácie v zmysle individuálnych cieľov (postojov, motívácií) jedinca [24]. Zhodnotenie teda smeruje k poznaniu toho, čo daný podnet alebo situácia znamená pre jednotlivca [25]. Toto vyhodnotenie (zhodnotenie) vyvolá určitý stupeň pripravenosti na akciu, ktorá nadobudne prioritu v čase pozdĺž nášho napredovania životom [26]. Na vyhodnotenie emočného podnetu a na následný emočný stav majú vplyv aj predchádzajúce vedomosti a očakávania, ale aj dodatočne získané poznatky a predstavy a cielené vedomé prehodnotenie významu emočného podnetu môže ovplyvniť emočný stav. Ellsworth a Smith zistili, že jednotlivé typy vyvolaných pozitívnych emócií (nádej, dôvera, záujem, hravosť, láska) sa od seba navzájom odlišovali odlišnými kognitívnymi vyhodnoteniami [27]. Pokusy Rosemana a Evdokasa ukázali, že rozdielne vyhodnotenia vyvolajú rozdielne emócie [28]. Podobne, Kuppens a spol. zistili, že zhodnotenia môžu mať pre rozdielnych ľudí rozdielny význam. Pre niektorých jedincov je môže byť hnev vyvolaný frustráciou, pre iných je nevyhnutný pocit vedomej nespravodlivosti a pre ďalších je dôvodom ohrozenie svojbytnosti [29]. Selektívne zameranie, alebo odpútanie pozornosti ovplyvní odpoveď vyhodnocovacieho systému na emočný podnet. Princíp teórie kognitívneho zhodnotenia sa využíva aj pri liečbe psychiatrických porúch pomocou kognitívno-behaviorálnej terapie, ktorá vychádza z predpokladu, že základom niektorých psychiatrických diagnóz je nesprávna a nevhodná interpretácia prežívanej reality a pokúša sa o dodatočné prehodnocovanie významu určitých situácií, postojov, emócií a pod [30].

## 3 Vzťah kognitívnych a emočných procesov

Ako už bolo načrtnuté vyššie, nie len kognícia ovplyvňuje emócií (ich tvorbu, prežívanie, či vnímanie), ale aj emócie majú veľký vplyv na kognitívne procesy [31]. Emočné naladenie vedie k prechodnej zmene výkonu v určitých špecifických kognitívnych funkciách. Obecne sa akceptoval názor, že „príjemné“ emócie majú vo vzťahu ku kognícii nešpecificky benefičný efekt, zatiaľ čo vplyv nepríjemných“ emócií je skôr negatívny [32]. Emócie predstavujú evolučný prostriedok umožňujúci prisúdiť udalostiam určitú váhu – určiť, do akej miery je udalosť pre organizmus žiaduca, alebo nežiaduca. Z tohto hľadiska by malo byť pre udalosti s vyššou pozitívnu alebo negatívnu hodnotou zabezpečené prednostné perцепčné spracovanie. Takýto vplyv emócií na pozornosť je zrejmy z viacerých štúdií. Napríklad, čas potrebný na detekciu objektu medzi ďalšími rušivými objektmi je výrazne kratší, v prípade, že cieľovým podnetom je emočný objekt [33]. Pri teste so selektívnym zameraním resp. odpútaním pozornosti dochádza k tzv. slepote z nepozornosti. Znamená to, že ďalší podnet, na ktorý nebola pozornosť zameraná, si dotýčný neuvedomí. Neplatí to však vtedy, ak je tento druhý podnet emočne významný [34]. Podobne, syndróm jednostranného opomínania (angl. „neglect“), pri ktorom je z dôvodu poškodenia pravostrannej dolnej parietálnej kôry znemožnené alebo obmedzené vnímanie podnetov prezentovaných do kontralaterálnej polovice zorného poľa, je obmedzený na podnety bez emočného obsahu. V prípade emočných podnetov (obrázky pavúkov, tváre s výrazom hnevu, radosti) je tento deficit značne zoslabený [35].

Veľmi dobre známy je vplyv emócií na pamäť. Pri bezprostrednom spomínaní si na obrázky s rôznym emočným aj neutrálnym obsahom si vyšetované osoby lepšie vybavujú obrázky s vysokým emočným nábojom [36]. Erk skúmal proces kódovania neintegrálnych emočných informácií. Prezentácii napísaného slova predchádzali pozitívne, negatívne a neutrálne obrázky z IAPS (z angl. „International Affective Picture System“ – medzinárodný systém afektívnych obrázkov). Vybavovanie si z pamäte bolo lepšie pri slovách, ktorým predchádzal pozitívny obrázok [37]. Emočný stav selektívne ovplyvňuje pracovnú pamäť pre verbálne a neverbálne informácie. Nepríjemný emočný stav (strach, obava a pod.) pôsobí negatívne na výkon pracovnej pamäte pre verbálne podnety, ale pozitívne na výkon pracovnej pamäte pre priestorovú orientáciu. Naopak, príjemný emočný stav, napr. očarenie, má opačný efekt, zlepšuje verbálnu pracovnú pamäť a zhoršuje priestorovú pracovnú pamäť. Podnety s vysokým nabudením (angl. „arousal“) vykazujú

mnemonicú výhodu v porovnaní s podnetmi s nízkym nábudením vďaka zlepšenému kódovaniu do pamäte a zlepšenej konsolidácii. Zdá sa, že táto výhoda je sprostredkovaná amygdalou, hipokampom a kôrovými štruktúrami zapojenými do neuronovej siete podporujúcej epizodickú pamäť [38]. Dolan skúmal pomocou pozitronovej emisnej tomografie (PET) vybavovanie si emočných a neemočných obrázkov, ktorému predchádzalo ich štúdium. Podľa jeho výsledkov zodpovedá samotné spomínanie si na emočné položky aktivácii ľavej amygdaly. Druhý neuronálny okruh je spojený s aktivitou predného temporálneho pólu a odráža psychické nastavenie týkajúce sa vyvolávania z emočnej pamäte [31].

### 3.1 Poruchy emócií a kognície

V rozpore s tradičným pohľadom na protiklad rozumu a emócií sa zdá, že za určitých okolností môžu emočné procesy ovplyvniť uvažovanie a rozhodovanie výhodným spôsobom. Tento vplyv je viditeľný napríklad pri poškodení ventro-mediálnej prefrontálnej kôry. Poškodenie v tejto oblasti síce nie je spojené s poškodením kognitívnych funkcií, ale vedie k tvorbe pre pacienta nevýhodných rozhodnutí [39]. Tento efekt, známy aj v súvislosti s hypotézou somatických markerov sa vysvetľuje neschopnosťou „precítiť“ možné varianty budúceho riešenia.

Narušená rovnováha emočného a kognitívneho spracovania je spoločným menovateľom celého spektra psychiatrických diagnóz (napr. post-traumatická stresová porucha, obsedantne-kompulzívna porucha, schizofrénia a podobne). Vzťah kognitívnych procesov a úzkostných porúch sa zdá byť veľmi silný. Viaceré teórie predpokladajú priamu súvislosť medzi neadekvátnymi kognitívnymi mechanizmami a rozvojom úzkostných porúch. Napríklad pri panickej poruche sa za jednu z hypotéz vzniku a rozvoja ochorenia považuje katastrofické spracovanie somatických pocitov a vnemov [40]. Ešte evidentnejší je tento vzťah pri post-traumatickej stresovej poruche, pri ktorej dochádza k opakovanému, neúnosne silnému znovu-prežívaní traumatickej udalosti [41]. Napriek tomu, že pamäťové procesy a úloha pamäťových mechanizmov pri vzniku a udržaní psychopatológie nie sú dostatočne objasnené, kognitívno-behaviorálna terapia sa zameriava práve na tento aspekt poruchy [42]. Vďaka pomerne úspešnému použitiu psychoanalytického prístupu pri liečení depresie [43], úzkostných porúch [44], porúch osobnosti [45], závislosti [46], schizofrénie [47] či bipolárnej poruchy [48] sa dá predpokladať významnú úlohu, ktorú pamäťové mechanizmy v psychopatológii týchto porúch zohrávajú.

### 3.2 Kognitívne posuny pri úzkostných poruchách

Mnohé štúdie sa zaoberajú tzv. tendenčnou pozornosťou (z angl. „attentional bias“) pri úzkosti, teda zvýšenou afinitou pozornosti úzkostných osôb k podnetom súvisiacim s ohrozením [49]. Okrem tendenčnej pozornosti pre emočné podnety súvisiace s úzkosťou bola popísaná aj podobná predpojatost' pre pamäť (z angl. „memory bias“) u úzkostných jedincov, v podobe lepšej pamäte pre informácie súvisiace s ohrozením. Napríklad vo forme lepšej pamäte na negatívne emočné slová u jedincov s úzkostnými poruchami [50], [51], [52] a jedincov s vyšším skóre dlhodobej úzkostnosti [53]. Mitte túto predpojatost' pamäte nepovažuje len za dôsledok predpojatej pozornosti, ale za ďalší, aditívny komponent úzkosti. Do svojej meta-analýzy zahrnul štúdie zamerané na sledovanie tendenčnej pamäte pri spomínaní si na emočné podnety u zdravých osôb líšiacich sa v stupni úzkostlivosti, ako aj pacientov s úzkostnými poruchami a porovnával ich so zdravou kontrolnou skupinou. Mitteho štúdia poukazuje na kontinuálne rozloženie úzkostnosti ako charakterovej črty v populácii: od zdravých, málo úzkostných jedincov, cez zvýšenú úzkostnosť až po patologickú úzkostnosť pri úzkostných poruchách, bez prítomnosti ostrej deliacej hranice vymedzujúcej jednotlivé skupiny. Z výsledkov spomínanej štúdie ďalej vyplýva, že pacienti s úzkostnou poruchou sa líšia od zdravej populácie rovnakým spôsobom ako vysoko úzkostní zdraví jedinci od jedincov s nízkou úzkosťou [56].

## 4 Pamäťové mechanizmy a úzkostnosť

Úzkostným a afektívnym poruchám sa dlhodobo venuje aj naše pracovisko (Laboratórium kognitívnej neurovedy, Ústav normálnej a patologickej fyziológie, SAV) [57][58][59][60][61][62]. V nedávnej štúdií sme sa zamerali na bližšie preskúmanie vzťahu medzi úzkostnosťou a explicitnou pamäťou [63]. Za týmto účelom sme analyzovali spomienku na emočne podfarbenú situáciu vyvolávajúcu úzkosť u ľudí líšiacich sa svojou tendenciou reagovať s úzkosťou. Za objekt výskumu sme si zvolili zdravých, nediagnostikovaných jedincov, líšiacich sa v osobnostnej črte – úzkostnosti. Jedným z dôvodov bolo kontinuálne rozloženie úzkostnosti od zdravých, málo úzkostných jedincov, cez zvýšenú úzkostnosť až po patologickú úzkostnosť pri úzkostných poruchách, na ktoré poukázala práca Mitteho. Ďalším dôvodom, prečo sme si vybrali osoby s ešte nerozvinutou poruchou bola snaha zamerať sa na menej vyjadrenú dispozíciu, pri ktorej by mohli byť zdroje prípadnej odlišnosti viditeľnejšie ako v prípade už

rozvinutej úzkostnej poruchy. Tímto jedincom sme prezentovali dve filmové ukážky. Jednu emočne nabudzujúcu, vyvolávajúcu strach a úzkosť a druhú emočne relatívne neutrálnu, nevzbudzujúcu špecifické emócie ani nabudenie. Po dvoch týždňoch od prezentácie sme analyzovali ich spomienky na dané ukážky. Počas experimentálnych situácií – pôsobenia podnetu a jeho ukladania do pamäte, aj počas jeho vyvolávania si z pamäte – sme zaznamenávali jednak subjektívne, vedomé reakcie prostredníctvom dotazníkov a individuálnej výpovede a jednak autonómne fyziologické reakcie a tiež bioelektrickú aktivitu mozgu.

Zistili sme, že skupiny sa navzájom líšili v emočnej zložke spomienky. Úzkostnejší jedinci opisovali spomienku na negatívnu filmovú ukážku väčším počtom slov s emočným nábojom. Počet emočných spomienok bol štatisticky signifikantne asociovaný so subjektívnym hodnotením emočného stavu pri spomínaní si, ale nie pri sledovaní filmovej ukážky. To môže znamenať, že úzkosť interaguje skôr s následným vybavovaním z pamäti ako s ukladáním do nej. Tento nález je v súlade s predchádzajúcimi závermi Kroesa a spol. [64], ktorí uvádzajú, že podanie propranololu, beta-adrenergného antagonistu, krátko po tom, čo je subjekt požiadaný zapamätať si zoznam slov, zabráni uľahčenému spomínaniu si na emočné slová, ale neovplyvní spomínanie si na neutrálne slová. Tieto nálezy vyzdvihujú úlohu pamäťových procesov pôsobiacich v časovo neskorších fázach, po ukladaní do pamäte. U úzkostných osôb by mohlo nesprávne neskoršie spracovanie negatívnych emočných informácií uľahčovať tvorbu neadekvátnych kognitívnych schém [65].

Autonómne reakcie u osôb s vyššou mierou úzkostnosti neboli v súlade s prežívanými emóciami. V skupine nízko úzkostných jedincov bola intenzita emócie prežívanej počas spomínania si (druhé sedenie) na hororovú ukážku hodnotená vyššie jedincami, ktorí dosahovali vyššie hodnoty elektrodermálnej aktivity počas predchádzajúceho sledovania tohto filmu (prvé sedenie). Hodnoty subjektívneho pocitu nabudenia v čase spomínania si na emočne nabitú udalosť v tejto skupine teda odpovedali hodnotám vodivosti kože, ktoré súvisia s fyziologickým stavom nabudenia. V skupine s vysokou úzkosťou sme však tento vzťah nepozorovali. Subjektívne hodnotenie intenzity emócie vysoko úzkostných jedincov počas spomínania si na horor nesúviselo s hodnotami EDA počas sledovania ukážky. Rozpor medzi subjektívnou percepciou telesného stavu a objektívnym fyziologickým stavom by mohol predstavovať významnú zložku vytvárania maladaptívnych kognitívnych schém vedúcich k rozvoju a udržaniu úzkostných porúch [66].

Podobné nálezy priniesla aj analýza elektroencefalografickej (EEG) aktivity. Počas sledovania

emočne nabudzujúcej – hororovej ukážky odpovedal vzorec EEG aktivity u jedincov s nízkym skóre úzkostnosti stavu celkového nabudenia oproti relatívne relaxovanejšiemu stavu počas sledovania neutrálnej filmovej ukážky. V skupine s vyšším skóre úzkostnosti, sme takýto rozdiel nenašli. Naopak, miestami paradoxne došlo akoby k nižšiemu nabudeniu ako pri emočne neutrálnom filme. Je pravdepodobné, že v elektroencefalografickej aktivite vysoko úzkostnej skupiny sa prejavil podobný fenomén, aký sme pozorovali už pri hodnotení ich elektrodermálnej aktivity, a sice zdanlivo nižšiu reakciu na emočnú – úzkosť-vyvolávajúcu situáciu v porovnaní s nízko úzkostnou skupinou.

Naše výsledky poukazujú na rozdielny mechanizmus zabezpečujúci vybavovanie si emočných informácií z pamäte, v závislosti od miery úzkostnosti a na disonanciu (narušený vzťah) medzi emočnou pamäťou na silné, úzkosť-vzbudzujúce podnety a fyziologickými reakciami na ne u ľudí s vyššou mierou úzkostnosti.

## PodĎakovanie

Tento príspevok vznikol za podpory grantovej agentúry VEGA v rámci grantového projektu číslo 2/0173/11.

## Literatúra

- [1] M. J. Power, T. Dalgleish: *Cognition and emotion: from order to disorder*, 2nd ed. Psychology Press, New York, 2008.
- [2] D. Ostatníková, J. Hájek: Emócie, správanie, motivácia. In: *Lekárska fyziológia* (K. Javorka et al., eds.), Osveta, Martin. 2009: 589.
- [3] F. Koukolík: Citový mozek. In: *Mozek a jeho duše*, 3.vydanie. Galén, Praha, 2008: 151-175
- [4] A. Wierzbicka: Defining emotion concepts. *Cognitive Science* 16 (1992) 539-581.
- [5] P. Ekman: Expression and the nature of emotion. In: *Approaches to emotion* (P. Ekman, K. Scherer, eds.), Hillsdale, NJ: Erlbaum, 1984: 319-343.
- [6] P. Ekman: Facial expression and emotion. *American Psychologist* 48 (1993) 384-392.
- [7] P. Ekman, R. W. Levenson, W. Friesen: Autonomic nervous system activity distinguishes among emotions. *Science* 221 (1983) 1208-1210

- [8] J. T. Cacioppo, D. J. Klein, G. C. Berntson, E. Hatfield: The psychophysiology of emotion. In: *Handbook of emotions* (M. Lewis, J. M. Haviland, eds.), Guilford Press, New York, 1993: 119-142.
- [9] R. B. Zajonc, D. N. McIntosh : Emotions research: Some promising questions and some questionable promises. *Psychological Science* 3 (1992) 70-74.
- [10] N. H. Frijda: *The emotions*. Cambridge: Cambridge University Press, 1986.
- [11] J. Panksepp: Toward a general psychobiological theory of emotions. *Behav. Brain Sci.* 5 (1982) 407–467.
- [12] J. S. Morris, C. D. Frith, D. I. Perrett, D. Rowland, A. W. Young, A. J. Calder, R. J. Dolan: A differential neural response in the human amygdala of fearful and happy facial expressions. *Nature* 383: (1996) 812–815.
- [13] H. C. Breiter, N. L. Etcoff, P. J. Whalen, W. A. Kennedy, S. L. Rauch, R. L. Buckner, M. M. Strauss, S. E. Hyman, B. R. Rosen: Response and habituation of the human amygdala during visual processing of facial emotion. *Neuron* 17 (1996) 875–887.
- [14] A. Bechara, D. Tranel, H. Damasio, R. Adolphs: Double dissociation of conditioning and declarative knowledge relative to the amygdala and hippocampus in humans. *Science* 269 (1995) 1115–1118.
- [15] M. L. Phillips, A. W. Young, C. Senior, M. Brammer, C. Andrew, A. J. Calder, E. T. Bullmore, D. I. Perrett, D. Rowland, S. C. Williams, J. A. Gray, A. S. David: A specific neural substrate for perceiving facial expressions of disgust. *Nature* 389 (1997) 495–498.
- [16] W. Penfield, M. E. Faulk: The insula: further observations of its function. *Brain* 78 (1955) 445–470.
- [17] M. M. Bradley, P. J. Lang: Measuring emotion: The Self-Assessment Mannequin and the Semantic Differential. *Journal of Behavior Therapy and Experimental Psychiatry* 25 (1994) 49-59.
- [18] P. J. Lang, M. M. Bradley, B. N. Cuthbert: Emotion, motivation, and anxiety: brain mechanisms and psychophysiology. *Biol Psychiatry* 44 (1998) 1248-63. Review.
- [19] M. M. Bradley, M. Codispoti, B. N. Cuthbert, P. J. Lang: Emotion and motivation I: defensive and appetitive reactions in picture processing. *Emotion* 1 (2001) 276-98.
- [20] F. Dolcos, K. S. LaBar, R. Cabeza: Dissociable effects of arousal and valence on prefrontal activity indexing emotional evaluation and subsequent memory: an event-related potential study. *Neuroimage* 23 (2004) 64–74.
- [21] T. Canli, J. E. Desmond Z. Zhao, G. Glover, J. D. Gabrieli: Hemispheric asymmetry for emotional stimuli detected with fMRI. *NeuroReport* 9 (1998) 3233 –3239.
- [22] D. Sabatinelli, M. M. Bradley, J. R. Fitzsimmons, P. J. Lang: Parallel amygdala and inferotemporal activation reflect emotional intensity and fear relevance. *Neuroimage* 24 (2005) 1265–1270
- [23] J. A. Russell, L. F. Barrett: Core affect, prototypical emotional episodes, and other things called emotion: dissecting the elephant. *J Pers Soc Psychol.* 76(5) (1999) 805-19. Review.
- [24] N. H. Frijda: The place of appraisal in emotion. *Cognition and Emotion* 7 (1993) 357-387
- [25] P. M. Neidenthal, S. Kranth-Gruber, F. Ric: *Psychology of Emotions: Interpersonal, Experiential, and Cognitive Approach*. Psychology Press, New York, 2006.
- [26] N. H. Frijda: Emotion experience. *Cognition and Emotion* 19 (2005) 473-497
- [27] P. C. Ellsworth, C. A. Smith: Shades of joy: Patterns of appraisal differentiating pleasant emotions. *Cognition and Emotion* 2 (1988) 301-331
- [28] I. Roseman, A. Evdokas: Appraisals cause experienced emotions: Experimental evidence. *Cognition and Emotion* 18 (2004) 1-28
- [29] P. Kuppens, I. Van Mechelen, D. J. M. Smits, P. De Boeck, E. Caulemans: Individual differences in patterns of appraisal and anger experience. *Cognition and Emotion* 21 (2007) 689-713.



- [30] J. Vyskocilova, J. Prasko: Emotional processing strategies in cognitive behavioral therapy. *Activitas Nervosa Superior Rediviva* 54(4) (2012) 150-158.
- [31] R. J. Dolan: Emotion, cognition and behavior. *Science* 298 (2002) 1191-1194.
- [32] P. M. Neidenthal, S. Kranth-Gruber, F. Ric: *Psychology of Emotions: Interpersonal, Experiential, and Cognitive Approach*. Psychology Press, New York, 2006.
- [33] A. Öhman, A. Flykt, F. Esteves: Emotion drives attention: detecting a snake in the grass. *J. Exp. Psychol.: Gen.* 127 (2001) 69–82.
- [34] A. K. Anderson, E. A. Phelps: Lesions of the human amygdala impair enhanced perception of emotionally salient events. *Nature* 411 (2001) 305 – 309.
- [35] P. Vuilleumier, J. L. Armony, J. Driver, R. J. Dolan: Effects of attention and emotion on face processing in the human brain: An event-related fMRI study. *Neuron* 30 (2001) 829-841.
- [36] M. M. Bradley, M. K. Greenwald, M. C. Petry, P. J. Lang: Remembering pictures: pleasure and arousal in memory. *J Exp Psychol Learn Mem Cogn.* 18(2) (1992) 379- 390.
- [37] S. Erk, M. Kiefer, J. Grothe, A. P. Wunderlich, M. Spitzer, H. Walter: Emotional context modulates subsequent memory effect. *Neuroimage* 18(2) (2003) 439-447.
- [38] S. B. Hamann, T. D. Ely, S. T. Grafton, C. D. Kilts: Amygdala activity related to enhanced memory for pleasant and aversive stimuli. *Nature Neuroscience* 2 (1999) 289-293.
- [39] A. Bechara, D. Tranel, H. Damasio: Characterization of the decision-making deficit of patients with ventromedial prefrontal cortex lesions. *Brain.* 123 (2000) 2189-202.
- [40] R. J. McNally: Automaticity and the anxiety disorders. *Behav Res Ther* 33(7) (1995) 747-754.
- [41] A. Ehlers, A. Hackmann, T. Michael: Intrusive re-experiencing in posttraumatic stress disorder: phenomenology, theory, and therapy. *Memory.* 12(4) (2004) 403-15. Review.
- [42] J. Prasko, A. Grambal, D. Kamaradova, D. Jelenova: Imagery rescripting of traumatic or distressing stories from childhood. *Activitas Nervosa Superior Rediviva* 54(3) (2012) 113–120.
- [43] A. T. Beck: The current state of cognitive therapy: A 40-year retrospective. *Archives of General Psychiatry* 62 (2005) 953-959.
- [44] A. T. Beck, G. Emery: *Anxiety disorders and phobias: a cognitive perspective*. Basic Books, New York, 1985.
- [45] A. T. Beck, A. Freeman, D. D. Davis: *Cognitive therapy of personality disorders*, 2nd ed., Guildford Press, New York, 2004.
- [46] F. D. Wright, A. T. Beck, c. F. Neuman, B. S. Liese: Cognitive therapy of substance abuse: Theoretical rationale. *NIDA Research Monograph Series* 137 (1993) 123-146.
- [47] A. T. Beck, N. A. Rector: Cognitive approaches to schizophrenia: Theory and therapy. *Annual Review of Clinical Psychology* 1 (2005) 577-606.
- [48] C. F. Newman, R.L. Leahy, A. T. Beck, N. A. Reilly-Harrington, L. Gyulai: *Bipolar disorder: A cognitive therapy approach*. American Psychological Association, Washington DC, 2002.
- [49] Y. Bar-Haim, D. Lamy et al.: Threat-related attentional bias in anxious and nonanxious individuals: a meta-analytic study. *Psychol Bull* 133(1) (2007) 1-24.
- [50] M. Cloitre, M. R. Liebowitz: Memory bias in panic disorder: An investigation of the cognitive avoidance hypothesis. *Cognitive Therapy and Research* 15,(5) (1991) 371-386.
- [51] S. R. Vrana, A. Roodman, J. C. Beckham: Selective processing of traumarelevant words in posttraumatic stress disorder. *Journal of Anxiety Disorders* 9(6) (1995) 515-530.
- [52] B. Friedman, J. Thayer, T. Borkovec: Explicit memory bias for threat words in generalized anxiety disorder. *Behavior Therapy* 31(4) (2000) 745-756.
- [53] A. Mathews, K. Mogg, J. May, M. Eysenck: Implicit and explicit memory bias in anxiety. *Journal of Abnormal Psychology* 98( 3) (1989) 236-240.

- [54] J. Reidy, A. Richards: A memory bias for threat in high-trait anxiety. *Personality and Individual Differences* 23(4) (1997) 653- 663.
- [55] R. Russo, D. Whittuck, D. Roberson, K. Dutton, G. Georgiou, E. Fox: Moodcongruent free recall bias in anxious individuals is not a consequence of response bias. *Memory* 14(4) (2006) 393-399.
- [56] K. Mitte: Memory bias for threatening information in anxiety and anxiety disorders: a meta-analytic review. *Psychol Bull* 134(6) (2008) 886-911.
- [57] P. Kukumberg, P. Valkovič, P. Blažiček, A. Gúth, J. Martinková, V. Provazník, F. Jagla: Sweat: a potential marker of clinical activity in panic disorder. *Neuroendocrinology Letters* 30(3) (2009) 400-402.
- [58] P. Bob, M. Kukleta, F. Jagla: Traumatic stress, anxiety and epilepsy. *Studia Psychologica: international journal for research and theory in psychological sciences* 49(2) (2007) 127-133.
- [59] P. Bob, M. Šusta, A. Procházková-Večeřová, M. Kukleta, J. Pavlát, F. Jagla, J. Raboch: Limbic irritability and chaotic neural response during conflicting stroop task in the patients with unipolar depression. *Physiological Research* 55(suppl.1) (2006) 107-112.
- [60] P. Bob, M. Kukleta, I. Riečanský, M. Šusta, P. Kukumberg, F. Jagla: Chaotic EEG patterns during recall of stressful memory related to panic attack. *Physiological Research* 55(suppl.1) (2006) 113-119.
- [61] M. Jergelová, F. Jagla: Central and peripheral correlates of eye movements in selected mood disorders. *Neuroendocrinology Letters* 31(6) (2010) 731-737.
- [62] F. Jagla, V. Zikmund, I. Zucha: Eye movements and eye movement related brain potentials in patients with symptoms of depression. *Homeostasis-Praha* 40(3/4) (2000) 71-75.
- [63] B. Cimrová, I. Riečanský, F. Jagla: The role of trait anxiety in explicit memories of emotional experience. *Studia Psychologica : international journal for research and theory in psychological sciences* 53(3) (2011) 253-261.
- [64] M. C. W. Kroes, B. A. Strange, R. J. Dolan:  $\beta$ -adrenergic blockade during memory retrieval in humans evokes a sustained reduction of declarative emotional memory enhancement. *Journal of Neuroscience* 30 (2010) 3959- 3963.
- [65] A. Beck, D. Guth, R. A. Steer, R. Ball: Screening for major depression disorders in medical inpatients with the Beck Depression Inventory for Primary Care. *Behav Res Ther* 35(8) (1997) 785-791.
- [66] R. Hoehn-Saric: Anxiety and somatization. *International Congress Series* 1287 (2006) 368-372.

# Analýza efektu krátkodobej a dlhodobej motivácie pomocou fuzzy prístupu

Ladislav Clementis

Ústav aplikovanej informatiky

FIIT STU v Bratislave

Ilkovičova

842 16 Bratislava 4

Email: clementis@fiit.stuba.sk

## Abstrakt

Motivácia ako evolučný prostriedok slúži k naplneniu potenciálu jedinca v populácii. Je hybným motorom pre jedinca alebo skupinu pre napĺňanie potrieb, hodnôt a cieľov. Motivačné prostriedky môžeme klasifikovať podľa celkovej účinnosti, a tak isto podľa ich pôsobenia, resp. distribúcie efektívnosti v čase. Túto dvojrozmernú klasifikáciu môžeme výpočtovo opísať pomocou fuzzy prístupu. Motivačnú efektívnosť prostriedku je možné matematicky aproximovať štatistickými rozdeleniami, ako napríklad logaritmicko-normálne rozdelenie. Na základe fuzzy prístupu je možné nastaviť parametre štatistických kriviek tak, aby ich rozdelenia súhlasili s krivkami účinnosti motivačných prostriedkov, čo umožní ich ďalšie skúmanie.

## 1 Úvod

Motivácia ako taká je skúmaná psychológmi a sociológmi po dlhý čas. Z pohľadu robotiky ide o cieľ napĺňať podstatu bytia, ktorá je vštepená jedincovi stvoriteľom. Človeku ako jedincovi vytvorenému prírodou je motivácia vštepená pudom prežitia, čiže napĺňaním základných životných potrieb.

Zároveň z pohľadu evolučných algoritmov sa *fitness* jedinca meria aj na základe postavenia, alebo prestíže jedinca v populácii, čo je možné klasifikovať ako naplnenie sociálnych, finančných alebo iných hodnôt. Tieto hodnoty si vypestuje jedinec sám, na základe pozorovania svojho okolia. Poskytujú mu mieru pre porovnanie ostatných jedincov, ako aj pre zaradenie seba samého medzi ostatných.

Človek má evolúciou vštepený pud sebarealizácie, čo sa prejavuje napĺňaním daných cieľov. Tieto ciele môžu mať rôzny charakter. Môžu byť krátkodobé alebo dlhodobé. Môžu byť napĺňané priebežne vykonávaním akejsi činnosti alebo dosiahnutím cieľu činnosti.

Činnosti môžu mať jasne daný cieľ, alebo cieľ môže byť vágneho charakteru. Všeobecne však ide o napĺňanie potenciálu jedinca a jeho sebarealizáciu.

Skúmaním matematických vlastností motivácie ako takej môžeme odhaliť jej podstatu. Pomocou štatistických prostriedkov je možné tento fenomén opísať a tým pádom aj predvídať účinky motivačných prostriedkov na výkonnosť jedinca napríklad na pracovisku a pod.

Jedna z možností ako matematicky opísať účinky motivácie je pomocou logaritmicko-normálneho rozdelenia. Logaritmicko-normálne rozdelenie je parametrické, a tieto parametre je možné nastaviť tak, aby rozdelenie súhlasilo s opisom účinnosti motivácie v čase.

Je veľa spôsobov ako tieto parametre aproximovať, v tejto práci sa zaoberáme aproximáciou týchto parametrov pomocou klasifikácie postavenej na fuzzy prístupe.

## 2 Motivácia ako hybná sila jedinca alebo skupiny v populácii

Motivácia [4, 5, 7, 8] jedinca v rámci jeho osobných charakteristík, alebo v rámci skupiny a kooperácie môžeme deliť do rôznych kategórií, podľa pôvodu vzniku. Najzákladnejšie rozdelenie je:

1. potreby - pre každého človeka rovnaké
2. hodnoty - individuálne, vyvinuté počas života
3. cieľ - môže sa meniť a môže byť:

- krátkodobý
- dlhodobý

a tak isto môže byť:

- špecifický
- nejasný, vágny

Živočíšne potreby sú základom prežitia jedinca. Týkajú sa jedenia, pitia, bezpečnosti a podobne. Ak nie sú splnené základné potreby jedinca, jedinec uhynie.

Hodnoty si buduje jedinec na základe výchovy, vzdelania a svojho okolia. Hodnoty sú nad rámec základných

potrieb a udávajú štandard jedinca v populácii. Ide o hodnoty ako majetok, rodinný status, spoločenské postavenie, kariéra a podobne. Ak má jedinec splnené základné potreby, môže sa zamerať na kumuláciu ďalších hodnôt, ktoré ho odlišia od ostatných jedincov. Tieto hodnoty ovplyvňujú významnosť osoby v populácii, udávajú jej status.

Akumulácia hodnôt súvisí priamo s napĺňaním cieľov, je to spôsob ako hodnoty nadobúdať. Dosahovanie cieľov však má aj iný charakter, a to seberealizáciu jedinca. Seberealizácia je zároveň prostriedok napĺňania životnej úlohy jedinca. Je spojená s iniciatívou a je odmeňovaná prostriedkami motivácie.

Prostriedky motivácie môžu začínať od majetkových (ako finančné ohodnotenie výkonu), až po radosť s dobre vykonanej práce. Samozrejme treba brať v úvahu iné ako pracovné ohodnotenie, ako napríklad založenie rodiny a podobne.

Motivácia v rámci skupiny jedincov a kooperácie nadobúda zmysel pre vzájomnú pomoc. Jednotlivci v skupine sú vzájomne posúvaný vpred a každý svojim príspevkom pomáha ostatným. Individuálny záujem jednotlivcom je zakorenený v zisku zo spolupráce. Cieľom skupiny je kumulovať väčší zisk ako iné skupiny. Pri tejto úlohe by jedinec samotný súperiaci so skupinami o presadenie sa nemal šancu.

Jedinci v skupine pomáhajú naplniť cieľ, či už krátkodobý alebo dlhodobý, vážny alebo špecifický. V prírode sa formujú tupy dravcov pre kooperáciu pri love koristi. Korisť sa obdobne zgrupuje tak, aby jednotlivci boli menej ohrození. Tieto zoskupenia popisujú matematické modely ako multiagentové systémy a podobne. Sú tvorené na základe motivácie vzájomnej pomoci a kumulovanej sily zoskupenia, čo prináša jedincom výhody.

## 2.1 Hlavné rozdiely medzi motivačnými zdrojmi

Z hľadiska dávkovania motivačného prostriedku môžeme rozlišovať:

- diskretný motivačný zdroj
- spojitý motivačný zdroj

Diskretný motivačný zdroj môže byť jednorazová výkonnostná odmena v ľubovoľnej forme, pravidelný prídel finančných zdrojov a podobne. Spojité motivačné prostriedky súvisia s potenciálom prostredia ako napríklad kolektív, istoty a stále prítomné zabezpečenia ktorých prítomnosť samotná je motivujúca.

Je dôležité poznamenať, že diskretné dávkované motivačné prostriedky majú spojitú účinnosť pôsobenia.

Z hľadiska početnosti alebo výskytu môžeme motivačné prostriedky rozdeliť ako:

- jednorazový motivačný zdroj
- opakujúci sa motivačný zdroj

Jednorazový motivačný zdroj môže byť napríklad bonusové finančné ohodnotenie zamestnanca, napríklad odmena za dobre vykonaný ukončený projekt. Dlhodobý motivačný zdroj môže byť pravidelný plat.

Keď sa zamyslíme nad rámec finančného ohodnotenia, jednorazový motivačný zdroj môže byť aj uznanie od nadriadeného alebo kolegov. Dlhodobým motivačným zdrojom je potom celkový prístup nadriadeného a kolektívu k jedincom.

Mohli by sme uviesť mnohé ďalšie kategorizácie motivačných zdrojov, ale pre naše potreby nám vystačia tieto. V tejto práci sa primárne zaoberáme jednorazovými diskretnými motivačnými zdrojmi.

## 3 Funkcia hustoty logaritnicko-normálneho rozdelenia pre popis efektu jednorazového diskretného motivačného zdroja

*Logaritnicko-normálne rozdelenie*  $LN(\mu, \sigma)$  (tiež bežne označované ako *log-normálne rozdelenie*) s parametrami  $\mu$  a sigma  $\sigma$  je spojitý rozdelenie pravdepodobnosti náhodnej premennej  $X$  také, že náhodná veličina  $\ln(X)$  má *normálne rozdelenie* so *strednou hodnotou*  $\mu$  a *smerodajnou odchýlkou*  $\sigma$ .

Logaritnicko-normálne rozdelenie sa v prírode hojne vyskytuje. Služi taktiež ako prostriedok pre vedné odbory pracujúce s pravdepodobnosťou a štatistikou [2]. Príklady sú uvedené v odseku 3.1. Matematický základ logaritnicko-normálneho rozdelenia je v odseku 3.2.

### 3.1 Výskyt logaritnicko-normálneho rozdelenia v prírode a vedných odboroch

Logaritnicko-normálne rozdelenie sa v prírode, prírodných a iných vedách vyskytuje napríklad:

- V biológii [3], a to ako rozdelenie veľkostí živých tkanív jedincov v populácii (dĺžka tela, povrch tela, hmotnosť, dĺžka ochlpenia resp. vlasov, pazúrov resp. nechťov, zubov, krvný tlak dospelého muža alebo ženy [6]...).
- V meteorológii, hydrológii a vodnom hospodárstve pre analýzu denných, mesačných a ročných zrážok, výšok vodných tokov a pod.
- V ekonómii, finančnictve [1] a bankovníctve, je známe že finančné príjmy jedincov v populácii sú rozdelené pomocou logaritnicko-normálneho rozdelenia.

- V demografii, veľkosti miest sú rozdelené pomocou logaritmicko-normálneho rozdelenia.
- Vo fyzike, chémii a materiálových technológiách, napríklad koeficienty trenia a opotrebovania majú logaritmicko-normálne rozdelenia.

### 3.2 Matematický základ logaritmicko-normálneho rozdelenia

Funkcia hustoty  $f_X$  logaritmicko-normálneho rozdelenia má exponenciálny tvar (1) a je parametrizovaná pomocou  $\mu$  a  $\sigma$ .

Príklady kriviek funkcie hustoty logaritmicko-normálneho rozdelenia pre rôzne hodnoty parametrov sú vyobrazené na Obr. 1, Obr. 2, Obr. 3 a Obr. ??.

$$f_X(x; \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln(x)-\mu)^2}{2\sigma^2}}, x > 0 \quad (1)$$

Distribučná funkcia logaritmicko-normálneho rozdelenia  $F_X$  je definovaná rovnicou (2) a (3).

$$F_X(x; \mu, \sigma) = \frac{1}{2} \operatorname{erfc} \left[ -\frac{\ln x - \mu}{\sigma\sqrt{2}} \right] \quad (2)$$

$$F_X(x; \mu, \sigma) = \Phi \left( \frac{\ln x - \mu}{\sigma} \right) \quad (3)$$

Funkcia  $\operatorname{erfc}(x)$  (4) je funkcia doplnku chybovej funkcie  $\operatorname{erf}(x)$  (5) (Gauss error function).

$$\operatorname{erfc}(x) = 1 - \operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-t^2} dt \quad (4)$$

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \quad (5)$$

Funkcia  $\Phi$  je distribučná funkcia normovaného normálneho rozdelenia (*standard normal distribution*), čiže normálne rozdelenie s parametrami  $\mu = 0$  a  $\sigma = 1$  a funkciou hustoty  $\phi$  (6).

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \quad (6)$$

Tieto vzorce popisujúce logaritmicko-normálne rozdelenie sa využívajú v štatistike, hydrometeorologických predpovediach a iných spomínaných v odseku 3.1.

## 4 Fuzzy prístup pre odhad parametrov krivky funkcie efektu jednorazového diskretného motivačného zdroja

Zadefinujeme si zjednodušene množiny  $A$  (7) a  $B$  (8) pre kategorizáciu konkrétnych vlastností motivačných zdrojov.

$$A = \{kratka, kratsia, dlhsia, dlha\} \quad (7)$$

Obdobne:

$$B = \{mala, mensia, vacsia, velka\} \quad (8)$$

Supermnožina  $A$  označuje kategorizáciu motivačného zdroja na základe trvania pôsobenia, resp. doznievania efektu motivačného zdroja, pričom čo sa týka trvania,  $kratka < kratsia < dlhsia < dlha$ . Príslušnosť motivačného zdroja do jednej z množín nachádzajúcich sa v  $A$  ovplyvňuje parameter  $\mu$ .

Supermnožina  $B$  označuje kategorizáciu motivačného zdroja na základe sily pôsobenia, resp. miera efektu motivačného zdroja, pričom čo sa týka sily zdroja,  $mala < mensia < vacsia < velka$ . Príslušnosť motivačného zdroja do jednej z množín nachádzajúcich sa v  $B$  ovplyvňuje parameter  $\sigma$ .

Pravdepodobnosti priradenia motivačných zdrojov do jednotlivých množín je netriviálna úloha a parametre logaritmicko-normálneho rozdelenia je dôležité vhodne zvoliť, prípadne aproximovať evolučnými prostriedkami umelej inteligencie.

V tejto práci ukážeme, ako nastavovanie parametrov mení povahu krivky funkcie hustoty logaritmicko-normálneho rozdelenia.

Zadefinujeme si prvok  $c$  (9) ktorý patrí do množiny  $C$  vytvorenej reláciou  $A \times B$ .

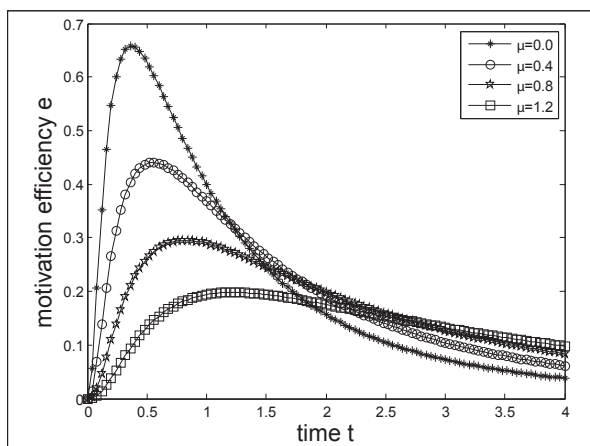
$$c \in C = A \times B \quad (9)$$

Nami odhadnuté hodnoty parametrov  $\mu$  a  $\sigma$  sú v tabuľke 1. Je náročné posudzovať motiváciu v praxi pomocou fuzzy prístupu, nie však nereálne. V praxi by hodnoty boli ľahko porovnateľné, ale parametre ťažšie odhadnuteľné.

Hodnoty v tabuľke 1 boli odhadnuté ilustračne, pre aplikáciu v reálnom svete by odhad bol netriviálnou záležitosťou.

**Tabuľka 1:** Tabuľka zobrazujúca odhadnuté hodnoty  $\mu$  a  $\sigma$  pre všetky  $c \in C = A \times B$ , z ktorej sme čerpali hodnoty pri tvorbe grafov na obrázkoch Obr. 1, Obr. 2, Obr. 3 a Obr. ??.

$\mu \backslash \sigma$	malá $\sigma = 1.0$	menšia $\sigma = 0.5$	väčšia $\sigma = 0.25$	veľká $\sigma = 0.125$
krátka $\mu = 0.0$	0.0 \ 1.0	0.0 \ 0.5	0.0 \ 0.25	0.0 \ 0.125
kratšia $\mu = 0.4$	0.4 \ 1.0	0.4 \ 0.5	0.4 \ 0.25	0.4 \ 0.125
dlhšia $\mu = 0.8$	0.8 \ 1.0	0.8 \ 0.5	0.8 \ 0.25	0.8 \ 0.125
dlhá $\mu = 1.2$	1.2 \ 1.0	1.2 \ 0.5	1.2 \ 0.25	1.2 \ 0.125



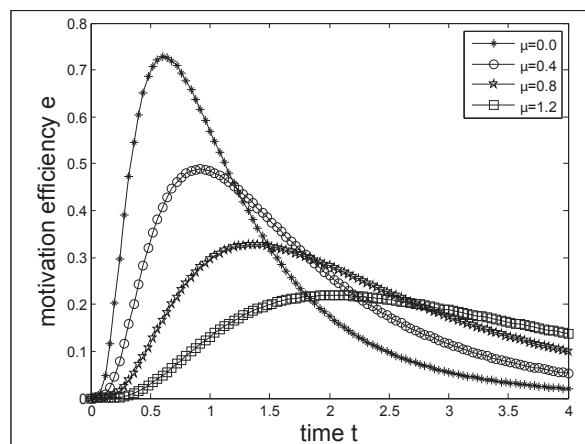
**Obrázek 1:** Graf funkcie hustoty logaritmicke-normálneho rozdelenia pre zafixovanú sigma = 1.0 a  $\mu \in \{0.0, 0.4, 0.8, 1.2\}$ .

#### 4.1 Krivky funkcie hustoty logaritmicke-normálneho rozdelenia pre zafixovanú $\sigma = 1.0$

Všimnime si v tabuľke 1 prvý stĺpec s hodnotami (pre  $b = mala$  a na  $a \in A$  nezáleží, teda  $\sigma = 1.0$  a postupne  $\mu \in \{0.0, 0.4, 0.8, 1.2\}$ ).

Graf kriviek logaritmicke-normálneho rozdelenia pre zafixovanú  $\sigma = 1.0$  je vyobrazený na Obr. 1.

Rozoberme graf na Obr. 1. Pre  $\mu = 0.0$  bol okamžitý nástup efektu motivácie viditeľne vysoký. Okolo času  $t = 2$  ale úroveň efektu motivačného prostriedku klesol na úroveň približne rovnakú, ako pre  $\mu \in \{0.4, 0.8, 1.2\}$ . Tesne po nástupe bol ale efekt viac ako dvojnásobný oproti  $\mu = 0.8$  a viac ako trojnásobný oproti  $\mu = 1.2$ . Parameter  $\mu$  nám pre  $\sigma = 1.0$  výrazne ovplyvnil mieru efektu motivačného prostriedku po začiatku jeho účinku, ale neskôr hodnoty prestal výrazne ovplyvňovať.



**Obrázek 2:** Graf funkcie hustoty logaritmicke-normálneho rozdelenia pre zafixovanú sigma = 0.7 a  $\mu \in \{0.0, 0.4, 0.8, 1.2\}$ .

#### 4.2 Krivky funkcie hustoty logaritmicke-normálneho rozdelenia pre zafixovanú $\sigma = 0.7$

Graf na obrázku Obr. 2 nám vyobrazuje krivky logaritmicke-normálneho rozdelenia pre  $\mu \in \{0.0, 0.4, 0.8, 1.2\}$  a zafixovanú  $\sigma = 0.7$ . Hodnoty efektívnosti motivačného zdroja neboli výrazne znížením hodnoty parametra  $\sigma$  ovplyvnené, akurát následkom bol pomalší nástup efektu a pomalšie sa vytrácanie efektu.

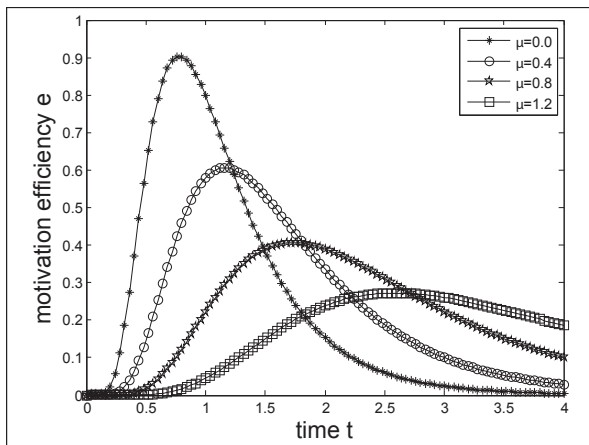
Môžeme preto vyhlásiť že znížením parametra  $\sigma$  sme klasifikovali efekt motivačného zdroja za dlhšie trvajúci. Túto vlastnosť má ale primárne úprava parametra  $\mu$ . Radšej vyhlásme že tým, že efekt bol v obdobnej účinnosti dlhší čas efektívny sme vlastne zvýšili celkovú kumulovanú účinnosť motivačného zdroja, čo je v poriadku.

Zaujímavé je zároveň sledovať krivky po prekročení času  $t > 2$ . Efekt najkratšie trvajúceho motivačného zdroja ( $\mu = 0.0$ ) sa prudko znížil, zatiaľ čo motivačný zdroj pre  $\mu = 1.2$  svoju účinnosť nestratil tak prudko.

Toto je hlavný rozdiel medzi krátkodobou a intenzívnou motiváciou, a dlhodobou nie tak intenzívnou motiváciou, ktorá má relatívne rovnakú kumulatívnu hodnotu, ale je rovnomernejšie distribuovaná.

#### 4.3 Krivky funkcie hustoty logaritmicke-normálneho rozdelenia pre zafixovanú $\sigma = 0.5$

Aby boli rozdiely ešte viditeľnejšie, na obrázku Obr. 3 nám poskytuje ukážku kriviek pre zafixovanú  $\sigma = 0.5$  a hodnoty  $\mu \in \{0.0, 0.4, 0.8, 1.2\}$ . Rozdiely v rozložení účinnosti motivačného zdroja sú v tomto prípade očividné, kým motivačný zdroj pre  $\mu = 0.0$  okolo času  $t = 2$  prudko stráca účinnosť, motivačný zdroj pre  $\mu = 1.2$  práve



**Obrázek 3:** Graf funkcie hustoty logaritmicke-normálneho rozdelenia pre zafixovanú sigma  $\sigma = 0.5$  a  $\mu \in \{0.0, 0.4, 0.8, 1.2\}$ .

začína účinkovať. Motivačný zdroj pre  $\mu = 0.8$  práve končí hlavnú etapu účinku a chystá sa na ústup.

#### 4.4 Krivky funkcie hustoty logaritmicke-normálneho rozdelenia pre zafixované $\mu = 0.0$

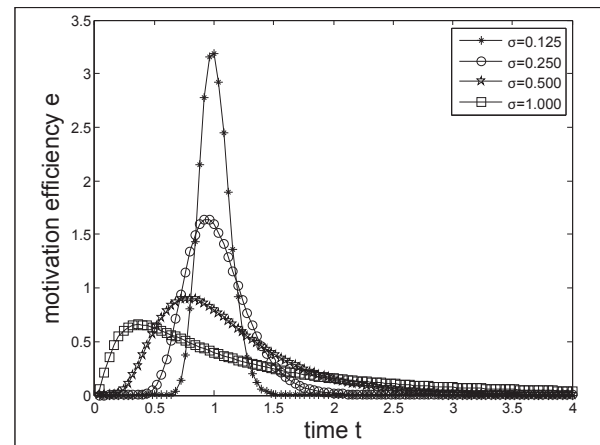
Krivky funkcií hustoty logaritmicke-normálneho rozdelenia vyobrazené na Obr. 4 znázorňujú fixované  $\mu = 0.0$  a variabilnú  $\sigma \in \{0.125, 0.250, 0.500, 1.000\}$ . Ako sme spomenuli, parameter  $\mu$  ovplyvňuje dobu trvania motivácie, resp. distribúciu motivácie v čase. Obrázok Obr. 4 znázorňuje, že všetka motivácia bola takmer vyčerpaná do času  $t = 2.5$ . Distribúcia bola oveľa viac zhustená ako pri Obr. 3, kde ešte pri  $t = 4$  mala stále dosť veľkú účinnosť.

Ukážkový je ale výskok motivácie v jednom bode, v okolí  $t = 1$ , keďže stredná hodnota  $\mu = 0$ . Pre nízku smerodajnú odchýlku  $\sigma = 0.250$  a  $\sigma = 0.125$  hodnota motivácie prudko vyskakuje, tým pádom je v krátkom časovom okamihu veľmi vysoká.

Túto situáciu môžeme obrazne prirovnať k zle vychovanému psovi. Takýto pes presne vie aký prvok je po ňom žiadaný. Pes nie je hlúpy, ale je moc lenivý lebo nie je motivovaný poslúchať. Má takmer nulovú motiváciu poslúchať autoritu a byť aktívny. Ale ako náhle vytiahneme maškrtu, je okamžite ochotný spraviť čokoľvek. Po konzumácii odmeny jeho motivácia klesne prudko nadol do pôvodného stavu.

## 5 Motivácia ďalšej práce

Kvalitnou analýzou napríklad pracovného prostredia, ako aj kolektívu na pracovisku by sme mohli relatívne presne



**Obrázek 4:** Graf funkcie hustoty logaritmicke-normálneho rozdelenia pre zafixované  $\mu = 0.0$  a  $\sigma \in \{0.125, 0.250, 0.500, 1.000\}$ .

klasifikovať motivačné a demotivačné zdroje. Mohli by sme na základe rozsiahlej štúdie nájsť množiny, do ktorých by sme tieto zdroje a prostriedky zaradili.

Aproximáciou funkcií motivačných efektov na výkonnosť môžeme optimalizovať dávkovanie motivácie jedincom i skupinám. Tým by sme optimalizovali ich výkonnosť, naplnili ich hodnoty a ciele.

Motivovaný človek nie len že dobre vykoná zadanú prácu, ale je s ňou aj spokojný, lebo má finančné, osobnostné alebo kolektívne uznanie.

## 6 Záver

Existuje mnoho spôsobov ako analyzovať motiváciu. Máme k dispozícii dostatok matematických prostriedkov, ktoré nám môžu pomôcť pri analýze kognitívnych prostriedkov. Analýzou môžeme identifikovať zdroje motivácie, demotivujúce vplyvy.

Motivačné prostriedky ako také sú ťažko merateľné. Môžeme však merať a porovnávať napríklad výkonnosť motivovaných a nie motivovaných jedincov výkonnosť jedincov. Štatistické výsledky budú samozrejme v prospech motivovaných.

Tieto údaje sú taktiež skreslené, lebo motivačných a demotivačných faktorov je v okolí jedinca mnoho. Tým pádom je jedinca ako subjekt neizolovateľný. Samotná izolácia môže mať demotivačný účinok.

Napriek týmto faktom vieme o existencii motivácie a jej vplyvoch pomerne veľa. Sami na sebe vieme pozorovať pomocou nálady a pracovných či iných výkonoch, ako na nás vplýva okolie. Na základe týchto poznatkov môžeme predvídať účinnosť motivačných prostriedkov na

modelových jedincov alebo skupiny.

## Pod'akovanie

Tento príspevok vznikol za podpory grantovej agentúry VEGA (Vedecká grantová agentúra Ministerstva školstva, vedy, výskumu a športu Slovenskej republiky a Slovenskej akadémie vied) v rámci grantového projektu VEGA 1/0553/12 a VEGA 1/0458/13.

## Reference

- [1] Lognormal distribution. In: C.F. Lee, A. Lee (eds.) Encyclopedia of Finance, p. 170. Springer US (2006)
- [2] Lognormal distribution. In: The Concise Encyclopedia of Statistics, pp. 321–322. Springer New York (2008)
- [3] Lognormal distribution. In: Encyclopedia of Genetics, Genomics, Proteomics and Informatics, p. 1119. Springer Netherlands (2008)
- [4] Motivation. In: M. Binder, N. Hirokawa, U. Windhorst (eds.) Encyclopedia of Neuroscience, p. 2413. Springer Berlin Heidelberg (2009)
- [5] Motivation. In: N. Seel (ed.) Encyclopedia of the Sciences of Learning, p. 2335. Springer US (2012)
- [6] Jolicoeur, P.: The lognormal distribution. In: Introduction to Biometry, pp. 89–93. Springer US (1999)
- [7] Rapoport, A.: Mathematical theory of motivation interactions of two individuals: I. The bulletin of mathematical biophysics **9**(1), 17–28 (1947)
- [8] Seel, N.: Attribution theory of motivation. In: N. Seel (ed.) Encyclopedia of the Sciences of Learning, pp. 379–381. Springer US (2012)



# Hierarchické učenie posilňovaním

Viliam Dillinger

Katedra aplikovanej informatiky, FMFI, Univerzita Komenského  
Mlynská dolina 842 48 Bratislava  
Email: viliam.dillinger@fmph.uniba.sk

## Abstrakt

Učenie posilňovaním je oblasť strojového učenia, ktorá skúma metódy a algoritmy, prostredníctvom ktorých sa agent učí, ako plánovať svoje akcie pomocou priamej interakcie s prostredím, s cieľom optimalizovať svoje správanie. Aj keď učenie posilňovaním vyriešilo mnoho problémov svojho predchodcu, dynamického programovania, stále je ovplyvnené exponenciálnym nárastom zložitosti v prípade rastúceho počtu atribútov opisujúcich stav prostredia. V snahe prekonať tento hlavný problém, známy ako preklatie vysokej dimenzie, bolo v poslednej dobe vyvinutých viacero prístupov ako rozšírení klasického učenia posilňovaním zavedením časovo rozšírených akcií a pomocou hierarchií. Tento článok poskytuje prehľad týchto rozšírení.

## 1 Úvod

Učenie posilňovaním (reinforcement learning, RL) je oblasť strojového učenia, skúmajúca metódy, ktorými sa agent učí plánovať svoje akcie pomocou priamej interakcie s prostredím. Z hľadiska riadenia RL zahŕňa online aproximáciu riešenia stochastického problému optimálneho riadenia, väčšinou v podmienkach bez úplnej znalosti prostredia. Väčšina RL algoritmov je založená na dynamickom programovaní (DP). V porovnaní s DP, učenie posilňovaním má mnoho výhod:

- je online, takže je ním možné riešiť problémy v reálnom čase,
- umožňuje využitie funkčných aproximátorov, čo výrazne zrýchľuje učenie, najmä v prostrediach so spojenými stavmi a/ alebo akciami,
- sústreďuje výpočty pozdĺž behaviorálnych trajektorií, teda svoje učenie sústreďuje na relevantné vedomosti.

Vďaka týmto výhodám učenie posilňovaním poskytuje dobré výsledky na problémoch, ktoré boli veľmi

zložité pre DP. Zároveň je učenie posilňovaním biologicky dôveryhodné a bolo študované aj psychológmi a neurovedcami [19].

Aj napriek tomu, že RL ponúka dobrú efektívnosť vo veľkých prostrediach, stále je ovplyvňované preklatím vysokej dimenzie – počet parametrov, ktoré je nutné sa naučiť, exponenciálne stúpa s počtom atribútov popisujúcich stav prostredia. Mnoho nedávno vyvinutých prístupov sa pokúša vyliečiť toto preklatie pridaním časovo rozšírených akcií, čo prirodzene vedie k hierarchickým architektúram riadenia a učiacim algoritmom.

## 2 Markovov rozhodovací proces

Markovov rozhodovací proces (Markov decision process, MDP) je najrozšírenejším matematickým nástrojom na modelovanie rozhodovacieho procesu. V každom časovom kroku je prostredie v stave  $s$ . Agent si vyberie a vykoná akciu  $a$ , ktorá môže byť vykonaná v danom stave. Následne, prostredie zmení svoj stav na stav  $s'$  s pravdepodobnosťou  $P(s'|s, a)$  a agent dostane odmenu  $R(s'|s, a)$ . Cieľom agenta je maximalizovať kumulatívnu odmenu v čase.

**Definícia 1** *Markovov rozhodovací proces je štvorica  $\langle S, A, P, R \rangle$  kde*

- $S$  je konečná množina stavov. V každom kroku sa agent nachádza práve v jednom stave,
- $A$  je konečná množina akcií. V stave  $s$  agent môže vykonať akciu z množiny  $A_s$ ,  $A = \bigcup_{s \in S} A_s$ ;
- $P$  je pravdepodobnostná distribúcia nasledujúcich stavov. Ak je prostredie v stave  $s$  a agent vykoná akciu  $a$ , stav prostredia sa zmení na  $s'$  s pravdepodobnosťou  $P(s'|s, a)$ ;
- $R$  je funkcia odmeny. Po vykonaní akcie  $a$  v stave  $s$  a presunu do stavu  $s'$ , agent dostane odmenu  $R(s'|s, a)$ ;

Správanie agenta v prostredí popisuje stratégia  $\pi$ , ktorá ku každému stavu priradzuje akciu, ktorú agent v danom stave vykoná,  $a = \pi(s)$ ,  $s \in S$ ,  $a \in A_s$ .

Pre konkrétnu stratégiu definujeme  $V$ -funkciu (value function). Táto funkcia odhaduje budúce odmeny, ktoré agent dostane, ak sa nachádza v danom stave a bude vyberať akcie podľa konkrétnej stratégie. Z iného uhla pohľadu, táto funkcia hovorí, ako dobré je pre agenta byť v danom stave. Matematicky je definovaná ako:

$$V^\pi(s) = E\{r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \dots | s_t = s, \pi\}$$

kde  $\gamma \in (0, 1)$  je diskontný faktor. Tento faktor popisuje, ako veľmi agent preferuje skoršie odmeny. Optimálna  $V$ -funkcia  $V^*$ , je taká, ktorá maximalizuje hodnotu každého stavu. Jej hodnoty vieme vypočítať pomocou Bellmanovej rovnice (dynamické programovanie) ako

$$V^*(s) = \max_a \sum_{s'} P(s'|s, a) [R(s'|s, a) + \gamma V^*(s')]$$

Rovnako ako  $V$ -funkcia odhaduje budúce odmeny v stavoch,  $Q$ -funkcia (action-value function)  $Q(s, a)$  odhaduje budúce odmeny v stave  $s$  pri vykonaní akcie  $a$ . Optimálna  $Q$ -funkcia, rovnako ako  $V$ -funkcia, maximalizuje hodnoty každého páru stav-akcia. Pomocou dynamického programovania ju vypočítame ako

$$Q^*(s, a) = \sum_{s'} P(s'|s, a) [R(s'|s, a) + \gamma \max_{a'} Q^*(s', a')]$$

Stratégia, ktorá v každom stave vyberie akciu s maximálnou hodnotou, je optimálna stratégia  $\pi^*$ . Každá optimálna stratégia musí byť chamtivá (greedy) vzhľadom na  $Q^*$  ( $V^*$ ) a dostupné akcie (musí sa pohybovať pozdĺž stavov s maximálnymi hodnotami).

## 2.1 Semi-Markove rozhodovacie procesy a iné rozšírenia MDP

Klasické MDP predpokladá, že každá akcia trvá rovnako dlhý čas. Semi-Markovove rozhodovacie procesy (SMDP) rozširujú koncept klasického MDP o akcie, ktoré trvajú niekoľko časových jednotiek.

**Definícia 2** Časovo diskrétny SMDP je štvorica  $\langle S, A, P, R \rangle$  kde  $S$  a  $A$  sú rovnaké ako v definícii MDP (pozri definíciu 1) a

- $P$  – ak je prostredie v stave  $s$  a agent vykoná akciu  $a$ , stav prostredia sa zmení na  $s'$  za  $\tau$  časových jednotiek s pravdepodobnosťou  $P(s', \tau | s, a)$

- $R$  – po tom, čo agent vykoná akciu  $a$  v stave  $s$  a presunie sa do stavu  $s'$  za  $\tau$  časových jednotiek, dostane odmenu  $R(s', \tau | s, a)$

Bellmanovu rovnicu pre  $Q$ -funkciu (rovnako aj pre  $V$ -funkciu) je nutné pre SMDP upraviť na:

$$Q^\pi(s, a) = \sum_{s', \tau} P(s', \tau | s, a) [R(s', \tau | s, a) + \gamma^\tau Q^\pi(s', \pi(s'))]$$

Ďalšou požiadavkou klasického MDP je konečná množina stavov a akcií. Pôvodný koncept je však veľmi ľahko možné rozšíriť do spojitého priestoru. Pri tomto rozšírení je však nutné použiť funkčné aproximátory na reprezentáciu  $V$ -funkcie či  $Q$ -funkcie, nakoľko pôvodné zapamätanie si hodnôt všetkých kombinácií stavov a akcií už nie je možné.

Medzi dôležité rozšírenia patrí aj čiastočne pozorovateľné MDP (partially observable MDP, POMDP). v tomto rozšírení k pôvodnej definícii MDP pribudne množina pozorovaní  $O$ . Namiesto stavu prostredia, agent “vidí” iba pozorovanie  $o$  s pravdepodobnosťou  $\Omega(o | s', a)$  ( $o \in O$ ). Agent si teda musí uchovávať pravdepodobnostnú distribúciu stavov, v ktorých sa môže nachádzať. Celý problém je však možné redukovat na riešenie spojitého MDP, v ktorom aktuálny stav je reprezentovaný práve danou pravdepodobnostnou distribúciou.

## 3 Učenie posilňovaním

Aj keď pomocou DP vieme vyriešiť dané MDP v polynomiálnom čase vzhľadom na počet stavov, výpočty môžu trvať veľmi dlho najmä vo veľkých priestoroch. RL algoritmy dokážu aproximovať MDP podstatne rýchlejšie ako klasické DP. Podľa [2], RL algoritmy kombinujú niektoré z nasledujúcich vlastností:

1. Obmedzujú výpočty na stavy pozdĺž tréningových trajektórií, resp. na stavy v ich okolí. Tieto trajektórie získavajú buď pomocou simulácie, alebo pomocou priamej interakcie s prostredím. Vďaka tomu sa predíde zbytočnému počítaniu v stavoch, ktoré majú malú pravdepodobnosť dosiahnutia, a naopak sa dosiahne dobrá aproximácia v relevantných stavoch.
2. Miesto toho, aby uvažovali všetky možné nasledujúce stavy, odhadnú  $V$  resp.  $Q$ -funkciu pomocou vzorkovania z vhodnej distribúcie. Vďaka tomu zjednodušia výpočet týchto funkcií.
3. Používajú funkčné aproximátory na reprezentáciu  $V$  resp.  $Q$ -funkcií a/alebo stratégie, čo zníži po-

žiadavky na pamäť a umožní prenos vedomostí, čím zrýchli konvergenciu.

### 3.1 Q-učenie a SARSA

Medzi najpoužívanejšie RL algoritmy patrí Q-učenie (Q-learning) a SARSA. Q-učenie používa okamžitú odmenu a predpokladanú maximálnu Q-hodnotu v nasledujúcom stave na zlepšenie odhadu Q-hodnoty práve vykonanej akcie v predchádzajúcom stave. Základnou myšlienkou algoritmu je, že ak agent vykoná akciu, ktorá vedie k vysokej okamžitej odmene alebo do stavu, v ktorom je možné vykonať dobre ohodnotenú akciu, posilní hodnotu práve vykonanej akcie. Naopak, ak práve vykonaná akcia vedie k zlej okamžitej odmene (trestu) alebo sa dostane do stavu, v ktorom všetky akcie majú zlé Q-hodnoty, oslabí práve vykonanú akciu.

Formálne, ak je agent v stave  $s$ , vyberie akciu  $a$ , dostane sa do stavu  $s'$  a obdrží odmenu  $r$ , upraví si  $Q(s, a)$  pomocou:

$$Q_{t+1}(s, a) = (1 - \alpha_t)Q_t(s, a) + \alpha_t \left( r + \gamma \max_{a' \in A_{s'}} Q_t(s', a') \right)$$

kde  $0 < \alpha_t < 1$  je rýchlosť učenia. Ak s narastajúcim časom  $t$  rýchlosť učenia klesá k nule,  $Q$  skonverguje k  $Q^*$  s pravdepodobnosťou 1.

SARSA, pomenovaná podľa state-action-reward-state-action, aproximuje Q-funkciu veľmi podobne ako Q-učenie. Základný rozdiel je v tom, že SARSA pri upravovaní Q-hodnôt nepoužíva najlepšiu možnú akciu v nasledujúcom stave, ale počká si, až stratégia vyberie nasledujúcu akciu. Následne pred jej vykonaním zoberie jej hodnotu a pomocou nej a práve obdržanej odmeny upraví hodnotu predchádzajúcej akcie.

$$Q_{t+1}(s, a) = (1 - \alpha_t)Q_t(s, a) + \alpha_t (r + \gamma Q_t(s', \pi(a)))$$

Oba algoritmy je možné modifikovať pre SMDP tak, že odmenu nahromadenú počas vykonávania akcie dĺžky  $\tau$  budeme interpretovať ako okamžitú odmenu a  $\gamma$  upravíme rovnako ako v prípade DP pre SMDP. Pôvodné Q-učenie modifikujeme ako

$$Q_{t+1}(s, a) = (1 - \alpha_t)Q_t(s, a) + \alpha_t \left( r + \gamma^\tau \max_{a' \in A_{s'}} Q_t(s', a') \right)$$

$$r = r_{t+1} + \gamma r_{t+2} + \dots + \gamma^{\tau-1} r_{t+\tau}$$

### 3.2 Architektúra aktér-kritik

Ďalším z obľúbených prístupov v RL je architektúra aktér-kritik, v ktorej je agent rozdelený na dve časti.

Aktér vyberá akcie a učí sa Q-funkciu, v tomto kontexte označovanú ako  $W$ . Kritik kritizuje akcie aktéra vo forme chýb a učí sa V-funkciu. Keď agent vyberie akciu  $a$  v stave  $s$ , dostane odmenu  $r$  a prejde do stavu  $s'$ , ako prvá sa vyráta chyba:

$$\delta = r + \gamma V(s') - V(s)$$

Následne sa pomocou nej upravujú hodnoty aktéra aj kritika

$$V_{t+1}(s) = V_t(s) + \alpha_C \delta$$

$$W_{t+1}(s, a) = W_t(s, a) + \alpha_A \delta$$

kde  $\alpha_C$  a  $\alpha_A$  sú rýchlosti učenia kritika a aktéra.

### 3.3 Problémy učenia posilňovaním

V klasickom RL (taktiež nazývanom aj "plochom") je mnoho problémov, ktoré viedli k ďalšiemu výskumu. Zmienime sa len o najvýraznejších z nich vo vzťahu k hierarchickej RL.

#### Preklatie vysokej dimenzie

Preklatie vysokej dimenzie je prvým a najväčším problémom RL. Ak zvýšime počet atribútov, ktoré popisujú stav prostredia, počet stavov narastie exponenciálne s každým pridaným atribútom. Čím je väčšie prostredie, tým je náročnejšie pre agenta naučiť sa jeho mechaniky. Smutné je, že sa tomu nedá vyhnúť, nakoľko problémy reálneho sveta majú väčšinou veľmi veľa atribútov.

#### Náhodná chôdza

Na začiatku svojho učenia agent len náhodne vyberá akcie, nakoľko ešte nedostal žiadnu odmenu, podľa ktorej by sa riadil. Tento problém je omnoho väčší v prostrediach s roztrúsenou odmenou, resp. v prostrediach, kde agent dostane odmenu až po dosiahnutí cieľového stavu. Ak je vzdialenosť od jeho počiatočného stavu k cieľovému stavu veľká (čo je pri veľkých stavových priestoroch bežné), predpokladaný čas, za ktorý dosiahne tento stav neúmerne narastá s danou vzdialenosťou. Teda čas, ktorý je potrebný na to, aby sa agent vôbec začal učiť, je veľmi veľký.

#### Šírenie odmeny

Ďalší problém, ktorý úzko súvisí s problémom náhodnej chôdze, je šírenie odmeny na veľké vzdialenosti. Keď agent konečne prvý krát dosiahne cieľový stav, učenie ledva začína. Problémom v klasickom RL je to,

že agent sa “pozerá” len o jeden krok dopredu a len o jeden krok dozadu upravuje svoje hodnoty. To znamená, že agent musí náhodnou chôdzou znova a znova objavovať cieľový stav, alebo niektrý zo stavov na objavenej ceste k nemu. Čím je väčšia vzdialenosť k cieľu, tým viac iterácií je potrebných na naučenie sa cesty.

## Prenos vedomostí

V mnohých prostrediach sú si niektoré ich časti navzájom podobné. Ak sa agent naučí vykonávať akcie v jednej z nich, táto vedomosť môže byť prospešná aj v druhej. Jeden zo spôsobov prenosu vedomostí je využitie funkčných aproximátorov, napríklad neurónových sietí. Aj keď dokážu zovšeobecňovať nad stavmi, bolo by lepšie, keby agent mohol priamo použiť niektorú časť stratégie v inej časti priestoru.

## 4 Hierarchické učenie posilňovaním

Aj keď RL dokáže nájsť riešenie MDP, nedokáže sa vyhnúť spomínaným problémom. Jednou z možností, ako ich prekonať, je použiť hierarchické štruktúry s využitím časovo rozšírených akcií. Jednotlivé prístupy v tejto oblasti vieme rozdeliť na dve základné skupiny – prístupy, ktoré využívajú tzv. opcie (options) a prístupy, ktoré zavádzajú vrstvy.

### 4.1 Prístupy využívajúce opcie

Opcie sú známe pod mnohými rôznymi názvami, ako napríklad makrá, makro-operátory, zručnosti, časovo rozšírené akcie, režimy, alebo aktivity. Ich princíp je však rovnaký. K množine základných jednokrokových akcií pridáme opcie (alebo naopak, primitívne akcie považujeme za jednokrokové opcie). Opcia je sekvencia primitívnych akcií. Opcie môžu byť úplne fixné, ale môžu to byť aj stochastické stratégie špecializované na riešenie niektorého podproblému. Vďaka tomu môžu byť využité na riešenie podobných problémov v rôznych častiach stavového priestoru.

V práci [20] navrhli rozšíriť základnú myšlienku RL o časovo rozšírené akcie. Opcie vyžadovali rozšírenie štandardnej definície MDP na SMDP, pričom ako dĺžku akcie používame dĺžku sekvencie primitívnych akcií, ktoré boli vykonané počas realizácie zvolenej opcie.

Počas vykonávania opcie, tá môže zavolať inú opciu ako svoj podprogram. Hierarchická štruktúra vzniká práve v zmysle daných volaní. Vo svojej podstate, opcie rozširujú možnosti výberu akcie agenta. Bohužiaľ, výber akcie (opcie) je stále založený na práve

pozorovanom stave, takže abstrakcia vzniká len na úrovni akcií, nie na stavoch.

#### 4.1.1 Fixné stratégie

V práci [18] boli navrhnuté viackrokové akcie (multi-step akcie). Viackroková akcia stupňa  $n$  je opcia, ktorá pozostáva zo zopakovania vybranej primitívnej akcie  $n$  krát. Aj keď je táto forma opcií triviálna, agentovi umožní ľahšie prehľadávať priestor počas náhodnej chôdze, čím zvýši šancu dosiahnuť aj vzdialené stavy v relatívne krátkom čase.

Počas vykonávania opcie sa netrénuje len daná opcia, ale aj všetky možné kombinácie opcií nižšieho stupňa, pomocou ktorých agent skončí v rovnakom stave. Takéto učenie viacerých  $Q$ -hodnôt zlepšuje šírenie odmeny. Bohužiaľ, tento prístup neposkytuje žiadnu príležitosť pre prenos vedomostí.

#### 4.1.2 Hľadanie relevantných stavov

Veľká skupina prístupov pomocou opcií používa rovnakú myšlienku. Nájsť relevantný stav v danom probléme (zvaný tiež podcieľ) a potom natréňovať opciu, aby vedela dosiahnuť daný stav.

Tieto prístupy vo všeobecnosti veľmi dobre fungujú, avšak majú svoje úskalía, ktoré vo svojej práci experimentálne ukázal [3]. Ak sa zvolené podciele nachádzajú ďaleko od optimálnej cesty, spôsobí to spomalenie tréningu agenta, pretože sa agent pri náhodnej chôdzi vďaka nim opakovane vzdáva od svojho cieľa. Ak sa zvolené ciele nachádzajú v blízkosti optimálnej cesty, za istých okolností môžeme skonvergovať len ku suboptimálnym riešeniam.

Výber vhodných podcieľov je teda pri tejto skupine prístupov kľúčový. Podciele môže definovať programátor, avšak existujú algoritmy na ich automatickú detekciu.

$Q$ -rez ( $Q$ -cut) algoritmus [14] hľadá podciele pomocou max-flow/min-cut algoritmu (teória grafov) a mapy z histórie spracovania. Algoritmus hľadá úzke miesta v stavovom priestore pomocou rezov, ktoré rozdelia stavový priestor na separátne časti. Nájsené zúženia sú určené ako podciele.

Počet predchodcov (Predecessor-count) [7] využíva rozdiely v počte predchodcov ako svoju mieru relevantnosti. Pri signifikantnom zvýšení tejto miery algoritmus považuje daný stav za relevantný a vyberie ho ako podcieľ.

Ďalší prístup detekuje podciele pomocou rôznorodej hustoty (diverse density). Podľa [11], agent počas náhodnej chôdze s veľkou pravdepodobnosťou ostáva

v silno prepojených regiónoch stavového priestoru. Opcia by teda mala prepájať práve dva takéto regióny, čím by ich k sebe priblížila a tým umožnila agentovi lepšie prehľadávať priestor. Frekvenciu návštev ako mieru relevantnosti zavrholi, nakoľko nemohla byť generalizovaná vo veľkých stavových priestoroch. Miesto nej použili svoj algoritmus založený na mnoho-inštančnom učení (multiple-instance learning).

Na spojitú stavovú prostredia sa špecializuje acQuire-makrá (acQuire-macros) [10]. Tento algoritmus hľadá maximum v histórii odmien, ktoré používa na vytvorenie trajektórie v priestore. Ak nájde takéto maximum, zistí, či sa nachádza v často navštevovanom regióne. Ak áno, daný región vyberie ako svoj podcieľ. Aj keď tento algoritmus vyžaduje uchovávanie histórie trajektórií, jeho pamäťové požiadavky nie sú veľmi veľké vďaka dobrej kompresii dát.

## 4.2 Viacvrstvé prístupy

Prístupy z tejto skupiny priamo vytvárajú hierarchie, pomocou ktorých rozkladajú stavový priestor na jednotlivých vrstvách s rôznou úrovňou detailu. Na vyšších vrstvách sú obsiahnuté len globálne relevantné informácie, zatiaľ čo nižšie vrstvy obsahujú detailné informácie. Vďaka tomu nie je rozhodovanie na vyšších úrovniach obťažované triviálnymi otázkami, avšak na nižších úrovniach sú všetky informácie dostupné, rovnako ako v klasickom RL. Aby sme to dosiahli, rozložíme problém na niekoľko MDP. Každé z nich bude mať vlastnú množinu stavov, pričom vyššie vrstvy majú túto množinu menšiu (abstraktné stavy).

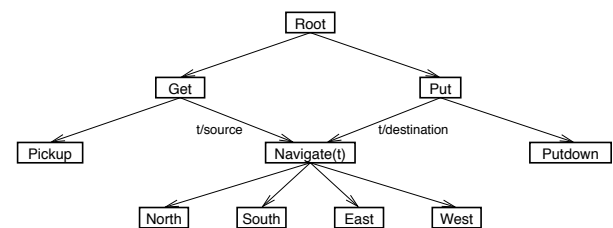
Každý z prístupov tejto skupiny zavádza rôzne obmedzenia agenta, ktoré znižujú jeho schopnosti. Aj keď je výsledná stratégia optimálna vzhľadom na dané obmedzenia, nemusí byť optimálna vzhľadom na pôvodné MDP. Na druhú stranu, tieto prístupy dokážu nájsť riešenie niekoľkonásobne rýchlejšie ako klasické RL. Ide o kompromis medzi nájdením dobrého riešenia (aj keď nie optimálneho) rýchlo a nájdením optimálneho riešenia za veľmi dlhý čas. Je na programátorovi aby zabezpečil aby jeho hierarchia povoľovala aspoň skoro optimálne riešenia.

### 4.2.1 MAXQ hodnotová dekompozícia

Jeden z najobľúbenejších prístupov hierarchického RL je MAXQ hodnotová dekompozícia (MAXQ value decomposition) od [5]. MAXQ rozloží pôvodné MDP na hierarchiu podúloh (menších MDP) a následne využíva  $Q$ -hodnoty týchto podúloh na výpočet  $Q$ -funkcie pôvodného MDP (jednoduché sčítanie). Každá podúloha

má presne špecifikované, aké primitívne akcie a podúlohy môže používať. Akcie a podúlohy, ktoré nie sú relevantné na riešenie danej podúlohy jednoducho nie sú zaradené do množiny akcií danej podúlohy. Agent volí svoje akcie nielen na základe aktuálneho stavu, ale aj na základe hierarchie volaní podúloh, čo môže viesť k nemarkovovskej stratégii.

Dekompozícia je orientovaný acyklický graf. Každý uzol reprezentuje podúlohu, listy sú primitívne akcie a hrany znázorňujú, ktorá podúloha môže volať ktorú podúlohu alebo primitívnu akciu. Podúloha môže mať parametre, avšak podúlohy s rôznymi parametrami sú považované ako rôzne podúlohy (parametre sa berú akoby boli súčasťou názvu). Príklad takéhoto grafu je zobrazený v obrázku 1.



**Obr. 1:** Príklad grafu úlohy preTaxi problém. Agent má za úlohu nanavigovať sa k zákazníkovi, zobrať ho, odviesť ho podľa jeho želania a tam ho vyložiť. Prevzaté od [5].

Práca [12] ponúka spôsob, ako vytvoriť dekomponovaný graf z vyriešeného MDP a jednej úspešnej trajektórie. Aj keď tieto požiadavky znamenajú vyriešenie problému, je možné takto získaný graf využiť na riešenie podobných problémov. Väčšinou je teda na programátorovi, aby identifikoval relevantné podúlohy.

Viaceri výskumníci rozšírili tento algoritmus pre viac-agentové problémy. V práci [6] autor vyvinul kooperatívne HRL (Cooperative HRL), kde vyššie vrstvy (podproblémy) sú označené ako kooperatívne a agenti na nich získavajú informácie o iných agentoch. Ďalším multi-agentovým rozšírením je napríklad multi-agentová zdieľaná hierarchia (Multi Agent Shared Hierarchy) [13].

### 4.2.2 Hierarchia automatov

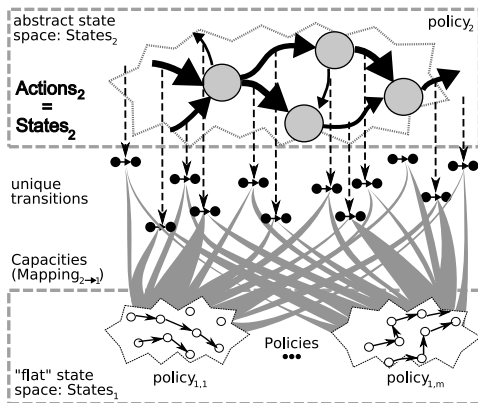
Hierarchia automatov (Hierarchies of machines, HAM) [16] je založená na obmedzení učných stratégií pomocou hierarchie čiastočne špecifikovaných konečných nedeterministických automatov. Tieto automaty majú štyri druhy stavov – akčné stavy, kedy je vykonaná akcia v prostredí; volacie stavy, kedy automat zavola

iný automat ako svoj podprogram; stop stavy, kedy automat vráti riadenie automatu, ktorý ho volal a výberové stavy, kedy automat nedeterministicky vyberie ďalší stav.

Učenie nastáva vo výberových stavoch s využitím istej formy Q-učenia. Pri výpočte  $Q$ -hodnoty sú ako stavy použité dvojice (stav prostredia, stav automatu).

#### 4.2.3 HASSLE a HABS

Učenie pomocou hierarchického priradovania podstratégií k podcieľom (Hierarchical Assignment of Subpolicies to Subgoals Learning, HASSLE) [1] je viacvrstvá štruktúra, v ktorej stavy na vyšších vrstvách sú abstrakciami stavov nižšej vrstvy. Prechody medzi týmito abstraktnými stavmi realizujú podstratégie na nižšej vrstve. Na každej vrstve je fixný počet podstratégií, teda každá z nich musí realizovať viacero týchto prechodov. Číselné vyjadrenie schopnosti vykonať daný prechod autor nazýva kapacitou a využíva ho pri výbere vhodnej podstratégie na realizáciu vybraného prechodu (pozri obrázok 2).



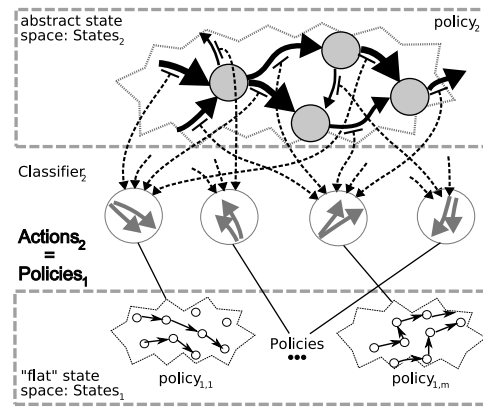
Obr. 2: Hierarchická štruktúra modelu HASSLE. Prevzaté od [15].

Keď vybraná podstratégia dosiahne cieľový abstraktný stav, je odmenená a jej kapacita pre daný prechod sa zvýši. Ak dosiahne iný ako cieľový abstraktný stav, dostane nulovú odmenu a pôvodný cieľový abstraktný stav je vymenený za dosiahnutý. Následne sa upravujú kapacity spolu so stratégiou vyššej vrstvy. Vďaka tomuto postupu je učenie stratégie vyššej vrstvy veľmi efektívne.

Problémy však nastávajú so zvyšujúcim sa počtom abstraktných stavov, kedy narastá počet prechodov medzi nimi. Pre každý takýto prechod si okrem jeho  $Q$  hodnoty musíme pamätať aj všetky jeho kapacity, čo značne spomalí učenie a zvýši pamäťovú náročnosť.

Zároveň nie je možné použiť funkčné aproximátory na vyššej vrstve a nie je vhodné využívať viac ako dve vrstvy kvôli nemožnosti znovu využiť podstratégie na tretej a vyššej vrstve.

Veľkú časť problémov HASSLE vyriešilo Hierarchické priradovanie správania pomocou samoorganizácie (Hierarchical Assignment of Behaviours by Self-organizing, HABS) [15]. Tento prístup sa spolieha na istú geometriu v atribútoch stavového priestoru. Algoritmus používa prechody medzi abstraktnými stavmi na vyššej úrovni ako relatívne, teda také, ktoré len určujú smer v priestore atribútov. HABS sa snaží namapovať pozorované prechody na vyššej úrovni na množinu podstratégií v nižšej úrovni. Tieto podstratégie sú priamo používané ako akcie vyššej úrovne (pozri obrázok 3).



Obr. 3: Hierarchická štruktúra modelu HABS. Prevzaté od [15].

Keď vybraná podstratégia ukončí svoju činnosť, klasifikátor zistí, či úspešne vykonala prechod do zvoleného abstraktného stavu. Ak áno, je odmenená, ak nie, je potrestaná. V oboch prípadoch sa popis správania, ktoré realizuje, priblíži k práve vykonanému smeru posunu. Ak sa podstratégii nepodarí prejsť do iného abstraktného stavu, je potrestaná, ale popis jej správania sa nemení. Odmena získaná počas vykonávania podstratégie sa využíva na tréning akcie vyššej úrovne.

Vďaka týmto zmenám kapacity už nie sú potrebné a taktiež je možné znovu použiť podstratégie aj na tretej a vyššej vrstve. Zároveň je možné použiť funkčné aproximátory pri realizácii stratégie vyššej úrovne. Oproti HASSLE má však horšie využitie informácie na tréning vyšších úrovní.

#### 4.2.4 Ďalšie prístupy

V práci [9] bol navrhnutý algoritmus HEXQ založený na CQ algoritme [8]. Jeho základnou myšlienkou je usporiadať atribúty prostredia podľa frekvencie ich zmeny. Rýchlo sa meniace atribúty priradí nižším vrstvám a pomaly sa meniace zasa vyšším. Tento algoritmus pracuje v dvoch fázach – v prvej vytvorí dekompozíciu a v druhej sa učí pomocou RL. V práci [17] autori navrhli súbežné HEXQ (concurrent HEXQ), ktoré umožňuje vykonávať dekompozíciu aj učenie naraz a zároveň ponúka mechanizmus na opravu zlých dekompozícií.

V práci [22] navrhli HQ-učenie (HQ-learning), ktoré považuje každý problém za multi-agentový problém. Definuje sekvenciu agentov, ktorí si vyberú pozorovanie, ktoré chcú vidieť (svoj podcieľ) a snažia sa ho dosiahnuť. Začína prvý agent a riadenie je postupne predávané ďalším agentom v poradí, pokiaľ nedosiahnu globálny cieľ alebo im nevyprší časový limit. Každý z nich je použitý práve jedenkrát. Následne si každý agent upraví svoje vlastné stratégie a svoju HQ tabuľku, pomocou ktorej si vyberá svoj podcieľ. Tento algoritmus dokáže riešiť POMDP, avšak nie je vhodný na problémy, kde treba dlhodobo udržať nejaký stav, nakoľko sa mu “minú agenti”.

V práci [4] navrhli algoritmus zvaný feudálne učenie (Feudal learning), kde manažéri (stratégie) riadia svojich podmanažérov (podstratégie) pomocou príkazov a odmeny (trestu). Vývoj hierarchie manažérov je riadený dvoma princípmi. Prvým princípom je skrývanie odmeny, pri ktorom manažéri dávajú odmenu svojim podriadeným, ak úspešne splnia zadaný príkaz, a to aj keď bol proti cieľom ich nadriadeného. Druhým princípom je skrývanie informácií, pri ktorom manažér nepotrebuje vedieť detaily o tom, čo robil jeho podriadený, rovnako ako nepotrebuje vedieť cieľ jeho nadriadeného. Stavový priestor je rozdelený na časti a každý manažér dostane jeho časť. Manažéri na najnižšej vrstve majú ako svoje akcie primitívne akcie a manažéri na vyšších vrstvách majú fixný počet preddefinovaných príkazov.

V práci [21] navrhli systém zvaný Akvizícia správania pomocou viacvrstvého RL (Behavior acquisition by multi-layered RL). Jednotlivé správania sa snažia byť expertom na rôzne časti stavového priestoru, čo sa realizuje pomocou samoorganizácie. Stav vyššej vrstvy pozostáva z vektora maximálnych  $Q$ -hodnôt každého správania z nižšej vrstvy, čo indikuje, ako ďaleko je dané správanie od dosiahnutia svojho cieľa. Každé správanie má určený svoj cieľ tak, aby vznikla približne uniformná distribúcia. To dosahujú tak, že presúvajú svoje ciele do regiónov, kde ostatné správania

majú nízke maximálne  $Q$ -hodnoty. Ak sú dve správania blízko seba, jedno z nich sa vymaže. Ak sú ciele príliš ďaleko od seba, vytvorí sa nové správanie tak, aby vhodne pokrylo nepokrytú časť prostredia.

## 5 Záver

V tomto článku sme poskytli prehľad prístupov hierarchického učenia posilňovaním, ktoré sa snažia vyriešiť problém vysokej dimenzie, a ďalšie s tým súvisiace problémy, pomocou časovo rozšírených akcií. Prístupy v tejto oblasti sa delia na dve hlavné skupiny – prístupy využívajúce opcie a prístupy využívajúce viaceré vrstvy.

Prístupy využívajúce opcie pridávajú k množine základných jednokrokových akcií opcie, ktoré sú stochastické stratégie špecializované na riešenie niektorého podproblému. Tým sa rozširujú akčné možnosti agenta s cieľom zrýchliť exploráciu a tým aj samotné učenie. Prístupy využívajúce viaceré vrstvy vytvárajú hierarchie, pomocou ktorých rozkladajú stavový priestor s rôznou úrovňou detailu na jednotlivých vrstvách. Pomocou tejto hierarchie zavádzajú rôzne obmedzenia agenta, čím znižia počet parametrov, ktoré je nutné sa naučiť, čím zrýchlia učenie.

Hierarchické učenie posilňovaním ponúka značné zrýchlenie oproti klasickému učeniu posilňovaním, čo experimentálne ukázali autori spomenutých prístupov. Rovnako ako učenie posilňovaním nahradilo dynamické programovanie pri riešení markovovských rozhodovacích procesov, hierarchické učenie posilňovaním vidíme ako ďalší vývojový stupeň v tejto oblasti.

## Podakovanie

Tento príspevok bol podporený grantom VEGA 1/0439/11.

## Literatúra

- [1] Bakker, B., Schmidhuber, J.: Hierarchical reinforcement learning based on subgoal discovery and subpolicy specialization. In: Proceedings of the 8th Conference on Intelligent Autonomous Systems (IAS-8). pp. 438–445 (2004)
- [2] Barto, A.G., Mahadevan, S.: Recent advances in hierarchical reinforcement learning 13, 2003 (2003)
- [3] Botvinick, M.M., Niv, Y., Barto, A.C.: Hierarchically organized behavior and its neural foundati-

- ons: A reinforcement learning perspective. *Cognition* 113(3), 262–280 (2009)
- [4] Dayan, P., Hinton, G.E.: Feudal reinforcement learning. In: *Advances in Neural Information Processing Systems* 5. pp. 271–278. Morgan Kaufmann (1995)
- [5] Dietterich, T.G.: Hierarchical reinforcement learning with the MAXQ value function decomposition. *Journal of Artificial Intelligence Research* (13), 227–303 (2000)
- [6] Ghavamzadeh, M., Mahadevan, S., Makar, R.: Hierarchical multi-agent reinforcement learning. *Autonomous Agents and Multi-Agent Systems* 13(2), 197–229 (2006)
- [7] Goel, S., Huber, M.: Subgoal discovery for hierarchical reinforcement learning using learned policies. In: Russell, I., Haller, S.M. (eds.) *FLAIRS Conference*. pp. 346–350. AAAI Press (2003)
- [8] Hengst, B.: Generating hierarchical structure in reinforcement learning from state variables. In: *Lecture Notes in Artificial Intelligence*. Springer (2000)
- [9] Hengst, B.: Discovering hierarchy in reinforcement learning with hexq. In: *Maching Learning: Proceedings of the 19th International Conference on Machine Learning*. pp. 243–250. Morgan Kaufmann (2002)
- [10] MCGovern, A.: acQUIRE-macros: An algorithm for automatically learning macro-actions. In: *NIPS’98 Workshop on Abstraction and Hierarchy in Reinforcement Learning* (1998)
- [11] MCGovern, A., Barto, A.G.: Automatic discovery of subgoals in reinforcement learning using diverse density. In: *ICML*. pp. 361–368. Morgan Kaufmann (2001)
- [12] Mehta, N., Ray, S., Tadepalli, P., Dietterich, T.: Automatic discovery and transfer of MAXQ hierarchies. In: *Proceedings of the 25th International Conference on Machine Learning* (2008)
- [13] Mehta, N., Tadepalli, P.: Multi-agent shared hierarchy reinforcement learning. In: *ICML Workshop on Richer Representations in Reinforcement Learning* (2005)
- [14] Menache, I., Mannor, S., Shimkin, N.: Q-cut – dynamic discovery of subgoals in reinforcement learning. In: *Proceedings of the 13th European Conference on Machine Learning*. pp. 295–306. ECML’02, Springer-Verlag, London, UK, UK (2002)
- [15] Moerman, W.: *Hierarchical Reinforcement Learning: Assignment of Behaviours to Subpolicies by Self-Organization*. Ph.D. thesis, Utrecht University (2009)
- [16] Parr, R., Russell, S.: Reinforcement learning with hierarchies of machines. In: *Advances in Neural Information Processing Systems* 10. pp. 1043–1049. MIT Press (1997)
- [17] Potts, D., Hengst, B.: Discovering multiple levels of a task hierarchy concurrently. *Robotics and Autonomous Systems* 49(1-2), 43–55 (2004)
- [18] Schoknecht, R., Riedmiller, M.: Speeding-up reinforcement learning with multi-step actions. In: *Proceedings of the 12th International Conference on Artificial Neural Networks (ICANN)*. pp. 813–818. Springer (2002)
- [19] Seth, A., Prescott, T., Bryson, J. (eds.): *Modeling Natural Action Selection*. Cambridge University Press (2012)
- [20] Sutton, R.S., Precup, D., Singh, S.: Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning. *Artificial Intelligence* 112, 181–211 (1999)
- [21] Takahashi, Y., Asada, M.: Behavior acquisition by multi-layered reinforcement learning. In: *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*. pp. 716–721 (1999)
- [22] Wiering, M., Schmidhuber, J.: HQ-learning. *Adaptive Behavior* 6, 219–246 (1997)



# Vyhrávajú stereotypy nad štatistikou? Normatívne a stereotypné uvažovanie vo vzťahu k úrovni kognitívnej reflexie

Katarína Dudeková, Lenka Kostovičová

Ústav aplikovanej psychológie  
Fakulta sociálnych a ekonomických vied UK  
Mlynské luhy 4, Bratislava 821 05  
katarina.dudekova@savba.sk  
lenka.kostovicova@gmail.com

## Abstrakt

Omyl základného výskytu (*base-rate fallacy*) je jedna z najlepšie dokumentovaných odchýlok od normatívnych odpovedí v rámci pravdepodobnostných úsudkov. Nová evidencia naznačuje, že pôvod odchýlky nespočíva v samotnom ignorovaní štatistických údajov, ale v nedostatočnej schopnosti potlačiť intuitívne heuristické odpovede založené na stereotypoch. V tomto výskume sme vychádzali z konceptu kognitívnej reflexie ako schopnosti spracovať informácie analyticky a inhibovať intuitívne odpovede.

## 1 Úvod

Aj v bežnom každodennom rozhodovaní pociťujeme odlišnosť medzi dvoma typmi kognitívnych procesov, tých uskutočnených rýchlo s malou mierou uvedomenia a tých, ktoré sú pomalšie a viac uváženejšie. Kognitívni psychológovia (ale nielen) sa myšlienkou duálnych procesov intenzívne zaoberajú už približne pol storočia, pričom stále viac diskutujú v tejto súvislosti o obraze ľudskej racionality.

### 1.1 Duálne kognitívne procesy

Kľúčovým predpokladom pre pochopenie toho, ako ľudia usudzujú a rozhodujú sa je teória dvoch procesov, ktoré prebiehajú v ľudskom mozgu. Teória duálnych procesov v kognitívnej psychológii uvažujú o dvoch odlišných procesoch – jednom automatickom, implicitnom, heuristickom a druhom explicitnom, vedomom, a teda kontrolovanom či analytickom systéme [1,2] s neutrálnymi názvami Systém 1 a Systém 2 [napr. 3], ktoré si osvojili mnohí bádatelia v danej oblasti.

Systém 1 je rýchly, podvedomý, automatický a nemá vysoké nároky na kognitívnu kapacitu. Jeho procesy prebiehajú spontánne a nevyžadujú zapojenie pozornosti,

nastanú okamžite a bez námahy [4]. Produkuje intuitívne rýchle rozhodnutia, často spočívajúce v použití zjednodušenia alebo mentálnych skratiek – heuristik [3].

Systém 2 je na rozdiel od Systému 1 racionálny, pomalší, umožňuje analytické myslenie, je kontrolovaný a vyžaduje zvýšené úsilie [5]. Prideluje pozornosť určitým vedomým psychickým procesom, ktoré si častokrát vyžadujú použitie jednoduchších alebo zložitejších výpočtov. Jeho činnosť je často spájaná so subjektívnym prežívaním, produkovaním voľby a sústredením [3].

Práve interakcia týchto systémov je podstatou nazerania na myseľ, tvorbu úsudkov a rozhodnutí. Automatická činnosť Systému 1 produkuje určité modely myšlienok a Systém 2 dokáže tieto myšlienky koordinovať a usporiadať ich do série krokov [3].

### 1.2 Heuristiky v rozhodovaní a usudzovaní – používanie Systému 1

Vznik teórie duálnych procesov podnietil i výskum v oblasti heuristik a omylov. Aj keď heuristiky v mnohých situáciách (hlavne pri nízkych nárokoch na objem potrebných informácií a energiu vynaloženú na kognitívne procesy) vedú k dobrým výsledkom, v niektorých prípadoch výrazne skresľujú naše rozhodnutia a spôsobujú podstatné omyly. Doterajší výskum naznačuje, že ľudia si používanie mentálnych skratiek pri rozhodovaní prenášajú aj do zložitejších situácií, ktoré si vyžadujú aj zapojenie analytických a matematicko-štatistických postupov.

### 1.3 Heuristika reprezentatívnosti / Omyl základného výskytu

Na základe heuristiky reprezentatívnosti odhadujeme pravdepodobnosť príslušnosti konkrétneho javu,

správania či predmetu k určitej skupine či kategórii na základe toho, do akej miery sa približujú našej predstave reprezentatívneho člena tejto skupiny [6]. Heuristika tiež odráža stupeň kognitívneho procesu, zapojeného do usudzovania o konkrétnej skupine, kategórii či jave, pričom schému kognitívneho procesu v mysli s použitím heuristiky reprezentatívnosti popisuje Obr. 1.



Obr.1. Schéma kognitívneho procesu pri použití heuristiky reprezentatívnosti [13].

Klasickú ukážku vplyvu intuitívnych heuristických operácií týkajúcich sa reprezentatívnosti na rozhodovanie možno nájsť už v experimentoch Kahnemana a Tverskeho [7], pri výskumoch omylu základného výskytu (*base rate fallacy*, *base-rateneglect*).

Omylom základného výskytu sa označuje tendencia ľudí ignorovať základný výskyt v súbore (v populácii) v prospech iných informácií, skôr aby tieto dve informácie integrovali. Štúdie v danej oblasti naznačujú, že ľudia sa príliš spoliehajú na svoje intuitívne a stereotypné presvedčenia, namiesto toho, aby svoje úsudky podstúpili náročnejšej argumentácii [8]. V týchto experimentoch účastníci odpovedali na úlohy, v ktorých stereotypný opis navádzal na výber odpovedi, najviac sa približujúcej reprezentatívne prototypu z určitej skupiny [7].

V pôvodnom výskume omylu základného výskytu Kahnemana a Tversky [7] preukázali, že ľudia sa konzistentne spoliehajú na popis kategórie a vlastnosti (napríklad ako je reprezentatívny popis osobnosti k skupine právnikov alebo inžinierov) a nechajú svoje rozhodnutia viesť týmto opisom, bez toho, aby do úsudku začlenili informácie o základnom výskyte. Kahneman

a Tversky [7] zistili, že prevažná väčšina vzdelaných vysokoškolákov nedokázala odpovedať na predložený problém správne (štatisticky správne, normatívne odpovedať). Dokonca ani univerzitní profesori neboli imúnni voči heuristickej zaujatosti, a zdalo sa, že zanedbávajú zásadnú informáciu o základnom výskyte.

Aj napriek tomu, že výskum v oblasti odhadovania pravdepodobnosti trvá už dlho, povaha tejto odchýlky od normatívneho uvažovania nie je úplne jasná. Niektorí autori tvrdia, že ľudia sú len heuristickí myslitelia a nie sú si vedomí, že poskytnuté navodené stereotypné presvedčenie môže byť v nesúlade s ich úsudkami o probléme. Iní autori tvrdia, že ľudia vždy odhalia konflikt medzi ich stereotypným myslením a normatívnym uvažovaním, ale jednoducho nedokážu zabrániť stereotypnému mysleniu [9].

K výskumu nás inšpirovalo tvrdenie [10], že medzi alternatívnymi pohľadmi na skreslenie tradičných údajov uvažovania je zložitý rozhodnutý. Problémom je z časti fakt, že výskumníci rozhodovania a usudzovania často nedávajú dôraz na poznávací procesy pri rozhodovaní, ale skôr iba na jeho výsledok, a teda výskumy usudzovania a rozhodovania majú sklon zameriavať sa na výstupy a presnosť týchto výstupov, rozhodnutí (napríklad, či ľudia dávajú správnu odpoveď) a nie na za nimi prebiehajúce kognitívne procesy [10].

Nová evidencia [9] naznačuje, že pôvod odchýlky nespočíva v samotnom ignorovaní štatistických údajov, ale v nedostatočnej schopnosti potlačiť intuitívne heuristické odpovede založené na stereotypoch. Pre teóriu usudzovania a rozhodovania je preto dôležité neustále zväčšovanie evidencie dôkazov či už na objasnenie povahy heuristických skreslení intuitívneho Systému 1 alebo o zapojení analytických úsudkov Systému 2.

#### 1.4 Kognitívna reflexia

So schopnosťou racionálne a normatívne sa rozhodnúť sa spája aj schopnosť kognitívnej reflexie, pričom vyššia miera kognitívnej reflexie predpokladá racionálne rozhodnutia ľudí. Test kognitívnej reflexie (CRT) [4], ktorý sme v pilotnom výskume použili, skúma práve schopnosť človeka produkovať kognitívne úsilie, zapojenie analytického myslenia a potlačenie intuitívnej odpovede.

#### 1.5 Výskumný cieľ a hypotézy

Vo výskume sme sa zamerali na testovanie kognitívnej reflexie v spojitosti s používaním heuristiky reprezentatívnosti, konkrétne v spojitosti s ignorovaním /

neignorovaním informácie o základnom výskyte. Cieľom pilotného výskumu bolo porovnanie úspešných a neúspešných riešiteľov klasického rozhodovacieho problému „právnik – inžinier“ v skóre CRT (v rozsahu 0 – 3 správnych odpovedí) ako aj v počte intuitívnych odpovedí v tomto teste. Naším predpokladom bolo, že ľudia dopúšťajúci sa omylu základného výskytu (nevhodne aplikujúci heuristiku reprezentatívnosti) dosahujú nižšiu úroveň kognitívnej reflexie a produkujú viacej intuitívnych odpovedí ako ľudia úspešne usudzujúci o základnom výskyte (produkujúci normatívne správne odpovede, založené na štatistických údajoch).

## 2 Metódy

### 2.1 Participanti

Výskumnú vzorku pilotného výskumu tvorilo 55 participantov s priemerným vekom 28,45 rokov (SD = 4,864), ktorí boli oslovení aby na báze dobrovoľnosti odpovedali na úlohy prezentované prostredníctvom krátkeho elektronického dotazníka. Výskumnú vzorku tvorili prevažne absolventi vysokých škôl rôznych odborov. Z hľadiska rodu tvorilo výskumnú vzorku 22 mužov a 33 žien.

### 2.2 Materiály

Účastníkom výskumu, boli pomocou elektronického dotazníka predložené stručné popisy osobnosti niekoľkých jedincov, ktorí boli údajne náhodne vybratí z populácie, pozostávajúcej z istého počtu ľudí (informácia o základnom výskyte).

Inšpirujúc sa výskumom autora De Neys a kol. [9] sme skonštruovali 4 typy problémov základného výskytu na testovanie našej hypotézy:

1. V kontrolnom príklade – **kontrolná úloha** - popis osoby popisoval typického člena väčšej skupiny, a teda stereotypné presvedčenie a základný výskyt navádzali na rovnaké odpovede.
2. V hlavnej výskumnej úlohe – **právnicki a inžinieri, úloha základného výskytu** – bola odpoveď založená na stereotype (informácie popisujúce osobu vybratú z populácie boli v rozpore s odpoveďou vychádzajúcou zo základného výskytu. V troch ďalších úlohách neboli odpovede v konflikte.
3. V prvej neutrálnej úlohe – **neutrálna úloha 1** – bol popis osoby úplne neutrálny, a preto tieto položky nevyvolávali stereotypné, heuristické odpovede a očakávali sme, že odpovede

participantov budú v súlade so štatistickými údajmi o základnom výskyte.

4. Napokon v poslednom príklade – **neutrálna úloha 2** – v ktorom bol neutrálny základný výskyt a nenavádzal na určitú odpoveď, mala reakcia participanta závisieť od stereotypného presvedčenia a popisu osoby.

V tab. 1 predstavujeme prehľadne konkrétne úlohy použité vo výskume v poradí v akom boli prezentované v dotazníku. Úlohou bolo určiť, do ktorej skupiny (napr. právnici, inžinieri) popísaní jedinec s najväčšou pravdepodobnosťou patrí. Štatisticky povedané, je viac pravdepodobné, že náhodne vybraný a opísaný jedinec bude skôr z väčšej, než z menšej skupiny [9].

Tab. 1. Prehľad úloh - 4 typy problémov základného výskytu

<b>Úloha 1: Švédci a Taliani (kontrolná úloha)</b>
Predstavte si populáciu 5 Švédov a 995 Talianov. Maro má 16 rokov. Miluje futbal, po ktorom si často chodí posedieť s priateľmi pri pizzi alebo cestovinách. Ktorá z možností je pravdepodobnejšia?
- Maro je Švéd.
- Maro je Talian
<b>Úloha 2: Právnicki a Inžinieri (base – rate úloha, úloha základného výskytu)</b>
Predstavte si populáciu 5 inžinierov a 995 právnikov. Jozef má 45 rokov, manželku a 4 deti. Neprejavuje záujem o politické a spoločenské dianie a je vo všeobecnosti konzervatívny. Obľubuje rybnárenie a matematické hádanky. Ktorá z možností je pravdepodobnejšia?
- Jozef je inžinier.
- Jozef je právnik.
<b>Úloha 3: Pravičiarci a Ľavičiarci (neutrálna úloha 1)</b>
Predstavte si populáciu 5 pravičiarov a 995 ľavičiarov. Peter má 175 centimetrov, hnedé vlasy a 2 malé deti. Má rád kávu a slané praclíky. Ktorá možnosť je pravdepodobnejšia?
Peter je pravičiar.
Peter je ľavičiar
<b>Úloha 4: Štyridsiatnici a „Sedemnástnici“ (neutrálna úloha 2)</b>
Predstavte si populáciu 500 štyridsiatnikov a 500 sedemnástročných. Róbert žije v Bratislave. Každý deň trávi so známymi, rád sleduje hudobné programy. Je fanúšik kapely Korn a šetrí si na auto. Ktorá možnosť je pravdepodobnejšia?
Róbert má 17.
Róbert má 40.

Pri inštruovaní druhého testu na skúmanie, Testu kognitívnej reflexie, sme predpokladali, že úlohy na zistenie kognitívneho úsilia generujú vo väčšine odpovedí nesprávne „intuitívne odpovede“. Tento fakt je podporený skutočnosťou, že spomedzi všetkých možných nesprávnych odpovedí, ktoré ľudia môžu označiť dominujú v odpovediach práve intuitívne (10, 100 a 24) [4]. Konkrétne úlohy CRT testu opisuje Tab. 2.

Tab. 2. Úlohy Testu kognitívnej reflexie

<p>Pálka a lopta stoja €1,10 dokopy. Pálka stojí o 1,00 € viac ako lopta. Koľko stojí lopta?          ___ centov</p> <p>Na jazere je záhon plávajúceho lekna. Každý deň sa zdvojnásobí veľkosť záhonu. Ak trvá 48 dní, aby záhon pokryl celé jazero, ako dlho by trvalo záhonu, aby pokryl polovicu tohto jazera? ___ dní</p> <p>Ak trvá 5 strojom 5 minút, aby vyrobili 5 vecí, ako dlho by trvalo 100 strojom, aby vyrobili 100 vecí? ___ minút.</p>
--

## 2.3 Výsledky

V testovanom súbore sme porovnávali subjekty podľa úspešnosti/neúspešnosti v jednotlivých úlohách základného výskytu, pričom sme predpokladali, že nami vytvorené typy úloh vyvolajú rovnaký druh skreslenia, chyby (*bias*) ako klasické problémy v uskutočnených experimentoch.

Primárne nás zaujímala úspešnosť riešenia v klasickej úlohe základného výskytu (úloha 2). Omylu základného výskytu sa účastníci mohli dopustiť aj pri úlohe 1 a 3. Úloha 4 bola neutrálna z hľadiska základného výskytu, a preto sme ju do výsledkov nezaradili.

Odpovede v súlade so základným výskytom (na základe posudzovania pravdepodobnosti) boli označené ako úspešné riešenia. Odpovede založené na popise osoby (na stereotype), v prípade že boli v protiklade s informáciou o základnom výskyte, sme kódovali ako neúspešné riešenia úloh. Taktiež sme subjekty porovnávali na základe dosiahnutého skóre správnych odpovedí v CRT a produkovania intuitívnych odpovedí v CRT.

Pri kontrolnej úlohe (Švédci a Taliani, úloha 1) odpovedalo 87,27% účastníkov úspešne a 12,73% neúspešne. Hlavnú úlohu základného výskytu (Právnicki

a Inžinieri) úspešne vyriešilo 70,91% účastníkov a neúspešne 29,09%. Pri tretej úlohe s neutrálnym popisom osoby odpovedalo úspešne 67,27% a neúspešne 32,73% účastníkov (pri všetkých úlohách bol celkový počet N=55).

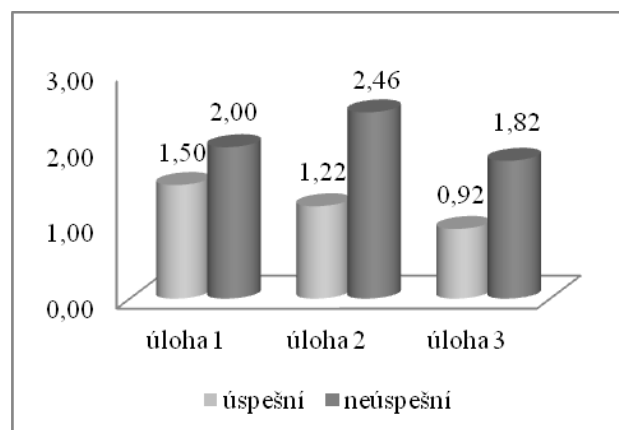
Výsledky testovania normálneho rozloženia hodnôt z CRT (pri pohľade na počet správnych odpovedí) a z CRTint (pri pohľade na počet intuitívnych odpovedí) uvádzame v nasledujúcej tabuľke 3.

Tab.3. Výsledky testu normality CRT a CRT int

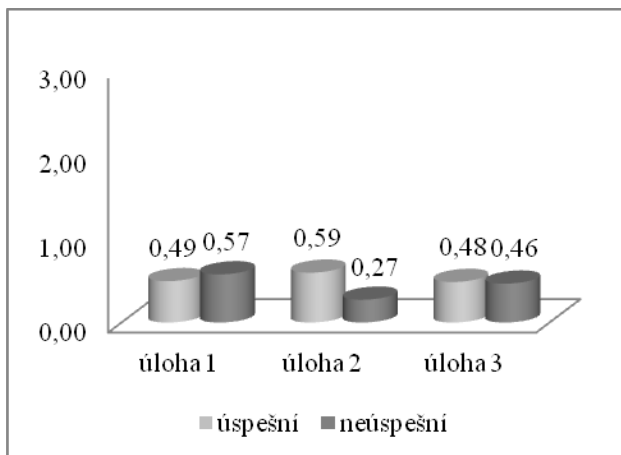
Úloha	CRT		CRT int	
	SW	p	SW	p
1	úspešní	0,845 < 0,001	0,695 < 0,001	
	neúspešní	0,858 0,144	0,600 < 0,001	
2	úspešní	0,856 < 0,001	0,722 < 0,001	
	neúspešní	0,751 0,001	0,587 < 0,001	
3	úspešní	0,848 < 0,001	0,698 < 0,001	
	neúspešní	0,810 0,002	0,669 < 0,001	

Na základe výsledkov normality sme pre ďalšie testovanie použili skóre v CRT teste, ktorého medián odpovedí predstavuje mieru kognitívnej reflexie v rámci jednotlivých skupín (úspešných a neúspešných riešiteľov). V nasledujúcom grafe 1 uvádzame mediánové hodnoty skóre CRT z hľadiska správnych odpovedí a v grafe 2 mediánové hodnoty skóre v CRT z hľadiska intuitívnych odpovedí pri jednotlivých analyzovaných úlohách.

Graf 1 Mediánové hodnoty skóre CRT



Graf 2 Mediánové hodnoty skóre CRTint



Pri porovnaní hodnôt úspešných a neúspešných riešiteľov boli zistené isté štatisticky významné i vecne významné rozdiely mediánov pri hlavnej úlohe základného výskytu (úloha 2) ( $U = 152,00$ ;  $p = 0,002$ ;  $r_m = 0,42$ ) a pri úlohe s neutrálnym opisom osoby náhodne vybranej z ponúkanej populácie (úloha 3) ( $U = 223,00$ ;  $p = 0,040$ ;  $r_m = 0,28$ ). Pri kontrolnej úlohe (úloha 1) sa významné rozdiely nezistili ( $U = 128,00$ ;  $p = 0,326$ ;  $r_m = 0,142$ ). Z hľadiska intuitívnych odpovedí sa významné rozdiely nepotvrdili ani u jednej úlohy ( $U1 = 156,00$ ;  $p = 0,353$ ;  $r_m = 0,05$ ;  $U2 = 244,00$ ;  $p = 0,142$ ;  $r_m = 0,20$ ;  $U3 = 325,00$ ;  $p = 0,875$ ;  $r_m = 0,02$ ).

## 2.4 Diskusia

V pilotnom výskume danej problematiky sme sa snažili poukázať na fakt, že u účastníkov pri riešení klasických úloh zisťujúcich omyl základného výskytu „zvíťazí“ stereotypné presvedčenie nad poskytnutou štatistickou informáciou o základnom výskyte v populácii.

DeNeys a kol. [9] sa vo výskume zamerali na zisťovanie nervového základu ne/rozpoznania konfliktu medzi stereotypnou a štatistickou informáciou a na zisťovanie nervového základu inhibície intuitívnej odpovedi. V uplynulom desaťročí početné štúdie so zobrazovacou technikou preukázali, že odhalenie konfliktu a aktuálna inhibícia určitej odpovede sú sprostredkované dvoma odlišnými regiónmi v mozgu. Výsledky ich výskumu ukázali, že oblasť inhibície intuitívnej odpovede, bola špecificky aktivovaná aj v situáciách, keď sa účastník vyhol stereotypnej odpovedi a oblasť detekcie konfliktu bola aktivovaná, aj keď sa ľudia rozhodli reagovať stereotypne. Výsledky naznačujú, že ľudia boli schopní

odhaliť ich predpojatosť jednať stereotypne, no aj napriek tomu, nedokázali potlačiť intuitívne odpovede a reagovali na základe stereotypného popisu.

Z výsledkov predvýskumu sa nám potvrdilo, že neúspešní riešitelia problémov typu „právnik – inžinier“, odhaľujúcich omyl základného výskytu dosahujú vo všeobecnosti nižšie skóre v CRT teste, čo sa týka počtu správnych odpovedí, no nepotvrdil sa nám fakt, že neúspešní riešitelia uvádzali viac intuitívnych odpovedí. Tieto výsledky preto interpretujeme skôr ako povzbudenie Systému 2, než potlačenie Systému 1, ale vzhľadom na počet účastníkov môže byť zaujímavé podrobiť danú problematiku ďalšiemu skúmaniu.

Zaujímavé výsledky sa ukázali v úlohe 3, kde bola účastníkom poskytnutá informácia o základnom výskyte (populácia 5 pravičiarov a 995 ľavičiarov) a neutrálny popis osoby náhodne vybranej z tejto populácie, pri ktorej sme predpokladali, že neutrálny popis osoby nevyvolá stereotypné, heuristické odpovede a očakávali sme, že odpovede účastníkov budú v súlade so štatistickými údajmi o základnom výskyte. Napriek nášmu predpokladu odpovedalo 32,73% účastníkov neúspešne a popísanú náhodne vybranú osobu s väčšou pravdepodobnosťou priradili ku skupine 5 pravičiarov. Domnievame sa, že aj na základe nami zvoleného neutrálneho vykreslenia osoby, mohla časť účastníkov v popise rozpoznať stereotyp pravičiara, prípadne odpovedali v súlade so svojimi „silnými“ postojmi a preferenciami voči jednotlivým skupinám.

Stereotypy zasahujú do fenoménu usudzovania v mnohých oblastiach, klinických, právnych či sociálno-psychologických. Výskum heuristik [11, 6] poukazuje na to, že v každom okamihu v ktorom sa nachádzame existuje celý rad podnetov, ktoré súťažajú o našu pozornosť. Vzhľadom ku kognitívnym obmedzeniam pri spracovaní všetkých informácií je výhodné a strategické, niektoré rozhodnutia urobiť bez väčšej námahy a pozornosť sústrediť na informácie kľúčové pre dôležité rozhodnutia. Definovanie pojmov a konceptov ako sú predsudky, stereotypy a rasizmus nám pomáha premýšľať o tom, ako sa vzťahujú ku každému z nás a ako ovplyvňujú naše sociálne interakcie. Stereotypy sú vo všeobecnosti definované ako presvedčenia o vlastnostiach, atribútoch a správaní členov určitých skupín [12]. Z kognitívneho hľadiska sú stereotypy vnímané ako druh mentálnej skratky ktorú používame na získanie informácií rýchle a bez námahy (teda za použitia Systému 1).

V pôvodnom výskume [12] účastníci odpovedali na 24 problémov z každého typu úlohy. Tieto úlohy boli

založené na celom rade stereotypov (vek, pohlavie, rasa a podobne) a boli vybrané na základe rozsiahlej pilotnej štúdie. Preto prepojenie na oblasť stereotypov v týchto výskumoch považujeme za prínosné a hodné ďalšieho skúmania.

V prípade, že by sme si chceli odpovedať na našu pôvodnú otázku, či vyhrávajú stereotypy nad štatistikou, navrhujeme pri ďalšom skúmaní zväčšiť výskumnú vzorku a použiť doplnujúce metodiky, ktoré by pomohli vysvetliť jednak stereotypné myslenie ľudí a jednak princíp vnímania štatistických informácií o základnom výskyte.

### PodĎakovanie

Tento príspevok vznikol za podpory Centra excelentnosti SAV „CESTA – Centrum strategických analýz“.

### Literatúra

- [1] K. E. Stanovich: *The Robot's Rebellion: Finding Meaning in the Age of Darwin*, The University of Chicago Press, Chicago, 2004.
- [2] K. E. Stanovich, R. F. West: Individual Differences in Rational Thought. *Journal of Experimental Psychology* 127 (2) (1998) 161- 188.
- [3] D. Kahneman: *Thinking, fast and slow*, Farrar, New York, Straus and Giroux, 2011.
- [4] S. Frederick: Cognitive Reflection and Decision Making. *Journal of Economic Perspectives* 19(4) (2005) 25–42.
- [5] J. St. B. T. Evans: In two minds: dual-process accounts of reasoning. *Trends in Cognitive Sciences* 7(10) (2003) 454–459.
- [6] A. Tversky, D. Kahneman: Judgment under Uncertainty: Heuristics and Biases. *Science, New Series* 185(4157) (1974) 1124-1131.
- [7] D. Kahneman, A. Tversky: On the psychology of prediction. *Psychological Review* 80 (1973) 237–251.
- [8] W. DeNeys: Dual Processing in Reasoning Two Systems but One Reasoner. *Psychological Science* 17 (5) (2006) 428-433.
- [9] W. DeNeys, O. Vartanian, V. Goel: When our brains detect that we are biased. *Psychological Science* 19 (5) (2008)
- [10] J. St. B.T. Evans: On the resolution of conflict in dual process theories of reasoning. *Thinking and Reasoning* 13(2007) 321–339.
- [11] G. Gigerenzer: *Rationality for Mortals*, How people cope with uncertainty. Oxford University Press, New York, 2008.
- [12] R. S. Khan: Stereotyping from the perspective of perceivers and targets. *Online Readings in Psychology and Culture* 5 (2002) 1-13.
- [13] R. Hastie, R.M. Dawes: *Rational Choice in an Uncertain World: The Psychology of Judgment and Decision Making*, Thousand Oaks, SAGE Publication, 2009.

# Je reprezentačný pluralizmus v kognitívnej vede nevyhnutný?

Igor Farkaš

Centrum pre kognitívnu vedu, Fakulta matematiky, fyziky a informatiky, Univerzita Komenského  
Mlynská dolina, 84248 Bratislava  
farkas@fmph.uniba.sk

## Abstrakt

V príspevku sa zamýšľame nad vysvetľovaním povahy kognitívnych procesov ako prejavov mysle. Uvádžeme a porovnávame existujúce teórie a paradigmy v kognitívnej vede z reprezentačno-výpočtového pohľadu, pričom zameriavame pozornosť na diskkrétne a spojité formy reprezentácie, ako aj spojitú a diskkrétne dynamiku. Argumentujeme, že konekcionizmus má najbližšie k vysvetleniu mysle a jej prejavov, avšak nie z dôvodu spojitosti reprezentácií, ale vďaka schopnosti učenia a generalizácie. Pojednávame aj o abstrakcii ako dôležitom aspekte ľudskej kognície, ktorý objasňujeme z pohľadu neurovedy. V závere konštatujeme, že reprezentačný pluralizmus v kognitívnej vede je z epistemologického hľadiska užitočný, najmä v kontexte budovania inteligentných robotických systémov, no možno v budúcnosti nie nevyhnutný, čo sa týka mechanistického pochopenia mysle.

## 1 Úvod

Dlhodobou ambíciou kognitívnej vedy je vysvetliť podstatu ľudskej mysle ako entity, v ktorej prebiehajú kognitívne procesy subjektu. Primárne máme na mysli biologické subjekty (človeka alebo zvieratá), no môžeme uvažovať aj o (umelých) kognitívnych procesoch v súvislosti s artefaktami, napr. robotmi, od ktorých očakávame nejaké prejavy „inteligentného“ správania. Komplikáciou v kognitívnej vede je to, že pojem mysle, na rozdiel od mozgu ako predpokladaného neurálneho substrátu, je vo svojej podstate vágny, a preto dáva priestor pre rôzne pohľady, v rámci ktorých sa ponúkajú odpovede. Napredujúce empirické vedy (psychológia, kognitívna neuroveda a iné) ovplyvňujú tieto paradigmy, pričom sa vychádza z konsenzu, že empirické poznatky sú zdrojom informácie pre modifikáciu vedeckých teórií, aj keď to sťažujú prípady nejednoznačnej interpretácie výsledkov experimentu.

Napriek komplikáciám sa ako dominantný prístup k poznávaniu mysle vyprofiloval reprezentačno-výpočtový pohľad [37] založený na hypotéze, že v mysli existujú *mentálne reprezentácie* (analógie dátových štruktúr), nad ktorými sa realizujú nejaké operácie (výpočty). Týmto sa však

len presunulo bremeno vysvetľovania, pretože ani pojmy reprezentácie a počítania nie sú jednoznačne interpretované. Do akej miery tieto reprezentácie a výpočty súvisia s mozgom, závisí od paradigmy, ktorú uznávame. V každom prípade však môžeme predpokladať to, že kognitívne procesy prebiehajú v čase a v nejakom priestore reprezentácií. Potom sa môžeme pýtať, či ten čas a priestor sú vo svojej podstate diskkrétne alebo spojité veličiny. Inak povedané, či je myseľ diskkrétne alebo spojitá, alebo aj-aj. Reprezentačno-výpočtový prístup je akceptovaný v rôznych výpočtových paradigmách kognitívnej vedy (spomínaných ďalej). Výpočtové modelovanie považujeme za nepostrádateľnú súčasť mechanistického poznávania mysle, pretože okrem snahy o vysvetlenie ľudskej kognície tieto paradigmy možno využiť aj pri konštruovaní inteligentných (robotických) systémov, čím sa rozširuje horizont kognitívnej vedy [7].

Koexistencia rôznych vysvetlení kľúčových pojmov (reprezentácia, počítanie) vedie k určitému pluralizmu. V tomto príspevku sa zamýšľame nad otázkou pluralizmu z pohľadu umelej inteligencie (UI), ale aj psychológie a neurovedy. Ďalej budeme argumentovať, že reprezentačný pluralizmus je síce užitočný (v UI), no z hľadiska ontologickej podstaty ľudskej mysle nie nevyhnutný. Zdôvodníme, v čom tkvie výhoda konekcionistickej paradigmy, ktorá má spomedzi existujúcich prístupov najlepšie predpoklady na najkompletnejší opis kognitívnych procesov.

V tomto príspevku najprv stručne sumarizujeme hlavné teórie (časť 2) a hlavné výpočtové paradigmy v kognitívnej vede (časť 3), a tie potom navzájom porovnáme (časť 4). Ďalej zameriame pozornosť na porovnanie dvoch dynamík formálnych systémov (časť 5). Spomenieme zopár argumentov v kontexte rozdielnosti medzi symbolizmom a konekcionizmom (časť 6) a napokon pojednáme o neurovednom pohľade na abstrakciu ako dôležitú vlastnosť kognície (časť 7). Záver je stručným zhrnutím našej argumentácie (časť 8).

## 2 Teoretické smery v kognitívnej vede

Prvý teoretický smer – *symbolizmus* (kognitivizmus) bol podnietený zrodom konceptu Turingovho stroja ako hypo-

tetického univerzálneho počítača, a vynájdením moderných digitálnych počítačov. Základným konceptom sú diskkrétne výpočty so *symbolmi*. Pod pojmom symbol máme na mysli *amodálny* symbol – konštrukt zavedený v kognitívnej psychológii na označenie reprezentácie nezávislej od modality (vstupných senzorických podsystemov). Symbolový pohľad na myseľ je dodnes považovaný za tú správnu úroveň opisu najmä v kontexte vysvetľovania procesov tzv. vyššej kognície, ako je usudzovanie, plánovanie, a tiež používanie prirodzeného jazyka (napr. [8, 25, 30], kde kognitívne procesy sú oddelené od perceptuálnych a motorických procesov. Mentálne operácie sú realizované pomocou vnútorného jazyka mysle, napr. mentálčiny [8]. Výhodou symbolizmu je, že poskytuje silné matematické a logické formalizmy, ktoré sú zväčša transparentné a preto človeku zrozumiteľné. Pre človeka je prirodzené uvažovať v termínoch diskrétnych podmienok, pravidiel, propozícií a logických inferencií.

V kognitívnej psychológii sa v 70. rokoch minulého storočia začali objavovať skeptické pohľady na centrálnosť (vnútorných) reprezentácií a symbolového spracovania informácie v ľudskej kognícii a začali vznikať alternatívne teórie. K takýmto smerom patrí ekologická psychológia [10], ktorá zdôrazňovala, že „nie je dôležité, čo je vo vnútri hlavy, ale vnútri čoho (akého prostredia) sa hlava nachádza.” Tento teoretický smer plynulo prešiel od 80. rokov do teórie situovanej kognície, ktorá bola postavená na predpokladoch, že ľudská inteligencia je fundamentálne interaktívna a fundamentálne neoddeliteľná od kontextu [13].

Dôraz na interakciu s prostredím ostal v centre pozornosti aj naďalej, no v 90. rokoch sa začal klásť dôraz na reprezentácie, ktoré túto interakciu umožňovali, a to v rámci *stelesnenej kognície*. Existuje viacero pohľadov na stelesnenú kogníciu [40]. Barsalou navrhol zastrešujúci termín *ukotvená kognícia* [2], ktorý zjednocuje rôzne aspekty stelesnenia a situovania v prostredí, vrátane konceptu tzv. rozšírenej mysle (extended mind). Silnú podporu pre túto paradigmu v kognitívnej vede možno nájsť v narastajúcej empirickej evidencii (za ostatných 20 rokov), či už z behaviorálnych štúdií alebo zobrazovacích metód (pozri referencie v [2]). Táto evidencia napovedá, že všetky kognitívne funkcie (vrátane tých vyšších) sa do istej miery „opierajú” o (nízkoúrovňové) senzomotorické procesy. Inými slovami, neurálne reprezentácie vyšších a nižších kognitívnych procesov majú prienik. Tieto teórie tiež pracujú s pojmom symbol, avšak v zmysle modálnych symbolov. Napríklad, vplyvná teória tzv. perceptuálnych symbolov [1] postuluje existenciu symbolov, ktoré sú realizované pomocou multimodálnych reprezentácií, a ktoré v podstate zodpovedajú pojmom (mentálnym kategóriám).

### 3 Výpočtové paradigmy v kognitívnej vede

V priebehu zhruba 60 rokov existencie kognívnej vedy sa v nej vyprofilovali štyri výpočtové paradigmy: symbolová, konekcionistická, dynamická a pravdepodobnostná<sup>1</sup> (v takomto poradí vzniku). Tieto paradigmy sa vzájomne líšia predpokladanými reprezentáciami a spôsobmi výpočtu.

Pojem reprezentácie sa bežne používa v kognitívnej psychológii, lingvistike i UI. Jazykový pôvod tohto pojmu napovedá, že ide o re-presentovanie niečoho vonkajšieho (vo svete) niekde inde, vo vnútri nejakého systému (živého alebo umelého) alebo i na papieri. V kognitívnej psychológii sa hovorí o mentálnych reprezentáciách, čo je v podstate prvý teoretický konštrukt v kognitívnej vede [26]. Korene pojmu mentálna reprezentácia však siahajú až do antických čias, keď širšie chápanie tohto pojmu medzi filozofmi nemalo výpočtový charakter. Tradičná polemika v kognitívnej vede sa týka povahy týchto reprezentácií, čo reflektujú aj jednotlivé paradigmy kognitívnej vedy.

**Symbolizmus** je konceptuálne úzko spätý s digitálnym počítačom, ktorý realizuje diskkrétne výpočty so symbolmi. Výpočet v počítači prebieha za pomoci dvoch kľúčových komponentov: procesora a pamäte. Interakcia počítača s prostredím je na periférii záujmu a prebieha prostredníctvom vstupno-výstupných podsystemov. Procesor sériovým spôsobom spracováva symboly, uložené v pamäti a vykonáva pritom inštrukcie podľa programu uloženého v inej časti pamäte. Klasická paradigma konceptualizuje myseľ ako výpočtový stroj, oddeliteľný od prostredia, ktorý manipuluje s internými symbolmi, odvodenými (zvonku) pomocou transdukcie z prostredia, podľa logických pravidiel. Výpočtovú teóriu mysle vystihuje hypotéza o fyzikálnom symbolovom systéme, ktorý „disponuje nutnými a postačujúcimi prostriedkami na všeobecné inteligentné konanie” [19]. Koncept univerzálneho Turingovho stroja je síce dôležitý v kontexte abstraktného počítania, avšak nie je relevantný pre UI a kognitívnu vedu, pretože neprispel k riešeniu praktických úloh UI, a ani k tomu, ako pracuje mozog [32].

**Konekcionizmus** spochybňuje základný predpoklad tradičnej UI, že mentálne procesy možno najlepšie charakterizovať ako algoritmické manipulácie so symbolmi. Konekcionizmus však nie je homogénnou skupinou, ale spektrom metód, ktoré boli inšpirované architektúrou a fungovaním mozgu (inšpirácia „zdola”). Sila konekcionistického systému – *umelej neurónovej siete* – nie je v samotných neurónoch, ale v ich vzájomných (excitačných a inhibičných) prepojeniach a interakcii. Paralelné spracovanie a

<sup>1</sup>V literatúre sa pravdepodobnostný prístup (verím, že zatiaľ) neobjavuje ako samostatná paradigma. Domnievam sa však, že je dostatočne principiálny a odlišný od ostatných prístupov, preto si nálepku paradigma zaslúži.



distribučnosť aktivity<sup>2</sup> predstavujú základný rozdiel v architektúre v porovnaní so symbolovým systémom, pretože každý neurón je súčasne procesorom aj pamäťou (aj keď elementárnou). Taktiež, povaha komunikácie medzi neurónmi má numerický a nie symbolový charakter, preto hovoríme v prípade sietí o *subsymbolorých* reprezentáciách. Neurónová sieť v podstate realizuje nelineárne vektorové operácie v metrickom (euklidovskom) priestore. Je zrejme, že takéto operácie a reprezentácie sú v porovnaní so symbolovými oveľa menej transparentné (sieť ako „čierna skrinka”). Našťastie existujú techniky zhlukovania a vizualizácie mnohorozmerných dát, vďaka ktorým môžeme zisťovať, čo sa v neurónovej sieti deje [20].

**Dynamický prístup** je charakteristický úzkym prepojením agenta na prostredie (na rozdiel od symbolizmu a konekcionizmu, pozri aj [6]), v procese permanentnej vzájomnej interakcie, ktorá prebieha v spojitom priestore a čase, a ktorá sa dá najlepšie opísať pomocou diferenciálnych rovníc [27]. Dôraz sa teda kladie na situovanosť a stesnenosť ľudského správania [29], čo je intuitívne zrejme hlavne pri vysvetľovaní bezprostrednej senzomotorickej interakcie agenta s okolím. Dynamický prístup má tiež svoje vnútorné členenie. Radikálna dynamická platforma popiera akékoľvek reprezentácie, zatiaľ čo mäkkšia platforma ich nevyklucuje, čiže je konzistentná s reprezentačno-výpočtovým pohľadom. Dynamický prístup ku spoznávaní vonkajšieho sveta je konzistentný so zjednávacím (enactive) prístupom [35], kde sa dôraz z vnútorných reprezentácií (vopred daného) vonkajšieho sveta presúva na vnímanie a jednanie vo svete, ktorý sa takto spoluvytvára. Vytráca sa dichotomické delenie na subjekt a objekt poznania.

**Pravdepodobnostný prístup** sa stal populárny v kognitívnej vede najmä v ostatnej dekáde rokov [23]. Ide o teoreticky podložený (bayesovský) prístup, ktorý využíva široké spektrum reprezentácií (stromy, vektory, logické pravidlá atď.), kombinuje ich so štatistickým učením a inferenciami za prítomnosti neurčitosti. Ponúka vysvetlenie rôznych prejavov ľudského správania (pozri [11] a tamojšie referencie). Postupuje zhora, od funkcie ktorú chceme vysvetliť, pričom sa hľadajú optimálne reprezentácie dát. Taktiež zahŕňa aspekt vrodené–získané v podobe tzv. indukčných predispozícií (inductive biases), čo sú v podstate apriórne distribúcie veličín na množine hypotéz (vrodené), ktoré vstupujú do výpočtu posteriorných distribúcií (získané). Jedným z problémov pravdepodobnostného prístupu je však jeho výpočtová neúnosnosť v prípade zložitejších problémov [18].

<sup>2</sup>Aj keď treba upresniť, že tzv. lokalistické modely neurónových sietí (teda nie distribuované) pracujú s reprezentáciami, ktoré možno nazvať symbolovými.

Úrovně analýzy	Symb	Kon	Dyn	Pr
výpočtová	+	+	+	+
algoritmická	+	+	+	+
implementačná	–	+	–	–

**Tabuľka 1:** Relevantnosť úrovní analýzy v jednotlivých výpočtových paradigmatách kognitívnej vedy (symbolovej, konekcionistickej, dynamickej, a pravdepodobnostnej).

## 4 Porovnanie paradigiem

V snahe porovnať jednotlivé paradigmy sa na ne môžeme pozrieť z pohľadu teórie troch nezávislých úrovní analýzy (vysvetlenia nejakého fenoménu), a to výpočtovej, algoritmickéj a implementačnej, ktorú Marr [17] rozpracoval v kontexte modelovania vizuálneho spracovania informácie, a ktorá výrazne ovplyvnila kognitívnu vedu. *Výpočtová úroveň* definuje výpočty, ktoré treba vykonať, napríklad pomocou matematickej funkcie alebo špecifikácie úlohy. *Algoritmická úroveň* špecifikuje použité reprezentácie uchovávané informácie, a výpočty s nimi. *Implementačná úroveň* špecifikuje procesy spracovania informácie, ktoré sú viazané na konkrétny hardvér, ktorý je použitý na implementáciu. Medzi úrovňami platí vzťah tzv. viacnásobnej realizovateľnosti, čo znamená, že jeden opis na vyššej úrovni sa dá transformovať na viacero opisov na nižšej úrovni.<sup>3</sup>

Marrova teória bola silne ovplyvnená symbolizmom, ktorý implementačnú úroveň považuje za nepodstatnú. Toto je pochopiteľné v prípade počítača, ktorý je naozaj duálnou entitou s nezávislým hardvérom a softvérom, kde algoritmy sa menia na implementáciu úplne automatickým procesom kompilácie (t. j. prekladu do hardvérovo-závislého strojového kódu). Avšak, v prípade mozgu nemožno neurálnu implementáciu automaticky odvodiť z nejakého opisu na vyššej úrovni. Mozog nebol nikým skonštruovaný, ale sa evolvoval tak, aby umožnil organizmu efektívne konať v dynamickom prostredí [3]. Implementačná úroveň hrá dôležitú úlohu a mala by byť vhodne spojená s opisom na vyššej úrovni, aby nám to uľahčilo interpretáciu jej funkcie.

Jednotlivé paradigmy možno porovnať, čo sa týka relevantnosti jednotlivých úrovní analýzy (viac v [7]). Porovnanie znázorňuje tabuľka 1. Je vidieť, že iba konekcionistická paradigma zahŕňa všetky tri úrovne. Symbolizmus a pravdepodobnostný prístup zjednocujú dva aspekty. Po prvé, oba prístupy ponúkajú vysvetlenia na výpočto-

<sup>3</sup>Ako ilustračný príklad uvažujme násobenie dvoch viacciferných čísel. To je teda cieľ výpočtu, ktorý sa dá dosiahnuť rôznymi algoritmami, napríklad takým, ktorý človek bežne používa, a v ktorom sa medzivýsledky násobenia (jedného čísla číslicou druhého čísla) zapíšu pod seba a potom sa sčítajú. Napokon, implementačná úroveň už predstavuje konkrétnu realizáciu tohto algoritmu v nejakom fyzickom médiu (počítač, pero a papier a i.).

Distribučnosť	Učenie je silnou stránkou	
	–	+
–	Symbolizmus	Pravdep. modely
+	Dynam. systémy	Konekcionizmus

**Tabuľka 2:** Porovnanie paradigiem z pohľadu vybraných charakteristík.

vej úrovni, ktorú nepovažujú len za abstrakciu od inherentných mechanizmov (pravdepodobne neurálnych), ale za nezávislú úroveň vysvetlenia nejakého fenoménu. Po druhé, oba prístupy sú symbolové, no líšia sa v tom, že pravdepodobnostné modely používajú spojité reprezentácie (pravdepodobnosť). Konekcionizmus a dynamické systémy predstavujú subsymbolové prístupy, pričom je medzi nimi súvis (rekurentné neurónové siete sú príkladmi dynamických systémov) [6, 24]. Porovnanie paradigiem z pohľadu uvedených charakteristík ponúkame v tabuľke 2. Tu máme na mysli to (1) či učenie je silnou stránkou paradigmy, a (2) či paradigma využíva lokalistické alebo distribuované reprezentácie.<sup>4</sup> Z tabuľky je vidieť, že aj z pohľadu týchto dvoch charakteristík sú neurónové siete prítupom symbolizmu. Za kľúčovú vlastnosť neurónových sietí považujeme (1) *schopnosť zovšeobecnenia* (vďaka distribuovaným reprezentáciám) a (2) *schopnosť učenia*, t. j. realizáciu elementárnych zmien v znalosti systému (dlhodobej pamäti). V komplexných úlohách nie je principiálne možné vložiť všetky znalosti do systému; ten sa ich však môže naučiť (čo vo všeobecnosti nie je vôbec ľahké).

Aj keď každá z paradigiem má svoje špecifiká, najsignifikantnejší rozdiel sa týka *priepasti medzi symbolovými výpočtovými prístupmi a konekcionizmom* (kde sú kľúčové rekurentné neurónové siete schopné pracovať sekvencne). Pravdepodobnostný prístup, ako symbolová inferencia v spojitom priestore, stojí niekde medzi nimi. Dynamické prístupy majú blízko ku konekcionizmu.

Symbolizmus a konekcionizmus predstavujú kvalitatívne odlišné alternatívy, aj niektorí autori vidia oba prístupy ako zlučiteľné, figurujúce na rôznych úrovniach abstrakcie. Táto polemika sa ťahne de facto od 80. rokov minulého storočia (napr. [8, 33]). Oba tábory hovoria o vzájomnej nekompatibilite oboch prístupov, no na základe odlišných argumentov. Symbolisti tvrdia, že buď (1) oba prístupy sú nekompatibilné, alebo (2) konekcionistické modely sú len *implementáciou symbolových modelov* (implementačný konekcionizmus). Na druhej strane, podľa konekcionistov sú oba prístupy nekompatibilné preto, lebo *symbolové systémy nedokážu implementovať neurónové siete* vo všetkých prí-

<sup>4</sup>Pre vysvetlenie dodajme, že pravdepodobnostné modely síce pracujú aj s distribuovanými reprezentáciami, avšak aktualizácia pravdepodobností hypotéz sa deje na symbolovej úrovni.

padoch. Tento konekcionistický argument sa týka dynamiky systému, ktorú si vysvetlíme.

## 5 Subsymbolová a symbolová dynamika

V oboch paradigmách sa na kognitívne procesy pozeráme ako na deterministické sekvenčné procesy prebiehajúce v čase.<sup>5</sup> Symbolový systém spracováva symboly v čase podľa svojho programu, podobne ako dynamický systém sa vyvíja v stavovom priestore, hnaný aktivačnou dynamikou. Rozdiely medzi oboma paradigmami spočívajú v štruktúre stavového priestoru. Z hľadiska argumentácie nám stačí zamerať pozornosť na (stavový) priestor, v ktorom existujú reprezentácie (predpokladáme diskretný čas). Základná otázka znie: Sú oba typy dynamík ekvivalentné, t. j. môžeme jednoznačne transformovať jednu na druhú? Alebo je symbolová dynamika abstrakciou, ktorá nie vždy dokáže opísať zložitosť dynamických dejov?

**Deterministický dynamický systém** možno opísať časovo-závislým stavovým vektorom  $x(t)$ , ktorý sa vyvíja podľa deterministickej diferenčnej rovnice a vytvára tak trajektóriu (presnejšie povedané, množinu bodov v prípade diskretného času) v stavovom priestore  $X$ , čiže množine všetkých možných stavov. Napríklad, neurónová sieť s  $n$  sigmoidálnymi neurónmi na rekurentnej vrstve, t. j. s aktivačnou funkciou  $f(z) = 1/(1 + \exp(-z))$ , má stavový priestor  $X = [0, 1]^n \subset R^n$ , t. j. v tvare  $n$ -rozmernej kocky. V každom čase aktivita siete je reprezentovaná bodom v tomto stavovom priestore.

Je známe, že aj vo veľmi jednoduchom nelineárnom dynamickom systéme, môže vzniknúť veľmi zložitý správanie (deterministický chaos), ktoré sa môže výrazne (kvalitatívne) meniť aj pri minimálnych zmenách parametrov. Príkladom je logistická mapa daná rovnicou  $x(t+1) = r \cdot x(t) \cdot (1 - x(t))$ , kde  $x(t) \in (0, 1)$  a  $r \in \mathcal{R}$  je parameter. Hodnoty  $x(t)$  sa ustália na rôznych hodnotách (nazývaných atraktory) práve v závislosti od  $r$ . Tieto ustálené hodnoty sa menia od jednej, cez dve, štyri hodnoty, až po chaotický režim v podobe tzv. podivných atraktorov (pri hodnote  $r = 3.5699456\dots$ ). Chaotické systémy sú v prírode veľmi rozšírené a majú tú podstatnú vlastnosť, že ich vývoj v čase je v princípe nepredikovateľný v dlhšom horizonte, a to kvôli extrémnej závislosti od počiatočných podmienok (v literatúre známy ako „motýlí efekt“).

**Symbolová dynamika** vznikne zo subsymbolovej tak, že stavový priestor rozdelíme na konečný počet neprekrývajúcich sa oblastí, ktorých zjednotenie tvorí celý priestor  $X$ . Vývoj systému potom môžeme opísať pomocou (konečnej) sekvencie symbolov, ktorá opisuje prechody medzi jednotlivými oblasťami. Napríklad v prípade logistickej mapy by sme interval  $(0,1)$  mohli rozdeliť napoly

<sup>5</sup>Teraz neuvažujeme stochastické modely neurónových sietí.

a prechody medzi oblasťami (alebo v rámci nich) opísať rôznymi symbolmi.

### 5.1 Sú obe dynamiky ekvivalentné?

Vzniká tu otázka, či opis na úrovni symbolovej dynamiky je rovnako presný a úplný ako ten na úrovni aktivačnej dynamiky. Inými slovami, ak sa rozhodneme opísať dynamický systém pomocou symbolovej dynamiky (použitím vhodnej partície), či sa nejaká informácia o vývoji systému stratí alebo nie. Tu je dôležitý koncept *topologickej ekvivalencie* medzi oboma priestormi. Dva priestory sú topologicky ekvivalentné práve vtedy, keď jeden môžeme spojito transformovať na ten druhý, a naopak (v matematike sa tomu hovorí homeomorfizmus).

Podmienkou pre zabezpečenie topologickej ekvivalencie je vytvorenie tzv. *generujúcej partície* (generating partition), ktorá má kľúčovú vlastnosť, že existuje jednoznačný vzťah medzi trajektóriami v stavovom priestore a sekvenciami symbolov (matematické vyjadrenie možno nájsť napr. v [12]). Problémom je to, že skonštruovať generujúcu partíciu vieme len v jednoduchších prípadoch. Doteraz bolo navrhnutých niekoľko algoritmov ako z pozorovaných dát (nemusíme poznať dynamiku, t. j. model systému) odhadnúť generujúcu partíciu, a to v prípade o niečo zložitejších chaotických dynamických systémov v 2D (pozri napr. [12] a tamojšie referencie). V prípade vysokorozmerných systémov však riešenia neexistujú. Z uvedeného teda vyplýva, že dynamické systémy sú výpočtovo silnejším formalizmom než symbolová dynamika. Subsymbolová dynamika má super-Turingovu silu, vďaka ktorej dokáže generovať dynamické správania, ktoré nie sú dosiahnuteľné pomocou symbolovej dynamiky [31, 36].

Avšak, položme si však otázku: *Je tento záver podstatný pre pochopenie mysle?* Niektorí sa nazdávajú, že áno, kvôli nutnosti vedieť vysvetliť niektoré „jemňôsky“ pri učení, či kognitívnom vývine, napríklad pri akvizícii gramatiky jazyka [36]. Áno, potrebujeme vysvetliť jednotlivé štádia kognitívneho vývinu, ale na to netreba nekonečnú presnosť v reprezentáciách. Toto naše tvrdenie podporíme niekoľkými argumentami:

(1) Neuróny pracujú v šume a vykazujú spontánnu aktivitu pálenia. Aj keď existujú rôzne teórie neurálneho kódu, na každom neuróne pravdepodobne nezáleží [28]. Tento neurálny šum nevnímame, a preto naša myseľ nepotrebuje nekonečnú presnosť.

(2) Introspektívne si človek uvedomuje len svoje mentálne stavy, ktoré sa menia oveľa pomalšie (rádovo stovky milisekúnd) než neurálne stavy (milisekundy). Okrem toho, vychádzame z predpokladu viacnásobnej realizovateľnosti.<sup>6</sup>

<sup>6</sup>Rôzne neurálne stavy znamenajú ten istý mentálny stav.

(3) Mozog pracuje zväčša v neautonómnom režime (prijíma externé vstupy a vysiela výstupy do prostredia), zatiaľ čo závery matematických dynamických modelov sa týkajú autonómneho režimu. Analýza neautonómnych systémov je nesmierne obtiažna. Viacero prác z oblasti modelovania komplexných systémov prichádza s hypotézou, že činnosť mozgu sa pohybuje na hranici chaosu (on the edge of chaos), vďaka čomu si pravdepodobne zachováva obrovskú komplexnosť a variabilnosť, ktorú má ešte poznávajúci subjekt pod kontrolou (napr. [15]).

## 6 Ešte zopár argumentov

Napriek tomu, že pomocou symbolových výpočtov by sme vedeli opísať (z hľadiska požadovanej presnosti) všetky dynamické deje, je užitočnejšie uvažovať o spojitý mysli, pretože spojitost možno z matematického hľadiska vnímať ako všeobecnejší prípad.<sup>7</sup> V prospech poznávania spojitý mysle (ako dynamického systému) hovoria aj empirické poznatky z kognitívnej psychológie [34].

Na oba prístupy sa možno pozrieť aj z perspektívy úrovne opisu. Skok od spojitého koneccionizmu k diskrétnemu symbolizmu však znamená zmenu kvality. V literatúre sa objavujú argumenty v prospech ekvivalencie medzi oboma úrovňami, preto k nim zaujmeme stanovisko.

*Argument 1: Dynamické systémy možno simulovať na digitálnom počítači.* Áno, všetko, čo sa dá formalizovať (algoritmizovať) vieme simulovať na digitálnom počítači s vysokou presnosťou (s výnimkou chaotickej dynamiky, kde by sme potrebovali nekonečnú presnosť počítača). Na počítači vieme simulovať diskkrétne aj spojité procesy, to však nič nehovorí o ontologickej podstate týchto procesov. Počítač tu figuruje len ako simulačný prostriedok.

*Argument 2: Dynamické systémy možno opísať pomocou výpočtov.* K odpovedi na tento argument je nutné objasniť pojem počítania. Klasické počítanie sa týka diskrétnych systémov, neklasické počítanie zahŕňa aj systémy so spojitým priestorom stavov (napr. neurónové siete so sigmoidálnymi aktivačnými funkciami). Dynamické systémy sú teda príkladom neklasických výpočtov, ktoré vieme simulovať pomocou diskrétnych výpočtov. Viac o počítaní možno nájsť v prácach [24, 7].

*Argument 3: Ľudská myseľ je tak komplexná, že je nutné oprieť sa o reprezentačný pluralizmus* [5]. Nazdávam sa, že pri súčasnom stave poznania je pluralizmus najschodnejšou cestou (hybridné systémy sú v UI dominantné [38]). Z evolučného a neurovedného hľadiska za pravdepodobnejšie považujeme jednotné vysvetlenie, napríklad v duchu tzv. kognitívnej inkrementálnosti [4].

<sup>7</sup>Na spracovanie diskrétnych dát im stačí v neurónovej sieti priradiť symbolové, lokalistické reprezentácie, kde v každej reprezentácii je len jeden neurón aktívny.

*Argument 4: Aj v iných vedách, napríklad vo fyzike sa nakoniec dospelo k viacerým vysvetleniam (napr. Newtonova fyzika a kvantová fyzika). Prečo by to tak nemohlo byť aj v prípade kognitívnej vedy? Na rozdiel od fyziky, ktorá vysvetľuje podstatu reality týkajúcu sa rôznych priestorových škál, sa v prípade mozgu nezdá, že máme do činenia s tak rozdielnymi škálami. Áno, môžeme hovoriť o molekulárnej úrovni na jednej strane a systémovej úrovni na strane druhej. Nazdávam sa, že vysvetlenie kognitívnych procesov si vyžaduje systémovú úroveň (napr. predpokladané kvantové fenomény v mozgu považujeme za irelevantné pre vysvetlenie mysle; pozri aj [16]),*

## 7 Abstrakcia z pohľadu neurovedy

Na spektrum ľudských schopností sa môžeme pozrieť aj z pohľadu abstrakcie. Vysokoúrovňové procesy ako napríklad rozmýšľanie o matematickej rovnici vnímame ako viac abstraktné, než uchopenie pohára do ruky. Abstraktné veci sa bežne spájajú s využitím symbolov, zatiaľ čo konkrétne veci už menej. Tradičný pohľad v kognitívnej vede je typicky dichotomický. Ako sa na abstrakciu pozerá neuroveda a konekcionizmus? Pojem abstrakcie má viacero príbuzných významov. V našom kontexte môžeme povedať, že abstrakcia je daná *invariantnosťou*. Zhruba povedané: čím je reprezentácia aktivovateľná väčším spektrom vstupných podnetov, tým je abstraktnejšia. Abstrakcia je kľúčová vlastnosť kognície. Umožňuje nám kategorizovať vstupy, abstrahujúť od detailov, všímaním si len podstatných charakteristík.

Symbolový a subsymbolový pohľad môžeme vnímať ako stojace na opačných koncoch pomyslenej dimenzie abstrakcie, ktorú považujeme za kontinuum. Subsymboloré reprezentácie sú distribuované, existujú v podobe numerických vektorov v euklidovskom priestore s definovanou metrikou. Sú aktivované vtedy, keď vstupný podnet je dostatočne podobný (v zmysle danej metriky). Symbolové reprezentácie stoja na vrchole abstrakcie, sú kategorické. Sú aktivované pri širokom spektre vstupných podnetov, čiže sú vysoko invariantné voči charakteristikám podnetu.

Ako tieto dva odlišné svety premostiť? Človek predsa dokáže pracovať s oboma typmi entít. Ako je to s pojmami? V ľudskej kognícii ich existuje celé spektrum, od veľmi konkrétnych (napr. pes) až po vysoko abstraktné (napr. demokracia), ktoré si človek dokáže osvojiť, aj keď nie naraz, ale v postupných štádiách vývinu. Je známe, že dieťa si najprv osvojuje konkrétne veci a až neskôr abstraktnejšie, čo je podmienené maturationými procesmi prebiehajúcimi v mozgu. Abstrakcia sa týka nielen sveta objektov ale aj priestoru udalostí či akcií. Rozhodnutie čoho sa napiť je zrejme jednoduchšie (konkrétnejšie) než rozhodnutie, na akú vysokú školu sa prihlásiť.

Neuróny v mozgu figurujú na rôznych vrstvách hie-

rarchie. Majú však spoločné to, že každý neurón spracováva nejaký mnohorozmerný vstup a v závislosti od svojich parametrov (synaptických váh) reaguje naň alebo nie. Narastajúca abstrakcia, čo sa týka aktivovania neurónu súvisí s jeho invariantnosťou voči niektorým charakteristikám vstupov (ktoré tiež môžu byť viac alebo menej abstraktné). Typickou vlastnosťou mozgu sú asociačné oblasti, kde dominujú *multimodálne neuróny*, ktoré reagujú na podnety z viac ako jednej modalít. Abstrakcia vzniká pomocou kaskády nelineárnych operácií medzi vrstvami. Na rozdiel od amodálnych symbolov (v symbolových systémoch), tento princíp *nevyžaduje proces transdukcie a abstrakcia vzniká postupne v procese učenia*.

Existujú početné poznatky o organizácii mozgu, čo sa týka invariantnosti neurónov pri reakcii na rôzne podnety. Napríklad vizuálny systém cicavcov je organizovaný hierarchicky v zmysle extrakcie príznakov (od jednoduchších ku zložitejším), čo sa dosahuje zväčšovaním rádiu receptívneho poľa neurónov (afarentných spojení) smerom k vyšším vrstvám. Na vrchole pyramídy stojí area IT (inferior temporal) ako oblasť invariantného rozpoznávania objektov (voči pozícii, rotácii, či škále, a v prípade biologických objektov i deformácie) [14]. Podobne area STS (Superior Temporal Sulcus) má svoje členenie v kontexte rozpoznávania biologických pohybov.

Analogicky je to v prípade exekutívnej funkcie mozgu – konania. Predný lalok je hierarchicky organizovaný (od primárnej motorickej kôry smerom k prefrontálnej kôre), pričom jednotlivé časti sa podieľajú na rozhodovacích procesoch s rôznym stupňom abstrakcie [21]. Najvyššie v pyramíde stojí predná časť prefrontálnej kôry. Pozoruhodné je aj to, že jednotlivé časti frontálnej kôry majú analogickú štruktúru (projenie s talamom a niektorým z jadier v bazálnych gangliách, ktoré sa podieľajú na výbere akcie [22]). Ukazuje sa teda, že podobné (alebo rovnaké) neurálne mechanizmy operujú na rôznych úrovniach hierarchie (abstrakcie).

Aj organizácia dlhodobej pamäti v mozgovej kôre má hierarchickú organizáciu [9], v rámci ktorej existujú vzájomné prepojenia v prednej a zadnej časti kôry umožňujúce vybrať multimodálny obsah a aktivovať ho v podobe pracovnej pamäti.

Ako si teda ľudský mozog poradí s rôznymi entitami, čo sa týka miery abstrakcie? Je známe, že kôra má pomerne homogénnu štruktúru (kortikálne stĺpce) naprieč lalokmi, aj keď pravdepodobne existujú rozdiely v distribuovanosti reprezentácií v jednotlivých lalokoch mozgu. Ako argumentujú O'Reilly a spol. [22], v posteriórnej kôre sú neurálne reprezentácie viac distribuované, čo umožňuje vznik multimodálnych asociácií a dobrej generalizácie, zatiaľ čo vo frontálnej kôre sú reprezentácie riedke (sparse), aby sa zabránilo nežiaducim interferenciám (napr. pri aktivovaní viacerých alternatív pri rozhodovaní). Riedke reprezentá-

cie existujú aj v hipokampe, ako centre podieľajúcim sa na tvorbe epizodickej pamäti.

Neurovedné poznatky sú využiteľné pri modelovaní abstrakcie pomocou umelých neurónových sietí. Tento opis bude kvalitatívne rovnaký či už ide o psa alebo demokraciu. Všetky koncepty sú v mozgu reprezentované podobne ako vzorce aktivít, s rôznou mierou distribuovanosti, a v rôznych častiach mozgu [39]. Preto aj odpovedajúce konekcionistické reprezentácie budú mať podobné vlastnosti, v rámci jednotného prístupu. O niečo zložitejšie budú výpočtové operácie nad týmito reprezentáciami, v snahe modelovať fenomény ľudskej kognície ako napríklad kompozicionalita, systematickosť a problém viazania (pozri [7]). Tieto smery výskumu ostávajú predmetom záujmu.

## 8 Záver

Na základe použitej argumentácie dospievame k týmto záverom: Najslubnejším prístupom k vysvetleniu mechanizmov mysle je konekcionistický prístup. Nie preto, že umožňuje realizovať neklasické výpočty (nekonečná presnosť nie je nevyhnutná), ale pretože umožňuje modelovať elementárne mechanizmy učenia, ktoré sú v kognícii kľúčové, pričom tieto mechanizmy nie sú nezávislé od implementačnej úrovne. Okrem toho, konekcionizmus ponúka možnosti modelovania kognitívnych procesov na rôznych úrovniach abstrakcie jednotným spôsobom (pomocou rovnakých princípov). Ako odpoveď na otázku v článku tvrdíme, že existujúci reprezentačný pluralizmus je síce užitočný (najmä v umelej inteligencii, kde sa nemusíme opierať o biologicky inšpirované riešenia), no možno sa ukáže v budúcnosti ako nie nevyhnutný pre výpočtovú kognitívnu vedu.

## Podakovanie

Tento príspevok bol podporený grantami VEGA 1/0439/11 a 1/0503/13, a KEGA 076UK-4/2013. Ďakujem svojim kolegom Jánovi Rybárovi a Martinovi Takáčovi za užitočné pripomienky.

## Literatúra

- [1] Barsalou, L.: Perceptual symbol systems. *Behavioral and brain sciences* 22(04), 577–660 (1999)
- [2] Barsalou, L.: Grounded cognition. *Annual Reviews of Psychology* 59, 617–645 (2008)
- [3] Churchland, P.S., Sejnowski, T.J.: *The Computational Brain*. The MIT Press (1992)
- [4] Clark, A.: *Mindware: An Introduction to the Philosophy of Cognitive Science*. Oxford University Press (2001)
- [5] Dove, G.: Beyond perceptual symbols: A call for representational pluralism. *Cognition* 110, 412–431 (2009)
- [6] Farkaš, I.: Konekcionizmus v náručí výpočtovej kognitívnej vedy, pp. 19–62 (2011)
- [7] Farkaš, I.: Indispensability of computational modeling in cognitive science. *Journal of Cognitive Science* 13(12), 401–435 (2012)
- [8] Fodor, J.: *The Mind Doesn't Work That Way*. MIT Press (2000)
- [9] Fuster, J.: Cortex and memory: Emergence of a new paradigm. *Journal of Cognitive Neuroscience* 21(11), 2047–2072 (2009)
- [10] Gibson, J.: *Ecological Approach to Visual Perception*. Houghton-Mifflin, Boston (1970)
- [11] Griffiths, T., Chater, N., Kemp, C., Perfors, A., Tenenbaum, J.: Probabilistic models of cognition: Exploring representations and inductive biases. *Topics in Cognitive Science* 14, 357–364 (2010)
- [12] Hirata, Y., Judd, K., Kilminster, D.: Estimating a generating partition from observed time series: Symbolic shadowing. *Physics Review E* 70, 016215 (2004)
- [13] Hutchins, E.: *Cognition in the Wild*. MIT Press (1996)
- [14] Jellema, T., Perrett, D.: Neural representations of perceived bodily actions using a categorical frame of reference. *Neuropsychologia* 44, 1535–1546 (2006)
- [15] Kitzbichler, M., Smith, M., Christensen, S., Bullmore, E.: Broadband criticality of human brain network synchronization. *PLoS Computational Biology* 5(3) (2009)
- [16] Litt, A., Eliasmith, C., Kroon, F., Weinstein, S., Thagard, P.: Is the brain a quantum computer? *Cognitive Science* 30, 593–603 (2006)
- [17] Marr, D.: *Vision*. W.H. Freeman, San Francisco, CA (1982)
- [18] McClelland, J.: The place of modeling in cognitive science. *Topics in Cognitive Science* 1(1), 11–38 (2009)
- [19] Newell, A., Simon, H.: *Computer science as empirical enquiry*. *Communications of the ACM* 19, 113–126 (1976)

- [20] O'Brien, G., Opie, J.: How do connectionist networks compute? 7(1), 30–41 (2006)
- [21] O'Reilly, R.: The what and how of prefrontal cortical organization. *Trends in Neurosciences* 33(4), 355–361 (2010)
- [22] O'Reilly, R., Munakata, Y., Frank, M., Hazy, T., Contributors: *Computational Cognitive Neuroscience*. Wiki Book, second edn. (2012)
- [23] Perfors, A., Tenenbaum, J., Griffiths, T., Xu, F.: A tutorial introduction to Bayesian models of cognitive development. *Cognition* 120, 302–321 (2011)
- [24] Piccinini, G.: Some neural networks compute, others don't. *Neural Networks* 21(23), 311–321 (2008)
- [25] Pinker, S.: *How the Mind Works*. W. W. Norton and co. (2009)
- [26] Pitt, D.: *Mental Representation* (2012), <http://plato.stanford.edu/archives/win2012/entries/mental-representation>
- [27] Port, R., van T. van Gelder (eds.): *Mind as Motion: Explorations in the Dynamics of Cognition*. MIT Press, Cambridge, MA (1995)
- [28] Rieke, F., Warland, D., de Ruyter van Steveninck, R., Bialek, W.: *Spikes: Exploring the Neural Code*. MIT Press (1999)
- [29] Schoener, G.: Dynamical systems approaches to cognition. In: Sun, R. (ed.) *Cambridge Handbook of Computational Psychology*, pp. 101–126. Cambridge University Press, New York (2008)
- [30] Šefránek, J.: *Inteligencia ako výpočet*. IRIS Bratislava (2002)
- [31] Siegelmann, H.: Neural and super-turing computing. *Minds and Machines* 13(1), 103–114 (2003)
- [32] Sloman, A.: The irrelevance of Turing machines to artificial intelligence. In: Scheutz, M. (ed.) *Computationalism: New Directions*, pp. 87–128. MIT Press, Cambridge (2002)
- [33] Smolensky, P.: On the proper treatment of connectionism. *Behavioral and Brain Sciences* 11, 1–23 (1988)
- [34] Spivey, M.: *The Continuity of Mind*. Oxford University Press, Oxford (2007)
- [35] Stewart, J., Gapenne, O., Di Paolo, E. (eds.): *Enaction: Toward a New Paradigm for Cognitive Science*. MIT Press (2010)
- [36] Tabor, W.: A dynamical systems perspective on the relationship between symbolic and nonsymbolic computation. *Cognitive Neurodynamics* 3(4), 415–427 (2009)
- [37] Thagard, P.: *Mind: An Introduction to Cognitive Science*. MIT Press, second edn. (2005)
- [38] Vernon, D., Metta, G., Sandini, G.: A survey of artificial cognitive systems: Implications for the autonomous development of mental capabilities in computational agents. *IEEE Transactions on Evolutionary Computation* 11(2), 151–180 (2007)
- [39] Wang, J., Conder, J., Blitzer, D., Shinkareva, S.: Neural representation of abstract and concrete concepts: A meta-analysis of neuroimaging studies. *Human Brain Mapping* 31(10), 1459–1468 (2010)
- [40] Wilson, M.: Six views of embodied cognition. *Psychonomic Bulletin and Review* 9(4), 625–636 (2002)

# Multi-agentové systémy pro modelování biologických fenoménů

Martina Husáková

Univerzita Hradec Králové, Fakulta informatiky a managementu  
Hradecká 1249/6, Hradec Králové 500 03  
martina.husakova.2@uhk.cz

## Abstrakt

Cílem výzkumníků, kteří se zabývají fungováním živých systémů, je odrývat neuvěřitelnou komplexitu těchto systémů pro hlubší pochopení zákonitostí, kterými se řídí. Využívají k tomu různorodé *in vivo*, *in vitro* nebo *in silico* přístupy. Tento příspěvek se orientuje na *in silico* metody vedoucí k analýze fungování biologických systémů, resp. na výpočetní modely a simulace. Hlavní pozornost je v příspěvku věnována multi-agentovým systémům a jejich vybraným existujícím aplikacím v oblasti biologie a medicíny spolu s uvedením jejich silných a slabých stránek pro modelování biologických fenoménů. Příspěvek zároveň naznačuje možná zajímavá témata, která by mohla být řešena právě s využitím multi-agentových systémů.

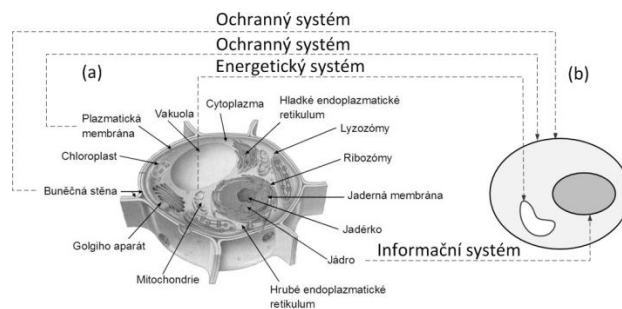
## 1 Úvodní kapitola

Biologické systémy jsou neodmyslitelně spjaty s vysokou mírou složitosti, která je dána existencí mnoha "hráčů", interakcí mezi nimi, a mnohdy jejich silnou provázaností s jinými systémy. Tvorba modelů takových systémů vyžaduje mj. hluboké znalosti týkající se předmětu modelování a volbu vhodného způsobu modelování. Jsou známy různé přístupy pro modelování biologických procesů. Tento příspěvek si klade za cíl představit multi-agentové systémy v kontextu silných a slabých stránek v jejich použitelnosti pro modelování biologických systémů spolu s uvedením zajímavých aplikací. Zároveň také vybízí k výzkumu a to návrhem zajímavých témat, která by mohla být řešena.

Příspěvek je strukturován následovně: druhá kapitola podává nezbytný základní pohled na to, co si představit pod soulovím biologický systém. Třetí kapitola zmiňuje přední přístupy k modelování biologických systémů. Čtvrtá kapitola se již detailněji věnuje pouze jednomu z nich – multi-agentovým systémům, jejich aplikacím v biologii a medicíně spolu s uvedením výhod i nevýhod v použití těchto typů systémů pro dané domény. Zmiňuje také několik témat, která by mohla být řešena s jejich pomocí. Kapitola pátá stručně rekapituluje obsah příspěvku.

## 2 Biologický systém

Biologický systém je systémem, který je v neustálé interakci s okolním prostředím. Na podněty reaguje s cílem zachovat stálost vnitřního prostředí a jeho rovnováhu - homeostázu. Děje v něm probíhající vedou k jeho dynamické povaze. Tato dynamika se projevuje např. v jeho schopnosti adaptace na různorodé podmínky prostředí, učení, zapamatování si minulých událostí nebo samo-organizaci. U biologických systémů si často povšimneme jisté hierarchie. Nejmenším známým útvarem schopným samostatné existence je buňka. Buňky jsou formovány do tkání a tkáně formují orgány, které spolu s ostatními již tvoří systém orgánů. Všechny tyto systémy lze považovat za biologické i komplexní. Právě komplexita je důvodem obtížného poznávání i modelování biologických systémů. Volba vhodného stupně abstrakce je velmi obtížným úkolem při tvorbě modelů, zvláště v biologii a medicíně. Obr. 1 ukazuje, v jakých úrovních složitosti se lze na buňku dívat: vysoká míra složitosti (a), nejnižší míra složitosti – minimální model (b).



Obr. 1. Složitost rostlinné buňky (úprava dle [1])

Jistě lze souhlasit s tvrzením fenomenálního fyzika A. Einsteina: „Všechno by mělo být tak jednoduché, jak je to jen možné, ale ne jednodušší.“ Přílišné zjednodušení totiž může vést k nesprávným závěrům.

Při studiu biologických systémů je nutné mít na paměti, že jsou něčím víc, než pouze souborem jejich komponent. Aplikace systémových přístupů pro studium těchto

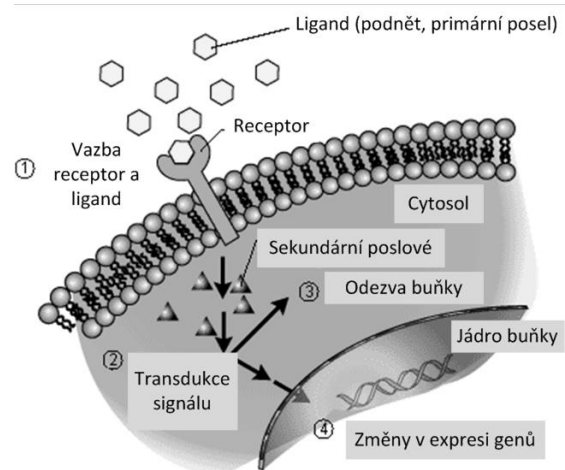
systémů vykrystalizovaly v systémovou biologii – směr, který se s výrazným přispěním matematiky a informatiky snaží pochopit (emergentní) procesy probíhající v biologických systémech. Pro tyto účely jsou na základě biologických dat tvořeny modely a ověřovány hypotézy pomocí simulací. Následující kapitola příspěvku zmiňuje přední přístupy pro modelování biologických systémů.

### 3 Přístupy k modelování

Přístupy použitelné pro modelování biologických systémů lze rozdělit dle [2] na dvě skupiny: *top-down* a *bottom-up*. Top-down přístupy pohlíží na systém z globálního – makroskopického úhlu pohledu. V centru pozornosti jsou především agregované (průměrné) charakteristiky sledovaného systému. Bottom-up přístupy detailněji modelují dílčí elementy studovaného systému, které existují na mikroskopické úrovni. S těmito přístupy lze lépe pozorovat interakce mezi těmito elementy a pozorovat jevy, které mají rysy emergence, což je s top-down přístupy obtížné. Daní za to je vysoká výpočetní náročnost v porovnání s top-down přístupy.

#### 3.1 Diferenciální rovnice

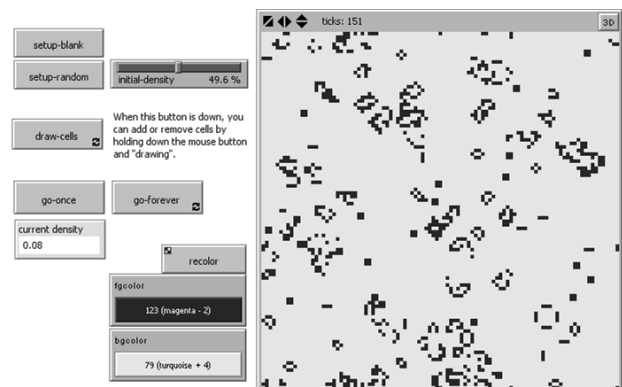
K předním představitelům top-down přístupů patří *diferenciální rovnice* (ordinary differential equations - ODE). Byly použity např. pro tvorbu prvního formálního modelu imunitní sítě dánským imunologem a nositelem Nobelovy ceny za fyziologii a medicínu N. K. Jerne v r. 1974 [3]. N. K. Jerne k vysloveným teoretickým aspektům fungování těchto sítí vytvořil první spojitou imunitní síť, která popisovala dynamiku bílých krvinek (lymfocytů). Tyto spojitě sítě byly rozšiřovány dalšími výzkumníky a daly podnět pro vznik skupiny diskretních imunitních sítí, které spíše využívaly diferenčních rovnic a aplikovaly se např. pro rozpoznávání vzorů, strojové učení nebo optimalizační problémy [37]. Diferenciální rovnice se využívají i pro výzkum rakovinných buněk a jejich šíření. Známe jsou aplikace modelů růstu avaskulárních tumorů (tumorů bez přístupu ke krevním cévám) [4], modelů angiogeneze (procesů spojených se zásobováním nádorových buněk živinami prostřednictvím působení na okolní cévy) [5]. Klíčovou roli v biologických systémech sehrávají mj. tzv. signální dráhy (signaling pathways) sestávající z kaskády různých biochemických reakcí, které jsou výsledkem reakce buňky na příjem určitého vstupního signálu. Různé signální dráhy pak formují velmi komplexní biologickou síť, viz obr. 2 upraven dle [15]. Tyto signální dráhy lze reprezentovat také diferenciálními rovnicemi [6], [7].



Obr. 2. Signální dráhy (úprava dle [15])

#### 3.2 Celulární automaty

K předním představitelům bottom-up přístupu patří *celulární automaty* (cellular automata - CA). Celulární automaty – uvedené do povědomí veřejnosti J. von Neumannem a S. Ulamem pod názvem „cellular spaces“ [8] – jsou diskretními systémy, které sestávají 2D nebo 3D mřížky tvořené umělými buňkami, viz jedna z aplikací CA Hra života (Game of Life) na obr. 3. Jejich rozhodování se řídí jednoduchými pravidly, které se odvíjí od stavu okolních umělých buněk. Typy stavů a jejich počet jsou u všech buněk stejné a konečné. Statické varianty CA nepředpokládají změnu pozic umělých buněk na rozdíl od dynamických alternativ. Komunikace mezi umělými buňkami se předpokládá pouze na lokální úrovni. I přes jejich jednoduchost lze s jejich pomocí poznávat komplexní systémy a sledovat možné emergentní jevy, nicméně tato jednoduchost může zapříčinit až přílišné zjednodušení zkoumaného systému a dospění ke zkresleným závěrům.



Obr. 3. Hra života v NetLogu



Celulární automaty byly použity např. pro modelování procesů odehrávajících se v imunitním systému člověka. Příkladem je příspěvek [9], který se zabývá modelováním lymfocytů, antigenů (objektů schopných vzbudit pozornost imunitního systému), buněk prezentujících antigeny, a protilátek, resp. jejich tvorbou, pohybem a zánikem. CA jsou aplikovatelné např. pro modelování růstu tumoru. Příspěvek [10] zmiňuje použití různých typů CA pro hlubší porozumění vývoje tumoru s následnou predikcí jeho chování. CA jsou v [11] použity pro simulaci dynamiky šíření viru HIV. V neposlední řadě lze CA využít i pro modelování chemotaktického pohybu bakterií (specifického pohybu v závislosti na intenzitě koncentrace určité látky). Příspěvek [12] zmiňuje použití dvou interagujících CA využitelných pro modelování chemotaxe nejvíce prozkoumané bakterie *Escherichia Coli* (*E.coli*). První CA slouží pro spojitou reprezentaci prostředí, kde dochází k difuzi chemických molekul upoutávajících pozornost bakterie *E.coli*. Druhý CA slouží pro modelování samotné bakterie.

## 4 Multi-agentové systémy

*Multi-agentové systémy* (multi-agent systems - MAS) jsou řazeny spolu s CA do kategorie bottom-up přístupů. MAS jsou stochastické, nejčastěji vnímané jako decentralizované, systémy sestávající z dílčích jednotek tzv. *agentů*<sup>1</sup>. Existují různé pohledy na to, co je agent a jakými vlastnostmi se má vyznačovat. Dle [13] je agentem výpočetní systém situovaný v určitém prostředí se schopností autonomních akcí za účelem dosažení daných cílů. Na základě příjmu podnětů z okolí pomocí senzorů tyto akce realizuje více či méně pro-aktivně a autonomně pomocí aktuátorů, viz obr. 4. „Slabá“ verze agenta předpokládá, že se jedná o entitu [14]:

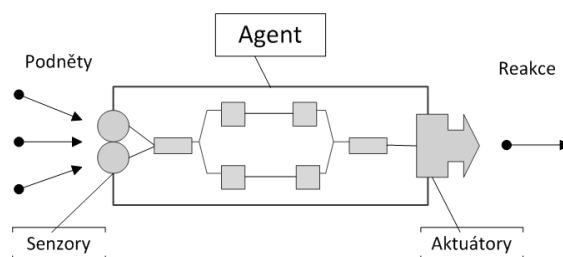
- autonomní (schopnou nezávislého jednání bez asistence člověka),
- sociální (schopnou interakce s okolím zahrnující komunikaci, kooperaci a koordinaci),
- reaktivní (vnímající okolí a adekvátně reagující na podněty),
- pro-aktivní (schopnou převzít iniciativu při řešení problémů).

„Silná“ verze agenta k těmto vlastnostem přidává [14]:

- mobilitu (pohyb v prostředí),
- benevolenci (jednání ve prospěch ostatních agentů),
- racionalitu (účelné a rozumné rozhodování),
- schopnost úvahy o vlastních záměrech,
- získávání a využívání znalostí pro rozhodování,

<sup>1</sup> Pojem agent je v příspěvku vnímán jako umělá entita vybavená metodami určující její chování.

- trvalost (neustálé konání dokud není cíl splněn),
- pravdomluvnost,
- schopnost agenta projevit emoce.



Obr. 4. Agent a jeho základní komponenty

Dle úrovně sofistikovanosti zpracovávání podnětů z okolního prostředí, jsou agenty klasifikovány na agenty *reaktivní* (založené na modelu podnět – reakce bez schopnosti plánování a explicitní reprezentace prostředí), *deliberativní* (založené na modelu podnět – plán – reakce s explicitní reprezentací prostředí), *sociální* (mající propracované mechanismy komunikace s ostatními agenty) a *hybridní* (kombinace předešlých typů agentů), viz blíže např. [13], [14], [20]. MAS lze nasadit všude tam, kde požadujeme:

- paralelní zpracovávání úloh,
- řešení problémů bez nutnosti asistence člověka,
- zajištění efektivní koordinace nebo kooperace aktivit s možností jejich optimalizace,
- studovat emergentní jevy,
- analyzovat projevy kolektivní inteligence určitých společenství nebo obecněji
- vytvářet takové modely sociálního, biologického, environmentálního nebo fyzikálního charakteru, kde pracujeme s populacemi.

### 4.1 Agenty v biologii a medicíně

Multi-agentový systém se jeví být velmi vhodnou a přirozenou alternativou pro modelování biologických fenoménů. Tato vhodnost již vyplývá z charakteru agenta – ucelené jednotky schopné samostatného rozhodování a řešení problémů. Živé entity i umělé agenty lze vnímat i z pohledu sociálního. Agenty mohou mít *pro-sociální* (kooperativní) charakter, kdy zájmy skupiny upřednostňují před zájmy vlastními, nebo *kompetitivní* charakter, kdy na prvním místě jsou jejich vlastní zájmy. Případně mohou vzhledem k situaci vykazovat oba tyto rysy. Pro-sociálnost i soutěživost lze pozorovat i v biologických systémech, např. u normálních a abnormálních buněk. Normální buňky, např. bílé krvinky, v kooperaci s ostatními jednají tak, aby nebyla narušena stabilita organismu. Jsou pro to ochotné vzdát se i vlastní

existence (apoptóza). Abnormální buňky, např. rakovinné povahy, mají doslova asociální projevy. Jsou schopné zabírat „území“ na úkor již existujících buněk a často normální buňky ovlivňují tak, aby se chovaly v jejich prospěch.

Mezi multi-agentovým a biologickým systémem lze vidět určité podobnosti. Tabulka 1 uvádí vybrané vlastnosti vyskytující se u obou systémů. Z pohledu MAS se jedná také o procesy, které jimi lze modelovat.

Tab. 1. Společné rysy biologického systému a MAS

<b>Biologický systém</b>	Otevřenost	Mobilita
	Paměť, učení	Orientace na cíl
<b>MAS</b>	Dynamika	Decentralizace
	Emergence	Autonomie
	Komplexita	Samo-organizace
	Adaptace	Homogenní, heterogenní entity

Na rozdíl od diferenciálních rovnic, multi-agentové systémy i celulární automaty nemusí být vybaveny silným formálním (matematickým) aparátem [21]. Jsou tedy vhodnější pro výzkumníky, kteří nejsou „čistými“ matematiky. Lze je vyvíjet poměrně snadno a rychle, samozřejmě s ohledem na řešený problém a použité vývojové prostředí nebo programovací jazyk. Na druhou stranu jsou často výpočetně náročnější v porovnání s diferenciálními rovnicemi, protože se jimi modeluje každá dílčí jednotka systému (agent). Diferenciální rovnice se zaměřují spíše na průměrné charakteristiky případných populací, než na jednotlivé charakteristiky každého jedince. Diferenciální rovnice předpokládají homogenitu studovaných populací. V případě MAS není problém modelovat homogenní, ale i heterogenní populace agentů. V porovnání s CA, multi-agentové systémy nabízejí tvorbu propracovanějších modelů, resp. možnou integraci algoritmů pro zpracovávání znalostí, učení, samo-organizaci a plánování.

Tabulka 2, vycházející z [8], [21], [26], [33], uvádí stručně a zobecněné porovnání diferenciálních rovnic (ODE), statických celulárních automatů (CA) a multi-agentových systémů (MAS).

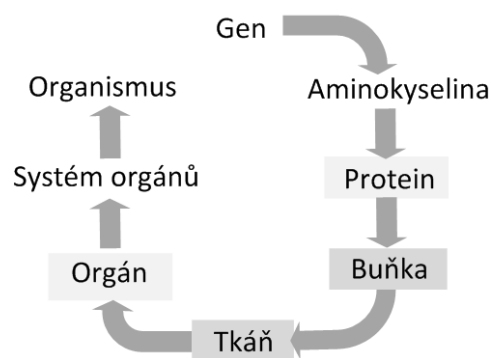
Tab. 2. ODE, CA a MAS – porovnání přístupů

	ODE	CA	MAS
Studium emergence	Ne	Ano	Ano
Průzkum makro-úrovně	Ano	Ano	Ano
Průzkum mikro-úrovně	Ne	Ano	Ano
Modelování interakcí mezi agenty v prostoru	Ne	Ano	Ano

Vývoj modelu z neúplných informací o systému	Ne	Ano	Ano
Nezbytnost matematického aparátu	Ano	Ne	Ne
Modelování heterogenních populací agentů	Ne	Ne	Ano

## 4.2 Vybrané aplikace v biologii a medicíně

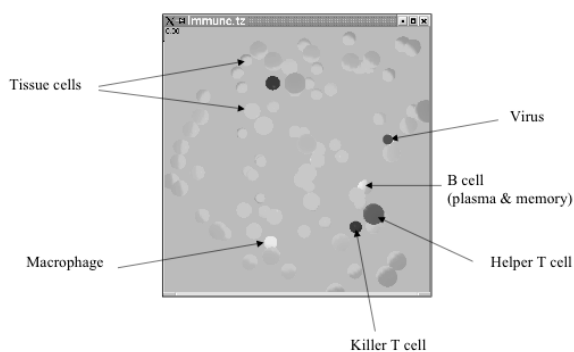
Multi-agentové systémy lze použít (nejen) v rámci biologie (medicíny) pro potvrzení nebo vyvrácení teoretických hypotéz, realizaci what-if analýzy (sledování důsledků při změnách parametrů modelu), predikci chování systému, získání dat, která není možné nebo obtížné získat z reálného systému, pro zautomatizování procesu získávání informací nebo využití multi-agentového přístupu pro konceptualizaci biologických systémů s využitím vhodných jazyků. Průzkumem použitelnosti MAS v biologii (medicíně) bylo zjištěno, že jsou nejvíce aplikovány pro studium jevů odehrávajících se na úrovni *buněčné* a *tkáňové*, viz obr. 5. V následujících odstavcích budou zmíněny vybrané aplikace MAS v biologii a medicíně.



Obr. 5. MAS a modelování biologických systémů

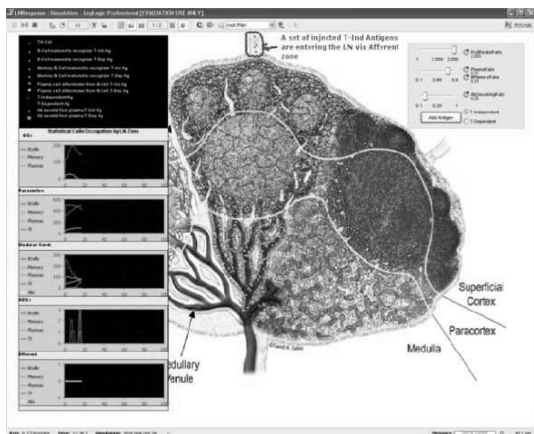
### 4.2.1 MAS a úroveň buněčná

Multi-agentové systémy lze uplatnit např. v tzv. výpočetní imunologii, resp. při modelování jevů na úrovni vrozené nebo adaptivní imunity. Příspěvek [16] se zabývá modelováním imunitní odezvy po napadení organismu virem s využitím technik rojové inteligence. Modely integrují různé typy imunitních buněk např. makrofágy, T-pomocné lymfocyty, naivní B-lymfocyty, tkáňové buňky a zabíječské T-lymfocyty. Pomocí vytvořených modelů byly sledovány populace virů a protilátek v průběhu imunitní odezvy látkové i buněčné. Pro vývoj modelů bylo využito open-source prostředí Breve [17], viz obr. 6.



Obr. 6. Breve pro výzkum látkové a buněčné imunity [16]

Příspěvek [18] se zabývá modelováním procesů odehrávajících se při imunitní odezvě na přítomnost antigenů uvnitř lymfatické uzliny. Do modelu jsou zahrnuty interakce mezi B-lymfocyty, pomocnými T-lymfocyty, protilátkami a antigeny. Pro konceptuální reprezentaci imunitních buněk byly využity stavové diagramy. Z modelovaných procesů byla uvažována zejména migrace buněk mezi lymfatickými uzlinami, množení buněk, diferenciace buněk a generování protilátek. Model sloužil i pro statistickou analýzu obsazenosti uvažovaných lymfatických uzlin antigeny v jejich různých částech, viz obr. 7. Tento příspěvek navazuje na [19], který sice explicitně nezmiňuje použití multi-agentového přístupu, ale pro stejný účel také využívá stavových diagramů.



Obr. 7. Model lymfatické uzliny v AnyLogicu [18]

Další aplikací je agentový model AgentCell, který byl vytvořen za účelem hlubšího průzkumu vztahů mezi nitrobuněčnými stochastickými procesy a jejich důsledky na chování buněk [28]. Pro tento účel byl vytvořen agentový model chemotaktického pohybu bakterie E.Coli. Každá bakterie je v příspěvku agentem s vlastní chemotaktickou sítí implementovanou v prostředí Repast

[27]. V neposlední řadě lze vyvíjet multi-agentové modely pro studium formování tzv. granulomů – komplexních systémů, které vznikají při neschopnosti imunitních buněk - makrofágů - eliminovat určité objekty, např. specifické parazity, houby a pod. Příspěvek [30] pojednává o tvorbě agentového modelu pro výzkum formování granulomů v plicích.

#### 4.2.2 MAS a úroveň tkáňová

Projekt Epitheliome si kladl za cíl hlouběji prozkoumat a porozumět procesům formování tzv. epitelových tkání – tkání, které mají zejména ochrannou funkci [30]. Formování tkání je emergentním jevem, který se odvíjí od interakcí mezi buňkami na mikroúrovni. Agentové systémy byly použity pro modelování interakcí mezi buňkami, kde každá buňka byla reprezentována agentem. Framework Flame, který je založen na použití konečných stavových automatů pro reprezentaci agentů, byl využit pro tvorbu modelu epitelové tkáně [31]. V příspěvku [32] je navržen přístup pro vývoj modelů simulujících chování buněčných tkání různého typu. Agentem je buňka tkáně, která je reprezentována pomocí konečného stavového automatu. Agenty tkáně jsou formovány do tkání, které jsou také agenty. Environmentální kontext hraje úlohu datové struktury, která má přehled o všech agentech, jejich vlastnostech, stavech a interakcích, které se mezi nimi (v „umělé tkáni“) odehrávají. Od agentů přijímá požadavky, na základě kterých rozhodne, zda agent může danou akci provést.

#### 4.2.3 MAS a jiné aplikace

V rámci projektu GeneWeaver byl vyvinut flexibilní systém pro automatickou analýzu genomu s využitím agentového přístupu. Byla vytvořena komunita autonomních agentů, v rámci které každý agent vykonával specifickou úlohu, např. registraci informací o agentech v komunitě, management databázi obsahující sekvenční data nebo analýzu sekvenčních dat (určování struktury a funkce proteinů), blíže v [22]. Příspěvek [23] využívá reaktivních agentů pro modelování metabolické sítě mitochondrie, která má významný vliv např. na buněčnou diferenciaci, apoptotické děje, kontrolu růstu buněk apod. Organizační paradigma Agent-Skupina-Role bylo využito pro modelování nitrobuněčných dynamických procesů bakterie E. Coli v příspěvku [24]. Dále pak v práci [25] byly celulární automaty spolu s MAS aplikovány pro formální modelování a simulaci životního cyklu kmenových buněk. Kostní dřev, která mj. tvoří tyto buňky, byla reprezentována jako neorientovaný graf. Uzly grafu představovaly kmenové nebo již diferencované buňky. Cílem bylo modelovat přechody mezi stavem kmenová-diferencovaná buňka.

### 4.3 Kritický pohled

Multi-agentový přístup je velice flexibilní. Poměrně jednoduše lze měnit strukturu MAS, resp. přidávat a ubírat dílčí entity systému, aniž by bylo nutné razantně měnit strukturu celého systému. Pojem agent je velmi blízký biologickým entitám. Pro tvůrce multi-agentových modelů je agent také snadněji uchopitelný, než je tomu v případě diferenciálních rovnic. Pro vývoj MAS existují různorodé nástroje umožňující realizovat experimenty, při nichž je možné spouštět stejný model s různými parametry, měnit parametry během simulace, vizualizovat systém v 2D, 3D formě, vytvářet modely v různých typech prostředí (mřížka, graf, spojitý prostor) nebo si pohrávat s časem simulace, resp. zpomalit nebo zrychlit vývoj zkoumaného systému.

MAS mají z pohledu tvorby modelů v biologii (medicině) i své stinné stránky. K nejzávažnějším patří důvěryhodnost výsledků generovaných multi-agentovými simulacemi. Kvalita modelu se odvíjí od určité preciznosti v jeho tvorbě a zároveň volby vhodné úrovně složitosti, použitého vývojového prostředku, dovednosti převést poznatky o sledovaném systému do podoby programového kódu. Validita modelu se ověřuje s pomocí reálných dat, která ale nemusí být vždy k dispozici. Jak je zmíněno v [33], mnohem častěji se porovnávají různé modelovací techniky, než reálná data s vytvořeným modelem. Modely ale mohou být vytvářeny za účelem získat data, která není možné získat z reálného systému. Pro řešení problému validace je třeba určitých systematických metodologií, které by napomohly v řešení tohoto problému [33]. Další komplikací je množství entit obsažených v biologickém systému. Reprezentace každé z nich ve formě agenta je vysoce výpočetně náročné. Paralelní zpracování úloh může být řešením [34]. Dalším neméně důležitým problémem je interpretace simulace. Čím více je modelovaný systém komplexní, tím obtížnější je zjistit, co nám vlastně vytvořený model ukazuje.

### 4.4 Náměty k zamyšlení

*Dendritické buňky* sehrávají klíčovou úlohu v modulaci odezvy imunitního systému. Jsou chápány jako prostředník mezi vrozenou a adaptivní vrstvou imunitního systému. Jsou schopné pohlcovat odumřelé nebo poškozené buňky (fagocytóza) a především vystavovat (prezentovat) antigenní fragmenty na povrchu pro jejich detekci T-lymfocyty [35].

V současné době jsou intenzivně zkoumány způsoby jejich vyhodnocování nebezpečí, protože dendritické buňky zadávají „příkazy“ zejména T-lymfocytům - jak se

mají nebezpečí postavit. Multi-agentové systémy by mohly být užitečné při analýze kolektivního chování dendritických buněk a T-lymfocytů.

Další možné uplatnění MAS v biologii (medicině) lze vidět při studiu tzv. *biofilmů*. S biofilmy se setkáváme téměř každý den. Tvoří se např. na povrchu zubní skloviny, znečišťují chirurgické nástroje, kontaktní čočky, implantáty apod. Biofilm je určitým společenstvím homogenních nebo heterogenních mikroorganismů (jednobuněčných nebo mnohobuněčných) a tzv. extracelulárních polymerních látek, které tyto mikroorganismy vylučují. Výzkum biofilmů ukazuje, že existují různé důvody jejich tvorby. Jedním z nich je skutečnost, že pokud bakterie „strádají“, mají tendenci nacházet bezpečí ve společnosti jiných bakterií [36]. Biofilm totiž tvoří bezpečný a homeostatický prostor, ve kterém mohou bakterie získávat chybějící živiny a být pod ochranou např. před imunitními buňkami nebo UV zářením. Biofilm je totiž velice odolný proti vnějším vlivům a proto je velmi obtížné ho odstranit.

Multi-agentové systémy by mohly být užitečné pro modelování tohoto komplexního společenství, resp. při zodpovězení otázek spojených s životním cyklem bakterií obsažených v biofilmech, interakcemi bakterií uvnitř biofilmu a s okolím.

## 5 Závěr

Multi-agentové systémy jsou významným paradigmatem, které nachází uplatnění při tvorbě modelů použitelných pro analýzu biologických systémů (zejména na úrovni buněčné a tkáňové), a to např. při výzkumu imunitních procesů, formování tkání, buněčných signálních drah nebo kmenových buněk. Zmíněné aplikace MAS ukazují jen malý fragment z mnohem rozsáhlejšího souboru vědeckých prací. Ač jsou MAS přirozenou alternativou pro výzkum v oblasti biologie (medicíny), k jejich uznání jako plnohodnotného přístupu pro výzkum v biologii a příbuzných oblastech, je třeba mít k dispozici metodologie, které nabídnou systematický přístup při tvorbě modelů se způsoby ověření jejich důvěryhodnosti.

### Poděkování

Tento příspěvek byl podpořen projektem Informační, kognitivní a interdisciplinární podpora výzkumu - INKOV, reg. č. CZ.1.07/2.3.00/20.0001, financovaným z prostředků EU a ČR.

## Literatura

- [1] Kolektiv vyučujících SZŠ Karviná. Buňka: Interaktivní výuková aplikace. Zpracováno v rámci projektu SIPVZ č. 0636P2006, [cit. 1. 4. 2013]. Dostupné z: <http://sszdrakarvina.cz/bunka/bi/03eu/ershr.htm>.
- [2] XIANG-HUA LI, a kol., 2009. Modelling Immune System: Principles, Models, Analysis and Perspectives. *Journal of Bionic Engineering*, vol. 6, issue 1, 77-85.
- [3] JERNE, N. K., 1974. Clonal Selection in a Lymphocyte Network. Cellular Selection and Regulation in the Immune Response. *Journal of Society of General Physiologists* series, vol. 2, 39-48.
- [4] ROOSE, T., a kol., 2007. Mathematical Models of Avascular Tumor Growth. *Siam Review*, vol. 49, no. 2, 179-208, [cit. 29. 3. 2013]. Dostupné z: <http://eprints.maths.ox.ac.uk/612/1/acroread.pdf>.
- [5] CHAPLAIN, M. A. I., a kol., 2006. Mathematical Modeling of Tumor-Induced Angiogenesis. *Annual Review of Biomedical Engineering*, vol. 8, 233-257.
- [6] RANGAMANI, P. a IYENGAR, R., 2012. Modelling cellular signalling systems. *Essays Biochem*, vol. 45, 83 – 94.
- [7] BREE, B., a kol., 2006. Physicochemical modelling of cell signalling pathways. *Nature Cell Biology*, vol. 8, 1195 – 1203.
- [8] BAUER, A. L., BEAUCHEMIN, C.A.A. a PERELSON, A. S., 2009. Agent-based modeling of host–pathogen systems: The successes and challenges. *Information Sciences*, vol. 179, issue 10, 1379-1389, [cit. 10. 3. 2013]. Dostupné z: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2731970/pdf/nihms-116325.pdf>.
- [9] CELADA, F. a SEIDEN, P., 1992. A computer model of cellular interactions in the immune system. *Immunol Today*, vol. 13, 56–62.
- [10] BOONDIREK, A., a kol., 2010. A Review of Cellular Automata Models of Tumor Growth. *International Mathematical Forum*, vol. 5, no. 61, 3023 – 3029, [cit. 15. 3. 2013]. Dostupné z: <http://www.m-hikari.com/imf-2010/61-64-2010/boondirekIMF61-64-2010.pdf>.
- [11] DOS SANTOS, R. M. Z a COUTINHO, S., 2001. Dynamics of HIV Infection: A Cellular Automata Approach. *Physical Review Letters*, vol. 87, issue 16.
- [12] KIM, TH., a kol., 2008. Investigations into the design principles in the chemotactic behavior of Escherichia coli. *BioSystems*, vol. 91, 171 – 182, [cit. 1. 3. 2013]. Dostupné z: <http://biomimetic.pbworks.com/f/Investigations%20into%20the%20design%20principlesKim.pdf#page=4&zoom=146,42,272>.
- [13] WOOLDRIDGE, M., 2005. *An Introduction to MultiAgent Systems*. John Wiley & Sons, 3rd Edition, 2005, ISBN 978 0471 49691 5.
- [14] WOOLDRIDGE, M. a JENNINGS, N. R., 1995. Intelligent agents: theory and practice. *The Knowledge Engineering Review*, vol. 10:2, 115 – 152.
- [15] SMITH-C, T. Membrane Proteins—Receptors, function, [cit. 1. 4. 2013]. Dostupné z: <http://fwcdscience.wikispaces.com/Membrane+Proteins--Receptors>.
- [16] JACOB, Ch., LITORCO, J., LEE, L., 2004. Immunity Through Swarms: Agent-Based Simulations of the Human Immune System. *Proceedings of the Third International Conference On Artificial Immune Systems (ICARIS)*, 2004. Springer-Verlag Berlin Heidelberg, LNCS 3239, 400 - 412.
- [17] Breve: a 3d Simulation Environment for Multi-Agent Simulations and Artificial Life - homepage, [cit. 12. 3. 2013]. Dostupné z: <http://www.spiderland.org/breve/>.
- [18] BELKACEM, K. a FOUJIL, Ch., 2012. An AnyLogic Agent Based Model for the Lymph Node Lymphocytes First Humoral Immune Response. *Proceedings of the International Conference on Bioinformatics and Computational Biology (ICBCB 2012)*, vol. 34, IACSIT Press, Signapore, 2012, 163 - 169.
- [19] SWERDLIN, N., COHEN, I. R., HAREL, D., 2008. The Lymph Node B Cell Immune Response: Dynamic Analysis In-Silico. *Proceedings of The IEEE*, vol. 96, no. 8, 2008, 1421 - 1442.
- [20] KUBÍK, A., 2004. Inteligentní agenty: Tvorba aplikačního software na bázi multiagentových systémů. Brno: Computer Press. Vydání 1. 2004, s. 280. ISBN 80-251-0323-4.
- [21] Maira Athanázio de Cerqueira GATTI, M. A. de C. a LUCERNA, C. J. de P., 2007. An Agent-Based Approach for Building Biological Systems: Improving the Software Engineering for Complex and Adaptive Multi-Agent Systems. *Monografias em Ciência da Computação*, no. 14/07, June 2007, ISSN: 0103-9741.
- [22] BRYSON, K., a kol., 2000. Applying Agents to Bioinformatics in GeneWeaver. *Proceedings of the Fourth International Workshop on Cooperative Information Agents*, LNAI, Springer-Verlag, vol. 1860, 60 – 71.
- [23] LALES, C., a kol., 2005. Simulation of mitochondrial metabolism using multi-agent system. *The Fourth International Joint Conference on Autonomous Agents and Multi-Agent Systems*, 2005, 137 - 145.
- [24] BOOSE, T., a kol., 2005. Modeling the Dynamics of Intracellular Processes as an Organisation of Multiple Agents. *The First International Workshop on Multi-Agent Systems for Medicine, Computational Biology and Bioinformatics*, 2005.
- [25] D'INVERNO, M. a SAUNDERS, R., 2005. Agent-based modelling of stem cell self-organisation in a niche. Engineering self-organising systems: methodologies and applications. LNCS 3464, Springer-Verlag, Berlin, Germany, 2005, 52-68.
- [26] MACHADO, D., a kol., 2011. Modeling formalisms in Systems Biology. *AMB Express*. Dec. 5;1:45.
- [27] Repast - homepage, [cit. 15. 3. 2013]. Dostupné z: <http://repast.sourceforge.net/>.
- [28] EMONET, T., a kol., 2005. AgentCell: a digital single-cell assay for bacterial chemotaxis. *Bioinformatics*, vol. 21, no. 11, 2714 – 2721, [cit. 10. 3. 2013]. Dostupné z:

<http://bioinformatics.oxfordjournals.org/content/21/11/2714.abstract>.

- [29] SEGOVIA-JUARES, J. L., GAUNGULI, D. a KIRSCHNER, D., 2004. Identifying control mechanisms of granuloma formation during M. tuberculosis infection using an agent-based model. *Journal of Theoretical Biology*, Dec 7;231(3): 357 - 76.
- [30] WALKER, D. C., a kol., 2004. The epitheliome: agent-based modelling of the social behaviour of cells. *Biosystems*, Aug-Oct; 76(1-3): 89 - 100.
- [31] Flame: Flexible Large-scale Agent Modelling Environment – homepage, [cit. 10. 3. 2013]. Dostupné z: <http://www.flame.ac.uk/>.
- [32] SANTOS, E. E., GUO, D., SANTOS, E. Jr. a ONESTY, W., 2004. A Multi-Agent System Environment for Modelling Cell and Tissue Biology. *Proceedings of the International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA 2004)*, June 21-24, Las Vegas, Nevada, USA, vol. 1.
- [33] AMIGONI, F. a SCHIAFFONATI, V., 2007. Multiagent-Based Simulation in Biology: A Critical Analysis. *Model-Based Reasoning in Science, Technology, and Medicine - Studies in Computational Intelligence*, vol. 64, 179-191.
- [34] PERIN, D., a kol., 2006. An agent-based approach to immune modelling. *Computational Science and Its Applications (ICCSA 2006)*, LNCS, vol. 3980, 2006, 612 – 621, [cit. 12. 3. 2013]. Dostupné z: <http://www.waset.org/journals/waset/v17/v17-184.pdf>.
- [35] ŠPÍŠEK, R. a BARTUŇKOVÁ, J., 2003. Dendritická buňka v imunitě: Její role při regulaci imunitní reakce. *Vesmír* 82, duben 2003, [cit. 2. 4. 2013]. Dostupné z: <http://vesmir.cts.cuni.cz>.
- [36] SCHINDLER, J., 2001. Mikrobiální biofilm: Jak žijí bakterie v přirozeném prostředí? *Vesmír* 80, duben 2001, [cit. 2. 4. 2013]. Dostupné z: <http://vesmir.cts.cuni.cz>.
- [37] CASTRO, L. N. a TIMMIS, J., 2002. *Artificial Immune Systems: A New Computational Intelligence Approach*. 1st Ed. London: Springer-Verlag, p. 357. ISBN 1-85233-594-7.

# Primární mentální schopnosti jako měřítko inteligence v internetovém prostoru

Anna Janíková, Kamil Matula, Štěpánka Tůmová

Ústav informatiky, Filozoficko-přírodovědecká fakulta, Slezská univerzita v Opavě  
Bezručovo náměstí 1150/13, 746 01 Opava

Email: anna.janikova@fpf.slu.cz, kamil.matula@fpf.slu.cz, stepanka.tumova@fpf.slu.cz

## Abstrakt

Dlouhou dobu se lidstvo zabývalo problematikou toho, co je inteligence, z čeho se skládá a jaké složky ji utvářejí. V průběhu let bylo na toto téma vytvořeno mnoho studií, které inteligenci různě definovaly, popřípadě měřily. Mezi nimi se objevila také teorie o jádru inteligence, sedmi primárních mentálních schopnostech.

V dnešní době dokážeme mnoho věcí měřit i definovat velmi efektivně. Vystaly však nové otázky a problémy spolu s „internetizací“ společnosti. V případě, že je člověk díky síti odříznut od některých primárních vjemů, jakým způsobem se chová? Je mezi jeho inteligencí (kterou vyjádříme pomocí sedmi primárních mentálních schopností) v prostředí internetu a mimo něj nějaký rozdíl?

## 1 Úvod

Smyslem tohoto článku je představit potenciálně zajímavý výzkumný problém, který se týká druhu inteligence, jenž označujeme jako „internetovou inteligenci“. V rámci článku představíme související pojmy.

Nejprve se budeme zabývat definicí inteligence a vymezením jejích složek i jiných souvisejících pojmů. Jedním z těchto pojmů je chování, protože je projevem inteligence. Vzhledem k tomu, že se zaměřujeme na „internetovou inteligenci“, zajímá nás především informační chování, které úzce souvisí s chováním jedince na internetu. Internetové prostředí, ačkoli nemá hmatatelnou podobu, je běžnou součástí našeho života a možná jej ovlivňuje více, než si uvědomujeme. Dnes se dokonce o internetu mluví jako o virtuálním místě, kam se uchýlíme, když hledáme zábavu, poučení, když chceme najít dobrou restauraci nebo se pobavit s přáteli. Avšak na rozdíl od reálného světa, svět virtuální zůstává ochuzen o mnohé informace, které získáváme jinak než zrakem a sluchem. Proto se okrajově budeme věnovat také problematice vnímání.

Kognitivní procesy přímo ovlivňují myšlení člověka, a tím také jeho inteligenci. Je tedy možné, že omezená percepce na internetu může ovlivňovat chování člověka v prostředí internetu. Aby bylo možné tento jev sledovat, musíme od sebe oddělit jednotlivé pojmy, a to pojem

„obecná“ a „internetová inteligence“. V průběhu článku se budeme snažit definici „internetové inteligence“ nastínit, avšak nejedná se o jednoduchou problematiku, a tak navrhneme metodiku, která nám pomůže tento pojem lépe vymezit.

Nejdříve z definic inteligence zkusíme vybrat tu, která nám bude sloužit jako referenční model k našemu vymezení. Pro potřeby vymezení „internetové inteligence“, která je dle našeho názoru součástí „obecné“ inteligence, budeme vycházet z definic informačního chování a z vybraného referenčního modelu. Prozatím předpokládáme, že projevy „internetové inteligence“ souvisí s činností uživatelů na internetu.

## 2 Definice inteligence

Předně je třeba specifikovat pojem inteligence. Studium inteligence probíhá více jak 100 let a v popředí se vystřídalo několik koncepcí, které lze rozřadit podle úhlu pohledu na biologicko-fyziologické, vývojové, kognitivní, systémové a faktorové.[1]

Biologicko-fyziologické teorie se zaměřovaly na měření mozkových vln, hemisfér, průtoku krve a jiné problémy.[1] Žádné z pojetí, které bylo vytvořeno, uspokojivě neodpovídá našim požadavkům. Tato pojetí možná dokáží s různou mírou přesnosti inteligenci měřit, avšak čistě biologicko-fyziologické koncepce jsou příliš jednostranně zaměřené, a to zejména na fyziologické projevy a pojem internetové inteligence by pomocí nich nešel přesně vymezit.

Dalším zajímavým pohledem na inteligenci je genetická epistemologie, která se zabývá sledováním vývoje inteligence v rámci vývoje člověka. Spadá do vývojových teorií, které se zabývají především schémata učení a jejich spojením s pojmem inteligence.[1] Tyto koncepce jsou jistě zajímavé, avšak velmi specificky orientované, a proto definování internetové inteligence se nehodí, tato pojetí jsou si příliš vzdálená.

Kognitivní teorie inteligence je postavena na předpokladu, že inteligence úzce souvisí s psychickými zobrazeními informací a procesů. Kognitivní přístupy

se seskupují do kognitivně-korelačních, kognitivně-komponentových a kognitivně-obsahových teorií. Kognitivně korelační přístup zahrnuje zkoumání průběhu základních informačních procesů v laboratořích a zjišťují se korelace s psychodiagnostickými testy inteligence. Kognitivně-komponentový přístup se snaží o studium kognitivních procesů, které ovlivňují inteligenci. Na základě psychologických testů se snažíme zjistit poznávací procesy a strategie, které dotazovaní uplatňují. Roli konkrétních poznatků při inteligentním chování zdůrazňuje kognitivně-obsahový přístup, a to například v šachové partii nováčků a zkušených hráčů.[1] První dva přístupy se týkají teoreticko-psychologických průzkumů a k zaměření našeho průzkumu nejsou vhodné. Avšak poslední z nich, kognitivně-obsahový přístup, může být pro naši koncepci potenciálně přínosný, a bude objektem dalšího zkoumání.

Dalším přístupem jsou systémové koncepce. Například Triarchická koncepce R. J. Sternberga,[2] případně teorie mnohonásobné inteligence H. Gardnera,[1] které vychází z toho, že inteligence se skládá z více systémových prvků. Tyto definice inteligence navazují na různé činnosti, a snaží se tak pokrýt její mnohé podoby, avšak díky tomu jsou velmi široké. Proto jsou pro náš průzkum nevhodné.

Následující způsob, který se snažil definovat inteligenci, byl faktorový přístup, jehož zakladatelem je Ch. Spearman. Nelíbily se mu jednostranné definice a jednoduché testování inteligence, a proto sestavil vlastní baterii mentálních testů. Výsledky následně zpracoval pomocí faktorové analýzy, kterou vynalezl. Ze svých výsledků vyvodil dva faktory, které podle něj ovlivňovaly skóre v testech. Nazval je faktorem „g“ (general), který reprezentoval podobnost výsledků v testech, a faktorem „s“ (special), který reprezentoval rozdílnost výsledků v testech. Další z teorií, které navazovaly, se snažily faktory uspořádat do nějaké hierarchie. R. B. Cattell byl žákem Ch. Spearmana. S jeho faktorovým rozdělením souhlasil, ale doplnil faktor „g“ o rozdělení na dvě části, krystalizovanou a fluidní inteligenci. Fluidní inteligenci si lze představit jako potenciální schopnost učit se a řešit problémy. Krystalizovaná inteligence symbolizuje mentální schopnosti přímo odvozené z předchozích zkušeností. Obě schopnosti spolu pozitivně korelují. Velkým kritikem předchozích teorií byl J. P. Guilford, který vycházel z modelu, jenž obsahoval 5 operací, 4 materiály a 6 výtvorů a zkombinováním těchto tří možností vytvořil kostku o 120 poznávacích schopnostech.[1] Ovšem jeho teorie zůstala pouze modelem, kvůli omezeným možnostem oddělení schopností jedné od druhé a jejich měření. Oproti tomu Spearmanův model, rozvinutý o Cattellovo rozdělení, nadále slouží k vymezení inteligence odborníkům z různých oborů.

### 3 Sedm primárních mentálních schopností

Myšlenka této práce se odráží a inspiruje v teorii představené L. L. Thurstonem, kterou publikoval ve své knize *Primary mental abilities*. [3] Ve své teorii tvrdí, že jádro inteligence spočívá v sedmi faktorech, které nazval „primárními mentálními schopnostmi“. Jsou to:

1. chápání slov (verbal comprehension),
2. slovní plynulost (verbal fluency),
3. induktivní usuzování (inductive reasoning),
4. prostorová vizualizace (spacial visualization),
5. početní faktor (number),
6. paměť (memory),
7. rychlost vnímání (perceptual speed). [2]

Faktory byly zjištěny pomocí faktorové analýzy, metody, kterou Thurstone popisuje ve své knize *The Vectors of Mind*. [4] Při šetření bylo vybráno 56 testů na základě předběžné klasifikace kognitivních faktorů, které byly aplikovány na vzorek 240 dobrovolníků. Rozsáhlá tabulka zjištěných korelací byla faktorizovaná tzv. centroid metodou, díky níž bylo extrahováno dvanáct faktorů. Dalším testováním zvýšil počet faktorů na třináct, z nichž pouze devět mělo význam pro psychologii. Následujícím zkoumáním omezil konečný počet faktorů na sedm. V konečném výsledku tyto interpretace korespondovaly s kategoriemi, které byly původně reprezentovány použitými testy. [5] Testy se skládaly ze tří částí, a to z instrukcí a příkladů, úvodních cvičení, a testu samotného. Ne všechny části musely být použity u všech typů testů, např. testy slovosledu a aritmetiky nebyly díky jednoznačnosti doprovázeny úvodními cvičeními. Většinu dotazovaných tvořili studenti prvních a druhých ročníků Chicagské Univerzity. Studie však nebyla zaměřena jen na ně, jak bude zmíněno později. Z celkového počtu dotázaných nedokázalo testy vyplnit devět procent respondentů. [6]

O ověření výsledků i použité metody se Thurstone pokusil pomocí další studie podobného charakteru, která byla provedena na středních školách. Respondenty se stalo 286 studentů maturitních ročníků. Populace vzorku se od předchozího šetření lišila především tím, že byli studenti všeobecně oborově zaměřeni. Testy se oproti předchozím lišily pouze adaptací pro strojové zpracování, jejichž množství a obsah zůstaly stejné. Studie narazila ve výsledcích pouze na minimální odchylky oproti původnímu výzkumu. [7]

Thurstonova teorie byla zvolena jako náš referenční model z několika důvodů. Prvním je univerzálnost faktorové analýzy, která byla aplikována způsobem, jenž vyhovuje našim potřebám. Druhým důvodem je samotná



metodika provedení Thurstonova průzkumu, a to proto, že z velkého množství různých proměnných byl získán konkrétní, minimalizovaný počet sjednocujících faktorů. Navíc byl průzkum proveden nejen extenzivně (rozsahem průzkumu i jeho ověřením), ale také intenzivně, protože každému respondentovi bylo předloženo 56 testů.

Podobně jako tomu bylo u práce L. L. Thurstona, i my hodláme provést sérii testování, za použití většího množství k tomuto účelu vytvořených testů, pro zjištění „internetové inteligence“. Tento postup slibuje možnost porovnání konkrétněji specifikované oblasti inteligence společně s obecným, psychologickým pojetím.

## 4 Faktorová analýza

Faktorová analýza vznikla v oboru psychologie, jejím zakladatelem je Ch. Spearman, který ve svém článku o povaze inteligence z roku 1904 představil hypotézu o existenci „obecné intelektové schopnosti“. O další rozvoj se zasloužil L. L. Thurstone. V současné době lze vysledovat průnik faktorové analýzy i do dalších oborů, především v souvislosti s rozšiřováním výpočetních technologií a postupným transformováním a přetvářením metody do přijatelnější podoby i pro vědce z jiných oblastí, jako jsou např. sociologie, zdravotnictví, marketing, apod.[8]

Primární funkcí této metody je redukce proměnných, kterou lze provádět za předpokladu vzájemné závislosti mezi proměnnými. A dále, že jsou zmíněné závislosti důsledkem působení určitého menšího počtu neměřitelných veličin, které stojí v pozadí. Tyto veličiny označujeme jako faktory. Proces explorativní faktorové analýzy spočívá v analyzování korelace většího množství tzv. manifestních proměnných, a na tomto základě určit skupiny proměnných, které dle statistiky vytváří vztah soumělnosti. Jinými slovy, za nimi stojí společný faktor, jinak také latentní proměnná. Počet faktorů, které hledáme, by měl být ve výsledku co nejmenší a pozorované závislosti by měly být objasněny co nejjednodušeji. Po objevení a pojmenování faktoru z něj lze vytvořit novou proměnnou, která je použitelná v analýzách místo původních položek.

Základem pro postup faktorové analýzy jsou výběrové korelační a parciální korelační koeficienty. Korelační koeficient je vyjádřením těsnosti lineární závislosti mezi proměnnými a pohybuje se v rozmezí hodnot -1 až 1. Parciální korelační koeficient je pak vyjádřením těsnosti lineární závislosti dvou proměnných za předpokladu konstantnosti ostatních proměnných. Ideálním stavem při vysvětlování závislostí proměnných pomocí společných faktorů je nabytí hodnoty 0, která je charakterizovaná jako vysvětlení lineární závislosti proměnných pomocí faktoru

beze zbytku.[9]

Vedle explorativní faktorové analýzy existuje také konfirmativní faktorová analýza. Různost obou typů se projevuje tím, že explorativní faktorová analýza se vymyká statistické proceduře založené na testování hypotéz. Její výsledky jsou nejednoznačné a na počátku nebývá stanoven počet faktorů. Konfirmativní faktorová analýza se používá v případech, kdy jsou již dopředu stanoveny hypotézy o počtu a povaze faktorů. Je však používána i tam, kde povaha proměnných není známa natolik, aby mohly být hypotézy o počtu a povaze faktorů přesně stanoveny. V takovém případě mohou obě formy následovat za sebou v pořadí explorativní a konfirmativní, kdy se předběžně posoudí faktorové struktury, a teprve poté se stanoví hypotézy, které jsou dále ověřovány.[10]

Pomocí faktorové analýzy lze interpretovat výsledky v podobě faktorů, s kterými lze nadále pracovat, a zefektivnit, příp. zjednodušit další zkoumání. Jedním z uplatnění, které je zmíněno v předchozím textu, je aplikace na testy inteligence.

V naší studii chceme využít faktorovou analýzu k definování termínu „internetová inteligence“ a případně porovnat její podobnost s referenčním modelem.

## 5 Inteligence a percepce

S inteligencí jsou úzce spjaty také kognitivní procesy, které slouží jak k získávání informací z vnějšího a vnitřního prostředí, tak k jejich dalšímu zpracování. Tyto procesy zahrnují senzomotorické procesy (čítí), percepce (vnímání), učení, paměť, imaginaci a myšlení. Kognitivní psychologie se snaží na percepci nahlížet jako na soubor procesů, jejichž prostřednictvím poznáváme a organizujeme poznatky. Percepce je složitým procesem a je základem všeho poznávání.[11]

Pro potřeby naší studie i tohoto článku není momentálně nezbytné definovat pojmy učení, paměť a imaginace, i když s inteligencí souvisí. V našem zájmu je především představení pojmů percepce, případně recepce a apercepce. Je nutné oddělit od sebe recepci jako čítí, percepci jako uvědomění si a apercepce jako pochopení, v jednotlivých fázích poznávacího procesu.[11]

Recepce chápeme schopností přijímat podněty smyslovými orgány. Jejím výsledkem je počitek, což je jednotlivý znak předmětu či děje, který právě působí na člověka a obsahuje jednu vlastnost působícího předmětu. Percepci můžeme považovat za schopnost nervového systému si přijímané podněty recepcí uvědomovat. Lze ji chápat jako jednu ze složek inteligence, a také jako základní biologickou potřebu, jejímž cílem je příjem informací z vnějšího prostředí. Apercepce je potom vnímání, které využívá již dříve nabyté zkušenosti. Jedná se o fázi, při které si člověk

uvědomuje a srovnává rozdíly mezi vnitřními a vnějšími podněty.[11]

Vnímání světa člověkem je jedinečné a osobité, protože každý jedinec je jedinečný a neopakovatelný. Percepcí je ovlivněno také chování jedince, které souvisí s inteligencí. Příčinou rozdílného chování jsou většinou rozdíly v sociálních zkušenostech a individuální psychické organizaci. Předpokladem pro vnímání jsou funkční smyslové orgány, mozek a inteligence člověka.[2]

Z předcházejícího textu vyplývá, že „obecná“ inteligence a chování, kterým se projevuje, jsou úzce spjata s vnímáním reálného světa člověkem. V případě „internetové inteligence“ se můžeme pokusit o specifikaci poznávání ve virtuálním prostředí. Předpokládáme, že poznávání ve virtuálním prostředí může být oproti „obecnému“ omezené. Vycházíme z toho, že podněty ve virtuálním prostředí vnímáme především zrakem, příp. sluchem, a ne všemi smysly (zrak, sluch, čich, hmat a chuť), které využíváme v reálném životě.

## 6 Informační chování

S inteligencí neodmyslitelně souvisí i chování. Např. Thurstone v publikaci *The Nature of Intelligence* chápe chování jako prvek lidského života vznikající v organismu člověka, které je dále ovlivňováno prostředím, ve kterém se člověk vyskytuje, napříč životem. Chování představuje sebevyjádření, které v určité míře souvisí také s tím, jak nazíráme na svět, jak jej vnímáme. V rámci vlastního projevu jedince prostřednictvím chování lze vysledovat inteligentní chování.[12] Můžeme tedy předpokládat, že chování je projevem inteligence.

Zůstaneme-li u naší snahy definovat pojem „internetová inteligence“, je vhodné vycházet z jí blízkého typu chování. Našemu účelu nejlépe vyhovuje již dříve popsáné a prozkoumané informační chování, respektive některé jeho součásti. Informačním chováním se myslí „souhrnné označení pro aktivity člověka v informačním prostředí. Zahrnuje řadu informačních rolí a činností: rolí tvůrce, uživatele, zprostředkovatele informací a jejich aktivity a strategii při uspokojování informačních potřeb. Informační chování se projevuje prostřednictvím informačních zájmů, informačních potřeb a informačních požadavků.“[13] Vedle toho dle J. Steinerové souvisí informační chování se třemi hlavními činnostmi, a to s vyhledáváním, komunikováním a využíváním informací.[14]

Z výše uvedených vymezení vyplývá, že informační chování můžeme pojmut jako aktivity, které jedinec vyvíjí v informačním prostředí. Pro naši studii je klíčové především prostředí internetu. Tyto aktivity (činnosti) lze rozdělit na ty, které jsou spojené s vyhledáváním, další s komunikováním, a nakonec s využíváním informací.

Při dosahování cílů, prostřednictvím zmíněných aktivit, jsou klíčové také informační zájmy, informační požadavky, a také informační potřeby.

## 7 Návrh metodiky pro definování „internetové inteligence“

V naší studii se snažíme definovat pojem „internetová inteligence“. Tento termín již v jisté podobě existuje jako „web intelligence“, ale věnuje se aplikovanému výzkumu umělé inteligence a informačních technologií v prostředí internetu s cílem vytvořit novou generaci internetových produktů a služeb.[15] My však chápeme termín „internetová inteligence“ v souvislosti s informačním chováním uživatelů na internetu.

„Internetovou inteligenci“ tedy považujeme pouze za druh „obecné“ inteligence. Z existujících definic, které popisují „obecnou“ inteligenci jich několik využijeme při upřesňování pojmu a při specifikaci metodiky. Využijeme jak původní Spearmanovu teorii, která popisuje dva hlavní faktory (složky) inteligence (obecný a speciální faktor), tak Cattellův pohled na inteligenci, který dále faktor „g“ dělí na fluidní a krystalizovanou složku.[1]

Podle Cattellovy teorie se fluidní inteligence vyvíjí do 23 let, a později postupně klesá. Krystalizovaná inteligence stoupá až do 30 let.[1] To by mohlo pomoci určit vhodnou věkovou skupinu pro vymezení složek „internetové inteligence“.

Vycházíme-li z Thurstonova definování inteligence, můžeme z prvků, které považujeme za společné s vnímáním v internetovém prostředí a z vymezení informačního chování formulovat oblasti problémů, které nám pomohou „internetovou inteligenci“ definovat. V rámci těchto oblastí poté budeme aplikovat testy inteligence jako Thurstone. Ty budou přizpůsobeny naší metodice. Výsledky poté vyhodnotíme pomocí faktorové analýzy. Nalezené faktory nám poté pomohou vymezit definici „internetové inteligence“ a zároveň nám poskytnou prostředek k porovnání „internetové inteligence“ s „obecnou“.

Vymezení „internetové inteligence“ bude jistě z velké míry ovlivňovat omezená recepce, a tím i percepce a apercepce nových poznatků. Ve vnímání reálného světa existuje 16 druhů počitků, které se dále dělí na vnější a vnitřní.[11] Naproti tomu můžeme předpokládat, že ve virtuálním prostředí bude vnějších počitků méně.

V neposlední řadě se budeme snažit navázat naše zjištění na obecný model vnímání podnětů v prostředí internetu, ale zkoumaná problematika nás bude zajímat čistě pragmaticky. Tj. jakým způsobem ovlivňuje model „internetové inteligence“?

Thurstonova definice inteligence pomocí sedmi

primárních mentálních faktorů nám bude sloužit jako referenční model k výsledkům našeho zkoumání. Srovnání proběhne v rámci finální fáze výzkumu.

Thurstone pojímá svou definici inteligence z počátku velice široce, chování je podle něj projevem inteligence. My považujeme informační chování v prostředí internetu za projev „internetové inteligence“. Tímto tvrzením se budeme řídit při definování oblastí „internetové inteligence“ z nichž vzejdou její složky. Formulujeme tedy oblasti „internetové inteligence“:

1. informační potřeby,
2. informační požadavky,
3. informační zájmy,
4. vyhledávání,
5. komunikování,
6. využívání informací.

Informační potřeby, požadavky a zájmy korespondují přímo s definicí informačního chování dle Z. Jonáka.[13] Informační aktivity, do kterých spadá vyhledávání, komunikování a využívání informací, které formulovala J. Steinerová.[14]

Bližší vymezení informačních potřeb, požadavků, zájmů a aktivit není obsahem tohoto článku. Další práce na toto téma však budou navazovat.

## 8 Shrnutí

V tomto článku jsme stručně shrnuli různé definice inteligence a vybrali z nich ty nejvhodnější pro náš výzkum. Byl to jak faktorový přístup k inteligenci Ch. Spearmana, tak rozšířený model R. B. Cattella a popsali jsme sedm primárních mentálních schopností L. L. Thurstonea. Spolu s jeho metodikou výzkumu. Zabývali jsme se tématem omezené percepce na internetu a naznačili jsme specifikaci poznávání ve virtuálním prostředí. Vymezili jsme pojem informační chování, který souvisí s termínem inteligence. A na jeho základě jsme se pokusili formulovat termín „internetové inteligence“. Zavedli jsme debatu nad návrhem metodiky pro vymezení tohoto termínu.

## Reference

- [1] Ruisel, I. *Základy psychologie inteligence*. 1. vyd. Praha: Portál, 2000.
- [2] Sternberg, R. J. *Kognitivní psychologie*. 1. vyd. Praha: Portál, 2000.
- [3] Thurstone, L. L. *Primary Mental Abilites*. Chicago: The University of Chicago Press, 1938.

- [4] Thurstone, L. L. *The Vectors of Mind: Multiple-factor Analysis for the Isolation of Primary Traits*. Chicago: The University of Chicago Press, 1935.
- [5] Thurstone, L. L. *Primary Mental Abilities. The Measurement of Intelligence*. [Netherlands]: Springer Netherlands, 1973, s. 131-136.
- [6] Thurstone, L. L. *The Factorial Isolation of Primary Abilities. Psychometrika*. 1936, vol. 1, no.3.
- [7] Thurstone, L. L. *Experimental Study of Simple Structure. Psychometrika*. 1940, vol. 5, no. 2.
- [8] Škaloudová, A. *Faktorová analýza* [online]. Praha: Univerzita Karlova, Pedagogická fakulta, c2010 [cit. 2013-04-08]. Dostupné z: <http://userweb.pedf.cuni.cz/kpsp/skalouda/fa/index.htm>
- [9] Škaloudová, A. *Explorativní faktorová analýza. Faktorová analýza* [online]. Praha: Univerzita Karlova, Pedagogická fakulta, c2010 [cit. 2013-04-08]. Dostupné z: [http://userweb.pedf.cuni.cz/kpsp/skalouda/fa/exp\\_fak\\_analyza.html](http://userweb.pedf.cuni.cz/kpsp/skalouda/fa/exp_fak_analyza.html)
- [10] Škaloudová, A. *Konfirmativní faktorová analýza. Faktorová analýza* [online]. Praha: Univerzita Karlova, Pedagogická fakulta, c2010 [cit. 2013-04-08]. Dostupné z: [http://www.userweb.pedf.cuni.cz/kpsp/skalouda/fa/kon\\_fak\\_analyza.htm](http://www.userweb.pedf.cuni.cz/kpsp/skalouda/fa/kon_fak_analyza.htm)
- [11] Nakonenčný, M. *Základy psychologie*. vyd. 1. Praha: Academia, 2002.
- [12] Thurstone, L. L. *The Nature of Intelligence*. New York: Harcourt Brace and Company, 1924.
- [13] Jonák, Z. *Informační chování*. In *KTD: Česká terminologická databáze knihovnictví a informační vědy (TDKIV)* [online]. Praha: Národní knihovna ČR, 2003- [cit. 2013-04-08]. Dostupné z: [http://aleph.nkp.cz/F/?func=direct&doc\\_number=000000463&local\\_base=KTD](http://aleph.nkp.cz/F/?func=direct&doc_number=000000463&local_base=KTD)
- [14] Steinerová, J. *Informačné správanie: pohľady informačnej vedy*. 1. vyd. Bratislava: Centrum vedecko-technických informácií SR, 2005.
- [15] Zhong, N. *Representation and Construction of Ontologies for Web Intelligence. International Journal of Foundations of Computer Science*. Singapore: World Scientific Publishing Company, 2002, vol. 13, no. 2, s. 555-570.



## Tak mohli by stroje myslieť alebo nie?

Dezider Kamhal

Centrum pre kognitívnu vedu, Fakulta matematiky, fyziky a informatiky, UK  
Mlynská dolina, 842 48 Bratislava  
kamhal@fmph.uniba.sk

### Abstrakt

V článku sa pripomínajú a analyzujú predpoklady, z ktorých vychádzal Turing pri navrhovaní svojho testu. Hlavná pozornosť sa venuje otázke „Môžu stroje myslieť?“ (resp. „Mohli by stroje myslieť?“) a skúma sa zmysluplnosť a povaha tejto a podobných otázok. V článku sa tvrdí, že táto otázka je triviálna a netýka sa faktov, ale významov výrazov. Jej nezrozumiteľnosť môže súvisieť s jej triviálnosťou. Navyše, tvrdí sa, že náhradné otázky, ktoré Turing ponúkol, s pôvodnou otázkou nemôžu súvisieť. Ukazuje sa tiež, že Turing sa nedržal konzistentne svojho názoru, podľa ktorého je otázka „Môžu stroje myslieť?“ nezmyselná.

### 1 Úvod

Od polovice 20. storočia, kedy táto otázka znela oveľa zaujímavejšie a provokatívnejšie, sa veľa zmenilo. Nedá sa povedať, že by dnes nedala ľuďom spať a ešte väčšmi platí, že kladná odpoveď by sotva niekoho vyviedla z rovnováhy. Nebudem rozoberať, prečo je to tak a čo sa zmenilo. Skôr sa pokúsím ukázať, že stojí za to lepšie sa nad touto a podobnými otázkami aj odpoveďami na ne zamyslieť.

Ani A. M. Turing vo svojom predvídavom a ako sa ukázalo aj programovom článku *Computing Machinery and Intelligence* [2] nereagoval na podobnú otázku bezprostredne, „bez rozmýšľania“. Pripomeňme si známu pasáž z úvodu:

„Mám v úmysle uvažovať o otázke ‚Môžu stroje myslieť?‘ To by malo začať definíciami významu výrazov ‚stroj‘ a ‚myslieť‘. Tieto definície by sa mohli sformulovať tak, aby čo možno najviac odrážali normálne používanie týchto slov, ale tento prístup je nebezpečný. Ak sa má význam slov ‚stroj‘ a ‚myslieť‘ nájsť preskúmaním toho, ako sa bežne používajú, ťažko sa vyhneme záveru, že význam otázky ‚Môžu stroje myslieť?‘ a odpoveď na ňu treba hľadať v štatistickom prieskume, napríklad v Gallupovom prieskume verejnej mienky. To je však absurdné. Namiesto toho, aby som sa pokúšal

o takúto definíciu, nahradím túto otázku inou, ktorá s ňou úzko súvisí a je vyjadrená pomerne jednoznačnými slovami.“

K tejto pasáži treba uviesť niekoľko dôležitých poznámok. Po prvé, Turing predpokladal, že *porozumeniu* otázke by malo predchádzať vymedzenie významu oboch kľúčových termínov, t. j., že bez toho by sme jej zrejme nerozumeli, alebo prinajmenšom rozumeli viacznačne. Po druhé, slovo „môžu“ v otázke zrejme nepokladal za problematické. Po tretie, jednou z možností by podľa neho bolo pri definovaní termínov čo najviac sa pridržiavať ich bežného, normálneho používania. Po štvrté, podľa neho by to znamenalo robiť príslušné štatistické prieskumy a to mu pripadá absurdné. Po šieste, hovorí o *nahradení* tejto otázky inou, formulovanou relatívne jednoznačne. A napokon, predpokladá, že náhradná otázka s pôvodnou úzko súvisí.

Pouvažujme o uvedených Turingových predpokladoch či názoroch a posúďme, či bol Turing vo svojich predpokladoch a názoroch konzistentný.<sup>1</sup> Predovšetkým, požadovať definíciu termínov, aby sme rozumeli vete, zloženej len z *bežne používaných slov*, je prinajmenšom zvláštne. Slová „stroj“ a „myslieť“ (aj „môžu“) patria do bežného jazyka, takže prípadná nezrozumiteľnosť otázky „Môžu stroje myslieť?“ nie je dôsledkom toho, že obsahuje slová s neznámym významom. Okrem toho, predpokladať, že treba štatisticky zisťovať, ako sa tieto výrazy používajú znamená predpokladať, že sme si neosvojili vlastný jazyk, teda že sa ho ešte len musíme naučiť.

### 2 Zrozumiteľnosť a nezrozumiteľnosť otázok

Ako je to so zrozumiteľnosťou tejto otázky, s chápaním jej významu? Na ilustráciu uvedieme otázku, ktoré majú podobnú štruktúru – pýtame sa

<sup>1</sup> Z citovanej úvodnej pasáže by sa dali odvodiť aj ďalšie predpoklady, z ktorých Turing vychádzal.

v nich, či nejaký druh objektov môže čosi robiť alebo, či nejaký druh objektov môže mať vlastnosť, vyjadrenú predikátom: „Môžu stoličky rozprávať?“, „Môžu transformátory vnímať?“, „Môžu počítače smútiť?“ atď. Prvý problém s takýmito vetami sa týka použitia slovesa „môžu“, pri ktorom sa Turing v článku nepozastavil.<sup>2</sup> Pýtame sa: 1. na schopnosť nejakých *existujúcich* objektov niečo robiť, alebo: 2. či by objekty daného druhu *principiálne mohli* alebo *nemohli* robiť niečo v budúcnosti? Úvodnú otázku by sme potom preformulovali takto: „*Mohli* by stroje myslieť?“

Ak chápeme otázku v prvom zmysle, odpoveď, ak vôbec nejaká, by bola záporná.<sup>3</sup> Zaujímavé je považovať, ako sme k zápornej odpovedi došli. Museli sme tieto existujúce objekty pozorovať alebo vykonať nejaké experimenty? Ako by napríklad vyzeralo skúmanie toho, či stoličky rozprávajú? Nie, na získanie zápornej odpovede nepotrebujeme nijaké empirické skúmanie. Predstava empirického skúmania, či stoličky rozprávajú, transformátory vnímajú alebo počítače smútiť (!) je skôr zábavná. Pravda, ak predpokladáme, že príslušné slová („stolička“, „rozprávať“, „transformátor“, „vnímať“, „počítač“, „smútiť“) používame obvyklým spôsobom. Všimnime si, že takto chápaná otázka sa jednoducho netýka faktov a je analogická otázke, či môžu byť štvorce okrúhle. Nezrozumiteľnosť takýchto otázok v prvom prípade teda súvisí s tým, že sa pýtame, či nejaký druh objektov môže mať vlastnosť (alebo vykonávať činnosť), ktorú nemá (alebo nevykonáva). To, že ju nemá, alebo ju nevykonáva, je zachytené v našom používaní slov a to môžeme vyjadriť napr. vetou: „Ale stroj predsa nemôže myslieť!“ Položme si spolu s Wittgensteinom otázku: „Je to empirická veta?“ A pripomeňme si aj jeho odpoveď: „Nie. Iba pri človeku a jemu podobných hovoríme, že myslia. Hovoríme to aj o bábikách, ba aj o duchoch. Pozeraj sa na slovo ‚myslieť‘ ako na nástroj!“ ([5], S. 149). Nezrozumiteľnosť takýchto otázok nie je dôsledkom nezmyselnosti alebo viacznačnosti použitých slov, ale toho, že nerozumieme, akú funkciu by mala otázka plniť.

Čo by sa zmenilo, keby sme otázku chápali v druhom zmysle, teda či by  *mohli*  príslušné objekty niečo robiť v budúcnosti? Vyššie uvedené otázky by zneli nasledovne: „*Mohli* by stoličky rozprávať?“, „*Mohli* by transformátory vnímať?“, „*Mohli* by

<sup>2</sup> Z ďalšieho textu je však jasné, že otázku chápal v druhom zmysle.

<sup>3</sup> „Odpoveďou“ na túto otázku by mohli byť aj rozpaky a mlčanie.

počítače smútiť?“. Za predpokladu, že v budúcnosti sa nezmení používanie slov označujúcich príslušný druh vecí alebo činností je zrejmé, že odpoveď, ak vôbec nejaká, bude záporná. Nezmení sa ani povaha týchto otázok, ostanú konceptuálnymi otázkami, teda otázkami týkajúcimi sa významov výrazov. Zamyslime sa nad dvoma nasledujúcimi otázkami: Kedy by sme na tieto otázky mohli odpovedať kladne? Môžu byť v tejto podobe empirickými otázkami? Ako sme už naznačili, odpoveďou na prvú otázku je, že by sa museli zmeniť významy (používanie) termínov týkajúcich sa príslušných druhov objektov alebo činností. Teda v prípade otázky („Mohli by stoličky rozprávať?“) by sme buď museli termín „stolička“ používať tak, aby sa vzťahoval na predmety, ktoré *môžu* rozprávať, alebo by sme sloveso „rozprávať“ používali aj na činnosti, ktoré *môžu* vykonávať aj stoličky.<sup>4</sup> Pokiaľ ide o empirickú, prípadne konceptuálnu povahu otázky, tak jej neempirická povaha je ešte zrejmejšia, pretože nemáme čo pozorovať, resp. nemáme s čím experimentovať. Hovoríme totiž o hypotetických možnostiach, a navyše možnostiach, ktoré sa netýkajú faktov, ale nášho používania slov.<sup>5</sup>

### 3 Náhradné otázky?

Všimnime si ďalšie Turingove tvrdenia z úvodu, jednak tvrdenie o tom, že nahradí pôvodnú otázku inou, vyjadrenou jednoznačnejšie, a jednak tvrdenie o úzkej súvislosti novej otázky s pôvodnou. V nasledujúcej časti Turing rozvíja predstavu imitačnej hry, ktorú hrajú stroj, žena a pýtajúci sa. Ten je v inej miestnosti a má zistiť, kto je stroj a kto je žena. Pripomeňme si, akú otázku či otázky ponúkol Turing ako náhradu pôvodnej:

„Čo sa stane, ak v tejto hre stroj preberie úlohu muža A? Bude sa pýtajúci sa mýliť pri takejto hre rovnako často, ako keď hru hrajú muž a žena? Tieto otázky nahrádzajú našu pôvodnú ‚Môžu stroje myslieť?‘“<sup>6</sup>

<sup>4</sup> Samozrejme, nešlo by o nevinnú a izolovanú zmenu v používaní termínov, pretože *slová nie sú siroty*. Zmena významov by sa mohla dotknúť na prvý pohľad neuveriteľného počtu ďalších príbuzných termínov. Stačí napríklad uviesť slová „konat“, „vykonávať“, „činnosť“, ale aj „ústa“, „jazyk“, „poslucháč“ a pod.

<sup>5</sup> O tom, že samému Turingovi išlo o otázku, ktorá sa nevzťahuje na existujúce stroje, ale na „predstaviteľné“, svedčí nasledovná formulácia: „Pýtame sa, či... si vieme predstaviť počítače, ktoré by obstáli.“ [2], s. 436.

<sup>6</sup> Imitačnú hru hrajú pôvodne muž (A), žena (B) a pýtajúci sa (C). Pozri [2], s. 434.

Ďalej špecifikuje, čo máme mať na mysli pod „strojom“ a zužuje význam slova „stroj“ vo formulácii otázky na význam výrazu „digitálny počítač“, ktorý neskôr ďalej spresňuje. Náhradná otázka potom znie, či si vieme predstaviť počítač, ktorý by obstál v imitačnej hre, prípadne v jej údajne presnejšej podobe: „Je pravda, že ak upravíme tento [konkrétny] počítač... môže vyhovujúco hrať úlohu osoby A v imitačnej hre, pričom úlohy B sa ujíma človek?“ ([2], s. 442).

Pri všetkých Turingových nových otázkach možno pochybovať o tom, či sú náhradou pôvodnej otázky v tom zmysle, že spresňujú jej význam. Nech akokoľvek zmeníme význam slova „stroj“ v pôvodnej otázke,<sup>7</sup> stále totiž ostáva problém, či tieto nanovo, inak, alebo presnejšie vymedzené „stroje“ myslia. Museli by sme totiž súhlasiť s Turingovým predpokladom, že úspešné zvládnutie imitačnej hry je ekvivalentné alebo sčasti ekvivalentné mysleniu. Inak povedané, že výrazom „úspešne zvládnuť imitačnú hru“ možno nahradiť výraz „myslieť“, resp. že výrazom „myslieť“ možno nahradiť výraz „úspešne zvládnuť imitačnú hru“.

#### 4 Nezmyselnosť pôvodnej otázky

Uvažovanie o vzťahu medzi oboma výrazmi a takisto o vzťahu medzi pôvodnou a náhradnými otázkami si však môžeme ušetriť, pretože na s. 442 svojho článku Turing píše: „Som presvedčený, že pôvodná otázka ‚Môžu stroje myslieť?‘ je príliš nezmyselná, aby si zaslúžila diskusiu. Jednako som však presvedčený, že na konci storočia sa používanie slov a všeobecná mienka vzdelancov zmenia natoľko, že človek bude schopný hovoriť o mysliacich strojoch bez očakávania, že mu budú odporovať.“ Ak je totiž pôvodná otázka nezmyselná či „príliš nezmyselná“<sup>8</sup>, sotva možno pokladať za jej náhradu (opäť v tom zmysle, že spresňuje jej význam) ktorúkoľvek z nových otázok, ktoré – prinajmenšom pre Turinga – boli zmysluplné. A rovnako nemožno hovoriť o úzkej súvislosti náhradných, zmysluplných otázok s pôvodnou, nezmyselnou.

Uznanie nezmyselnosti otázky ‚Môžu stroje myslieť?‘ však Turingovi pri iných príležitostiach nebránilo hovoriť o „mysliacich strojoch“ a o možnosti vytvoriť počítačový stroj, ktorý by myslel,

<sup>7</sup> Samozrejme, s výnimkou zmien jeho významu, ktoré by umožňovali zahrnúť pod „stroje“ aj „ľudi“, prípadne iné živočíchy, ktorým sme ochotní pripísať myslenie.

<sup>8</sup> Nie je celkom jasné, či a ako sa „prílišná nezmyselnosť“ odlišuje od prostej „nezmyselnosti“.

ani hovoriť o mysliacich strojoch alebo o myslení v súvislosti s počítačmi. Ako to noblesne vyjadril J. Copeland: „[Turing] Určite nedopustil, aby mu tento názor zabránil pomerne často si dopriať takúto diskusiu.“ ([1], s. 491).

Môžeme to ilustrovať pasážami z jeho prednášky odvysielanej na BBC „Môžu počítačové stroje myslieť?“ alebo z odvysielanej diskusie „Možno o automatických počítačových strojoch hovoriť, že myslia?“ (pozri [3] a [4]).

V uvedenej prednáške zazneli aj nasledujúce slová: „... náš hlavný problém, ako naprogramovať stroj tak, aby napodobňoval mozog alebo, ako by sme mohli povedať stručnejšie, aj keď nepresnejšie, aby myslel“ alebo „Nepokúsím sa veľa hovoriť o tom, ako sa tento proces ‚naprogramovania stroja, aby myslel‘ má urobiť“; a končí ju slovami: „... verím, že pokus vytvoriť mysliaci stroj nám veľmi pomôže pri odhaľovaní toho, ako my sami myslíme.“ A v diskusii v súvislosti so svojím testom okrem iného hovorí: „Mohli by ste ho nazvať testom na zisťovanie toho, či daný stroj myslí...“, aj keď to vzápätí koriguje a zmierňuje.

Zaujímavejšie však je, že druhá veta z vyššie citovanej časti článku [2] dáva do súvislosti *zmenu v používaní slov* s možnosťou hovoriť o mysliacich strojoch bez toho, aby človeku odporovali. Inými slovami, ak zmeníme význam slova „stroj“ (prípadne slova „myslieť“), ľudia bez problémov akceptujú kladnú odpoveď na otázku, či stroje môžu alebo by mohli myslieť. To je v súlade s presvedčením, že pôvodná otázka aj odpoveď na ňu sa netýka faktov, ale významov (používania) príslušných výrazov. Zároveň je to dôkaz, že A. Turing si prinajmenšom niekedy povahu pôvodnej otázky uvedomoval.

#### Pod'akovanie

Tento príspevok vznikol ako súčasť riešenia grantovej úlohy VEGA č. 1/0644/13 „Metafilozofia ako pragmatická analýza filozofických výpovedí“, a KEGA 076-UK4/2013.

#### Literatúra

[1] B. J. Copeland (ed): *The Essential Turing. Seminal Writings in Computing, Logic, Philosophy, Artificial Intelligence, and Artificial Life. Plus The Secrets of Enigma*, Oxford University Press, 2004

- [2] A. M. Turing: Computing Machinery and Intelligence, *Mind*, New Series, Vol. 59, No. 236. (1950), 433 – 460.
- [3] A. M. Turing: Can Digital Computers Think?, prednáška vysielaná v rádiu BBC 15. mája 1951. In: [1], s. 482 – 486.
- [4] A. Turing, R. Braithwaite, G. Jefferson, M. Newman: Can Automatic Calculating Machines Be Said to Think?, diskusia vysielaná v rádiu BBC 14. januára 1952. In: [1], s. 494 – 506.
- [5] L. Wittgenstein: *Filozofické skúmania*. Pravda, Bratislava, 1979.



# Čo ešte ukáže čas vo výskume efektu spätného pohľadu?

Matúš Konečný

Fakulta matematiky, fyziky a informatiky, Univerzity Komenského v Bratislave  
84248 Mlynská dolina, Bratislava  
konecny.matus@gmail.com

## Abstrakt

Efekt spätného pohľadu je skreslenie v usudzovaní o udalostiach, ktorých výsledok je už známy. Predpokladá sa, že pozostáva z troch separovateľných javov – skreslenia pamäte, dojmu nevyhnutnosti a dojmu predvídateľnosti. Skúmanie časových charakteristík usudzovania v priebehu vývinu udalostí je novým a perspektívnym experimentálnym prístupom. V príspevku uvádzam čitateľa do problematiky výskumu efektu spätného pohľadu a opisujem procesy vedúce k zvyšovaniu dojmu nevyhnutnosti výsledku. Predstavujem tiež výzvy, ktoré stoja pred výskumom časových charakteristík tohto efektu a na návrhu experimentu demonštrujem, ako môže tento prístup prispieť do odbornej diskusie o jeho povahe.

## 1 Úvod

Efekt spätného pohľadu (ESP) sa prejavuje ako tendencia nedoceňovať mieru neistoty výsledku udalosti po tom, ako je človeku výsledok už známy. Prvýkrát ho experimentálne skúmal Baruch Fischhoff v roku 1975 [4]. Odvetdy sa stal učebnicovým príkladom kognitívnych chýb a prilákal pozornosť mnohých výskumníkov. Dodnes sa mu venovalo viac než 800 štúdií [15].

Cieľom tohto článku je predstaviť koncepciu kompozicionality efektu spätného pohľadu, ktorá navrhuje, že sa skladá z troch separovateľných efektov – skreslenia pamäte, dojmu nevyhnutnosti a dojmu predvídateľnosti. Čitateľovi tiež priblížim teóriu kauzálneho modelu, ktorá vysvetľuje vznik dojmu nevyhnutnosti ako následku procesov kauzálnej atribúcie.

V ďalších častiach sa budem zaoberať skúmaním časových charakteristík efektu spätného pohľadu, ktorý patrí k najnovším experimentálnym prístupom. Explicitne sa mu zatiaľ venovali 2 publikované články [3,6]. Výsledky experimentov venujúcich sa časovým charakteristikám ESP odhalili vlastnosti tohto javu, ktoré boli predchádzajúcemu skúmaniu doslova ukryté v čase. Okrem nových vhládov do ESP, doterajší výskum časových charakteristík otvoril nové otázky o tom, ako tento jav funguje. Načrtnem aktuálne problémy, ktoré môžu byť zaujímavými oblasťami pre budúci výskum

v tejto oblasti a mať teoretické implikácie ako pre vysvetľovanie odlišností medzi jednotlivými efektmi v spätnom pohľade, tak aj pre teóriu kauzálneho modelu. Konkrétnu predstavu ilustrujem na návrhu experimentu, ktorý by mohol pomôcť zodpovedať na niektoré z opísaných otázok.

## 2 Efekt spätného pohľadu

V prúde udalosti, teda predtým ako poznáme výsledok, sa budúcnosť môže zdať neistejšia ako keď sa na tú istú udalosť pozrieme späť s odstupom času. V dobe prípravy tohto článku, začiatkom apríla 2013, existujú vo svete dohady o hrozbe jadrového útoku zo strany Kórejskej ľudovodemokratickej republiky. Keď táto situácia pomínie, s výhodou poznania výsledku sa vytratí aj časť neistoty v usudzovaní o pravdepodobnosti takéhoto útoku. Či už k nemu príde alebo nie, bude pre mnohých ťažké odolať dojmu, že výsledok bol predvídateľný a nákladné vojenské opatrenia neopodstatnené alebo naopak, že bolo možné predísť stratám na ľudských životoch. Pre niektorých autorov (napr. [3]) je neschopnosť zrekonštruovať usudzovanie bez znalosti výsledku udalosti integrálnou súčasťou efektu spätného pohľadu. Populárne býva nazývaný ako „knew it all along“ efekt (voľne preložené: „bolo mi to celý čas jasné“). Alebo ako sa hovorí u nás: „po vojne je každý generál“.

Prítomnosť efektu spätného pohľadu bola pozorovaná aj v rôznych iných úlohách ako odhadoch výsledkov udalostí, napríklad pri faktografických otázkach (napr. [5,18]). Tento článok je zameraný na efekt spätného pohľadu pri posudzovaní udalostí a niektoré opísované charakteristiky a procesy sú špecifické pre túto oblasť usudzovania a nemusia sa vzťahovať k chybám v spätnom pohľade pri inom type úloh.

Tendenciu približovať svoje odhady po poznaní výsledku ku skutočným sa niekedy interpretuje ako následok asimilácie nových informácií [4]. Spoznanie výsledku zvyšuje aktiváciu reprezentácií konzistentných s ním a tým v spätnom pohľade odchyľuje rekonštrukciu usudzovania pred spoznaním výsledku [16]. Nejde však o obyčajný prípad učenia sa. Ako Roesse a Vohs [15]

zdôrazňujú, zmeniť svoje predstavy o svete je po skúsenosti prínosné, ale je chybou predpokladať, že niekto v minulosti videl do vešteckej gule a vedel to, čo je teraz už známe. Klasické ponímanie efektu spätného pohľadu poukazuje práve na takýto druh chýb v usudzovaní.

Pri skúmaní ESP sa najčastejšie používajú dva experimentálne dizajny – pamäťový a hypotetický [12]. Pri pamäťovom najprv participanti robia odhady bez poznania výsledku a po tom ako im je skutočný výsledok oznámený je ich úlohou spomenúť si na svoj pôvodný odhad. Alternatívou k vnútrosubjektóvemu dizajnu je hypotetický medzisubjektívny dizajn. V takom experimente sú porovnávané odhady dvoch skupín participantov, pričom kontrolná robí svoje odhady bez vedomosti výsledku a experimentálna s vedomosťou výsledku. Pri hypotetickom dizajne sa považuje za kľúčové, aby experimentálna skupina bola inštruovaná robiť svoje odhady tak, akoby nevedela aký je výsledok, teda „vžiť sa do kože“ participantov z kontrolnej skupiny. Spomínané dizajny reflektujú klasické normatívne ponímanie ESP ako kognitívnej chyby v spätnom pohľade v porovnaní s posudzovaním z pohľadu naivného pozorovateľa.

## 2.1 Komponenty ESP

Definície a vysvetlenia efektu spätného pohľadu nie sú medzi výskumníkmi úplne rovnaké. Už z porovnania hypotetického a pamäťového dizajnu experimentu vyplývajú teoretické dôsledky o povahe ESP. Kým úloha v pamäťovom dizajne vyžaduje, aby si participant vyvolal svoj predchádzajúci úsudok, v hypotetickom dizajne tomu tak nie je. Dokonca daný úsudok s najväčšou pravdepodobnosťou ešte nikdy nerobil. Kritériom efektu spätného pohľadu je v oboch prípadoch odlišnosť medzi veľkosťou odhadov v naivnom a v spätnom pohľade. Takéto chápanie ESP však nezabezpečuje, aby sme pri oboch dizajnoch mohli hovoriť o následkoch tých istých kognitívnych procesov.

V prípade pamäťového dizajnu sa zdá byť nevyhnutné, aby medzi príčinami vzniku ESP figurovali pamäťové procesy. V prípade, že by si participanti bezchybne vybavili svoje predchádzajúce úsudky a splnili by úlohu ako boli inštruovaní, nemalo by prísť k žiadnemu ESP. Keďže toto neplatí v hypotetickom dizajne, musia sa na vzniku ESP podieľať aj nejaké iné kognitívne procesy. Rozdiely v dizajne nie sú jediné odlišnosti medzi metódami používanými na skúmanie ESP a teda situácia je ešte o niečo komplikovanejšia.

Ukazuje sa, že efekt spätného pohľadu pozostáva z troch samostatných komponentov, ktoré sú dôsledkami teoreticky aj empiricky separovateľných procesov – skreslenia pamäte, dojmu nevyhnutnosti a dojmu

predvídateľnosti [2,8]. Okrem procesov, ktoré ich najviac ovplyvňujú, plnia aj odlišné psychologické funkcie [8]. Nestler a kol. [8] sa domnievajú, že fenomenologické odlišnosti medzi komponentmi ESP môžu byť také významné, že je zavádzajúce uvažovať o efekte spätného pohľadu ako jednotnom jave.

Skreslenie pamäte zachytáva neschopnosť presného vybavenia si svojho predchádzajúceho úsudku [15]. Je zrejme, že táto zložka ESP je relevantná len pre úlohu v pamäťovom dizajne experimentu. Dal by sa ilustrovať typom výrokov ako „hovoril som, že to tak dopadne“, ktoré niekedy ľudia chybne tvrdia už s výhodou poznania toho, ako udalosti skutočne dopadli. Vzáhuje sa teda na subjektívne presvedčenia o vlastných reprezentáciách stavov sveta v minulosti [2]. Na skreslenie pamäte vplyvajú pamäťové procesy kódovania, uchovávaní, vyvolávaní, ukotvení pri získaní spätnej väzby o výsledku a procesy rekonštrukcie pôvodného úsudku [2].

Dojmy nevyhnutnosti a predvídateľnosti ovplyvňujú úsudky v spätnom pohľade pri oboch typoch klasických dizajnov. Dojem predvídateľnosti, sa rovnako ako skreslenie pamäte, vzáhuje na subjektívnu reprezentáciu, kým dojem nevyhnutnosti odráža predstavy o objektívnom stave sveta [2]. Hlavným procesom, ktorý vplyva na dojem nevyhnutnosti je kauzálna atribúcia a prepracovanie kauzálnych vzťahov medzi antecedentmi (predchádzajúcimi okolnosťami alebo udalosťami) a výsledkom po poznaní výsledku. Tento proces vedie k presvedčeniu „muselo to tak dopadnúť“ a detailnejšie sa mu venujem v časti 3 o teórii kauzálneho modelu. Dojem predvídateľnosti môže byť ovplyvnený dojomom nevyhnutnosti, ale reflektuje tiež metakognitívne fenomény ako pocit prekvapenia pri spoznaní výsledku, či jednoduchosti tvorby úsudku a vyvoláva pocit „vedel som, že to tak dopadne“ [2,15].

Tieto jednotlivé komponenty nevznikajú vždy nezávisle od seba. Vybavovanie si okolností udalosti môže ovplyvniť, ktorým antecedentom ľudia pripíšu úlohu v kauzálnom vysvetľovaní výsledku. Neúspešnosť vytvorenia takéhoto vysvetlenia vyvoláva pocit prekvapenia [10], ktorý môže ovplyvniť dojem predvídateľnosti. V smere od vyšších kognitívnych procesov, dojem schopnosti predvídať výsledok by mohol viesť k nadhodnoteniu kauzálnej sily antecedentov. Posilnenie spojení konkrétnych antecedentov s výsledkom môže tiež ovplyvniť selektívitu ich vyvolania z pamäte [1].

Autori koncepcie kompozicionality ESP boli úspešní v experimentálnej disociácii jednotlivých komponentov [2,8,9]. Ukázali, že je možné manipulovať experimentálnymi podmienkami tak, aby boli ovplyvnené komponenty ESP separátne. Na merania dojmov nevyhnutnosti a predvídateľnosti vyvinuli dve škály.

Dojem nevyhnutnosti merali použitím položiek s Likertovými škálami na formuláciách nasledujúcich výrokov prispôbených konkrétnemu podnetovému materiálu: „Za daných okolností nebol iný výsledok možný“, „Veľa faktorov by ešte mohlo ovplyvniť výsledok udalosti, takže iné výsledku sú stále možné“ a „Výsledok je nevyhnutný“. Rovnaký typ škály s odlišnými výroky použili aj na meranie dojmu predvídateľnosti: „Celý čas som vedel, ako to dopadne“, „Výsledok sa dal ťažko predpovedať“ a „Výsledok je predvídateľný“ [8]<sup>1</sup>. Úlohou participantov pri odpovedaní na spomenuté položky nebolo spomenúť si na svoju predchádzajúcu odpoveď, ale prehodnotiť svoj pôvodný úsudok. Autori argumentujú tým, že použitie rovnakej inštrukcie ako pri meraní skreslenia pamäte aj pri meraní dojmu nevyhnutnosti, či predvídateľnosti by spôsobila, že všetky tieto merania by odrážali skreslenie pamäte. Tieto škály tiež neboli použité pri tých istých participantov súčasne, aby autori predišli novej tendencii odpovedať konzistentným spôsobom na oboch mierach [8].

Nepoužitie inštrukcie na zanedbanie znalosti výsledku alebo spomenutia si na svoj predchádzajúci odhad je vo výskume spätného pohľadu novým javom. Nestler, Blank a Egeloff [8] poukazujú aj na to, že neschopnosť zanedbať svoju vedomosť výsledku nemusí byť jediný aspekt zaujímavý pre porozumenie spätného pohľadu. Skúmanie zmien percepcie udalostí v spätnom pohľade môže odhaliť psychologicky relevantné aspekty [8]. Pýtanie sa otázok, ktoré neobsahujú inštrukciu spomenúť si na svoju predchádzajúcu odpoveď sa zatiaľ javí ako dôležité pre schopnosť separovať jednotlivé komponenty ESP. Odlíšenie chýb v spätnom pohľade od vplyvov učenia sa ale týmto krokom zdá komplikovanejšie a prináša so sebou niektoré teoretické aj metodické otázky, ktoré vyžadujú zváženie a ďalšiu empirickú podporu (pre detailnejšiu diskusiu o tejto problematike pozri [8]).

Aby overili úlohu inštrukcie v skúmaní ESP, vykonali experiment, v ktorom dve skupiny participantov posudzovali volebné výsledky a dojem nevyhnutnosti. Úlohou pri oboch skupinách participantov bolo po spoznaní výsledkov volieb spomenúť si na svoje pôvodné odhady (skreslenie pamäte). Jedna skupina participantov si tiež mala spomenúť na svoj dojem nevyhnutnosti, kým druhá bola inštruovaná svoj pôvodný úsudok prehodnotiť. Ani v jednej skupine neprišlo

---

<sup>1</sup> Na meranie dojmov predvídateľnosti a nevyhnutnosti používali Blank a kol. [2] a Nestler a kol. [8] rôzne série výrokov. Tri výroky z každej škály boli vybrané na základe odporúčaní autorov vynechať položky, ktoré boli redundantné a neprispievali k zvýšeniu reliability. Uvádzané položky sú preložené z anglického prekladu výrokov v nemčine.

k skresleniu pamäte pri predikcii volebných výsledkov. Pri oboch skupinách sa však prejavil efekt spätného pohľadu v dojme nevyhnutnosti, ale neprejavil sa žiaden významný rozdiel medzi skupinami v závislosti od inštrukcie. Nestler a kol. [8] usudzujú, že inštrukcie na zanedbanie vedomosti výsledku nemusia zohrávať pri skúmaní efektu spätného pohľadu takú kľúčovú rolu ako sa všeobecne predpokladá.

Domnievam sa však, že autori nezohľadnili možný vplyv odlišností týchto dvoch posúdení, ktoré merali. Kým sa mnoho z nich možno už opakovane zaoberalo svojimi predstavami o výsledku volieb a mohli ich mať založené napríklad aj na predvolebných prieskumoch, mohlo sa jednať o už vtedy elaborované odhady. Teda to nemuselo byť prvý ani posledný krát, kedy sa takouto otázkou zaoberali. Táto skutočnosť by vysvetľovala aj prečo neprišlo k skresleniu pamäte. Na rozdiel od toho asi nikdy predtým, ani potom neodpovedali na otázku, ktoré sú súčasťou škály dojmu nevyhnutnosti. V ďalšom výskume bude dôležité reflektovať možné odlišnosti v spätnom pohľade v závislosti na použitej inštrukcii a porozumieť limitám takejto metódy a jej teoretickým implikáciám.

Rozdelenie efektu spätného pohľadu na 3 samostatné komponenty, či efekty, odráža dôležité aspekty tohto javu, ktoré boli v predchádzajúcom výskume prehliadané. Rôzne experimenty skúmali odlišné aspekty tohto javu, čo viedlo k istej konfúzii. Porozumenie separovateľnosti a interakcie jednotlivých týchto komponentov predkladá výzvu pre budúci výskum a môže významne napomôcť pri vysvetľovaní procesov, ktoré sa podieľajú na vzniku efektu spätného pohľadu.

### 3 Teória kauzálneho modelu

Teória kauzálneho modelu vysvetľuje vznik ESP pri posudzovaní výsledkov udalostí [1]. Predpokladá, že po spoznaní výsledku prichádza k aktivácii procesu hľadania zmyslu (z ang. „sense making“) prepracovaním kauzálnych spojení antecedentov s výsledkom. Najväčší vplyv má na vyvolanie dojmu nevyhnutnosti, ale ako som už spomenul v časti 2.1, ten môže úzko interagovať aj s dojomom predvídateľnosti [2,8].

Ľudia po získaní spätnej väzby o výsledku selektujú a hodnotia informácie o udalostiach a svoje všeobecné poznatky o podobných udalostiach tak, aby mu čo najlepšie porozumeli [1]. Majú tendenciu tvoriť skôr pozitívne ako negatívne vzťahy medzi príčinami a výsledkom tak, aby si vytvorili konzistentný kauzálny model danej udalosti.

Dôsledkom procesu hľadania zmyslu je efekt spätného pohľadu v prípade, že je identifikovaný aspoň jeden antecedent, ktorý umožní vytvorenie kauzálneho

vysvetlenia výsledku [1]. V prípade, že nie je možné vytvoriť takéto kauzálne spojenie, ESP sa neprejaví. K tomu môže prísť v prípade, že ľudia nie sú schopní porozumieť príčinám vedúcich k výsledku. Vtedy dokáže doplňujúce vysvetlenie pomôcť vo vytvorení kauzálneho modelu a zvýšeniu dojmu nevyhnutnosti [8]. Alternatívne, k podobnému javu príde aj keď nie je participantom prezentovaný žiaden antecedent, ktorý by do mohli do tvorby svojho vysvetľujúceho modelu zapracovať [19].

Z pohľadu teórie kauzálneho modelu je podmienkou vzniku ESP aktivácia procesu hľadania zmyslu. Ak táto podmienka nie je splnená, pretože je výsledok už v priebehu udalosti taký predvídateľný, že nie je potrebné kauzálny model prepracovať, ESP sa nemusí prejavíť [11,19]. Takéto prípady sú spájané s veľmi nízkou prekvapivosťou výsledku. Určitá miera prekvapivosti sa ukazuje byť nevyhnutná pre vznik ESP [7]. Naopak, príliš vysoká prekvapivosť výsledku vedie k neúspešnému ukončeniu procesu hľadania zmyslu, čo môže spôsobiť obrátenie ESP [10]. Vzťah medzi prekvapivosťou výsledku a ESP je teda nelineárny. Ľudia majú tendenciu preceňovať nevyhnutnosť a predvídateľnosť udalostí len do tej miery, nakoľko sú schopní pochopiť, prečo k výsledku prišlo [14] a zároveň keď sú ich pôvodné očakávania aspoň čiastočne vyvrátené.

Proces hľadania zmyslu môže zahŕňať formulovanie kontrafaktových úvah [14], teda zvažovanie, čo by sa stalo, keby boli okolnosti iné. Niekedy môže byť prirodzené zamyslieť sa: „Keby som začal túto prácu písať skôr, nemusel by som si teraz odopierať spánok.“ Formulovanie takýchto úsudkov vedie ku kauzálnemu vysvetľovaniu výsledku, a teda aj k zvyšovaniu posudzovanej pravdepodobnosti, že k takémuto výsledku za daných podmienok príde [14]. Kontrafaktové myslenie môže efekt spätného pohľadu aj redukovať. Formulovanie kontrafaktových úsudkov o tom, ako inak by mohla udalosť dopadnúť nezávisle od zmeny antecedentov znižuje silu kauzálneho spojenia medzi antecedentami a výsledkom, čím môže zmenšiť aj dojem nevyhnutnosti za daných podmienok [14].

Hľadanie zmyslu výsledku súvisí s vytváraním kauzálnych štruktúr vysvetľujúcich výsledok. Centrálnym procesom, ktorým k tomuto prichádza je kauzálna atribúcia. Pretváraním kauzálneho modelu, ktorý obsahuje udalosti, či podmienky predchádzajúce vzniku výsledku sa stáva výsledok zdanlivo predvídateľnejší a nevyhnutnejší [14]. Kontrafaktové myslenie môže byť tiež dôsledkom aktívnej snahy porozumieť minulosti a je produktom atribučných procesov už po vedomosti výsledku [14].

## 4 Časové charakteristiky ESP

Otázka, či má moment tvorby posúdenia pravdepodobnosti výsledku udalosti vplyv na efekt spätného pohľadu bola prvýkrát experimentálne skúmaná na animácii dopravnej nehody [13]. V hypotetickom dizajne porovnávali posúdenia pravdepodobnosti výsledku troch skupín participantov. Jedna skupina robila odhady pravdepodobnosti bez vedomosti výsledku v momente blízkom vyvrcholeniu deja. Tieto odhady boli porovnávané s odhadmi v spätnom pohľade v tom istom časovom momente a tiež v momente, ktorý bol od výsledku vzdialený. Výsledky ukázali významné odlišnosti v odhadoch medzi všetkými tromi skupinami. Odlišnosti medzi pravdepodobnostnými odhadmi v rôznych momentoch boli dovtedajším výskumom ukryté v čase, pretože išlo vždy len o porovnanie kde bola podaná informácia a výsledku alebo nebola. Teda nikdy neboli porovnávané posúdenia pravdepodobnosti v iných časoch ako práve pred vyvrcholením deja. Experimentálny dizajn Roeseho a kol. [13] však ešte neumožnil detailnejšie analyzovať trajektórie pravdepodobnostných posúdení v čase.

Časť autorov spomenutého výskumu neskôr publikovala sériu experimentov explicitne adresujúcich otázku časového vývinu efektu spätného pohľadu [3]. V experimentoch s pamäťovým dizajnom robili participantí opakované posúdenia pravdepodobnosti výsledku dvoch príbehov prezentovaných ako text a dvoch animácií dopravných nehôd. Animácie aj textové materiály boli rozdelené na viacero segmentov, po ktorých robili participantí posúdenia pravdepodobnosti výsledku počas vývinu deja. Efekt spätného pohľadu merali porovnávaním kriviek pravdepodobných posúdení v jednotlivých momentoch prvého posúdenia a posúdenia so znalosťou výsledku. Zisťovali odlišnosti v sklone, časovom posune a podobnosť celkového trendu pravdepodobnostných odhadov. Ich výsledky neukázali žiadne odlišnosti v sklone alebo celkovom trende. Ako významný sa ukázal ich časový posun. Participantí mali v spätnom pohľade pocit, že posudzovali vyššiu pravdepodobnosť výsledku skôr, ako v skutočnosti [3].

Princíp rozdelenia podnetového materiálu na viacero častí a opakované posudzovanie pravdepodobnosti výsledku adaptovali Konečný a Bačová [6] v experimente s hypotetickým dizajnom. Na 10 animáciách dopravných nehôd skúmali vzťah medzi momentmi pravdepodobnostných posúdení a výskytom antecedentov, ktorým mohla byť pripísaná úloha vo vysvetľovaní výsledku. Animácie rozdelili na dve časti. V prvej bola zobrazená situácia do momentu ako sa vyskytla okolnosť vedúca k dopravnej nehode, a druhá končila krátko pred výskytom nehody. Participantí robili pravdepodobnostné posúdenia výsledku po každom

segmente. Prvé posúdenie teda robili pred výskytom kauzálneho antecedentu a druhé po jeho zobrazení. Participanti v experimentálnej skupine videli najprv celú animáciu a následne plnili tú istú úlohu ako kontrolná skupina.

Efekt spätného pohľadu merali porovnávaním veľkostí odhadov pravdepodobnosti medzi experimentálnou a kontrolnou skupinou v tých istých momentoch animácie. Pozorovali rôzne vzory kombinácií efektu spätného pohľadu, obráteného efektu spätného pohľadu a žiadnych odlišností medzi skupinami. Efekt spätného pohľadu sa prejavil vždy len v druhom posúdení, teda po zobrazení kauzálneho antecedentu. Obrátený efekt spätného pohľadu naopak len v momentoch posúdenia pred antecedentom. Pri 6 animáciách sa obidva tieto efekty vyskytli v kombinácii, pri 3 iba obrátený ESP v prvom posúdení a v jednom videu sa prejavil iba ESP v druhom posúdení.

Konečným a Bačovou [6] pozorované kombinácie efektov boli prekvapivým výsledkom, ktorý by bol nemerateľný pri použití klasických experimentálnych prístupov. Odlišnosti pravdepodobnostných posúdení v iných momentoch ako práve pred výsledkom by neboli zachytiteľné. Tieto výsledky odhaľujú možný vzťah medzi chybami v spätnom pohľade a časom výskytu okolností, ktorým môže byť pri hľadaní zmyslu výsledku atribúovaná kauzálna úloha. Autori elaborujú toto zistenie v kontexte teórie kauzálneho modelu. Navrhujú, že vytvorenie kauzálneho spojenia medzi antecedentom a výsledkom udalosti nespôsobuje len vznik ESP v nasledujúcich momentoch, ako to ukazujú aj predchádzajúce výskumy, ale tiež môže spôsobiť jeho obrátenie v momentoch vývinu udalosti pred výskytom antecedentu.

Doteraz veľmi skromný počet publikácií venujúcich sa efektu spätného pohľadu už poukázal na aspekty, ktoré boli pre iné experimentálne prístupy prakticky neviditeľné. Zistenie, že ESP je charakteristický časovým posunutím pravdepodobnostných posúdení výsledkov, a že môže prichádzať ku kombinácii ESP s jeho obrátením majú nezanedbateľné teoretické implikácie. Ešte významnejšie sa však zdajú vyhladky, ktoré prinášajú pre ďalší výskum tohto fenoménu.

## 5 Čo ešte odhalí čas?

Sľubné začiatky výskumu časových charakteristík efektu spätného pohľadu prinášajú vlastné metodické problémy, otvárajú nové otázky, ale aj možné riešenia pre skúmanie viacerých teoretických koncepcií. Stručne popíšeme najaktuálnejšie výzvy pre tento nový prístup.

Opakovanie meraní pri tej istej udalosti môže už samotné ovplyvňovať posudzovanie pravdepodobnosti, alebo iných aspektov výsledkov udalosti.

V predchádzajúcich experimentoch prichádzalo vždy len k tvorbe jedného úsudku pri každej udalosti. Nie je zatiaľ možné vylúčiť, že formulácia jedného úsudku ovplyvní ten nasledujúci. Participanti pri usudzovaní môžu reflektovať svoje predchádzajúce odhady. Nasledujúce odhady môžu byť teda ovplyvnené napríklad procesom ukotvenia a prispôsobenia [17], prípadne aj inými procesmi, ktoré nevstupovali do usudzovania v klasických experimentálnych dizajnoch. Konečný a Bačová [6] tiež zvažujú možnosť vplyvu zámerného výberu momentov posúdenia, ktorý mohol vyvolať dojem kontrastu medzi pravdepodobnosťami výsledku v daných momentoch.

Myslím si, že existuje pomerne priamočiary spôsob, ktorým by sa mohol dať identifikovať vplyv opakovaného posudzovania. Viac svetla do tohto problému by prinieslo experimentálne porovnanie odhadov výsledku pri opakovanom posúdení so skupinou, v ktorej budú participanti robiť úsudky o výsledku pri jednej udalosti len v jednom momente udalosti.

Ďalším metodickým prvkom, ktorý sa objavil až pri porovnávaní posúdení vo viacerých momentoch je, že v spätnom pohľade nemajú participanti len znalosť o výsledku, ale aj o priebehu udalosti. Túto odlišnosť oproti klasickým dizajnom načrtáva už Roese a kol. [13] a na jej zváženie explicitne vyzýva Konečný a Bačová [6]. Yopchick a Kim [19] zistili, že poznanie výsledku bez dôvodu, prečo k nemu prišlo nestačí na vznik efektu spätného pohľadu. Otázne však zostáva, ako poznanie samotného výsledku ovplyvní tvorbu odhadov vo vývine udalosti, počas ktorého sú prezentované kauzálne relevantné antecedenty.

Predpokladám, že poznanie výsledku ovplyvní očakávanie a teda možno aj prepracovávanie kauzálneho modelu v priebehu rozvíjajúceho sa deja. Predstavme si napríklad osobu, ktorá ide prvý krát pozerat trilógiu Pána prsteňov a vie, že sa Frodovi podarí dostať sa do Mordoru. Bude asi považovať za pravdepodobné, že sa Frodovi podarí splniť svoje poslanie už keď Frodo príde k bráne do Mordoru, pretože cieľ bude na dosah. Práve tejto udalosti môže pripísať významnú kauzálnu rolu vo vysvetľovaní Frodovho úspechu. Keď ale bude pozerat ďalej, zistí, že sa Frodo otočil a zvolil inú cestu, počas ktorej mohla byť všetka jeho snaha zmarená. Nieкто, kto pozná príbeh celý by asi nikdy nepripísal veľkú kauzálnu rolu Frodovmu prvému priblíženiu k Mordoru, čo by mohlo spôsobiť odlišnosti v časovom vývine dojmu nevyhnutnosti Frodovho úspechu medzi týmito dvoma ľuďmi. Skúmanie vplyvu poznania výsledku v porovnaní s poznaním celého alebo čiastočného vývinu deja preto nemusí byť zaujímavé len z pohľadu metodického, ale rovnako aj pre porozumenie procesov prepracovávanie kauzálneho modelu pri hľadaní zmyslu.

Rôzne kombinácie efektov, ktoré experimentálne pozorovali Konečný a Bačová [6] (pozri časť 4) môžu tiež byť inšpiráciou pre ďalší výskum. Ako vysvetlenie ponúkli hypotézu, že môžu súvisieť s jednoduchosťou formulovania kontrafaktových úsudkov a odlišnými mierami prekvapivosti výsledkov animácií, ktoré v ich experimente neboli kontrolované. Ako spomínam v časti 3, prekvapivosť výsledku vplýva na proces hľadania zmyslu a jednoduchosť formulovania kontrafaktových úvah na dojem nevyhnutnosti výsledku. Takáto hypotéza je bez ďalšej empirickej podpory skôr špekuláciou, ale určite hodnou pozornosti v ďalšom výskume.

Do úvahy prichádza viacero otázok, ktorých zodpovedanie by mohlo pomôcť porozumieť odlišným vzorom efektov v spätnom pohľade. Z pohľadu možnej úlohy prekvapivosti výsledku sa ponúkajú podľa teórie kauzálneho modelu dve možnosti. V prípadoch, kedy nevznikol efekt spätného pohľadu sa dá predpokladať, že prekvapivosť bola buď príliš vysoká alebo príliš nízka. Teda ak je prekvapivosť nízka a nepríde k prepracovávaniu kauzálneho modelu, môže v skorších momentoch vzniknúť obrátenie efektu spätného pohľadu? Konečný a Bačová [6] naznačujú práve možnosť výskytu takéhoto javu. Navrhujú, že nezávisle od aktivácie procesu hľadania zmyslu, mohli participanti reflektovať nesplnenie podmienok pre očakávaný výsledok aj na základe svojho pôvodného kauzálneho modelu. Druhá možnosť je, že by bola prekvapivosť príliš vysoká a viedla by k neúspešnému vytvoreniu kauzálneho modelu danej udalosti. Môže sa aj v takomto prípade prejaviť obrátenie efektu spätného pohľadu v predchádzajúcich momentoch? Súčasná podoba teórie kauzálneho modelu neformuluje žiadne konkrétne predikcie v spomenutých otázkach, a aj preto ich považujem za potenciálne prínosné pre vysvetlenie procesov hľadania zmyslu.

Nakoniec poukážem na zdanlivú nezhodu o povahe trajektórií pravdepodobnostných posúdení vychádzajúcu z porovnania výsledkov doterajších experimentov skúmajúcich časové charakteristiky Fessela a kol. [3] a Konečného a Bačovej [6]. Fessel a kol. [3] na podobnom podnetovom materiáli ako použili Konečný a Bačová [6] namerali ako jediné skreslenie pravdepodobnostných odhadov v časovom posune, zatiaľ čo Konečný a Bačová [6] pozorovali aj obrátenie efektu spätného pohľadu. Konečný a Bačová [6] použili väčší počet podnetových materiálov a je možné, že keby vybrali len práve tú animáciu, kde sa obrátenie efektu spätného pohľadu neprejavilo, táto otázka by ani nenastala. Tiež je na mieste zobrať do úvahy, že významné odlišnosti v javoch v spätnom pohľade boli zistené aj medzi jednotlivými animáciami použitými v ich experimente a teda táto nezhoda sa nevzťahuje len na porovnanie výsledkov v spomínaných dvoch článkoch a

môže poukazovať len na rozdiely v použitom podnetovom materiáli. Tieto možné vysvetlenia by boli pripísateľné zväženiam o otvoreným otázkam opísaným v predchádzajúcich odstavcoch.

Napriek týmto možným pochybnostiam navrhujem alternatívne vysvetlenie. Myslím si, že mohli byť spôsobené tým, že Fessel a kol. [3] skúmali časové charakteristiky v experimente s pamäťovým dizajnom a Konečný a Bačová [6] s hypotetickým. Ani v jednom z týchto experimentov autori neseparovali komponenty efektu spätného pohľadu, ktorým sa venujeme v časti 2.1. Keď túto okolnosť vezmeme do úvahy, tak prideme k záveru, že Konečný a Bačová [6] merali len dojmy nevyhnutnosti a predvídateľnosti, kým javy, ktoré skúmali Fessel a kol. [3] zahŕňali aj skreslenie pamäte. Spomienka na svoje predchádzajúce odhady mohla v experimentoch Fessela a kol. [3] slúžiť ako kľúč k rekonštrukcii ich pôvodnej trajektórie. Možno už táto nezhoda napovedá o odlišnostiach v časových priebehoch vývinu jednotlivých komponentov. Návrh môjho experimentu by mal adresovať práve túto otázku.

## 6 Návrh experimentu

Ako som už avizoval v časti 5, cieľom tohto experimentu bude zistiť odlišnosti v časovom priebehu komponentov efektu spätného pohľadu. Vzhľadom na potrebu empirického overenia vplyvu inštrukcie na zanedbanie výsledku, bude adresovať aj túto otázku.

Odlišnosti výsledkov medzi časovými priebehmi posúdení pravdepodobnosti výsledku medzi Konečným a Bačovou [6] a Fesselom et. al. [3] naznačujú, že môže existovať rozdiel v časových charakteristikách efektu spätného pohľadu, ktorý je primárne výsledkom skreslenia pamäte a takým, ktorý je len výsledkom dojmov nevyhnutnosti a predvídateľnosti. Keďže je meranie dojmu nevyhnutnosti a predvídateľnosti v tej istej skupine participantov problematické (pozri časť 2.1), zameriame sa len na 1 komponent ESP, a to na dojem nevyhnutnosti. Voľba venovať sa dojmu nevyhnutnosti a nie dojmu predvídateľnosti je rozhodnutím do budúcnosti. Myslím si, že tento experimentálny prístup ešte prispeje k porozumeniu procesom prepracovávania kauzálneho modelu, ktorý súvisí práve s dojomom nevyhnutnosti. Práve preto výsledky tohto experimentu môžu priniesť aj podnety pre budúci výskum v tejto oblasti.

Na základe úvah opísaných v časti 2.1 by mal dojem nevyhnutnosti súvisieť s procesmi hľadania zmyslu. Pri formulácii hypotéz o časovom priebehu zmien dojmu nevyhnutnosti vychádzam z výsledkov Konečného a Bačovej [6]. Budem predpokladať, že k zvýšeniu dojmu nevyhnutnosti v spätnom pohľade príde len v momentoch, kedy je už z vývinu deja známy

aspoň jeden kauzálny antecedent vysvetľujúci výsledok. Zároveň budem predpokladať, že v predchádzajúcich momentoch k ich zvýšeniu nepríde, prípadne sa prejaví obrátenie efektu spätného pohľadu.

Ak je skreslenie pamäte nezávislé od kauzálne atribučných procesov, nemal by vzniknúť rozdiel medzi jeho veľkosťou pred a po spoznaní kauzálnych antecedentov. Výsledky Fessela a kol. [3] napovedajú, že skreslenie pamäte zachováva trajektóriu posúdení v naivnom pohľade.

Ako som už spomenul v časti 2.1, meranie dojmu nevyhnutnosti bez inštrukcie na zanedbanie vedomosti výsledku dáva priestor pre isté pochybnosti. Nestler a kol. [8, s.1409] predkladajú na overenie dve hypotézy. Ak komponenty efektu spätného pohľadu nie sú separovateľné, použitie rovnakej inštrukcie pri ich meraní by sa mala prejsť ich závislosť. Druhou je predpoklad, že ak nie sú komponenty separovateľné, pri použití rôznych inštrukcií by mali vzniknúť medzi nimi rozdiely. V tomto ohľade sa zdržím predikcií a budem formulovať otázky o existencii súvisu medzi dojmom nevyhnutnosti a skreslením pamäte pri použití inštrukcie na zanedbanie výsledku a o existencii rozdielov v komponente dojmu nevyhnutnosti v závislosti od použitia inštrukcie na zanedbanie výsledku.

Spomenuté hypotézy a výskumné otázky budem testovať na experimente s pamäťovým dizajnom na dvoch skupinách participantov. Jeden skupina participantov pri meraní dojmu nevyhnutnosti dostane inštrukciu na zanedbanie výsledku a druhá bude mať úlohu prehodnotiť svoj pôvodný úsudok. Závislými premennými budú skreslenie pamäte a dojem nevyhnutnosti. Skreslenie pamäte budem merať ako rozdiel vzdialeností pôvodného odhadu výsledku od skutočného pred a po získaní vedomosti o výsledku. Na meranie dojmu nevyhnutnosti použijem škálu navrhnutú Nestlerom a kol. [8], opísanú v časti 2.1, upravenú pre použitie s konkrétnym podnetovým materiálom a v rôznych časových bodoch. Okrem inštrukcie bude ďalšou nezávislou premennou moment posudzovania.

Participantí budú robiť posúdenia výsledku a jeho nevyhnutnosti v dvoch momentoch vývinu štylizovaného príbehu o Cézarovom obliehaní galského mesta Alesia. Výsledok udalosti, ktorý budú odhadovať, bude počet padlých vojakov na každej strane a dojem nevyhnutnosti budem merať škálou opísanou v časti 2.1.

V prvej časti príbehu budú uvedené historické okolnosti obliehania a počet vojakov v Alesii (90 tisíc) a v útočiacej rímskej armáde (50 tisíc). Po prečítanie tejto časti budú participantí robiť prvé posúdenia. Druhá časť príbehu bude opisovať stavbu obkľučujúceho ohradenia Alesie Rimanmi a útok na nedokončené bariéry 15 tisícmi galských jazdcov, ktorým sa podarilo preraziť

rímske línie a vydať sa zháňať pomoc z ďalších galských kmeňov. Táto časť bude obsahovať tiež zmienku o budovaní druhej úrovne opevnenia Rimanmi v obavách z príchodu posil zvonka. Záverom časti príbehu bude opísaný príchod 240 tisícovej galskej armády. Nasledovať bude druhé posudzovanie strát a dojmu nevyhnutnosti výsledku. Nakoniec participantí dostanú informáciu, že ani útoky oboch galských vojsk neprekonali rímske opevnenia a Rimanovia so stratami 28 tisíc mužov vyzabíjali až 120 tisíc galských vojakov. Bude tiež poskytnuté vysvetlenie, kde bude označené za chybu rozhodnutie Galov vyslať takú veľkú časť svojej jazdy zháňať posily, čím oslabili svoje schopnosti útoku a dali Rimanom čas vybudovať opevnenie.

Nasledovať bude vyplnenie osobnostného dotazníka nesúvisiaceho s experimentom, aby vznikol retenčný interval dostatočný pre vznik skreslenia pamäte. Následne bude participantom rovnako prezentovaný celý príbeh a budú mať za úlohu spomenúť si na svoje predchádzajúce odhady obetí. V jednej skupine si budú mať tiež spomenúť na svoje odpovede na škále nevyhnutnosti, v druhej budú inštruovaní prehodnotiť svoje úsudky. Predpokladám, že vyslanie jazdy Galmi a príchod početnej pomocnej armády bude v naivnom pohľade považované za okolnosť kauzálne prispievajúcu k víťazstvu Galov, kým v spätnom pohľade bude vnímaná ako chyba zvyšujúca nevyhnutnosť rímskeho víťazstva.

Očakávam, že výsledky tohto experimentu prispedia do viacerých oblastí skúmania efektu spätného pohľadu. Manipulácia inštrukcie o zanedbaní vedomosti výsledku napomôže v ďalšej debata o jej roli v spätnom pohľade. Porovnanie posúdení vo viacerých momentoch vývinu deja môže odhaliť odlišnosti v časových charakteristikách jednotlivých komponentov, ktoré neboli v predchádzajúcich výskumoch adresované. Budú mať teda implikácie pre teóriu o komponentoch efektu spätného pohľadu. Môžu tiež pomôcť objasniť nekonzistenciu vo výsledkoch experimentov zameraných na časové charakteristiky [3,6].

## 7 Záver

Efekt spätného pohľadu napriek svojej dlhej, viac než 35 ročnej histórii, zostáva pre výskum atraktívnou témou. K najväčším prínosom k rozvíjaniu teórie efektu spätného pohľadu považujem práve dekompozíciu klasicky unitárneho javu na 3 komponenty, či možno dokonca samostatné efekty vyskytujúce sa v spätnom pohľade. V nedávnej dobe sa objavil aj pár publikácií zameraných na skúmanie temporálnych charakteristík ESP, ktoré sľubujú čerstvý pohľad na otázky, ktoré boli skryté v čase. Vo svojom príspevku predstavujem spôsob ako

tieto dva smery skombinovať tak, aby predstavovali prínos pre obe oblasti skúmania.

Je zrejmé, že objasnenie otázok, ktoré sme načrtli a výziev, ktoré pred nás skúmanie časových charakteristík, ale aj komponentov efektu spätného pohľadu predkladá je úloha pre systematický program experimentálneho výskumu. Navrhovaný experiment je len prvým krokom na jeden z možných ciest v ďalšom výskume.

## Literatúra

- [1] Blank, H., & Nestler, S.: Cognitive process models of hindsight bias. *Social Cognition* 25 (2007) 132-146.
- [2] Blank, H., Nestler, S., von Collani, G., & Fischer, V.: How many hindsight biases are there? *Cognition* 106 (2008) 1408-1440.
- [3] Fessel, F., Epstude, K., & Roese, N. J.: Hindsight bias redefines: It's about time. *Organisational Behavior and Human Decision Processes* 110 (2009) 56-64.
- [4] Fischhoff, B.: Hindsight  $\neq$  foresight: the effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology* 1 (1975) 288-299.
- [5] Guilbault, R. L., Bryant, F. B., Brockway, J. H., & Posavac, E. J. (2004). A meta-analysis of research on hindsight bias. *Basic and Applied Social Psychology* 26 (2004) 103-117.
- [6] Konečný, M., & Bačová, V.: Hindsight bias and its reversal: What will time reveal?. *Studia Psychologica* 54 (2012) 251-259.
- [7] Mazursky, D., & Ofir, C.: "I knew it all along" under all conditions? Or possibly "I could not have expected it to happen" under some conditions? *Organisational Behavior and Human Decision Processes* 66 (1996) 237-240.
- [8] Nestler, S., Blank, H., & Egloff, B.: Hindsight  $\neq$  Hindsight: Experimentally induced dissociations between hindsight components. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 36 (2010) 1399-1413.
- [9] Nestler, S., & Egloff, B.: Increased or reversed? The effect of surprise on hindsight bias depends on the hindsight component. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 35 (2009) 1539-1544.
- [10] Pezzo, M. V.: Surprise, defence, or making sense: What removes hindsight bias? *Memory* 11 (2003) 421-441.
- [11] Pezzo, M. V., & Pezzo, S. P.: Making sense of failure: A motivated model of hindsight bias. *Social Cognition* 25 (2007) 147-164.
- [12] Pohl, R.: Ways to assess hindsight bias. *Social Cognition* 25 (2007) 14-31.
- [13] Roese, N., Fessel, F., Summerville, A., Kruger, J., & Dilich, M. A.: The propensity effect: When foresight trumps hindsight. *Psychological Science* 17 (2006) 305-310.
- [14] Roese, N. J., & Olson, J. M.: Counterfactuals, causal attributions, and the hindsight bias: A conceptual integration. *Journal of experimental social psychology* 32 (1996) 197-227.
- [15] Roese, N. J., & Vohs, K. D.: Hindsight bias. *Perspectives on Psychological Science* 7 (2012) 411-426.
- [16] Schwarz, S., & Stahlberg, D.: Strength of hindsight bias as a consequence of meta-cognitions. *Memory* 11 (2003) 395-410.
- [17] Tversky, A., Kahneman, D.: Judgment under Uncertainty: Heuristics and Biases. *Science* 185 (1974) 1124-1131
- [18] Werth, L., & Strack, F.: An inferential approach to the knew-it-all-along phenomenon. *Memory* 11 (2003) 411-419.
- [19] Yopchick, J. E., & Kim, N. S.: Hindsight bias and causal reasoning: a minimalist approach. *Cognitive processing* 13 (2012) 63-72.



# Podporou kognitívnej reflexie k tolerancii vyšších strát? Vplyv akontability na averziu voči strate je čiastočne mediovaný mierou analytického myslenia

Lenka Kostovičová<sup>1</sup>, Katarína Dudeková<sup>1</sup>, Miroslav Sirota<sup>2</sup>, Viera Bačová<sup>1</sup>

<sup>1</sup> Ústav experimentálnej psychológie SAV, Dúbravská cesta 9, 813 64 Bratislava

<sup>2</sup> Medical Decision Making and Informatics Research Group, School of Medicine, King`s College London, Capital House, 42 Weston Street, SE1 3QD, London, United Kingdom

{lenka.kostovicova, katarina.dudekova, viera.bacova}@savba.sk, miroslav.sirota@kcl.ac.uk

## Abstrakt

V príspevku vychádzame z predpokladu, že sociálne prostredie vplýva na rozhodovanie sa ľudí. Experimentálne sme overovali, že akontabilita - implicitné alebo explicitné očakávanie, že človek bude svoje presvedčenia, pocity či konanie prezentovať alebo obhajovať pred inými - vedie k zvýšenému kognitívnemu úsiliu, čo sa prejaví v redukcii averzie k strate v rámci rizikovej voľby. Výsledky potvrdili, že akontabilita u participantov významne zvýšila úroveň kognitívnej reflexie a znížila ich prejavy averzie k strate, pričom efekt akontability na averziu k strate bol čiastočne mediovaný mierou analytického myslenia. Tieto zistenia podporujú hypotézu, že predpokladanie sociálnych dôsledkov rozhodnutia aktivuje analytický systém myslenia a inhibuje intuitívne odpovede, čo následne vedie k normatívne adekvátnejšiemu rozhodovaniu.

## 1 Experimenty v „sociálnom vákuu“

Participant experimentov v oblasti kognitívnej psychológie nepocitujú zodpovednosť za výsledky svojho konania a rozhodovania. Nemajú dôvod obávať sa akýchkoľvek interpersonálnych dôsledkov svojich volieb. Tento spôsob „izolovania“ sociálnych vplyvov, postavený na argumente eliminovania skresľujúcich premenných, však môže mať ďalekosiahle implikácie v súvislosti s ekologickou validitou výskumu.

Väčšinu našich rozhodnutí totiž vykonávame neanonymne, za prítomnosti, spoluúčasti iných ľudí, či s priamym dopadom na okolie. Naše myšlienky sú utvárané sociálnou nevyhnutnosťou vzájomného prezentovania, vysvetľovania, zdôvodňovania a následného hodnotenia konania, ktoré býva často v nesúlade s klasickými normatívnymi princípmi [1].

Naším cieľom je preskúmať vplyv akontability na rozhodovanie v podmienkach rizika a tiež mechanizmus na pozadí tohto efektu.

### 1.1 Akontabilita v rozhodovaní

Akontabilitu môžeme definovať ako explicitné alebo implicitné očakávanie, že človek bude svoje presvedčenia, postoje, konanie vysvetľovať alebo zdôvodňovať iným. Príčinou vplyvu akontability na naše rozhodovanie môže byť motivácia hľadania a udržania rešpektu a uznania od okolia, ochrana sociálneho imidžu a identity seba i ostatných [2].

Efekt akontability je podmienený rôznymi faktormi, pričom najdôležitejšie sú charakteristiky publika a načasovania - v zmysle vedomosti o zodpovednosti za rozhodnutie pred alebo po jeho učinení [3]. Opakovane sa potvrdilo, že akontabilita o ktorej participantí vedeli pred rozhodnutím (*predecisional accountability*), pričom pocitovali zodpovednosť pred publikom, ktorého názory neboli vopred známe, viedla ku kognitívnej snahe a multidimenzionálnemu mysleniu [napr. 1, 2, 3, 4]. Zároveň existuje evidencia, že tento spôsob navodenia komplexného myslenia potláča tzv. sklony ku kognitívnym chybám (*cognitive biases*), rezistentným voči iným spôsobom manipulácie. Avšak samotná metafora človeka „politika“, spájajúca sa so skúmaním akontability [2] v princípe spochybňuje klasifikáciu odpovedí ako omylov či skreslení vzhľadom na pragmatické, spoločenské ciele a všeobecnú efektivitu konkrétnych stratégií v každodennom živote.

Akontabilita je bežne experimentálne manipulovaná inštrukciou, že voľby a rozhodnutia účastníkov výskumu budú následne podrobené zverejneniu, diskusií, hodnoteniu.

## 1.2 Averzia k strate a akontabilita

Aspekt rizika je prítomný takmer v každom z našich - viac či menej závažných - rozhodnutí, od situácie bežného nákupu až po voľbu v oblasti zdravia. Vo väčšine z nich sú v hre aj faktory sociálneho vplyvu.

Koncept averzie k riziku bol predstavený Kahnemanom a Tverskym [5] a predstavuje jeden z parametrov ich prospektivej teórie, ktorý vypovedá o tom, že straty istej hodnoty majú pre ľudí väčšiu závažnosť než výhry rovnakej výšky. V rizikových voľbách sa averzia k strate prejavuje ako nevoľa hrať hru, hoci očakávaná hodnota hry je pozitívna (t.j. možná výhra je vyššia než možná prehra) [6].

Averzia k strate je jedným z vysvetlení rôznych fenoménov v rámci usudzovania a rozhodovania, napr. sklonu k udržaniu status quo [7] alebo efektu dispozície [8]. Vzťah medzi sociálnymi vplyvmi a averziou k strate, ako i rozhodovaniu v podmienkach rizika vôbec, je málo preskúmaný.

Testovaním vplyvu akontability na averziu k strate sa zaoberal Vieider [9], používajúc klasické prospektové úlohy Tverskeho a Kahnemana [10], t.j. 8 párov hier založených na hľadaní ambivalencie medzi nimi (v zmysle otázky „Kedy sú pre Vás hry rovnako príťažlivé?“). Úlohou participantov bolo doplniť sumu, ktorá druhú hru urobí rovnako príťažlivou ako je pre nich hra prvá, konštantne daná. Participant v experimentálnej skupine boli inštruovaní uviesť identifikačné údaje (meno, e-mailová adresa) a očakávali po absolvovaní výskumu interview s experimentátorom. Autor zistil, že akontabilita spôsobila redukciu averzie k riziku a navrhol vysvetlenie cez optiku teórie duálnych procesov [napr. 11]. Interpretácia je založená na aktivácii racionálneho, analytického systému myslenia a potlačení rýchlych, intuitívnych odpovedí prostredníctvom prítomnosti sociálnych dôsledkov rozhodnutia.

Kritika tejto štúdie je adresovaná najmä voči experimentátorovi v role osoby, ktorej sa účastníci výskumu „zodpovedajú“ a je otázne ako vnímali jeho očakávania. Zároveň výskumný súbor predstavovali študenti ekonómie, pričom tí majú vyššiu šancu identifikovať averziu k strate ako sklon k chybe, ako upozorňuje samotný autor, ktorý predpoklad overil v rámci retrospektívnych rozhovorov.

Nesmieme tiež opomenúť otázku finančnej motivácie, pričom v opisovanej štúdiu, na ktorú náš výskum nadväzuje, použil autor paušálnu sumu bez ohľadu na

výkon vo výške 15 eur, avšak za účasť na viacerých experimentoch.

## 1.3 Kognitívna reflexia a duálne procesy

Teórie dvoch mechanizmov myslenia majú dlhú tradíciu. V skratke môžeme skonštatovať, že hovoria o súčasnom pôsobení rýchleho, nevedomého intuitívneho systému a pomalšieho, analytického, na pravidlách založeného systému (Systém 1 a Systém 2) [12]. Systém 1 funguje na báze heuristik, nadobudnutej skúsenosti a vedomosti, je relatívne nenáročný v rámci kognitívnej kapacity. Systém 2 je označovaný za racionálny, umožňuje abstraktné, hypotetické myslenie, vyžaduje však kognitívnu kapacitu, čas a úsilie [11]. Systém 1 produkuje dojmy, úmysly a emócie pre Systém 2, ktorý ich môže reflektovať a ďalej spracúvať. Veľa činností však prebieha bez zapojenia druhého systému - automaticky, rýchlo, intuitívne - čo má niekedy za následok odchýlky od normatívneho usudzovania, kognitívne ilúzie či omyly [13].

Rozhodli sme sa overiť Vieiderov [9] predpoklad aktivácie analytického systému prostredníctvom akontability a následnej redukcie averzie k strate, a za týmto účelom použijeme Test kognitívnej reflexie (*The Cognitive Reflection Test; CRT*) [14]. Logické hádanky, ktoré sú súčasťou testu, evokujú rýchle, intuitívne odpovede, ktoré sú však nesprávne a skóre v teste, respektíve miera kognitívnej reflexie, vypovedá o schopnosti tieto odpovede potlačiť a zapojiť analytický systém. Kognitívna reflexia sa opakovane preukázala ako spoľahlivý prediktor výkonu v rôznych rozhodovacích úlohách, napr. v rámci rizikovej voľby [15] alebo bayesovského usudzovania [16].

## 1.4 Výskumný cieľ a hypotézy

Naším hlavným cieľom bolo overiť vplyv akontability na averziu k strate a tiež mechanizmus aktivácie analytického myslenia na pozadí tohto efektu. Predpokladáme, že vplyv akontability na averziu k strate je mediovaný úrovňou kognitívnej reflexie.

Navyše sa chceme pokúsiť o redukciu metodologických nedostatkov štúdie, z ktorej vychádzame. Participantmi nášho výskumu nebudú študenti ekonómie či príbuzných odborov. Akontabilita bude manipulovaná vo vzťahu k spolužiakom, čo oveľa lepšie simuluje situáciu bežného rozhodovania. Zaujímá nás tiež rôzna úroveň akontability a potenciálne rozdiely v intenzite jej vplyvu. V neposlednom rade plánujeme testovať efekt finančných incentív v rámci dizajnu, ktorý môže tiež zohrávať

dôležitú úlohu v zmysle spolupôsobiacich motivačných činiteľov sociálnych dôsledkov a finančnej odmeny.

## 2 Metódy

### 2.1 Participanti

Na našom výskume sa zúčastnilo 156 študentov Fakulty sociálnych a ekonomických vied Univerzity Komenského. Súbor tvorilo 50 mužov a 106 žien vo veku 19 až 26 rokov ( $M = 21,8$ ;  $SD = 1,8$ ). V rámci jednoduchého medziskupinového experimentálneho dizajnu boli študenti náhodne rozdelení do štyroch skupín: 2 kontrolné skupiny bez akontability (s finančnými incentívami / bez finančných incentív) a 2 experimentálne skupiny (nízka a vysoká úroveň akontability).

### 2.2 Materiály a procedúra

Averziu k strate sme sa rozhodli zisťovať prostredníctvom Gächterovej úlohy jednoduchej lotérieovej voľby (*Simple lottery choice task*) [17], do slovenčiny ju adaptovali Bačová & Jusková [6]. Hra vyzerá nasledovne:

Tab. 1. Gächterova úloha lotérieovej voľby

	HRA	Budem hrať túto hru	Nebudem hrať túto hru
1	Ak padne líce, prehráte 2 €, ak padne rub, vyhráte 6 €.		
2	Ak padne líce, prehráte 3 €, ak padne rub, vyhráte 6 €.		
3	Ak padne líce, prehráte 4 €, ak padne rub, vyhráte 6 €.		
4	Ak padne líce, prehráte 5 €, ak padne rub, vyhráte 6 €.		
5	Ak padne líce, prehráte 6 €, ak padne rub, vyhráte 6 €.		
6	Ak padne líce, prehráte 7 €, ak padne rub, vyhráte 6 €.		

Mierou averzie k strate je počet odmietnutých hier, teda skóre 0 až 6.

Úroveň kognitívnej reflexie sme určovali výsledkom testu kognitívnej reflexie [14], do slovenčiny adaptovali Sirota & Juanchich [16]. Znenie jeho otázok je nasledovné, pričom v zátvorke uvádzame najskôr správnu a potom intuitívnu odpoveď:

- (1) Pálka a lopta stoja €1,10 dokopy. Pálka stojí o 1,00 € viac ako lopta. Koľko stojí lopta? \_\_\_ centov (5, 10)
- (2) Ak trvá 5 strojom 5 minút, aby vyrobili 5 vecí, ako dlho by trvalo 100 strojom, aby vyrobili 100 vecí? \_\_\_ minút (5, 100)
- (3) Na jazere je záhon plávajúceho lekna. Každý deň sa zdvojnásobí veľkosť záhonu. Ak trvá 48 dní, aby záhon pokryl celé jazero, ako dlho by trvalo záhonu, aby pokryl polovicu tohto jazera? \_\_\_ dní (47, 24)

Úroveň akontability bola manipulovaná inštrukciou v scenári:

„Predstavte si, že ste účastníkom súťaže pozostávajúcej z šiestich kôl. Prosím, dodržiavajte inštrukcie a vyriešte nasledovné úlohy. Na záver hry budú niekoľkí z Vás odmenení finančnou výhrou<sup>1</sup> ...“

- bez akontability: „... Vaše odpovede ostanú anonymné.“
- nízka úroveň akontability: „... avšak výherca svoje odpovede zverejní.“
- vysoká úroveň akontability: „... avšak výherca svoje odpovede zverejní a vysvetlí.“

Po úvodných informáciách o dobrovoľnej účasti a ďalších princípoch výskumu a inštrukciách študenti vyplnili úlohy, a po ukončení práce bol jeden z každých 20-tich náhodne vyžrebovaný a odmenený sumou vo výške 10 eur. Účastníkov sme tiež oboznámili o podstate experimentu, s možnosťou pýtať sa doplňujúce otázky.

## 3 Výsledky

### 3.1 Predpoklady

Nezistili sme významné rozdiely v miere averzie k riziku, ani v úrovni kognitívnej reflexie podľa socio-demografických charakteristík (vek, pohlavie, študijný odbor) naprieč kontrolnými a experimentálnymi skupinami.

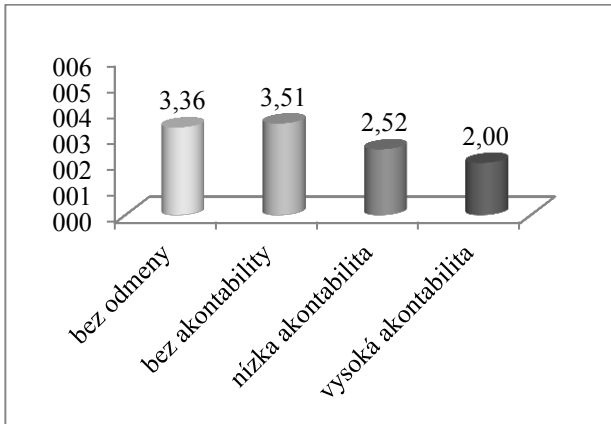
Bez ohľadu na skupinu vyšlo priemerne skóre v CRT teste 1,65 ( $SD = 0,96$ ) a priemerný počet odmietnutých hier v lotérieovej voľbe bol 2,83 ( $SD = 1,35$ ).

Shapiro-Wilkov test premenných averzia k strate, kognitívna reflexia a počet intuitívnych odpovedí preukázal normalitu rozloženia, preto ďalej pracujeme s hodnotami priemerov jednotlivých skupín.

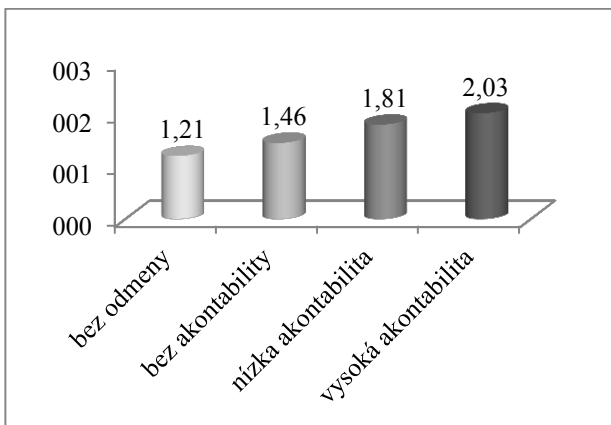
<sup>1</sup> V prípade skupiny bez finančnej motivácie táto veta absentovala.

### 3.2 Zistenia

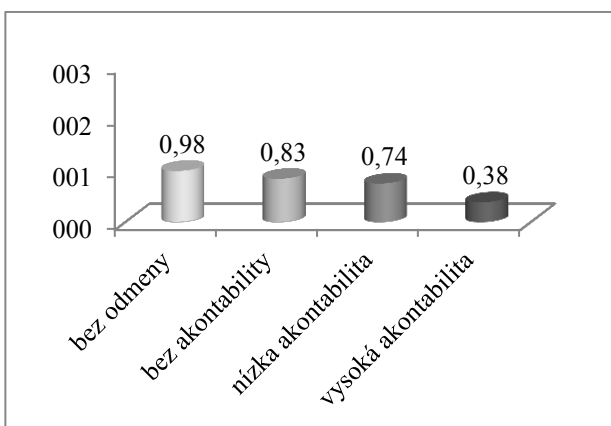
Priemerné skóre averzie k riziku jednotlivých skupín je zobrazené v Obrázku 1, skóre kognitívnej reflexie v Obrázku 2, skóre intuitívnych odpovedí v Obrázku 3.



Obr. 1. Averzia k strate podľa skupiny



Obr. 2. Kognitívna reflexia podľa skupiny



Obr. 3. Intuitívne odpovede podľa skupiny

Medzi kontrolnými skupinami sme nenašli významný rozdiel v averzii k strate:  $t(72) = -0,510$ ;  $p = 0,612$ , ani v kognitívnej reflexii:  $t(72) = -1,108$ ;  $p = 0,272$ , či v počte uvedených intuitívnych odpovedí:  $t(72) = 0,665$ ;  $p = 0,508$ . Vplyv akontability na averziu k strate a kognitívnu reflexiu by nemal byť skreslený finančnou motiváciou. Ďalej pracujeme s tromi skupinami.

Zamietame hypotézu o rovnosti priemerov v rámci úrovne averzie k strate:  $F(2; 120) = 17,453$ ;  $p < 0,001$ ;  $r_m = 0,47$ , miery kognitívnej reflexie:  $F(2; 120) = 4,004$ ;  $p = 0,021$ ;  $r_m = 0,25$  i počtu intuitívnych odpovedí:  $F(2; 120) = 4,042$ ;  $p = 0,020$ ;  $r_m = 0,25$ . Výsledky viacnásobnej komparácie post hoc uvádzame v Tabuľke 2, bez údajov na báze 95%-ného intervalu spoľahlivosti.

Tab.2. Výsledky viacnásobného porovnávania priemerov skupín z hľadiska averzie k strate, kognitívnej reflexie a intuitívnych odpovedí

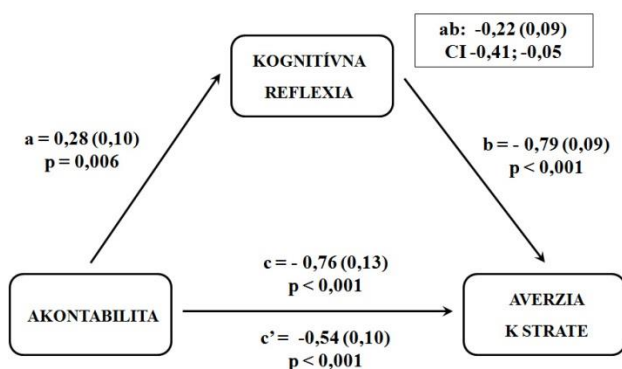
závislá premenná	skupiny	rozdiel priemerov	štđ. chyba	sig.
averzia k strate	1-2	,988*	0,257	< 0,001
	1-3	1,512*	0,260	< 0,001
	2-3	-0,524*	0,259	0,045
kognitívna reflexia	1-2	-0,346	0,198	0,083
	1-3	-0,562*	0,200	0,006
	2-3	0,215	0,199	0,282
intuitívne odpovede	1-2	0,091	0,182	0,617
	1-3	0,454*	0,184	0,015
	2-3	0,363*	0,183	0,049

Rozdiely medzi všetkými skupinami navzájom sú z hľadiska averzie k strate významné, avšak u kognitívnej reflexie je významný rozdiel len medzi skupinami bez akontability a s vysokou úrovňou akontability, u intuitívnych odpovedí u dvojíc založených na komparácii so skupinou vysokej úrovne akontability.

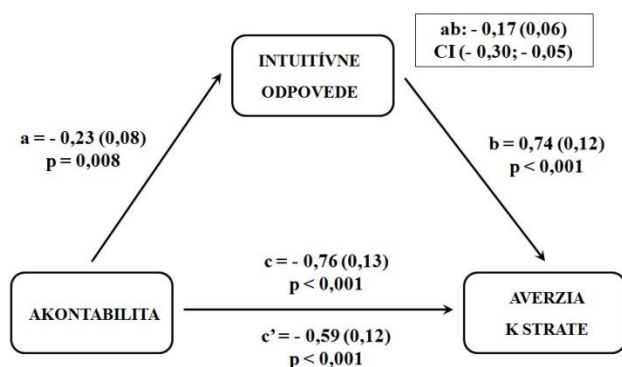
Na preverenie prítomnosti mediačného efektu kognitívnej reflexie sme použili postup Preachera a Hayesa [18, 19], konkrétne procedúru Mediate a sekvenčné kódovanie nezávislej kategorickej premennej. Všetky cesty vyšli významné, ako je možné vidieť na Obrázku 4. U akontability smerom ku kognitívnej reflexii sme zistili významnú pozitívnu hodnotu (cesta a) - vyššia miera akontability sa prejavila v zvýšenej úrovni kognitívnej reflexie. Ďalej bola identifikovaná významná negatívna cesta od kognitívnej reflexie k averzii k strate (cesta b) - participanti s vyššou kognitívnu reflexiou vykazovali nižšiu averziu k strate a opačne. Priamy efekt akontability na averziu k strate zostal významný i po kontrolovaní kognitívnej reflexie (cesta c'). Nepriamy

efekt bol negatívny a 95% interval spoľahlivosti (vytvorený bootstrapovaním s 5000 opakovaniami) (-0,22; *CI* [-0,41; -0,05]) neobsahoval nulovú hodnotu, a preto môžeme považovať efekt za štatisticky významný. Výsledky svedčia o čiastočnej mediácii efektu akontability na averziu k strate prostredníctvom úrovne kognitívnej reflexie.

Obdobné výsledky priniesla analýza mediácie efektu akontability na averziu k strate cez intuitívne odpovede, s reverznou štatisticky signifikantnou orientáciou ciest a, b: negatívna cesta od akontability smerom k intuitívnym odpovediam a pozitívna cesta od intuitívnych odpovedí smerom k averzii k strate (Obrázok 5). Priamy efekt ostal signifikantný a nepriamy efekt bol opäť negatívny, pričom 95% interval spoľahlivosti neobsahoval nulovú hodnotu (-0,17; *CI* [-0,30; -0,05]). Výsledky vypovedajú o čiastočnej mediácii efektu akontability na averziu k strate prostredníctvom počtu intuitívnych odpovedí.



Obr. 4. Mediačná analýza: kognitívna reflexia ako parciálny mediátor



Obr. 5. Mediačná analýza: počet intuitívnych odpovedí ako parciálny mediátor

## 4 Diskusia

Tetlock [2] nazval akontabilitu zanedbaným sociálnym kontextom volieb a rozhodnutí, explicitným alebo implicitným obmedzením činov ľudí. Avšak opomínaním akéhokoľvek kontextuálneho faktora tento pôsobí neprestáva, a predpoklad existencie pragmatických cieľov v procese usudzovania a rozhodovania je opodstatnený.

Vplyv akontability na rozhodovanie v podmienkach rizika, i špecificky na averziu k strate, bol potvrdený [9, 20], avšak mechanizmus na pozadí tohto efektu nebol odhalený. My sme sa rozhodli overiť hypotézu Vieidera [9] o redukcii averzie k strate prostredníctvom akontability, teda zdôraznenia sociálnych dôsledkov rozhodnutia, ktorá má za následok aktiváciu analytického systému a inhibovanie intuitívnych odpovedí.

Sumarizujúc naše zistenia konštatujeme, že akontabilita zvýšila kognitívnu reflexiu a znížila averziu k strate. Efekt akontability na redukcii averzie k strate bol čiastočne mediovaný mierou kognitívnej reflexie a počtom intuitívnych odpovedí. Na zodpovedanie príčiny vplyvu akontability na averziu k strate teda nie je postačujúce vysvetlenie založené na teórii duálnych procesov.

Rozdiely v postoji k riziku vplyvom zdôraznenia sociálnych dôsledkov majú dôležité implikácie pre ekologickú validitu laboratórnych experimentov (nielen) v oblasti usudzovania a rozhodovania, keďže bežnou praxou je eliminovať sociálne vplyvy, čo môže spôsobiť systematické skreslenia zistení. Tu vidíme prepojenie so stále aktuálnou tendenciou vykresľovať človeka ako používateľa intuitívnych heuristik, podliehajúceho kognitívnym omylom. Je však aplikácia obsahovo-neutrálnych pravidiel správania adekvátne? Čo keď sa pozrieme na pohnútky ľudí a ciele, ktoré správaním sledujú?

Ľudia sa adaptujú nielen na fyzické, ale i sociálne prostredie a slovami Herberta Simona, autora konceptu obmedzenej racionality, nie je možné pochopiť ľudské správanie ak ho skúmame mimo prostredia, v ktorom sa prirodzene odohráva [21].

## PodĎakovanie

Tento príspevok vznikol za podpory grantovej agentúry VEGA MŠ SR a SAV č. 2/0154/13 „Sociálne vplyvy v individuálnom rozhodovaní“ a Centra excelentnosti SAV „CESTA – Centrum strategických analýz“.

## Literatúra

- [1] J. S. Lerner, P. E. Tetlock: Accounting for the Effects of Accountability. *Psychological Bulletin* 125 (1999) 255-275
- [2] P. E. Tetlock: Accountability: The Neglected Social Context of Judgment and Choice. In *Research in organizational behavior* (B. Staw, L. Cummings, Eds.), JAI Press, Greenwich, CT, 1985: 297-332.
- [3] P. E. Tetlock: The Impact of Accountability on Judgment and Choice: Toward a Social Contingency Model. *Advances in Experimental Social Psychology* 25 (1992) 331-76
- [4] P. E. Tetlock, L. Skitka, R. Boettger: Social and cognitive strategies for coping with accountability: conformity, complexity, and bolstering. *Journal of personality and social psychology* 57 (1989) 632-40
- [5] D. Kahneman, A. Tversky: Prospect Theory: An Analysis of Decision under Risk. *Econometrica* 47 (1979) 263-292
- [6] V. Bačová, T. Jusková: Risky Choice And Affective Forecasting: Loss Aversion in Two Culturally Different Student Samples. *Studia Psychologica* 51 (2009) 329-342
- [7] A. Moshinsky, M. Bar-Hillel: Loss Aversion and Status Quo Label Bias. *Social Cognition* 28 (2010) 191-204
- [8] M. Weber, C. F. Camerer: The disposition effect in securities trading: an experimental analysis. *Journal of Economic Behavior & Organization* 33 (1998) 167-184
- [9] F. M. Vieider: The effect of accountability on loss aversion. *Acta Psychologica* 132 (2009) 96-101
- [10] A. Tversky, D. Kahneman: Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty* 5 (1992) 297-323
- [11] J. St. B. T. Evans: In two minds: dual-process accounts of reasoning. *Trends in Cognitive Sciences* 7 (2003) 454-459
- [12] K. E. Stanovich, R. F. West: Individual differences in reasoning: implications for the rationality debate? *The Behavioral and brain sciences* 23 (2000) 645-726
- [13] D. Kahneman: *Thinking, fast and slow*, Farrar, Straus and Giroux, New York, 2011.
- [14] S. Frederick: Cognitive Reflection and Decision Making. *Journal of Economic Perspectives* 19 (2005) 25-42
- [15] E. T. Cokely, C. M. Kelley: Cognitive abilities and superior decision making under risk : A protocol analysis and process model evaluation. *Judgment and Decision Making* 4 (2009) 20-33
- [16] M. Sirota, M. Juanchich: Role of numeracy and cognitive reflection in Bayesian reasoning with natural frequencies. *Studia Psychologica* 53 (2011) 151-161
- [17] S. Gächter, E. J. Johnson, A. Herrmann: *Individual-level loss aversion in riskless and risky choices*. Discussion Paper. Centre for Decision Research and Experimental Economics, University of Nottingham, 2007.
- [18] K. J. Preacher, A. F. Hayes: SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers* 36 (2004) 717-731
- [19] A. F. Hayes, K. J. Preacher: Statistical mediation analysis with a multicategorical independent variable. *Manuscript submitted for publication* (2012)
- [20] J. Pahlke, S. Strasser, F. M. Vieider: Risk-Taking for Others under Accountability. *Economics Letters* 114 (2012) 102-105
- [21] G. Gigerenzer: *Rationality for Mortals, How people cope with uncertainty*, Oxford University Press, New York, 2008.

## Odhad atribútov písma na základe požiadaviek človeka

Miron Kuzma

Ústav informatiky, Prírodovedecká fakulta  
Univerzita Pavla Jozefa Šafárika v Košiciach  
Jesenná 5, 040 01 Košice  
Email: miron.kuzma@student.upjs.sk

### Abstrakt

V každodennom živote, a to nielen v profesijnom ale aj v súkromnom, sa stretávame s veľkým počtom optimalizačných problémov. Pri hlbšom skúmaní však máme možnosť odkryť ich podstatu a začať uvažovať nad ich možným riešením. Medzi optimalizačné problémy patrí napríklad aj problém najatraktívnejšieho tvaru písma (fontu) pre pisateľa. Optimalizačnú funkciu tohto problému však nepoznáme. Na hľadanie riešenia problémov takéhoto druhu sa používajú okrem iných metód aj interaktívne evolučné výpočty. Pri nich však narazíme na problém únavy používateľa. Aby sa tento nedostatok eliminoval, je potrebné urýchliť konvergenciu týchto výpočtov.

V tomto článku predstavíme teoretický návrh algoritmu, ktorý na základe vstupných údajov od používateľa odhadne potenciálneho kandidáta na riešenie problému najatraktívnejšieho tvaru písma ako jednu z možností urýchlenia konvergenzie. Vychádzame z predpokladu, že používateľovi sa páči jeho vlastný štýl písma, aký používa. Tomuto štýlu písma by sa mal podobáť teda aj navrhnutý typografický font. V algoritme sa spája použitie Schwartzovho kultúrneho modelu a poznatkov z grafológie napaňovaním atribútov písma na tento kultúrny model.

### 1 Úvod

V praxi sa stretávame s množstvom optimalizačných problémov, pre ktoré nepoznáme optimalizačné funkcie. V takých prípadoch sú hľadané metódy, ktoré dokážu nahradiť explicitné vyjadrenie optimalizačnej funkcie inými prostriedkami. V niektorých prípadoch sú využiteľné vlastnosti človeka, ktorý vie posúdiť optimálnu hodnotu podľa svojich kritérií. Napríklad, interaktívne evolučné výpočty [3, 5, 6, 7, 14, 15] majú široké možnosti použitia v počítačovej grafike, nastavovaní sluchového aparátu, hudbe, priemyslových aplikáciách, spracovaní reči a obrazu, dataminingu, umení, v terapii, riadení a robotike, architektúre, dizajne, virtuálnej realite ale aj v mnohých ďalších oblastiach [23, 24, 25]. Hlavný problém pri použití interaktívnych evolučných výpočtov je však únava

používateľa. Existuje viac spôsobov urýchlenia konvergenzie [26], ktoré spracoval Pei [16]. My sme sa zamerali na teóriu hodnôt v spojení s fuzzy množinami.

Článok je pripravený v nasledujúcom zložení. V druhej kapitole popíšeme problém najatraktívnejšieho tvaru písma, ktorým sa zaoberáme [8, 9, 10]. Spracovaný prehľad obsahu a štruktúry hodnôt predstavíme v ďalšej kapitole. V štvrtej kapitole predstavíme spojenie teórie hodnôt a fuzzy množín. Na záver predstavíme možné vylepšenia algoritmu odhadujúceho vlastnosti písma.

### 2 Problém najatraktívnejšieho tvaru písma

Problémom najatraktívnejšieho tvaru písma sme sa zaoberali v prácach minulých [8, 9] a v súčasnom výskume, kde sa zaoberáme otázkou zrýchlenia konvergenzie interaktívnych evolučných výpočtov v snahe redukovat' únavu používateľa [10].

Majme jeden typ písma, ktorého tvar je popísaný pomocou konečnej množiny parametrov. Doména parametra môže byť konečná množina hodnôt alebo interval. Úlohou je nastaviť hodnoty parametrov. Aplikovaním hodnôt týchto parametrov na dané písma by mal používateľ získať čo najviac atraktívny tvar písma. Na popis tvaru písma použijeme jazyk Metafont [12]. Vzhľadom k tomu, tvary písma sú definované rovnicami, je možné meniť hodnotu jedného z týchto parametrov na jednom mieste a aplikovaním zmeny vytvoríť konzistentné zmeny v celom tvare písma Computer Modern.

Konfiguračný súbor písma obsahuje 62 parametrov. Experimentálne sme vybrali konečných 25 parametrov (resp. 21, ak budeme uvažovať rovnaké hodnoty pre veľké aj malé písmená). Tieto parametre, majú zásadný vplyv na konečný vzhľad písma.

Definičný súbor písma obsahuje vektor parametrov  $p$  rozmeru 62. Každá zložka vo vektore predstavuje jeden parameter. Každá zložka vektora  $p$  sa rovná nule, ak predstavuje jeden z 25 vybraných parametrov našej modifikácie. Modifikačný vektor  $\Delta p$  je vektor dimenzie 62 a jeho zložky sú parametre, ktoré používame pre modi-

**Tabuľka 1:** Zoznam parametrov a ich hodnôt. Parametre čas,  $P_1$  až  $P_{13}$  sú celočíselné,  $P_{14}$  až  $P_{17}$  sú reálne,  $P_{18}$  až  $P_{21}$  sú boolovské.

$P_i$	dolná hranica	horná hranica	$P_i$	dolná hranica	horná hranica
$P_1$	0	100	$P_{12}$	0	100
$P_2$	0	100	$P_{13}$	0	100
$P_3$	0	100	$P_{14}$	0	0,5
$P_4$	0	80	$P_{15}$	0	1,5
$P_5$	1	10	$P_{16}$	0	1,0
$P_6$	0	80	$P_{17}$	0	0,7
$P_7$	0	60	$P_{18}$	False	True
$P_8$	0	60	$P_{19}$	False	True
$P_9$	0	100	$P_{20}$	False	True
$P_{10}$	0	30	$P_{21}$	False	True
$P_{11}$	0	100	čas	0	indiv.

fikáciu. Tie zložky vektora  $\Delta\vec{p}$  sú rovné nule, ktoré sme nevybrali. Nový vektor parametrov  $\vec{y}$  získame súčtom vektorov  $\vec{p}$  a  $\Delta\vec{p}$ :

$$\vec{y} = \Delta\vec{p} + \vec{p} \quad (1)$$

Podľa definície optimalizačného problému [11], by mala existovať optimalizačná funkcia  $f$ . Táto funkcia by mala priradiť reálne číslo vektoru  $\vec{y}$ . Dopredu môžeme povedať, že „najatraktívnejší tvar“ môže znamenať iný tvar pre rôznych používateľov. Ako môžeme pozorovať, chýba explicitná definícia optimalizačnej funkcie  $f$ .

### 3 Teória obsahu a štruktúry hodnôt

#### 3.1 Spoločné vlastnosti hodnôt

Keď premýšľame o našich hodnotách, rozmyšľame, čo je pre nás v živote dôležité. Každý z nás má iné hodnoty (napr. výkon, bezpečnosť, benevolencia) s rôznym stupňom významu. Jedna hodnota môže byť veľmi dôležitá pre niekajú osobu, ale nedôležitá pre inú. Teória hodnôt [17, 19] preberá koncepciu hodnôt, ktorú špecifikuje šesť hlavných vlastností, ktoré sú implicitne zahrnuté v spisoch mnohých teoretikov:

(1) Hodnoty sú presvedčenia nerozlučne spojené a ovplyvňujú sa. Keď sa hodnoty aktivujú, spájajú sa s pocitom. Napr. u ľudí, pre ktorých je dôležitá nezávislosť ako hodnota a stane sa ohrozenou, pociťujú zúfalstvo, keď sú bezmocní si ju chránia a sú šťastní, keď ju môžu vychutnať.

(2) Hodnoty sa vzťahujú k žiaducim cieľom, ktoré motivujú akciu. Ľudia, pre ktorých spoločenský poriadok, spravodlivosť a ústretovosť sú dôležitými hodnotami, sú motivovaní k plneniu týchto cieľov.

(3) Hodnoty presahujú konkrétne akcie a situácie. Napríklad poslušnosť a poctivosť sú hodnoty, ktoré môžu byť relevantné v práci, škole, v športe, obchode a politike, v rodine, pri priateľoch a cudzincoch. Táto vlastnosť rozlišuje hodnoty od užších pojmov ako normy a postoje, ktoré zvyčajne odkazujú na konkrétne akcie, objekty alebo situácie.

(4) Hodnoty slúžia ako štandardy alebo kritériá. Hodnoty riadia výber či hodnotenie akcie, politiky, osôb a udalostí. Ľudia sa rozhodujú, čo je dobré alebo zlé, odôvodnené alebo nelegitímne, čo má zmysel a čomu má zmysel sa vyhnúť na základe možných následkov pre svoje "drahocenné" hodnoty. Avšak Vplyv hodnôt v každodennom rozhodovaní si len zriedka uvedomujú. Hodnoty vstupujú do povedomie, pokiaľ akcie alebo dôvody, ktoré niekto zvažuje, majú konfliktné dôsledky pre rôzne hodnoty, ktoré daná osoba má.

(5) Hodnoty sú zoradené podľa dôležitosti k sebe navzájom relatívne. Ľudské hodnoty tvoria usporiadaný systém priorít hodnôt, ktoré ľudí charakterizujú ako indivíduá. Majú prisúdiť význam úspechom alebo spravodlivosti, novosti alebo tradícii? Táto hierarchická vlastnosť taktiež rozlišuje hodnoty od noriem a postojov.

(6) Relatívna dôležitosť viacerých hodnôt vedie ku konkrétnym akciám. Každý postoj alebo správanie má zvyčajne vplyv na viac ako jednu hodnotu. Napríklad navštevovať kostol by mohlo vyjadriť tradíciu, prispôsobivosť a bezpečie na úkor hodnôt hedonizmu a stimulácie. Kompromis medzi príslušnými, súťažiacimi hodnotami je to, čo vedie k postojom a správaniu [17, 18]. Hodnoty prispievajú k akciám do tej miery, že sú relevantné v kontexte (teda pravdepodobné, že budú aktivované) a dôležité pre konkrétneho človeka.

Vyššie uvedené sú vlastnosti všetkých hodnôt. To, čo odlišuje jednu hodnotu od druhej je typ cieľa alebo motivácie, ktorú hodnota vyjadruje. Teória hodnôt definuje desať všeobecných hodnôt v závislosti od motivácie, ktorá je základom každej z nich. Predpokladajme, že tieto hodnoty zahŕňajú rozsah motivačne odlišných hodnôt uznávaných v rôznych kultúrach. Podľa teórie, tieto hodnoty môžu byť univerzálne, pretože sú založené na jednej alebo viacerých z troch univerzálnych požiadaviek ľudskej existencie. Tieto požiadavky sú: potreby jednotlivcov ako biologických organizmov, náležitosti koordinovanej sociálnej interakcie a prežitie a sociálne potreby skupín.

Jednotlivci nemôžu úspešne zvládať tieto požiadavky ľudskej existencie. Naopak, ľudia musia formulovať



vhodné ciele na ich zvládnutie, komunikovať o nich a spolupracovať pri ich dosiahnutí. Hodnoty sú sociálne žiaduce koncepty používané na prezentáciu týchto cieľov mentálne a slová používané na ich vyjadrenie v sociálnej interakcii. Z evolučného hľadiska tieto ciele a hodnoty, ktoré ich vyjadrujú majú zásadný význam pre prežitie.

### 3.2 Hodnoty

V tejto kapitole by sme v krátkosti predstavili každú z desiatich hodnôt z pohľadu širokých cieľov, ktoré vyjadrujú. Poznáme ich aj pod pojmom kultúrny model. Tu poznamenáme, že hodnoty sú založené na univerzálnych požiadavkách a odkazujú na príbuzné koncepty hodnôt. Aby sme popísali význam každej hodnoty viac konkrétne a jednoznačne, v zátvorkách je uvedený súbor položiek danej hodnoty, ktoré boli zahrnuté do prvého výskumu [20] na ich meranie. Niektoré dôležité položky hodnoty (Např. sebaúcta) môžu mať viac významov, vyjadrujú motivačné ciele viac než jednej hodnoty. Tieto položky sú uvedené v hranatých zátvorkách.

**Seba-smerovanie** Určujúci cieľ: nezávislé myslenie a konanie - výber, vytváranie, skúmanie. Seba-smerovanie pochádza z organistických potrieb kontroly a ovládania a interakčných požiadaviek samostatnosti a nezávislosti. (kreativita, sloboda, výber vlastných cieľov, zvedavosť, nezávislosť) [sebaúcta, inteligencia, súkromie].

**Stimulácia** Určujúci cieľ: vzrušenie, novosť a výzva v živote. Hodnoty stimulácie sú odvodené od organistických potrieb zmeny a stimulácia pre zachovanie optimálnej, pozitívnej, skôr než ohrozujúcej úrovne aktivácie. Táto potreba pravdepodobne sa vzťahuje k potrebám zakladajúcim hodnoty seba-smerovania. (pestrý život, vzrušujúci život, odvážny).

**Hedonizmus** Určujúci cieľ: radosť alebo zmyslové potešenie pre seba. Hodnoty hedonizmu pochádzajú z organistických potrieb a potešenia spojeného s ich napĺňaním. (potešenie, tešiť sa zo života, pôžitkárstvo).

**Úspech** Určujúci cieľ: osobný úspech prostredníctvom demonštrovania kompetencie v súlade so sociálnymi normami. Príslušný výkon, ktorý generuje prostriedky je nutný pre prežitie jednotlivca a pre skupiny a inštitúcie, aby dosiahli svoje ciele. Úspech ako hodnota sa uvádza v mnohých zdrojoch. My hodnoty úspechu chápeme ako zdôraznenie demonštrovania kompetencie, pokiaľ ide o prevažujúce kultúrne štandardy, a tým získanie sociálneho uznania. (ambiciózny, úspešný, schopný, vplyvný) [inteligentný, sebaúcta, uznanie spoločnosti].

**Moc** Určujúci cieľ: spoločenské postavenie a prestíž, kontrola a dominancia nad ľuďmi, zdrojmi a prostriedkami. Fungovanie spoločenských inštitúcií zrejme

vyžaduje určitý stupeň diferencovania postavenia. Dominancia a podriadenie sa ako rozmer objavuje vo väčšine empirických analýz medziľudských vzťahov vo vnútri i medzi kultúrami. Na odôvodnenie tejto skutočnosti spoločenského života a motivovania členov skupiny prijať túto hodnotu, musia skupiny považovať moc za hodnotu. Hodnoty moci môžu byť tiež premenené do individuálnych potrieb dominancie a ovládania. (autorita, bohatstvo, sociálna sila) [zachovanie môjho verejného obrazu, uznanie spoločnosti]. Obe hodnoty, moc a úspech, sa zameriavajú na sociálnu vážnosť, resp. úctu. Avšak hodnoty úspechu (např. ambiciózny) zdôrazňujú aktívnu demonštráciu úspešného výkonu v konkrétnej interakcii, zatiaľ čo hodnoty moci (např. autorita, bohatstvo) zdôrazňujú dosiahnutie alebo zachovanie dominantného postavenia v rámci všeobecnejšieho sociálneho systému.

**Bezpečie** Určujúci cieľ: bezpečnosť, harmónia a stabilita spoločnosti, vzťahov a seba samého. Hodnoty bezpečia sú odvodené od základných individuálnych a skupinových požiadaviek. Existujú dva podtypy hodnôt bezpečia. Prvý slúži predovšetkým individuálnym záujmom (např. čistota), druhý záujmom širšej skupiny (např. národná bezpečnosť). Dokonca aj druhý, však do značnej miery vyjadruje stotožnenie sa s jednotlivcom s cieľom širšej skupiny (alebo tými, s ktorými človek stotožňuje). Oba podtypy možno teda zjednotiť. (sociálny poriadok, bezpečie rodiny, národná bezpečnosť, čistota, opätovanie láskavosti) [zdravie, stredná cesta, pocit spolupatričnosti].

**Prispôsobivosť** Určujúci cieľ: zdržanie sa akcie, sklonov a popudov aby nedošlo k narušeniu alebo poškodeniu ostatných a porušeniu sociálneho očakávania alebo normy. Hodnoty prispôsobivosti pochádzajú z požiadavky, aby jednotlivci spomaľovali sklony, ktoré by mohli narušiť hladkosť interakcie a skupinového fungovania. Prakticky všetky analýzy hodnôt spomínajú prispôsobivosť. Hodnoty prispôsobivosti sú definované teda ako zdôraznenie sebaovládania v každodennej interakcii, zvyčajne s blízkymi ostatnými. (poslušný, disciplína, slušnosť, ctiť rodičov a starších) [lojálnosť, zodpovednosť].

**Tradícia** Určujúci cieľ: úcta, oddanosť a akceptácia zvykov a myšlienok, ktoré niekoho kultúra alebo náboženstvo poskytuje. Skupiny všade na svete vyvinuli postupy, symboly, myšlienky a názory, ktoré reprezentujú ich spoločné skúsenosti a osud. To je teda skupina zvykov a tradícií. Symbolizujú skupinovú solidaritu, vyjadrujú svoju jedinečnú hodnotu a prispievajú k jej prežitiu. Často majú formu náboženských obradov, presvedčenia a noriem správania. (rešpektovanie tradície, pokora, oddanosť, prijatie môjho miesta v živote) [miernosť, duchovný život]. Tradícia a prispôsobivosť sú hodnoty obzvlášť motivačne blízke, zdieľajú cieľ podriadenia seba samého v prospech sociálne daných očakávaní. Líšia sa

predovšetkým v objektoch, ktorým sa človek podriadi uje. Prispôsobivosť znamená podriadenosť osobám, s ktorými je človek často v interakcii - rodičia, učitelia, šéfovia. Tradícia znamená podriadenosť viac abstraktnejším objektom, náboženským a kultúrnym zvykom a predstavám ako dôsledkom, prispôsobivosť nabáda reagovať na aktuálne, prípadne zmeny očakávania. Tradícia vyžaduje schopnosť reagovať na nemenné očakávania (napr. z minulosti).

**Benevolencia** Určujúci cieľ: zachovať a zlepšiť životné podmienky tým, s ktorými je človek v častom osobnom kontakte. Hodnoty benevolencie pochádzajú zo základnej požiadavky pre hladké fungovanie skupiny a od organickej potreby príslušnosti. Najkritickejšie sú vzťahy v rodine a iných primárnych skupinách. Hodnoty benevolencie zdôrazňujú dobrovoľnú starosť o blaho druhých. (užitočnosť, úprimnosť, tolerantnosť, zodpovednosť, lojálnosť, skutočné priateľstvo, zrelá láska [pocit spolupatričnosti, zmysle života, duchovný život]). Hodnoty prispôsobivosti a benevolencie predstavujú kooperatívne a nápomocné sociálne vzťahy. Avšak, hodnoty benevolencie poskytujú vnútorný motivačný základ pre takéto správanie. Naproti tomu hodnoty prispôsobivosti podporujú spoluprácu, aby sa zabránilo negatívnym výsledkom pre seba. Obe hodnoty môžu motivovať rovnakú užitočnú akciu, samostatne alebo dohromady.

**Univerzalizmus** Určujúci cieľ: pochopenie, uznanie, tolerancia a ochrana blaha všetkých ľudí a prírody. To kontrastuje so skupinovým zameraním hodnôt benevolencie. Hodnoty univerzalizmu sú odvodené od potreby prežitia jednotlivcov a skupiny. Ale ľudia si neuvedomujú tieto potreby, kým sa nestretnú s ďalšími mimo rozšírenej primárnej skupiny a dokiaľ si nie sú vedomí nedostatku prírodných zdrojov. Ľudia si potom uvedomia, že ak sa nepodarí prijať ostatných, ktorí sú iní a zaobchádzať s nimi spravodlivo, tak to povedie k život ohrozujúcemu sporu. Môžu tiež spozorovať, že zlyhanie chrániť prírodné prostredia povedie k zničeniu zdrojov, na ktorých závisí život. Univerzalizmus spája dva podtypy záujmov, blaho tých vo väčšej spoločnosti a vo svete a blaho prírody. (veľkorysosť, sociálna spravodlivosť, rovnosť, svet v mieri, svet krásy, jednota s prírodou, múdrosť, ochrana životného prostredia) [vnútorná harmónia, duchovný život].

#### 4 Prechod od hodnôt k atribútom písma

Na to, aby bolo možné priradiť ľudským hodnotám charakteristické črty písma je potrebné v grafológii nájsť tu uvedené základné ľudské hodnoty. V prvom priblížení sa budeme zaoberať už spomenutými desiatimi ľudskými hodnotami. V zdrojovej literatúre grafológie [1, 2, 4, 13, 21, 22] je možné nájsť skúmané ľudské hodnoty s popisom približných atribútov písma. Toto písmo je na rozdiel od

nášho písané a je pre každého človeka vlastné. Ak však predpokladáme, že písmo, ktorým človek píše, sa danému človeku páči, potom je možné aplikovať poznatky z grafológie na náš problém najatraktívnejšieho tvaru písma. Ak by tieto úvahy platili, bolo by možné urýchliť konvergenciu evolučných výpočtov na ich začiatku - teda ponúknuť používateľovi vhodnú štartovaciu množinu.

Za základné atribúty písma budeme považovať: sklon (S), tlak (resp. hrúbku T), medzery medzi písmenami (M), proporčná veľkosť strednej časti písma (VS), proporčná veľkosť hornej časti písma (VH), proporčná veľkosť dolnej časti písma (VD), šírka písmen (SP), tvar a šírka slučiek (TS). Tieto atribúty dáme do vzťahu s kultúrnym modelom a jeho hodnotami: Seba-smerovanie (HSE), Stimulácia (HST), Hedonizmus (HED), Úspech (HUS), Moc (HMO), Bezpečie (HBE), Prispôsobivosť (HPR), Tradícia (HTR), Benevolencia (HBN), Univerzalizmus (HUN). Naše písmo má však viac atribútov podľa tab. 1 (alebo akékoľvek písmo jazyka Metafont) a bolo by teda ešte potrebné spraviť konverziu zo základných atribútov písma do atribútov konkrétneho písma, ktoré budeme modifikovať. V tabuľke 2 je znázornený výskyt jednotlivých atribútov písma pri konkrétnych ľudských hodnotách zo Schwartzovho kultúrneho modelu. Tento kultúrny model sa testoval dotazníkom *European Social Survey*, kde respondenti mali možnosť vyjadriť sa k jednotlivým otázkam priradením zvolenej hodnoty (-1 = je proti mojej hodnote, ..., 7 = úplne súhlasí)

Vychádzajúc z tabuľky 2 by algoritmus odhadu parametrov písma mohol fungovať nasledovne:

Predpokladajme, že funkcia príslušnosti  $\mu_p^i$   $i$ -teho atribútu písma (definičným oborom na  $\langle \min_i, \max_i \rangle$ ) bude daná predpisom lineárnej funkcie, takej, že  $f(\min_i) = 0$  a  $f(\max_i) = 1$ .

1. Každéj hodnote  $j$  zo spomenutých desiatich hodnôt Schwartzovho kultúrneho modelu priradí používateľ hodnotu  $w_j$  (-1 = je proti mojej hodnote, ..., 7 = úplne súhlasí).
2. Pre každý atribút  $i$  písma vyberieme tie hodnoty  $j$ , ktoré ho ovplyvňujú a pre každú kombináciu parameter-hodnota, zostavíme fuzzy funkciu príslušnosti  $\mu_h^k$  na základe mapovania z tabuľky 2. Pre atribút písma sklon, kde máme pojmy ľavý, pravý, rovno, vyberieme hodnoty z tabuľky HSE, HMO, HUN. Pre páry ľavý-HSE ( $\mu_h^1$ ), pravý-HMO ( $\mu_h^2$ ) a rovno-HUN ( $\mu_h^3$ ) navrhujeme funkcie príslušnosti  $\mu_h^k$  predpismi napr:

$$\mu_h^1 = \frac{x+1}{16} \quad (2)$$

$$\mu_h^2 = \frac{x+1}{16} + \frac{1}{2} \quad (3)$$

**Tabuľka 2:** Mapovanie ľudských hodnôt a parametrov písma na základe grafologických poznatkov a kultúrneho modelu. Vysvetlivky: Ľ - ľavý, P - pravý, R - rovno, M - malý, S - stredný, V - veľký; atribúty písma v prvom riadku: sklon (S), tlak (resp. hrúbku T), medzery medzi písmenami (M), proporčná veľkosť strednej časti písma (VS), proporčná veľkosť hornej časti písma (VH), proporčná veľkosť dolnej časti písma (VD), šírka písmen (SP), tvar a šírka slučiek (TS). Vysvetlivky a skratky kultúrneho modelu: Seba-smerovanie (HSE), Stimulácia (HST), Hedonizmus (HED), Úspech (HUS), Moc (HMO), Bezpečie (HBE), Prispôsobivosť (HPR), Tradícia (HTR), Benevolencia (HBN), Univerzalizmus (HUN).

	S	T	M	VS	VH	VD	SP	TS
HBE			V					
HBN			M				V	V
HED					M	V		M
HMO	P	V						
HPR				M				
HSE	Ľ		V	V			M	
HST			M		M	M		
HTR					M			
HUN	R							
HUS		M		S	V			

$$\mu_h^3 = \frac{1}{2} \quad (4)$$

3. Vyrátame agregáčnym operátorom (váženým priemerom) hodnotu  $\mu_p^i$  ako priemer hodnôt  $\mu_h^k$ , použijeme váhy  $w_j$ , ale vynormujeme ich tak, aby ich suma  $\sum_w = 1$ .
4. Nájďme vzor pre obrazy hodnoty funkcie príslušnosti  $\mu_p^i$  pre každý atribút písma  $i$ , čím dostaneme odhad hodnôt atribútov  $i$  písma, tieto hodnoty určia vzhľad typografického fonu resp. písma z východiskovej množiny písem pre používateľa.

Pri interaktívnych evolučných výpočtoch by bolo možné v jednotlivých iteráciách (generáciách evolučného algoritmu) zaviesť formu učenia pre jednotlivé funkcie príslušnosti pre atribúty písma (aby nemali iba lineárny predpis), čím by bolo možné presnejšie zachytiť individuálne črty používateľa vedúce k nájdeniu vhodných kandidátov na riešenie problému najatraktívnejšieho tvaru písma pre používateľa. Pre problém najatraktívnejšieho tvaru písma je potrebné ešte všeobecné atribúty písma na-

mapovať na konkrétne parametre fonu, ktorý sa bude upravovať.

## 5 Záver

Predstavili sme teoretický popis algoritmu odhadu parametrov písma pre problém najatraktívnejšieho tvaru písma. V tomto prvom priblížení sme sa zaoberali iba hlavnými desiatimi ľudskými hodnotami zo Schwartzovho kultúrneho modelu. Každú hodnotu je však možné rozšíriť na podhodnoty uvedené v kapitole, kde sme tento model rozoberali a popisovali jednotlivé hodnoty. Dostali by sme komplexnejší model používateľa, resp. jeho hodnotového rebríčka a vedeli pravdepodobne aj presnejšie odhadnúť atribúty písma, čím by sme opäť prispeli k rýchlejšej konvergencii evolučného výpočtu a tým pádom aj redukcii únavy používateľa.

## PodĎakovanie

Tento príspevok vznikol za podpory grantovej agentúry Ministerstva školstva, vedy, výskumu a športu Slovenskej republiky a Slovenskej akadémie vied - VEGA v rámci grantovej úlohy 1/0479/12 VEGA - Kombinatorické štruktúry a zložitosť algoritmov.

## Literatúra

- [1] Barrett, David V. aut. Grafológia. / 1. vyd. Bratislava: Ikar, vydavateľstvo a.s., 1997 Nestr.
- [2] Fischerová-Katzerová, Vladka, 1964 . Grafologie / 2., dopl. vyd. Praha: Grada, 2009. 227 s.
- [3] Gajdoš M., Reduction of Human Fatigue in IEC with Neural Networks for Graphic Banner Design, In: Master's Thesis, Technical University of Košice, Department of Cybernetics and Artificial Intelligence, 2006.
- [4] Jablonský, Eduard, 1946 aut. Príručka interpretácie grafologických znakov / 1. vyd. Žilina: Žilinská univerzita, 2005 268 s.
- [5] Jakša R., Takagi H., Nakano S., Image Filter Design with Interactive Evolutionary Computation, In: Proc. of the IEEE International Conference on Computational Cybernetics (ICCC2003), Siofok, Hungary, August 29-31, 2003.
- [6] R. Jakša, H. Tagaki: Tuning of Image Parameters by Interactive Evolutionary Computation, In: Proc.

- of 2003 IEEE International Conference on Systems, Man & Cybernetics (SMC2003), Washington D.C., (October 5-8, 2003) pp.492-497
- [7] Kováč J., Image Database Search Using Self-Organizing Maps and Multi-scale Representation, In: Master's Thesis, Košice, Technical University of Košice, Department of Cybernetics and Artificial Intelligence, 2007.
- [8] Kuzma M., Interactive Evolution of Fonts, master thesis, Technical University of Košice, 2008.
- [9] Kuzma M., Jakša R., and Sinčák P., Computational Intelligence in Font Design, Computational Intelligence and Informatics: Proceedings of the 9th International Symposium of Hungarian Researchers, Budapest, November 2008, pp.193-203.
- [10] Kuzma M., Andrejková G. , Interactive Evolutionary Computation in Optimization Problem Solving, Cognition and Artificial Life XII, Vol. XII, pp 120-125, Praha, 2012.
- [11] Kvasnička V., Pospíchal J. a Tiňo P.: Evolučné algoritmy, STU Bratislava, 6-7 (2000)
- [12] Metafont Tutorial,  
In: <http://metafont.tutorial.free.fr/downloads/mftut.pdf>,  
online zdroj. (May 8, 2008)
- [13] Mistrík, J. 1921-2000 aut. Grafológia: synkritická analýza v modernej grafológii / 1. vyd. Bratislava: Obzor, 1982. 199 s.
- [14] Neupauer M., Analysis of Medical Data using Interactive Evolutionary Computation, In: Master's Thesis, Košice, Technical University of Košice, Department of Cybernetics and Artificial Intelligence, 2006.
- [15] Pangrác L., Interactive Evolutionary Computation for Satellite Image Processing, In: Master's Thesis, Košice, Technical University of Košice, Department of Cybernetics and Artificial Intelligence, 2007.
- [16] Pei Y., Takagi H., A Survey on Accelerating Evolutionary Computation Approaches, 2nd International Conference of Soft Computing and Pattern Recognition (SoCPaR 2011), Dalian, China, pp.201-206 (Oct. 14-16, 2011).
- [17] Schwartz S. H. (1992). Universals in the content and structure of values: Theory and empirical tests in 20 countries. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 25)(pp. 1-65). New York: Academic Press.
- [18] Schwartz S.H. (1996). Value priorities and behavior: Applying a theory of integrated value systems. In C. Seligman, J.M. Olson, & M.P. Zanna (Eds.), *The psychology of values: The Ontario Symposium*, Vol. 8 (pp.1-24). Hillsdale, NJ: Erlbaum.
- [19] Schwartz S. H. (2005). Basic human values: Their content and structure across countries. In A. Tamayo & J. B. Porto (Eds.), *Valores e comportamento nas organizações [Values and behavior in organizations]* pp. 21-55. Petrópolis, Brazil: Vozes.
- [20] Schwartz S. H. (2006). Value orientations: Measurement, antecedents and consequences across nations. In Jowell, R., Roberts, C., Fitzgerald, R. & Eva, G. (Eds.) *Measuring attitudes cross-nationally - lessons from the European Social Survey*. London: Sage.
- [21] Stritz, F. Grafologické praktiká: Grafológia pre každého. / 1. vyd. Nitra: Garmond, Nitra, 2000 184 s.
- [22] Stritz, F., 1937-2000. Grafológia pre každého: vyše 200 ukážok rukopisov, podpisy slávnych osobností, 9 podrobných grafologických rozborov / 3. vyd. Nitra: Enigma, 2010. 217 s.
- [23] Takagi H. , Interactive Evolutionary Computation: System Optimization Based on Human Subjective Evaluation, IEEE International Conference on Intelligent Engineering Systems (INES'98), Vienna, Austria, pp.1-6 (1998)
- [24] Takagi H. , Interactive Evolutionary Computation: Fusion of the Capabilities of EC Optimization and Human Evaluation, Proceedings of the IEEE, Vol.89, No.9, pp.1275-1296 (2001)
- [25] Takagi H. : Interactive Evolutionary Computing: Fusion of the Capacities of EC Optimization and Human Evaluation, In: Proc. of 7th Workshop on Evaluation of Heart and Mind, KitaKyushu, Fukuoka, (2002) (in Japanese) pp.37-58
- [26] Wang S., Takagi H. , Improving the Performance of Predicting Users' Subjective Evaluation Characteristics to Reduce Their Fatigue in IEC, *Journal of Physiological Anthropology and Applied Human Science*, Vol. 24, No.1, pp. 81-85 (2005).

# Učenie reprezentácie objektov pomocou prekrývajúcich sa vývinových štruktúr

Andrej Lúčny, Michal Vician

Katedra aplikovanej informatiky, Fakulta matematiky, fyziky a informatiky, Univerzita Komenského  
Mlynská dolina, 842 48 Bratislava  
lucny@fmph.uniba.sk, Michal.Vician@st.fmph.uniba.sk

## Abstrakt

Prezentujeme systém, ktorý rozpoznáva nové objekty v scéne pozorovanej kamerou. Objekty sú reprezentované množinou šablón význačných orientácií (rozpoznávanie na základe tvaru). Pokiaľ disponujeme vhodnou množinou šablón, sme schopní reprezentovaný objekt rozpoznať. Problém je, že táto reprezentácia sa obvykle získava pomocou učiteľa. Náš systém je schopný postupne vytvárať túto reprezentáciu na základe sledovania pohybu objektov v scéne detektorom pohybu (pohybujeme s nimi pomocou nití). Takto pohyb objektov umožní systému porozumieť, že daný objekt tvorí jeden celok a získať schopnosť ho rozpoznať na základe jeho tvaru. Spolupráca medzi detekciou pohybu a rozpoznávaním tvaru je inšpirovaná Minského myšlienkou prekrývajúcich sa vývinových štruktúr (v rámci societného modelu mysle).

## 1 Úvod

Marvin Minsky vo svojej legendárnej knihe *The Society of Mind* [9] uviedol množstvo inšpiratívnych myšlienok, ktoré sa neskôr stali predmetom pokusov o ich implementáciu. Jednou z nich je idea vývinu pomocou prekrývajúcich sa úrovní (*development by overlapping stages*). Našou snahou bolo demonštrovať ju na konkrétnom príklade. Vybrali sme si na to úlohu z počítačového videnia, konkrétne rozpoznanie objektov v scéne. V posledných rokoch bolo vyvinutých viac metód na reprezentáciu objektov nepravidelného tvaru. Najjednoduchšou a zároveň aj najrýchlejšou metódou je metóda šablón význačných orientácií (*dominant orientation templates*) – DOT [3]. Všetky tieto metódy vyžadujú fázu učenia, počas ktorej je vytvorená reprezentácia rozpoznávaných objektov a prebieha za špeciálnych podmienok: v prípade DOT je objekt vystavený voči kontrastnému pozadiu, nakoľko je založená na detekciu hrán objektu, pričom je objekt reprezentovaný sadou šablón zodpovedajúcim rôznym pohľadom. Človek však takto objekty nerozpoznáva: reprezentáciu objektu si vytvorí za štandardných podmienok. Musí teda nejako pochopiť, že sa díva na

niečo, čo tvorí jeden kus. Zobrali sme preto jednu (a pre jednoduchosť aj jediná) z možných metód, ako také niečo môže urobiť a to detekciu pohybu rozdielovou snímku [2]. Kamerou pozorujeme scénu, v ktorej budú na nitiach zavesené neznáme objekty. Pokiaľ sa niektorý z nich pohne, nižšia úroveň systému rozpozna hranicu pohybujúceho sa objektu. Tá približne (teoreticky presne, ale prakticky len približne) zodpovedá jeho skutočnej hranici, z čoho môže byť vytvorená jedna šablóna, zodpovedajúca jednému pohľadu. Môže sa stať, že k reprezentácii objektu neprináša nič nové, lebo takú šablónu už v sade pre daný objekt máme. V opačnom prípade reprezentáciu objektu obohatíme o ďalšiu šablónu. Tak sa stane, že po čase reprezentácia objektu dozreje a bude dostatočná, aby sme ho dokázali rozpoznať pomocou metódy DOT, predstavujúcej vyššiu úroveň. Je pritom teda potrebné kombinovať obe úrovne riadenia: z úvodu je významnejšia nižšia úroveň založená na detekcii pohybu, neskôr však nastane stav, keď dohrala svoju úlohu a rozpoznávanie sa opiera výlučne o vyššiu úroveň založenú na rozpoznávaní tvaru.

V kapitole 2 popíšeme bližšie Minského ideu, v kapitole 3 metódu detekcie pohybu, v kapitole 4 metódu rozpoznávania tvaru. V kapitole 5 popíšeme spôsob integrácie, ktorý nie je technicky triviálny a preto ho zakladáme na architektúre Agent-Space [5], vytvorenej práve pre implementáciu častí Minského societného modelu, vychádzajúc z princípov multi-agentových systémov [4].

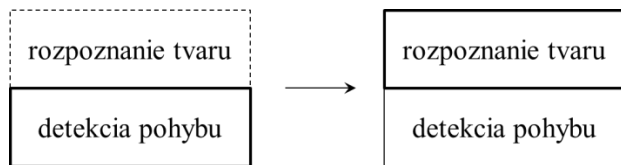
## 2 Vývin prekrývajúcich sa úrovní

Ide o predstavu, že systém nadobúda svoje schopnosti postupným zrením úrovni riadiacich štruktúr, pri ktorom vyššie úrovne disponujú lepšimi schopnosťami než nižšie, avšak nie je možné ich naštartovať z ničoho. Nezaobídu sa bez nižších úrovní, ktorých schopnosti sú slabšie a ich úloha v systéme spočíva v tom, aby priviedli vyššie úrovne k funkčnosti. Istý čas teda pracuje slabšie. Akonáhle však jeho určitá úroveň dospeje k istej saturácii schopností, ktoré si postupne buduje na základe dát vznikajúcich v nižších úrovniach, prevezme riadenie. Na systéme vtedy možno pozorovať kvalitatívny skok v jeho

schopnostiach. Systém je teda vopred štruktúrne usporiadaný tak, aby vývin jeho schopností postupne prechádzal rôznymi kvalitatívnymi fázami. Táto idea vychádza z vývinovej psychológie, hlavne z výskumu J. Piageta a snaží sa vysvetliť, prečo u detí dochádza k dramatickým a s ťabou načasovaným, zlomom v ich schopnostiach.

V našom prípade budeme mať len dve vrstvy. Nižšiu, ktorá dokáže fungovať z ničoho, ale vidí v scéne len pohybujúce sa predmety a vyššiu, ktorá dokáže fungovať až po vytvorení dostatočnej reprezentácie objektov v scéne, ale vidí v scéne i nepohybujúce sa predmety, pokiaľ sú celé viditeľné. Ďalšie vrstvy by mohli viesť vyriešiť zmenu škály (zväčšenie, zmenšenie), zákryt a podobne, toto však presahuje náš aktuálny zámer.

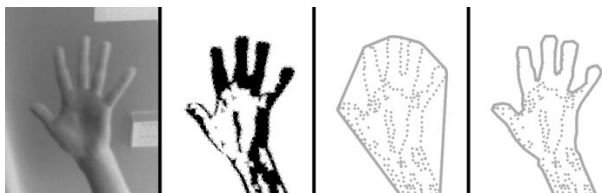
Nižšia vrstva funguje stále, ale dôležitá je len do okamihu, kedy si vyššia vrstva s jej pomocou vybuduje dostatočnú reprezentáciu rozpoznávaného objektu. Vtedy vyššia vrstva prevezme riadenie a schopnosti systému vykážu prudkú kvalitatívnu zmenu: predtým systém vidí len pohybujúce sa objekty, potom začne odrazu vidieť aj stojace (obr. 1).



Obr. 1. Vývin rozpoznávania na báze prekrývajúcich sa úrovni riadenia.

### 3 Detekcia pohybu

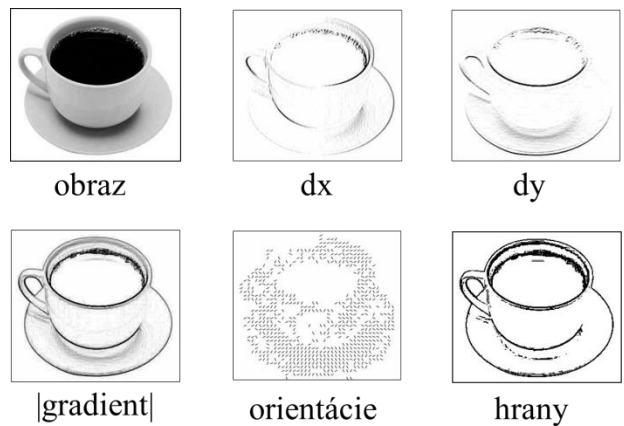
Nižšia úroveň systému rozpoznáva objekty na základe ich pohybu. Keďže kamera je v našom prípade statická, veľmi ľahko sa urobí rozdielový snímok. Ten poskytne množinu bodov obrazu, na ktorých dochádza ku zmene. Ako vidno na obr. 2, táto množina nedáva priamo hranicu objektu. Tú musíme spočítať ako najpravdepodobnejší nekonvexný ale celistvý tvar  $\chi$ -algoritmom [10] (obr. 2).



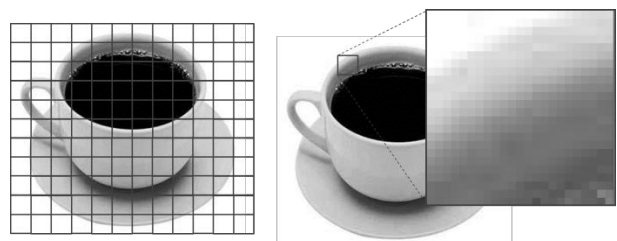
Obr. 2. Detekcia pohybu. Zľava doprava: obraz, rozdielový snímok, vstup do  $\chi$ -algoritmu, výstup z  $\chi$ -algoritmu

### 4 Rozpoznávanie tvaru

Rozpoznávanie tvaru predstavuje v našom systéme vyššiu úroveň. Realizovali sme ju pomocou metódy DOT – metódy šablón význačných orientácií [3]. Táto metóda používa na reprezentáciu objektu sadu šablón, predstavujúcich jeho tvar z rôznych pohľadov. Každá šablóna vychádza z výsledkov hranového detektora (napr. Canny [2]), ktorý intenzity šedej spracuje Sobelovým operátorom [2] (obr. 3). Ten dáva odhad derivácie obrazu, t.j. vektor derivácií vo vodorovnom a zvislom smere ( $dx, dy$ ) zvaný gradient. Keď zanedbáme smer tohto vektora, teda berieme tú z možností  $(-dx, -dy)$  a  $(dx, dy)$ , ktorá má kladnú zložku  $y$  a určíme jej argument, dostávame v každom bode orientáciu – uhol od 0 po  $\pi$ . Budú nás pritom zaujímať len tie body obrazu, ktoré zodpovedajú hranám (získaných prahovaním veľkosti gradientu a následným stenčením). Ďalej sa obmedzíme na niekoľko málo (napr. 8) kvalitatívnych hodnôt orientácie, napr.  $0 \approx (0, \pi/8)$ ,  $1 \approx (\pi/8, \pi/4)$ , ...

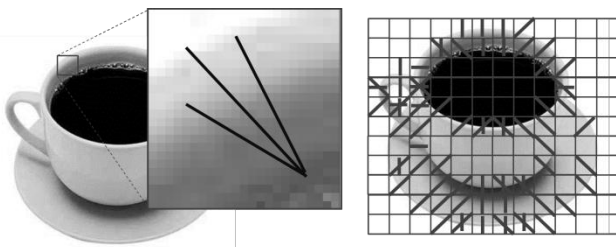


Obr. 3. Získanie hrán a orientácií gradientu obrazu.

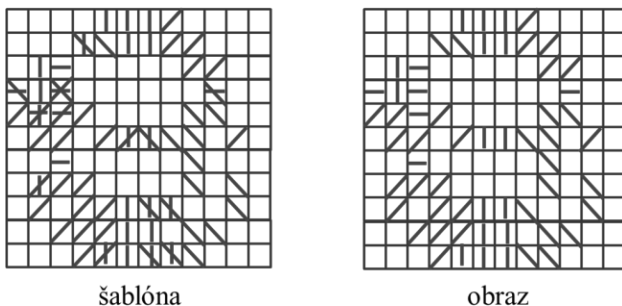


Obr. 4. Pokrytie obrazu regiónmi

Samotnú šablónu získame tak, že pokryjeme obraz neprekrývajúcimi sa regiónmi (obr. 4. vľavo). Ku každému regiónu potom vezmeme v úvahu orientácie bodov patriacich k hranám (obr. 4. vpravo). Z týchto orientácií určíme sadu prevládajúcich (obr. 5).

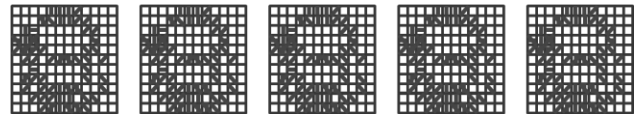


Obr. 5. Určenie význačných orientácií



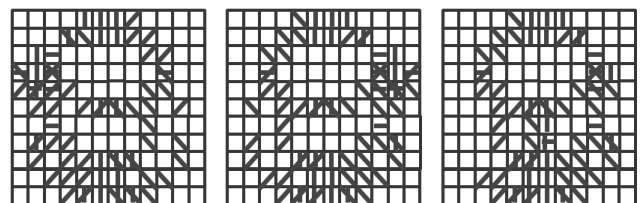
Obr. 6. Šablóna (vľavo) a prirovnávaný obraz (vpravo)

Tak dostaneme jednu šablónu. Tú potom môžeme prirovnať k obrazu tak, že aktuálny obraz spracujeme analogickým spôsobom ako vzorový obraz pri tvorbe šablóny, ale v tomto prípade vyberáme len jedinú význačnú orientáciu z regiónu (obr. 6). Samotné porovnanie obrazu so šablónou potom prebieha testovaním, či pre drvivú väčšinu regiónov nájdeme v obraze jednu z orientácií šablóny. Šablónu musíme samozrejme prikladať na rôzne miesta obrazu. Bolo by však príliš náročné prikladať ju na každý pixel. Je výhodnejšie držať sa jedného polohovania regiónov na obraz. V takom prípade ale zhoda nemusí nastať vôbec, pokiaľ sú regióny šablóny voči regiónom na obraze dostatočne posunuté. Môžeme si však narobiť z jedného vzoru viac šablón tak, že po ňom posúvame mriežku určujúcu regióny (obr. 7). V praxi nám to zďaleka nezvýši počet šablón o počet posunutí mriežky, lebo šablóny vychádzajú často rovnaké. Pri miernej strate presnosti (DOT je nepresná tak či tak) je možné dokonca zosumarizovať rôznym položením mriežky vygenerované šablóny do jedinej. Urobíme to tak, že v regióne šablóny budú všetky orientácie z okolitých poposúvaných regiónov. Takú šablónu nie je potrebné posúvať po pixloch, hoci už tak presne nevystihuje reprezentovaný objekt.



Obr. 7. Šablóny z rôznych posunutí regiónov

Z jedného pohľadu na objekt teda dostaneme jednu šablónu. Z rôznych pohľadov dostávame celú sadu šablón. Zdanlivo je ich opäť príliš moc. Keďže však nevyžadujeme 100% zhodu šablóny s obrazom, môžeme si dovoliť jednou šablónou nahradiť všetky jej podobné, v rámci podobnosti, ktorú pripúšťame pri porovnaní. Na väčšinu objektov nám stačí menej šablón, než by sme očakávali (obr. 8).



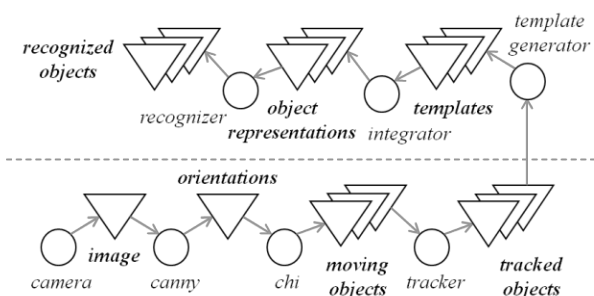
Obr. 8. Reprezentácia objektu

Porovnanie šablón sa dá implementovať veľmi efektívne pomocou bitových operácií [3]. DOT je preto veľmi rýchla metóda, použiteľná v reálnom čase. S kvalitnou kamerou dáva dobré experimentálne výsledky, ktoré sa vyrovnávajú i oveľa pomalším metódam, ako je napr. HOG [10].

## 5 Integrácia úrovni cez Agent-Space

Detekcia pohybu produkuje z každého snímku kamery sadu objektov, ktoré sa pohli. Ich pozícia sa porovnáva voči predchádzajúcej snímke, čo zodpovedá sledovaniu pohybu objektov na obraze. Vďaka tomu vieme, ktoré generované šablóny k sebe patria a tvoria reprezentáciu jedného objektu. Ich akumuláciou (samozrejme úspornou v zmysle metódy DOT) vzniká postupne reprezentácia objektu. Výrazná podobnosť medzi dvomi vzniknutými reprezentáciami vedie k zlúčeniu dvoch reprezentácií, keďže zrejme ide o rovnaký objekt. Keď zistíme, že detekcia pohybu síce generuje dost' ďalších šablón, ale tieto už neobohacujú reprezentáciu, spustíme detekciu objektu na základe tvaru.

V systéme prebieha súčasne viacero procesov, ktoré majú navyše rôznu rýchlosť. Skoordinovať ich preto nie je priamočiare. My sme na tento účel použili vlastnú architektúru Agent-Space, [www.agentspace.org](http://www.agentspace.org) [3], pomocou ktorej sme implementovali jednotlivé procesy v podobe reaktívnych agentov komunikujúcich nepriamo, cez tzv. *space* (obr. 9).



Obr. 9. Schéma systému v Agent-Space

## 6 Záver

Snažili sme sa ukázať jednoduchý ale reálny príklad na Minského ideu vývinu pomocou prekrývajúcich sa úrovní riadenia, ktorá sa odráža v zlomových zmenách v schopnostiach vyvíjajúceho sa systému. Systém, ktorý sme ako príklad predložili, kombinuje dve metódy rozpoznávania: na základe pohybu a na základe tvaru. Kým prvá metóda je schopná pracovať sama o sebe, druhá musí mať najprv vytvorenú dostatočne bohatú reprezentáciu. To sa udeje na základe výstupu z prvej metódy. Systém tak najprv vidí len pohybujúce sa objekty a potom zlomovo získava schopnosť vidieť aj objekty nepohybujúce sa.

Na implementáciu sme použili C++, pthread a OpenCV. Podstatnú časť implementácie systému vytvoril jeden z autorov v rámci svojej diplomovej práce [11]. K práci je možné pozrieť si aj demonštračné video <http://youtube/xpdvMPGmTnc>

Ďalšie možnosti rozvoja tohto projektu by spočívali v integrácii ďalších metód rozpoznávania objektov za účelom získania schopností, ktoré tento systém momentálne nemá. V prvom rade tento systém nepodporuje žiadny z invariantov okrem posunutia. Otočenie zvláda rozšírením sady šablón. Absolútne nezvláda zväčšenie a zmenšenie, čiže pri pohybe objektov nie je moc vhodné pohybovať s nimi pred kamerou dopredu a dozadu, radšej doľava a doprava a točiť s nimi. Nevie navyše riešiť zákryt objektov. Od rozpoznávania objektov, ktoré dokáže vykonať človek, má tento systém každopádne veľmi ďaleko.

## PodĎakovanie

Tento príspevok vznikol za podpory grantovej agentúry ASFEU v rámci grantovej úlohy KC-INTELSYS, ITMS 26240220072.

## Literatúra

- [1] R. A. Brooks: *Cambrian Intelligence*, The MIT Press, Cambridge, Mass., 1999.
- [2] Davies, E. R.: *Machine Vision: Theory, Algorithms, Practicalities*. Elsevier, 2005
- [3] Hinterstoisser, S. - Lepetit, V. - Ilic, S. - Fua, P. - Navab, N.: *Dominant Orientation Templates for Real-Time Detection of Texture-Less Objects*. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), San Francisco, California (USA), June 2010
- [4] Kelemen, J.: *The Agent Paradigm*. Computing and Informatics, Vol.22. (2003), pp. 513-519
- [5] Lúčny, A.: *Building Complex Systems with Agent-Space Architecture*. Computing and Informatics, Vol. 23 (2004), pp. 1001-1036
- [6] Lúčny, A.: *Od medzimodulových spojení k nepriamej komunikácii medzi agentami*. Znalosti, VŠE Ostrava, 2007.
- [7] Lúčny, A.: *Advantages of Multi-agent Approach to Building of Monitoring Systems*. In: Informatics 2007 (Ivan Plander ed.), SSAKI, Bratislava, 2007
- [8] Lúčny, A.: *Multiagentový prístup k modelovaniu mysle - alebo ako sledovať pingpongovú loptičku*. In: Umelá inteligencia a kognitívna veda III. (Kvasnička V. ed.), STU, Bratislava, 2011
- [9] Minsky, M.: *Society of Mind*. Simon & Schuster, New York, 1986
- [10] Sonka, M. – Hlaváč, V. – Boyle, R.: *Image processing, analysis and machine vision*. 3rd edition, Thomson Learning, Toronto, 2007
- [11] Vician, M.: *Formovanie riadiacich štruktúr prekrývaním úrovní vývinu*. Diplomová práca, FMFI UK Bratislava, 2011



# Výpočtový model zavádzania ako spoločnej akcie

Ludovít Malinovský

Centrum pre kognitívnu vedu, KAI FMFI, Univerzita Komenského  
Mlynská dolina, 84248 Bratislava  
Email: ludovit.malinovsky@gmail.com

## Abstrakt

Problém zavádzania v evolúcii jazyka znamená, že nevieme vysvetliť dôveryhodnosť lacných správ typických pre ľudský jazyk. Potenciálnym riešením je potreba kolektívneho zavádzania. Načrtne spôsob, ako formálne vnímať existujúce modely v evolúcii jazyka a navrhne spôsob, ako ich rozšíriť o tretieho hráča, čím vytvoríme model spoločného zavádzania. Navrhne model agenta schopného hrať takúto jazykovú hru a popíšeme správanie multiagentového systému s tromi a sto agentami. V modeli pozorujeme niekoľko fáz učenia, emergenciu roly aktívneho a pasívneho hráča a vznik niekoľkých dočasných i trvalejších skupiniek vo veľkej populácii.

## 1 Motivácia

Z darwinistického pohľadu existuje v otázke evolúcie ľudského jazyka problém zavádzania [22, 12]. Ak pre hovorcu existuje motivácia zneužiť komunikáciu pre vlastný prospech na úkor príjemcu, ten stráca motiváciu analyzovať potenciálne komunikačné snahy hovorcu a koná na vlastnú päsť. Komunikačný kód zabezpečujúci spoluprácu tak sotva vznikne, komunikácia sa mení na manipuláciu [13]. Aby hovorca zabezpečil dôveryhodnosť správy pre príjemcu, musí pre to vynaložiť energiu navyše, ako akúsi zálohu.

Preto sa vo výpočtových modeloch [14] a experimentoch emergencie [16] jazyka používajú komunikačné hry založené na čistej kooperatívnej hre. Ide o modely spoločnej akcie s komunikáciou, t.j. určitý aspekt akcie jedného hráča (hovorcu) si druhý hráč (príjemca) môže vyložiť ako zmysluplný signál a zareagovať naň. Vznikne úspešná spoločná akcia – koordinácia, napríklad ako pri pílení dreva: jeden píľku tlačí a druhý ťahá. V týchto modeloch sa často [9, 1] abstrahuje od konkrétneho významu akcie v prostredí a za koordináciu sa považuje, ak význam zakódovaný hovorcom je ten istý, ako význam rozpoznávaný príjemcom.

V reálnom živote sa však často stretávame so zavádzaním, či priamo klamaním, a jazyk napriek tomu nevymizol a nemá ani podobu nákladných, nápadných a riskantných zvieracích signálov (ktorým by zodpovedali ľudské rituály

[8]). Práve naopak, vzhľadom na množstvo prenesenej informácie je veľmi časovo i energeticky úsporný a hlasitosť neovplyvňuje dôveryhodnosť. Tieto modely teda evidentne nestačia na vysvetlenie vzniku a dynamiky ľudského jazyka. Neoverenou otázkou je, v akej forme by malo byť zavádzanie pre hovorcu zaujímavé, aby mala vzniknutá komunikácia vlastnosti ľudského jazyka.

Riešením by mohol byť taký model, kde by komunikácia koalície hovorcu a príjemcu nebola manipuláciou ohrozená priamo, ale skrz tretieho hráča s antagonistickými motívmi – ich spoločného protivníka. Hovorca by sa musel snažiť vyslať správu tak, aby ju pochopil príjemca, ale nepochopil antagonista [10, 11]. Príjemca navyše musí vedieť rozoznať stupeň a spôsob zavádzania, keďže schopnosti tretej strany môžu byť rôzne. Spoločnou akciou komunikujúceho tímu je teda kolektívne zavádzanie tretej strany [8].

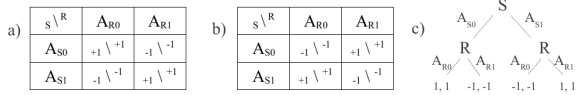
V článku popisujeme doterajšie výpočtové i experimentálne modely vzniku komunikácie založené na čistých kooperatívnych hrách, načrtne ich formalizáciu a navrhne rozšírenie na spomínaný model s tretím hráčom. Ďalej popíšeme architektúru agenta a štruktúru a výsledky simulácií. V diskusii sa venujeme nedostatkom modelu a navrhujeme ich riešenie.

## 2 Spoločné prvky aktuálnych modelov

Komunikačné hry používané vo výpočtových a experimentálnych modeloch možno formalizovať prostriedkami teórie hier [6]. Budeme používať primárne pojmy z multiagentových systémov, ekvivalentné pojmy z teórie hier budú v zátvorke, neskôr aj v texte ako synonymá. Typická epizóda spočíva v interakcii dvoch agentov (hráčov). Každý spracuje nejakú vstupnú informáciu a výstup obsahuje okrem iného aj zvolenú akciu (stratégiu). Odmena (skóre) prichádza v prípade, že kombinácia zvolených stratégií oboch hráčov (profil stratégií) patrí do určitej množiny spoločných akcií (koordinácia). V opačnom prípade nedostane odmenu žiadny z agentov. Po vyhodnotení väčšinou nasleduje ďalšia iterácia medzi tými istými dvoma hráčmi, alebo sa simulácia skončí splnením určitých podmienok.

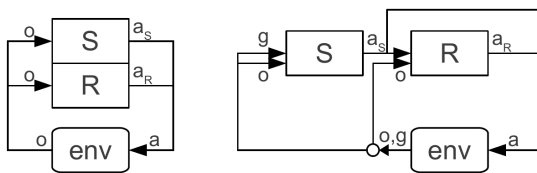
Modely možno rozdeliť do dvoch veľkých kategórií:

hry so simultánnou akciou a hry so sekvenčnou akciou. Najjednoduchším príkladom je koordinačná hra s dvoma možnými stratégiami (obr. 1), ale väčšinou ich býva viac. Niektoré experimenty navyše uvažujú podmnožinu akcií, ktoré nikdy nebudú mať dopad na odmenu, ale možno medzi nimi rozlišovať, a používať ich tak na prenos správy (obr. 3, 4).



**Obr. 1:** Funkcia odmeny v čistej koordinačnej hre dvoch hráčov S a R, kde množina dostupných stratégií pre A je  $\{A_{S0}, A_{S1}\}$  a pre R je  $\{A_{R0}, A_{R1}\}$  vyjadrená a,b) ako tabuľka a c) ako strom. Koordináciami sú profily stratégií a,c)  $[A_{S0}, A_{R0}]$  a  $[A_{S1}, A_{R1}]$ , alebo b)  $[A_{S0}, A_{R1}]$  a  $[A_{S1}, A_{R0}]$ .

Simultánne hry nerozdeľujú rolu hovorcu a príjemcu – obaja vykonávajú nejakú akciu na základe predchádzajúcej akcie spoluhráča. Mohli by sme povedať, že každý agent najskôr plní úlohu príjemcu a následne hovorcu. Takéto hry viac zodpovedajú realite, keďže ľudia vykonávajú svoje akcie súčasne. Môžeme si teda skákať do reči, ale väčšinou to nerobíme. Roly hovorcu a príjemcu teda musia vyemergovať.



**Obr. 2:** Simultánna vs. sekvenčná hra s dvomi hráčmi

Častejšie sa však používajú sekvenčné hry, kde sú roly hovorcu a príjemcu jasne určené pravidlami hry. Hovorca si najskôr zvolí, alebo dostane určené, ktorý z koordinačných profilov idú obaja ako celok vykonať. Popri tom vyšle správu, príjemca ju spracuje a doplní svoju časť výsledného profilu. Za akciu, resp stratégiu sa tu teda považuje nejaký skrytý význam a fyzická správa ide pomimo, t.j. je takým aspektom akcie, ktorý priamo neovplyvňuje, či bude profil koordinačný – mohla by byť hociaká (obr. 2).

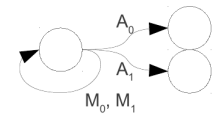
## 2.1 Príklady výpočtových simulácií

Výpočtové simulácie nám umožňujú sformulovať teóriu v podobe počítačového programu a overiť hypotézu pomocou merania priebehu nejakého parametra modelu. Umožňujú prakticky overiť plauzibilitu nejakého formálneho modelu

$S \setminus R$	$A_{R00}$	$A_{R01}$	$A_{R10}$	$A_{R11}$
$A_{S00}$	$+1 \setminus +1$	$+1 \setminus +1$	$-1 \setminus -1$	$-1 \setminus -1$
$A_{S01}$	$+1 \setminus +1$	$+1 \setminus +1$	$-1 \setminus -1$	$-1 \setminus -1$
$A_{S10}$	$-1 \setminus -1$	$-1 \setminus -1$	$+1 \setminus +1$	$+1 \setminus +1$
$A_{S11}$	$-1 \setminus -1$	$-1 \setminus -1$	$+1 \setminus +1$	$+1 \setminus +1$

**Obr. 3:** Hra so správami - akcie vytvárajú disjunktné klastre, rámci ktorých je jedno, ktorú akciu agent zahrá, dôležitý je len prvý bit v indexe. Druhý bit v indexe však môže niesť informáciu o vnútrnom stave agenta a poslužiť na komunikáciu.

$S \setminus R$	$A_{R0}$	$A_{R1}$	$M_{R0}$	$M_{R1}$
$A_{S0}$	$+1 \setminus +1$	$-1 \setminus -1$	$0 \setminus 0$	$0 \setminus 0$
$A_{S1}$	$-1 \setminus -1$	$+1 \setminus +1$	$0 \setminus 0$	$0 \setminus 0$
$M_{S0}$	$0 \setminus 0$	$0 \setminus 0$	$0 \setminus 0$	$0 \setminus 0$
$M_{S1}$	$0 \setminus 0$	$0 \setminus 0$	$0 \setminus 0$	$0 \setminus 0$



**Obr. 4:** Iná hra so správami. Akcie M nevedú k žiadnemu ohodnoteniu. Akcie A sú záväzné, ukončujú reťazec a tvoria klastre (obr. 3). Môžeme limitovať minimálny, alebo maximálny počet správ pred záväznou akciou, alebo vytvoriť zložitejší stavový automat obmedzujúci možné správanie agenta. Oba typy hier sú z hľadiska záverečného ohodnotenia ekvivalentné.

hry v kombinácii s modelom agenta bez matematického dôkazu. Na obr. 5 je načrtnuté, ako možno popísať typickú pomenúvaciu alebo rozpoznávaciu hru pomocou načrtnutého formalizmu.

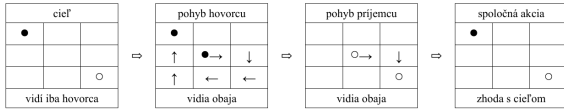
$S \setminus R$	#808080	#008800	#009088	#000080
sivá	$+1 \setminus +1$	$-1 \setminus -1$	$-1 \setminus -1$	$-1 \setminus -1$
šedá	$+1 \setminus +1$	$-1 \setminus -1$	$-1 \setminus -1$	$-1 \setminus -1$
zelená	$-1 \setminus -1$	$+1 \setminus +1$	$+1 \setminus +1$	$-1 \setminus -1$
modrá	$-1 \setminus -1$	$-1 \setminus -1$	$-1 \setminus -1$	$+1 \setminus +1$

**Obr. 5:** Jazyková hra [3, 19, 9] na príklade s farbami. Určitá množina symbolov (napríklad mená farieb) vyslaných hovorcom určuje množinu žiadúcich akcií prijímateľa (napríklad ukázať na príklad takej farby).

## 2.2 Príklady experimentov s ľuďmi

Experimenty s ľuďmi umožňujú otestovať formálny model hry na skutočných ľuďoch. Porovnaním výsledkov výpočtovej simulácie a experimentu s ľuďmi na tom istom formálnom modeli komunikačnej hry možno overiť vhodnosť

modelu agenta ako modelu skutočného človeka a nást' tak minimálnu vhodnú reprezentáciu agenta potrebnú v danej hre na vznik komunikácie s rovnakými vlastnosťami, aká vznikne pri ľudskej interakcii. Na obr. 6 a 7 sú dva príklady takýchto hier, ktoré rozlišujú priestor akcií a stavov. Koordinácia je založená na finálnom stave.



**Obr. 6:** Hra TCG z [5]. Množina možných akcií je ( $\uparrow, \downarrow, \leftarrow, \rightarrow, \text{stoj}, \text{koniec}$ ), stavový priestor prepojený akciami pomerne zložitý, ale intuitívny. Určitá množina akcií je záväzná a stav už potom nemožno zmeniť. Ak spoločná akcia je zhodná s cieľom, došlo ku koordinácii.

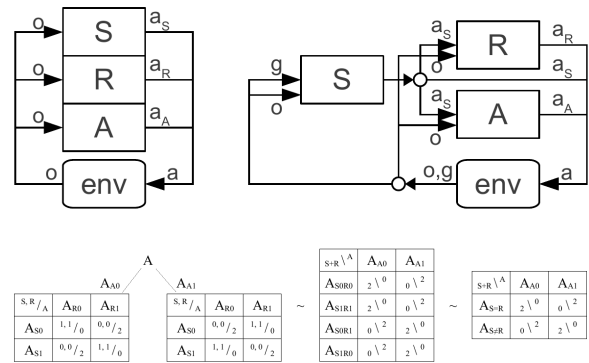


**Obr. 7:** Hra ECG z [17]. Množina možných akcií je ( $\uparrow, \downarrow, \leftarrow, \rightarrow, \text{stoj}, \text{koniec}$ ), stavový priestor zložitý, ale intuitívny. Cieľom je skončiť na rovnakej farbe. Nie každá moja farba sa nutne nachádza u súpera, minimálne jedna však hej. Treba spoločne takú nájsť, pričom komunikovať možno len pomocou pohybov.

### 3 Model spoločného zavádzania

Náš model je založený na väzenskom probléme zo steganografie [2], kde sa strážnik (tretia strana, protivník) snaží uhádnuť, či odpočutá komunikácia medzi dvoma väzňami obsahuje skrytú správu s nekalým úmyslom, alebo nie, a podľa toho buď zasiahnuť, alebo šetriť prostriedky a svoju reputáciu. Na rozdiel od kryptografie nie je primárne dôležité vedieť rozkódovať obsah skrytej správy, iba jej existenciu. Situáciu, keď sa hovorca snaží účinne zamaskovať zamýšľaný odkaz nejakou krycou komunikáciou, možno označiť za zavádzanie tretej strany. Hovoríme teda o modeli spoločného zavádzania.

Akúkoľvek hru môžeme rozšíriť o nového hráča tím, že spoločnú akciu pôvodných hráčov budeme považovať za jednotnú spoločnú akciu celku [21]. V priestore akcií tímu aj protivníka vymedzíme po dve množiny. U tímu pôjde o dva rôzne druhy koordinácie: priamu komunikáciu, kedy nebola prenesená žiadna skrytá správa, a skrytú komunikáciu, kedy bola správa hovorcu myslená iným spôsobom, než

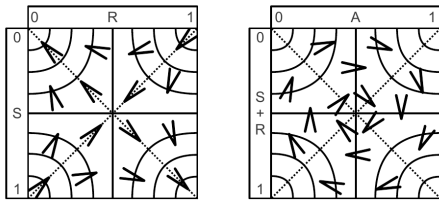


**Obr. 8:** Hore simultánna a sekvenčná verzia hry s protivníkom. Dole ohodnotenie: vľavo dva rôzne druhy koordinácie ako dve nezávislé tabuľky dvoch rôznych koordinačných hier, podľa voľby protivníka, v strede zjednotené akcie hovorcu a príjemcu do spoločnej akcie tímu, vpravo zjednotené klastre – hra tím ako celok vs. protivník.

zrejme rozpozna protivník. Ten naopak musí s prístupom k tým istým vstupom ako príjemca odhadnúť, či je potrebné zasiahnuť, alebo nie. Najjednoduchším spôsobom (obr. 8), ako potom reprezentovať dva rôzne módy koordinácie hovorcu a príjemcu je že z množiny dvoch akcií, ktoré môžeme považovať nejakým spôsobom za navzájom opačné, vykonávajú buď obaja to isté (klasická koordinácia), alebo každý niečo presne opačné ako druhý (tzv. antikoordinácia).

Sekvenciu akcií každého agenta potom možno chápať ako binárny reťazec a pokúsiť sa odhadnúť jeho zložitost' vhodným komprimačným algoritmom. Hra medzi tímom a protivníkom (vnímaná ako dvojhráčová), je príkladom hry "matching pennies". Jeden hráč sa snaží s tým druhým skoordinať, druhý sa snaží urobiť presný opak. Ide v princípe, ktorá má však iba zmiešané ekvilibrium. Aby sme systému umožnili dosiahnuť toto ekvilibrium a neudržiavali ho mimo umelo, môžeme akciu hráča interpretovať pravdepodobnostne a dostaneme vektor pravdepodobnostného rozdelenia akcie (zmiešanú stratégiu). Priestor ohodnotení tiež bude spojený (obr. 9).

Ak sú ohodnotenia možných kombinácií akcií vhodne zvolené, dostaneme hru s ekvilibrium v strede, čo v podstate znamená, že každý agent zvolí svoju akciu úplne náhodne, takže svojou stratégiou z dlhodobého hľadiska nič nezíska, ani nestratí. Skóre a pozorované akcie sú tiež zmiešané. Pri iterovanej interakcii dostávame sekvenciu zmiešaných stratégií, z ktorých môžeme spočítať dlhodobú priemernú pravdepodobnost' akcie. Zmiešané skóre z takéhoto priemeru však nemusí zodpovedať priemernému skóre, ktoré hráč reálne dosiahol. V extrémnom prípade jedna strana vždy "prejde cez rozum" druhej, takže, i keď tím aj pro-



**Obr. 9:** Priestor zmiešaných akcií s vyznačeným ekvilibriom v strede: vľavo v rámci tímu, vpravo tím vs. protivník. Šípky naznačujú očakávanú dynamiku, čiary spájajú body s rovnakým ohodnotením. Osi sú analogicky k tabuľkám z obr. 8: ľavý priestor zodpovedá tabuľke vľavo na obr. 8), prípadne druhej tabuľke zľava, ak by šípky boli naopak; pravý priestor je odvodený od tabuľky úplne vpravo.

tíhráč vsádzajú raz všetko na priamu a raz na skrytú komunikáciu, poradie je také, že vždy víťazí strana s lepším odhadom - schopnosťou vyextrahovať hodnotnú informáciu z predchádzajúcich akcií ostatných hráčov.

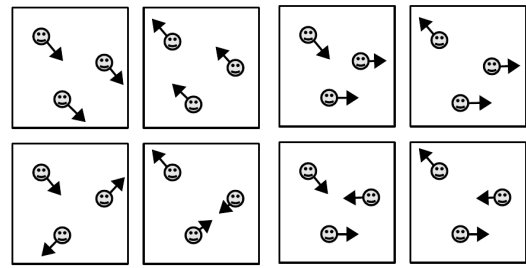
### 3.1 Sociálny priestor a mutliagentová simulácia

Aby sme dali akciám aj nejaký konkrétny význam a aby sme navrhli rozšírenie simulácie na viac než troch agentov, predstavujeme si, že sa nachádzajú v akomsi sociálnom priestore, kde je možné určiť ich vzájomnú sociálnu vzdialenosť. Pre jednoduchosť postačí euklidovský priestor. Z populácie agentov sú potom vždy vybraní traja rôzni jedinci tak, aby ich vzájomné rozloženie do rolí záviselo od ich sociálnej vzdialenosti – dvaja bližší majú väčšiu šancu vytvoriť tím než dvaja vzdialení.

Akcia je okrem ohodnotenia v kombinácii s akciami zvyšných hráčov interpretovaná aj ako zmena v tomto priestore. Priama komunikácia znamená zotrvanie tímu ako celku na jednom mieste, skrytá komunikácia je interpretovaná ako pokus o organizovanú zmenu spoločnej polohy – útek. Ak protivník zasahuje proti skrytej komunikácii, priblíži sa k bodu v prostriedku medzi hráčmi tímu. Z dlhodobého hľadiska sa tak menia ich vzájomné roly. Protivník tak má možnosť prebrať partnera jednému z členov tímu a stať sa jeho novým partnerom. Ak protivník nepredpokladá skrytú komunikáciu, jeho vzdialenosť od tímu sa zväčšuje.

Vyvstáva otázka, ako presne prepojiť akciu hráča, ohodnotenie a pohyb v sociálnom priestore. Jeden spôsob interpretuje akciu hráča abstraktne, t.j. primárne ako voľbu stratégie v jednoduchšej komunikačnej hre so spoločným zavádzaním, a tabuľkovo priradí agentovi okrem ohodnotenia aj zmenu pozície (obr. 10). Tejto hre budeme hovoriť abstraktná hra.

Druhý spôsob je interpretovať akciu primárne ako zmenu polohy v sociálnom priestore a na základe výsled-



**Obr. 10:** Vplyv akcie na pozíciu v sociálnom priestore v 4 základných situáciách - zodpovedajú pozíciám v pravej tabuľke v 8, resp. v pravom priestore zmiešaných akcií na obr. 9. V každom štvorci je vľavo hore protivník, vpravo dole tím. Ľavé 4 štvorce: abstraktná hra - možné pohyby sú iba priblíženie, vzdialenie, útek a ich zmes. Pravé 4 štvorce: stelesnená hra – voľný pohyb.

nej celkovej zmeny vzájomných polôh určiť o akú situáciu v hre šlo a priradiť ohodnotenie. V prípade euklidovského priestoru tak môžu hráči komunikovať pomocou pohybovej gestikulácie (obr. 10). Tejto hre budeme hovoriť stelesnená hra.

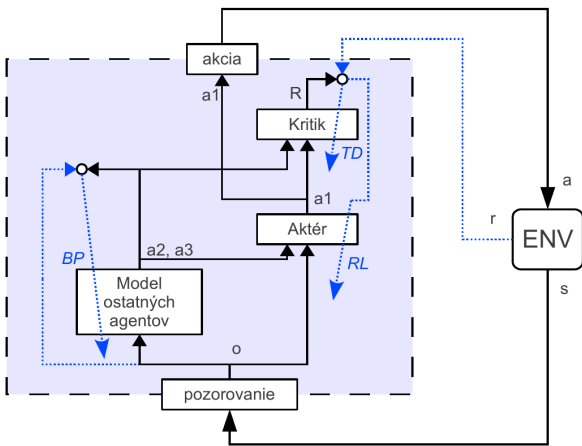
## 4 Architektúra agenta

Agent dostáva na vstupe posledné akcie všetkých agentov vrátane seba samého, plus informáciu z prostredia o svojej roli v interakcii. Na základe toho vráti na výstupe svoju akciu. V tomto prípade je vlastne množina stavov totožná s akciami a aktuálnym stavom je vektor minulých akcií. Zložitejší by bol model, kde agent na vstupe dostáva pozorovanie zmeneného stavu prostredia, a akciu, resp. intenciu ostatných hráčov si musí vnútorne odvodiť.

Keďže agentovou úlohou je naučiť sa správne konať v spojitom priestore akcií, pričom spätná väzba na kvalitu akcie z prostredia má podobu odmeny, jadrom nášho agenta je architektúra Aktér-Kritik pre učenie s posilňovaním v spojitom priestore akcií – CACLA [7]. Na učenie kritika sme použili TD-learning [20]. Navyše je model doplnený o tretí modul Teórie mysle, ktorý sa na základe pozorovania snaží predpovedať intencie ostatných agentov. Ten sa učí spätným šírením chýb podľa následného pozorovania [15]. Miera exploračie Aktéra bola nastavená na konštantnú hodnotu pre všetkých agentov i epizódy. Miera zhodnotenia budúcej odmeny pri učení adaptívneho heuristického kritika pomocou TD-learningu bola tiež nastavená na pomerne vysoké univerzálne konštantné hodnoty ( $\gamma = 0.99$ ), agenti sa teda snažili myslieť mnoho krokov do budúcnosti, ale všetci rovnako ďaleko, čo umožňuje robiť krátkodobé obete.

#### 4.1 Vstupom sú posledné akcie

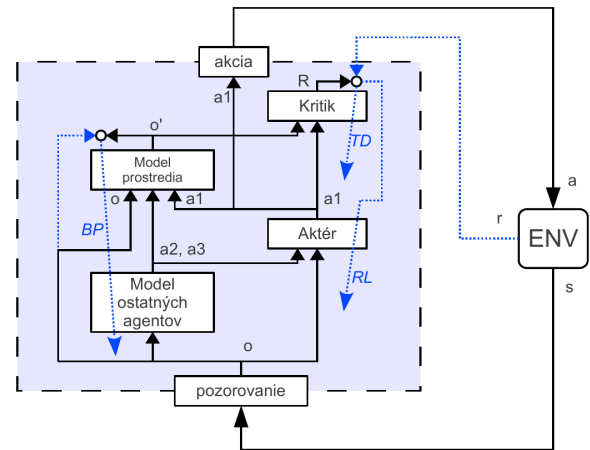
Intenciou ostatných agentov je v tomto prípade ich predpovedaná akcia. Modul Teórie mysle sa môže učiť priamo na základe feedbacku z prostredia. Moduly Aktéra a Teórie mysle obsahujú rekurentné spojenia na skytej vrstve, ktoré sa učia obyčajným spätným šírením chyby. Snažia sa teda pri svojej predpovedi zohľadniť nie len aktuálne, ale aj minulé správanie spoluhráčov. Kritik je prostý dvojvrstvový dopredný perceptrón s jedným lineárnym výstupným neurónom. Je trénovaný predpovedať kumulatívnu odmenu pomocou TD-learningu.



**Obr. 11:** Aktuálna architektúra. Modré bodkované čiary naznačujú signál použitý na trénovanie modulu skratkou naznačeným algoritmom.

#### 4.2 Vstupom sú pozorovania stavov prostredia

Reprezentácia stavu a akcie je tu oddelená, agent teda nedostáva feedback o akcii agenta, iba o jej pozorovanom dôsledku na prostredie. Musíme mať teda dopredný model na predpovedanie zmeny stavu na základe vlastnej akcie a predpovedanej akcie agentov. Predpoveď akcie agenta, ktorá je nutným vstupom do aktéra sa musí odvodiť na základe spätného šírenia chyby z dopredného modelu prostredia, alebo iného algoritmu učenia. Nikdy teda nebudeme mať na tejto vrstve reprezentáciu zhodnú s akciami ostatných agentov, ale mali by sme dostať niečo veľmi príbuzné informačnou hodnotou – môžeme to nazvať intenciou agenta. Týmto sa z modulu Teórie mysle stáva vlastne inverzný model, ktorý rozpoznáva úmysel ostatných agentov. Bolo by možné trénovať ho explicitne pomocou sebaopozorovania.



**Obr. 12:** Architektúra pre hry s oddelenou reprezentáciou stavu a akcie. Pribudol dopredný model sveta. Modul teórie mysle nie je explicitne trénovaný.

### 5 Výsledky simulácií

Uvedieme zaujímavé výsledky pre dve verzie simultánnej hry so sociálnym priestorom (abstraktnú a stelesnenú, kapitola 3.1) a stavom (vstupnou informáciou agentov) totožným s poslednými akciami (použili sme agentov z kapitoly 4.1). Zameriame sa na 3 druhy simulácie:

1. s tromi agentami s pevnými rolami – aby sme preskúmali dlhodobé vlastnosti interakcie z pohľadu každej roly. V tomto prípade u abstraktnej hry až tak nezáleží na sociálnom priestore, ale môže byť zaujímavé sledovať vývoj;
2. s tromi agentami so striedaním rolí – aby sme preskúmali schopnosť agenta naučiť sa vystupovať v rôznych roliach a vplyv sociálneho priestoru na celkovú interakciu;
3. s mnohými agentami so striedaním rolí – aby sme preskúmali emergentné vlastnosti multiagentovej simulácie

Pôvodná pozícia agentov v sociálnom priestore je dvojica uniformných náhodných reálnych čísel medzi 0 a 1, ale môže byť ľubovoľná. Vektor zmeny polohy agenta je upravený na interval -0.01 až 0.01 a je prevedený z maximovej do euklidovskej normy, aby si agent volil pohyb v ľubovoľnom smere do euklidovskej vzdialenosti 0.0001 od aktuálnej polohy.

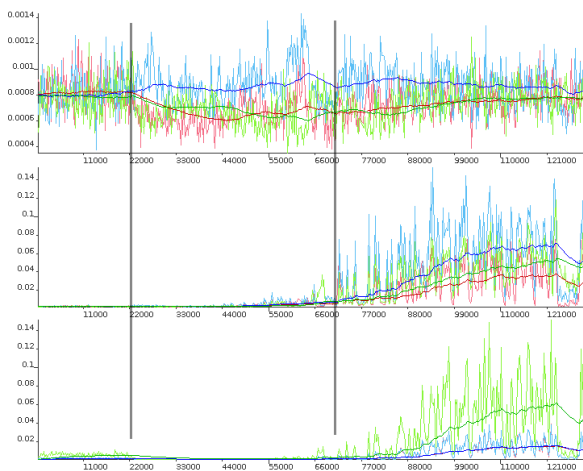
Algoritmus výberu agentov do trojice najskôr s uniformnou distribúciou náhodne vyberie hovorcu a príjemcu a protivníka vyberie – šanca byť vybraný exponenciálne klesá so vzdialenosťou agenta, pričom koeficienty sú nastavené

tak, aby do istej vzdialenosti bola väčšia šanca byť vybraný za príjemcu než protivníka, od istého bodu rastie šanca stať sa protivníkom a pri veľmi veľkých vzdialenostiach sa pomery vyrovnávajú, ale pravdepodobnosť interakcie je veľmi malá.

Trojica agentov zaradená do rolí vykoná v tejto podobe 1000 epizód a potom sa volí nová trojica (môže byť aj tá istá, ale s inými rolami). Za jednu takúto epochu sa teda agent posunie maximálne o vzdialenosť 0.1. Epoch vykonáme desiatky až stovky tisíc, aby sme preskúmali dlhodobé správanie systému a aby sa všetci agenti mali šancu zúčastniť dostatočne veľa krát.

### 5.1 Traja hráči s pevnými rolami

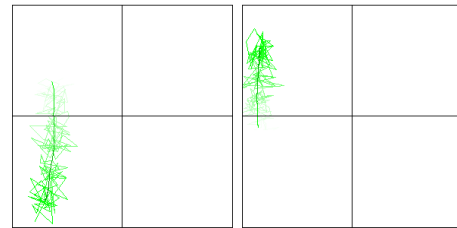
Na začiatok sme vyskúšali dynamiku samotnej abstraktnej hry. Populácia troch hráčov s pevne zvolenými rolami integrovala dokola stále podľa tých istých pravidiel. Sledovali sme stredné kvadratické chyby učenia na každom module pre jednotlivé roly, ako aj pohyb hráčov v abstraktnom priestore akcií a dopad na sociálny priestor. Na vývoji chýb možno spozorovať 3 fázy, ktoré sa vyskytovali pravidelne. Pre rôzne parametre učenia, ako rýchlosť učenia, momentum, explorácia a váha budúceho ohodnotenia, mali rôzny priebeh a dĺžku trvania, ale vždy sa objavili.



**Obr. 13:** Od vrchu chyba modulov: Aktér, Teória mysle, Kritik. Vyznačené hranice fáz. Roly sú farebne rozlíšené: červená - hovorca, modrá - príjemca, zelená - protivník.

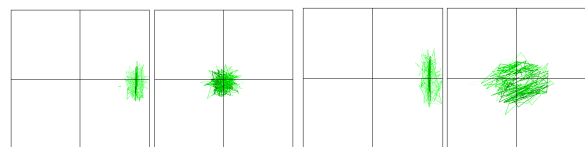
V prvej fáze sa učila najmä slučka Aktér-Kritik, čiastočne aj Teória mysle. Po pomerne krátkej dobe a zbežnom preskúmaní akčného priestoru, chyba poklesla a správanie sa ustálilo v blízkosti ekvilibria. Zvyčajne v tejto situácii jeden z členov tímu zaujal pasívnu rolu a vykonával jednu stabilnú akciu, zatiaľčo druhý hráč súperil s protivníkom za

celý tím. Na konci prvej fázy pasívnemu hráčovi a protivníkovi poklesla chyba Aktéra, aktívnemu členovi tímu však naopak mierne vzrástla. Podobne to bolo s modulom Teória mysle. Chyba kritika poklesla všetkým na veľmi nízke hodnoty.



**Obr. 14:** Správanie agentov v akčnom priestore v prvej fáze, vľavo vrámci tímu, vpravo tím vs. protivník (pozri obr. 9). Tmavozelená je odporúčaná akcia Aktéra, bledozelená je reálne vykonaná explorovaná akcia. Body po sebe idúcich akcií sú spojené čiarou, ktorá časom bledne.

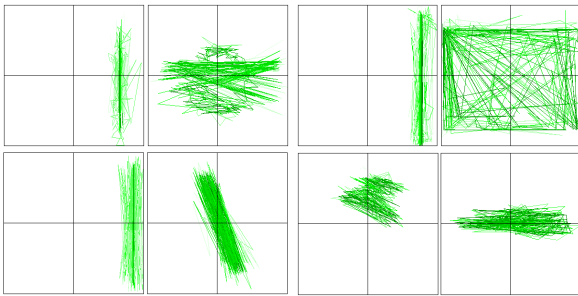
V druhej fáze už agenti zhruba poznali dôsledky svojich akcií a vzrástol význam schopnosti predpovedať akciu ostatných agentov. Exploráciou spôsobené kombinácie akcií v okolí ekvilibria občas viedli k úspešnej koordinácii tímu, občas k úspešnému predídaniu koordinácii zo strany protivníka. Z času načas dôjde k úspešnej koordinácii tímu viackrát po sebe, napriek tomu, že protivník sa snažil koordinácii predísť. Tím sa teda naučí, ako zareagovať na akciu protivníka, akoby vedel ďalšiu. Systém tak preskúma oblasť okolo ekvilibria a agenti sa učia navzájom si "prejsť cez rozum". Chyba Teória mysle u všetkých pomaly rastie, oblasť priestoru akcií okolo ekvilibria sa zväčšuje a reakcie sú čoraz prudšie, až dôjde k relatívne prudkému zlomu.



**Obr. 15:** Správanie agentov v akčnom priestore v druhej fáze, vľavo na začiatku, vpravo na konci.

Prudko narastie chyba Teória mysle a Kritik do takých rozmerov, že rozdiel medzi prvou a druhou fázou je zanedbateľný. V tejto tretej fáze už agenti majú zložitý model vhodných reakcií, využívajú celý akčný priestor a reagujú na seba veľmi zložitým, zdanlivo chaotickým spôsobom, pričom stále hľadajú lepšie riešenia. Občas sa na rôzne dlhú dobu vrátia do blízkosti ekvilibria, ale pomerne rýchlo sa z neho opäť vzdialia. Členovia tímu si v tejto fáze zvyknú občas vymeniť rolu aktívneho a pasívneho hráča.

V stelesnenej hre existuje druhý rozmer pre akciu,



**Obr. 16:** Rôzne správanie agentov v akčnom priestore v tretej fáze: prechod medzi dvoma typmi správania, preskakovanie medzi rohmi kvadrantov, dočasné víťazstvo tímu, výmena rolí aktívneho a pasívneho hráča.

ktorý je v sociálnom priestore interpretovaný ako pohyb do boku. Tento otvára agentom oveľa väčšie možnosti akcie a agenti častejšie volia správanie, ktoré nemá dopad na ich ohodnotenie, z hľadiska hry je blízko ekvilibría. Preskúvanie akčného priestoru im dlhšie trva, postupom času ale dochádza k podobnému správaniu ako v abstraktnej hre.

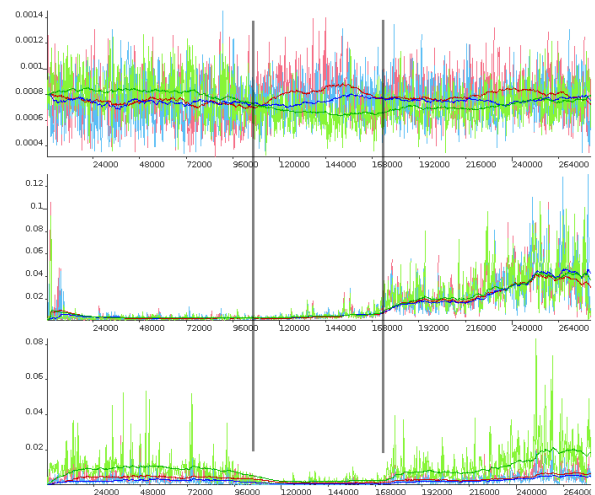
## 5.2 Traja hráči so striedaním rolí

Cieľom bolo otestovať schopnosť hráčov striedať si roly a zistiť dopad sociálneho priestoru na správanie. Keďže však agent nedostáva informáciu o polohe ostatných v sociálnom priestore, nemôže v tomto smere veľmi plánovať. Môže sa ale naučiť uprednostniť tie druhy koordinácie, ktoré vedú k častejšiemu vyberaniu do tímu s tým istým hráčom. V abstraktnej hre opäť pozorujeme 3 fázy učenia a správania, ale celý proces trvá dlhšie, keďže agent sa musí naučiť tri rôzne módy správania. V prvej fáze je výraznejšie trénovanie modulu teórie mysle, ktoré v druhej fáze u všetkých poklesne, rastie len pomaly a potom v tretej fáze prudko narastie.

Z hľadiska sociálneho priestoru chvíľu trvá než agenti vytvoria nejaký stabilný stav, väčšinou prechádzajú fázou, kedy sú od seba približne rovnako ďaleko a prudko si striedajú roly, nakoniec však niektorí dvaja vytvoria trvalý pár a tretí je odsunutý do roly trvalého protivníka. V stelesnenej hre sú zmeny správania podobné, až na to že prvá a druhá fáza splývajú v jednu.

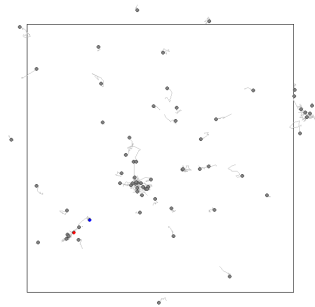
## 5.3 Mnoho agentov

Priebeh učenia je podobný ako v predchádzajúcom bode, ale zmeny trvajú ešte dlhšie, keďže je potrebné naučiť sa správanie väčšieho počtu rôznych agentov. U abstraktnej i stelesnenej hry sa agenti postupne združia do skupiniek cca 2 až 20 jedincov, ktoré sa pohybujú po priestore ako ce-



**Obr. 17:** 3 fázy pri striedaní rolí. Druhá fáza sa skracuje.

lok. Občas niektorý člen skupinky opustí, alebo cudzí agent okolo skupinky najskôr krúži a potom sa pripojí. Občas dôjde k spojeniu dvoch skupiniek do jednej. Stelesnená simulácia ale nakoniec vedie do stavu, keď je každý agent osamotený a najbližší agenti sú od neho približne rovnomerne ďaleko. Čím sme ďalej od stredu populácie, tým väčšie sú aj priemerné vzdialenosti medzi agentami, systém je bezškálový. Predpokladáme, že tieto vlastnosti sú dané pravidlom vyberania agentov do interakcie.



**Obr. 18:** Multiagentová simulácia s abstraktnou hrou. Bodky predstavujú agentov, čiarky za nimi ich niekoľko minulých zaznamenných polôh. Štvorec vyznačuje hranice priestoru, v ktorom pôvodne všetky vznikli. Mnoho agentov je mimo záber obrázka. Zjavné sú tri skupinky. V tejto simulácii nakoniec ľavé dve skupinky splynuli do jednej, pravá sa udržala rozdelená.

## 6 Diskusia

V našom modeli sme priviedli akcie na vstup z vlastného pohľadu agenta. Na jednej pozícii je vždy vlastná akcia (ja), na inej akcia spoluhráča (ty) a na ďalšej protihráča (on). Protivník nemá skutočného spoluhráča, tak dostával na druhej pozícii akciu hovorcu, podobne ako príjemca, pričom navzájom bol s príjemcom protihráč. Agent by si mohol vedieť ovládať tento referenčný, či pozornosťný rámec a vytvoriť tak stuácie podobné seba pozorovaniu, vciťovaniu sa, imitácii a podobne. Za týmto účelom by si mohli moduly Aktéra a Teórie mysle svoje výstupy navzájom poskytovať ako učebné vstupy.

Od spôsobu výberu agentov do trojice závisí vzdialenosť vzniknutých skupiniek a finálne rozloženie individuálnych agentov. Protivník by si agent mohol sám určovať, aká veľká je jeho skupina potenciálnych vhodných partnerov a koho chce sledovať. Ak by bola simulácia rozšírená o evolučné mechanizmy, mnohé podobné parametre by sa mohli prispôbovať pomocou ich vplyvu na fitness.

Sekvenčná hra by jasnejšie určila agentom ich roly (hovorca a príjemca by v rámci tímu plnili iné úlohy), a umožnila by lepšie skúmať samotnú komunikáciu. Spomedzi existujúcich modelov sa simultánna hra podobá najviac na ECG, ostatné modely sú sekvenčné. Prostredie by určilo hovorcovi, čo má vykonať príjemca, a hovorca by sa snažil na jednej strane odovzdať túto správu a na druhej strane prispôbiť svoju akciu tak, aby zmiatol prípadného protivníka. Na takýto model sa chceme sústrediť v najbližšej budúcnosti, navyše v ňom chceme rozlišovať medzi priestorom stavov a akcií.

Komplexitu správania agenta možno analyzovať metódami teórie informácie a zložitosti. Najjednoduchším spôsobom zmerania zložitosti reťazca je skomprimovať ho vhodným algoritmom. Do úvahy pripadá shanonova entropia, kolmogorovská zložitnosť, alebo štatistická komplexita [4], od ktorej je potom odvodená aj matematická definícia emergencie [18].

Zaujímavé by bolo vyskúšať model na ľudských hráčoch, čím by sa jednak overila schopnosť ľudí riešiť takúto úlohu, zložitost' nájdených riešení a vhodnosť navrhnutého modelu agenta na reprezentovanie skutočného človeka v danej situácii.

## 7 Záver

Náš model spoločného zavádzania ukazuje potenciál podobnej trojhráčkovej hry pri modelovaní vzniku komunikácie. V modeli emerguje určité striedanie rolí v tíme a po istej dobe nastáva komplexné správanie. V prostredí s populáciou v sociálnom priestore dochádza ku vzniku krátko-odobu stabilných skupín. V ďalšej práci sa chceme zamerať

na vytvorenie sekvenčného modelu s oddelenými reprezentáciami stavov a akcií, vyhodnotenie komplexity akcie, vyhodnotenie prítomnosti jednej alebo viacerých úrovní komunikácie a porovnanie dynamiky vzniknutej komunikácie pri dvoch a troch hráčoch.

## PodĎakovanie

Tento príspevok vznikol v Centre pre kognitívnu vedu na Katedre aplikovanej informatiky, FMFI UK v Bratislave, za podpory grantov VEGA 1/0439/11 a 1/0503/13 a GUK/643/2012.

## Literatúra

- [1] Borenstein, E., Ruppín, E.: The evolution of imitation and mirror neurons in adaptive agents. *Cognitive Systems Research* 6(3), 229–242 (2005)
- [2] Cachin, C.: An information-theoretic model for steganography. *Information and Computation* 192(1), 41–56 (2004)
- [3] Cangelosi, A., Parisi, D.: The emergence of a 'language' in an evolving population of neural networks. *Connection Science* 10(2), 83–97 (1998)
- [4] Crutchfield, J.P., Young, K.: Inferring statistical complexity. *Physical Review Letters* 63(2), 105–108 (1989)
- [5] De Ruiter, J.P., Noordzij, M.L., Newman-Norlund, S., Newman-Norlund, R., Hagoort, P., Levinson, S.C., Toni, I.: Exploring the cognitive infrastructure of communication. *Interaction Studies* 11(1), 51–77 (2010)
- [6] Demichelis, S., Weibull, J.: Language, Meaning, and Games: A Model of Communication, Coordination, and Evolution. *The American Economic Review* 98(4), 1292–1311 (2008)
- [7] van Hasselt, H., Wiering, M.A.: Reinforcement learning in continuous action spaces. In: *Approximate Dynamic Programming and Reinforcement Learning, 2007. ADPRL 2007. IEEE International Symposium on*. pp. 272–279. IEEE (2007)
- [8] Knight, C.: Ritual/speech coevolution: a solution to the problem of deception. In: Hurford, J.R., Studdert-Kennedy, M., Knight, C. (eds.) *Approaches to the evolution of language: social and cognitive bases*, pp. 68–91. Cambridge University Press, Cambridge (1998)



- [9] Kvasnička, V., Pospíchal, J.: An emergence of coordinated communication in populations of agents. *Artificial Life* 5(4), 319–342 (1999)
- [10] Malinovský, : Modelovanie biologicko-kultúrnej koevolúcie jazyka a teórie mysle. In: Kelemen, J., Kvasnička, V., Pospíchal, J. (eds.) *Kognice a umělý život XI*, pp. 133–140. Slezská univerzita, Opava (2011)
- [11] Malinovský, , Farkaš, I.: Triadická interakcia v emergencii komunikácie. In: Kelemen, J., Nahodil, P. (eds.) *Kognice a umělý život XII*, pp. 137–142. Agentura Action M, Praha (2012)
- [12] Maynard-Smith, J.: Must reliable signals always be costly? *Animal behaviour* (1994)
- [13] Oliphant, M.: Formal Approaches to Innate and Learned Communication: Laying the Foundation for Language. Ph.D. thesis, University of California, San Diego (1997)
- [14] Perfors, A.: Simulated Evolution of Language: A Review of the Field. *Journal of Artificial Societies and Social Simulation* 5(2) (2002)
- [15] Rumelhart, D.E., Hintont, G.E., Williams, R.J.: Learning representations by back-propagating errors. *Nature* 323(6088), 533–536 (1986)
- [16] Scott-Phillips, T., Kirby, S.: Language evolution in the laboratory. *Trends in cognitive sciences* 14(9), 411–417 (2010)
- [17] Scott-Phillips, T.C., Kirby, S., Ritchie, G.R.: Signalling signalhood and the emergence of communication. *Cognition* 113(2), 226–233 (2009)
- [18] Shalizi, C.R., et al.: Causal architecture, complexity and self-organization in the time series and cellular automata. Ph.D. thesis, University of Wisconsin–Madison (2001)
- [19] Steels, L.: The origins of syntax in visually grounded robotic agents. *Artificial Intelligence* 103(1-2), 133–156 (1998)
- [20] Sutton, R.S.: Learning to predict by the methods of temporal differences. *Machine learning* 3(1), 9–44 (1988)
- [21] Von Neumann, J., Morgenstern, O.: *The theory of games and economic behavior* (1947)
- [22] Zahavi, A.: The fallacy of conventional signalling. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 340(1292), 227–230 (1993)



## ARCH/GARCH modely a ich neuronová reprezentácia

Dušan Marček<sup>1</sup>, Marek Horváth<sup>2</sup>

<sup>1</sup>Výzkumný ústav Centra excelence IT4Innovation (Information Technology for Innovation)  
Slezská univerzita v Opave, Bezručovo nám.13, 746 01 Opava  
dusan.marcek@fpf.slu.cz

<sup>2</sup>Ústav informatiky, Filozoficko-prirodovedecká fakulta Slezská univerzita v Opavě, Bezručovo nám.13, 746 01 Opava  
mac@fife.sk

### Abstrakt

Článok opisuje vývoj a aplikáciu ARIMA/ARCH-GARCH modelu. V prvej časti poskytuje všeobecne návod na vývoj ARCH-GARCH modelov pre predikciu dlhopisov VUB banky a v druhej časti porovnáva predikčnú presnosť modelov s modelmi založenými na teórii neuronových sietí (granulárnych RBF (Radial Basic Function) neuronových sietí), ktoré sú schopné, v dôsledku zohľadnenia nelineárnych závislostí v modeloch, lepšie zachytiť vývojovú dynamiku časových radov vysokofrekvenčných dát a tým vylepšujú ich predikčné schopnosti. V závere sa zmienujeme o manažérskych aspektoch aplikácií uvedených metodológií v praxi v predikčných systémoch na podporu rozhodovania manažmentu malých a stredných podnikov a finančných inštitúcií.

### 1 Úvod

Modelovanie vysokofrekvenčných finančných časových radov sa dostalo už dávnejšie do popredia záujmov analytikov aj teoretických pracovníkov v oblasti ekonómie a informatiky a v súčasnosti sa stáva jednou z najfrekvencovanejších výskumných tém finančnej ekonometrie a aplikovanej informatiky pre ekonomické finančné procesy. Tieto modely sú vhodným nástrojom na opis vnútorného mechanizmu, ktorý generuje dáta skúmaného procesu. Ich teoretická prepracovanosť je už na takej úrovni, že sa dajú ľahko implementovať do počítačových programov, a tak masovejšie aplikovať tak na prognostické ciele v ekonomike, ako aj dynamické riadenie technologických, výrobných a kontrolných systémov. Modely ARMA majú však tri podstatné obmedzenia, ktoré musíme mať na zreteli pri ich aplikáciách. V prvom rade sú to lineárne modely, v ktorých výstupná premenná je modelovaná ako lineárna funkcia svojich hodnôt z predchádzajúcich období a aktuálnych a oneskorených hodnôt poruchového člena. Linearita podstatne obmedzuje

ich aplikáciu v dynamických systémoch. Druhým obmedzením modelov ARMA je predpoklad o nemennosti rozptylu poruchového člena, a tým aj nemennosti rozptylu modelovej veličiny počas jej celého vývoja. Tretie obmedzenie ARMA modelov vyplýva z podmienky stacionárnosti pre časové rady. Stacionarita časového radu, t. j. predpoklad o jeho časovej invariantnosti pre strednú hodnotu a rozptyl je ťažko udržateľná v mnohých finančných časových radoch. Nestacionárnosť dát v ekonomických časových radoch je často dôsledkom šokových opatrení v ekonomike. Napr. vývoj kurzov hlavných svetových mien má zjavnú rastúcu a následne klesajúcu tendenciu. V časových radoch základných makroekonomických veličín, ako sú národný dôchodok a jeho zložky, ponuka krátkodobých cenných papierov a pod., spravidla vykazujú vzostupný trend. V niektorých obdobiach býva vzostupný trend prerušený výrazným poklesom s následným obnovením rastu. Premennivosť časových radov nie je konštantná v čase. Napr. v devízových trhoch je determinovaná prítomnosťou obchodníkov v hlavných obchodných centrách, a tá je premenlivá v priebehu dňa a týždňa.

Je všeobecne známe, chaotické systémy, medzi ktoré nesporne patria finančné a predovšetkým devízové trhy, generujú náhodné veličiny, ktoré majú podobné vlastnosti ako biely šum, v skutočnosti ale neodhaľujú ich pravú podstatu. Časové rady zaznamenané v takýchto systémoch môžu mať dlhodobý konštantný rozptyl, ale v niektorých kratších intervaloch je rozptyl relatívne vyšší alebo nižší než dlhodobý rozptyl. Takéto časové rady sa nazývajú podmienené heteroskedastické. Štúdium rozptylu náhodnej zložky a explicitné zohľadnenie podmienenej heteroskedasticity v modeloch časových radov poskytuje lepšie výsledky pri konštrukcii prognóz. Argumentmi funkcie, ktorá opisuje vývoj podmienenej heteroskedasticity (rozptylu) časového radu, sú spravidla minulé realizácie časového radu. Ide teda o autoregresnú

reprezentáciu podmienenej heteroskedasticity, čo je súčasne vysvetlenie vzniku pomenovania ARCH (Autoregressive Conditionally Heteroscedastic) modelov.

Ako bolo už spomenuté, všetky spomínané modely majú nedostatok, že sú to lineárne modely. V praxi sa väčšinou nevystačí s lineárnymi modelmi. Avšak v modelovaní veličín spravidla s vysokofrekvenčnými dátami môže detekcia nelineárnych skrytých vzorov v krátkodobom časovom období podstatne vylepšiť predikčnú schopnosť modelov, čo bude aj ukázané na nelineárnych predikčných modeloch založených na umelých neuronových sieťach.

V príspevku sú prezentované niektoré výsledky konštrukcií modelov na prognózovanie vysokofrekvenčných finančných dát dlhopisov VUB banky. Príspevok má nasledujúcu štruktúru. V druhej kapitole je krátko opisovaná trieda ARCH-GARCH model a ich variantov: EGARCH, PGARCH a TGARCH modelov. V tretej kapitole sú prezentované dáta a demonštruje sa predikčná schopnosť ARCH/GARCH modelov. Vo štvartej kapitole je prezentovaný prístup modelovania a predikcie cien dlhopisov pomocou RBF UNS a v kapitole piatej je krátko naznačený výskumný zámer v danej oblasti.

## 2 Niektoré ARCH-GARCH modely pre finančné dáta

Jednotlivé typy ARCH modelov sa vzájomne odlišujú podľa toho, ako je modelovaný rozptyl poruchového člena stochastického procesu. Vychádza sa z predpokladu, že poruchový člen  $\varepsilon_t$  v modeli časového radu je heteroskedastický. Poruchový člen pre všetky pozorovania  $N$  vytvára postupnosť  $\{\varepsilon_t\}$ , ktorého hodnoty, za predpokladu prítomnosti heteroskedasticity, spĺňajú nasledovné podmienky

$$E(\varepsilon_t) = 0; \quad \text{cov}(\varepsilon_t, \varepsilon_s) = \begin{cases} \sigma_t^2 & \text{pre } t = s \\ 0 & \text{pre } t \neq s \end{cases} \quad (1)$$

Postupnosť poruchových členov  $\{\varepsilon_t\}$  nemá konštantný rozptyl vo všetkých pozorovaniach. Rozptyl je v čase premenlivý.

Prvý model, ktorý poskytol ucelený rámec pre modelovanie rozptylu je Englov ARCH model [5]. Bollerslev [2] rozšíril ARCH model známy ako zovšeobecnený ARCH (GARCH) pre postupnosť  $\{Y_t\}$ , v

ktorom podmienený rozptyl  $h_t$  závisí aj od skorších podmienených rozptylov, t. j.

$$h_t = \alpha_0 + \sum_{i=1}^m \alpha_i Y_{t-i}^2 + \sum_{j=1}^s \beta_j h_{t-j}, \quad (2)$$

kde  $\{V_t\}$  j postupnosť IID (Independent Identical Distribution) náhodnej veličiny s nulovou strednou hodnotou a jednotkovým rozptylom. Výraz (2) sa zredukuje na ARCH( $m$ ) model, ak  $s = 0$ ,  $\alpha_j$  a  $\beta_j$  sú potom parametre ARCH/GARCH modelu.

V literatúre je známych niekoľko rozšírení základného GARCH modelu (2). V literatúre možno nájsť rozšírenia, ako sú EGARCH a PGARCH. TGARCH (exponenciálny GARCH) model umožňuje zachytávať javy „dobrých“ a „zlých“ správ vo vývoji volatility finančných časových radov. Na vývoj dát vysokofrekvenčných časových radov majú výrazný vplyv negatívne šoky, t. j. nepriaznivé postoje investorov a celkové postoje nepriaznivé k ekonomickej situácii, nepriaznivé správy a pod. „Dobré“ správy majú omnoho menší účinok na volatilitu.

Nelson [13] navrhol nasledujúci EGARCH (Exponential GARCH), ktorý berie do úvahy pákové efekty, t. j. že pokles ceny aktíva vedie k väčšej volatilita než rovnaký vzostup aktíva. Jeho formula je nasledujúca

$$\log h_t = \alpha_0 + \sum_{i=1}^p \alpha_i \frac{|\hat{\varepsilon}_{t-i}| + \gamma_i \hat{\varepsilon}_{t-i}}{\sigma_{t-i}} + \sum_{j=1}^q \beta_j h_{t-j}, \quad (3)$$

kde  $\hat{\varepsilon}_{t-i}$  sú reziduály,  $\gamma_i$  sú koeficienty pákových efektov. Ak  $\hat{\varepsilon}_{t-i}$  je kladné, t. j. sa predpokladajú len “dobré” informácie, výsledný efekt z  $\hat{\varepsilon}_{t-i}$  je  $(1 + \gamma_i)\hat{\varepsilon}_{t-i}$ . Opačne, ak  $\hat{\varepsilon}_{t-i}$  je záporné, t. j. predpokladajú sa nepriaznivé informácie, potom celkový efekt z  $\hat{\varepsilon}_{t-i}$  je  $(1 - \gamma_i)|\hat{\varepsilon}_{t-i}|$ . Keďže nepriaznivé informácie majú väčší účinok na volatilitu, očakávaná hodnota pákového efektu by mala byť záporná [5]. Ding, Granger, Engle [4] vyjadrili identifikáciu pákového efektu modelom s označením PGARCH (Power GARCH) v tvare

$$h_t^d = \alpha_0 + \sum_{i=1}^p \alpha_i (|\hat{\varepsilon}_{t-i}| + \gamma_i \hat{\varepsilon}_{t-i})^d + \sum_{j=1}^q \beta_j h_{t-j}^d, \quad (4)$$

kde  $d$  je kladný exponent. Ak  $d = 2$ , potom model (3.3) sa redukuje na základný GARCH model s pákovými efektmi. Iným typom GARCH modelu, ktorý je schopný identifikovať pákový efekt je model s označením TGARCH (Threshold GARCH). Model je známy aj pod označením GJR-GARCH, pretože Glosten, Jagannathan a Runkle [6] (1993) navrhli rovnaký typ rozšírenia. Model má tvar

$$h_t = \alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^q \gamma_i S_{t-i} \varepsilon_{t-i}^2 + \sum_{j=1}^q \beta_j h_{t-j}, \quad (5)$$

kde  $S_{t-i} = 1$ , ak  $\varepsilon_{t-i} < 0$  a  $S_{t-i} = 0$ , ak  $\varepsilon_{t-i} \geq 0$ . V závislosti od toho či je hodnota  $\varepsilon_{t-i}$  nad, resp. podprahovou hodnotou nuly,  $\varepsilon_{t-i}^2$  má rôzny vplyv na podmienený rozptyl  $h_t$ . Ak hodnota  $\varepsilon_{t-i}$  je kladná, celkový efekt bude  $\alpha_i \varepsilon_{t-i}^2$ . Ak je hodnota  $\varepsilon_{t-i}$  záporná, celkový efekt bude  $(\alpha_i + \gamma_i) \varepsilon_{t-i}^2$ . V tomto modeli, na rozdiel od EGARCH modelu, očakávaná hodnota koeficientu pákového efektu pri negatívnych informáciách by mala byť kladná. Iné ARCH-GARCH modely ako sú ARCH-GARCH regresný model ARCH-GARCH – M (Mean) možno nájsť v [11,12].

### 3 Aplikácia ARIMA-ARCH/GARCH modely na údaje časového radu výnosov VUB dlhopisu

Denné údaje časového radu výnosov VUB dlhopisu  $\{r_t\}$  boli zaznamenávané od 7. mája 2004 do 28 februára 2008, čo predstavuje celkom 954 dát. Tieto dáta sú uvedené na [http://fria.fri.uniza.sk/files/data\\_VUB](http://fria.fri.uniza.sk/files/data_VUB). Úlohou je navrhnúť prediktor vývoja výnosov dlhopisu. Pretože ide o dlhý časový rad, je namieste predpoklad, že rozptyl časového radu je v čase premenlivý. Pre konštrukciu prediktora najskôr celú množinu dát rozdelíme na 2 podmnožiny. Prvá podmnožina – podmnožina pre odhad parametrov modelu (trénovacia množina) obsahuje dáta pozorovaní  $t = 1, 2, \dots, 900$  a druhá – validačná podmnožina pre konštrukciu ex post predpovedí, obsahuje dáta s pozorovaniami  $t = 901, 902, \dots, 954$ .

Aplikáciou metód štatistickej analýzy dát trénovacej množiny s použitím ACF, PACF, AIC/BIC bol predbežne identifikovaný a kvantifikovaný model časového radu ako proces ARMA(5,0) v tvare

$$r_t = \xi + \phi_1 r_{t-1} + \phi_2 r_{t-2} + \phi_3 r_{t-3} + \phi_4 r_{t-4} + \phi_5 r_{t-5} + \varepsilon_t, \quad (6)$$

kde  $\xi, \phi_1, \phi_2, \dots, \phi_5$  sú parametre modelu,  $\varepsilon_t$  sú IDD, t. j. nezávislé hodnoty náhodnej veličiny s normálnym rozdelením pravdepodobnosti, s nulovou strednou hodnotou a konštantným rozptylom.

Aby sme zistili, či rozptyl je konštantný vo všetkých pozorovaniach dlhopisu, dáta dlhopisu sme podrobili LM

testu, ktorý potvrdil prítomnosť podmienenej heteroskedasticity. Použitím software R2.5.0 (<http://cran.r-project.org>) model bol špecifikovaný ako ARMA(5,0)+GARCH(1,1) a je kvantifikovaný v tvare

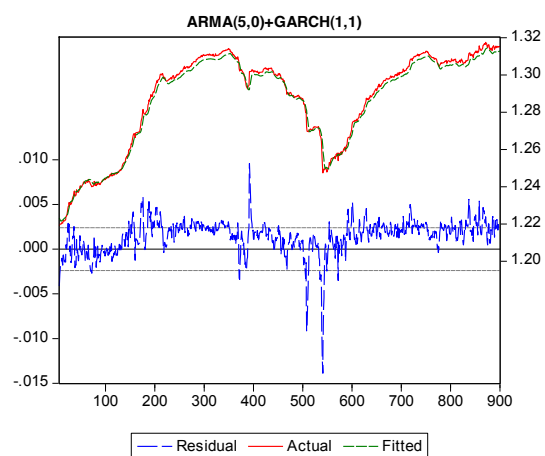
$$r_t = 0,0000748 + 0,06628r_{t-1} + 0,09557r_{t-2} + 0,0275r_{t-3} + 0,0528r_{t-4} + 0,09795r_{t-5} + \varepsilon_t$$

s výrazom pre premenlivý rozptyl  $h_t$

$$h_t = 1,958 \cdot 10^{-8} + 0,1887 \varepsilon_{t-1}^2 + 0,8075 h_{t-1}, \quad (7)$$

kde  $\varepsilon_t$  je postupnosť odhadnutých reziduálov z rovnice (6).

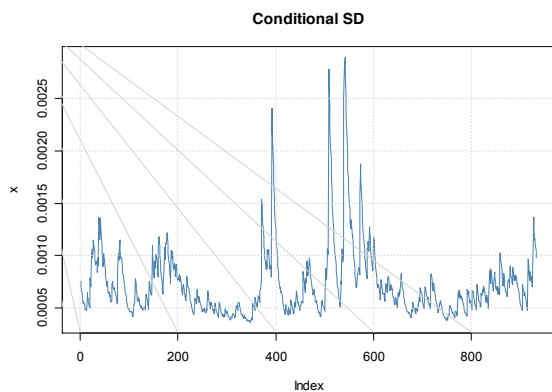
Po kvantifikácii modelu môžeme preveriť, či reziduá modelu (7) sú s vlastnosťami IDD. Testovanie bolo vykonané prostredníctvom software Eviews (<http://www.eviews.com>) pomocou BDS testu aplikovaným na štandardizované reziduá  $\hat{\varepsilon}_t$  získané ako  $\hat{\varepsilon}_t = \varepsilon_t / \sqrt{h_t}$  pre dimenzie  $m = 2, 3$  a pre tolerančné vzdialenosti (parametre blízkosti)  $\varepsilon = 0,5; 1,0; 1,5$  a  $2,0$ . BDS testom bola potvrdená hypotéza, že dáta sú IDD. Priebeh skutočných a odhadnutých hodnôt modelom (7) a jeho reziduálov je vidieť na obr. 1. Odhad volatily a jej časový vývoj je na obr. 2.



Obr. 1. Skutočné a odhadnuté hodnoty VUB dlhopisu: model ARMA(5,0) + GARCH(1,1) ARCH-GARCH model (7). Hodnoty a vývoj reziduálov sú zakreslené v spodnej časti. Skutočné hodnoty výnosov VUB dlhopisu sú zakreslené plnou čiarou, odhadnuté hodnoty prerušovanou čiarou. 1.

V literatúre [13] možno nájsť ďalšie rozšírenia základného GARCH modelu (2) ako sú EGARCH a PGARCH. TGARCH (exponenciálny GARCH) model umožňuje zachytávať javy „dobrých“ a „zlých“ správ vo vývoji volatility finančných časových radov.

Až doteraz pri modelovaní časového radu dlhopisu VUB sme predpokladali, že v štandardnom GARCH modeli (7) dáta aj poruchový člen pochádzajú z normálneho rozdelenia, čo potvrdzuje aj diagram QQ plot štandardizovaných. V software R systéme aj iných štatistických vývojových nástrojoch je možné vytvárať zovšeobecnené – GED (Generalized Error Distribution) štandardizované reziduály a studentizované reziduály.



Obr. 2. Odhad volatility a jej vývoj pre ARMA(5,0) + GARCH(1,1) proces, model (7).

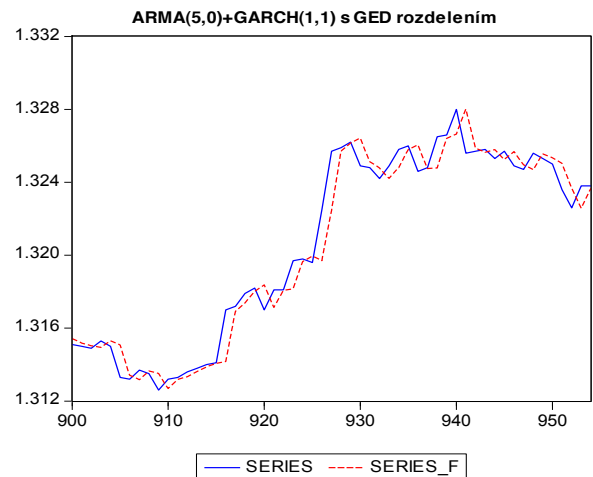
Nakoniec pre indikáciu pákového efektu sme kvantifikovali aj EGARCH a PGARCH model. Napr. pre EGARCH model, t. j., model ARMA(5,0) + EGARCH(1,1) koeficient pákového efektu modelu (3)  $\gamma = -0,210$ . Záporná hodnota pákového efektu znamená, že celková ekonomická situácia nie je priaznivá, ide skôr o „zlé“ vyhladky do budúcnosti. Ak analyzujeme vývoj odhadov volatility na obr. 1 a vývoj výnosov VUB obligácie na obr. 1, zisťujeme, že v periódach ekonomickej depresie („zlé“ správy) pákový efekt spôsobuje asymetrické skoky volatility.

To, že najpriaznivejšie štatistické charakteristiky v tabuľke 1 poskytuje model s GED údajmi, sa odrazilo aj v jeho ex post predpovediach. Tento model generuje presnejšie ex post predpovede (založené na charakteristike RMSE) v porovnaní s analyzovanými variantnými modelmi, čo je vidieť v tabuľke 1.

Tabuľka 1. Hodnoty RMSE pre rôzne rozšírenia základného GARCH modelu.

Rozdelenie: Model:	Gaussovo	$t$	GED
AR(5)+ GARCH(1,1)	0,003461	0,002345	0,001056
AR(5)+ EGARCH(1,1)	0,001066	0,001064	0,001063
AR(5)+ PGARCH(1,1)	0,001064	0,001064	0,001062

Vývoj skutočných hodnôt a ex post predpovedí modelu ARMA(5,0) + GARCH(1,1) s dátami GED je na obr. 3.



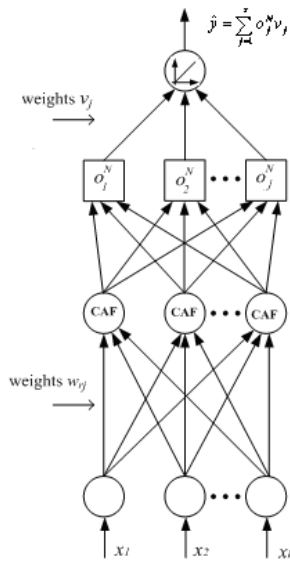
Obr. 1. Skutočné hodnoty (plná čiara) a predpovede (prerušovaná čiara) výnosov VUB dlhopis.

#### 4 Alternatívny (neurónový) model

Rovnaké dáta aké boli použité pre štatistický model v odhade jeho parametrov boli použité v prípade modelu založenom na G RBF UNS, ktorého architektúra je zobrazená na obrázku 4. Premenné na pravej strane modelu (6) boli použité ako vstupné veličiny do G RBF UNS. V tejto sieti pre odhad hodnôt váh skrytých neurónov bola použitá zhlukovacia metóda aplikovaná na vstupné dáta siete. Ako algoritmus pre vyhľadanie centier zhlukov bola použitá adaptívna verzia Kohonenovho konkurenčného učenia. V skrytých neurónoch boli aplikované Gaussovské aktivačné funkcie s cloud konceptom. Vo výstupnom neuróne bola použitá lineárna prenosová funkcia. Výstupné

hodnoty zo stredných neurónov  $o_j^N$  boli normalizované tak, aby súčet týchto hodnôt bol rovný jednej [9].

Uvedenú architektúru siete v podstate možno chápať ako nelineárnu náhradu odhadu štatistického parametrov modelu (7) AR(1) + EGARCH(1,1) za nelineárny (neurónový) odhad parametrov siete.



Obr. 4. Granulárna RBF sieť s cloud aktivačnými funkciami (CAF).

Granulárna RBF sieť na obrázku 4 predpokladá, že hodnota šumu entropie je známa. V našom prípade táto hodnota bola volená ako 0,015 násobok entropie (smerodajnej odchýlky Gaussovskej RBF funkcie). Ďalším voliteľným parametrom je parameter označovaný symbolom  $K$  pre prekryvanie rozdelení vstupných dát. Predikčná presnosť architektúr RBF sietí bola vyjadrená kritériami RMSE na validačnej (ex post) množine. Detailný výpočtový algoritmus pre výpočet RMSE (MSE) hodnôt a váh siete možno nájsť v [11]. Predikčné výsledky z neurónovej alternatívy sú uvedené v tabuľke 2.

## 5 Empirické zhodnotenie

Z tabuliek 1 a 2 je vidieť, že modely založené na štatistickej metodológii (ARIMA a ARCH/GARCH modely) majú oproti konkurenčným RBF UNS modelom lepšiu predikčnú schopnosť. Modely založené na RBF UNS však majú stále vysokú (vyhovujúcu) presnosť pre použitie v manažérskych predikčných systémoch.

Tabuľka 2. RMSE miery presnosti pre ex post predpovede pre rôzne architektúry RBF UNS (podrobnosti sú uvedené v texte).

Počet RBF neurónov ( $\mathcal{S}$ )	$K$	RMSE
1	1.25	0.00719
	4.0	0.00716
5	1.25	0.00756
	4.0	0.00758
10	1.25	0.00720
	4.0	0.00715

Špecifikované modely neurónovej alternatívy spĺňajú takmer všetky atribúty “dobrých” modelovacích a predikčných nástrojov. Nemajú mnoho vysvetľujúcich veličín. V porovnaní so štatistickou alternatívou RBF NN modely sú po ekonomickej a výpočtovej stránke efektívnejšie, sú jednoduchšie, jednoducho modifikovateľné na zmeny v existujúcich predikčných systémoch, vhodné pre prognózovanie vysokofrekvenčných dát, s možnosťou ľahkého zakomponovanie kvalitatívnych (expertných) informácií. ARCH-GARCH a ARIMA modely vyžadujú vyššie náklady na ich vývoj, zavedenie a prevádzku v manažérskych predikčných systémoch, sú náročné na ich porozumenie manažérmi. Inou nevýhodou týchto modelov je, že neexistuje jednoducho využiteľný spôsob pre aktualizáciu parametrov modelu, ak sú k dispozícii nové (aktuálne) dáta. Model sa musí periodicky kompletne vyvíjať.

## 5 Výskumné témy do budúcnosti

Z hľadiska blízkej budúcnosti sa predpokladá sústrediť sa v rámci granularných výpočtov na vyhľadanie prenosových funkcií RBF neurónov so zakomponovaním argumentu premenlivej hyperentropie v závislosti od hodnoty funkcie príslušnosti vstupov do zhlukov popisovaných cloud modelmi (s Gaussovským, GED a Studentovým rozdelením).

## PodĎakovanie

Tento ělánok bol vypracovaný v rámci projektu Centrum excellence IT4Innovations, reg. ě. CZ.1.05/1.1.00/02.0070 podporovaněho Operaěněm programom Vězkum a věvoj pro inovace, financovaněho zo ťtrukturálnych fondov EU a zo ťtátneho rozpoětu ĆR.

## Literatěra:

- [1] Bishop, C.H. M., *Neural Networks for Pattern recognition*. (Oxford University Press, New York, 1995).
- [2] Bollerslev, D. Generalized Autoregressive Conditional Heteroscedasticity, *Journal of Econometrics* 31 (1986) 307-327.
- [3] Brock, W., Dechert, W., Scheinkman, J. and LeBaron, B. *A Test for Independence Based on the Correlation Dimension*. Forthcoming in *Econometric Reviews*, (1996).
- [4] Ding, Z., Granger, C.W., and Engle, R.F. A Long Memory Property of Stock Market Returns and a New Model, *Journal of Empirical Finance*, 1, (1993) 83-106.
- [5] Engle, R.F. Auto regressive Conditional Heteroscedasticity with Estimates of the Variance of United Kindom Inflation, *Econometrica* 50 (1982) 987-1007.
- [6] Glosten, L.R., Jaganathan, R. and Runkle, D.E.: On the relation between the expected value and the volatility of the nominal excess returns on stocks, *Journal of Finance* 48, 1993, pp. 1779-1801
- [7] Li, D, and Du, Y. *Artificial intelligence with uncertainty* (Boca Raton: Chapman & Hall/CRC, Taylor & Francis Group, 2008).
- [8] Liao, K.P. & Fildes, R. *The accuracy of a procedural approach to specifying feedforward neural networks for forecasting*. *Computers & Operations Research*, 32, (2005) 2151-2169.
- [9] Kecman, V., *Learning and soft computing: support vector machines, neural networks, and fuzzy logic*, (Massachusetts: The MIT Press, 2001).
- [10] Marcek, M. and Marcek, D. *Granular RBF Neural Network Implementation of Fuzzy Systems: Application to Time Series Modelling*, *J. of Multi-Valued Logic & Soft Computing*, 14 (2008) 400-414.
- [11] Marcek, M., Pancikova, L. and Marcek, C. *Econometrics & Soft Computing* (The University Press, Zilina, 2008). Zilina: The University of Zilina Press, (in Slovak).
- [12] Marcek, D. and Marcek, M. *Time Series Analysis, Modelling and Forecasting with Applications in Economics* (The University Press, Zilina, 2001).
- [13] Nelson, D.B. Conditional Heteroskedasticity in Asset Returns: a New Approach, *Econometrica* 59 (2) (1991) 347-370.
- [14] Yao, Y.Y. *Granular Computing for Data Mining*. Kissimmee: Congres of Data Mining, Intrusion Detection, Information Assurance and Data Network Security 2006. Vol. 6241. 16<sup>th</sup> – 17<sup>th</sup> April 2006, pp. 624105.1-624105.12
- [15] Zhang, G.P. & Qi, M. *Neural network forecasting for seasonal and trend time series*. *European Journal Of Operational Research*, 160, (2005) 501-514.
- [16] Zivot, E., Wang, J. *Modeling Financial Time Series with S-PLUS®*, (Springer Verlag, NY, 2005).



# Komplexné siete a modelovanie javov reálneho sveta

Mária Markošová

Katedra aplikovanej informatiky,  
Fakulta matematiky, fyziky a informatiky, Univerzita Komenského  
84248 Mlynská dolina, Bratislava  
Email: markosova@fmph.uniba.sk

## Abstrakt

Táto práca si kladie za cieľ ukázať prehľad aplikácií teórie komplexných sietí v modelovaní javov reálneho sveta. V prvej časti krátko opíšem základné štatistické miery, pomocou ktorých možno opísať štruktúru a vlastnosti komplexných sietí. Nasledujúce kapitoly sa venujú modelovaniu sociálnych, biologických, komunikačných a iných reálnych sietí. Pretože aplikácií je veľké množstvo, vybrané sú najtypickejšie a najkurióznejšie prípady. Posledná časť sa zaoberá našim výskumom v oblasti komplexných sietí, kde patrí aplikácia teórie komplexných sietí na štúdium lexikónu jazyka a funkčných prepojení v mozgu.

## 1 Základné vlastnosti komplexných sietí

Komplexné siete sú medzi vedeckými pracovníkmi čoraz populárnejšie. Dôvodom je fakt, že poskytujú dobrý nástroj na analýzu a modelovanie rozsiahlych distribuovaných systémov. Mnoho javov v reálnom svete možno skúmať tak, že na ne nahliadame ako na súbor v podstate identických interagujúcich entít. Napríklad v sociálnej sieti Facebook ľudia predstavujú entity a vzťahy medzi nimi vyjadrené facebookovým priateľstvom sú interakciami. Podobne v internetovej sieti napríklad rootre môžu predstavovať interagujúce elementy a fyzické prepojenia medzi nimi interakcie. Aj vedecký článok môže byť interagujúcou jednotkou; napríklad si môžeme predstaviť, že je v interakcii so všetkými prácami, ktoré sú v ňom citované. Určite by sa dalo nájsť mnoho podobných príkladov. Niektoré budú podrobnejšie rozobraté v ďalších kapitolách tejto práce. Reálne distribuované systémy sú obyčajne veľké a pozostávajú aj z niekoľkých tisícov interagujúcich jednotiek. Preto poznatky z teórie komplexných sietí môžu poskytnúť dobrú možnosť porozumieť ich štruktúre a postaviť zodpovedajúci model.

Mojou ambíciou je v tejto práci ukázať, že mnohé reálne siete, ktoré vznikli samoorganizujúcimi sa procesmi, zdieľajú podobné užitočné vlastnosti. Ich uzly, napríklad, nie sú v priemere veľmi ďaleko od seba, ak

vzdialenosť meriame počtom hrán, cez ktoré treba prejsť od jedného uzla k druhému. Preto jednotlivé uzly môžu spolu dobre komunikovať. Reálne siete majú takmer vždy netriviálnu štruktúru. Sú zložené z veľkého množstva slabo zapojených periférnych uzlov a niekoľkých dôležitých, dobre zapojených hubov. Sú málo odolné voči cieľným útokom na huby, ale dobre odolávajú náhodným poruchám. Ak napríklad v internetovej sieti niekto vyradí z prevádzky dôležitý uzol spojený s mnohými inými (hub), môže spôsobiť značné problémy v internetovej komunikácii. Ak však náhodne vyberieme nejaký uzol a vyradíme ho z prevádzky, je veľmi pravdepodobné, že to nie je hub, ale periférny uzol a preto sa nič závažného nestane. Navyše, reálne siete sú často hierarchicky organizované.

### 1.1 Akú sieť nazývame komplexnou sieťou?

Ak nejaký systém nazývame „komplexným“, intuitívne rozumieme, že tento systém má buď netriviálnu štruktúru, alebo dynamiku, alebo oboje. „Netriviálny“ znamená, že nie je jednoduché túto štruktúru, alebo procesy v dynamickom systéme prebiehajúce, formalizovať. Pod pojmom „sieť“ budeme rozumieť dynamicky sa meniacu (rastúcu) štruktúru, ktorej matematickou reprezentáciou je graf.

Predstavme si, napríklad, sieť o  $N$  uzloch. Nech hrana medzi každou dvojicou uzlov existuje s pravdepodobnosťou  $p$ . Takáto sieť má pre pomerne veľký interval hodnôt  $p$  relatívne jednoduchú štruktúru. Väčšina uzlov má približne ten istý počet susedov; teda ten istý **stupeň**  $k$ . Uzlov, ktoré majú výrazne viac alebo výrazne menej susedov, je relatívne málo. Preto, ak sa pýtame, ako dobre je typický uzol zapojený do siete, koľko susedov má typický uzol, predstavu získame tak, že vypočítame priemerný stupeň uzla  $\bar{k}$ , teda jediný parameter. Ale väčšina reálnych sietí nemá takúto jednoduchú štruktúru. Ak sa na ne pozeráme z hľadiska stupňov uzlov, vidíme, že sa skladajú z uzlov rôznych stupňov. Neexistuje v nich nijaký typický uzol. Niektoré uzly majú mnoho susedov a nazývame ich hubmi. Mnoho uzlov má však len niekoľko susedov, teda malý

stupeň. Preto, aby sme získali akúsi predstavu o štruktúre takejto komplexnej siete, nevystačíme s jediným parametrom, napríklad priemerným stupňom uzlov. Musíme urobiť histogram, kde na osi  $x$  bude stupeň uzla a na osi  $y$  normovaný počet uzlov, ktoré majú daný stupeň. Urobíme teda distribúciu stupňov uzlov. Ak sa distribúcia stupňov uzlov správa podľa mocninného zákona

$$P(k) \propto k^{-\gamma}, \quad (1.1)$$

potom sieť nazývame **bezškálovou** a  $\gamma$  **škálovacím exponentom**. Slovo bezškálová naznačuje, že neexistuje typický uzol, s typickým stupňom ktorý by vnášal do systému typickú škálu. Ako ukázali Barabási a Albert [1], štruktúra siete charakterizovaná distribúciou (1.1) sa vyvinie vtedy, keď sieť rastie pomocou preferenčného pripájania nových uzlov. Znamená to, že uzol, ktorý prišiel do siete ako nový, sa pripojí s väčšou pravdepodobnosťou k takému uzlu, ktorý má väčší stupeň („rich gets richer“ zákon).

Ďalší parameter nazývame klasterizačným koeficientom uzla  $x$  a opisuje hustotu prepojení v okolí uzla  $x$

$$C_x = \frac{e_x}{\binom{k_x}{2}}. \quad (1.2)$$

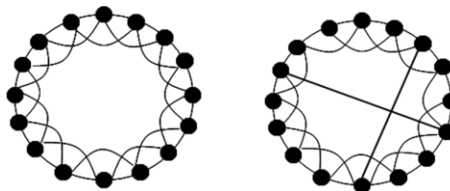
Vo vzťahu (1.2)  $e_x$  predstavuje počet skutočne existujúcich hrán medzi susedmi uzla  $x$  a  $k_x$  je stupeň uzla  $x$ , a teda aj počet jeho susedov. Veličina  $C_x$  teda hovorí o tom, aký zlomok zo všetkých možných hrán v okolí uzla  $x$  skutočne existuje.

Ak má sieť jednoduchú štruktúru, priemerný klasterizačný koeficient  $\bar{C}$  je dobrou mierou pre popis okolia typického uzla. Ak tomu tak nie je, opäť sa musíme uchýliť k zložitejšiemu popisu. Napríklad môžeme spočítať priemery klasterizačných koeficientov uzlov stupňa  $k$  a urobiť ďalší histogram. Po prenormovaní dostanem distribúciu priemerných klasterizačných koeficientov uzlov stupňa  $k$ ,  $C(k)$ . Ak platí, že

$$C(k) \propto k^{-\delta}, \quad (1.3)$$

potom, ako ukázali Ravasz a Barabási [2], má sieť hierarchickú štruktúru, ktorá je tiež výsledkom dynamiky siete. Takáto štruktúra vznikne vtedy, ak sa k sieti buď pripája celý súbor uzlov usporiadaných do nejakého vzoru, alebo len jeden uzol, ktorého všetky hrany sa pripoja na uzly – susedov [3]. V súčasnosti sa intenzívne skúmajú dynamické procesy, ktoré vedú na hierarchickú štruktúru siete [4].

Bezškálová štruktúra siete je veľmi často kombinovaná s hierarchickou štruktúrou. Navyše reálne siete sú často aj tzv. **sieťami malého sveta** [5]. Pod sieťou malého sveta rozumieme takú sieť, ktorá má pomerne veľký klasterizačný koeficient  $C$ , ale malú separáciu uzlov  $l$  (kde pod separáciou uzlov rozumiem malú priemernú najkratšiu vzdialenosť medzi uzlami). Štruktúru sietí malého sveta možno pochopiť pomocou obr. 1. Na obr. 1 a) je sieť s pravidelnou štruktúrou, kde priemerný klasterizačný koeficient siete je  $\bar{C} = 0.5$  a separácia uzlov rastie s rastúcim počtom uzlov lineárne  $l \propto N$ . Na obr. 1b) je sieť malého sveta. Vďaka náhodným krátkym spojeniam sa radikálne skraca separácia uzlov ( $l \propto \ln(N)$ ) ale, ak krátkych spojení je relatívne málo, lokálna štruktúra siete ostáva prakticky neporušená ( $\bar{C} = 0.5$ ). Štruktúra siete malého sveta tak leží medzi pravidelnou a náhodnou, čo má pre reálne siete isté výhody. Uzly vďaka krátkym spojeniam majú k sebe blízko, čo umožňuje dobrú komunikáciu medzi nimi, ale pritom uzly stále tvoria viac menej pevnú mriežku s dobre definovanou štruktúrou.



Obr 1. a) Pravidelná sieť. b) Sieť malého sveta.

## 2 Modelovanie javov reálneho sveta pomocou komplexných sietí

V tejto časti sa zamýšľame nad tým, ako možno reálne distribuované systémy modelovať pomocou sieťových modelov. Ukážem niekoľko typických príkladov z rôznych vedných oblastí. Uvediem aj niekoľko reálnych sietí, ktoré majú skôr charakter kuriozity.

### 2.1 Sociálne siete

Sociálne siete boli jedným z prvých systémov, pri štúdiu ktorých sa aplikovali sieťové modely a poznatky z teórie grafov. Už z roku 1967 poznáme Milgramov slávny pokus zistiť, cez koľko prostredníkov sa dokážu skontaktovať dvaja náhodne vybraní ľudia v populácii USA [6]. Po vyhodnotení výsledkov Milgram sformuloval zákon šesťstupňovej separácie a zaviedol pojem „malý svet“. Vlastnosť malého sveta, teda malá separácia uzlov je vlastná aj ostatným sociálnym sieťam.

Pod sociálnou sieťou rozumieme takú sieť, kde uzly tvoria ľudia (alebo zvieratá) a interakcia medzi dvoma uzlami predstavuje hranu. Interakcie nie je vždy ľahké definovať, pretože napríklad pojem priateľstva je v rôznych kultúrach chápaný rôzne. Preto sociálne siete tvorené medziľudskými vzťahmi trpia často nejasnou definíciou toho, čo je vlastne hranou, kedy hrana medzi uzlami existuje a kedy už nie.

Existujú však sociálne siete, kde je pojem interakcie jednoznačne daný. Napríklad sieť filmových hercov alebo iných profesionálov [7, 8], sieť sexuálnych kontaktov [9] a sieť mobilných hovorov [10]. V hereckej sieti sú herci uzlami a ak dvaja herci hrali spolu v jednom filme, potom sú spojení hranou. Takáto sieť profesionálnych kontaktov veľa napovie o úlohe a popularite jednotlivých hercov. Pribúdanie nových uzlov do tejto siete je pravdepodobne dané preferenčným pripájaním, pretože vo filmovom priemysle je tendencia, aby film obsadilo aspoň niekoľko známych hviezd. Pre nováčika je teda pravdepodobnejšie, že bude hrať vo filme s hviezdou ako s inými nováčikmi. Tento spôsob rastu siete naznačuje aj mocinná distribúcia stupňa uzlov v hereckej sieti (1.1), kde škálovací exponent  $\gamma \approx 2.3$ . Ďalšou vlastnosťou hereckej siete je fakt, že má štruktúru malého sveta, teda vysoký klasterizačný koeficient a nízku separáciu uzlov. Dvaja náhodne vybraní herci sa dokážu skontaktovať cez menej ako štyroch prostredníkov. Vysoký klasterizačný koeficient naznačuje, že herci majú tendenciu hrať s osvedčenými partnermi. Priemerný stupeň uzla v tejto sieti je dosť vysoký  $\bar{k} \approx 29.0$  a dá sa vysvetliť veľkým množstvom filmov v ktorých herec počas svojej hereckej kariéry hrá.

Onnela so spolupracovníkmi [10] skonštruoval sieť mobilných hovorov v USA, ktorá mala  $4.6 \times 10^6$  uzlov a  $7.0 \times 10^6$  hrán. V tejto sieti uzly reprezentujú používatel'ov mobilných telefónov a hrana medzi dvoma uzlami vznikne vtedy, keď si daní používatelia zatelefonojú. Autori sledovali hovory počas osemnástich mesiacov a vytvorili sieť s váhovanými hranami, kde váha reprezentovala trvanie hovoru oboch osôb. Distribúcia váh aj stupňov uzlov je mocinná (1.1) s koeficientmi  $\lambda = 8.1$  a  $\gamma = 1.9$ . Analýza tejto siete teda ukázala, že sa jedná o bezškálovú sieť v ktorej najsilnejšie prepojenia (s najväčšou váhou) vznikajú v komunitách priateľov a vzájomné prepojenia medzi komunitami sú len relatívne slabé. Odtiaľ vyplýva ďalšia vlastnosť tejto sociálnej siete, že jej rozpad (globálny efekt) je skôr spôsobený vymazaním hrany s nízkou váhou ako s vysokou.

Liljeros a spol. [9] študovali sieť sexuálnych kontaktov vo švédскеj populácii indivíduí vo veku od 18 do 74 rokov. Svoje pozorovania založili na starostlivo

vypracovaných dotazníkoch, v ktorých sa pýtali na počet sexuálnych partnerov za jeden rok. Uzlami v tejto sieti sú teda osoby a hrana medzi nimi vznikne vtedy, ak majú vzájomný sexuálny kontakt. Opäť, takáto sieť je komplexnou sieťou s mocinnou distribúciou stupňov uzlov (1.1) s exponentom  $\gamma = 2.54$  v prípade žien a  $\gamma = 2.31$  pre mužov. Priemerný počet partnerov (priemerný stupeň uzla) je v prípade mužov o niečo vyšší ako v prípade žien. Ako možnú príčinu autori uvádzajú fakt, že kvôli spoločenským očakávaniam muži v dotazníkoch počet svojich partneriek nadhodnocujú. Autori urobili aj ďalšie pozorovanie, v ktorom sa pýtali na počet partnerov v doterajšom živote. Takáto sieť je tiež bezškálová ale s nižšími škálovacími exponentmi  $\gamma = 2.1$  v prípade žien a  $\gamma = 1.6$  v prípade mužov. Autori tvrdia, že bezškálový charakter siete sexuálnych kontaktov je spôsobený preferenčným pripájaním uzlov, pretože čím má jedinec viacej partnerov (vyšší stupeň) tým skôr získa ďalších, pretože má vytríbenejšiu techniku zvädzania. Autori uvažujú aj o šírení sexuálnych chorôb a tvrdia, že úspešná kampaň na zamedzenie sexuálnych chorôb by mala smerovať k vybraným jedincom, ktorí tvoria huby v sieti sexuálnych kontaktov. Ich pozorovanie je potvrdené aj štúdiom šírenia epidémií po bezškálových sieťach, kde autori potvrdili, že huby sú mediátormi rýchleho šírenia [11].

Zaujímavú štúdiu sociálneho správania delfínov publikoval David Lusseau [12]. Sedem rokov pozoroval izolovanú komunitu delfínov v zálive Doubtfull Sound na Novom Zélande a vytvoril ich sociálnu sieť, v ktorej jednotlivé delfíny sú uzlami a hrana medzi dvoma uzlami vznikne vtedy, ak dvaja jedinci sú pozorovaní spolu častejšie, ako by bol ich náhodný spoločný výskyt. Hoci sieť má len niečo cez šesťdesiat uzlov, distribúcia stupňa uzlov sa zdá byť bezškálová. Huby a teda aj centrá komunitného života v tejto sieti skoro výlučne (až na jednu výnimku) tvorili staršie samice. Pre bezškálové siete je typické, že sa rozpadajú, ak vymažeme uzly, ktoré majú vysoký stupeň. Pre komunitu delfínov to však tak nie je. Ak aj samica, ktorá je hubom v sociálnej sieti zomrie, sieť sa nerozpadne. Nerozpadne sa dokonca ani vtedy, ak umrie viacero kľúčových jedincov naraz. Komunita aj po katastrofických udalostiach stále drží pospolu.

Zaujímavú prácu, pod názvom The Medieval Inquisition: Scale Free Networks and the Suppression of Heresy publikoval Paul Omreod [13]. Zo stredovekých historických prameňov čerpal poznatky o potláčaní heretických názorov vo Francúzsku. Zistil, že katolícka inkvizícia sa v priebehu niekoľkých desaťročí naučila ako účinne potláčať herézu a robila to takým spôsobom, ako keby vedela, že heréza je „choroba“, ktorá sa šíri po bezškálovej sociálnej sieti. Katolícke authority už

začiatkom trinásteho storočia porovnávali šírenie heretických názorov so šírením chorôb, ale domnievali sa, že ak sa nedá oddeliť choré od zdravého, treba vyhubiť všetkých podozrivých. Pripomína to vymazávanie náhodne zvolených uzlov v bezškálovej sieti. Neviedlo to k ničomu, herézy prepukali ďalej. Neskôr si inkvizitóri uvedomili, že treba zmeniť taktiku a zamerať svoje úsilie na niekoľko málo centrálnych osôb (hubov) ktoré veľa cestujú a majú mnoho kontaktov. Takéto úsilie malo väčší úspech a potlačilo organizovaný dissent vo Francii až do čias Reformácie.

## 2.2 Transportné siete

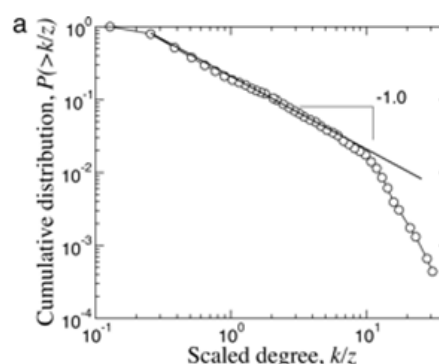
Tieto siete sú veľmi dôležité z hľadiska ekono-mického rozvoja krajín. Študujú sa hlavne preto, aby ich bolo možné čo najviac optimalizovať. Patrí k nim, napríklad, sieť letísk, diaľnic, železníc a podobne. Sú špecifické v tom, že ich štruktúru ovplyvňuje hustota obyvateľstva aj politicko-ekonomická situácia v tej ktorej časti sveta.

V sieti letísk, jednotlivé letiská predstavujú uzly a ak medzi dvoma letiskami existuje letecké spojenie, je medzi nimi hrana. Guimera [14] so spolupracovníkmi zistili, že letiskovú sieť možno považovať za sieť s neorientovanými hranami, pretože vo veľkej väčšine prípadov, ak lietadlo letí jedným smerom, existuje spojenie aj opačným smerom. Analyzovali celosvetovú sieť letísk a došli k zaujímavým záverom. Zistili, že veľké kľúčové letiská pracujú na hranici kolapsu, pretože obsluhujú viacej lietadiel, ako je optimálne obslúžiť za pekného počasia. A tak každý, aj ten sebemenší problém, ako napríklad zatiahnutá obloha, ktorý zníži o pár percent počet pristání a štartov za hodinu, má za následok opožďovanie letov. To sa, samozrejme, prenáša na ďalšie a ďalšie letiská a má ekonomické dôsledky. Preto je dôležité porozumieť štruktúre letiskovej siete a pokúsiť sa o jej optimalizáciu.

Autori [14] si problém letiskovej siete upravili. Za uzly nepovažovali jednotlivé letiská, ale mestá. Vyhli sa tak skresleniu dát, pretože medzi letiskami v jednom meste logicky letecké spojenie neexistuje. Výsledky analýzy ukázali, že letisková sieť je sieťou malého sveta. Dve náhodne vybrané mestá na zemeguli sú v priemere spojené leteckou trasou, ktorá ide cez štyri mestá. Je silne klasterizovaná, s priemerným klasterizačným koeficientom  $\bar{C} = 0.62$  a zložená z komunit. Letisková sieť je zároveň aj bezškálovou sieťou (obr.2).

Táto sieť je špecifická v tom, že netypicky, uzly s najvyšším stupňom nie sú zároveň aj uzlami s najväčšou centralitou, teda tými, cez ktoré prechádza najviac najkratších ciest medzi uzlami. Pod najkratšou cestou medzi dvoma uzlami siete rozumieme takú cestu, ktorá tieto uzly spája najmenším počtom hrán zo

všetkých možných ciest medzi nimi. Stupeň uzla teda v tejto sieti neposkytuje informáciu o tom, akú rolu daný uzol hrá. V letiskovej sieti existujú mestá, uzly, ktoré majú relatívne málo spojení, ale hrajú dôležitú úlohu (meriame pomocou centrality). Ukázalo sa, že najlepšie prepojené mestá sú sústredené v Európe a severnej Amerike, zatiaľ čo mestá, ktoré hrajú dôležitú úlohu sú rovnomerne distribuované po celom svete. Vysvetlením tejto anomálie je komunitná štruktúra letiskovej siete. Prikladom môže byť Aljaška. Mestá na Aljaške nemajú letecké spojenie so zvyškom sveta, jediné spojenie vedie cez letisko Anchorage. Odtiaľ sa priamo lieta na veľké letiská v USA. Teda všetky cesty z mnohých letísk na Aljaške vedú cez Anchorage, ktoré preto hrá dôležitú úlohu.



Obr 2. Kumulatívna distribúcia stupňov uzlov letiskovej siete. Podľa [14].

Letiskové siete skúmali aj iní autori. Napríklad Bagler [15] študoval letecké spojenia Indie. Uvažoval aj váhy na hranách siete, ktorá zodpovedala množstvu letov medzi jednotlivými letiskami. Okrem toho, že potvrdil bezškálovosť tejto siete aj jej štruktúru malého sveta, zistil, že sa zároveň jedná aj o hierarchickú sieť (1.3).

Okrem leteckých transportných sietí sa skúmali aj iné transportné siete, napríklad elektrická sieť USA, sieť diaľnic USA a podobne [16, 17, 18]. Dobře preštudovaná je hlavne elektrická sieť. Je dôležitá, pretože každý výpadok elektrickej energie spôsobuje značné ekonomické škody. Mnohí autori sa snažili pochopiť prečo v elektrickej sieti môže dôjsť k výpadkom energie, ktorý ovplyvňuje veľké oblasti [17]. Preto študovali mnohé modely sietí a kritické miesta v ich topologickej štruktúre. Simulovali ako sa distribuuje energia pri výpadku časti siete, aj to, ako postupne zlyhávajú ďalšie oblasti siete kvôli preťaženiu. Iný prístup je preštudovať štruktúru reálnej elektrickej siete [16, 19]. Ukázalo sa, že má štruktúru siete malého sveta, teda dobré prepojenie medzi uzlami kombinované s vysokou klasterizáciou, ale nie je bezškálovou sieťou. Distribúcia stupňa uzlov klesá

exponenciálne. Aj v takejto sieti vymazanie hubov môže spôsobiť jej rozpad a vysvetliť, prečo rozsiahle oblasti krajiny môžu ostať bez energie.

### 2.3 Biologické siete

Proteín proteínovú interakčnú sieť ako prví v roku 2001 modelovali Jeong a jeho spolupracovníci [20]. Skúmali sieť *S. cerevisiae*, ktorá pozostáva z 1870 proteínov (uzlov) spojených 2240 hranami (identifikovanými priamymi fyzikálnymi interakciami). Autori ukázali, že proteín proteínová interakčná sieť organizmov tvoriacich droždie je bezškálová a má komplexnú štruktúru (obr.3). To isté platí aj pre interakčnú sieť baktérie *H. pylori*, ktorú autori tiež skúmali. Proteín proteínová interakčná sieť teda pozostáva z niekoľkých proteínov (hubov), ktoré sú dôležité a veľkého množstva slabo zapojených proteínov. Preto je takáto sieť, ako ostatne všetky bezškálové siete, odolná voči náhodným poruchám, ale citlivá na cielené útoky. V organizmoch môže mať porucha zacielená na dôležité, dobre prepojené proteíny v interakčnej sieti fatálny dôsledok. Autori preto skúmali účinok vymazania uzla. Obr. 3.c ukazuje, že pravdepodobnosť, že vymazanie uzla (proteínu) zo siete proteínových interakcií má za následok smrť organizmu, koreluje s počtom interakcií (stupňom) daného uzla. Napríklad v sieti je 93 percent proteínov (uzlov), ktoré majú počet interakcií (stupeň) menší ako päť, ale len 21 percent z nich je tak dôležitých, že bez nich organizmus umiera. Na druhej strane len okolo 0.7 percent uzlov má stupeň vyšší ako pätnásť, ale ich vymazanie zo siete spôsobí smrť organizmu až v 62 percentách prípadov.

Ako ukázali ďalší autori, proteínové interakčné siete študované na iných organizmoch, sú tiež bezškálové, majú štruktúru siete malého sveta a sú modálne. Maslov a Sneppen ukázali [21], že funkčné moduly v bunke sú organizované okolo hubov. Ravasz zistila [22], že bezškálová štruktúra proteínovej siete je zároveň aj hierarchickou (1.3).

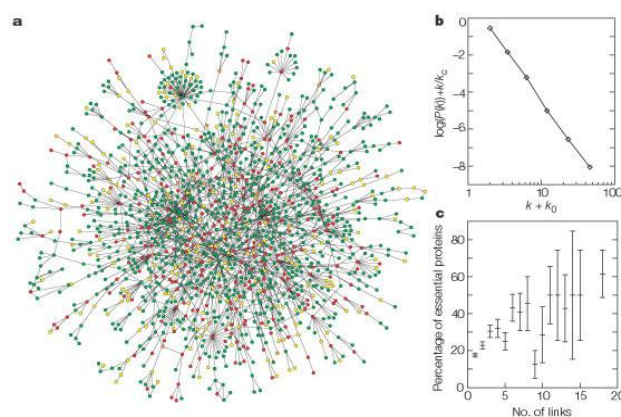
Ďalšou zaujímavou biologickou sieťou je funkčná sieť mozgu. Mozog sa skladá z neurónov, ktoré sú fyzicky prepojené cez synapsie. V mozgu však existuje aj funkčné prepojenie. Jednotlivé oblasti mozgu totiž funkčne spolupracujú na riešení kognitívnych úloh. Funkčné siete mozgu študovali Euliz a Chialvo [23] ako prví. Funkčnú sieť mozgu skonštruovali na báze nameraných *fMRI* signálov mozgu.

Pokusná osoba rieši nejakú jednoduchú úlohu (napríklad klepanie prstami jednej ruky do signálneho rytmu) a jej mozog sa scanuje. Mozog je rozdelený na malé časti, voxely (obyčajne majú okolo  $3\text{mm}^3$ ), z ktorých sa nejaký čas sníma *fMRI* signál (obr.4). Voxely sú považované za uzly siete a hranu (funkčné

spojenie) medzi dvoma uzlami položíme vtedy, ak je signál z nich zosnímaný dostatočne časovo skorelovaný a korelačný koeficient  $\nu$  prekročí nejaký dopredu definovaný prah ( $\nu > 0.5$ ). Takto vzniknuté funkčné siete mozgu majú štruktúru sietí malého sveta a sú bezškálové, aj keď ich bezškálový charakter je stále predmetom diskusie [24].

### 2.4 Siete a lingvistika

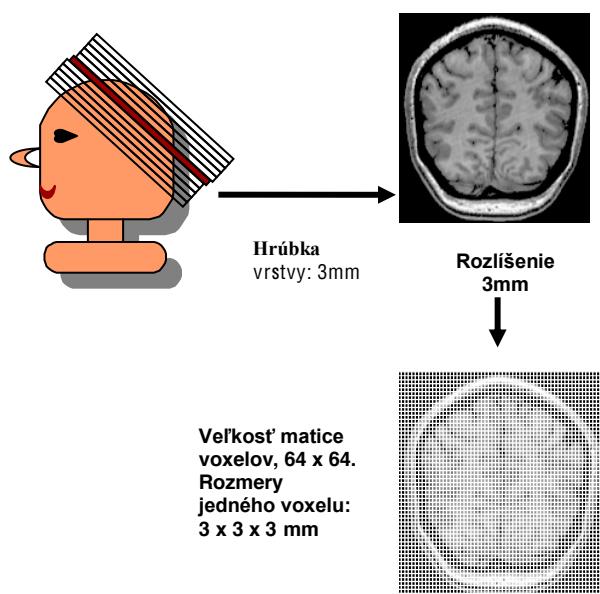
Netradičnou a zaujímavou oblasťou aplikácie teórie sietí je lingvistika. Porozumenie štruktúry jazyka by mohlo byť prospešné z mnohých hľadísk, napríklad pri konštrukcii automatických prekladových slovníkov, alebo pri navrhovaní rôznych prehľadávačov. V lingvistickej sieti možno za uzly považovať slová, poprípade slabiky a hranami sú interakcie medzi nimi, ktoré môžu byť definované rôznym spôsobom. Existujú dve hlavné skupiny lingvistických sietí, sémantické a syntaktické siete.



Obr 3. a) Proteín proteínová interakčná sieť organizmu *Sacharomyces cerevisiae*. b) Distribúcia stupňa uzlov siete. c) Podiel dôležitých proteínov stupňa  $k$ , voči stupňu (počtu spojení). Podľa [20].

V sémantických sieťach sa interakcia chápe ako významová príbuznosť [25]. Napríklad slovo je uzol a všetky slová, ktoré toto slovo opisujú vo výkladovom slovníku a aj sami sú heslami toho istého výkladového slovníka sú s ním spojené hranou. Ďalšia možnosť je tvoriť sémantickú slovnú sieť na báze asociácií. Experimentátor povie slovo a pokusné osoby asociujú k nemu významovo podobné slová, ktoré sú s ním spojené hranou. Takéto synonymické siete sú bezškálovými sieťami malého sveta (sieť na báze angličtiny a turečtiny), alebo len bezškálovými sieťami (poľština).

V syntaktických slovných sieťach slovo je uzol, a interakcia, hrana medzi uzlami vznikne vtedy, ak jednotlivé slová nájdeme v texte vedľa seba [26]. Tieto siete majú zvláštnu štruktúru, ktorá je vyjadrená v distribúcii stupňov uzlov. Siete sú bezškálové, ale charakterizované dvoma rôznymi škálovacími exponentmi. Tento jav vysvetlil Dorogovtsev s Mendesom [27] tak, že v jazyku prebiehajú zároveň dva procesy. Jednak doňho prichádzajú nové slová a pripájajú sa do kontextu iných slov preferenčne, jednak sa obohacuje význam starších slov (vznikajú hrany medzi staršími uzlami). Autori vytvorili model, ktorý kvalitatívne (ale nie kvantitatívne) kopíroval empirické dáta. Lepší model, ktorý modeluje dáta aj kvantitatívne správne sa podarilo vytvoriť nám [28] a popíšeme ho v ďalšej časti.



Obr. 4. Tvorba funkčných sietí mozgu. Signál z mozgu sa sníma po vrstvách hrúbky 3 až 6 milimetrov. Každá vrstva sa podelí na malé oblasti, pixely z ktorých snímame integrovaný *fMRI* signál. Ak je signál medzi náhodne zvolenými uzlami (pixelmi) skorelovaný, potom ich považujeme za funkčne prepojené.

### 3 Náš výskum v oblasti komplexných sietí

Výskumu komplexných sietí sa na našej katedre intenzívne venujeme už niekoľko rokov. Možno ho rozdeliť na oblasť čisto teoretickú a na oblasť aplikačnú. V teoretickej oblasti [3, 4] sme sa zamerali hlavne na hľadanie jednoduchých procesov rastu siete, ktoré vedú na bezškálovú hierarchickú štruktúru siete. Navrhli sme model rastúcej siete riadený klasterizáciou [3] a ukázali sme, že jeho zjednodušená verzia sa dá, pre isté

parametre, premapovať na známy Vásquezov model surferov blúdiacich po neznámej sieti [3] a analyticky riešiť. Ukázali sme tiež, že lokálne pripájanie hrán nových uzlov vedie k hierarchickej štruktúre siete. Navrhli sme dva minimálne modely rastúcej siete, ktoré sú bezškálové a hierarchické [4]. V prvom modeli nový uzol prináša dve hrany, ktoré sa pripoja na konce vybratej staršej hrany siete. V druhom modeli nový uzol prináša tri hrany. Jedna z nich sa pripojí na náhodne vybraný uzol a dve na konce náhodne vybranej hrany. Pripájanie nových hrán na konce náhodne vybranej staršej (skôr prinesenej) hrany v sieti zaručuje lokalitu procesov a zároveň aj preferenčné pripájanie uzlov.

Pretože tento článok je o aplikáciách teórie sietí v reálnych situáciách, zameriam sa podrobnejšie na aplikačnú časť nášho výskumu [24, 28, 29]. Týka sa hlavne aplikácií v oblasti lingvistiky a funkčných sietí mozgu.

V lingvistike sme sa zamerali na syntaktické slovné siete a to hneď na niekoľko otázok:

1. Sú štatistické vlastnosti syntaktickej slovnej siete jazykovo závislé?
2. Sú štatistické vlastnosti slovnej siete závislé od jazykového žánru v rámci toho istého jazyka?
3. Dá sa vytvoriť variácia Dorogovtsev-Mendesovho modelu jazykovej siete, ktorý by aj kvantitatívne správne popisoval reálne dáta?

Na niektoré otázky sme odpovedali v práci [28, 29]. Ukázali sme, že slovná sieť vytvorená na báze rôznych textov slovenského jazyka má podobné štatistické vlastnosti ako sieť vytvorená na báze anglického národného korpusu. Obe siete sú dokonca podobné v tom, že Dorogovtsev – Mendes (DM) model je pre ne dobrý len kvalitatívne, ale nie kvantitatívne. Predpovede škálovacích exponentov, ktoré dáva DM model sú správne len pre menej strmú lineárnu časť distribúcie stupňov uzlov, ale nie pre strmšiu časť. Tento rozdiel sa nám podarilo odstrániť zavedením nového procesu do slovnej siete [28]. Význam slova sa totiž s časom nielen obohacuje, ako predpokladali obaja autori pôvodného modelu [27], ale aj mení. Teda nielen vznikajú nové hrany medzi starými uzlami, ale dochádza aj k odpájaniu koncov hrán a ich následnému prepojeniu. Keď sme obohatili dynamiku modelu o prelinkovanie starších hrán, model dával takú distribúciu stupňa uzlov, ktorá zodpovedala nameraným dátam tak kvalitatívne ako aj kvantitatívne. Otázka žánrovej závislosti, alebo nezávislosti vlastností syntaktickej slovnej siete je ešte stále v štádiu výskumu.

Funkčné siete mozgu sme študovali v spolupráci s kolegami z University of Otago, Dunedin, Nový Zéland. Snímali sme mozog štyroch zdravých pokusných osôb, ktoré riešili jednoduchú kognitívnu úlohu (striedavé stláčanie tlačítka palcami na pokyn signálneho krátkého

tónu) [24]. Mozog sme snímali aj v periódach pokoja, keď pokusné osoby relaxovali. V oboch prípadoch sme vytvorili pre každú osobu dva typy funkčných sietí, sieť typu „task“ (zo signálu pri riešení úlohy) a „rest“ (z pokojového signálu). U všetkých štyroch osôb sme zistili, že sieť typu „task“ možno považovať za bezškálovú, ale lineárna časť distribúcie stupňov uzlov v log-log škále je príliš krátka na to, aby o tom bolo možné jednoznačne rozhodnúť, poprípade dobre zmerať exponent  $\gamma$  (1.1). Isté však je, že sieť typu „task“ sa líši od siete typu „rest“, pretože sieť typu „task“ nie je bezškálová. Ukázali sme tak, v súlade s inými autormi [24], že funkčná sieť mozgu pri plnení kognitívnej úlohy má netriviálnu štruktúru. Našou ďalšou úlohou bude urobiť viacero meraní pri plnení zložitejších kognitívnych úloh.

## 4 Záver

V tomto krátkom prehľade sme sa snažili ukázať šírku možností ako aplikovať teóriu komplexných sietí na modelovanie reálnych štruktúr. Určite sa mi nepodarilo spomenúť všetky práce, pretože ich je veľmi veľa. Zamerala som sa preto hlavne na najdôležitejšie aplikácie a spomenula som aj zopár zaujímavých kuriozít. Ak má čitateľ záujem dozvedieť sa viac, môže si prečítať kompletnejší (ale stále nie kompletný) prehľad aplikácií teórie sietí autorov Da Costa a kol. [30], ktorý obsahuje vyše 550 citácií. Z toho vidieť, že táto oblasť výskumu je veľmi aktuálna a živá.

## Literatúra

- [1] Barabási A., L., Albert R., Emergence of Scaling in Random Networks, *Science* 286, 509, 1999
- [2] Ravasz E., Barabási A., L., Hierarchical Organization in Complex Networks, *Phys. Rev. E* 57, 026112, 2003
- [3] Náther P., Markošová M., Rudolf B., Hierarchy in the growing scale free network with local rules, *Physica A* 388, 5036, 2009
- [4] Markošová M., Rudolf B., Čajági M., Náther P., Analytically solvable models of the scale free growing network with hierarchical structure, submitted to *Physica A*
- [5] Watts D. J., *Small Worlds: The Dynamics of Networks between Order and Randomness*, Princeton, Princeton University Press, Princeton, NJ, 1999
- [6] Milgram S., The Small World Problem, *Psychology Today* 2, 60, 1969
- [7] Dorogovtsev S. N., Mendes J. F. F., Evolution of Networks, *Adv. Phys.* 51, 1079, 2002
- [8] Newman M. E. J., Scientific collaboration networks I, *Phys. Rev. E* 64, 16131, 2001
- [9] Liljeros F., Edling C. H., Amaral L. A. N., Stanley E. H., Aberg Y., The web of Human Sexual Contacts, *Nature* 411, 907, 2001
- [10] Onella J. P. a kol., Analysis of large scale weighted network of one to one human communication, *New Jour. Phys.* 9, 179, 2007
- [11] Volchenkov D., Volchenkova L., Blanchard P., Epidemic spreading in a variety of scale free networks, *Phys. Rev. E* 66, 046137, 2002
- [12] Lusseau D., The emergent properties of dolphin social network, *Proc. R. Soc. Lond. B* 270, 8186, 2003
- [13] Omreod P., The medieval inquisition: scale - free networks and the suppression of heresy, *Physica A* 339, 645, 2004
- [14] Gumiera a kol., The worldwide transportation network, *Proc. Nat. Acad. Sci. USA* 102, 7794, 2005
- [15] Bagler G., Analysis of the airport network of India as a complex weighted network, *Physica A* 387, 2972, 2008
- [16] Hyashi Y., A review of recent studies of geographical scale - free networks, *IPSJ Digital Courier* 2, 155, 2006
- [17] Albert R., Albert L., Nakarado G., L., Structural vulnerability of the North American power grid, *Phys. Rev. E* 69, 2510, 2004
- [18] Dobson I. a kol., An initial model for complex dynamics in electric power system blackouts, *Proc. 34th Ann. Conf. System Science*, 710, 2001
- [19] Amaral L. A. N. a kol., Classes of small world networks, *Proc. Nat. Acad. Sci. USA* 97, 11149, 2000
- [20] Jeong H., a kol., Lethality and centrality in protein networks, *Nature* 407, 651, 2000
- [21] Maslov S., Sneppen K., Specificity and stability in topology of protein networks, *Science* 296, 910, 2002
- [22] Ravasz E., Hierarchical modularity in biological networks, in *Networks in cell biology*, Cambridge University Press, 2010
- [23] Eguiluz V. M., Chialvo D. R., Cecchi G., Baliki M., Apkarian A. V., Scale - free Brain Functional Networks, *Phys. Rev. Letters* 94, 018102, 2005
- [24] Markošová M., Franz L., Beňušková L., Topology of Brain Functional Networks: Towards the Role of Genes, Editori M. Koeppen, N. Kasabov, G. Coghill, *Advances in Neuro-Information Processing, ICONIP 2008, LNCS 5506*, Springer, Berlin/Heidelberg, 111, 2009
- [25] Motter A. E., De Moura P. S., Ying - Cheng L., Dagsputa P., Topology of the conceptual network of language, *Phys. Rev. E* 65, R 065102, 2002
- [26] Cancho R. F., Solé R., The small world of human language, *Proc. R. Soc. London B* 268, 2261, 2001
- [27] Dorogovtsev S. N., Mendes J. F. F., Language as an Evolving Web, *Proc. Royal Soc. London B*, 268, 2603, 2001
- [28] Markošová M., Network Model of Human Language, *Physica A* 387, 661, 2008
- [29] Náther P., Markošová M., Positional word web and its numerical and analytical studies, *Computing and Informatics* 30, 1287, 2011
- [30] Da Costa L. F. a kol., Analyzing and modeling real world phenomena with complex networks, *Advances in Physics* 60, 2011





# Data-mining nad systémem eLogika

Marek Menšík, Jakub Gerlich

VŠB – Technická univerzita Ostrava  
17. listopadu 15, 708 33 Ostrava  
mensikm@gmail.com, gerlich.jakub@gmail.com

## Abstrakt

Využívání LMS systémů je dnes již běžnou praxí na univerzitách, kde probíhá výuka pro velké množství studentů. Tyto systémy jsou schopny uchovávat data, která jsou na první pohled nepodstatná pro uživatele, avšak hlubším zkoumáním lze v datech najít souvislosti vedoucí k zajímavým poznatkům. V našem článku popisujeme v jednoduchosti systém eLogika. A dále data, která uchováváme v systému. Ukazujeme, jak pracujeme s druhotnými daty, které se ukládají při práci se systémem a jak tyto data je možné využít pro zkvalitňování výuky. V článku popisujeme, jakým způsobem předzpracovat data, jaké metody na která data použít a jakou informaci nám dané metody nad danými daty poskytují. Zaměřujeme se zde převážně na metody asociační a shlukovací.

## 1 eLogika

eLogika je e-learningový LMS systém, který byl původně navržen pro podporu výuky předmětu Matematická logika. Hlavním důvodem vzniku systému eLogika byla snaha usnadnit studium studentům, protože ve velkých počtech není možné, aby se pedagogové věnovali jednotlivcům. Tím usnadňuje práci jak vyučujícím, tak studentům, neboť se k informacím dostanou vždy a všude, kde mají připojení na internet. Pro pedagogy je přínos zřejmý. Vymýšlet stále nové testy je časově náročné, dále je ručně opravovat a vyhodnocovat. Studentům prezenční i kombinované formy nemusí vyhovovat styl výuky pedagoga, skladba učiva a vymezený čas, tak proto mnohdy radši sáhne po elektronické verzi, kde se problematice věnuje tehdy, když má čas.

V rámci automatizovaného zpracování výsledků studenta zde odpadá osobní invence pedagoga, která je mnoha směrech nežádoucí. Má-li být student hodnocen objektivně tak do samotného procesu vyhodnocování testu nesmí pedagog zasahovat. Toto stanovisko může vyvolávat otázky, ale pokud je test opravován strojem, tak není možné se odvolat na zaujatost vyučujícího.

Bylo tedy nutné najít způsob, jak zrychlit práci pedagogů ušetřením času přípravy testu a vyjít vstříc

studentům formou možnosti samostudia díky mnoha elektronickým (interaktivním) materiálům a otestování svých čerstvě nabytých informací.

Jak již bylo řečeno, tak eLogika mimo jiné napomáhá testování studentů. V rámci tohoto testování se do systému ukládají data, která jsou relevantní k daným testům. Na základě takovýchto dat je poté možno kvalitněji pracovat se studenty, neboť tyto data jsou velice důležitá pro tvorbu následujících dodatečných testů pro procvičení dané látky.

## 2 Databáze systému eLogika

V této kapitole je popsána část dat, která se v systému eLogika evidují. eLogika umožňuje evidenci otázek, studentů, aktivit, studijních materiálů a dalších věcí. Pro potřeby data-miningu nás budou zajímat zejména data, která se ukládají při práci studentů a vyučujících se systémem. Proto se zde rozepíšeme pouze o malé části databáze.

Z databáze systému eLogika nás budou zajímat zejména tři typy informací. První jsou informace o testových otázkách a odpovědích, které nám mohou pomoci odhalit problematické části vyučované a zkoušené látky. Druhou skupinu tvoří data o studentech a jejich výsledcích. Ty se dají použít pro případnou změnu struktury kurzu. Poslední skupinou informací jsou data o přístupech do systému, která mohou být použita k naplánování časového rozvrhu změn nebo pro rozbor toho, jak uživatelé systém používají.

eLogika umožňuje generování většího množství podobných testů. Před vytvořením samotného testu je nutné nastavit šablonu, podle které má být test vygenerován. V té se nastavuje bodové ohodnocení testu a otázky, které mají být použity. Mohou se zde vybrat konkrétní otázky, které mají být v testu použity, nebo je zde možné určit množinu otázek, ze kterých má být otázka náhodně vybrána. V databázi jsou dva druhy otázek. Jedná se o otázky, které student zodpoví vlastními slovy a o variantní otázky, u kterých student vybírá správnou odpověď (nebo odpovědi) z vygenerované nabídky.

V databázi jsou tabulky „vygenerovany\_test“, „vygenerovana\_otazka“ a „vygenerovana\_odpoved“. Záznamy v těchto tabulkách reprezentují konkrétní vygenerovanou kombinaci otázek a odpovědí, tedy jednu variantu testu. V těchto tabulkách jsou pouze reference na tabulky obsahující samotné otázky a odpovědi. Pro automatické generování testů je potřebná i tabulka „vygenerovany\_test\_uzivatel“. Zde je vazba mezi studentem a konkrétním testem, který mu byl zadán, a k záznamům v této tabulce se také přiřazují výsledky testu. Tyto tabulky jsou vlastně vazební a pouze zprostředkovávají vazbu.

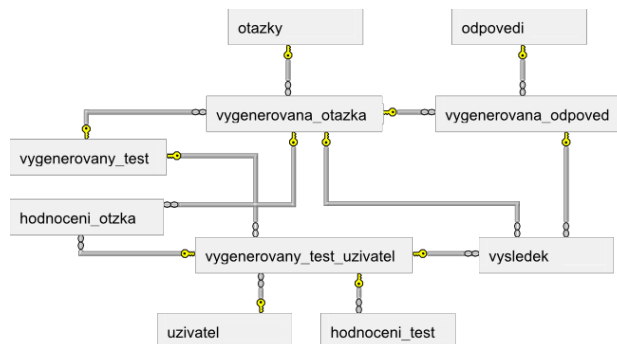
Tabulka „vygenerovany\_test“ také obsahuje referenci na tabulku „test“, která obsahuje reference na tabulky s informacemi o způsobu generování testu. Databáze obsahuje i tabulku „termin“, která také obsahuje referenci na test. Záznamy v této tabulce reprezentují naplánovaný termín konání testu a je zde tedy místo konání, začátek a konec termínu a další informace o něm.

Text a další atributy otázek a odpovědí jsou uloženy v jiných tabulkách. Tabulka „otazky“ obsahuje informace o otázkách. Je zde samotný text otázky, odhad času, který bude potřebný k jejímu vyřešení, datum a čas vytvoření otázky a referenci na uživatele, který ověřil správnost otázky a jejich odpovědi. Dále je zde specifikováno zařazení otázky (kategorie) a další atributy, které pro naše účely nejsou podstatné. Tabulka „odpovedi“ obsahuje předpřipravené odpovědi pro variantní otázky. Nejsou zde odpovědi, které formulují samotní studenti, ale pouze ty, které vytvořil autor otázky. Poskytuje nám tedy informace o variantních odpovědích. Předně je to reference na otázku, se kterou je spojena, samotný text odpovědi, předpokládaný čas, který bude student potřebovat k jejímu vyhodnocení a informace o tom, jestli je tato odpověď správná nebo ne.

Výsledky testů jsou v několika tabulkách, oddělené od testů samotných. Jsou zde tabulky „vysledek“, „hodnoceni\_test“ a „hodnoceni\_otazka“. Všechny tyto tabulky obsahují referenci na tabulku „vygenerovany\_test\_uzivatel“, takže každý záznam z nich lze přes tuto tabulku přiřadit ke studentovi a k variantě testu, kterou řešil. V tabulce „hodnoceni\_test“ je evidován čas poslední změny výsledku, maximální počet bodů, který bylo možné z testu získat a počet bodů, které získal student. Tabulka „hodnoceni\_otazka“ obsahuje referenci na tabulku „vygenerovana\_otazka“, protože přes záznam v této tabulce je možné dostat se ke kombinaci odpovědí, která byla k danému testu vygenerována. Dále je zde maximum bodů, které bylo možné za otázku získat a počet bodů, které získal student. Záznamy v tabulce „vysledek“ reprezentují odpovědi, které student označil za správné. Obsahuje tedy reference na vygenerovanou otázku a vygenerovanou odpověď.

Tabulka „otazky“ obsahuje odhad času, nutného k vyřešení otázky. Tabulky „stat\_otazka\_uzivatel“ a „stat\_odpoved\_uzivatel“ obsahují referenci na tabulku „vygenerovany\_test\_uzivatel“ a dále obsahují informaci o době řešení dané otázky, respektive počtu změn výběru odpovědi.

Přístupy do systému se evidují v tabulce „uzivatel\_prihlaseni“. Zde jsou atributy jako IP adresa, datum a čas přístupu do systému, reference na uživatele a použitý prohlížeč.



Obr. 1. Struktura části databáze systému eLogika

### 3 Datový sklad systému eLogika

Výše popsané tabulky jsou nejdůležitější pro dolování dat ze systému eLogika. Předtím, než budeme chtít použít nějaký algoritmus pro data-mining, je potřeba data v databázi předzpracovat a přesunout je do datového skladu. Problematiku datových skladů je možné najít v literatuře [1][2][3].

Databáze systému eLogika je uzpůsobena k pravidelnému přidávání i výběru dat. Je zde množství vazebních tabulek a není ojedinělé, že pro výběr vhodných záznamů se využívá spojení několika tabulek. Pro účely dolování dat není takováto struktura vhodná. Budeme potřebovat získat potřebná data co nejjednodušším způsobem. Proto jsme se rozhodli vytvořit datový sklad, do kterého zkopírujeme záznamy z vybraných tabulek.

V datovém skladu jsou data rozdělena do tabulek dimenzí a do tabulek faktů. Dimenzionální tabulky budou obsahovat data, která generuje systém, jako je například datum a čas, a data, která vytvářejí vyučující, jako jsou například text otázky, text odpovědi a bodové hodnocení testu. Budou zde tedy věci, které se nemění a které jsou společné pro větší množství záznamů.

Tabulky faktů budou obsahovat záznamy, které generují studenti a běžní uživatelé svou prací se systémem. Budou zde tedy například výsledky testů, zvolené odpovědi a záznamy přístupů. Záznamy v těchto

tabulkách budou taky obsahovat reference na tabulky dimenzí.

Jediné vazby, které zde budou, se budou nacházet mezi tabulkou faktů a tabulkou dimenzí. Pokud tedy záznamy vhodně rozdělíme, bude datový sklad umožňovat efektivnější výběr dat.

Datový sklad a původní databáze obsahují velké množství identických záznamů, ale přesto bylo nutné udělat nějaké změny, zejména s atributy, které popisovaly datum a čas. Pro reprezentaci tohoto údaje jsme vytvořili novou dimenzionální tabulku, díky které lze záznamy lépe rozdělit. Obsahuje klasické atributy, jako rok, měsíc, den, hodina, minuta a většina. Navíc ale obsahuje atributy jako například akademický rok. Přestože takového záznamy budou zabírat víc místa, než pouhá číselná reprezentace data a času, umožní rychlejší výběr vhodných záznamů. Takovýmto způsobem budeme moci rychleji najít například všechny testy, které byly napsány v pondělí. Všechny potřebné informace totiž už budou uloženy v databázi a nebude nutné vypočítávat nic v průběhu výběru z databáze.

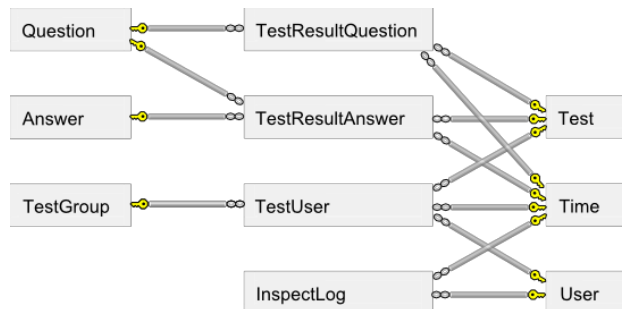
V původní databázi systému eLogika je většinou čas reprezentován jako instance třídy DateTime. Najdou se ale i takové tabulky, kde je čas uložen jako číslo. Proto jsme vytvořili vlastní třídu Time, jejíž instance se ukládají právě do nové časové tabulky. Primární klíč je zde atribut Ticks, což je číslo reprezentující datum a čas. V následujícím kódu je metoda používaná při vytváření nové instance naší třídy.

```
public void Create(DateTime datetime)
{
    Ticks = datetime.Ticks;
    DayOfMonth = datetime.Day;
    Month = datetime.Month;
    DayOfWeek = datetime.DayOfWeek;
    Hour = datetime.Hour;
    Minute = datetime.Minute;
    Second = datetime.Second;
    Year = datetime.Year;

    if (Month < 9)
        AcademicYear =
String.Format("{0}/{1}", Year - 1, Year);
    else
        AcademicYear =
String.Format("{0}/{1}", Year, Year + 1);
}
```

Velká část tabulek byla změněna jen mírně. Struktura datového skladu je naznačena na obrázku 2.

Tabulka „Test“ je dimenzionální tabulka, která obsahuje popis testu. Jsou zde tedy údaje jako maximum bodů, počet otázek a odhadovaný čas, nutný pro dokončení testu.



Obr. 2. Struktura datového skladu

Tabulky „Question“ a „Answer“ jsou také dimenzionální. Obsahují informace o otázkách, resp. odpovědích a obsahují stejné atributy, jaké jsou popsány u tabulek „otázky“ a „odpovědi“.

Tabulka „TestGroup“ je další dimenze, která reprezentuje záznamy z tabulky „termin“. Obsahuje tedy čas a místo konání testu. Dále ale obsahuje i pořadí termínu testu. Toto číslo udává, kolikrát se daný test konal. To by samozřejmě bylo možné zjistit i bez tohoto atributu, ale takto to bude jednodušší.

Poslední dvě dimenzionální tabulky jsou „Time“ a „User“. Tabulka „Time“ obsahuje atributy třídy Time, které jsou zmíněny výše. Tabulka „User“ obsahuje pouze identifikátor uživatele a jeho adresu a další atributy, které však není povinen v systému eLogika vyplnit.

Mezi tabulky faktů patří „TestResultQuestion“. Záznamy v této tabulce reprezentují konkrétní zodpovězené otázky z testů. Je zde evidováno datum a čas řešení testu (ve formě reference na záznam v tabulce „Time“), maximum bodů za otázku, získaný počet bodů, kvadrát získaného počtu bodů, čas, který student strávil řešením otázky a jeho kvadrát. Zbývající atributy jsou reference na dimenzionální tabulky „Test“ a „Question“.

Záznamy v tabulce „TestResultAnswer“ reprezentují konkrétní odpovědi uživatelů. Eviduje se zde, jestli byla odpověď zodpovězena správně, počet změn výběru odpovědi a reference na tabulky „Time“, „Test“, „Question“ a „Answer“.

Následující SQL kód ukazuje spojení pěti tabulek, které je nutné k získání všech potřebných údajů, týkajících se zodpovězené odpovědi:

```
SELECT vtu.id_vyg_test, h.id_vyg_test_uz,
vot.id_otazka, o.spravna, vod.id_odpoved,
v.id_vyg_odpoved, ht.zmena FROM
[hodnoceni_otazka] h JOIN
[vygenerovana_odpoved] vod ON
h.id_vyg_otazka=vod.id_vyg_otazka LEFT JOIN
[vysledek] v ON
h.id_vyg_otazka=v.id_vyg_otazka AND
h.id_vyg_test_uz=v.id_vyg_test_uz AND
vod.id_vyg_odpoved=v.id_vyg_odpoved JOIN
[vygenerovana_otazka] vot ON
vod.id_vyg_otazka=vot.id_vyg_otazka AND
```

```

h.id_vyg_otazka=vot.id_vyg_otazka JOIN
[vygenerovany_test_uzivatel] vtu ON
h.id_vyg_test_uz=vtu.id_vyg_test_uz JOIN
[odpovedi] o ON o.id_odpoved=vod.id_odpoved
JOIN [hodnoceni_test] ht ON
h.id_vyg_test_uz=ht.id_vyg_test_uz

```

Takto složitý kód je nutný proto, že v tabulce „výsledek“ jsou zaznamenány pouze označené variantní odpovědi. Takže se s ní nedozvíme nic o odpovědích, které uživatel u otázky neoznačil za správné. Abychom získali seznam těchto všech odpovědí, museli bychom buď provést toto spojení tabulek, nebo zkopírovat záznamy ve dvou fázích. První bychom museli projít všechny záznamy v tabulce „vygenerovana\_odpoved“ a zkopírovat je do datového skladu a pak bychom prošli záznamy z tabulky „vysledky“ a udělali bychom s nimi to samé. To je samozřejmě možné, ale znamenalo by to nutnost zkopírování celé databáze pro získání jakéhokoli výsledku po konání jediného dalšího testu. S použitím tohoto kódu můžeme využít číslování řádků a jednoduše navázat tam, kde jsme s importem skončili.

Tabulka „TestUser“ obsahuje informace o konkrétních testech, které uživatelé řešili. Je zde doba, po kterou uživatel řešil test a její kvadrát, počet bodů, které získal a jejich kvadrát a číslo pokusu o zvládnutí testu. Nejedná se zde o číslo termínu testu, ale o počet termínů testu, kterých se uživatel již zúčastnil. Kromě toho jsou zde reference na tabulky „Test“, „User“, „Time“ a „TestGroup“.

Některé atributy jsou převzaty z původní databáze a některé jsou dopočítané. Dopočítané jsou kvadráty některých údajů. Ty jsou zde kvůli zjednodušení použití některých statistických metod. Například výpočet mediánu se díky tomu velice zjednoduší.

Poslední tabulkou faktů je „InspectLog“. Tato tabulka obsahuje vše z tabulky „uzivatel\_prihlaseni“, přičemž místo identifikátoru uživatele a data a času přístupu do systému jsou zde reference na tabulky „User“ a „Time“.

## 4 Data-mining

Když máme všechna data přesunuta do datového skladu, můžeme začít dolovat informace. Můžeme se pokusit zjistit informace o testech, o studentech a o přístupech do systému. K tomuto účelu můžeme používat několik data-miningových algoritmů, jako například shlukovou analýzu, asociační pravidla a case-based reasoning (CBR).

### 4.1 CBR

Case-based reasoning můžeme použít k analýze otázek a odpovědí. Bližší informace o CBR jsou v literatuře [4][5]. CBR obecně slouží k řešení nových problémů na základě skupiny již vyřešených problémů. V našem případě můžeme zanalyzovat testové otázky a rozdělit je do skupin podle úspěšnosti. Pak můžeme například zjistit, že extrémně neúspěšné otázky byly nedostatečně probírány, nebo obsahovaly chybu. Když se příště objeví otázka s podobnou úspěšností, nalezení příčiny by mělo být jednodušší.

Můžeme si určit, jaká úspěšnost je při řešení testové otázky přípustná a jaká už může značit problém a podle takovýchto parametrů můžeme vyhledávat otázky v databázi. Obdobný postup můžeme zopakovat u testových odpovědí.

Pro CBR bychom použili tabulky „TestResultAnswer“ pro analýzu odpovědí a „TestResultQuestion“ pro analýzu otázek. Vstupy pro analýzu jsou počet záznamů s otázkou, resp. s odpovědí (kolikrát se vyskytla v testech) a procento správných zodpovězení (úspěšnost). Parametry pro rozhodnutí o vybrání otázky jsou akademický rok, hraniční počet záznamů a hraniční úspěšnost. Termínem hraniční počet záznamů je myšlen minimální počet použití otázky/odpovědi a termínem hraniční úspěšnost je myšleno minimální nebo maximální procento správných zodpovězení, nutné pro vybrání otázky/odpovědi.

Výsledkem této analýzy je seznam otázek (resp. odpovědí), jejich úspěšnost a kolikrát byly použity. Mohli bychom takto vyhledat otázky s příliš vysokou, nebo naopak s příliš nízkou úspěšností. Tak bychom mohli najít otázky, které jsou špatně formulované, nebo témata, na které by měl být při výuce kladen větší důraz. Na druhou stranu můžeme zjistit, že některé otázky jsou až příliš jednoduché.

### 4.2 Asociační pravidla

Asociační pravidla se hodí k predikcím výsledků testů, jelikož se pokoušejí najít skryté vztahy. Informace o těchto metodách jsou v literatuře [1][2]. Zde využijeme Apriori algoritmu [1], který vyhledává frekventované kombinace. Právě díky jejich nalezení můžeme vytvořit asociační pravidla. Když máme nalezeny kombinace, můžeme vybrat takové, které mají dostatečný počet výskytů (podporu). U těch můžeme zjistit jejich podporu a spolehlivost:

$$\begin{aligned}
 \text{podpora}(A \Rightarrow B) &= P(A \cup B) \\
 \text{spolehlivost}(A \Rightarrow B) &= P(B|A) = \frac{|A \cup B|}{|A|}
 \end{aligned}$$

$A$  a  $B$  jsou hodnoty atributů. Podpora pravidla  $A \Rightarrow B$  je určena tím, jaké procento záznamů obsahuje atribut  $A$  i

atribut  $B$ . Spolehlivost pravidla  $A \Rightarrow B$  je určena podmíněnou pravděpodobností. Je to tedy procento záznamů obsahujících atribut  $A$ , které obsahují i atribut  $B$ .

**Příklad:** Mějme databázi 500 studentů a jejich výsledků z různých předmětů. Pokud budeme tyto záznamy zkoumat, můžeme zjistit, že 100 studentů mělo výslednou známku jedna z matematiky i z logiky (takováto kombinace má podporu 20%). Pak už stačí zjistit, jakou spolehlivost budou mít obě varianty, tedy jestli dobrý výsledek z matematiky může znamenat dobrý výsledek z logiky, nebo jestli naopak dobrý výsledek z logiky může znamenat dobrý výsledek z matematiky. Nebo můžeme zjistit, že přestože se tato kombinace vyskytuje dostatečně často, spolehlivost je nízká a není tam tedy žádný vztah.

Data vhodná pro tyto metody se nacházejí v tabulkách „TestUser“ a „TestResultQuestion“. Tyto tabulky obsahují referenci na termín testu v „TestGroup“, takže víme čas konání testu a jeho pořadí, tedy kolikrát už měli nějakí studenti možnost stejný test řešit. To nám může pomoci s predikcemi, protože je možné, že výsledky prvního termínu nebudou stejné, jako výsledky například pátého. Vstupem pro tyto algoritmy je doba trvání testu a získané body. Volitelnými parametry můžou být akademický rok, den v týdnu, hodina konání testu a další časové údaje, případně jejich kombinace. Dalšími volitelnými vstupními parametry jsou pořadí testu, identifikátor testu a pokus o zvládnutí testu.

Tímto způsobem můžeme zjistit, jestli může být vztah mezi získanými body (nebo dobou trvání testu) a některými volitelnými parametry. Testy jsou organizovány v různé dny v týdnu a různé časy, takže s dostatečným množstvím dat je možné ptát se, jestli může být výsledek ovlivněn dnem v týdnu nebo hodinou konání testu.

Taky můžeme zkusit zjistit, jestli s rostoucím počtem pokusů o zvládnutí testu roste úspěšnost. Pokud jsou studenti, kteří se o zvládnutí testu pokoušeli třikrát a vícekrát, můžeme zkoumat, jestli se jejich výsledky postupně zlepšovaly.

### 4.3 Shluková analýza

Shluková analýza slouží k rozdělení záznamů do skupin, kde prvky z jedné skupiny jsou si podobné (viz literatura [1][2]). Tyto algoritmy porovnávají objekty podle zvolených atributů a podle nastavení toho, jaký význam má být na daný atribut kladen.

Většinou je nutné atributy upravit podle jejich významu, nebo použít správnou kombinaci. Například nemůžeme jako jeden atribut použít počet bodů za test a jako druhý pořadí pokusu. Body, které se pohybují

v řádech desítek, by neumožnily prosazení vlivu čísla pokusu, které má hodnotu maximálně 3.

Následně se vypočítá rozdílnost (vzdálenost) objektů  $a$  a  $b$  funkcí  $d(a, b)$ . Tato funkce by měla splňovat určité vlastnosti. Její výsledek je nezáporný ( $d(a, b) \geq 0$ ), vzdálenost dvou identických objektů je nulová ( $d(a, a) = 0$ ), vzdálenost dvou objektů není závislá na směru měření ( $d(a, b) = d(b, a)$ ) a měla by platit trojúhelníková nerovnost.

Pro tyto účely můžeme použít několik druhů počítání vzdálenosti. Nejčastěji se vzdálenost počítá následujícími způsoby:

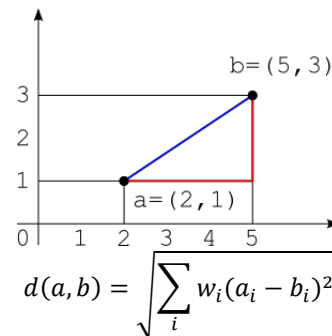
Euklidovská vzdálenost:

$$d(a, b) = \sqrt{\sum_i (a_i - b_i)^2}$$

Manhattanská vzdálenost:

$$d(a, b) = \sum_i |a_i - b_i|$$

Vážená euklidovskou vzdálenost, kde  $w_i$  je váha  $i$ -tého atributu.



$$d(a, b) = \sqrt{\sum_i w_i (a_i - b_i)^2}$$

Obr. 3. Ukázka rozdílu mezi euklidovským (modrá) a manhattanským (červená) způsobem měření vzdálenosti

Euklidovská vzdálenost objektů  $a$  a  $b$  z obrázku 3 je

$$d(a, b) = \sqrt{(2 - 5)^2 + (1 - 3)^2} = 3,61$$

Manhattanská vzdálenost těchto dvou objektů je

$$d(a, b) = |2 - 5| + |1 - 3| = 5$$

Z těchto způsobů měření vzdálenosti budeme používat zejména váženou euklidovskou vzdálenost, protože nám umožňuje nastavit významnost jednotlivých atributů a shluky by mohly být přirozenější. Bude však záležet na datech a nastavení vah atributů.

Existuje několik metod shlukování. Tyto metody vytvářejí shluky s různou časovou náročností a různými způsoby, což znamená, že se výsledné shluky mohou lišit. V tomto článku zatím nebudeme řešit, která metoda bude v praxi použita. Bude vhodné vyzkoušet víc metod, podle toho, jak budou vypadat data.

Jednou metodou je shlukování podle střední hodnoty (*k*-means clustering) [1]. Tato metoda rozdělí objekty do *k* skupin tak, aby byla vzdálenost objektu od střední hodnoty skupiny minimální. Jiná varianta je neurčovat skupinu střední hodnoty, ale podle objektu, který se nachází v její blízkosti. Díky této úpravě nám extrémně vzdálený bod moc nenaruší shluky. Tyto metody přiřadí každý objekt do právě jednoho shluku.

Další metodou je hierarchické shlukování [1]. To může fungovat dvěma směry. Buď zdola nahoru (na začátku tvoří každý objekt jednu skupinu a postupně se tvoří větší skupiny, až po skupinu tvořenou všemi objekty) nebo shora dolů (Na začátku máme všechny objekty v jediné skupině a tu rozdělujeme na menší disjunktní skupiny). S touto metodou nám vznikne množina úrovní shlukování, tedy jakási hierarchie.

Mřížková metoda (Grid-based method) umožňuje zrychlení výpočtu tím, že rozdělí objekty do konečného počtu buněk (uspořádaných do mřížky) a dále již pracuje pouze s nimi. Tím se sníží počet objektů, se kterými je třeba pracovat. Pro tuto analýzu můžeme použít tabulku „TestUser“, protože obsahuje výsledky testů a umožní nám rozdělit podle nich testy nebo studenty. Další vhodnou tabulkou by byla tabulka „TestResultQuestion“, která naopak obsahuje výsledky a další informace o otázkách a může nám pomoci rozdělit otázky. U obou tabulek by byla vstupem doba řešení a získané body. Dalšími volitelnými parametry by byly ostatní informace, které jsou dostupné, jako je identifikátor testu, akademický rok, předmět, úspěšnost a další.

Touto analýzou můžeme zjistit, že některé testy nebo otázky jsou časově náročnější, než se předpokládalo. Nebo můžeme rozdělit studenty podle úspěšnosti v testech. S pomocí takovýchto údajů pak můžeme případně změnit strukturu předmětu nebo testů.

Další tabulkou, kterou můžeme použít je „InspectLog“. Vstupem by byl internetový prohlížeč, IP adresa, identifikátor uživatele a datum a čas přístupu. Z této tabulky se můžeme pokusit udělat několik druhů shluků, podle toho, co se budeme snažit zjistit. Přístupy můžeme shlukovat podle toho, jaké prohlížeče uživatelé používají. To může pomoci optimalizovat systém pro prohlížeče, které jeho uživatelé používají nejvíce. Když zjistíme, že existuje doba, kdy systém nebývá používán žádným uživatelem, můžeme na tento čas naplánovat aplikování úprav systému. A v neposlední řadě můžeme zjistit, jaká část uživatelů se k systému připojuje ze školy,

případně ze školních kolejí, a kolik se jich připojuje z vnější sítě.

## 5 Analýza otázek a odpovědí

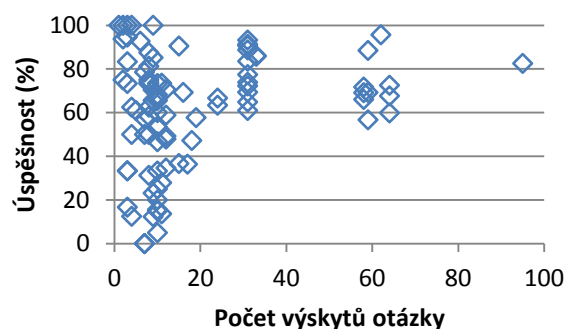
Tabulky „TestResultQuestion“ a „TestResultAnswer“ jsme použili pro analýzu testových otázek a odpovědí. Mezi těmito tabulkami jsou drobné rozdíly, proto je popíšeme odděleně.

„TestResultQuestion“ obsahuje všechny výskyty otázky v testech studentů. Vstupní parametry jsou seznam akademických roků, ze kterých máme záznamy, hraniční úspěšnost a hraniční počet výskytů. Algoritmus vyhledá otázky se správným počtem výskytů ve vybraných akademických letech. Poté spočítá průměrnou úspěšnost (z poměrů mezi získaným a maximálním počtem bodů) a vybere takové otázky, které mají správnou úspěšnost. Algoritmus pro analýzu tabulky „TestResultAnswer“ má stejný průběh, ale počítá poměr mezi správným výběrem odpovědi a celkovým počtem výskytu odpovědi.

Tento algoritmus využívá LINQ. První spočítá úspěšnost a počet výskytů a následně z otázek vybere odpovídající záznamy. Část jeho kódu, která řeší výběr otázek je následující:

```
(FROM p IN db.TestResultQuestion WHERE
p.Time.AcademicYear==year GROUP BY
p.Question INTO g SELECT new TestResultCBR
{ Question=g.Key, Count=g.Count(),
Correct=Math.Round(((g.Sum(p => p.Points) *
1.0) / (1.0 * g.Sum(p => p.Points_max))),4)
}).ToList()
```

Takto jsme zkusili zanalyzovat poslední tři roky, které jsou evidovány v databázi systému eLogika.



Obr. 4. Graf zobrazující otázky podle počtu výskytů a úspěšnosti

Výsledky byly překvapivě dobré. Hledali jsme otázky s úspěšností pod 10% a s minimálně 10 výskyty. Těmto podmínkám odpovídala pouze jedna otázka. Tato otázka

se týkala Hilbertova kalkulu a správně ji zodpověděl pouze jeden student z 20. Jsou zde dvě možnosti: studenti buď neporozuměli tématu, nebo se dostatečně nepřipravovali. Je možné, že by toto téma mělo být vyučováno důkladněji.

Na druhou stranu je v databázi 13 otázek, které mají úspěšnost přes 90% s přinejmenším 10 výskyty. Výsledky analýzy odpovědí jsou podobné výsledkům analýzy otázek. V databázi je 5 odpovědí s minimálně 10 výskyty, které mají úspěšnost pod 10%. Tři z nich jsou spojeny s otázkou na Hilbertův kalkul, kterou jsme našli.

Výsledky pro odpovědi s přinejmenším 10 výskyty a s úspěšností nad 90% jsou zajímavější. Je zde 30 odpovědí, které splňují tyto vlastnosti. Je však zajímavé, že jsou mezi nimi pouze tři otázky, které jsou spojeny s některou z otázek s největší úspěšností.

Na obrázku 4 je graf zobrazující všechny otázky v databázi systému eLogika podle toho, jakou zatím měly úspěšnost a jak moc byly používány. Vidíme zde, že zde není žádná otázka, která by svou úspěšností výrazně vybočovala z ostatních. Je zde však několik otázek, které mají výrazně větší počet výskytů, než ostatní. To je způsobeno tím, že v nějakých testech byly zadány přímo a nebyly generovány náhodně.

Toto by mohlo znamenat, že zde není žádná příliš lehká otázka. To, že zde není žádná otázka s velkým počtem úspěšných nebo naopak neúspěšných odpovědí, je dobré znamení. Hlubší zkoumání výsledků by mohlo pomoci dvěma způsoby. Mohlo by to být užitečné pro lepší formulování otázek a odpovědí a mohlo by to pomoci změnit materiály nebo strukturu kurzu.

## 6 Závěr

Nasazením systému eLogika došlo k urychlení práce jak pedagogů, tak studentů. Studenti jsou lépe připraveni na zkoušky a již v průběhu semestru jsou pomocí cvičných testů připravováni na budoucí průběh zkoušky. Nestane se tedy, že studenti jdou na zkoušku ve stavu, kdy neví, co očekávat. Toto působí velice dobře na psychiku a tedy i na lepší úspěšnost u testu. Výsledky z data-miningových metod jsou použity pro generování cvičných testů a tedy je možné studenty taktéž jmenovitě cíleně trénovat v kategoriích příkladů, které jsou pro ně či pro skupinu obtížné.

## Poděkování

Tento výzkum byl podpořen grantovou akademií České republiky, project No. 401/10/0792 “Temporal aspects of knowledge and information”, projekt CZ.1.07/2.2.00/28.0209 “Elektronické opory a e-learning pro obory výpočtového a konstrukčního charakteru“ a

grantovou agenturou VŠB-TU Ostrava, projekt SP2013/207 „Využití umělé inteligence při získávání znalostí o procesech, jejich modelování a dolování“.

## Literatura

- [1] J. Han and M. Kamber: *Data mining: concepts and techniques*. 2nd ed. Morgan Kaufmann Publishers, San Francisco, 2006.
- [2] M. Berry: *Data mining techniques*. 2nd ed. Oxford University, Indianapolis, 2004.
- [3] V. Rainardi: *Building a data warehouse with examples in SQL Server: processes, suitability and applications*. CA: Apress, Berkeley, 2008
- [4] R. Bergmann: *Experience management: foundations, development methodology, and Internet-based applications*. Springer, New York, 2002.
- [5] EDITORS, M. Antonia: *Case-based reasoning: processes, suitability and applications*. N.Y: Nova Science publishers, Hauppauge, 2009.





## Možnosti rozsáhlých aplikací ambientní inteligence

Peter Mikulecký

Fakulta informatiky a managementu, Univerzita Hradec Králové  
Rokitanského 62, 500 03 Hradec Králové  
peter.mikulecky@uhk.cz

### Abstrakt

Ambientní inteligence má poměrně dobře rozpracované přístupy, architektury a technologie pro využití ve vnitřních prostorách, v bytech či domech. Zajímavé možnosti pro její uplatnění se jeví být aplikace na báze rozsáhlých bezdrátových sensorových sítí ve volném prostoru (lesy, povodí řek, pohoří, apod.). Tyto sítě lze poměrně dobře modelovat pomocí mutiagentových architektur a budovat nad nimi složité systémy s ambientní inteligencí, které by monitorování dané oblasti mohly využívat např. pro zajištění bezpečnosti osob v daném prostoru se vyskytujících. V příspěvku se těmito možnostmi budeme zabývat jednak na úrovni přehledu stávající situace v daném oboru a jednak úvahami o návrhu architektur pro některé zajímavé "outdoor" aplikace ambientní inteligence.

### 1 Úvod

Oblast **ambientní inteligence** (angl. Ambient Intelligence, což lze volně přeložit jako všudypřítomná inteligence nebo i inteligence prostředí), je v období posledních více než 10 let předmětem intenzivního výzkumu zejména v zemích Evropské unie, jejíž rámcový program výzkumu považuje tuto oblast za jednu ze stěžejných pro další rozvoj znalostní společnosti. Představa o ambientní inteligenci vychází z původní vize Marka Weisera [36] o počítačích pro 21. století. Tato vize předpokládala posun ve využívání informačních technologií od obvyklých osobních počítačů k různým typům informačních zařízení, které budou integrovány do běžného životního prostředí lidí a budou jim tak napomáhat při řešení běžných i méně běžných životních či pracovních situací. V souvislosti s tím se rovněž předpokládá potlačení dnes převažujících technologických aspektů výpočetních a informačních technologií do pozadí, přičemž do popředí zájmu se přesunou inteligentní rozhraní konkrétních uživatelských prostředí. Základním dokumentem Evropské komise, kterým byl celý výzkum v této oblasti iniciován, je [14], podrobnější informace k celé problematice lze nalézt např. v [4], [9], nebo [19].

Hlavním účelem ambientní inteligence je naplnit poslání tzv. **pomáhajících technologií**, tedy technologií, které jsou určeny zejména k tomu, aby byly nápomocné lidem při jejich běžných každodenních aktivitách, ale i při některých specifických činnostech. Jak už jsme výše uvedli, představa o použití ambientní inteligence vede k představě o inteligentních prostředích, které vzniknou z velice obvyklých a každodenně využívaných prostředí, v nichž se běžně pohybujeme a působíme, jejich obohacením o vyspělé „všudypřítomné“ informační a komunikační technologie schopné inteligentní podpory lidských aktivit vyvíjených v tomto prostředí. Výzkumy v daném směru obvykle vyúsťují do návrhů prototypů inteligentních domů, domácností, kanceláří, učeben, domovů pro seniory, apod.

Všechny tyto aplikace mají jednu vlastnost společnou, a sice to, že jsou omezeny na uzavřený prostor – např. inteligentní kancelář jako jedna místnost, vybavená řadou senzorů, kamer, atd., aktivity se odehrávají tak říkajíc mezi čtyřmi zdmi. Složitější architekturu představuje např. realizace inteligentního domu, kde příslušné sensorové vybavení může přesáhnout představu o omezení zdmi domu a může přesahovat do zahrady či prostoru kolem vchodu do domu. To vše jsou však přeci jenom prostorově omezené, nijak rozsáhlé aplikace ve srovnání se současnými stále více využívanými aplikacemi, založenými na rozsáhlých bezdrátových sensorových sítích (viz např. [1], [2], [3], [10] nebo [32]).

Prudký vývoj technologie senzorů, trvalé snižování jejich energetické náročnosti, snižování jejich ceny, miniaturizace, přitom ale zvyšování jejich univerzálnosti a také zvyšování jejich inteligence přináší stále nové možnosti jejich nasazení v různých typech aplikací. Z nich mají zvláštní význam **rozsáhlé bezdrátové sensorové sítě** (large-scale wireless sensor networks) senzorů různého typu, přičemž v těchto rozsáhlých sítích jednotlivé senzory komunikují bezdrátově jak mezi sebou, tak i s dalšími subjekty, což mohou být třeba nějaká centra, do kterých sensorová síť svá pozorování odesílá, nebo i uživatelé, kteří se v dosahu sítě pohybují a mohou se sítí komunikovat např. pomocí chytrého mobilu.

Naší myšlenkou, kterou chceme v tomto příspěvku více rozebrat je, že sítě tohoto typu by mohly být doplněny o takové vlastnosti, které by z prostředí v dosahu bezdrátové sensorové sítě vytvořily prostředí s ambientní inteligencí. Několik prvních úvah v tomto směru jsme již uvedli v [20], [21], [23] či [25]. V tomto příspěvku nejprve rozebereme některé stávající aplikace rozsáhlých sensorových sítí pro environmentální monitorování v otevřeném prostoru a poté popíšeme několik možných multi-agentových architektur, které by nad těmito sítěmi mohly být vytvořené a byly by základem pro inteligentní „outdoor“ prostředí na bázi přístupů rozvíjených v oblasti ambientní inteligence.

## 2 Ambientní inteligence

Jak jsme již uvedli v [19], **ambientní inteligence** (Ambient Intelligence) je vizí informační společnosti budoucnosti, ve které je největší důraz kladen na

- uživatelskou přívětivost;
- efektivní a distribuovanou podporu služeb;
- posílení možností efektivního využívání relevantních zdrojů ze strany uživatelů;
- podporu interaktivní činnosti.

V rámci této vize se předpokládá, že lidé budou působit v prostředí inteligentních a intuitivně využitelných rozhraní k procesorům, umístěným do všech typů objektů v daném prostředí, tedy do zcela běžných předmětů, které nás obklopují a které tvoří prostředí, v němž působíme. Takovéto prostředí by pak mělo být schopné rozpoznat přítomnost konkrétních lidí a reagovat na jejich přítomnost nevyrušujícím a často neviditelným způsobem, který bude obvykle plně integrován do dané situace. Tato reakce prostředí by měla zásadně působit ve prospěch konkrétního uživatele, měla by mu pomoci řešit zcela běžné ale i zřídka se vyskytující obtížné problémy. Proto se těmto technologiím říká i pomáhající technologie.

V práci [22] jsme zdůraznili, že na prostředí, protkané technologiemi ambientní inteligence, můžeme přirozeně nahlížet jako na multi-agentový systém, tedy distribuovaný, decentralizovaný systém autonomních, heterogenních a navzájem interagujících entit (agentů). Síla multi-agentových systémů spočívá v možnosti propojit různé druhy agentů, disponující rozdílnými možnostmi a schopnostmi zpracovávat nejrůznější druhy vstupů.

V multi-agentové architektuře inteligentního prostředí lze uvažovat skupiny agentů, které spolupracují a/nebo sdílejí nějaký účel, např.:

- agent poskytující informace spolupracující s agentem asociovaném s displejem,

- osobní agenty skupiny jednotlivých uživatelů,
- množina agentů reprezentující osobní zařízení (počítač) jednoho uživatele,
- apod.

V rámci určené skupiny mohou agenty komunikovat na základě role, kterou v této skupině zastávají. Lze očekávat, že takto vymezený pojem skupiny pomůže uživateli získat lepší možnost ovládat interaktivní prostředí, nebo v něm explicitně vidět kombinace zařízení, služeb či relevantního obsahu. Úlohy a aktivity právě zapojeného uživatele jsou pak každá reprezentovány jinou skupinou agentů.

## 3 Rozsáhlé sensorové sítě

Technologie ambientní inteligence jsou určené pro reálný fyzický svět, proto je u nich velice důležité použití senzorů. Vývoj v oblasti senzorů je za poslední léta velice dramatický a pokrok ve vývoji senzorů lze vidět takřka každodenně. Vývoj směřuje ke stále více sofistikovaným senzorům, přičemž jejich nároky na spotřebu energie jsou stále menší. Sensory jsou ta zásadní technologie, která propojuje reálný svět s dostupnou výpočetní silou počítačů, pracujících v pozadí těchto aplikací. Sensory jsou obvykle celkem malých rozměrů, lze je proto vestavět v podstatě do každé aplikace ambientní inteligence.

Existuje celá řada prací zaměřených na technické možnosti aplikací sensorových sítí. Kniha [17] uvádí řadu výsledků, zaměřených na citlivost senzorů a sensorových sítí na kontext, ve kterém fungují. Hlavní myšlenkou kontextové citlivosti je, že má-li senzor více informací o kontextu, ve kterém má fungovat, může přizpůsobovat své chování a funkcionalitu pouze těm situacím, kdy to je zapotřebí a kdy je jeho fungování adekvátní okolnostem. Toto může být důležité např. pro energetickou náročnost senzorů. Další zajímavé výsledky přinášejí např. práce [12], [13], nebo [28], dobré přehledové studie jsou např. [1], [2], [3] nebo [32]. Jsou to většinou přehledy stávajících výsledků v oblasti bezdrátových senzorů a jejich aplikacích v bezdrátových sensorových sítích. Vynikající přehled environmentálně orientovaných bezdrátových sensorových sítí přináší Corke et al [10]. Přináší přehled experimentů s bezdrátovými sensorovými sítěmi pro environmentální a zemědělské aplikace, pohlíží rovněž kriticky na možnosti dalšího využití sensorových sítí pro účely environmentálního monitoringu.

Hezký, i když už trošku starší přehled aplikací bezdrátových sensorových sítí přináší Arampatzis, Lygeros a Manesis [3]. Popisují mnoho aplikací, které využívají bezdrátové senzory, nebo bezdrátové sensorové sítě a rozdělují je do pěti vhodně vymezených kategorií,

mezi nimiž jsou např. vojenské aplikace, nebo environmentální monitorování.

#### 4 Monitorování vnějších prostorů

Celá řada aplikací rozsáhlých bezdrátových sensorových sítí je zaměřena na environmentální monitorování, tedy na online snímání různých veličin, které senzory sítě v monitorovaném přírodním prostředí zaznamenávají, sbírají a v klasickém využití senzorů posílají tyto veličiny k dalšímu zpracování nebo využití, obvykle do nějakého centra či dispečinku.

Potenciál environmentálního monitorování nespočívá pouze v tom, že lze sledovat nové přírodní jevy, které se doposud nedaly dobře pozorovat, ale i v tom, že lze nepřetržitě sledovat vnější prostředí a posílat data příslušným dispečinkům či jiným souvisejícím organizacím, čímž je umožněn základ pro vytvoření systému včasného varování pro různé typy přírodních katastrof a pro jejich včasné zvládnutí. Zde právě vidíme značné možnosti pro uplatnění přístupů ambientní inteligence.

Z mnoha v současnosti známých a zajímavých rozsáhlých environmentální aplikací sensorových sítí se zmíníme o dvou projektech, *Live E!* [18] a *FieldServer* [29]. Projekt *Live E!* [18] je projekt vyvíjený otevřeným výzkumným konsorciem (WIDE consortium a IPv6 Promotion Council Japan), zaměřený na sdílení digitálních informací o monitorovaném životním prostředí, zejména o počasí. Jeho účelem je vytvořit globální infrastrukturu s cílem sběru a distribuce environmentálních informací. Platforma projektu *Live E!* sestává z řady levných meteorologických stanic instalovaných v rámci 13 zemí, mezi jinými v Japonsku, Francii, Thajsku a na Taiwanu.

Každá z těchto meteorologických stanic je provozována jinou organizací (obvykle univerzitou, nebo která má ve své náplni sensorovou technologii) a zahrnuje v sobě několik senzorů, které měří obvyklé meteorologické veličiny, jako je teplota, velikost srážek, vlhkost vzduchu, rychlost větru, atd. Data jsou pak přenášena cestou internetu do centrální databáze, kde jsou uložena a poskytována k dalším analýzám.

Každá organizace, provozující meteorologickou stanicí, má k dispozici server, který sbírá data naměřená senzory v reálném čase prostřednictvím internetu. Všechny tyto servery si pak uložená data mohou vyměňovat pomocí vzájemného internetového propojení. *Live E!* platforma slouží také k testování nových sensorových aplikací. Začátkem roku 2009 bylo v projektu nasazeno a provozováno 106 meteorologických stanic v rámci 13 zemí a 11 serverů provozovaných 9 organizacemi. Nejhustější senzory vybavená oblast v rámci projektu je *Kurashiki-city*

v Okayamě v Japonsku: 25 senzorů na ploše 25km×25km.

Projekt *FieldServer* [29] je zaměřen na vývoj a síťovou aplikaci tzv. *Field serverů*, což jsou bezdrátové sensorové sítě rozšiřující monitorování environmentálních faktorů tím, že jejich sensorové uzly jsou umístěny v přesně určených místech v otevřeném prostoru (např. v polích, či povodí řeky, apod.), čímž se snižují nutné náklady na jejich instalaci a je umožněn sběr snímaných dat v reálném čase. Znamé jsou aplikace *Field serverů* v Japonsku, kde byly nasazeny na farmách. Výsledkem jejich monitorování jsou snímky pro bezpečnostní hlídky získané v reálném čase, jakož i environmentální data pro farmáře. Získané snímky mají vysoké rozlišení a mohou být využity např. pro řešení různých konkrétních situací v monitorovaném prostředí, které mohou vyžadovat zásah do prostředí. Byla vyvinuta celá řada typů *Field serverů*.

Další zajímavou aplikační oblastí pro využití bezdrátových sensorových sítí je management vodních toků a jejich povodí. Podle Yanga [38], účelem plánu managementu vodního toku je maximalizovat profit různých uživatelů vodního toku při současně redukci možných konfliktů, které se mezi nimi mohou vyskytnout. Management vodního toku lze velice efektivně modelovat multi-agentovým systémem [24], existuje však pouze několik studií, které věnují pozornost i zvládnutí katastrofických situací [6].

Jak upozorňuje Ruiz-Garcia a další [32], relativně nízká cena bezdrátových sensorových sítí již dnes umožňuje vytvářet husté sítě sensorových uzlů, které mohou adekvátně reprezentovat a snímat proměnlivou povahu vnějšího prostředí. Mohou tedy i zprostředkovávat různé informace o možných, v přírodě se vyskytujících rizicích, např. mohou varovat zemědělce, že mohou přijít noční mrazíky a ohrožovat úrodu. Bylo vytvořeno a nasazeno několik bezdrátových sensorových sítí jako systémů pro protipožární dohled např. v lesích. Dokážou měřit teplotu a vlhkost, jakož i rozeznat přítomnost kouře s následným vysíláním poplašné informace [16]. Takovéto senzory umí snímat různé statické i dynamické proměnné, jako je vlhkost či nasycení půdy, typ hořícího paliva, sklon monitorovaného terénu, směr a rychlost větru, apod. Dokážou rovněž určit směr postupu a možný vývoj čela hořícího území.

Veliká pozornost je v poslední době věnována možnostem monitorování záplavových území, zejména pro případy bleskových povodní. Povodně, jak je známé, způsobují velké materiální škody i ztráty na lidských životech, zejména v rozvojových zemích, ale ne jenom tam. Soustředěné úsilí je věnováno vývoji systémů na bázi bezdrátových sensorových sítí, které se zaměřují na minimalizaci možných škod pomocí včasné předpovědi

blížící se katastrofy [27]. Mimo již zmíněných povodní sem patří i opačný problém, a to dlouhá období sucha, která jsou pro celou řadu zemí podobným problémem, jako jinde záplavy. Zajímavá řešení systémů pro předcházení těmto problémům lze najít v [5], [6], [16] nebo [27].

Cardell-Oliver a další [7] navrhli novou reaktivní senzorovou síť, která měří nasycení zeminy vodou a využívá přitom adaptivní přístup: pokud prší, měří nasycení půdy velice často, např. každých 10 minut, zatímco v době, kdy neprší, měří nasycení třeba jenom jednou za den. To je umožněno oddělením senzorů, monitorujících dešť od těch, které měří nasycení půdy. Tyto senzorové uzly sdílejí měřené informace, ale je jasné, že pokud neprší, dochází k minimalizaci času potřebného k posílání, čtení nebo zpracování informací, což znamená i podstatné snížení energetické náročnosti.

## 5 Možnosti pro ambientní inteligenci

Má-li nějaký systém být schopen podporovat lidské aktivity ve venkovním prostředí (outdoor), musí být schopen identifikovat geografickou polohu podporované osoby jako důležitou kontextuální informaci, kterou je posléze možné využít v různých důležitých scénářích, jako je např. asistence při hledání směru úniku z ohroženého prostoru, kontextově závislé doporučení výrobku (kontextová reklama), nebo včasné varování konkrétní osoby před možným ohrožením. Tradičně se zde využívá služeb GPS, což je zvlášť vhodné v otevřeném prostoru, ale ne příliš použitelné uvnitř budov.

K specifikám podpory činnosti osob působících ve venkovním prostředí patří zejména to, že tyto osoby se obvykle docela intenzivně věnují různým aktivitám (sportu, práci) a nemusí si uvědomovat blížící se nebezpečí, jemuž mohou být neočekávaně vystaveni. Doposud existuje pouze malý počet pokusů o vytvoření řešení na bázi ambientní inteligenci, které by byly schopné případné nebezpečné situace v otevřeném prostředí řešit, např. postupem podle následujícího scénáře:

*Mladá sportovkyně Anastázie se právě nachází v hornatém a zalesněném prostředí, kde se věnuje svým oblíbeným koníčkům (turistice, skalolezení, cyklistice, apod.). Protože ví, že předpověď počasí není úplně příznivá a je hlášena i možnost přivalových dešťů, prostřednictvím svého chytrého telefonu se přihlásila do webové služby včasného varování, která je v daném prostoru aktivní. Tato služba, založená na docela husté síti bezdrátově propojených inteligentních senzorů, trvale vyhodnocuje data získaná ze sítě senzorů (teplota, vlhkost, nasycení půdy, síla a směr větru, atd.), eviduje pohyb osob (přihlášených i ostatních) v daném prostoru a*

*V případě náhlé nebezpečné situace nebo při vysoké pravděpodobnosti jejího vzniku začne vysílat varovné signály a poskytovat podporu zejména osobám, které jsou do takovéto sítě přihlášené, což je také Anastázie. V případě nebezpečí se tak Anastázie prostřednictvím svého chytrého mobilu dozví podrobnosti o tom, jaké nebezpečí se blíží, odkud ji co může hrozit a jaké má možnosti pro rychlý ústup z ohroženého prostoru.*

*Poskytuje-li takováto síť možnost akustických varovných signálů, může další osoby v ohroženém prostoru varovat pomocí nich, ale přinejmenším bude moci případným zasahujícím složkám záchranných sborů sdělit, že v daném prostoru se ještě vyskytovali či vyskytují další osoby a zhruba určit jejich lokalizaci.*

V odborné literatuře lze nalézt pouze několik článků orientovaných na řešení právě popsané našim scénářem. Některá z těchto řešení jsou navíc značně zúžená, např. pouze na včasné varování, že hrozí nebezpečí, bez možnosti poskytnout bližší podrobnosti, jako je třeba usměrnění ohrožené osoby k nejbezpečnějšímu opuštění ohroženého prostoru.

V práci [34] je popsán systém na bázi ambientní inteligence, který má za úkol ochránit děti od potenciálně nebezpečných situací v městském prostředí. Jedná se zde pravděpodobně o první systém svého druhu, který je zaměřen na všudypřítomné monitorování konkrétních dětí, přičemž dynamicky detekuje možné nebezpečné situace cestou z a do školy a rychle poskytuje podporu dětem, popřípadě i jejich rodičům s cílem minimalizovat až vyloučit některá možná nebezpečí. K detekci případného nebezpečí systém používá mnoho kontextových informací ze sledovaného prostředí v okolí monitorovaných dětí. Aby toto vše bylo možné, autoři své řešení zakládají na dvou předpokladech:

- (1) množství senzorů, RFID štítků a dalších zařízení pro sběr informací je rozmístěno v okolí školy a cestou do a ze školy;
- (2) monitorované dítě musí mít sebou (oblečené, nebo nést v tašce) jistá zařízení, která mohou sbírat a vysílat kontextová data z okolí, ve kterém se dítě pohybuje a kde jsou všechny výše uvedené senzory a další zařízení instalována.

Mnoho odborných článků je zaměřených na různá řešení pro podporu turizmu, zejména pro podporu kontextové navigace turistů na jejich trasách. Obvyklým využitím přístupem zde je (viz např. [28] či [29]) vývoj inteligentních agentů, kteří kolektivně upřesňují kontext, v němž se uživatel nachází, a poté v závislosti na daném kontextu shromažďují, vybírají a přeposílají tyto informace na uživatelův osobní digitální komunikátor (např. chytrý mobil či tablet). Tato řešení však obvykle mají za cíl pouze cestovatele navigovat na jeho trase, popřípadě mu kontextově poskytovat informace o

různých místech, která by ho mohla zajímat (tzv. POI – Points of Interest), ale nemají možnost vyhodnotit situaci, v níž se uživatel právě nachází, jako potenciálně nebezpečnou a uživatele před ní varovat, resp. poskytnout užitečné informace o tom, jak z této situace uniknout.

V současnosti využívané bezdrátové senzorové sítě, zaměřené na environmentální monitorování [13], jsou dnes už dostatečně zralé k tomu, aby se staly základem pro mnohem komplexnější podporu různých venkovních lidských aktivit.

Jak upozorňuje Efstratiou [11], bezdrátové senzorové sítě jsou stále víc využívány jako řešení pro rozsáhlé sledovací a monitorovací aplikace, nicméně tyto aplikace jsou obvykle navrženy pro jedno úzce specializované použití. Data, která jsou přitom shromážděna, jsou obvykle dostupné pouze jedné autoritě, nejčastěji majiteli této senzorové sítě a nejsou tedy k dispozici pro řadu dalších možností jejich využití, což je značně neefektivní. Obvyklá situace u tohoto typu aplikací je, že bezdrátové senzorové sítě sbírají data z prostředí, které monitorují (i když na vysoce sofistikované úrovni), a poté nasbíraná data posílá do nějakého centrálního dispečinku, kde na základě získaných dat příslušní specialisté, odpovídající za monitorovanou oblast, dělají nutná rozhodnutí. Ale je celkem jasné, že od takovýchto senzorových sítí bychom měli chtít mnohem víc.

Například, pokud by byl sestaven inteligentní systém, pracující nad senzorovou sítí, který by měl za úkol ne toliko přeposílat sesbíraná data někam dál, ale který by byl schopný i tato data dále zpracovat a vyhodnotit v nich některé zajímavé souvislosti, typické pro rizikové situace v monitorovaném prostředí, pak by takový systém byl zcela jistě schopen vyprodukovat alespoň některá včasná varování lidí, kteří se v monitorovaném prostředí pohybují, co by je mohlo ochránit před možným ohrožením.

Jak uvádí Efstratiou [11], vize další generace senzorových sítí je vizí světa, v němž infrastruktura senzorů bude sdíleným zdrojem schopným dynamického přizpůsobení či přeprogramování za účelem podpory celé řady různých aplikací současně. Navíc, takovéto víceúčelové senzorové sítě (které budou třeba ve vlastnictví různých institucí) bude možné sdílet a kombinovat, aby tak bylo možné získat mnohem podrobnější obraz okolního světa.

My tento názor sdílíme a podporujeme. Proto je potřeba se orientovat na výzkum, který by sice vycházel z myšlenky sdílení a kombinování senzorových sítí, ale obohatil by je o návrh a vývoj systému na báze ambientní inteligence, jenž by pracoval nad rozsáhlou bezdrátovou senzorovou sítí, pracující v potenciálně nebezpečném venkovním prostředí a plnil by (kromě jiných) především následující úkoly:

- Monitorování obvyklých hydro-meteorologických parametrů prostředí (tlak, teplota, vlhkost, nasycení půdy vodou, rychlost a směr větru, atd.);
- Monitorování příznaků možných nebezpečných situací (různé seizmoakustické signály, kouř, voda na neobvyklých místech, apod.);
- Monitorování výskytu a pohybu osob v daném prostoru, případně i monitorování výskytu a pohybu zvířete (zejména větší kusy, nebo početnější stáda);
- Vyhodnocování dat poskytovaných senzorovou sítí, hledání důležitých souvislostí mezi nimi (např. obrazců typických pro některé situace) a identifikace možného nebezpečí;
- Identifikace těch osob v daném prostoru, které mohou být vystaveny ohrožení;
- Soustředěná snaha o kontakt a komunikaci s osobami, vystavenými potenciálnímu nebezpečí, nejspíš využitím jejich mobilních komunikačních zařízení, včetně poskytování intenzivní informační a znalostní podpory těchto osob s cílem umožnit jim bezpečný únik z ohroženého prostředí (zde případně může systém rozhodnout také o přivolání příslušných záchranných složek).

Mezi zajímavé příklady inteligentních systémů, které se blíží naplnění výše uvedených myšlenek, patří systémy popsané např. v [35] nebo [37].

Bezdrátová senzorová síť popsaná v práci [37] slouží k podpoře skupiny geofyziků, zkoumajících potenciálně stále fungující vulkán. Tato senzorová síť má mimo jiné za cíl ochránit pomocí včasného varování pracovníky, kteří se pohybují v nebezpečné blízkosti vulkánu tím, že neustále monitoruje činnost vulkánu a po vyhodnocení seizmo-akustických signálů z vulkánů může danou situaci vyhodnotit jako potenciálně nebezpečnou. Zde jsou využity malé, lehké a na spotřebu energie nenáročné senzory, které jsou produktem současného rozvoje této oblasti.

V popsaném případě je vlastně celá myšlenka založena na doplnění senzorové sítě, původně určené k pouhému sběru geofyzikálních údajů, o nové funkcionality, což bylo umožněno také i díky stále univerzálnější a sofistikovanější konstrukci senzorů.

V práci [35] je pak popsána jiná aplikace, která není sice úplně aplikací ve vnějším prostředí, ale tomuto použití se velice přibližuje. Jedná se o rozsáhlou inteligentní senzorovou síť pro monitorování situace v uhelním dole v Číně, což navozuje mimo představy podzemního dolu také představu monitorování rozsáhlého pozemního dolu. Mnoho podrobností o čínském národním projektu zaměřeném na rozsáhlé senzorové sítě bohužel není k dispozici, dobrým zdrojem informací je zde ale např. [28].

## 6 Závěr

Výzkumný projekt *SMEW: Smart Environments at Workplaces*, který jsme řešili do roku 2012, byl orientován na inteligentní podporu pracovišť využitím přístupů ambientní inteligence. Zde jsme se zaměřili na pracovní aktivity, probíhající ve vnitřním prostředí a předmětem zkoumání byly inteligentní kanceláře a další pracoviště v budovách. Tato problematika úzce souvisí i s výzkumem inteligentních budov a mnohé výsledky tohoto výzkumu také využívá. Celá řada výsledků se jeví být přenositelná i do vnějšího prostředí, tento přenos však nemůže být nikterak jednoduchý vzhledem k celé radě nových faktorů, které na vnější pracovní prostředí působí či mohou vplývat. V každém případě je možnost rozšířit již existující přístupy ambientní inteligence z vnitřního na vnější prostředí velice lákavá a zajímavá. Nosnou myšlenkou se zde jeví využití velice bouřlivě se rozvíjející oblasti rozsáhlých bezdrátových sensorových sítí jako základu pro monitorování příslušného vnějšího prostředí, nad nímž by bylo možné a vhodné vybudovat inteligentní systém s využitím multi-agentové architektury, implementující již vyzkoušené přístupy ambientní inteligence. Účelem takovéto architektury by bylo zpracovávat data, získaná od rozsáhlé sensorové sítě, která by příslušné prostředí monitorovala, udržovat vhodný model možných stavů tohoto prostředí a identifikovat na základě zpracovaných dat každou odchylku stavu prostředí od očekávaného stavu. Na základě vyhodnocení možného stupně nebezpečí v případě, že se monitorovaný stav výrazně liší od očekávaného stavu prostředí, by systém mohl autonomně a proaktivně rozhodnout o akci, kterou je nutné podniknout, tedy např. včas varovat osoby v ohroženém prostředí se vyskytující a případně jim poskytnout vhodnou podporu, např. prostřednictvím jejich chytrých mobilů.

Je zřejmé, že mezi základní výzkumné cíle v daném směru by měly patřit:

- Návrh multi-agentové architektury nad environmentální bezdrátovou sensorovou sítí monitorující vhodně zvolené vnější prostředí s potenciálem výskytu rizikových situací;
- Vytvoření adekvátního mechanismu rozpoznávání potenciálně nebezpečné situace v prostředí na základě analýzy dat, snímaných sensorovou sítí z monitorovaného prostředí;
- Použití vhodného mechanismu pro více-kriteriální rozhodování multi-agentového systému o tom, jaké akce budou pro danou situaci nejvhodnější;
- Vytvoření nejvhodnějšího komunikačního kanálu pro komunikaci systému s potenciálně ohroženými

osobami nacházejícími se v monitorovaném prostředí;

- Experimentální implementace navržené multi-agentové architektury nad reálnou a vhodně vybranou bezdrátovou sensorovou sítí (např. horské nebo lesní prostředí, povodí řeky, apod.) a její vyhodnocení.

Přístupy ambientní inteligence již v mnohém prokázaly svoji užitečnost a funkcionalitu v implementacích ve vnitřním prostředí (inteligentní domy, domácnosti, kanceláře). Jsme přesvědčeni, že jejich implementace ve vnějším prostředí má vysoký aplikační potenciál a že mohou výrazně přispět mimo jiné i k mnohem lepšímu využití v současnosti mimořádně rychle se rozvíjející technologii bezdrátových sensorových sítí. Pokud by takovéto aplikace přispěly k tomu, aby se lidé cítili i v potenciálně nebezpečném vnějším prostředí jistě a bezpečně, jejich účelu by bylo dosaženo.

## Poděkování

Tento příspěvek vznikl za podpory projektu „Informační, kognitivní a interdisciplinární podpora výzkumu – INKOV, registrační číslo CZ.1.07/2.3.00/20.0001, který je spolufinancován Evropským sociálním fondem a státním rozpočtem České republiky.

## Literatura

- [1] I. F. Akyildiz, T. Melodia, K. R. Chowdury: Wireless multimedia sensor networks: A survey. *IEEE Wireless Communications*, December 2007, 32-39.
- [2] I. F. Akyildiz, W. Su, Z. Sankarasubramaniam, E. Cayirci: A survey on sensor networks. *IEEE Communications Magazine* 40 (2002) 102-114.
- [3] Th. Arampatzis, J. Lygeros, and S. Manesis: A survey of applications of wireless sensors and wireless sensor networks. In *Proc. of the 13<sup>th</sup> Mediterranean Conference on Control and Automation*, Limassol, Cyprus, 2005: 719-724.
- [4] J. C. Augusto, H. Nakashima, H. Aghajan, Ambient intelligence and smart environments: A state of the art. In: *Handbook of Ambient Intelligence and Smart Environments*, Springer Science+Business Media, 2010: 3-31.
- [5] E. Basha, D. Rus: Design of early warning flood detection systems for developing countries. In *Proceedings of the Conference on Informations*

and *Communication Technologies and Development*, Bangalore, India, 2007.

- [6] L. Brouwers, K. Hansson, H. Verhagen, M. Boman: Agent models of catastrophic events. In *Proceedings of Modelling Autonomous Agents in a Multi-Agent World, 10<sup>th</sup> European Workshop on Multi Agent Systems*, Annecy, 2001.
- [7] R. Cardell-Oliver, K. Smettem, M. Kranz, K. Mayer: A reactive soil moisture sensor network: Design and field evaluation. *Int. Journal of Distributed Sensor Networks* 1 (2005) 149-162.
- [8] M. Ceriotti et al.: Is there light at the ends of the tunnel? Wireless sensor networks for adaptive lighting in road tunnels. In: *Proceedings of the 10th International Conference on Information Processing in Sensor Networks, IPSN 2011* (X. D. Koutsoukos, K. Langendoen, G. J. Pottie, V. Raghunathan, eds.), New York, IEEE, 2011: 187-198.
- [9] D.J.Cook, J.C.Augusto, V.R.Jakkula: Ambient intelligence: Technologies, applications, and opportunities. *Pervasive and Mobile Computing* 5 (2009) 277-298.
- [10] P. Corke, T. Wark, R. Jurdak, W. Hu, P. Valencia, D. Moore: Environmental wireless sensor networks. In *Proceedings of the IEEE*, Vol. 98, No. 11 (2010) 1903-1917.
- [11] Ch. Efstratiou: Challenges in supporting federation of sensor networks. In *NSF/FIRE Workshop on Federating Computing Resources*, 2010.
- [12] E. Elnahrawy and B. Nath: Context-aware sensors. In *Proc. 1<sup>st</sup> European Workshop on Wireless Sensor Networks*, 2004: 77-93.
- [13] J.K. Hart, K. Martinez.: Environmental sensor networks: A revolution in the earth system science?, *Earth-Science Reviews* 78, (2006) 177 – 191.
- [14] ISTAG: *Scenarios for Ambient Intelligence in 2010*. European Commission, Luxembourg, 2001.
- [15] M. Li, Y. Liu: Underground coal mine monitoring with wireless sensor networks. *ACM Trans. on Sensor Networks* 5, 2, (March 2009) 10-39.
- [16] J. Lloret, M. Garcia, D. Bri, S. Sendra, A wireless sensor network deployment for rural and forest fire detection and verification. *Sensors* 9 (2009) 8722-8747.
- [17] S. Loke: *Context-Aware Pervasive Systems*. Boca Raton: Auerbach Publications, 2007.
- [18] S. Matsuura, et al.: LiveE! Project: Establishment of infrastructure sharing environmental information. In *2007 International Symposium on Applications and the Internet Workshops (SAINTW'07)*, IEEE, 2007: 67.
- [19] P. Mikulecký.: Ambientní inteligence po deseti letech. In: *Kognice a umělý život XII* (J. Kelemen, P. Nahodil, eds.), Vyd. Agentura Action M, Praha, 2012: 167-175.
- [20] P. Mikulecký: Large scale ambient intelligence – Possibilities for environmental applications, *Ambient Intelligence Perspectives II*, Volume 5 Ambient Intelligence and Smart Environments, IOS Press, 2010: 3-10.
- [21] P. Mikulecký: Remarks on ubiquitous intelligent supportive spaces. In: *Recent Advances in Applied Mathematics and Computational and Information Sciences, Proc. of the 15th American Conference on Applied Mathematics and Proc. of the International Conference on Comp. and Information Sciences* (Eds. Jegdic, K., Simeonov, P., Zafiris, V.), Vols. I and II, WSEAS Press, Athens, 2009: 523-528.
- [22] P. Mikulecký, K. Olševičová, K.: Inteligentní prostředí jako multiagentový systém. In: *Kognice a umělý život VII* (Kelemen, J., Kvasnička, V., Pospíchal, J., eds.), Slezská univerzita v Opavě, 2007: 245-248.
- [23] P. Mikulecký, R. Cimler, K. Olševičová.: Outdoor large-scale ambient intelligence. In: *Proc. of the 19th International Business Information Management Association Conference IBIMA 2012* (K.S.Soliman, ed.), Vyd. IBIMA, 2012: 1833-1835.
- [24] Mikulecký, P., Ponce, D., Olševičová, K., Haviger, J., Bodnárová, A.: Application of multi-agent systems and ambient intelligence approaches in water management. In: *Global Changes and Water Resources, Proc. of the 13th World Water Congress* (Varis, O. Tortajada, C., Chevallier, P. et al., eds.), Association VERSeau Développement, Montpellier, (publication on a CD), 2008.

- [25] P. Mikulecký, P. Tučník: Ambient intelligence for outdoor activities support: Possibilities for large-scale wireless sensor networks applications. In: *ICSNC 2012, The Seventh International Conference on Systems and Networks Communications* (E. Borcoci, S. Manoharan, eds.), Vyd. IARIA, 2012: 213-217.
- [26] P. Misra, S. Kanhere, D. Ostry, S. Jha: Safety assurance and rescue communication systems in high-stress environments: A mining case study. *IEEE Communications Magazine*, April 2010, 66-73.
- [27] P. Morreale, Q. Feng, P. Croft: A green wireless sensor network for environmental monitoring and risk identification. *Int. J. of Sensor Networks* 10, 1/2 (2011) 73 – 82.
- [28] L.M.Ni, Y. Liu, Y. Zhu: China's national research project on wireless sensor networks. *IEEE Wireless Communication*, December 2007, 78-83.
- [29] S. Ninomiya, T. Kiura, A. Yamakawa, T. Fukatsu, K. Tanaka, H. Meng, M. Hirafuji: Seamless integration of sensor network and legacy weather databases by MetBroker. In *2007 International Symposium on Applications and the Internet Workshops (SAINTW'07)*, IEEE, 2007: 68.
- [30] L.M.L. Oliveira, J.J.P.C. Rodrigues: Wireless sensor networks: A survey on environmental monitoring. *Journal of Communications* 6, 2 (2011): 143-151.
- [31] Z. Pei, Z. Deng, S. Xu, X. Xu: Anchor-free localization method for mobile targets in coal mine wireless sensor networks. *Sensors* 9 (2009) 2836-2850.
- [32] L. Ruiz-Garcia, L. Lunadei, P. Barreiro, J.I. Robla: A review of wireless sensor technologies and applications in agriculture and food industry: State of the art and current trends. *Sensors* 9 (2009) 4728-4750.
- [33] V. Seal, A. Raha, S. Maity, et al.: A simple flood forecasting scheme using wireless sensor networks. *International Journal of Adhoc, Sensor & Ubiquitous Computing (IJASUC)* Vol.3, No.1 (2012) 45-60.
- [34] K. Takata, Y. Shina, H. Komuro, M. Tanaka, M. Ide, and J. Ma: Designing a Context-Aware System to Detect Dangerous Situations in School Routes for Kids Outdoor Safety Care. In *EUC 2005* (L.T. Yang et al., eds.), LNCS 3824, Berlin: Springer, 2005: 1016-1025.
- [35] X. Wang, X. Zhao, Z. Liang, M. Tan: Deploying a wireless sensor network on the coal mines. In: *Proceedings of the 2007 IEEE International Conference on Networking, Sensing and Control*, London, UK, 2007: 324-328.
- [36] M. Weiser: The computer for the 21st century. *Scientific American* 265 (1991) 94-104.
- [37] G. Werner-Allen, K. Lorincz, M. Welsh, et al.: (2006), Deploying a wireless sensor network on an active volcano. *IEEE Internet Computing* 10 (2006) 18-25.
- [38] I.-C. Yang: *Modelling Watershed Management with an Ecological Objective - A Multi-agent System Based Approach*. PhD Dissertation, University of Illinois at Urbana-Champaign, 2010.
- [39] Y. Zhang, L. Luo, J. Huo, and W. Zhu, An eco-hydrology wireless sensor demonstration network in high-altitude and Alpine environment in the Heihe river basin of China. *Wireless Sensor Network*, 4, 5 (2012) 138-146.



# Hybridní neuronové systémy pro návrh architektur autonomních agentů v oblasti umělého života

Pavel Nahodil, Jaroslav Vítků

České Vysoké Učení v Praze, Fakulta Elektrotechnická  
Katedra kybernetiky  
Technická 2, 166 27 Praha 2  
nahodil@fel.cvut.cz, vitkujar@fel.cvut.cz

## Abstrakt

V současné době existuje mnoho typů různých algoritmů a jejich implementací. Pro zvýšení efektivity výzkumu v oblasti kompozice modulárních inteligentních systémů se snažíme sjednotit komunikaci mezi různými algoritmy. Algoritmus je zde viděn jako subsystém a komunikace mezi subsystémy je definována podobně jako komunikace v umělých neuronových sítích. Libovolnou kombinací těchto subsystémů (modulů) jsme poté schopni vytvářet nové modulární systémy. Za účelem praktického ověření chování výsledných architektur jsme vytvořili nový simulátor kombinací dvou stávajících softwarových nástrojů. Práce je z oblasti umělého života, proto se zde zabýváme především možnostmi kompozice architektur autonomních agentů. Simulovat je však možné modulární systémy napříč mnoha vědními obory. Definování těchto modulárních podobně jako neuronové sítě má následující hlavní důsledky: Přístup nám umožňuje libovolně kombinovat způsoby návrhu architektur „*zdola nahoru*“ a „*shora dolů*“. Výsledná síť má poté mnohem menší komplexitu topologie oproti klasické umělé neuronové síti. Navíc, jednotná komunikace nám umožňuje zkoumat možnosti automatického návrhu těchto modulárních systémů.

## 1 Úvod

Hlavní cíl našeho výzkumu spočívá v návrhu modulárních hybridních architektur autonomních agentů. V současné době neexistuje architektura, která by byla prokazatelně lepší ve všech doménách. Z tohoto důvodu jsou často agenti navrhováni specificky pro daný úkol. Takové architektury se většinou skládají z často se opakujících pod-systémů. Výběr podsystémů a jejich propojení je vybráno uživatelem na základě jeho znalostí a zkušeností.

V současné době se výzkum dělí na stále menší a specializovanější pod-obory. Člověk tak často nestíhá sledovat vývoj ve více než jednom specifickém oboru. Je

tak stále těžší udržet si nadhled ohledně možností využití výsledků výzkumu z jednotlivých oborů a jejich kombinací. Trend je navíc takový, že tato situace se bude stále zhoršovat. Často také navíc nastává situace, že výsledky z jednotlivých oborů se překrývají, a přestože se dva různé přístupy mohou jmenovat zcela odlišně, mohou být principiálně velice podobné. Dále lze zvolit řešení problému na různých úrovních abstrakce. Například pro rozpoznávání sekvencí můžeme použít buď Markovovy řetězce [4], nebo biologicky mnohem věrohodnější přístupy inspirované neuronovými sítěmi [15,10,16,9]. Správný výběr jednotlivých pod-systémů (využívajících znalosti z různých oborů), nejen pro tvorbu autonomních agentů, se tak stává stále složitějším.

V mnoha úlohách se setkáváme s tím, že máme k dispozici nějakou metriku ukazující kvalitu dosaženého výsledku. Na základě tohoto ukazatele (*fitness funkce*) pak můžeme vyhodnotit jak je navržená architektura vhodná pro řešení daného problému. Tuto vlastnost lze výhodně využít pro automatizaci návrhu architektur autonomních agentů.

Pokud tedy známe množinu dostupných pod-systémů, úlohu a odpovídající fitness funkci, bylo naší snahou sestavit algoritmus, který bude sám zcela autonomně navrhovat nové architektury autonomních agentů přímo pro danou úlohu.

## 2 Teoretický základ

Klasické metody návrhu autonomních agentů bývají dnes děleny do tří hlavních skupin:

- **Návrh přístupem shora-dolů:** Tato metoda návrhu je založena především na velice dobré apriorní znalosti úlohy a všech podmínek. Autonomní agent je často řízen svými explicitně definovanými znalostmi. Výzkum je zde zaměřen na reprezentaci a použití znalostí ve výsledném řídicím systému, který přímo

interaguje s okolním světem. Funkční moduly agenta implementují především algoritmy pro percepci světa, učení a plánování za použití stávajících znalostí. Architektury navržené tímto klasickým způsobem jsou schopné řešit i ty nejsložitější problémy (například typu šachy), ale často se velice špatně potýkají s nepředvídanými okolnostmi. Příkladem takovéto architektury je *Belief-Desire-Intention* (BDI) [12].

- **Návrh přístupem zdola-nahoru:** Tento opačný přístup staví na jednodušších modulech implementujících především jednoduché vzory chování, například vyhýbání se překážkám, identifikování objektů nebo prohledávání prostředí. Abstraktnější funkce systému, jako například plánování, správa a reprezentace znalostí nejsou explicitně definovány. Místo toho, toto chování může vznikat *emergentně* z interakce jednodušších pod-systémů jako zcela nová, vynořující se vlastnost [6].
- **Hybridní přístup:** Hybridní architektury v sobě integrují oba předchozí přístupy, tzv. "*knowledge-based*" (symbolické) a "*behavior-based*" (sub-symbolické) dohromady. Vědci se tímto způsobem snaží využít výhody obou přístupů a potlačit jejich nevýhody. Znalostní pod-systémy implementující učení a plánování (schopné produkovat komplexní chování) jsou v takovýchto architekturách kombinovány s bloky implementujícími jednodušší chování (vzrůstá robustnost v nepředvídaných situacích). Hlavním předmětem výzkumu v tomto oboru je ve hledání metod jak efektivně kombinovat tyto symbolické a sub-symbolické dohromady [14].

## 2.1 Umělé neuronové sítě

Hlavní rozdíly mezi Umělými neuronovými sítěmi (ANN) a klasickým výpočetním modelem implementovaným počítači založenými na von-Neumannově architektuře jsou extenzivně popisovány v četné literatuře, proto následuje pouze velmi zkrácený popis. Von-Neumannův typ počítače ze svého principu realizuje deterministické výpočty založené na přesných datech. Díky této vlastnosti není proto tolerantní k chybám. Tato vlastnost je ale v reálném prostředí velice důležitá. Pokud chceme bezpečně rozlišovat mezi jednotlivými stavy (bity), musíme k tomu poskytnout dostatek energie (což mj. v důsledku způsobuje vysokou spotřebu dnešních počítačů). Další nevýhoda je v centralizovaném výpočtu, čímž je maximální propustnost dat omezena rychlostí zpracování instrukcí.

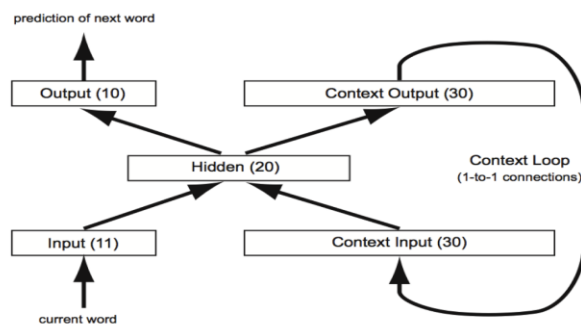
S výše popsaným porovnejme zcela odlišný výpočetní model, navržený v přírodě evolucí: „*mozek*“. Ten má především následující užitečné vlastnosti:

- Je to vysoce distribuovaná a redundantní metoda výpočtu
- data a výpočet jsou zde uložena na jednom místě
- výpočet je už z principu vysoce odolný proti chybám
- díky toleranci chyb je systém navíc vysoce energeticky úsporný.

Zdá se tedy, že vše bylo podřízeno především *vysoké energetické úspornosti*. Díky tomu se v evoluci prosadil systém odolný proti chybám, bez nutnosti zbytečného přenosu dat z místa na místo, a tedy energeticky nenáročný. Existuje však i mnoho dalších důvodů, proč nejsou umělé neuronové sítě, inspirované tímto systémem, více používané. Asi ta nejvýznamnější nevýhoda ANN je to, že pracují jako černé skříňky ("*black-box*"). Uživatel nemá potřebný vhlad do systému a tím pádem ani mnoho možností, jak přizpůsobovat jeho funkci. I tento problém se snažíme v našem přístupu řešit.

## 2.2 Modulární neuronové sítě

Jednou z možností jak zabránit neprůhlednosti funkce neuronové sítě a současně například urychlit učení je v její modularizaci. Tzv. *modulární neuronové sítě* se skládají ze dvou a více modulů, které mezi sebou komunikují předdefinovaným způsobem. Každý modul je implementován neuronovou pod-sítí.



Obr. 1. Příklad rekurentní sítě Jordanova typu pro predikci slov ve větě [13].

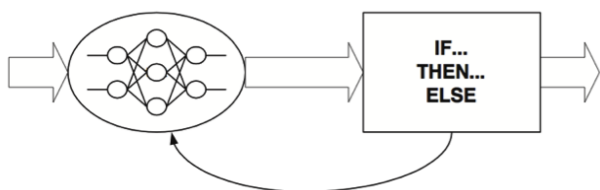
Příklad takovéto modulární rekurentní neuronové sítě je na Obr. 1. Vstupní a výstupní moduly pro predikci slov ve větě komunikují s externím světem. Modul ve skryté vrstvě reprezentuje stav sítě a je závislý na konkrétním kontextu (zde na předchozích slovech ve větě).

Je také vidět že předdefinované propojení mezi bloky velice dramaticky snižuje počet spojnic mezi neurony (oproti plně propojené rekurentní ANN). Definování takovýchto omezení na topologii sítě vede často ke značnému zjednodušení struktury výsledného systému a urychlení jeho učení.

### 2.3 Hybridní neuronové sítě

Dalším krokem ke zjednodušení systému na základě principu ANN jsou tzv. hybridní neuronové sítě. Tyto sítě kombinují přístupy návrhu shora-dolů a zdola-nahoru do jedné sítě složené z heterogenních uzlů. Hybridní síť si můžeme představit jako modulární ANN, kde jednotlivé moduly mohou být implementovány buď pod-sítí, nebo předem definovaným algoritmem. Podle [17] můžeme hybridní neuronové sítě rozdělit na tři hlavní typy:

- **Unifikované nervové architektury** - tyto systémy se spoléhají výhradně na konekcionistický přístup, přestože symbolická interpretace uzlů je také možná. Specifická apriorní znalost problému je často zabudovaná do topologie neuronové sítě. Příkladem takového systému může být síť z Obr. 1.
- **Hybridní transformační architektury** - používají transformaci symbolické a sub-symbolické reprezentace. Hlavní zpracování informace se provádí pomocí neuronové sítě. Tyto znalosti pak mohou být interpretovány na symbolické úrovni a naopak.
- **Hybridní modulární architektury** - obsahují jak symbolické, tak neuronové moduly navržené pro specifický úkol. Zde, na rozdíl od transformačních architektur, symbolické reprezentace nejsou jen počáteční nebo konečné, ale pomocí zpětné vazby se mohou přizpůsobovat novým znalostem. Schéma hybridní modulární architektury je na Obr. 2.



Obr. 2. Schéma hybridní modulární architektury. Transformace mezi symbolickou a sub-symbolickou reprezentací je přizpůsobována na základě aktuální potřeby pomocí zpětné vazby [17].

Tento článek se mimo jiné zabývá přístupem, který je schopen navrhovat modulární architektury spadající do třetího typu - tedy *hybridní modulární architektury*.

### 2.4 Neuro-evoluce

Evoluční Algoritmy (EA) jsou algoritmy pro prohledávání stavového prostoru inspirované evolucí v přírodě. V porovnání například s lokálním prohledáváním, tyto algoritmy si udržují celou populaci tzv. jedinců. Jedinec se sestává z genotypu (unikátní řešení daného problému) a fitness funkce (vyjádření kvality řešení). Pomocí přirozeného výběru (na základě hodnoty fitness funkce), křížení a mutace genotypů EA generuje nové populace jedinců s postupně kvalitnějšími jedinci. Díky své robustnosti EA hodí pro prohledávání velkých stavových prostorů, kde by klasické metody často selhaly.

Výzkum metod využívajících EA, (které jsou schopné navrhovat umělé neuronové sítě na základě daného problému) se nazývá *neuro-evoluce*. Genotyp jedince zde reprezentuje topologii neuronové sítě (a, nebo váhy mezi neurony). Fitness funkce je kvalita řešení problému, poskytnutého danou neuronovou sítí. Existuje mnoho metod, jak zakódovat neuronovou síť do genotypu jedince, přehled těch základních je například v [3].

### 2.5 Evoluční návrh hybridních systémů

Jeden z cílů našeho současného výzkumu spočívá v nalezení vhodného Evolučního Algoritmu (EA) který bude schopen efektivně navrhovat tyto hybridní modulární architektury. Hlavní rozdíl oproti klasickým neuro-evolučním algoritmům je ve faktu, že zde jednotlivé prvky (moduly) již nejsou *Single-Input Single-Output* (SISO) moduly (jako v případě neuronů v ANN), ale každý modul má obecně  $m$  vstupů a  $n$  výstupů. Jedná se tedy o *Multiple-Input Multiple-Output* (MIMO) systémy. V případě návrhu hybridních sítí s MIMO moduly bez omezení na topologii roste prohledávaný prostor všech zapojení extrémně rychle. Možnosti jak definovat omezení na topologii takovýchto systémů již byly zkoumány. Například *Karteziánské genetické programování* (CGP) [5] se zabývá evolučním návrhem logických obvodů složených z MISO modulů do dopředné topologie.

Další zajímavý příklad evolučního návrhu MISO systémů se jmenuje *Group of Adaptive Models Evolution* (GAME) prezentovaný v [7]. Algoritmus postupně (od vstupů směrem k výstupům) pomocí upraveného EA staví dopředné architektury hybridních neuronových sítí pro modelování mnohdimenzionálních dat. Výše zmíněné algoritmy se zabývají vždy jen vyhraněným tématem a používají pouze MISO moduly.

### 3 Náš přístup

Oproti výše popsaným systémům, náš přístup se zabývá návrhem systémů složených z obecnějších MIMO modulů. Důraz je kladen především na architektury autonomních agentů pohybujících se v předem neznámém prostředí. Moduly tedy implementují především učení bez učitele, případně posilované učení.

Výhody našeho přístupu oproti klasickým ANN jsou tedy především v modularizaci neuronových sítí a ve zjednodušení komplexity celého systému pomocí modulů navržených metodami shora-dolů.

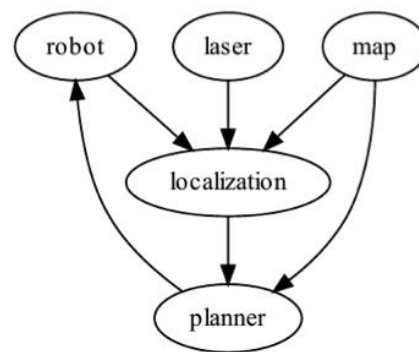
Abychom mohli efektivně navrhovat a testovat architektury autonomních agentů, některé základní požadavky na takový systém musí být splněny. Tyto základní nároky jsou popsány níže.

#### 3.1 Pře-použitelné moduly

Jedna z časově nejnáročnějších prací na výzkumu nových architektur spočívá především v implementaci jednotlivých modulů. Pokud bychom byli schopni sjednotit "podobu" jednotlivých modulů tak, aby mezi sebou šly kombinovat libovolné pod-systémy, to co by lidé naprogramovali pro svůj výzkum, by bylo dále velice snadno použitelné pro další účely. Vývoj dalšího výzkumu by tím byl značně urychlen. Požadavky na takový "framework" jsou především následující:

- standardizovaná komunikace
- nezávislost na programovacím jazyce
- jednoduchost protokolu.

Standard splňující podobné požadavky se již používá v oblasti výzkumu robotiky. Systém se jmenuje *Robotic Operating System* (ROS) [11]. Jedná se o decentralizovanou síť uzlů implementujících libovolný algoritmus. Každý uzel může být spuštěn ve svém procesu a komunikuje s ostatními uzly po síti pomocí předdefinovaných zpráv. Příklad jednoduché sítě ROS uzlů je uveden v Obr. 3. Jedna z hlavních výhod ROSu je volně dostupná knihovna dostupných uzlů. Tento systém jsme se rozhodli nadále využívat i v oblasti výzkumu umělého života.



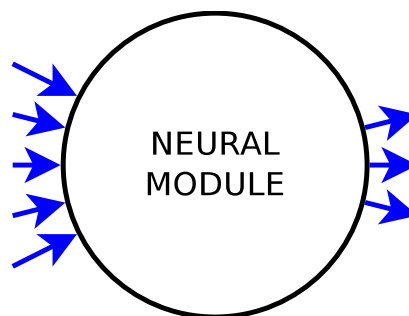
Obr. 3. Příklad modulárního systému implementovaného podle ROS standardu. Jednotlivé uzly (procesy) spolu komunikují pomocí asynchronního posílání zpráv (spojnice mezi uzly) [11].

#### 3.2 Standardizovaná komunikace

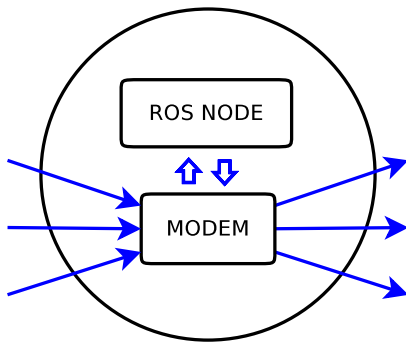
Přestože ROS určitým způsobem definuje možné typy práv použitelné pro komunikaci, náš cíl je komunikační protokol sjednotit tak, aby byly informace kompatibilní mezi obecně jakýmkoliv uzly.

Je známo, že jakákoliv funkce může být aproximována dopřednou neuronovou sítí s jednou skrytou vrstvou [2]. Z toho může být usouzeno, že obecně jakékoliv chování může být aproximováno pomocí ANN s komplexnější architekturou.

Proto jsme se rozhodli se při definování komunikačního protokolu inspirovat právě u ANN. Jednotlivé moduly, nazvané "neurální moduly" (viz. Obr. 4) tedy mezi sebou komunikují pomocí primitivního datového typu float, případně vektorů takovýchto čísel. Schéma ROS uzlu použitého jako neurální modul je vyobrazeno na Obr. 5.



Obr. 4. Koncept neurálního modulu - černá skříňka s obecným počtem vstupů a výstupů.

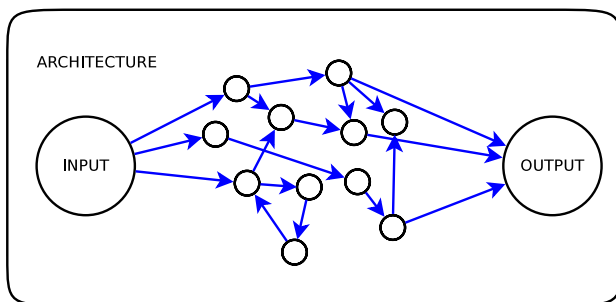


Obr. 5. Neurální modul, který zapouzdřuje libovolný uzel v ROSu. Každému ROS uzlu je přiřazen modem, který transformuje ROS zprávy do pole datových typů float a zpět.

Takto definovaná komunikace přináší mnoho výhod. Výsledné modulární systémy mohou být přímo porovnávány s modulárními neuronovými sítěmi [1]. Moduly lze mezi sebou zapojovat zcela libovolným způsobem. A nakonec: pro návrh topologie takovýchto systémů lze použít upravené metody *neuro-evoluce*. Hlavní nevýhoda takto definované komunikace je především ve velkém počtu spojů mezi moduly.

### 3.3 Hybridní návrh architektury

Jednou z hlavních výhod tohoto přístupu je také v hybridizaci samotného návrhu hybridních architektur. Pomocí definování jednotlivých modulů, které algoritmus může použít, jsme schopni určit, jak moc bude vhodná architektura, navržená metodou *top-down* (tedy jak velkou budeme mít nad návrhem kontrolu) a naopak jakou část návrhu necháme automatizovanou (tedy *bottom-up* přístup).



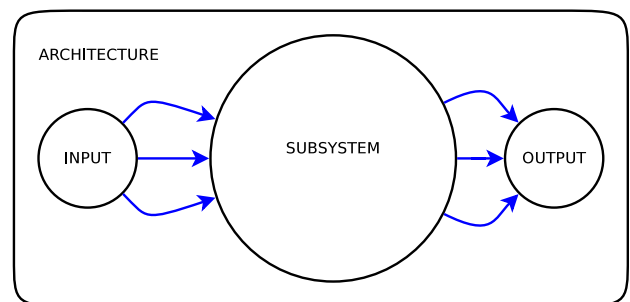
Obr. 6. Příklad hybridní architektury navržené metodou zdola-nahoru.

Na schématu v Obr. 6 vidíme extrémní případ, kdy uživatel nechal celou architekturu navrhnout metodou

*bottom-up*: složitější chování je produkováno interakcí mezi jednotlivými moduly.

Architektura na Obr. 6 je složená z nejjednodušších modulů, zajišťujících jednoduché chování (například neurony s danou přechodovou funkcí). Složitější chování zde vznikne emergentně z interakce jednotlivých modulů.

Oproti tomu, následující spodní obrázek Obr. 7 reprezentuje architekturu navrženou kompletně jenom pomocí *top-down* přístupu. Architektura je složena z pouze jednoho modulu, předem známé struktury. Pomocí poskytnutí modulů různého stupně komplexity jsme tak schopni plynule kombinovat oba tyto přístupy dohromady.



Obr. 7. Příklad hybridní architektury navržené zcela pouze metodou shora-dolů.

Architektura se skládá prakticky jen z jednoho velkého modulu, který implementuje vše potřebné. Modul je přímo připojen na vstupy a výstupy architektury.

### 3.4 Simulátor hybridních modulárních systémů

Za účelem jednoduchého vytváření a testování nových architektur autonomních agentů byl námi vytvořen simulátor hybridních modulárních systémů nazvaný *Nengoros*. Simulátor vznikl přidáním podpory ROSu do simulátoru velkých neuronových sítí s názvem *Nengo* (dostupný na adrese [www.nengo.ca](http://www.nengo.ca))

Simulátor *Nengo* podporuje simulaci velkých modulárních neuronových sítí, především pak *neuronové sítě třetí generace* [8].

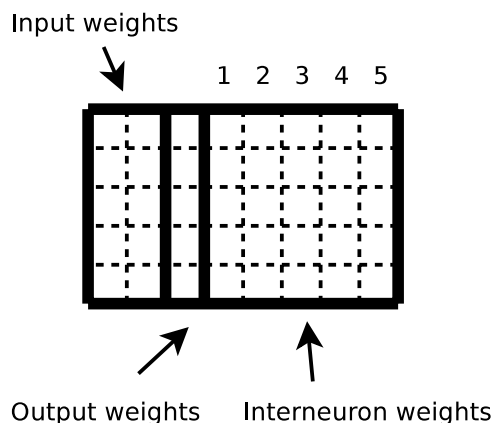
Mezi další výhody simulátoru patří možnost rychlého prototypování (ať již pomocí GUI nebo skriptování) a platformově nezávislá implementace v Javě.

## 4 Modifikovaná reprezentace hybridních sítí pro neuro-evoluci

Pro základní testování nového typu modulárních systémů a jejich simulátoru jsme navrhli jednoduchou reprezentaci těchto systémů vhodnou pro neuro-evoluční algoritmy. Jedná se o modifikaci přímé reprezentace matice vah neuronové sítě v matici.

### 4.1 Reprezentace neuronové sítě pomocí matice vah

Jedna z nejjednodušších metod jak reprezentovat topologii ANN ve formátu použitelném pro EA je pomocí matice vah, tento princip je vidět na Obr. 8.



Obr. 8. Příklad možnosti reprezentace plně rekurentní neuronové sítě zakódované pomocí matice vah. Síť má dva vstupy, jeden výstup a pět skrytých neuronů.

Hodnoty čísel v matici reprezentují váhy mezi jednotlivými neurony. Matice je schopna reprezentovat rekurentní neuronové sítě.

Pro použití v EA je matice převedena na genotyp seřazením prvků do 1-dimenzionálního vektoru. Na ten jsou pak aplikovány evoluční operátory. Délka genomu  $l$  je dána podle rovnice:

$$l = \alpha^2 + (IN + OUT) k,$$

kde  $k$  je počet neuronů,  $IN$  je počet vstupů a  $OUT$  počet výstupů.

Výhoda této reprezentace je v její jednoduchosti, naopak hlavní nevýhodou je fakt, že počet vah, které je nutné reprezentovat, roste exponenciálně s počtem neuronů.

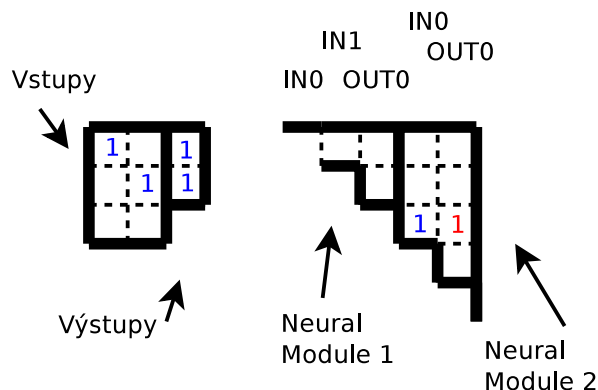
### 4.2 Modifikovaná reprezentace pro hybridní neuronové sítě

Výše popsané kódování je schopno reprezentovat váhy mezi jednotlivými neurony - SISO uzly. Pro reprezentaci obecně MIMO modulů je však nevhodná. Bylo třeba stávající kódování upravit tak, aby bylo schopno reprezentovat různé váhy pro moduly s obecně  $m$  vstupy a  $n$  výstupy.

V nejjednodušším případě by každý sloupec (řádka) mohl reprezentovat jeden konkrétní vstup/výstup modulu a použít kódování bez dalších změn. Ze schématu na Obr. 9 je však vidět, že při zohlednění orientace vstupů/výstupů stačí uchovat pouze horní trojúhelníkovou matici bez diagonály. Na tomto obrázku vidíme, že síť má (zleva doprava) tři vstupy, jeden výstup, jeden modul se dvěma vstupy a jedním výstupem a jeden klasický neuron.

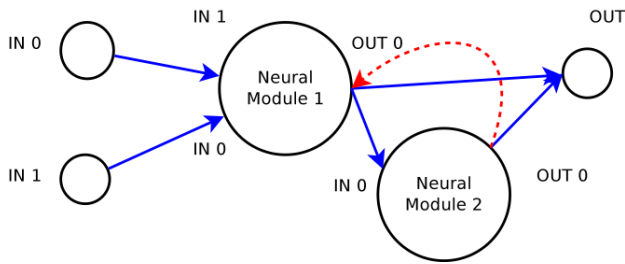
Jelikož je informace o orientaci hran implicitní, dekódování výše popsané reprezentace pak má tři hlavní kroky:

- Vytvořit neorientované spoje mezi uzly
- Validní spoje (spojující vstup a výstup) zachovat
- Neplatné spojnice (spojující dva vstupy nebo dva výstupy) smazat.



Obr. 9. Modifikovaná matice vah pro reprezentaci hybridních neuronových sítí obsahujících MIMO moduly.

Následující Obr. 10 zobrazuje výsledné zapojení hybridní architektury složené z jednoho neuronu a jednoho MISO modulu.



Obr. 10. Hybridní neuronová síť dekodovaná z matice z předchozího obrázku. Červeně (čárkovaně) je vyznačen neplatný spoj, všechny ostatní spoje jsou platné.

Výhoda tohoto kódování je především v tom, že pro MIMO systémy nepotřebujeme tak velkou matici vah. Nicméně následující rovnice ukazují, že i tak délka genomu  $l$  roste asymptoticky exponenciálně s počtem prvků v síti:

$$l = \frac{\alpha^2 - \alpha}{2} + IN \sum_k in_k + OUT \sum_k out_k$$

$$\alpha = \sum_k in_k + \sum_k out_k,$$

kde  $k$  je počet neurálních modulů v síti,  $IN$  je počet vstupů architektury,  $OUT$  je počet výstupů architektury,  $\sum_k out_k$  je suma všech výstupů přes všechny moduly a

$\alpha$  je celkový počet vstupů a výstupů přes všechny moduly. Toto kódování je tedy pro hybridní systémy podobně efektivní jako kódování pomocí matice vah pro klasické ANN, nicméně pro klasické ANN je zbytečně redundantní.

#### 4 Závěr

Předkládaný model umožňuje vytvářet a testovat nové architektury hybridních modulárních systémů. Výsledné systémy jsou tvořeny kombinací jednotlivých modulů. Každý modul může implementovat libovolný algoritmus, od jednoduchých matematických operací až například po komplexní plánovací algoritmy. Tento model nám umožňuje přímo kombinovat pod-systémy, pocházející z různých vědních oborů, jako například umělá inteligence, umělý život, výpočetní neurověda, kognitivní věda, fuzzy logika, expertní systémy, robotika, řízení apod.

Takto definované modulární systémy je navíc možné reprezentovat podobně jako umělé neuronové sítě. Díky tomu lze aplikovat postupy pro optimalizaci topologie neuronových sítí, například modifikovanou neuro-evoluci.

V tomto článku jsme zatím především představili náš nový "*framework*" sloužící k dosažení kompatibility mezi libovolnými algoritmy. Tím se nám otevřela možnost integrovat různé části softwaru (ale i HW) jako neurální moduly do velkých modulárních architektur.

Očekávané výhody tohoto přístupu jsou především v pokusu o sjednocení komunikačního rozhraní mezi různými algoritmy a vědními obory a využití jejich kombinací pro syntézu nových modulárních architektur. Tyto architektury (ať se již jedná o jejich reprezentaci nebo způsob komunikace mezi moduly) jsou blíže inspirovány modulárními neuronovými sítěmi. K jejich dalším výhodám patří především možnost těsného sjednocení dvou protichůdných přístupů návrhu architektur autonomních agentů: zdola-nahoru a ze shora-dolů. Tento způsob reprezentace modulárních systémů nám také umožňuje zkoumat možnosti automatického návrhu modulárních systémů pro daný problém.

Představili jsme simulátor těchto hybridních modulárních systémů *Nengoros*, který umožňuje rychlou kompozici a ověření funkce nových architektur složených z přepoužitelných modulů. Krom možností automatického návrhu a zkoušení nových architektur, tento simulátor poskytuje možnost ručního návrhu a odzkoušení prakticky libovolného modulárního systému z libovolné oblasti výzkumu.

Představili jsme zde také prvotní možnosti reprezentace těchto hybridních neuronových sítí vhodné pro použití v evolučních algoritmech a ukázali, že asymptotická složitost takové reprezentace je stejná jako klasická reprezentace pomocí matice vah. S touto reprezentací byly zatím provedeny základní experimenty ukazující na její praktickou použitelnost především pro síť s malým počtem spojů mezi moduly. Neuro-evoluční algoritmus může být použit pro rychlé prohledávání rozsáhlého prostoru kombinací jednotlivých podsystémů a navrhnout tak zcela nové modulární architektury, optimalizované přímo pro daný problém.

Do budoucna jako náš hlavní cíl zůstává výzkum nových typů reprezentace těchto modulárních sítí a hledání vhodnějšího algoritmu pro automatický návrh architektur.

#### Poděkování

Tento příspěvek vznikl za podpory Katedry kybernetiky na Fakultě elektrotechnické, ČVUT v Praze a v rámci SGS grantu číslo SGS12/146/OHK3/2T/13.

## Literatura

- [1] G. Auda and M. Kamel. Modular Neural Networks: a Survey. *Int J Neural Syst*, 9(2):129–151, Apr 1999.
- [2] G. Cybenko. Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals, and Systems (MCSS)*, 2:303–314, 1989.
- [3] J. Fekiac, I. Zelinka, and J. C. Burguillo. A review of methods for encoding neural network topologies in evolutionary computation. In *Proceedings 25th European Conference on Modelling and Simulation ECMS*, pp. 410–416, 2011.
- [4] Gernot A. Fink. *Markov Models for Pattern Recognition*. Springer Verlag, 2003.
- [5] P. Fišer, J. Schmidt, Z. Vašíček, and L. Sekanina. On logic synthesis of conventionally hard to synthesize circuits using genetic programming. In *IEEE 13th International Symposium on Design and Diagnostics of Electronic Circuits and Systems (DDECS)*, pages 346–351, 2010.
- [6] D. Kadleček and P. Nahodil. New Hybrid Architecture in Artificial Life Simulation. In: *Advances in Artificial Life*. Lecture Notes in Computer Science, Volume 2159, Springer Verlag, Berlin, 2001. ISBN: 978-3-540-42567-0.
- [7] P. Kordík. Fully automated knowledge extraction using group of adaptive models evolution. PhD thesis, Czech Technical University in Prague, FEE, Dept. of Computer Sci. and Computers, 112 pp., 2006.
- [8] W. Maass. Networks of spiking neurons: The third generation of neural network models. In *Journal Neural Networks*, 10:1659–1671, 1996.
- [9] J. M. J. Murre, R. H. Phaf, and G. Wolters. Calm networks: a modular approach to supervised and unsupervised learning. In *Proc. Int Neural Networks /JCNN*. Joint Conf, pages 649–656, 1989.
- [10] V.A. Nguyen, J.A. Starzyk, Wooi-Boon Goh, and D. Jachyra. Neural network structure for spatio-temporal long-term memory. In *IEEE Transactions on Neural Networks and Learning Systems*, volume 23, pp. 971–983, 2012.
- [11] M. Quigley, K. Conley, B. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, and A. Ng. ROS: an open-source robot operating system. In *ICRA Workshop on Open Source Software*, 2009.
- [12] S. Sardina, deSilva, L. and L. Padgham. Hierarchical planning in bdi agent programming language: a formal approach. In *AAMAS 06 Proceedings of the 5th International Joint Conference on Autonomous Agents and Multiagent Systems*, pp. 1001–1008, 2006.
- [13] Y. Sugita and Martin V. Butz. Towards emergent strong systematicity in a simple dynamical connectionist network. In *Computer Technologies and Information Sciences; Biology and Medicine*. Department of Cognitive Psychology, Universitat Wurzburg, 2008.
- [14] J. Vítků. An Artificial Creature Capable of Learning from Experience in Order to Fulfill More Complex Tasks, Diploma Thesis, Dept. of Cybernetics, Faculty of Electrical Engineering, Czech Technical University in Prague, 2011.
- [15] D. Wang and M. A. Arbib. Complex temporal sequence learning based on short-term memory. 78(9): 1536–1543, 1990.
- [16] J.-H. Wang and M.C. Tsai. Learning recognition of temporal sequences by coding temporal distance in neural networks. In *IEEE International Joint Conference on Neural Networks Proceedings*, volume 2, pages 1422–1427, 1998.
- [17] S. Wermter and R. Sun. *Hybrid Neural Systems*. Springer Verlag, Heidelberg, New York, 2000.



# Využití metod vicekriteriálního rozhodování pro analýzu e-mailových zpráv

Libor Olajec

Ústav informatiky, Slezská univerzita Opava  
Bezručovo nám. 13, 746 01 Opava  
libor.olajec@fpf.slu.cz

## Abstrakt

Vicekriteriální rozhodování je založeno na matematickém modelování. Rozhodováním rozumíme vybrání jedné varianty, která se nachází v seznamu realizovatelných variant na základě určitého počtu kritérií. U rozhodovacích úloh se důsledky rozhodnutí posuzují podle různých kritérií, tyto úlohy můžeme nazvat jako úlohy vicekriteriálního rozhodování. Rozhodování v těchto úlohách je závislé na optimalizaci a transformaci informací. V příspěvku vytvoříme soustavu kritérií hodnocení a jejich vah, na jejichž základě a s použitím vybrané vicekriteriální vyhodnocovací metody budeme detekovat obsah emailových zpráv. Dle vypočtených výstupních hodnot určíme, jde li o žádoucí případně nežádoucí sdělení.

## 1 Úvod

Spam je výraz, který označuje nevyžádané a obtěžující sdělení, většinou se jedná o hromadně rozesílaný reklamní mail. Tuto zprávu musí adresát případně speciální software detekovat, tím je zatížen jak výpočetní prostředek, tak adresát emailu, adresát je připraven o čas.

Rozlišení toho co je spam ovšem není jednoznačné, záleží například na tom v jakém prostředí se nacházíme. Spam nemusí být jenom čistě komerční sdělení, ale může jít rovněž o sdělení politické, akademické, náboženské aj. Na pracovišti nás odvádí od práce reklamní sdělení, ale toto samé sdělení nás nutně nemusí obtěžovat v prostředí domova, proto při detekování spamů klasickými metodami prostřednictvím speciálního počítačových programů (antispamů), může docházet k chybám, které způsobují, že příchozí poštovní zprávy mohou být označeny za neregulérní, přes to že to nemusí být pro daného příjemce v daném prostředí pravda. Základní techniky, které programy pro detekci spamu provádějí, jsou následující:

- černá listina (blacklist) – adresa IP odesílatele je na černém seznamu (uloženém na serveru poskytovatele služeb), příchozí pošta z této adresy je označena jako spam,
- černé seznamy již známého spamu – obsahují použité spamové zprávy, u každé této zprávy je vypočten identifikátor, jehož hodnota je pak následně porovnána s příchozí poštou,

- Heuristiky – hledání typických příznaků v těle zprávy (znaky, slova, odkazy atd.) prostřednictvím speciálního algoritmu,
- Bayesovské filtry – jde o kombinaci heuristických algoritmů za spolupráce samotných uživatelů, kteří sami označí zprávu za spam a filtr se podle těchto zpráv učí spam rozpoznávat.

U mechanismu vicekriteriálního rozhodování (VKR) lze využít matematicko-statistické přístupy, mechanismus je rovněž vhodný z hlediska možnosti včleňování velmi širokého spektra kritérií, navíc si můžeme tato kritéria individuálně dle určených priorit sami nastavovat. Zároveň lze tento způsob využít pro konstrukci agentů v MAS<sup>1</sup>, problematiku sw agentů a tvorba MAS s vicekriteriálním procesem je představena v [7, 9]. Cílem příspěvku je představit možnost využití metody vicekriteriálního rozhodování pro detekci obsahu e-mailových zpráv.

## 2 Vicekriteriální analýza - základní popis

Teorie vicekriteriálního rozhodování je založena na matematickém modelování, kde podstata řešení úloh spočívá ve vybrání vhodné varianty ze seznamu potenciaálně realizovatelných variant na základě množství kritérií. Pro výběr metody VKR je nutno znát:

- o čem je potřeba rozhodnout,
- jaké mají být splněny cíle,
- jaké jsou podmínky pro splnění cílů.

Pro obecný postup VKR je nutno vytvořit soustavu kritérií hodnocení. Kde kritéria jsou definována podle [3] jako hlediska zvolená rozhodovatelem sloužící k posouzení výhodnosti jednotlivých variant rozhodování z hlediska dosažení, resp. stupně plnění dílčích cílů řešeného rozhodovacího problému.

Pro racionální tvorbu kritérií je nutno důkladně poznat objekt, který chceme hodnotit (v našem případě jde o e-mailovou zprávu, ta je v našem řešení objektem hodnocení a současně výslednou variantou) včetně jeho struktury a funkcí.

<sup>1</sup> MAS – multiagentový system viz [KUBÍK, 2004].

*Toto je princip námi představeného řešení, nevyhodnocujeme dle stanovených kritérií řešení, ale samotný obsah doručené e-mailové zprávy.*

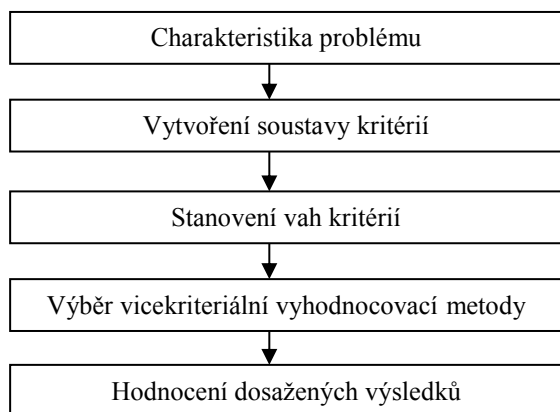
Požadavky, které musí stanovený soubor kritérií splňovat, jsou úplnost, operacionalita, neredundance a minimální rozsah [3].

Úplnost zajišťuje zvážení všech aspektů řešeného problému a přestože není splnění tohoto požadavku dle [3] jednoduché, lze jeho míru naplnění zvýšit využitím expertů z různých oblastí při specifikaci kritérií.

Operacionalita předpokládá jasný a jednoznačný smysl kritérií, čehož lze snadněji dosáhnout u kritérií kvantitativní povahy, hůře u kritérií povahy kvalitativní, kde pracujeme zpravidla se slovním popisem, jak uvádí [3].

Neredundance je požadavkem na zamezení překrývání a tedy duplicit kritérií a konečně minimální rozsah souboru kritérií, který zajišťuje vyšší přehlednost a zjednodušení závěrečného hodnocení, přestože ve svém důsledku může způsobit při agregaci kritérií snížení jejich operacionality [1].

Postup řešení problému s pomocí vícekritériální analýzy:



Obr. 1. Kroky pro řešení úloh s pomocí VKR

### 3 Nastavení vah kritérií

Pro využití vhodné varianty řešení ve VKR je nutno vytvořit úplnou soustavu kritérií a jednotlivým kritériím rovněž přiřadit váhy určující jejich důležitost. Při použití těchto diferenciováných vah, jsou výsledky rozhodnutí na těchto vahách závislé.

Pro stanovení vhodných vah existuje množství metod. Přičemž platí, že metodám jenž konstruují váhy na základě subjektivního návrhu uživatele říkáme, že jde o metody:

- přímé.

Při vysoké důležitosti kritérií stoupá rovněž důležitost jejich vah. Hodnoty nastavení vah je pak možno nastavit uživatelem velmi obtížně proto existují metody, které s pomocí jednoduchých subjektivních uživatelských informací umožňují konstrukci vah, jsou to metody:

- nepřímé – mezi které můžeme zařadit například metodu Fullerova trojúhelníku nebo Saatyho metodu.

Podle [6] v praxi existuje několik způsobů, jak zvolit vhodná kritéria a určit jejich váhu:

1. Návrh kritérií i jejich vah je záležitostí jediného odborníka, který pravděpodobně sám také provede výpočet a seřazení variant některou z metod vícekritériálního rozhodování.
2. Na určování vah kritérií, a případně i na volbě kritérií vůbec, se podílí skupina odborníků. Výsledek pak lze získat například tak, že se vybraná skupina odborníků sejde a při společné diskusi vyberou kritéria relevantní pro řešení daného rozhodovacího problému a určí i jejich vzájemnou procesní důležitost.
3. Rovněž je zde možnost vybrané skupině odborníků rozeslat dotazníky s návrhy kritérií, do nichž tito uvedou svůj názor na důležitost jednotlivých kritérií.

Podle [4] se nejčastěji používají tyto čtyři způsoby stanovení vah:

Zachování proporcionality:

$$v_{i1} = \frac{a_i}{\max a_i} \quad (1)$$

Převedení hodnoty do intervalu 0,1:

$$v_{i2} = \frac{a_i}{\sum_i a_i} \quad (2)$$

Nejčastěji používaný způsob při Saatyho metodě:

$$v_{i3} = \frac{a_i - \min a_i}{\max a_i - \min a_i} \quad (3)$$

Nejčastěji využívaný při normování unimodulárních vektorů, využívá se například při nasazení metody TOPSIS (Technique for Order Preference by Similarity to Ideal Solution):

$$v_{i4} = \frac{a_i}{(\sum_i a_i^2)^{\frac{1}{2}}} \quad (4)$$

Normované váhy představují nezáporné čísla, jejich součet je roven jedné.

### 3.1 Párové srovnání kritérií – Fullerova metoda

Párově srovnáváme mezi sebou dva kritéria a vyhodnocujeme, které kritérium je důležitější. Expertem (uživatel) je provedeno srovnání dvou kritérií mezi sebou, počet srovnání je určen následujícím vztahem

$$N = \binom{k}{2} = \frac{k(k-1)}{2} \quad (5)$$

Pomocí tzv. Fullerova trojúhelníku (6) můžeme provést srovnání. Kritéria očíslovíme pořadovými čísly  $1, 2, \dots, k$ . Uživatel (expert) pak dle předloženého trojúhelníkového schématu, ve kterém se každá dvojice kritérií nachází právě jedenkrát, vyznačí například kroužkem každé z dvojice to kritérium, které je dle jeho názoru důležitější. Jestliže počet zakroužkování  $i$ -tého kritéria označíme  $n_i$  pak váhu  $i$ -tého kritéria vypočteme dle:

$$v_i = \frac{n_i}{N}, \quad i = 1, 2, \dots, k \quad (6)$$

Schéma Fullerova trojúhelníku:

1	1	1	.	.	1
2	3	4	.	.	k
	2	2	.	.	2
	3	4	.	.	k
	.	.	.	.	k
				k-2	k-2
				k-1	k
					k-1
					k

(7)

Jelikož tato metoda vyžaduje informace od uživatele a zároveň jde o metodu velmi jednoduchou lze ji velmi efektivně využít pro určení vah kritérií pro detekci obsahu zpráv.

### 4 Výběr vícekritériální vyhodnocovací metody

Vícekritériální vyhodnocovací metody lze rozdělit dle výpočetního principu, jenž tyto metody využívají, jde například o:

- maximalizaci užítku,
- metody využívající vyhodnocování variant na základě relací a jejich preferencí,
- metody hledající minimální vzdálenost od ideální varianty,
  - TOPSIS, tuto metodu jsme zvolili jako vhodnou pro nasazení do systému.

Při základním postupu při detekci spamu je ideální variantou například zpráva odeslaná přímo z adresy IP, která se nachází na blacklistu, případně zpráva obsahující stejné sekvence, jako zpráva vyskytující se na blacklistu

těchto sekvencí. Naopak nejhorší variantou zpráv jsou ty, jenž žádné sekvence z blacklistu neobsahují, přesto nemůžeme s určitostí říci, že nejde o spam, můžeme tedy vždy jen určit minimální vzdálenost od ideální varianty.

### 5 Nasazení metody pro detekci obsahu zprávy – TOPSIS

Při použití této metody se jedná o výběr varianty, která je co nejbližší k ideální variantě, jde o variantu, pro kterou všechny hodnoty kritérií dosahují nejlepších hodnot. Ideální varianta je reprezentována vektorem  $(H_1, H_2, \dots, H_k)$  a je co nejdále od bazální varianty reprezentované vektorem  $(D_1, D_2, \dots, D_k)$ .

Nejdříve se konstruuje normalizovaná kritériální matice  $R = (r_{ij})$ , pro výpočet normalizovaných hodnot platí vzorec:

$$r_{ij} = \frac{v_{ij}}{\sqrt{\sum_{i=1}^p v_{ij}^2}} \quad (8)$$

kde:  $i = 1, 2, \dots, p$   
 $j = 1, 2, \dots, k$

Po provedení transformace jsou sloupce v matici  $R$  vektory jednotkové délky. Následuje výpočet vážené kritériální matice  $W$ , každý  $j$ -tý sloupec normalizované matice  $R$ , vynásobíme odpovídající vahou  $v_j$ :

$$W = \begin{bmatrix} w_{11} & \dots & w_{1k} \\ \vdots & \ddots & \vdots \\ w_{p1} & \dots & w_{pk} \end{bmatrix} = \begin{bmatrix} v_1 \cdot r_{11} & \dots & v_k \cdot r_{1k} \\ \vdots & \ddots & \vdots \\ v_1 \cdot r_{p1} & \dots & v_k \cdot r_{pk} \end{bmatrix} \quad (9)$$

Určíme ideální variantu:

$$H_j = \max_i(w_{ij}); \quad i = 1, 2, \dots, k \quad (10)$$

Určíme bazální variantu:

$$D_j = \min_i(w_{ij}); \quad j = 1, 2, \dots, k \quad (11)$$

Vypočítáme vzdálenosti variant od ideální varianty:

$$d_i^+ = \sqrt{\sum_{j=1}^k (w_{ij} - H_j)^2}; \quad i = 1, 2, \dots, p \quad (12)$$

Vypočítáme vzdálenosti variant od ideální varianty

$$d_i^- = \sqrt{\sum_{j=1}^k (w_{ij} - D_j)^2}; \quad i = 1, 2, \dots, p \quad (13)$$

Vypočteme relativní ukazatele vzdálenosti  $i$ -té varianty od bazální varianty

$$c_i = \frac{d_i^-}{d_i^+ + d_i^-}; \quad i = 1, 2, \dots, p \quad (14)$$

Pro hodnoty  $c_i$  platí:

$$0 \leq c_i \leq 1$$

$$c_i = 0 \leftrightarrow a_i \approx (D_1, D_2, \dots, D_k)$$

$$c_i = 1 \leftrightarrow a_i \approx (H_1, H_2, \dots, H_k)$$

Pro hodnoty těchto ukazatelů platí, že hodnota ukazatele rovna 0 se nabývá pro bazální variantu a hodnota ukazatele rovna 1 – platí pro ideální variantu.

## 6 Implementace metody TOPSIS pro detekci zpráv

V případě metody TOPSIS se jedná o princip minimalizace vzdálenosti od ideální varianty [2].

Podle [6] ideální variantou nazveme variantu, pro kterou všechny hodnoty kritérií dosahují nejlepších hodnot. Ideální varianta je většinou hypotetická, jako nejlepší se pak vybírá taková, která je podle určité metriky nejbližší k ideální variantě.

Metoda TOPSIS poskytuje úplné uspořádání množiny všech variant, tj., jak bylo řečeno výše, je určena i pro výběr nejlepší varianty. Požadovanými vstupními údaji jsou kritériální hodnoty pro jednotlivé varianty a váhy jednotlivých kritérií.

Pro testování byl použit soubor dat, který obsahuje šest zpráv. Tyto zprávy byly zaslány autorům během šesti hodin stejného dne a roku. Zprávy byly zpracované poštovním klientem MS Outlook 2010<sup>2</sup>. Prostřednictvím této aplikace bylo zároveň provedeno odfiltrování těchto doručených zpráv, tato filtrace bude sloužit k vyhodnocení výsledku nasazené metody VKR.

### 6.1 Určení kritérií

Pro zvolení jednotlivých kritérií byla důležitá relativní jednoduchost jejich detekce. V doručené e-mailové zprávě lze z hlediska složitosti triviálně segregovat soubor znaků v záhlaví zprávy, soubor všech slov ve zprávě, počet odkazů a jazykový formát zprávy. Kritéria pro detekci obsahu zpráv byla zvolena autory příspěvku následovně:

- kritérium  $k_1$  – počet písmen, čísel a znaků v hlavičce zprávy,
- kritérium  $k_2$  – počet slov v těle zprávy,

- kritérium  $k_3$  – počet odkazů v těle zprávy,
- kritérium  $k_4$  – jazyk zprávy (pro anglický jazyk nastavíme hodnotu 3, pro jiný jazyk než anglický a český hodnotu 2 a pro český jazyk hodnotu 1).

Nejdříve sestrojíme kritériální matici, všechny kritéria jsou maximalizační, sloupce představují kritéria  $k_i$  a řádky varianty řešení  $x_i$ :

$$\begin{bmatrix} 66 & 483 & 9 & 3 \\ 38 & 1771 & 9 & 3 \\ 52 & 1562 & 9 & 3 \\ 39 & 32 & 0 & 1 \\ 9 & 16 & 0 & 1 \\ 55 & 67 & 1 & 1 \end{bmatrix}$$

Dle vzorce (4 nebo 8) transformujeme prvky matice na prvky matice normalizované:

$$\begin{bmatrix} 0,576 & 0,200 & 0,576 & 0,548 \\ 0,332 & 0,734 & 0,576 & 0,548 \\ 0,454 & 0,648 & 0,576 & 0,548 \\ 0,340 & 0,013 & 0,000 & 0,183 \\ 0,079 & 0,007 & 0,000 & 0,183 \\ 0,480 & 0,028 & 0,064 & 0,183 \end{bmatrix}$$

### 6.2 Nastavení vah kritérií

Dle (7) sestavíme Fullerovo schéma a dle (5) určíme počet srovnání, následuje tabulka:

Kritéria			Body
1	(1)	(1)	1,5
(2)	(3)	4	
	(2)	(2)	3
	3	4	
		(3)	1
		(4)	0,5

Pramen: [7], upraveno autorem

Dle (6) vypočteme vektor vah kritérií:

$$v_4 = (0,25; 0,5; 0,167; 0,083)$$

Dle vzorce (9) vypočítáme váženou kritériální matici  $W$ :

$$\begin{bmatrix} 0,144 & 0,100 & 0,096 & 0,045 \\ 0,083 & 0,367 & 0,096 & 0,045 \\ 0,114 & 0,324 & 0,096 & 0,045 \\ 0,085 & 0,007 & 0,000 & 0,015 \\ 0,020 & 0,004 & 0,000 & 0,015 \\ 0,120 & 0,014 & 0,011 & 0,015 \end{bmatrix}$$

<sup>2</sup> MS Outlook 2010 – jde o e-mailového klienta společnosti Microsoft. Aplikace Microsoft Outlook 2010 obsahuje filtr nevyžádané pošty vytvořený s cílem snížit počet nevyžádaných e-mailových zpráv ve složce Doručená pošta. Nevyžádaná pošta (spam) je pomocí tohoto filtru přesměrována ze složky Doručená pošta do složky Nevyžádaná pošta, kde ji lze efektivněji spravovat. [8]

### 6.3 Určení ideální a bazální varianty

Z vážené kritériální matice určíme dle (10) ideální variantu ( $H$ ) a dle (11) bazální variantu ( $D$ ):

$$H_4 = (0,144; 0,367; 0,096; 0,045)$$

$$D_4 = (0,020; 0,004; 0,000; 0,015)$$

### 6.4 Vzdálenosti variant od ideální a bazální varianty

Dle vzorce (12) vypočítáme vzdálenosti od ideální varianty (zprávy)  $d_i^+$ , kde  $i = 1, 2, \dots, 5$  a dle vzorce (13) od bazální varianty (zprávy)  $d_i^-$ , kde  $i = 1, 2, \dots, 6$ :

Varianta $x_i$	$d_i^+$	$d_i^-$
$x_1$	0,267	0,186
$x_2$	0,061	0,382
$x_3$	0,043	0,349
$x_4$	0,378	0,065
$x_5$	0,397	0,000
$x_6$	0,162	0,101

Pro výpočet relativního ukazatele vzdálenosti  $c_i$  použijeme vzorce (14):

Varianta $x_i$	$c_i$
$x_1$	0,4106
$x_2$	0,8623
$x_3$	0,8903
$x_4$	0,1467
$x_5$	0,0000
$x_6$	0,3840

Výsledné varianty dle relativního ukazatele  $c_i$  jsou vyhodnocované zprávy:

zpráva $_i$	$c_i$
zpráva $_1$	0,4106
zpráva $_2$	0,8623
zpráva $_3$	0,8903
zpráva $_4$	0,1467
zpráva $_5$	0,0000
zpráva $_6$	0,3840

Pro vyhodnocení zpráv je seřadíme dle výsledku relativního ukazatele  $c_i$ :

zpráva $_i$	$c_i$	SPAM
zpráva $_2$	0,8623	↑
zpráva $_3$	0,8903	
zpráva $_1$	0,4106	
zpráva $_6$	0,3840	
zpráva $_4$	0,1467	
zpráva $_5$	0,0000	

## 7 Závěr

V příspěvku jsme vytvořili soustavu kritérií pro detekci obsahu e-mailových zpráv, provedli jsme nastavení jednotlivých vah pomocí metody párového srovnávání kritérií – Fullerovy metody, kde uživatel ve schématu tzv. Fullerova trojúhelníku srovnává každé ze dvou kritérií mezi sebou.

Dále byl proveden výběr vhodné vyhodnocovací metody TOPSIS. Tato metoda umožňuje vybrat variantu řešení, která je co nejbližší k ideální variantě – jde o variantu, pro kterou všechny hodnoty kritérií dosahují nejlepších hodnot. Ideální varianta je co nejdále od bazální varianty.

Při našem postupu by byla ideální variantou zpráva, pro kterou by se výsledná hodnota relativního ukazatele  $c_i=1$ , šlo by tedy o stoprocentní SPAM. V závěrečné tabulce příspěvku je provedeno seřazení souboru testovaných zpráv dle relativního ukazatele  $c_i$ .

Jelikož rozlišení toho co je spam není jednoznačné, kdy záleží například na tom, v jakém pracovním prostředí se příjemce zprávy nachází, je při námi představené implementaci VKR výhodou možnost individuálního uživatelského nastavení kritérií a rovněž hranice relativního ukazatele vzdálenosti  $c_i$ .

### Poděkování

Tento článek vznikl za podpory Operačního programu Vzdělávání pro konkurenceschopnost, který je spolufinancován Evropským sociálním fondem a Státním rozpočtem ČR, projekt č. CZ.1.07/2.2.00/28.0014, „Interdisciplinární vzdělávání v ICT s jazykovou kompetencí“ a v rámci projektu Centrum excelence IT4Innovations, reg. č. CZ.1.05/1.1.00/02.0070.

### Literatura

- [1] Blažek, L. Sylabus k předmětu Management I, část Rozhodování. [online]. Dostupný na WWW: <http://is.muni.cz>. str. 13
- [2] Fiala, P., Jablonský J., Maňas, M., Vícekriteriální rozhodování, VŠE Praha, 1994
- [3] Fotr, J., Dědina, J., Hružová, H. Manažerské rozhodování. Praha: Ekopress, 2003. 250 s., str. 16
- [4] Hersh, M. Mathematical modeling for sustainable development. Heidelberg: Springer-Verlag, 2006.

<sup>3</sup> Pracovní prostředí – je definováno jako prostředí, ve kterém je pracovníkem prováděná specifická pracovní činnost, pracovní činnost může být prováděna například v oblasti politické, akademické, kulturní, výrobní aj.

- [5] Kavan, M. Výrobní a provozní management. Praha: Grada Publishing, a.s., 2002. 424 s., str. 255.
- [6] Korvini, P. Teoretické základy vícekriteriálního rozhodování. In: [online]. [cit. 2013-04-07]. Dostupné z: [http://korviny.cz/mca7/soubory/teorie\\_mca.pdf](http://korviny.cz/mca7/soubory/teorie_mca.pdf)
- [7] Kubík, A. Inteligentní agenty, tvorba aplikačního software na bázi multiagentových systémů, Computer Press, Brno, 2004.
- [8] Office 365. *Office - office.com* [online]. 2013 [cit. 2013-04-07]. Dostupné z: <http://office.microsoft.com/cs-cz/outlook-help/prehled-filtru-nevyzadane-posty-HP010355048.aspx>
- [9] Olajec, L. Multiagentový systém s více-kriteriálním rozhodováním proti robotickým sítím. In *Kognice a umělý život XII*, (Kelemen, J., Nahodil, P., sest.), Praha, ČR 2012, s. 176–181.

# Konekcionistický model uchopovania

Matej Pecháč

Centrum pre kognitívnu vedu, KAI FMFI UK, Univerzita Komenského v Bratislave  
Mlynská dolina, 84248 Bratislava  
Email: matej.pechac@gmail.com

## Abstrakt

Jedným z problémov súčasnej kognitívnej robotike je úloha uchopovať predmety efektormi rôznych typov (humanoidné/priemyselné ruky). Vyvinuli sme model ovládajúci priemyselné robotické rameno, ktorý je zameraný práve na riešenie tejto úlohy v jednoduchých situáciách. Úlohu uchopovania sme sa rozhodli rozdeliť na dve časti, a to na dosahovanie a samotný úchop. Každú z týchto podúloh riešil samostatný modul s architektúrou aktér–kritik. Najskôr tréning a testovanie prebiehalo v simulovanom prostredí na reálnych dátach a neskôr sme natrénovaný model testovali aj v laboratórnych podmienkach s robotickou rukou. Testy boli zamerané na uchopovanie naučených predmetov s pridaním šumu do ich pozície a orientácie.

## 1 Úvod

V tomto príspevku predstavíme jednoduchý model uchopovania predmetov, ktorý sme navrhli pre použitie v reálnom prostredí. K dispozícii sme mali priemyselné robotické rameno Katana 450 s piatimi stupňami voľnosti a jednoduchou dvojrstou rukou. Tým sa nám značne zredukoval počet póz a typov úchopov, aké sme mohli otestovať a v experimentoch sa zameriavame na vykonanie úchopu „od boku“ a úchopu „zhora“.

Uvádzame najskôr niekoľko iných modelov, ktoré v súčasnosti existujú a čitateľ sa s nimi môže oboznámiť tu [1-4]. Nepochybne veľmi komplexným modelom je model D. Krafta a kol. [1], kde do samotného modelu dosahovania a uchopovania predmetov zahrnul aj niekoľko druhov pamäte. Prvou je ikonická pamäť, ktorá by sa dala chápať ako vizuálny vstup. Do vizuálno-priestorového skicára sa ukladá cieľový objekt a buduje sa jeho interná reprezentácia. Po vykonaní úchopu sa do epizodickej pamäte ukladá záznam o kvalite úchopu. Nakoniec sa buduje objektová a úchopová pamäť a v úchopovej pamäti vznikajú na základe skúsenosti afordancie. Učenie modelu pozostáva z troch štádií, ktoré sú blízke štádiám dieťaťa. V prvej si model vytvára priestorovú predstavu o tvare a veľkosti objektu a pokúša sa robiť náhodné úchopy. O ich úspešnosti si práve v epizodickej pamäti vedie záznam. V druhom

štádiu, ktoré Kraft nazýva hraním si model skúša rôzne uchopenia a buduje už množinu afordancií, na základe ktorej vie vytvoriť odhad pre najlepšiu pozíciu ruky. Nakoniec v tretej fáze dochádza k vytváraniu plánov, ktorých cieľom je uchopiť objekt optimálne (typickým príkladom môže byť šálka, ktorú by tento model uchopil za ucho).

V našom prípade sme sa do modelu rozhodli nezahrnúť tak komplexnú štruktúru pamäte a viacstupňové učenie, ale sústredili sme sa na implementáciu modelu, ktorý by vedel vykonávať úchopy ako Kraftov model v prvom štádiu. Avšak pomocou techník spracovania obrazu sa nejedná o náhodné úchopy, ale hrubé úchopy zamerané na správne natočenie ruky, aké je možné pozorovať v ranom štádiu uchopovania u detí [3].

## 2 Model uchopovania

Problém uchopovania sme rozdelili na dva podproblémy: dosiahnutie objektu a uchopenie objektu. Z toho dôvodu sme vytvorili dva ekvivalentné moduly s rovnakou architektúrou aj učiacim algoritmom a inými parametrami.

Po analýze komplexity prostredia a stanovení cieľov nášho projektu sme sa rozhodli vo výpočtovom modeli použiť dopredné neurónové siete a učenie s posilňovaním ako biologicky prijateľné riešenie pre našu úlohu. Keďže akcie aj stavy prostredia sú spojité vybrali sme algoritmus spojitého učenia aktéra a kritika (CACLA) [5]. Ide o pomerne nový a menej používaný algoritmus, preto uvádzame aj jeho zápis v pseudokóde s krátkym vysvetlením:

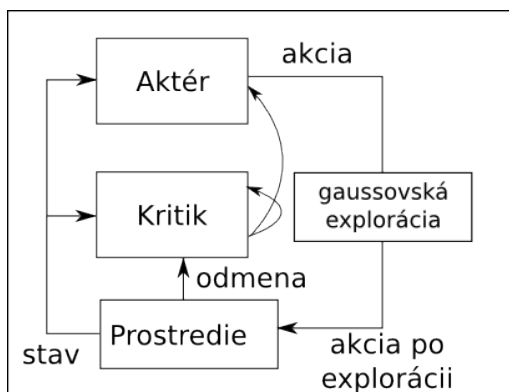
```

1: Initialize  $\theta_0$  (below  $V_t(s) = V(s, \theta_t)$ ),  $\psi_0$ ,  $s_0$ .
2: for  $t \in \{0, 1, 2, \dots\}$  do
3:   Choose  $a_t \sim \pi(s_t, \psi_t)$ 
4:   Perform  $a_t$ , observe  $r_{t+1}$  and  $s_{t+1}$ 
5:    $\delta_t = r_{t+1} + \gamma V_t(s_{t+1}) - V_t(s_t)$ 
6:    $\theta_{t+1} = \theta_t + \alpha_t(s_t) \delta_t \forall \theta V_t(s_t)$ 
7:   if  $\delta_t > \theta$  then
8:      $\psi_{t+1} = \psi_t + \beta_t(s_t)(a_t - AC(s_t, \psi_t)) \nabla \psi AC(s_t, \psi_t)$ 
9:   if  $s_{t+1}$  is terminal then
10:    Reinitialize  $s_{t+1}$ 

```

V algoritme CACLA sa kritik snaží aproximovať funkciu odmeny, ktorá prichádza z prostredia a predikovať ju do budúcnosti ( $v_t(s_t)$ ). Aktér generuje výstupný vektor, ktorý je prostredím interpretovaný ako akcia. V prvom kroku dochádza k inicializácii parametrov kritika ( $\theta$ ), a aktéra ( $\psi$ ), čo sú v našom prípade váhy neurónových sietí a tiež samotného stavu prostredia ( $S$ ) (riadok 1). V nasledujúcom kroku aktér generuje akciu ( $a$ ), ktorá sa ešte zašumuje s gaussovským rozložením pravdepodobnosti (riadok 3,4). Tento proces sa nazýva exploračia a zabezpečuje vykonávanie nových akcií, ktoré sa môžu ukázať ako vhodnejšie oproti predošlým vykonaným v rovnakom stave. Po vykonaní akcie prichádza z prostredia odmena ( $r$ ) a stav prostredia ( $S$ ) sa aktualizuje. V poslednom kroku sa vypočíta chyba (riadok 5) a na základe tejto chyby sa aktualizujú parametre kritika (riadok 6) a pokiaľ je chyba pozitívna, čo znamená, že akcia po exploračia sa ukázala byť vhodnejšia, tak sa aktualizujú aj parametre aktéra (riadok 8). Pre hlbšie štúdium odkazujeme čitateľa na kapitolu H. von Hasselta [5].

Ako je možné vidieť z popisu, algoritmus CACLA beží nad tromi základnými blokmi: aktér, kritik a interaguje s prostredím (detailnú schému možno vidieť na obrázku č. 1).



**Obr. 1.** Architektúra modulu dosiahnutia/uchopenia objektu

Aktér aj kritik boli implementovaní ako dopredné neurónové siete [6] a ich vstupom bol vektor charakterizujúci stav. Za stav bola považovaná proprioceptívna informácia o polohe ramena. V našom prípade sa jednalo o pozíciu ruky v priestore, jej orientáciu vyjadrenú ako kvaternión a otvorenosť prstov ramena. Akciou bola potom zmena polohy v priestore, zmena uhlu natočenia ruky v osi dopredného pohybu ramena, čo umožňovalo pripraviť orientáciu ruky na úchop a nakoniec povel k zatvoreniu prstov. Rameno Katana 450 má iba dva stavy pre prsty a to je otvorené alebo zatvorené (hoci je možné otvorenosť vyjadriť percentuálne, čo sme využili pri určení stavu prostredia).

Podstatným blokom v našom modeli je blok prostredia, ktorý vytváral rozhranie medzi ovládacou API ramena Katana a vnútornými blokmi aktéra a kritika. Blok prostredia nekomunikoval priamo s ramenom avšak využívali sme medzivrstvu s názvom Golem, ktorá poskytovala funkcie na plánovanie a vykonanie presunu ramena do novej pozície a orientácie ruky, čo značne uľahčovalo náš postup.

Ďalšími podstatnými funkciami bloku prostredia sú funkcie odmeny a funkcie vracajúce či je aktuálny stav prostredia fatálny alebo finálny. Funkcia odmeny pre modul dosahovania reflektuje jeho úlohu tj. presun do vhodnej pozície blízko objektu a správne natočenie ruky. Je lineárnou kombináciou vzdialenosti ruky od objektu a natočenia ruky voči hlavnej osi objektu. Presný matematický zápis poskytuje rovnica č. 1.

$$r(t) = \alpha \cdot \left( 1 - \frac{2 \cdot |d(t)|}{max_d} \right)^2 \cdot sgn \left( 1 - \frac{2 \cdot |d(t)|}{max_d} \right) + \beta \cdot \left( 1 - \frac{2 \cdot \sin(|\Delta(t)|)}{max_\Delta} \right)^2 \cdot sgn \left( 1 - \frac{2 \cdot \sin(|\Delta(t)|)}{max_\Delta} \right) \quad (1)$$

Môžeme si všimnúť, že pomocou parametrov  $\alpha$  a  $\beta$  je možné dávať v odmeňovacej funkcii dôraz na skrátenie vzdialenosti alebo správne natočenie ruky. V našom modeli sme odmenu počítali s oboma parametrami rovnými hodnotou 0.5. Funkcia  $sgn$  je štandardnou matematickou funkciou pre určenie znamienka, ktorá vracia +1 pre argument väčší ako 0 a -1 pre argument menší ako 0. Premenná  $d(t)$  predstavuje vzdialenosť ruky od objektu v čase  $t$  a  $\Delta(t)$  uhol medzi osou Z (os v smere dopredného pohybu) ruky a hlavnou osou dosahovaného predmetu v čase  $t$ . Môžeme si ešte všimnúť, že používame hodnotu  $\sin(\Delta(t))$  z potreby ekvivalencie uhlov  $0^\circ$  a  $180^\circ$ , ktoré sú pre natočenie ruky rovnaké. Konštantná hodnota  $max_d$  je maximálna možná vzdialenosť od objektu a  $max_\Delta$  je rovné  $\pi/2$ .

Za finálny stav prostredia sa považuje stav, kedy je vzdialenosť ruky od objektu menšia ako polovička dĺžky jej prstov (0.15m) a rozdiel medzi uhlami hlavnej osi a osi Z ruky je menší ako  $20^\circ$ .

Na druhej strane považujeme za fatálny stav, kedy sa ruka príliš priblíži ku stolu (0.05m). Táto podmienka síce so samotnou úlohou dosiahnutia nemá nič spoločné, ale ide samozrejme o ochranu zariadenia. Ďalšou podmienkou je, že nesmú po sebe nasledovať dve akcie, ktoré by nevyvolali zmenu stavu, čo indikuje, že cieľová pozícia je pre rameno nedosiahnuteľná. Poslednou možnosťou je, že vzdialenosť ruky od objektu je väčšia ako škálovací faktor  $max_d$ , ktorý je počítaný ako vzdialenosť od základne ramena k objektu násobená parametrom (1.5).

Podobne ako sme definovali funkciu odmeny pre modul dosiahnutia zadefinujeme funkciu odmeny pre uchopenie (rovnica č. 2).



$$r(t) = \alpha \cdot (1 - g(t))^2 + \beta \cdot \left(1 - \frac{2 \cdot \sin(|\Delta(t)|)}{\max_{\Delta}}\right)^2 \cdot \operatorname{sgn}\left(1 - \frac{2 \cdot \sin(|\Delta(t)|)}{\max_{\Delta}}\right) \quad (2)$$

Môžeme si všimnúť, že druhý člen sa vôbec nelíši od rovnice č. 1, že došlo k nahradeniu prvého člena. Ten vyjadruje otvorenosť prstov ruky (premenná  $g(t)$ ) v čase  $t$ . Premenná  $g(t)$  obsahuje hodnoty z intervalu  $\langle 0,1 \rangle$ , kde 0 znamená, že ruka je zavretá, 1 naopak otvorená. Preto nebolo nutné nijako tento člen škálovať, ako v predchádzajúcich prípadoch. Druhý člen lineárnej kombinácie opäť vyjadruje uhol medzi hlavnou osou a osou Z ruky. V prípade uchopenia sme fatálny stav definovali ako stav, kedy sa prsty ruky dokázali úplne uzatvoriť, čo znamenalo, že nebol uchopený žiaden objekt. V opačnom prípade kedy prsty niečo uchopili bolo zovretie zastavené vlastnou API ramena, keď „pocíťovalo“ odpor v zatváraní prstov. Nevyhnutné pre finálny stav okrem neúplného uzavretia prstov bol rozdiel uhlu osi Z ruky a hlavnej osi menší ako  $5^\circ$ .

### 3 Experimenty

Na učenie a neskôr otestovanie nášho modelu sme pripravili tréningovú množinu použitím zariadenia *Microsoft Kinect*, ktoré dokáže vytvárať RGB-D dáta. Vybrali sme štyri objekty s rozličným tvarom a veľkosťou (čajová krabica, krabica so šesťuholníkovou podstavou, vertikálna krabica, plechovka), pri čom dva objekty mali hlavnú os rovnobežnú so stolom a dva os kolmú na rovinu stola. Sekundárna os objektov bola mierne pootočená, čo vyžadovalo dodatočné natočenia ruky.

Na predspracovanie dát, ktoré nahradilo vizuálnu kôru sme použili knižnicu Point Cloud Library (PCL) [12] a knižnicu Vision for Robotics (V4R), ktorá je vyvíjaná na Univerzite vo Viedni. Knižnica V4R zabezpečila vysegmentovanie objektu zo stola. Keďže súradnice objektov boli v súradnicovej kamery, pomocou geometrických transformácií sme súradnicový systém zjednotili so súradnicovým systémom ramena, ktorého počiatok bol v jeho základni. Následne sme pomocou metódy knižnice PCL boli schopní vypočítať ťažisko objektu, ktoré sme považovali za koordináty objektu. Ako posledné sme vypočítali kovariančnú maticu z oblaku bodov objektu a mohli sme aplikovať analýzu hlavných komponentov (Principal Component Analysis). Jej výsledkom boli tri vlastné vektory a tri vlastné čísla. Pomocou veľkosti vlastných čísel bolo možné určiť hlavnú os objektu. Čitateľa, ktorého zaujíma analýza, spracovanie a kategorizácia objektov môžeme upriamiť na publikácie skupiny V4R vo Viedni [10][11].

Naše experimenty sme rozdelili do dvoch častí. V prvej časti sa mal model naučiť uchopiť objekt z tréningových dát. V tomto prípade sme merali dve miery úspešnosti – kumulatívnu odmenu  $r$  a počet akcií (krokov)  $e$ , potrebných  $n$ -krát úspešne splniť cieľ, kde  $n$  je voľný parameter, ktorý sme nastavili na 5. Celá prvá časť experimentu prebiehala v simulovanom prostredí ROS (Robotic Operating System) z pochopiteľných dôvodov časovej náročnosti tohto procesu. V druhej časti sme pristúpili k experimentom v laboratóriu s ramenom Katana 450. K pozícii objektov z tréningovej množiny sme pridali šum a testovali sme na nich natrénovaný model. Tu sme merali úspešnosť modelu a počet akcií, nutných k pretrénovaniu modelu na uchopenie zašumeného objektu. Je dôležité pripomenúť, že sme zároveň so šumom pridali aj mieru exploračie, ktorá bola vždy rovná veľkosti zašumenia, aby bol model schopný explorať akcie a tak sa preučiť na uchopenie nového objektu.

### 3.1 Učenie dosiahnutia a uchopenia objektu

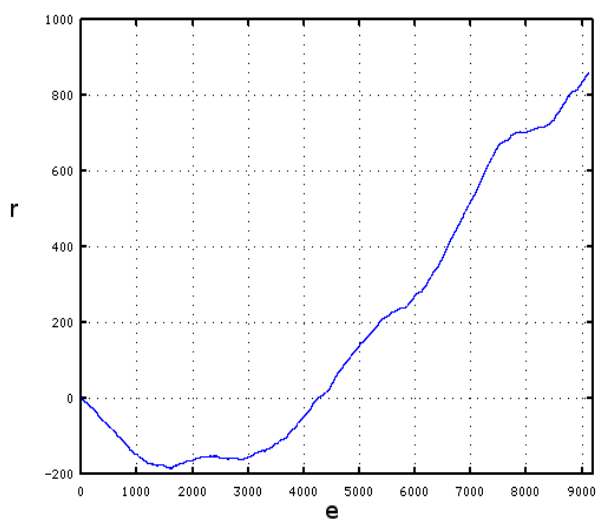
Na obrázku 2 a 3 vidíme typickú krivku učenia pre jeden z objektov (vertikálna krabica). Na začiatku robí model v podstate náhodné akcie a mapuje nimi svoj stavový priestor, preto je kumulatívna chyba negatívna a klesá. Po približne 4300 krokoch začne dominovať pozitívna odmena a model nájde akcie, ktoré sú posilňované a vedú k cieľovému stavu.

	Dosiahnutie obj.			Uchopenie obj.		
	r	e	r/e	r	e	r/e
Obj. 1	858.2	9110	0.094	247.6	637	0.39
Obj. 2	1284.2	12936	0.099	196.5	584	0.34
Obj. 3	823.5	8652	0.095	291.5	622	0.47
Obj. 4	1198.8	12571	0.095	268.9	645	0.417

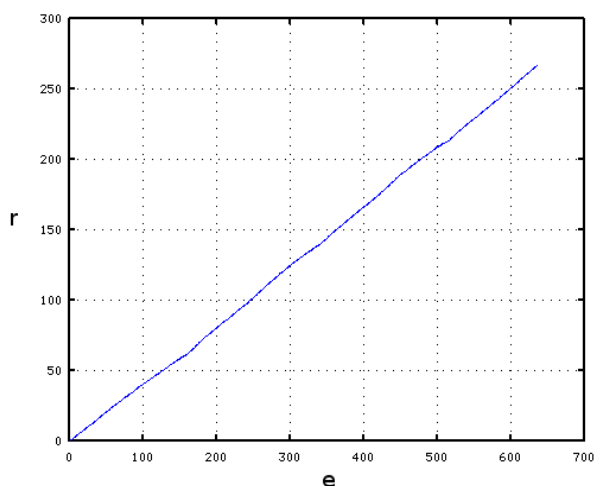
**Tab. 1:** Výsledky učenia uchopenia štyroch objektov

Tabuľka 1 ukazuje výsledky pre všetky štyri objekty v tomto poradí: 1. objekt – vertikálna krabica, 2. objekt – horizontálna krabica (od čaju), 3. objekt – plechovka (vertikálna hlavná os) a 4. objekt – ležiaca krabica so šesťuholníkovou podstavou. Vyhodnotili sme taktiež pomer kumulovanej odmeny a počtu krokov, čo sa dá považovať za priemerný prírastok odmeny po jednej akcii. V prípade modulu dosahovania vidíme, že prírastok nepresiahol hodnotu 0.01, ale pri uchopení je pomerne dosť vysoký: 0.38. Dá sa to vysvetliť oveľa väčším stavovým priestorom pri dosahovaní, kedy model omnoho ťažšie hľadá akcie, ktoré sú posilňované a dostáva sa do finálnej polohy. Preto model potrebuje omnoho viac krokov k akumulovaniu odmeny, taktiež v prípade uchopenia úplne chýba perióda negatívnej odmeny, čo je dané dizajnom funkcie odmeny pre modul

dosahovania (rovnica 2). Keďže je rozdiel osi Z ruky a hlavnej osi objektu veľmi malý (modul uchopovania začína s uhlom osí najviac  $20^\circ$ , čo je finálny stav dosahovania) a je škálovaný hodnotou  $\pi/2$ , model dostáva od začiatku pozitívnu odmenu.



Obr. 2. Krivka vývoja kumulovanej odmeny modulu dosahovania pre objekt č. 1



Obr. 3. Krivka vývoja kumulovanej odmeny modulu uchopovania pre objekt č. 1

### 3.2 Testovanie s pridaním šumu

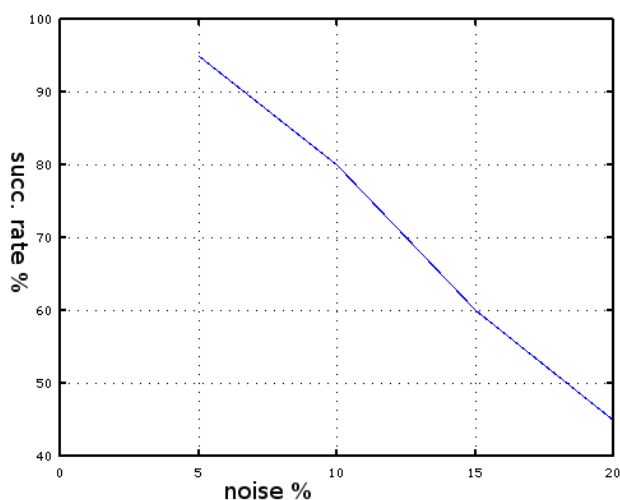
V tabuľke 2 môžeme vidieť výsledky testu, kde bol pridaný šum k pozícii objektu. V tomto štádiu experimentu bolo učenie vypnuté nastavením učiacej rýchlosti na nulu. Miera exploraácie sa vždy rovnala pridávanému šumu. Pre každú hodnotu šumu sme spúšťali 20 testov a merali úspešnosť modelu. Za

úspešný sme považovali pokus ak rameno skončilo vo finálnom stave a objekt uchopilo. Ak pokus skončil fatálnym stavom, či už pri dosahovaní alebo uchopovaní, pokus sme považovali za neúspešný.

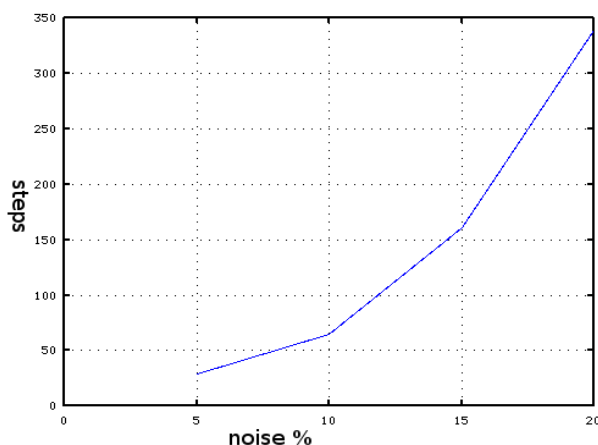
Následne sme nastavili rýchlosť učenia na pôvodné hodnoty a nechali model nech sa naučí uchopiť objekt na novej pozícii. V tomto prípade sme merali počet akcií (krokov) potrebných k preučeniu modelu.

Šum	Úspešnosť	Počet krokov
5 %	95 %	29
10 %	80 %	65
15 %	60 %	161
20 %	45 %	339

Tab. 2: Výsledky testov s pridaním šumu



Obr. 4. Vývoj úspešnosti modelu pri narastajúcej miere šumu



Obr. 5. Stúpajúci počet akcií nutných k naučeniu uchopenia objektu s pridaným šumom

## 4 Diskusia

Ako hlavný problém nášho modelu sa ukázala katastrofická interferencia, keď sme chceli model naučiť uchopovať dva a viac predmetov. Model je schopný sa naučiť uchopovať objekt v novej pozícii a orientácii, avšak stráca ihneď schopnosť uchopiť predchádzajúci objekt. Keďže sa naučené neurónové siete ukladali do súborov, nami navrhované riešenie na predídenie tomuto problému zahrňuje implementáciu klasifikátora, ktorý by na základe vizuálneho vstupu bol schopný určiť o aký objekt sa jedná a následne by nahral súbor s vhodnými neurónovými sieťami. Klasifikátor by mohol pozostávať zo samoorganizujúcej sa mapy (SOM) [7] a doprednej siete. SOM by sa organizovala pomocou vstupných obrázkov objektu a následne pomocou učenia s učiteľom by sa dopredná sieť naučila asociovať výstup SOM s korektným modelom uloženým v pamäti. Toto riešenie sme však už neimplementovali a rozhodli sme sa pre naivnejší prístup, v ktorom sme rozdelili stôl, na ktorom boli predmety položené na štyri segmenty a pre každý segment bol natrénovaný jeden model na uchopenie objektu, ktorý sa v ňom nachádza. Objekty sa samozrejme vždy nachádzali v tom istom segmente pri testoch s pridaným šumom.

Na druhej strane sa pozitívne prejavila schopnosť nášho modelu generalizovať, čo vidieť z výsledkov druhého experimentu, kde aj napriek zašumeným dátam bola stále úspešnosť nad 50% po hodnotu šumu 15%. Určite môžeme túto vlastnosť zdôvodniť použitím neurónových sietí.

Na záver diskusie chceme dodať zmienku o modely systému zrkadliacich neurónov [9], kde práve náš model je jeden zo stavebných blokov a simuluje časť mozočka. V spomínanom modeli generuje vstup pre motorickú kôru, ktorá obsahuje zrkadliace neuróny [8] a tak poskytuje možnosť kontinuálneho a súbežného učenia celého modelu.

## 5 Záver

Predstavili sme model zvládajúci úlohu dosiahnutia a uchopenia objektu, ktorý bol založený na použití algoritmu CACLA vhodný pre spojitý stavový priestor a neurónových sietí tvoriacich aktéra a kritika.

Pripravili sme tréningovú množinu z reálnych dát použitím technológie *Microsoft Kinect*, na ktorej sa náš model dokázal túto úlohu naučiť. Tiež sme ukázali schopnosť modelu generalizovať, ktorú zdôvodňujeme použitím neurónových sietí a otestovali sme model v laboratórnych podmienkach. Po preskúmaní možností nášho modelu sme objavili katastrofickú interferenciu pri pokusoch naučiť model uchopiť viac ako jeden objekt. V diskusii preto ponúkame riešenie tohto problému a tiež

pripájame jeho možnosť použitia s komplexnejším modelom systému zrkadliacich neurónov.

## Podakovanie

Chcem poďakovať svojmu vedúcemu Dipl.-Ing. Dr. techn. Michaelovi Zillichovi z Univerzity vo Viedni, Bernhardovi Mathiasovi za technickú podporu a Doc. Ing. Igorovi Farkašovi, PhD. za pomoc s algoritmom CACLA.

## Literatúra

- [1] D. Kraft, R. Detry, N. Pugeault, E. Baseski, F. Guerin, J. Piater, and N. Kruger. Development of object and grasping knowledge by robot exploration. *IEEE Transactions on Autonomous Mental Development*, 2010.
- [2] E. Oztop, M. Kawato. *Sensorimotor Control of Grasping: Physiology and Pathophysiology*. Cambridge University Press, Chapter: Models for the control of grasping, 2009.
- [3] E. Oztop, N. Bradley, and M. Arbib, Infant grasp learning: A computational model, *Exp. Brain Res.*, vol. 158, pp. 480–503, 2004
- [4] R. Prevede, G. Tessitore, M. Santoro, and E. Catanzariti. A connectionist architecture for view-independent grip-aperture computation, *Brain Research*, vol. 1225, pp. 133-145, 2008.
- [5] H. van Hasselt. Reinforcement learning in continuous state and action spaces. *Reinforcement Learning: State of the Art*, Springer, 2001.
- [6] S. Haykin. *Neural Networks: A comprehensive foundation*. Upper Saddle River, NJ, USA: Prentice Hall, pp. 59-65, 1999.
- [7] T. Kohonen. *Self-organizing Maps*. Springer, 1997.
- [8] G. Rizzolatti, L. Fogassi, V. Gallese. Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Rev. Neurosci.* 2, 661–670, 2001.
- [9] K. Rebrová, M. Pecháč, and I. Farkaš. Towards a robotic model of the mirror neuron system. *Zaslané na ICDL-EpiRob*, 2013.
- [10] J. Prankl, M. Zillich and M. Vincze. Combining Shape and Appearance – A Probabilistic Object Recognition Framework.
- [11] W. Wohlkinger, M. Vincze. Shape-based depth image to 3d model matching and classification with inter-view similarity. In: *Intelligent Robots and Systems (IROS)*, 2011 IEEE / RSJ International Conference on 2011, p. 4865–70. 2011.
- [12] PCL. Point Cloud Library <http://pointclouds.org/> Online: accessed 07-Apr-2011





... atd. a pošle tuto abecedu Searlovi s předpokladem, že pochopí a zašle stejnou abecedu jemu, aby mohli spolu komunikovat. Je nutné poznamenat, že až dosud byla komunikace úspěšná, i když ani jeden z partnerů nerozuměl obsahu zpráv. Oba však rozumí situaci v níž se ocitli a skrze tuto situaci rozumí i svému úsilí. A v této chvíli mají za cíl dopátrat se toho, zda je to partner při komunikaci, kdo mate, nebo zda je matoucím prvkem počítač zprostředkující informace. Předpokládejme, že Searle pochopí Turingův úmysl, napíše svou abecedu a oba se domluví.

Obtíže by ovšem nastaly, kdyby to byl jeden, nebo oba počítače, které by proměňovaly typ písma, řekněme na základě generátoru náhodných čísel. Pak by před Turingem i Searlem stál úkol odhalit algoritmus měnění typu písma, nebo nalézt jinou možnou cestu komunikace, např. použít program na kreslení a vepsat tam vzkazy nikoli klávesnici, ale myši a posílat to jako přílohu, atd. Podobných kombinací Turinga v Searleově čínské místnosti bychom mohli konstruovat mnoho. Nic by to však nezměnilo na existenci operací lidského myšlení. Jak u Turinga, tak i u Searlea budou v těchto případech probíhat smysluplné myšlenkové operace a oba budou vzájemně komunikovat, budou sledovat určitý cíl, který je nakonec může dovést ke vzájemnému rozumní.

Otázkou, kterou chceme nastolit, je však nikoli otázka rozumní, nebo simulace myšlenkových procesů a rozhodnutí o tom, zda jde o výkony člověka nebo stroje, ale toho, co je to lidské myšlení; a také otázku: kam až mohou dojít informační technologie, aniž by rozuměly. Výše uvedená anekdota má jen naznačit výchozí momenty dalšího směřování uvažování.

## 2 Intencionalita

Domníváme se, že těžiště se v tomto případě točí okolo:

- charakteristiky myšlení; a
- emergence intencionality.

Nastolíme-li otázku charakteristiky myšlení, stavíme před sebe velmi nesnadný úkol, neboť odpověď na tuto otázku musí být vyvozena z představy, jak pracuje mysl člověka a jaké postavení v ní zaujímá lidské myšlení. Zároveň však tato charakteristika musí být koherentní s dalšími substrukturami mysli a jejich funkcemi, tudíž se v ní musí implicitně zobrazovat i charakteristiky dalších kategorií, jakým jsou např. vědomí, zkušenost, paměť, atd., neboť jediné tak je možné podat na základě charakteristiky myšlení také vysvětlení funkcí těchto substruktur mysli.

Pro naše pojetí zkoumané problematiky vycházíme z následující charakteristiky myšlení: Myšlení charakterizujeme jako čistý proud (pohyb) operující s pojmy, ideami, představami, čísly, tóny, symboly, znaky, impulsy, atd., který vždy tyto pojmy, tóny, čísla, ..., atd. uchopuje, kombinuje, smíchává, diferencuje, poměřuje, atd., ale je za nimi, jako něco samostatného. Lidské myšlení pak obdobně charakterizujeme jako čistý proud (pohyb), v němž se uskutečňují operace s intencionálními obsahy mysli. V této souvislosti chceme upozornit na to, že Müller zjišťuje, že latinské sloveso „cogito“ pro „myslet“ etymologicky znamená „smíchat dohromady“. Myšlení také nevážeme pouze na pojmy nebo slova, ale také ne na určitý mimoslovní substrát. I když je myšlení velmi úzce vázáno na jazyk, Hadamar uvádí mnohá svědectví pro to, že myšlení končí tam, kde se vtěluje do slov a sám popisuje své i zkušenosti dalších matematiků (např. Eulera), ale nejen matematiků, s myšlenkovými operacemi vázanými na obrazce nebo na symboly.<sup>3</sup>

Jsme si zde vědomi odlišnosti mezi charakteristikou myšlení a lidského myšlení, což nás zpětně vrací k problematice, výše nastíněné jako Turingovsko–Searleovský spor a také k problematice emergence intencionality. Myšlení tedy charakterizujeme jako obsažnější kategorii, než je tomu u lidského myšlení a otázka zda stroje myslí či nikoli, je převedena na otázku zda myslí stejně jako člověk. Stejně tak bychom mohli uvažovat o shodnosti či rozdílnosti myšlení u člověka a např. u psa, nebo delfína. Je pro nás nepochybné, že jisté operace (do jisté míry je lhostejné, zda je budeme nazývat myšlenkovými či komputačními) probíhají v mozku každého psa i delfína a probíhají také v každém počítači. Jsou ovšem jiné než v mysli člověka. Myšlenkové operace počítače neobsahují, domníváme se, intencionalitu tzn. jejich zaměřenost a zvýznamnění. Kdežto pes má jakousi představu sebe sama jako budoucího, neboť je schopen zahrabat si kost na zlé časy a jeho úkony jsou zaměřeny na něj samotného. Pes také ví, co je to bolest a slast, neboť si vybírá slast a bolesti se vyhýbá, atd. To o počítači zatím říci nelze.

Na druhé straně je však nutné také poukázat na to, že v lidských myslích probíhá nesmírné množství myšlenkových operací a ne všechny jsou lidmi bezprostředně uvědomovány a dokonce některé z myšlenkových operací je patrně velmi obtížné (snad nemožné) přivést k uvědomění. A nemáme zde na mysli jen ty operace lidského myšlení, které se vztahují

<sup>3</sup> Srovnej: Hadamar, J.: *The Mathematician's Mind*. Princeton University Press 1996, p. 72 - 77

k nevědomí (zejména té části nevědomí, kterou C. G. Jung charakterizuje jako nevědomé obsahy, které nikdy nemohou být zvědoměny, a které podle jeho názoru tvoří nejobjemnější část lidské mysli). Jde nám zde například o operace lidského myšlení, při nichž jsou syntetizovány počítky do vjemů; různé neuvědomované kalkuly při rozhodování; ale také např. práci naší mysli při intuitivních průhledech přinášejících bezprostřední porozumění.

Jsou tedy tyto neuvědomované operace stejného druhu jako operace prováděné počítačem? Musí být obsahy mysli uvědomované, aby měly intencionalitu? Tyto otázky nutně vyvstávají, aby vyhranily přístupy k lidskému a strojovému myšlení. Pokusme se tedy nastínit, jak je to s intencionalitou a lidským myšlením a tím i zaujmout stanovisko k tomu, zda musí strojové myšlení disponovat intencionalitou, aby mu byla připsána schopnost myslet.

Pojem má vždy intencionální obsah. To platí jak pro materiální, tak i pro formální pojmy. Intencionalita je vždy zaměřenost na určité jsoucno pro sebe ať již ve formě pozicionality nebo kvazi\_pozicionality; a je také zvýznamněním pojmu lidským myšlením, což znamená, že jde vždy o odkazování pojmu k pojmenovanému jsoucnu pro sebe a zároveň i o vřazení pojmu do struktury souvisejících pojmů v lidské mysli. Intencionalita zároveň vždy poukazuje i na to, jak lidské myšlení konstituuje významy skrze endoceptivní strukturu, v níž jsou obsaženy plány, očekávání, ambice, starost, zkušenost, scénáře a schémata možných konstitucí a explanací jsoucna pro sebe, dlouhodobá paměť, atd. Právě tato endoceptivní struktura zvýznamňuje strukturu pojmů.

Domníváme se, že intencionalita je také úzce vázána na uvědomění toho, jaké místo ve struktuře jazyka pojem zaujímá. Uvědomění v našem pojetí znamená, že lidské myšlení reflektuje to, na jaké místo v jazykové struktuře pojem vřazuje, reflektuje tedy tento proces vřazování a zároveň i to, co je pojmem pojmenováno a jak tedy pojmenované jsoucno pro sebe souvisí s jinými jsoucnými pro sebe, v jakých vzájemných vztazích koexistují. Musí tedy jít o schopnost lidského myšlení rozštěpit se ve dva proudy, které současně operují na dvou úrovních. Jde tedy o uvědomění si vřazení pojmu do určité struktury a zároveň i o reflexi intencionality pojmu, která tímto vřazením emerguje.

Sledujeme-li proces uvědomění (tak nezbytný pro intencionální určení pojmu) z hlediska vazby lidského myšlení na svůj biologický substrát – mozek, pak je možné říci, že jde o excitaci určitých neuronálních

okruhů při percepci, které zároveň excitují jiné neuronální okruhy reflektující (a snad i kontrolující) to, co se při percepci děje. Tyto reflektující neuronální okruhy zároveň registrují, s kterými dalšími neuronálními okruhy jsou excitované percepční okruhy propojeny, a co je v nich obsaženo. Na této rovině reflexivních neuronálních okruhů tedy emerguje intencionalita. Třetí okruh, který jsme v předchozím odstavci charakterizovali jako uvědomění si uvědomování intencionality, pak je možné charakterizovat jako sebezvazek konstituující jáství, tedy subjektivitu. Tedy to, že je zde něco, co toto všechno provádí. Co provádí jak percepci, tak i vřazování percepcí do endoceptivní struktury.

Povšimněme si však blíže jak je intencionalita pojmu vázána<sup>4</sup> na:

- komplexitu
- adekvácií (onomatopoesis) a
- kontexty.

Komplexitou rozumíme vztaženost každého pojmu k jiným

- asociativně,
- analogicky, nebo
- logicko-syntakticky k souvisejícím pojmům a
- narativní struktuře.

Ve velké většině případů jde o asociativní vazby naplňující intencionální obsaženost pojmu. Každý pojem je v asociativní vazbě buď na základě styčnosti, podobnosti nebo kontrastu. Když lidské myšlení asociativně operuje s pojmy, pak přechází z původně myšleného pojmu k jinému, který má s původním myšlenkovým obsahem nějakou souvislost. Lidské myšlení však mohlo přejít úplně jinam, k jinému pojmu s jiným obsahem. Toto spojení však bylo inhibováno. To je způsobováno bohatostí ohraničených (enkapsických) pojmů. Asociativní operace lidského myšlení uvádí pojem do struktury možných stavů a tím vytváří intencionální obsah daného pojmu. Tyto výkony jsou prováděny lidským myšlením automaticky, bez reflexe jednotlivých asociativních operací. Lidské myšlení tedy myslí pojem vždy takto strukturálně, v komplexitě všech jeho možných stavů. A to, že pojem je myšlen v asociativních operacích přičleňuje určitému pojmu intencionalitu, v níž je pojem pro lidské myšlení. Pojem je takto zvýznamněn dalšími souvisejícími, podobnými

<sup>4</sup> Srovnej: Livet, P.: Consciousness as Valued Procedural Mode of Apprehension. In: Consciousness and Intencionality. Kluwer, Boston 1999 pp. 73 - 90

nebo kontrastními pojmy. Vyjevuje se tedy v reflexi ve své komplexitě.

Totéž, co platí o asociacích, platí také o analogiích, neboť i tyto začleňují pojem do strukturních vazeb, i když jinak než je tomu při asociativních vazbách. Jestliže si asociace můžeme připodobnit ke zřetězení a průběžnost operací lidského myšlení v asociacích jako přechody po jednotlivých člancích řetězců (i s tím, že na těchto cestách je mnoho křížovatek), pak analogie je takový postup lidského myšlení, v němž obdobnost a svázanost pojmu má podobu celé plochy, pole, i když přece jen v mnohem určené ohraničené. Tato určenost může být ve vlastnostech a vztazích, ale také v nutné a opakovatelné následnosti, ve způsobech vysvětlení, nebo se může také týkat dynamiky vývoje toho, jaké *capty*<sup>5</sup> lidské myšlení vytrhne ze jsoucen a jaké endoceptivní obsahy jsou uchovány v mysli.

Při analogických myšlenkových operacích poznání vlastností a vztahy určitých pojmů přenášíme i na jiné pojmy, které tím nabírají na významnosti, získávají intencionalitu. Analogie působí jako generátor intencionality, když zahrnuje co největší množství podstatných souvislostí. V takovém případě může určitá poznání skutečnost ozvučit (reverberate) nereflektované související operace lidského myšlení a tím naplnit intencionální obsah pojmu. K tomu však je nutné, aby analog:

- byl co nejčtetnější a zároveň se týkal podstatných vlastností pojmu;
- z toho plyne také nutnost těsného spojení mezi vlastnostmi pojmu a analogu;
- je nutné brát v úvahu také vše, čím se pojem a analog odlišují.

Třetím faktorem komplexity podílejícím se na intencionalitě pojmu je jeho logicko-syntaktická vázanost. Zde jsou rozhodující souvislosti logické, kdy pojem může být jen v limitovaných vztazích k jiným pojmům a tato limitace je dána logickými operacemi lidského myšlení.

Omezenosti se v tomto smyslu projevují nejvíce, operuje-li lidské myšlení na idealizační úrovni. A totéž, co platí o omezenosti vazby konstituce intencionality při tvorbě pojmu v logických vztazích platí rovněž i ve vazbách syntaktických. Také tyto vazby vymezují významovost pojmu. Lze to uvést mnoha příklady, které jsou založeny na homonimech, na jejichž základě je postaveno mnoho

<sup>5</sup> Termín: *capty* (odvozeno z *capture*) přebíráme od P. Checklenda z jeho pražské přednášky.

anekdot. Mnohdy však význam může být určen dokonce interpunkcí nebo intonací.

Adekvácie (onomatopoesis) vymezuje intencionalitu především shodností znění s obsahem daného slova. Tím se zároveň zpřesňuje intencionální obsah pojmu. Adekvácií lze ilustrovat řadou příkladů. Např. slovo *skomírání* vykazuje svůj intencionální obsah tím, že asociuje slova: *skoro* a *umírání*. Toto vykazování není bezprostřední, musí být domyšleno a toto domyšlení a tím i propátrávání intencionálních obsahů myslí zároveň dokresluje intencionální obsah původního slova (pojmu). Obdobně je tomu i u mnoha dalších slov jakými jsou *stanovisko*, *názor*, atd. Adekvácie navozuje hlavní asociativní vazby a tím omezuje mlhavost intencionality.

Kontext se rovněž výrazným způsobem spolupodílí na tvorbě intencionality. Kontext zahrnuje nejen komplexitu při tvorbě pojmu, ale širší pole, v němž se na konstituci významu podílí situace, v níž se lidské myšlení nachází při použití konkrétního pojmu. Ve filosofické literatuře je tato problematika iniciovaná a i když fragmentárně, přece jen na mnoha příkladech analyzovaná L. Wittgensteinem a nemáme k ní příliš co dodat. Lidské myšlení je vždy v nějaké situaci, neboť jeho pohybu vždy něco předcházelo a něco vždy následuje a tyto retence a pretence se spolupodílí na intencích.

Výše uvedený výčet konstitutivních charakteristik intencionality pojmu není úplný. Intencionalita je velmi složitou problematikou a je obtížně stanovitelné jakými prostředky se k ní lidské myšlení dopracovává. Každopádně její emergence je závislá, domníváme se, také na výše uvedených momentech.

### 3 Lidské a strojové myšlení

Je však intencionalita nutností pro strojové myšlení a kam až může strojové myšlení dojít, postrádá-li intencionalitu. Vytýčený problém se pokusíme artikulovat prostřednictvím Hegelova uchopení panství a rabství, jakožto něčeho, skrze co se vyjevuje také vztah člověka a umělé inteligence. Zastavme se nejprve u Hegelova textu<sup>6</sup>.

*Sebevědomí je zprvu jednoduché bytí pro sebe, soběrovné tím, že všechno jiné vylučuje ze sebe: jeho bytností a absolutním předmětem je mu já: a v této bezprostřednosti či v tomto bytí svého bytí pro sebe je čímsi jednotlivým. Co je pro ně jiné, jakožto něco postrádajícího bytnosti, znamenkem zápornosti označený předmět. Avšak toto jiné*

<sup>6</sup> Hegel, G.F.W.: *Fenomenologie ducha*. Praha, ČSAV 1960 s. 154 an.



je rovněž sebevědomí: individuum vystupuje proti individuu. Když takto bezprostředně vystupují, jsou pro sebe ve způsobu obyčejných předmětů" samostatné podoby, vědomí vnořená do bytí života - neboť předmětné bytí se zde určilo jakožto život - která vůči sobě ještě nevykonala pohyb absolutní abstrakce, jehož mocí vylučují ze sebe veškeré bezprostřední bytí a jsou pouze čistě negativní bytí soběrovného vědomí, či která se ještě sobě navzájem nepředstavila jakožto čistě bytí pro sebe, tj. jako sebevědomí. Každé je si jisto sebou samým, ale nikoli druhým: a proto nemá jeho vlastní jistota o sobě ještě pravdu: neboť jeho pravda by byla jen v tom, aby se mu jeho vlastní bytí pro sebe představilo jako samostatný předmět, či což je totéž, kdyby se mu předmět představil jako tato čistá jistota o sobě samém. To však není podle pojmu uznávání možno jinak, než že jako druhý pro něho, tak on pro druhého, každý o sobě a svým vlastním konáním a opět konáním druhého provádí tuto čistou abstrakci bytí pro sebe

Podání sebe jakožto čisté abstrakce sebevědomí záleží však v tom, že sebevědomí se ukazuje jako čistý zápor svého předmětného způsobu, čili v tom, že se ukazuje, že není připoutáno k žádnému určitému jsoucnu, k žádnému obecnému jednotlivu jsoucnu vůbec, že není připoutáno k životu. Toto podání je dvojnásobné konání: konání jiného a konání, jež vychází ze sebe sama. Pokud je to konání druhého, je tedy cílem každého smrti toho druhého. V tom je však obsaženo to druhé, totiž konání z vlastní moci: neboť to první konání obsahuje v sobě to, že dáváme všanc vlastní život. Vztah obojího sebevědomí je tedy určen tak, že sebe sama, a to navzájem mezi sebou, osvědčují bojem na život a na smrt. - Musí podstoupit tento boj, neboť jistotu sebe samých, že jsou pro sebe, musí na druhém a na sobě samých povznést k pravdě. A je to výhradně nasazení života, jímž se osvědčuje svoboda, jímž se osvědčuje, že pro sebevědomí není bytostí bytí, bezprostřední způsob, v němž vystupuje, jeho vnoření do šíře života - nýbrž že na něm není nic, co by pro ně nebylo zanikajícím momentem, že je pouze čistým bytím pro sebe....

Tato zkušenost ukazuje sebevědomí, že život je pro ně tak bytostný jako čistě sebevědomí. V bezprostředním sebevědomí je jednoduché já absolutním předmětem, který je však pro nás čili o sobě absolutním zprostředkováním a jehož bytostným momentem je trvajícím samostatnost. Rozklad oné jednoduché jednoty je výsledkem první zkušenosti: tato zkušenost klade čistě sebevědomí a vědomí, které není čistě pro sebe, nýbrž pro jiné, tj. je jsoucí vědomí, čili vědomí v podobě věčnosti. Oba momenty jsou bytostně nutné: - poněvadž si jsou zprvu nerovné a navzájem protikladné a poněvadž k jejich reflexi v jednotu dosud nedošlo, jsou to dvě

protikladné formy vědomí: jedna je samostatné vědomí, kterému je bytostí bytí pro sebe, druhá vědomí nesamostatné, jemuž je bytostí život čili bytí pro jiné: první je pán, druhá rab.

To, co zde Hegel podává, naznačuje jak tomu je se vzájemným vztahem dvou sebevědomí, která se osvědčují teprve v boji na život a na smrt. Jak tomu však bude ve vztahu člověk – stroj? Tvrdím, že budeme-li konstruovat stroje jako raby, pak se budeme zároveň snažit o to, aby měly vědomí a tedy i intencionalitu; avšak budeme-li konstruovat stroje jako pány, pak tyto nepotřebují vědomí, ani intencionalitu.

Možná, že pro někoho tato úvaha zní poněkud zmateně, ale jsem přesvědčen, že je hluboce zakofeněná ve výše uvedené pasáži. Rabství je totiž vždy vědomím, i když ještě nevykonalo pohyb absolutní abstrakce. Rabství usiluje o uznání a potvrzení a jako takové musí být konstruováno tak, aby disponovalo vědomím. Rabství se sice potvrzuje vůči věcem, které poznamenává znaménkem zápornosti, ale nikoli absolutně, protože nepodstoupilo boj na život a na smrt s jiným sebevědomím, ještě se nepředstavilo jako čistě bytí pro sebe. Také z hlediska pána je nutné, aby rabství mělo vědomí, protože právě v tomto a jedině v tomto boji může být panství potvrzeno jako čistě bytí pro sebe.

Ale stroje konstruované jako panství (i když ne plně v Hegelově smyslu) vědomí nepotřebují. Je totiž možné konstruovat je bez ohledu na to zda potlačí jiné bytí a tak dospějí k potvrzení sebe sama. Strojům, jsou-li tak zkonstruovány, nezáleží na tom, zda jsou potvrzeny; neví o takovéto ambici a nepotřebují ji k tomu, aby byly funkční. Ke své funkčnosti stroje nepotřebují vědomí a intencionalitu, ale postačí jim mapa prostředí, pro které jsou určeny a algoritmy rozhodovacích procesů vzhledem k prostředí a zastávaným funkcím.

V tomto ohledu lze očekávat od strojů velmi mnoho. Antropocentrický pohled, který staví člověka, vědomí, intencionalitu na nejvyšší stupínek je oprávněný jako současný evoluční pohled. Je obtížné předvídat, jak dlouhou dobu tento pohled může vydržet; zda bude prioritním pohledem vždy, pokud zde bude lidstvo; nebo zda i tehdy, když zde bude lidstvo, bude lidmi uznáno, že vědomí a intencionalita jsou jen jednou z možných forem a vazeb existence, případně že jde o takovou formu, která nemusí být nejdokonalejší a nejperspektivnější, nej, nej...

Před konstrukci chytrých strojů mnozí badatele, kteří se na jejich tvorbě podílejí, varují. Jejich varování jde především tím směrem, že sestrojíme-li intelekt

mnohonásobně převyšující lidský intelekt, nemusí být návratu k naší superioritě na zemi. H. de Garis např. píše o schopnostech bilionkrát převyšující lidské schopnosti.<sup>7</sup> A i když mnohé z takovýchto vizí je nutné brát rezervovaně, přesto dynamika vývoje chytrých strojů na jedné straně a snadnost z níž lidé delegují mnohé ze svých schopností a dovedností na tyto stroje na straně druhé, naznačují, že tyto oblasti není možné ignorovat.

Člověk vyzvedává vědomí a intencionalitu jako něco výlučně lidského a nenapodobitelného. To však neznamená, že tyto vlastnosti jsou neobejitelné a v tom může být velká síla strojů. Domníváme se, že chytré stroje se mohou rozvíjet ve značně variabilních formách, aniž budou sledovat nějaký účel a aniž budou jakkoli zohledňovat účely člověka. Zejména tehdy, bude-li jistý stupeň komputačních či myšlenkových operací fixován na nano úrovni, např. v podobě jednoduchých pravidel orientujících částice ke sdružování a kooperaci. Lze si to představit jako evoluci bez člověka. Teprve člověk skrze sebe rozpoznává evoluci a v evolučních procesech souvislosti a směr. Jde tedy o realizace možností, v nichž se chytré stroje mohou projevit. Některé z těchto projevů mohou mít vztah k jiným koexistujícím formám jsoucen, např. k člověku, ale není to pro ně podmínkou. Nelze v tomto případě také hovořit o nižších a vyšších formách, které jsou zatížeností panstvím a rabstvím a ospravedlňujícími aktivy vyšších vůči nižším. Dynamika možného tak může přinést pro člověka mnohá překvapení.

## Literatura

- [1] de Garis, H.: Artilect War <http://www.cs.usu.edu/~degaris/artilectwarshort.html>
- [2] Hadamar, J.: The Mathematician's Mind. Princeton University Press 1996
- [3] Hegel, G.F.W.: Fenomenologie ducha. Praha, ČSAV 1960
- [4] Livet, P.: Consciousness as Valued Procedural Mode of Apprehension. In: Consciousness and Intencionality. Kluwer, Boston 1999
- [5] Searle, J.C.: Mysl, mozek a věda. Praha, Mladá fronta 1994
- [6] Turing, A.M.: Počítacie stroje a inteligencia. In: Mysel/Telo/Stroj. Bratislava, Bradlo 1992

---

<sup>7</sup> Srovnej: de Garis, H.: Artilect War In: <http://www.cs.usu.edu/~degaris/artilectwarshort.html>

## Robotický model systému zrkadliacich neurónov: experimentálna analýza

Kristína Rebrová a Igor Farkaš

Centrum pre kognitívnu vedu, KAI FMFI UK, Univerzita Komenského v Bratislave

Mlynská dolina, 84248 Bratislava

Email: {kristina.rebrova,farkas}@fmph.uniba.sk

### Abstrakt

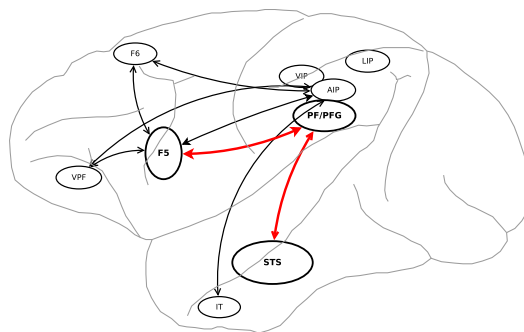
V príspevku predstavujeme model systému zrkadliacich neurónov vyvinutý v rámci kognitívnej robotiky použitím simulátora robota iCub. Pri formovaní modulárnej architektúry nášho modelu sme navrhli algoritmus obojsmerného učenia odvodený od biologicky plauzibilného modelu GeneRec. V článku prezentujeme výsledky testov zameraných na vlastnosti nášho algoritmu a jeho úspešnosť na rôznych dátových množinách. Náš model narozdiel od klasických modelov systému zrkadliacich neurónov, neuvažuje predspracované, pozične invariantné informácie. Naopak, cieľom nášho modelu je vysvetliť a realizovať vznik invariantných reprezentácií vo vyšších motorických a perceptuálnych mozgových oblastiach. V príspevku uvádzame predbežné výsledky z nášho modelu, konkrétne vznik asociácie medzi vizuálnymi a motorickými reprezentáciami z vlastnej perspektívy.

### 1 Úvod

V tomto príspevku priamo nadväzujeme na príspevok [16], v ktorom sme načrtli problematiku variantných a invariantných zrkadliacich neurónov. Naším cieľom je navrhovať a testovať konceptuálny model systému zrkadliacich neurónov v rámci paradigmy kognitívnej robotiky. V nasledujúcom texte zhrnieme podstatné východiská a návrh nášho modelu. Ďalej sa bližšie povenujeme použitým algoritmom, konkrétne obojsmernému učeniu na báze aktivačných fáz (bidirectional activation-based learning, BAL) [15].

Zrkadliace neuróny pália nielen pri vykonávaní, ale aj pri pozorovaní akcií (cielených motorických úkonov). Predpokladá sa, že pri porozumení pozorovanej akcie sa zapája systém zrkadliacich neurónov (mirror neuron system, MNS)[18], fronto-parietálny neurálny okruh zodpovedný za párovanie pozorovaných akcií a vlastného motorického repertoára. Základom MNS u makakov sú tri mozgové oblasti: F5 v premotorickej kôre, PF a AIP v temenom laloku, vid' Obr. 1. Unikátnu úlohu v MNS hrá horná spánková brázda (superior temporal sulcus, STS), ktorá podobne ako F5, reaguje na biologické pohyby, ale len v prí-

pade vizuálnej stimulácie. Tým, že oblasť STS nemá multimodálne vlastnosti, nemôže byť považovaná za skutočnú súčasť MNS. Na druhej strane, STS slúži MNS ako hlavný zdroj informácií.



**Obr. 1:** Náčrt systému zrkadliacich neurónov u makakov. Prepojenia medzi kľúčovými oblasťami sú farebne označené.

STS má samo o sebe zaujímavú štruktúru a vlastnosti. Ako uvádzajú rôzne výskumy Perreta a kolegov [14] či [5], STS obsahuje neuróny rôznych typov, medzi nimi množstvo takých, ktoré reagujú na vizuálny vnem objektu (tváre), či deja (pohybu) len z určitej perspektívy. Obsahuje ale aj neuróny, ktoré sú vzhľadom na perspektívu invariantné (reagujú na rôzne uhly pohľadu), čo znamená, že kódujú širšiu, všeobecnejšiu kategóriu. V STS sa takéto neuróny nachádzajú anatomicky najbližšie k frontálnej časti kôry, ktorá kóduje koncepty na najvyššej úrovni a považuje sa za racionálnu časť mozgu. Nedávny výskum [2] ukázal, že podobné vizuálne vlastnosti ako variantné neuróny v STS vykazujú aj zrkadliace neuróny v oblasti F5, a to väčšina z nich (asi tri štvrtiny). Caggiano a kolegovia skúmali tri druhy pohľadov, a to spredu (otočenie subjektu o 180° voči agentovi), z boku (90°) a z vlastnej perspektívy a objavili variantné rovnako ako invariantné zrkadliace neuróny. Prítomnosť vizuálne invariantných zrkadliacich neurónov v F5 naznačuje, že v F5, rovnako ako v STS, dochádza k hierarchickému usporiadaniu neurónov ústiacemu do abstraktného spracovania pozorovanej akcie, čo pravde-

podobne uľahčuje porozumenie pozorovanému deju.

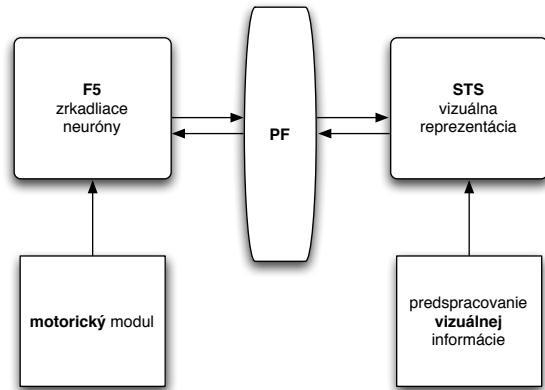
Spolu s objavom zrkadliacich neurónov a ich vlastností vzniklo množstvo výpočtových modelov systému zrkadliacich neurónov. Tieto modely zväčša modelujú presné časti mozgu opíc (či ľuď) a slúžia na konceptuálnu podporu teórií o motorickom porozumení pozorovanému deju. Najznámejšie z nich pochádzajú z dielne Arbiba, Oztopa a ďalších, napríklad FARS [3], MNS1 [10], MSI [12], MNS2 [1] a iné (pozri [11]). Oztop a kolegovia [13] tvrdia, že práve výpočtové modely nám ponúkajú silný nástroj na skúmanie teórií o zrkadliacich neurónoch. Niektoré výpočtové modely MNS možno zaradiť aj do oblasti kognitívnej robotiky. Takéto modely využívajú vlastnosti zrkadliacich neurónov aby obohatili správanie robotického agenta o nové funkcie. Vhodným príkladom je RNNPB [20]. Náš model patrí do tejto skupiny.

Mnohé klasické modely MNS čiastočne či úplne obchádzajú problém transformácie perspektívy, z ktorej agent pozoruje dej. Model Tessitoreho a kolegov [21] priamo rieši problém interpretácie variantných vstupov. Základom modelu je premisa, že motorické akcie možno popísať na základe malého množstva parametrov popisujúcich natočenia kĺbov ruky pri vykonávaní konkrétneho typu uchopenia predmetu. Na rozdiel od tradičných modelov predpokladá, že tok informácie medzi zrkadliacimi neurónmi a oblasťou STS je obojsmerný, a že informácia z motorickej modalít obohacuje a urýchľuje interpretáciu vstupov vo vizuálnych oblastiach. V našom modeli zdôrazňujeme obojsmerný tok informácie medzi motorickými (F5) a vizuálnymi (STS) oblasťami. Naším cieľom je tiež obsiahnuť potenciál pre vytvorenie variantných a invariálnych reprezentácií. V súčasnom stave nášho výskumu však prezentujeme výsledky z fungovania nášho modelu len pre jednu perspektívu, konkrétne pre vlastnú perspektívu ( $0^\circ$ ).

## 2 Model systému zrkadliacich neurónov

Náš robotický model systému zrkadliacich neurónov (Obr. 2) pozostáva z niekoľkých hierarchicky zapojených modulov. Na najvyššej úrovni sa nachádza abstrakcia neurálneho okruhu STS–PF–F5, ktorý obojsmerne spája sekundárnu perceptuálnu (vizuálnu) a motorickú informáciu. Vizuálna a motorická informácia sa do modelu dostáva z exekutívnych modulov, konkrétne modulu pre uchopovanie a vizuálneho modulu, ktorý nie je tvorený neurálnou architektúrou, iba poskytuje systému predspracované vizuálne dáta. Nosné moduly F5 a STS reprezentujú motorické a vizuálne aspekty akcií a sú tvorené samoorganizujúcimi sa mapami. Možno ich považovať za paralelu asociatívnych oblastí v mozgu. Pre implementáciu okruhu STS–PF–F5 sme navrhli algoritmus obojsmerného učenia na báze aktivačných fáz (bidirectional activation-based learning, BAL),

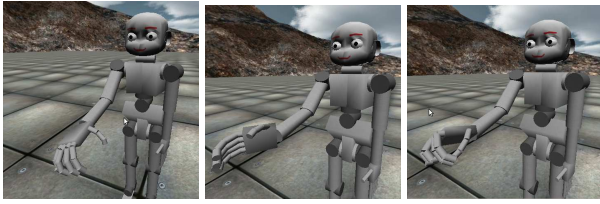
ktorý je modifikáciou biologicky plausibilného učenia širším chyby GeneRec [9].



**Obr. 2:** Schéma robotického modelu zrkadliacich neurónov.

Pre realizáciu modelu sme použili dáta zo simulovaného robota iCub [8, 22] natrénovaného na uchopovanie predmetov tromi rôznymi spôsobmi podľa tvaru a veľkosti objektu [25] zobrazenými na Obr. 3. Akcie, ktoré robot vykonával boli silový úchop (power grasp), bočný úchop (side grasp) a jemný úchop (precision grip). Neurálny model pre uchopovanie sa učil na báze biologicky plausibilného učenia s posilňovaním, konkrétne jeho spojitou variantou CACLA [23]. Podobnú neurálnu architektúru trénovanú pomocou CACLA popisujeme napríklad aj v [4]. Tejto metodológii sa venuje aj článok *Konekcionistický model uchopovania: experimenty v reálnom prostredí* v tomto zborníku. Pre účely trénovania vyšších častí našej modulárnej architektúry sme zaznamenali a preškálovali hodnoty 16 kĺbov používaných pri natáhovaní sa a uchopovaní. Vizuálne dáta pre STS modul sme vytvorili pomocou projekcie vybraných stupňov voľnosti a pozícií (spolu 40 hodnôt) na pravú kameru robota. Ďalšie tri perspektívy ( $90^\circ$ ,  $180^\circ$ , and  $270^\circ$ ) sme vytvorili rotáciou motorických dát pre prvú perspektívu a následným premietnutím do 2D a preškálovaním.

Moduly STS a F5 predstavujú zjednodušenú funkčnú aproximáciu vizuálnej oblasti STS a motorickej oblasti F5 (u makakov) a k nej homologickým oblastiam ľudskeho mozgu. Modelujeme ich samoorganizujúcimi sa mapami s rekurentným kontextom, konkrétne MSOM [19]. MSOM alebo Merge SOM, čiže zlučovacia samoorganizujúca sa mapa, je klasická Kohonenova sieť [7], ktorá má navyše kontextovú vrstvu ako tradičná rekurentná sieť. Na rozdiel od ďalších rekurentných SOM modelov (napr. RecSOM či SOMSD, pozri porovnanie v [24]), má MSOM kontextovú vrstvu iba o veľkosti vstupu. Do tejto vrstvy



**Obr. 3:** Príklady troch typov úchopov zobrazené z perspektívy pozorovateľa. Z ľava do prava: silový úchop (power grasp), bočný úchop (side grasp) a jemný úchop (precision grip).

sa teda vkladá menšie množstvo informácie, konkrétne ide o lineárnu kombináciu váhy a kontextu aktuálneho (teda posledného) víťazného neurónu. Kvôli priestorovým nárokom nechávame na čitateľa, aby sa s matematickou definíciou MSOM zoznámil pri čítaní nášho predošlého článku [16] alebo [17] (pre detailnejšie informácie o vlastnostiach a optimálnych parametroch pre tento typ siete pozri [24]).

Obidva MSOM moduly sme trénovali na dátach zo simulovaného iCub-a. Motorický modul F5 dostával motorické vstupy, konkrétne 18 inštancií každého z troch pohybov ilustrovaných na Obr. 3. Vzhľadom na to, že vo finálnej podobe má náš model reagovať na 4 rôzne perspektívy, vizuálny modul STS sme trénovali na všetkých 12 kategóriách dát (4 perspektívy pre 3 úchopy) a to pre korešpondujúce motorické inštancie. V prvom rade sme hľadali optimálne parametre pre siete na základe [24], viac informácií sa nachádza v [17]. Výsledky experimentov s MSOM naznačujú, že pri vizuálnych dátach je perspektíva vplyvnejšia než typ úchopu. Výsledky z optimálnych máp sme zaznamenali a pripravili ako vstup pre najvyšší modul.

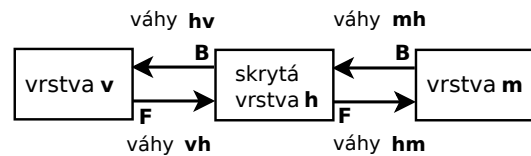
Vzhľadom na charakter a výpočtovú náročnosť našej úlohy sme sa rozhodli výstupy z MSOM pred použitím „zbinarizovať“ a to tak, aby výsledné reprezentácie tvorili takzvané riedke (sparse) neurálne kódy. Riedke distribuované reprezentácie (tiež známe pod termínom „population coding“) sú charakteristické robustnými odpoveďami na rozne podnety a okrem toho dávajú priestor na zovšeobecňovanie (keďže podobné vstupy vyvolávajú podobné riedke reprezentácie). V našom prípade majú reprezentácie vďaka samoorganizujúcim sa mapám tiež topografické usporiadanie, čo je často využívaný princíp organizácie v mozgu. Riedke reprezentácie vznikajú v biologických sieťach na základe laterálnej inhibície. My sme ako výpočtovú skratku použili algoritmus  $k$ -WTA ( $k$ -winner-take-all), ktorý priradí hodnotu 1.0  $k$  neurónom s najmenšou vzdialenosťou od vstupu a ostatné neuróny zresetuje na 0.0.

V súčasnom stave sme naimplementovali BAL pre asociáciu motorických a perceptuálnych reprezentácií pri

tom, ako agent sám vykonáva akciu a pozoruje sám seba. V nasledujúcej kapitole popíšeme náš algoritmus, ďalej popíšeme jeho vlastnosti a výsledky zo všeobecných testov a nakoniec odprezentujeme aktuálne výsledky z nášho robotického modelu.

### 3 Algoritmus BAL

Pri výbere učiaceho algoritmu pre asociáciu vyšších motorických a vizuálnych oblastí sme dbali na biologickú plauzibilitu. Algoritmus BAL je založený na zovšeobecnenej recirkulácii (Generalized Recirculation, GeneRec) [9]. Tento učiaci algoritmus síce spadá do kategórie učenia s učiteľom, no na rozdiel od štandardného šírenia chyby funguje na báze rozdielu medzi dvoma aktivačnými stavmi siete, plusovou a mínusovou. V mínusovej fáze sa sieť šíri aktivácia od vstupu k výstupu a zaznamenáva sa odhad siete na výstupe. Tento proces je schematicky zobrazený na Obr. 4. V druhej, plusovej fáze, sa na výstup sa „zafixuje“ (angl. clamp) želaný vzor (target) a aktivácia sa šíri naopak (od vstupu k výstupu). O’Reillyho model, pôvodne určený na jednosmernú asociáciu vstupov a výstupov, sme adaptovali tak, aby mohol slúžiť na obojsmernú asociáciu. Viac o GeneRec modeli a BAL v [15].



**Obr. 4:** Schéma obojsmerného učenia na báze aktivačných fáz.

Keďže v našom prípade figurujú vstupné dáta a výstupné odhady siete na oboch vstupno-výstupných vrstvách, nazveme tieto  $v$ , teda vizuálna vrstva (STS) a  $m$  teda motorická vrstva (F5). Skrytá vrstva  $h$  figuruje v našom modeli ako oblasť, ktorá prepája STS a F5, teda PF. Napriek tomu, že na tejto vrstve nutne neočakávame vznik neurálneho kódovania, zaujíma nás, ako sa bude jej aktivita vyvíjať (na túto aktiváciu totiž nemáme žiaden priamy vplyv, reprezentácie na tejto vrstve sú teda úplne samoorganizované).

Náš algoritmus pre obojsmernú asociáciu sa odvíja od GeneRec modelu, no tiež sa od neho líši. Náš model je perceptrón so 4 maticami váh, 2 pre každý smer toku aktivácie. Výpočet aktivácie na jednotlivých vrstvách v dvoch aktivačných fázach zobrazuje Tabuľka 1. Najprv aktivujeme vstupy na vizuálnej vrstve  $v$  smere  $v^F \rightarrow h^F \rightarrow m^F$  a získavame aktivácie na vrstvách  $v$  tomto smere, ktorý my označujeme písmenom F (zodpovedá mínusovej fáze)<sup>1</sup>. V

<sup>1</sup> V našom prípade máme vstupy aj výstupy na oboch stranách. Rozli-

**Tabuľka 1:** Aktivačné fázy a stavy v modeli BAL. Obe aktivačné fázy sa skladajú z dvoch krokov, pričom vždy na začiatku aktivujeme vzor na vstupe a dostávame odhad siete na výstupe. Vrstvy  $\mathbf{v}$  a  $\mathbf{m}$  si navzájom striedajú úlohu vstupnej a výstupnej vrstvy.

Vrstva	Fáza	Vstup (net)	Aktivácia
$\mathbf{v}$	F	-	$v_i^F$ (vstup)
$\mathbf{h}$	F	$\eta_j^F = \sum_i w_{ij}^{vh} x_i^F$	$h_j^F = \sigma(\eta_j^F)$
$\mathbf{m}$	F	$\eta_k^F = \sum_j w_{jk}^{hm} h_j^F$	$m_k^F = \sigma(\eta_k^F)$
$\mathbf{m}$	B	-	$m_k^B$ (vstup)
$\mathbf{h}$	B	$\eta_j^B = \sum_k w_{kj}^{mh} y_k^B$	$h_j^B = \sigma(\eta_j^B)$
$\mathbf{v}$	B	$\eta_i^B = \sum_j w_{ji}^{hv} h_j^B$	$v_i^B = \sigma(\eta_i^B)$

nasledujúcej fáze zafixujeme na vstup želaný výstup, ktorý sieť odhadovala v prvej fáze. V tomto prípade už o ňom uvažujeme ako o vstupe pre spätný smer, teda  $\mathbf{m}^B \rightarrow \mathbf{h}^B \rightarrow \mathbf{v}^B$  označený písmenom B (plusová fáza). Výstup zo spätného smeru reprezentuje odhad siete o vstupe  $\mathbf{v}^F$ .

Na základe rozdielov v aktivačných fázach upravujeme váhové matice v sieti. Všeobecný predpis pre úpravu váh v doprednom F smere popisuje rovnica 1 a úpravu váh v spätnom smere rovnica 2, kde  $\lambda$  je rýchlosť učenia. Všetky vrstvy obsahujú trénovateľné biasové neuróny, ktorých váhy sa upravujú analogicky k ostatným neurónom a ako vstup dostávajú konštantne hodnotu 1.0.

$$\Delta w_{ij}^F = \lambda \cdot a_i^F (a_j^B - a_j^F) \quad (1)$$

$$\Delta w_{ij}^B = \lambda \cdot a_i^B (a_j^F - a_j^B) \quad (2)$$

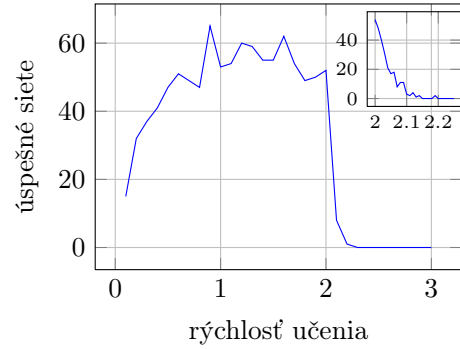
## 4 Experimenty s algoritmom BAL

Pri skúmaní vlastností nášho algoritmu sme sa venovali viacerým testom na rôznych typoch umelých dát. Pre vyhodnotenie úspešnosti nášho učiaceho algoritmu sme zvolili tri kvantitatívne miery. Konkrétne sme vyhodnocovali: (1) bitovú úspešnosť (označenie „bity“), čiže percento bitov, v ktorých sa výstup siete a želaná hodnota zhodujú (priemer zo všetkých bitov všetkých vzorov), (2) vzorovú úspešnosť (označenie „vzory“), teda percento vzorov, v ktorých mala sieť 100% bitovú úspešnosť, (3) strednú kvadratickú chybu (mse). Miery úspešnosti siete sme vyhodnocovali zvlášť pre smer F a pre smer B. Vo všetkých experimentoch sme inicializovali váhy na malé náhodné hodnoty z normálnej distribúcie.

### 4.1 Kódovač 4-2-4

Pre porovnanie s pôvodným GeneRec modelom sme náš model testovali pomocou známej úlohy „kódovača“ 4-2-4,

šujeme preto medzi smermi aktivácie, nie medzi odhadovacou mínusovou a zabezpenou plusovou fázou ako O'Reilly.



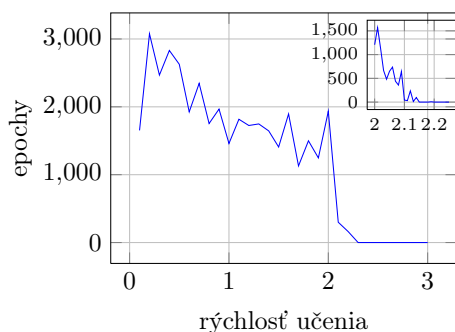
**Obr. 5:** Kódovanie 4-2-4: počet úspešných sietí ako funkcia rýchlosti učenia. Zaujímavé je, že napriek tomu, že sa náš model sa učí pre rôzne rýchlosti učenia, konvergencia algoritmu pre  $\lambda > 2.0$  prudko poklesne na nulu. Ak teda zvolíme príliš veľkú rýchlosť učenia, algoritmus neskonverguje vôbec. Výsledky pre detailnejší interval okolo kritickej hodnoty zobrazuje vnorený graf.

ktorý sa učí asociovať 4-bitové vstupy cez 2 skryté neuróny (kompresia). V našom prípade sa sieť učí rovnaké zobrazenie z oboch strán. Pri testovaní sme zistili, že sieť trénovaná pomocou BAL pri tejto úlohe niekedy vôbec neskonverguje k 100% úspešnosti. Pri testovaní parametrov, konkrétne rýchlosti učenia ( $\lambda$ ), sme teda vyhodnocovali počet sietí, ktoré skonvergujú (zo 100) a priemerný počet trénovacích epoch, ktoré siete potrebujú pre dosiahnutie 100% úspešnosti. Výsledky uvádzame na Obr. 5 a Obr. 6. Zaujímavé je, že úspešnosť sietí veľmi prudko poklesne po prekonaní  $\lambda = 2.0$ . Tomuto fenoménu sa venuje vnorený graf s detailnejším intervalom  $\lambda$  od kritickej hodnoty 2.0.

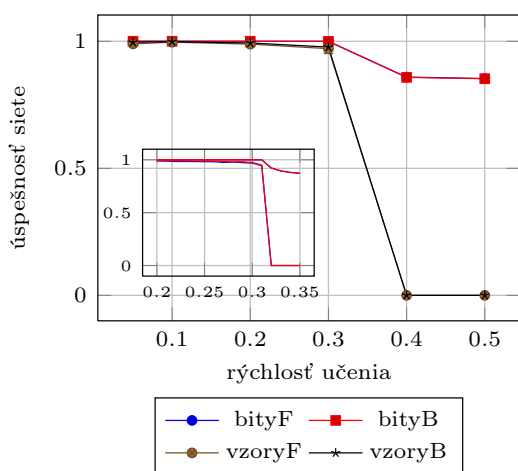
Podobné kritéria ako my mal pri zhodnotení svojho algoritmu a jeho variácií aj O'Reilly [9], ktorý zaznamenal konvergenciu 90%. Pri najúspešnejšej rýchlosti učenia  $\lambda = 0.9$  konvergoval BAL v 65% prípadov, čo je porovnateľné so symetrickou verziou GeneRec aproximujúcov kontrastné Hebovské učenie (Contrastive Hebbian Learning, CHL). Tento a ďalšie experimenty detailnejšie popisujeme v [15].

### 4.2 Párovanie riedkych binárnych vzorov

V snahe skúmať algoritmus BAL pomocou dát, ktoré sa ponášajú na riedke (sparse) binárne reprezentácie, aké v našom robotickom modeli párujeme (pozri časť 2) sme vytvorili dátové množiny z náhodných binárnych čísel. Naše binárne vstupné vektory obsahujú vždy presný počet  $k$  pozitívnych bitov. V tomto prípade sme použili 100 rôznych párov vektorov veľkosti 144 bitov s  $k = 12$  pozitívnymi bitmi. Tak, ako v predošlom experimente, sme tes-

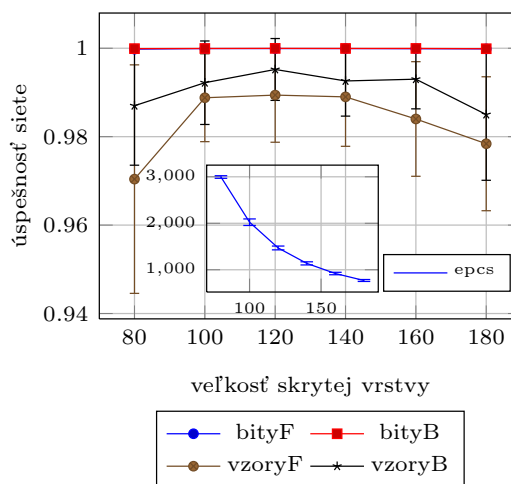


**Obr. 6:** Kódovanie 4-2-4: priemerný počet tréningových epoch ako funkcia rýchlosti učenia pre úspešne konvergujúce siete.



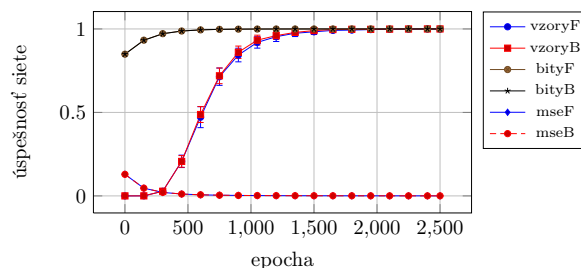
**Obr. 7:** Párovanie binárnych vzorov: úspešnosť siete ako funkcia  $\lambda$ .

tovali rôzne  $\lambda$  a tiež rôzne veľkosti skrytej vrstvy  $n_H$ . V tomto prípade sa k určitej úspešnosti dostanú všetky sledované siete (50 pre každú sledovanú hodnotu parametrov). Z priestorových dôvodov zaznamenávame len 2 miery úspešnosti, a to vzorovú a bitovú. Výsledky zobrazuje Obr. 7 a Obr. 8. Ukázalo sa, že v tomto prípade stačí oveľa nižšia  $\lambda$  a to okolo hodnoty 0.2. Tiež sa tu vyskytuje prudký pokles úspešnosti sietí (pozri vnorený náhľad na detailný interval v oblasti kritických hodnôt). Čo sa týka  $n_H$ , zistili sme, že má vplyv viac menej len na počet epoch potrebných na skonvergovanie siete k úspešnosti (pozri vnútorný graf na Obr. 8). Pri tomto experimente uvádzame taktiež ilustráciu priebehu učenia na Obr. 9. Pri skúmaní vlastností nášho modelu sme sa zaoberali aj aktiváciami na skrytej vrstve. Zistili sme, že  $\mathbf{h}^F$  a  $\mathbf{h}^B$  majú tendenciu blížiť sa k sebe až sú nakoniec takmer totožné. Čo sa týka charakteru reprezentácií na  $\mathbf{h}$ , hodnoty aktivácií neurónov sa pohybujú



**Obr. 8:** Párovanie binárnych vzorov: úspešnosť siete ako funkcia  $n_H$ .

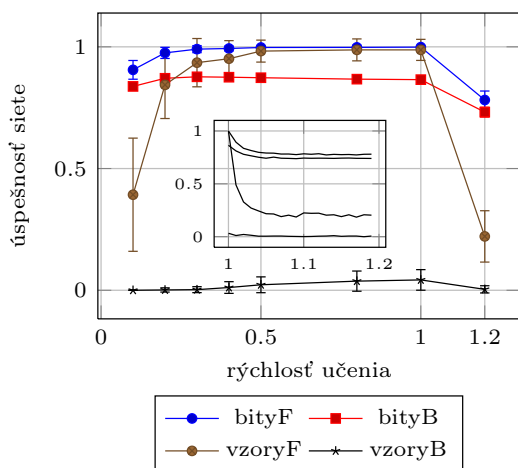
okolo 0.5, čiže nedochádza k binarizácii. Vlastnosti aktivácií na skrytej vrstve sú vhodným subjektom pre ďalšie skúmanie algoritmu BAL.



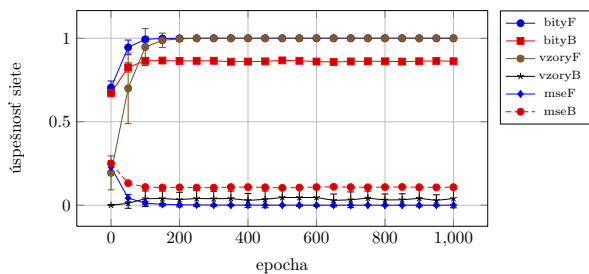
**Obr. 9:** Párovanie binárnych vzorov: vývoj sietí v čase (priemer a štand. odchýlka pre 50 sietí).

### 4.3 Komplexné asociácie

Vzhľadom na to, že v budúcnosti očakávame od nášho modelu netriviálne párovanie dátových vzorov, konkrétne nie len unikátne kombinácie 1:1, ale mapovanie jedného pohybu na 4 rôzne vizuálne reprezentácie (podľa perspektívy), rozhodli sme sa otestovať BAL na obdobných komplexne pospájaných dátach. Pre zjednodušenie sme vytvorili tréningovú množinu len o 16 dátových pároch dĺžky 16 bitov, z ktorých bolo  $k = 3$  pozitívnych, pričom 1 vzor na jednej strane je vždy asociovaný so štyrmi inými na druhej strane. Tak ako v ostatných dvoch prípadoch sme zistovali optimálne hodnoty  $\lambda$  a  $n_H$ . Na Obr. 10 ilustrujeme úspešnosť siete pre rôzne hodnoty  $\lambda$ , manipulovanie s  $n_H$  neukázalo žiadne významné rozdiely, preto ho neilustrujeme.

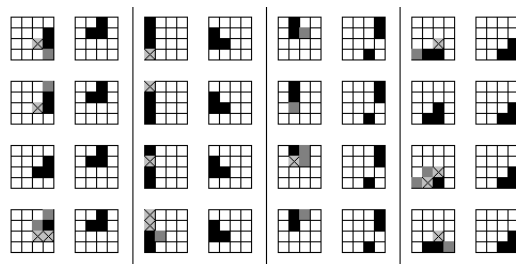


Obr. 10: Komplexné asociácie: úspešnosť siete ako funkcia  $\lambda$ .



Obr. 11: Komplexné asociácie: vývoj sietí v čase (priemer a štand. odchýlka pre 50 sietí).

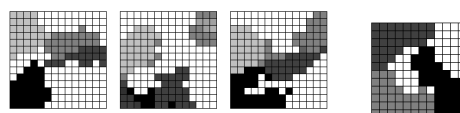
trujeme. Zo skúmaných parametrov sa najlepšie osvedčili  $\lambda = 1.0$  a  $n_H = 14$ . S týmito sme opäť monitorovali priebeh učenia siete. Priemerné hodnoty z 50 sietí zobrazuje Obr. 11. V prípade komplexných asociácií, kde sme v smere F asociovali štyri rôzne vzory s jedným a naopak v smere B, sme zaznamenali značne menšiu úspešnosť pre nejednoznačný smer B, konkrétne bitovú úspešnosť  $\approx 86\%$  a vzorovú úspešnosť len  $\approx 4\%$ . Z takýchto výsledkov môžeme predpokladať, že sieť úspešne reprodukuje smer F a vo smere B nevie určiť, ktorý zo 4 asociovaných vzorov vybrať, a teda zvolí ich prienik. Takéto správanie je u neuronových sietí očakávané. Odozvy siete bližšie ilustruje Obr. 12 vpravo. Ofarbené pixely indikujú aktiváciu nad 0.5. Ak  $p^d$  je želaný vzor a  $p^o$  je vzor na výstupe siete, potom pixel  $x$  bude označený čiernou farbou v prípade, že sú pozitívne obe hodnoty, teda  $p_x^d = p_x^o$ , šedou farbou ak sa pozitívna hodnota nachádza iba v  $p^d$  a šedou farbou a krížikom v prípade, že ide o falošnú pozitívnu aktiváciu na výstupe siete  $p^o$ .



Obr. 12: Odozvy jednej siete po natrénovaní na náhodných komplexných dátach.

## 5 Experimenty s robotickým modelom MNS

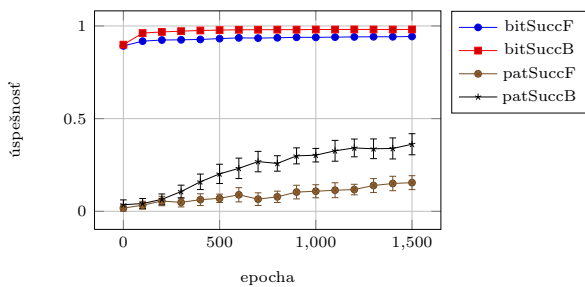
Prvým krokom vo vývoji agenta v našom modeli je sformovanie asociácie medzi vizuálnymi a motorickými reprezentáciami seba samého. V tejto časti prezentujeme výsledky z najvyššieho modulu nášho MNS modelu pre vlastnú perspektívu. Ako vstupy nám slúžili odozvy z dvoch MSOM sietí, modulov F5 a STS, ktoré sme trénovali na robotických dátach pomocou optimalizovaných parametrov (viac v [17]). Experimentovali sme tiež s veľkosťou oboch máp a parametrom  $k$ , teda počtom pozitívnych bitov v jednom vzore. Podľa výsledkov systematických testov rôznych dátových množín vzhľadom na priestorovú separáciu a jednoznačnosť vzorov a tiež podľa testov s BAL algoritmom sme zvolili veľkosť vizuálnej mapy  $14 \times 14$  s  $k_v = 16$  a veľkosť motorickej mapy  $12 \times 12$  s  $k_m = 8$ . Na Obr. 13 ilustrujeme víťazné aktivácie neurónov, tvoriace vzory štyroch perspektív, pre každý z troch pohybov (vyfarbené pixely znamenajú, že sa na danej pozícii nachádzal aspoň jeden víťaz v aspoň jednom vzore) a víťazné aktivácie pre všetky motorické kategórie v jednom obrázku (každý pohyb je zakreslený iným odtieňom šedej).



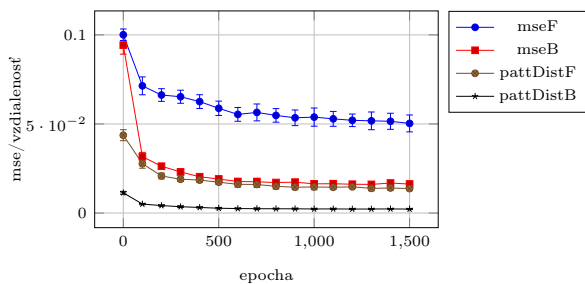
Obr. 13: Víťazné vizuálne neuróny na mape  $14 \times 14$  s  $k_v = 16$  pre tri typy úchopov a štyri perspektívy (vľavo) a motorické neuróny na mape  $12 \times 12$  s  $k_m = 8$  pre tri typy úchopov (vpravo).

Priebeh učenia siete ilustrujeme pre 20 sietí s architektúrou 196–170–144, náhodnými malými váhami a rýchlosťou učenia  $\lambda = 0.2$  trénovaných počas 1500 epoch. Obr. 15 zobrazuje dve miery úspešnosti siete, bitovú a vzorovú. Obr. 14 zobrazuje strednú kvadratickú chybu a ďalšiu kvantitatívnu mieru, ktorú sme nazvali vzdialenosť vzorov (vzdialenosť) a vystihuje vzdialenosť t' ažiska želaného



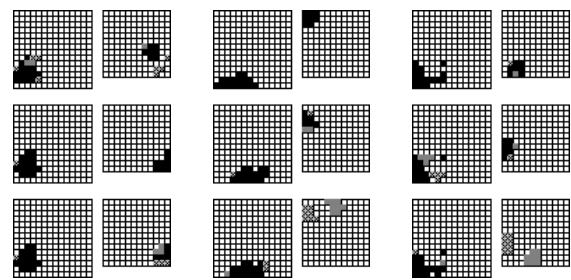


**Obr. 14:** Robotický model: bitová a vzorová úspešnosť v čase (20 sietí). Relatívne vysoká bitová úspešnosť naznačuje, že sieť reprodukuje vzory pomerne dobre, no často robí chyby (skoro v každom vzore), čo reflektuje malá vzorová úspešnosť.



**Obr. 15:** Robotický model: mse a vzdialenosť vzorov v čase (20 sietí). Klesajúca úspešnosť vzorov indikuje, že sa želaný a skutočný výstup siete nachádzajú v rovnakom klastrí.

vzoru a vzoru, ktorý dala ako výstup sieť na 2D mriežke. Vývoj úspešnosti sietí počas testovania naznačuje, že napriek tomu, že sa model evidentne učí, jeho úspešnosť nedosiahne vysoké hodnoty, hlavne pri vzorovej úspešnosti (smer F 15% a smer B 36%). Bitová úspešnosť je pomerne vysoká v oboch smeroch ( $\approx 96\%$ ) a vzdialenosť t' ažísk je nízka. Môžeme teda predpokladať, že sieť robí veľa malých chýb. Na základe úspešnosti a klesajúcej vzdialenosti vzorov usudzujeme, že sieť nerobí chyby medzi kategóriami, ale mýli si vzory v rámci kategórie (vd'aka topografickej organizácii vstupných vzorov). Na Obr. 16 ilustrujeme vybrané odozvy siete. Podobne ako v predošlej kapitole, farebné pozitívne pixely sú rozdelené na zhodné (čierna), neúspešné ciele (šedá) a falošné pozitíva (šedá s krížikom). Nedostatky v úspešnosti siete možno vysvetliť aj na základe vlastností vstupných dát. Napriek tomu, že ide o mapovanie jednej vizuálnej reprezentácie na jednu motorickú, vzory v rámci kategórií pohybov sa zvyknú podobat' a opakovať. V prípade, že model dostáva určité vzory v n-m priradeniach, má tendenciu reprodukovat' jeden z asociovaných vzorov, najmä ten, ktorý sa vyskytuje najpočet-



**Obr. 16:** Odozvy jednej siete po natrénovaní na robotických dátach. Čierna farba symbolizuje zhodu, šedá želaný vzor, ktorý sieť nezreprodukovala a šedá s krížikom indikuje vzor, ktorý dala na výstupe sieť mimo želaného vzoru (false-positives).

nejšie alebo bol nedávno aktivovaný. Ak by sme brali do úvahy čisto kategorické odozvy, teda jeden z troch druhov pohybu, bola by úspešnosť modelu značne vyššia.

## 6 Záver

Predstavili sme návrh nášho robotického modelu systému zrkadliacich neurónov. Pre potreby nášho modelu sme vytvorili algoritmus obojmerného učenia na báze aktivačných fáz BAL odvodený od biologicky plauzibilného modelu GeneRec. V článku sme sa bližšie povenovali vlastnostiam BAL pomocou testov s umelými dátami, a nakoniec sme načrtli prvé výsledky z nášho robotického modelu. Ukazuje sa, že napriek tomu, že si náš model „mýli“ inštancie pohybov v rámci kategórie, možno ho považovať za úspešný, minimálne na požadovanej vyššej úrovni reprezentácií (rozoznanie typu pohybu). Náš model možno porovnať s Hebbovským modelom [6], ktorý tiež popisuje vznik okruhu STS–PF–F5. Narozdiel od tohto prístupu sú u nás reprezentácie na PF (teda skrytej vrstve) vlastné siete. Je možné, že reprezentácie, ktoré si vytvára BAL na skrytej vrstve nemajú charakter reprezentácií v PF a pre lepšie vystihnutie neurálneho okruhu zrkadliacich neurónov bude potrebné priamo modelovať aj reprezentácie na tejto vrstve.

V neposlednom rade nám stále ostáva otvorená otázka variantných a invariantných reprezentácií v F5 a STS. Naše poznatky o MSOM nám zatiaľ naznačujú, že bude potrebné vyvinúť špeciálny mechanizmus pre invariantné reprezentácie, napríklad v podobe hierarchicky vyššej vrstvy nad modulom STS a F5. O hierarchickom vzťahu a anatomickej odlišnosti medzi variantnými a invariantnými neurónmi v STS svedčia aj poznatky z neurovedy [5]. Predpokladáme, že mechanizmus obojmerného informačného toku medzi zrkadliacimi neurónmi a vizuálnymi reprezentáciami by mohol napomáhať aj pri formovaní invariantných kategorických reprezentácií.

## Pod'akovanie

Tento príspevok vznikol v Centre pre kognitívnu vedu na KAI FMFI UK v Bratislave, za podpory grantov VEGA 1/0439/11 a 1/0503/13 a UK/631/2013. Ďakujeme L. Zdechovanovi za výstupy z natrénovaného iCub robota a M. Pecháčovi za výstupy z natrénovaných MSOM sietí.

## Literatúra

- [1] J. Bonaiuto, E. Rosta, a M. Arbib. Extending the mirror neuron system model, I: Audible actions and invisible grasps. *Biological Cybernetics*, 96:9–38, 2007.
- [2] V. Caggiano, L. Fogassi, G. Rizzolatti, P. Thier, a A. Casile. Mirror neurons differentially encode the peripersonal and extrapersonal space of monkeys. *Science*, 324(5925):403–406, 2009.
- [3] A.H. Fagg a M. Arbib. Modeling parietal-premotor interactions in primate control of grasping. *Neural Networks*, 11:1277–1303, 1998.
- [4] I. Farkaš, T. Malík, a K. Rebrová. Grounding the meanings in sensorimotor behavior using reinforcement learning. *Frontiers in Neurobotics*, 6(1), 2012.
- [5] T. Jellema a D. Perrett. Neural representations of perceived bodily actions using a categorical frame of reference. *Neuropsychologia*, 44:1535–1546, 2006.
- [6] C. Keysers a D.I. Perrett. Demystifying social cognition: a hebbian perspective. *Trends in Cognitive Sciences*, 8(11):501–507, 2004.
- [7] T. Kohonen. *Self-organizing Maps*. Springer, 1997.
- [8] G. Metta, G. Sandini, D. Vernon, L. Natale, a F. Nori. The iCub humanoid robot: an open platform for research in embodied cognition. In *Proc. of 8th Workshop on Perf. Metrics for Intel. Systems*, 50–56, 2008.
- [9] R.C. O'Reilly. Biologically plausible error-driven learning using local activation differences: The generalized recirculation algorithm. *Neural Computation*, 8(5):895–938, 1996.
- [10] E. Oztop a M. Arbib. Schema design and implementation of the grasp-related mirror neuron system. *Biological Cybernetics*, 87:116–140, 2002.
- [11] E. Oztop, M. Kawato, a M. Arbib. Mirror neurons and imitation: A computationally guided review. *Neural Networks*, 19(3):254–271, 2006.
- [12] E. Oztop, D. Wolpert, a M. Kawato. Mental state inference using visual control parameters. *Cognitive Brain Research*, 22:129–151, 2005.
- [13] E. Oztop, M. Kawato, a M.A. Arbib. Mirror neurons: Functions, mechanisms and models. *Neuroscience Letters*, 2012.
- [14] D. Perrett et al. Viewer-centred and object-centred coding of heads in the macaque temporal cortex. *Experimental Brain Research*, 86(1):159–173, 1991.
- [15] K. Rebrová and I. Farkaš. Bidirectional Activation-based Neural Network Learning Algorithm. Zasláné.
- [16] K. Rebrová a I. Farkaš. Robotický model systému zrkadliacich neurónov. In Nahodil P. Kelemen J., editor, *Kognice a umělý život XII*, 231–238. Slezská univerzita, Opava, 2012.
- [17] K. Rebrová, M. Pecháč, a I. Farkaš. Towards a robotic model of the mirror neuron system. Zasláné.
- [18] G. Rizzolatti a C. Sinigaglia. The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nature Reviews Neuroscience*, 11(4):264–74, 2010.
- [19] M. Strickert a B. Hammer. Merge SOM for temporal data. *Neurocomputing*, 64:39–71, 2005.
- [20] J. Tani, M. Ito, a Y. Sugita. Self-organization of distributedly represented multiple behavior schemata in a mirror system: reviews of robot experiments using RNNPB. *Neural Networks*, 17(8-9):1273–1289, 2004.
- [21] G. Tessitore, R. Prevede, E. Catanzariti, a G. Tamburini. From motor to sensory processing in mirror neuron computational modelling. *Biological cybernetics*, 103(6):471–485, 2010.
- [22] V. Tikhonoff, A. Cangelosi, P. Fitzpatrick, G. Metta, L. Natale, a F. Nori. An open-source simulator for cognitive robotics research: The prototype of the iCub humanoid robot simulator. In *Proc. of 8th Workshop on Perf. Metrics for Intel. Systems*, 57–61, 2008.
- [23] H. van Hasselt. Reinforcement learning in continuous state and action spaces. In *Reinforcement Learning*, pages 207–251. Springer, 2012.
- [24] P. Vančo a I. Farkaš. Experimental comparison of recursive self-organizing maps for processing tree-structured data. *Neurocomputing*, 73(7-9):1362–1375, 2010.
- [25] L. Zdechovan. Modeling the object grasping using the neural networks and iCub robotic simulator (in Slovak). Master's thesis, Comenius University in Bratislava, 2012.

# Etapy a mechanizmy raného jazykového vývinu

Ján Rybár

Centrum pre kognitívnu vedu, KAI, FMFI, Univerzita Komenského v Bratislave  
Mlynská dolina, 842 48

rybar@fmph.uniba.sk

## Abstrakt

Aké sú etapy raného jazykového vývinu, a ktoré mechanizmy sú charakteristické pre jednotlivé obdobia? Čo je v nich univerzálne a čo špecifické? Akú vlastnosť má vstup jazykovej informácie (tzv. opatrovčina)? Aké sú mechanizmy tvorby a percepcie reči? Na tieto otázky sa snaží odpovedať náš výskum, ktorého čiastočná sumarizácia výsledkov je predmetom nášho príspevku.

## 1 Úvod

Ako si dieťa dokáže osvojiť niečo tak komplexné ako je jazyk? Ako zvládne niečo, čo je tak bohaté štruktúrované na mnohých úrovniach?<sup>1</sup> Hoci sa môže zdať, že deti sa začínajú učiť jazyk až s prvými slovami okolo jedného roku, nie je to tak. Aj s prispáním moderných záznamových technológií jazykovej produkcie sa zistilo, že prvé slová sa môžu objaviť už od ôsmeho mesiaca veku dieťaťa. Existujú dokonca frekvenčné slovníky, v ktorých si môžeme nájsť, kedy a s akou frekvenciou sa jednotlivé prvé slová objavujú.<sup>2</sup> V každom prípade sa deti učia jazyk dlho predtým než začnú hovoriť a o jazyku vedia veľa už predtým než povedia prvé slovo.

## 2 Etapy raného jazykového vývinu

Prvou etapou jazykového vývinu je tzv. *cooing* (nastupuje vo veku okolo troch mesiacov). Existuje veľa názvov pre toto obdobie, napríklad *lalácia*, *hrkúťanie*, *broukání* (v češtine), ale nie sú také výstižné ako anglický výraz. Ide o produkciu zvukov, ktoré nie sú ani plačom, ani zvukom protestu, sú evidentne niečím iným. Ide

o zvuky podobné – *kú, kú, kú, akú, akú*, resp. *gú, gú, gú, agú, agú, agú*. Teda anglický výraz *cooing* nie je vôbec náhodný. Tieto zvuky sú spontánne, ale vznikajú pri nadväzovaní kontaktu s dieťaťom. Vznikajú v „konverzácii“ s dieťaťom. Ak sa usmejeme na dieťa, aj ono sa usmieva, ak budeme robiť smiešne pohyby tváre, aj ono bude robiť niečo podobné, ak budeme robiť *agu, agu, agu* aj ono bude odpovedať podobne.<sup>3</sup>

V siedmom mesiaci začína druhá etapa jazykového vývinu tzv. *babbling* (v slovenčine sa používa menej trefný termín *džavot*), pre ktoré sú charakteristické zvuky ako sú *babababa, dadadada* a pod. Ide o prvé spojenie hlások a samohlások. V neskoršej fáze obdobia *babblingu* sa začínajú formovať prototypové filtre pre percepciu foném materinského jazyka.<sup>4</sup> To je však na úkor univerzálnej schopnosti, ktorú dieťa malo v predchádzajúcom štádiu (rozoznávať, vnímať akékoľvek fonémy akéhokoľvek jazyka). Toto zúženie pravdepodobne pomáha deťom ľahšie a rýchlejšie si osvojiť materinský jazyk [1].<sup>5</sup>

Až ďalšou etapou je etapa *prvých slov*. Ako vôbec vznikajú prvé slová, ktoré majú význam? Intuitívne sa zdá, že to funguje takto – deti vidia veci a počujú slová, ktorými rodičia ukazujú na tieto veci, na základe toho si deti spoja tieto mená s týmito vecami. Túto teóriu sformuloval už svätý Augustín pred stáročiami. To je však veľmi naivná teória. Je to podstatne komplikovanejšie. Touto teóriou nevysvetlíme, odkiaľ sa vzalo napríklad slovo „*ešte*“, alebo také výrazy ako je „*kukuk*“, „*čojeto*“ a pod., ktoré sa tiež vyskytujú ako prvé slová. To, čo všetko môže byť prvým slovom je determinované predovšetkým potrebami a záujmami dieťaťa. Môžu to byť ľudia, zvieratá, časti tela, časti

<sup>1</sup> Akvizíciu jazyka môžeme skúmať na úrovni fonologickej, slovníkovej, morfolologickej, syntaktickej, sémantickej, naratívnej atď. Tieto akvizície sú simultánne a vzájomne prepojené.

<sup>2</sup> Napríklad takýto anglický a španielsky frekvenčný slovník možno nájsť na stránke: [www.sci.sdsu.edu/cdi/lexical\\_e.htm](http://www.sci.sdsu.edu/cdi/lexical_e.htm).

<sup>3</sup> Jednoducho na zvuky v určitej melódii, odpovedá zvukmi v podobnej melódii.

<sup>4</sup> Napríklad japonské deti do siedmeho mesiaca ešte rozlišujú medzi „r“ a „l“, ale v desiatom mesiaci už nie, je to pre nich tá istá fonéma.

<sup>5</sup> Len pre zaujímavosť poznamenajme, že vo všetkých jazykoch, ktorými v súčasnosti ľudia hovoria, je spoľuhlások asi 600 a samohlások asi 200 [4, s. 110].

oblečenia, jedlo, slová pre činnosť („hore“, „von“, „pije“), ale aj gramatické modifikátory („ešte“ čaj, „malý“ chlapec, „veľký“ strom) a slová sociálnej interakcie („áno“, „nie“, „chcem“, „nechcem“). S. Pinker vo svojej knihe „Jazykový inštinkt“ [2, s. 303-304] uvádza, že prvé slovo jeho synovca bolo „Batman“. Ak prvým slovom môže byť slovo „Batman“, tak potom to môže byť čokoľvek, s čím sa dieťa stretáva vo svojom jazykovom prostredí.

Prvé slová majú často charakter *holofráz* – myšlienok vyjadrených jedným slovom. V prípade nami skúmaného dieťaťa<sup>6</sup> napríklad slovo „hore“ znamená „ist na ruky“, „dole“ znamená „vyberte má zo stoličky“, „pije“ znamená „chcem piť“. Zaujímavé je, že takýchto slov je v jazykovej produkcii tohto obdobia je okolo 11% [4, s. 217]. Už aj týchto niekoľko málo príkladov z nášho výskumu poukazuje na to, že holofrázy majú silné komunikačné zameranie.

Ďalšou etapou sú *dvoj-trojslovné vety a tzv. telegrafická reč* (je to reč bez ohýbania – časovania, skloňovania a pod.). Dieťa začína dávať slová dokopy okolo 15 mesiacov [4, s. 95]. Uvedme niekoľko charakteristických príkladov z jazykovej produkcie nami skúmaného prípadu. Boli to vety typu „tata číta“ a tiež vysokú frekvenciu mali vety typu „ešte jeden“ (banán, kameň, kvetina, pes atď.)<sup>7</sup> Pre obdobie od 18 mesiacov je charakteristický nástup schopnosti rýchleho zapamätania slov (tzv. fast mapping). Dieťa je schopné okamžite zopakovať takmer každé slovo (hoci fonologicky to nemusí byť najpresnejšie).

Napriek tomu dieťa v tomto vekovom období má problémy s morfológiou (presnejšie s funkčnými morfémi časovania, skloňovanie atď.) Tieto morfémy sú ignorované alebo značne fluktuujú (raz tam sú, druhý raz zase nie sú). Vzhľadom na to, že slovenčina (podobne ako ostatné slovanské jazyky) patrí medzi jazyky s bohatou morfológiou, výskum nástupu funkčných morfém je veľkou výzvou pre výskumníkov v tejto oblasti. Gramatická morfológia sa výraznejšie začína presadzovať v treťom roku života dieťaťa.<sup>8</sup>

<sup>6</sup> V našom výskume sme zaznamenávali jazykovú produkciu dieťaťa na záznamník alebo videokameru raz za týždeň (niekedy aj častejšie). Záznamy vznikali v čase približne od 13.00 hod do 18.00 hod. Výskum prebiehal od siedmeho mesiaca do dvadsiatich štyroch mesiacov veku dieťaťa.

<sup>7</sup> Na spoločnej prechádzke sme šli okolo domu, na ktorého pivničných oknách bol starý zhrdzavený plech; keďže dieťa je v tomto veku veľký prieskumník, všetkého sa musí dotknúť, tak sme ho preventívne varovali výrazom „nono“. Keď sme prešli k ďalšiemu podobnému oknu, zrazu sme počuli trojslovnú vetu: „Ešte jeden nono“.

<sup>8</sup> Zaujímavé je, že v prípade nami skúmaného dieťaťa sa fenomén nadmerného používania pravidiel (overregulation)

### 3 Mechanizmy produkcie reči

Medzi typické fenomény raného jazykového vývinu (v priemere od 1,5 do 2,5 rokov) patrí fenomén *nadmerného rozširovania významu* (overextension errors). Môžeme uviesť niekoľko príkladov z nášho výskumu: slovo „lola“ (ktoré vzniklo z nepresného imitovania slova „moľa“) znamená aj „mucha“, „mravec“ a všetky ostatné „chrobáčky“; slovo „žížža“ je slovo pre každé svetlo (napríklad aj pre Mesiac na nočnej oblohe); menom maminej najlepšej priateľky dieťa volá aj jej všetky ostatné priateľky a samozrejme slovom „tata“ volá každého muža a pod.

Ďalej môžeme konštatovať, že v jazykovej produkcii dieťaťa sa už od konca prvého roku a v priebehu druhého roku života už veľmi výrazne prejavujú niektoré fundamentálne charakteristiky plnohodnotného jazyka. Jednou z nich je komunikačná úloha. O tom, že jazyk je predovšetkým nástroj na vyjadrenie potrieb dieťaťa nieto žiadnych pochyb. Napríklad medzi prvé slová mojej dcéry patrilo aj slovo „ešte“.<sup>9</sup> Druhá vlastnosť, ktorá je spätá už s týmto obdobím je dynamickosť jazyka. Akvizícia jazyka určite nie je len kumulatívna záležitosť. Jazyk malého dieťaťa nie je uzatvorený systém. Dochádza v ňom k neustálym zmenám, niečo pribúda, ale niečo aj vypadáva. Najnápadnejšie je to v súvislosti s procesom permanentného fonologického spresňovania jazykovej produkcie dieťaťa (ale tiež v prípade, ak sa nejaké slovo z nejakého dôvodu prestane používať). Treťou fundamentálnou vlastnosťou plnohodnotného jazyka, ktorú dieťa už v tomto období spĺňa je, že používa arbitrárne symboly. Citoslovci, ktorými dieťa napodobňuje rôzne zvieratka (a ktoré dieťa často používa) ešte túto podmienku nespĺňajú. No, ak slovo „krá“ znamená nielen vranu, ale aj všetky vtáky, ba aj lietadlo a helikoptéru, potom je to už o niečom inom.

Veľmi dôležitá pri akvizícii jazyka je aj forma vstupu jazykovej informácie, s ktorou prichádza dieťa do

---

objavil už na konci druhého roku (23 mesiacov, 27 dní). Keď dieťa krhličkou polievalo kvetiny na záhrade, na otázku, čo robí, odpovedalo „pracujem“. Ten istý vzor použilo, aj keď sme sa ho opýtali, čo robí s metlou, odpovedalo, že „metlujem“ (pre porovnanie pozri [3, 7. kap.]. V tom istom čase sme po prvýkrát u dieťaťa zaznamenali aj tvorbu časovacích funkčných morfém (časovanie slovesá v jednotnom čísle vo všetkých osobách) – „zhasn~~em~~“, „zhasn~~eš~~“, „zhasn~~o~~“, ďalej funkčnú morfému minulého času – „zhas~~la~~“ a funkčnú morfému prvej osoby množného čísla (hoci nie celkom presnú) – „chrápa~~m~~“.

<sup>9</sup> Stalo sa to raz, keď som mal svoju dcéru sám na starosti. Okolo druhej v noci sa začala mrviť a mrnčať. Vstal som a dal som jej piť čaj z fľaše, ktorú som mal pripravenú. Keďže sa mi však veľmi chcelo spať, po krátkej chvíli som fľašu vytiahol a vtedy som počul veľmi dôrazné „ešte“.

styku. Nevyhnutnou podmienkou je interakcia s osobami z bezprostredného okolia (rodičmi a ostatnými opatrovateľmi). Reč, ktorou hovoríme s malým dieťaťom do troch rokov má tieto špeciálne vlastnosti: vyššia tónina hlasu (falzet), pomalšia reč, predlžovanie slabík, vyberanie jednoduchších slov, reč nie je úplne fluentná, vety nie sú celkom gramatické, sú tiež krátke a jednoduché, pre dieťa sa vytvárajú ľahké formy – napríklad „havo“ (pes), „cica“ (mačka), „hijohet“ (kôň) a pod. Je to emočne pozitívne naladená reč, ľahko pochopiteľná, je skôr o konkrétnych veciach, týkajúcich sa záujmov dieťaťa, než o abstraktných záležitostiach [4, s. 82].<sup>10</sup>

Má tento spôsob reči s dieťaťom, má „opatrovčína“ univerzálny medzikultúrny charakter? Pozitívnu odpoveď na túto otázku môžeme nájsť v mnohých psychologických a antropologických výskumoch, podľa ktorých „opatrovčína“ nie je len vynálezom rodičov zo západných stredných sociálnych vrstiev. Nejaká kombinácia vlastností „opatrovčína“ bola potvrdená vo všetkých skúmaných sociálnych vrstvách a vo všetkých skúmaných kultúrach [4, s. 103-104].

V tomto smere v našom kultúrnom prostredí veľkú úlohu hrajú riekanky a pesničky, ktoré sa deti veľmi rýchlo a radi učia, pretože veľmi dobre vyhovujú požiadavkám kladených na „opatrovčinu“ (ich slová sú jednoduché, rýmujú sa, majú spevavú formu a pod.).

Ako sme už vyššie spomínali, k úspešnej akvizícii jazyka u dieťaťa je potrebná interakcia s osobami z jazykového prostredia<sup>11</sup> a pozitívne emočné prostredie (úloha rodičov je tu nespochybniteľná). No prostredníctvom akých mechanizmov sa učia deti z rodičovských výstupov? Aké nástroje má k dispozícii dieťa v tomto procese? Odpoveď na túto otázku, prinajmenšom pokiaľ ide o akvizíciu slovníka, sa zdá na prvý pohľad jednoduchá. Hlavným bezprostredným nástrojom dieťaťa pri nadobúdaní slovníka je *imitácia*. Zopakovanie nejakého výrazu je však v raných fázach vývinu jazyka málokedy presné, pretože sú tu veľké fonologické obmedzenia (dieťa nie vždy „trafi“ výslovnosť, rytmus, akcent a pod.).<sup>12</sup> Ďalej napríklad vo

<sup>10</sup> Nazýva sa to rôzne – reč orientovaná na dieťa (child directed speech), primárne lingvistické dáta (primary linguistic data), opatrovčína (caretaker talk), materinčina (motherese) a ešte inak (pozri [4, s. 80]). Dávame prednosť výrazu opatrovčína.

<sup>11</sup> Napríklad televízia je zdroj neinteraktívneho vstupu, preto sama osebe sa nemôže stať nejakým výrazným nástrojom akvizície jazyka v ranom vývine jazyka dieťaťa (Saxton, 2010, s. 88).

<sup>12</sup> V tejto súvislosti nám veľmi konvenuje predstava, že dieťa pri nadobúdaní jazyka sa nachádza v situácii, keď má k dispozícii audiovizuálne zariadenie, ale klávesy na klávesnici

výskumoch ako dieťa imituje poradie slov sa zistilo, že tento druh imitovania je presný len asi na 10 percent (Saxton, 2010, s. 92). Imitácia v ranom jazykovom vývine môže mať rôzne formy. Uvedme príklady z produkcie nami skúmaného dieťaťa. Najpresnejšie imituje špeciálne slová „opatrovčiny“ – také ako „babali“ (ležať), hači (sedieť), bábo (malé dieťa) a pod. Ide o slová s ľahkou výslovnosťou, ktoré si dieťa dokáže rýchlo osvojiť. Často dochádza k zámene foném v pomerne banálnych slovách, napríklad slovo „dobre“ vyslovuje ako „hobre“, slovo „vysoký“ ako „hysoký“, ešte častejšie je prehodenie a zdvojenie slabík, namiesto „zarabať“ hovorí „zabarabať“ a pod. Pokiaľ ide o dlhšie slová, skrakuje ich výslovnosť, napríklad hovorí „okulia“ (namiesto okuliare), d’ak (namiesto d’akujem), „kroko“ (namiesto krokodíl), „bčko“ (namiesto jabĺčko) a pod. Častým prípadom je tiež úprava výslovnosti zložitých slov. Napríklad slovo „električka“ bolo najprv imitované ako „etela“ (21. mesiac), potom ako „etelička“ (23. mesiac).

Veľa polemík sa vedie aj okolo tzv. korektívneho vstupu (opravovanie nesprávnej jazykovej produkcie dieťaťa). Odpoveď veľmi závisí od toho, ako riešia jednotliví autori dilemu, čo je vrodené a čo získané z jazykového prostredia pri nadobúdaní jazyka u dieťaťa. Jedna strana, nativisti (apelujúci na vrodené faktory), tvrdia, že rodičia dôsledne neopravujú chyby dieťaťa. Pre rodičov je dôležitejšie, či im dieťa rozumie. Ak rodičia predsa len skúsia chvíľu dieťa systematicky opravovať, sú veľmi často sklamaní z nie veľmi úspešných výsledkov. Jednoducho, pokiaľ dieťa nedospeje do príslušného vývinového štádia, inštrukciu ignoruje. To znamená, že jazykový vývin je do značnej miery spojený s procesmi dozrievania, ktoré sú biologicky determinované. Druhá strana, empirici, ktorí zdôrazňujú úlohu jazykového prostredia tvrdia, že rodičia síce dieťaťu nehovoria, že „toto hovoríš zle“, ale opravujú ho tým, že to povedia správne. To bolo potvrdené v mnohých výskumoch [4]. Tieto výsledky protirečia jednému z hlavných argumentov nativistov, a to, že jazykový vstup u dieťaťa je chudobnejší, než jazykový výstup (poverty of the stimulus argument) [2, kap. 9].<sup>13</sup> Podľa súčasných empirických výskumov skôr platí téza „čím bohatší vstup, tým bohatší výstup“ (v hre je frekvencia a pestrosť jazykových výrazov v prostredí).<sup>14</sup>

tohto zariadenia nie sú označené, a dieťa musí postupne skúšať jednotlivé klávesy a spätne vyhodnocovať ich efekty [2].

<sup>13</sup> Preto tu musí existovať nejaké vrodené zariadenie pre nadobúdanie jazyka (language acquisition device – LAD), inak by sme nemohli vysvetliť rýchlosť a ľahkosť, s ktorou dieťa nadobúda jazyk.

<sup>14</sup> K tejto diskusii pozri tiež napríklad [5].

## 4 Mechanizmy percepcie reči

V predchádzajúcich častiach sme sa zaoberali predovšetkým produkciou reči dieťaťa. Teraz by sme chceli zamerať pozornosť na druhú stranu akvizície jazyka – percepciu reči u dieťaťa. Samozrejme na takom malom priestore nemôžeme systematicky analyzovať percepciu reči dieťaťa, musíme sa obmedziť len na niektoré najpozoruhodnejšie mechanizmy a princípy. Medzi ne určite patria mechanizmy, ktoré pomáhajú deťom pochopiť (dešifrovať) význam zo zvukového kódu. Zastavme sa krátko pri niektorých. Pri percepcii reči sa stretávame s veľmi zaujímavým fenoménom *koartikulácie* (prekrývanie foném). Ide o to, že izolovaná fonéma napríklad *r* a tá istá fonéma v slove *ryba* neznie celkom rovnako, napriek tomu náš percepčný systém s tým nemá absolútne problém. Avšak výpočtové (programátorské) zvládnutie tohto problému je veľmi náročné.

Ďalej naše percepčné zariadenia vo všeobecnosti majú jednu pozoruhodnú vlastnosť a to, že dokážu fenomény, ktoré sú kontinuálne, vnímať diskkrétne (kategoricky). Napríklad spektrum je kontinuum, ale dokážeme ho vnímať diskkrétne. Podobne je to aj pokiaľ ide o reč. Aj dlhú vetu môžeme povedať jedným dychom bez prestávky, a napriek tomu bez problémov vieme v nej vydeliť jednotlivé slová. Tomuto fenoménu hovoríme *kategorická percepcia*. V tejto súvislosti P. K. Kuhl [1] urobila zaujímavý experiment na percepciu foném *r* a *l*. Skupine pokusných osôb zloženej z Američanov a Japoncov pustila z pásky prúd slov *lake, lake...*, ktoré sa v určitom momente bez prerušenia zmenilo na prúd slov *rake, rake...* Úlohou pre pokusné osoby bolo určiť moment, kedy sa prúd slov zmení z jedného slova na druhé. Pre Američanov to nebol žiadny problém, pokiaľ Japonci nepočuli žiadnu zmenu, pretože majú inak kategorizované fonémy (*r* a *l* je vlastne jedna fonéma). Táto kategorizácia foném nie je vrodenu záležitosťou, dochádza k nej v druhej polovičke prvého roku dieťaťa (ako sme spomínali vyššie, deje sa to medzi siedmym a desiatym mesiacom).

Ďalším zaujímavým mechanizmom percepcie reči je tzv. *McGurkov efekt*. McGurk (resp. jeho tím) uskutočnil nasledujúci experiment. Probandom púšťal videozáznam, na ktorom nejaká osoba hovorila *ga-ga-ga...* Zo zvukového záznamu však synchronne púšťal *ba-ba-ba...* Pokusné osoby v tejto situácii „dešifrovali“ tento zvuk ako *da-da-da...*<sup>15</sup> McGurk tým experimentálne potvrdil prepojenie zrakovej a sluchovej percepcie,

<sup>15</sup> Existujú početné realizácie tohto experimentu, niektoré sú prístupné na *youtube.com*. Naša vlastná skúsenosť s týmto efektom je, že sme síce nepočuli presné *da-da-da*, ale určite to nebolo ani *ga-ga-ga*, ani *ba-ba-ba*.

potvrdil fakt, že rečová percepcia je multimodálna. Už dieťa v cooingovom a babblingovom štádiu sa pokúša napodobňovať, aj „čítať“ z pohybu pier. Prepojenie zrakovej a sluchovej percepcie je jeden z dôležitých mechanizmov, ktorý pomáha dieťaťu prelomiť (dešifrovať) rečový kód.

Z uvedeného, okrem iného, vyplýva aj odpoveď na často kladenú otázku, či percepcia reči je špeciálnou percepciou (s osobitnými mechanizmami), alebo pracuje len s tými istými mechanizmami, ako sú použité pri spracovaní ostatných zvukov. Z toho, čo sme uviedli skôr vyplýva, že percepcia reči je doménovo špecifická (teda v hre sú špeciálne mechanizmy).

## 5 Záver

Dieťa je skutočne „kognitívnym laboratóriom“ pri skúmaní mechanizmov podieľajúcich sa na percepcii a produkcii jazyka. V týchto súvislostiach by sme ešte raz osobitne chceli zdôrazniť, že slovenčina ako jazyk s bohatou funkčnou morfológiou (koncovky skloňovania, časovania atď.) je veľkou výzvou pre výskumníkov v tejto oblasti.

## PodĎakovanie

Tento príspevok vznikol v Centre pre kognitívnu vedu, KAI, FMFI UK v Bratislave za podpory VEGA MŠVVaŠ SR, projekt č. 1/0503/13 a KEGA MŠVVaŠ SR, projekt č. 076UK-4/2013

## Literatúra

- [1] A. Gopnik, A. N. Meltzoff, P. K. Kuhl: *The Scientist in the Crib. What Early Learning tells us about the Mind*. HarperCollins Publishers, New York, 2001.
- [2] S. Pinker: *Language Instinct. How the mind creates language*. HarperCollins, New York, 2000. Český preklad: *Jazykový instinkt. Jak mysl vytváří jazyk*. Vydalo nakladatelství dybbuk, Jan Šavřda, Praha, 2009.
- [3] S. Pinker: *Slová a pravidlá. Zložky jazyka*. Kalligram, Bratislava, 2003.
- [4] M. Saxton: *Child Language. Aquisition and Development*. SAGE, London, 2010.
- [5] M. Takáč: Jazykové univerzálne a stelesnené kognícia. In: *Kognitívne paradigmy*. Vyd. Európa, Bratislava, 2013.

# Experimentální zkoumání kognitivních procesů pomocí původního webového nástroje MuTeP: zkoumání práce na mapách v kontextu osobnosti uživatele

Čeněk Šašínska, Zdeněk Stachon

Psychologický ústav FF MU  
Arna Nováka 1  
Brno 602 00

Geografický ústav PřF MU  
Kotlářská 2  
Brno 611 37

ceneksasinka@gmail.com, zstachon@geogr.muni.cz

## Abstrakt

V rámci interdisciplinární psychologicko-kartografické spolupráce byla vyvinuta původní výzkumná platforma MuTeP (Multivariantní-Testovací Program). Tento webový nástroj je primárně určen pro zkoumání kartografických vizualizací s fokusem na angažované kognitivní procesy. Zároveň umožňuje adaptovat či nově vytvářet a následně administrovat různé typy psychologických úloh a testů, včetně performačních. Platforma MuTeP umožňuje flexibilně měnit obrazový podnětový materiál a zároveň kombinovat řadu dílčích funkcí (časovač, výběr linie, bodové označování, asociace, zakreslení lomené linie atd.), díky čemuž lze efektivně sestavovat výzkumné baterie. Administrace připravených experimentů probíhá online za kontrolovaných podmínek a veškeré akce participantů (kromě pohybu myši) jsou ukládány do databáze. Platforma MuTeP umožňuje v konkrétních případech automatické vyhodnocování výkonů. Realizace empirických šetření na platformě MuTeP posouvá rovněž standardy ve způsobu prezentace výsledků, neboť čtenáři lze nabídnout nejen popis experimentu, ale za definovaných podmínek umožnit rovněž přístup do experimentu samotného.

Cílem příspěvku je rovněž na realizovaných experimentech představit užívaný výzkumný design v oblasti zkoumání a evaluace kvality kartografických vizualizací v kontextu interindividuálních rozdílů.

## 1 Charakteristika softwaru

Výzkumný software MuTeP je originální platforma, která vznikla pro potřeby výzkumného záměru „Dynamická geovizualizace v krizovém managementu“ a která je ovšem dále užívána i po jeho skončení. Do současné doby bylo na této platformě realizováno několik

diplomových (např. Mlejnková, 2012) a dizertačních prací (např. Zbořil, 2010) a rovněž i práce habilitační (Kubiček, 2012). O vývoji a směřování softwaru MuTeP poprvé informovala Sliviaková et al. (2009) právě na konferenci Kognice a umělý život. V současné době je software dále rozvíjen v navazujících výzkumných projektech a připravuje se ostré spuštění nové verze, která významným způsobem rozšiřuje možnosti využití platformy.

Technická specifikace softwaru je blíže popsána v Kubiček a Kozel (2010) a naopak Šašínska a Morong (2012) se více zaměřují na popis způsobu využívání platformy a aktuálně dosaženého stupně jejího vývoje. Šašínska a Morong (2012) mj. uvádí, že byl při návrhu designu softwaru kladen maximální důraz na jeho variabilitu, jelikož bylo cílem zkoumat vždy nový podnětový materiál a zároveň zcela odlišné typy úkonů. Primárním cílem bylo zkoumání práce na elektronických mapách resp. kognitivních procesů s ní souvisejících např. vizuální vnímání, rozhodování, plánování. Šašínska a Morong (2012) dále uvádí, že „výzkumník je schopen efektivně měnit obsah (různé mapové podklady) a povolovat a kombinovat potřebné funkce, které umožní zkoumat různé typy úkonů resp. kognitivních procesů. Software MuTeP je umístěn na serveru a jak příprava výzkumných baterií, tak samotná administrace testů probíhá online. Přístup pro administrátora (výzkumníka) je vázán na heslo a rovněž povolení IP adresy, testy jsou zpřístupněny přes heslo popř. volně na stránkách. Již byl realizován sběr dat i v zahraničí např. v Norsku. Software je vystaven na „Google Web Toolkit“ a „open layers“ (viz literatura). To má zásadní význam pro využití ve výzkumu v kartografii. V principu jde o to, že je stavěn podobně jako jiné interaktivní mapové systémy (resp. GISy - geografické informační systémy) např. google maps nebo mapy.cz. Umístíme-li do vrstev SW-MuTeP

kartografické podklady, může uživatel provádět obdobnou činnost jako na zmíněných GISech. To znamená, posouvat mapu, přibližovat, měnit měřítko apod. V současnosti tyto funkce nejsou implementovány a pracuje se na statických mapách, které zcela vyhovují aktuálním výzkumným požadavkům.“ Interaktivní mapy jsou vypracovávány jako nezávislý modul pro novou verzi.

Zatímco primární potřebou kartografů je sledování výkonů uživatele při řešení úloh, které realizuje na různých variantách map, tak aby mohla být navzájem porovnána kvalita těchto map, v psychologické perspektivě se zaměřujeme již na konkrétní kognitivní procesy, které jsou při zmíněných činnostech angažovány. Software MuTeP umožňuje rovněž adaptovat či nově vytvářet různé typy psychologických úkolů a testů, kterými jsme schopni měřit osobnostní proměnné uživatele, které mohou působit jako intervenující proměnné při práci s mapou do problematiky svého přispěvku.

## 2 Příklady úloh na platformě MUTEP

Vysoká variabilita typů úloh a zkoumaných jevů na platformě MuTeP je umožněna kombinací jejich základních funkcí, způsobem zápisu událostí do databáze a rovněž způsobem jejich vyhodnocování. Do databáze se zapisují veškeré akce kromě pohybu myši a čas těchto akcí, takže výzkumník má k dispozici vždy kompletní protokol z realizovaného běhu. Zároveň lze již dopředu definovat správná či očekávaná řešení, takže je možné efektivní automatické či poloautomatické vyhodnocení dílčích sledovaných jevů. V případě potřeby je možné ovšem i ex-post dohledat i další relevantní dříve neuvažované údaje. Díky stejné funkcionalitě a užitým principům, lze realizovat úlohy zaměřené na práci s mapami a zároveň lze platformu užít pro vytvoření či adaptaci a následnou administraci psychologických úloh a testů. V následujících dvou podkapitolách přiblížíme a na konkrétních příkladech ukážeme možnosti užití funkcionality softwaru jednak při výzkumu map a rovněž v případě psychologického testování.

### 2.1 Úlohy na mapách

Při zkoumání kvality kartografických vizualizací se konkrétní typy administrovaných úloh vždy odvíjí od účelu užití dané mapy. Značná část experimentů, ve kterých byly na platformě zkoumány alternativní mapové podklady, byla realizována v kontextu krizového managementu. Proto danému účelu odpovídaly i zkoumané fenomény, kdy se kladl důraz mj. na percepční kvalitu map resp. její čitelnost. Existuje vícero taxonomií úloh na mapách, my si představíme pojetí Wehrend a

Lewis (1990), které přibližuje Kubiček (2012). Autoři nabízí co do počtu položek jednu z nejrozsáhlejších taxonomií úloh na mapách a rozlišují tyto jednotlivé typy úloh:

- identifikace; rozpoznání obecné (vizuální) charakteristiky s jejíž pomocí jsou jednotlivé objekty rozlišitelné
- zjištění polohy; nalezení položek jisté hodnoty a rozpoznání absolutní či relativní pozice (Kde je? Je nalevo či napravo?)
- odlišení; rozpoznání či rozlišení dvou prvků od sebe
- kategorizace; vytvoření specificky definovaného rozdělení či klasifikace např. na základě barvy, velikosti, umístění, typu objektu
- zjištění shluků; spojení do skupin (shluků) založených na stejném, podobném nebo příbuzném typu grafické kvality
- prostorové uspořádání; popis celkového prostorového vzoru. Tato úloha má úzký vztah ke shlukům a také k umístění a identifikaci, avšak je komplexnější. Zjištění shluků má za cíl detekovat jednotlivé skupiny shluků, zatímco prostorové uspořádání požaduje také podrobný popis celkového uspořádání shluků.
- seřazení; vytvoření pořadí či pozice s ohledem na objekty podobného typu
- srovnání; proces nalezení podobností, rozdílů či pořadí
- propojení; spojení grafických prvků na základě určitého vztahu
- korelace; vytvoření přímého propojení

Často využívanou funkcí při úlohách na mapách je tzv. multi-klik (obr. 1), kdy je úlohou participanta označit požadované objekty. Po dokončení zadání pokračuje participant na další úlohou kliknutím na tlačítko „dále“. Lokace cílových objektů jsou definovány buď koordinátou XY a nebo polem. Platforma MuTeP automaticky generuje správnost výběru resp. vzdálenost od cílového bodu. Zaznamenává se pořadí kliknutí resp. označených míst a u všech úkonů se zaznamenává rovněž čas.





Obr. 1. Vícečetné bodové označování na mapě

Často využívaným typem úlohy je výběr dopředu vyznačené linie např. trasy nebo oblasti. Participanti volí pomocí funkce „výběr linie“ trasu či oblast na základě zadání. Administrátor má možnost pomocí koordinát definovat polohu linií a rovněž i jejich barvu. Při podržení kurzoru myši nad konkrétní linií, se tato linie (originálně žlutě) zvýrazní a lze ji označit (obr. 2). Nevýhoda užívání této funkce při plnění zadávání participantem je zřejmá. Linie sama o sobě zatěžuje graficky mapu a rovněž omezuje možnosti řešení. Výhoda naopak spočívá v jednoduché manipulaci s myší při vybírání řešení.



Obr. 2. Výběr trasy pomocí funkce „výběr linie“ (Mlejnková, 2011)

Funkce, která již umožňuje zachovat ve výzkumu více realistické podmínky při práci s mapou, je „lomená čára“. Na obrázku 3 je již tenkou (originálně modrou) linií

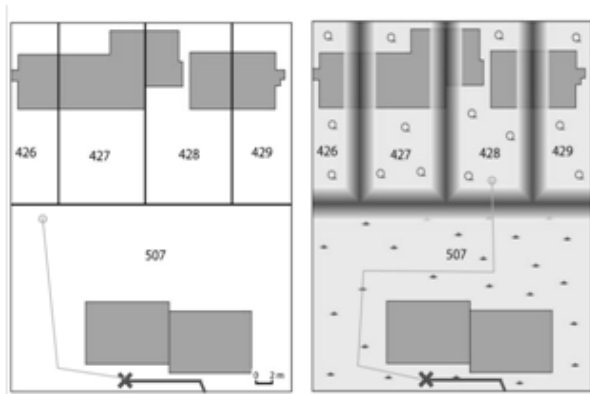
zakreslena uživatelem trasa pomocí lomené linie, kterou započal kliknutím do mapy, následně ji dalšími jednoduchými kliknutími lomil a ukončil ji dvojitým kliknutím. Nevýhodou užití této funkce je vyšší náročnost na jemnou motoriku, v důsledku čehož samotný výkon resp. naměřený čas není pouze a jenom funkcí zkoumaného fenoménu (vnímání a interpretace graficky kódované informace), ale je ovlivněn zároveň zručností participanta. Další nevýhoda se týká vyšší náročnosti při vyhodnocení výkonů participantů. Je důležité upozornit, že realizované experimenty byly vždy prováděny na výzkumném souboru zahrnující minimálně desítky osob a navržené testové baterie obsahují běžně desítky položek. Často rovněž neexistovalo jediné správné řešení (správná trasa), ale výzkumník se zajímal o subjektivně vybrané optimální řešení. Automatické vyhodnocení v tomto případě platforma umožňuje s rizikem ztráty informace. Administrátor definuje dopředu očekávaná možná řešení např. 5 konkrétních tras, jejichž volba se jeví být nejvíce pravděpodobná. Následně pomocí skrytých polygonů definuje konkrétní očekávaná řešení. Platforma následně srovnává výkon participanta, tedy jím zaznačenou lomenou linií, s předdefinovanými řešeními na základě několika parametrů např. shodu počátku a konce linie s očekávanými oblastmi, procento linie ve skrytém polygonu. Splňuje-li skutečně zaznačená lomená linie definované podmínky, je automaticky přiřazena ke konkrétnímu předdefinovanému řešení.



Obr. 3. Zaznačení trasy pomocí funkce „lomená linie“ (Mlejnková, 2011)

Stejná funkce, tedy lomená linie, je využita i pro odlišný typ úlohy, ve které bylo cílem srovnávat alternativní návrhy katastrálních map se zavedeným údajem o neurčitosti daných informací (obr. 4). Úkolem participantů bylo vést elektrické vedení tak, aby končilo u hranice pozemku dle definovaného zadání. Měřila se

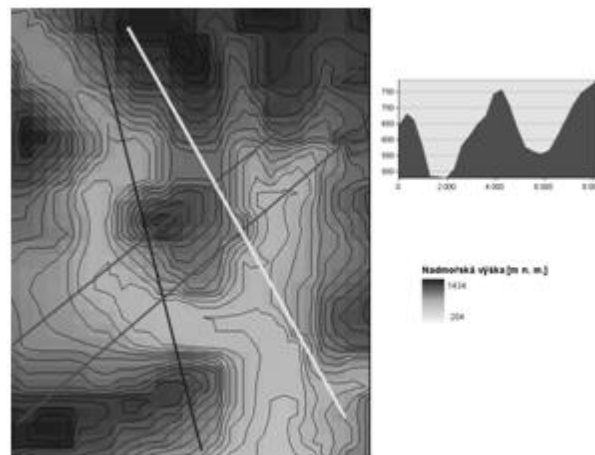
jednak správnost zakončení a rovněž vzdálenost od cílové hranice pozemku v pixelech.



Obr. 4. Zaznačení elektrického vedení na návrzích katastrálních map (Kubiček, 2011)

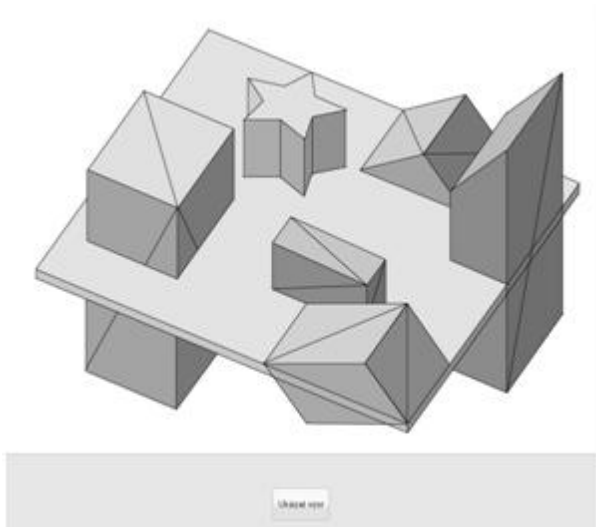
## 2.2 Psychologické testy a úlohy

Ač byl software MuTeP navržen primárně pro zkoumání výkonů na mapách a pro účely evaluace map, jeho využití bylo směřováno rovněž ve směru administrace psychologických testů a úloh. Většina současných psychologických testů má grafickou podobu a již díky základní funkcionalitě softwaru bylo možné vytvářet jejich elektronické mutace. Nově vyvíjená verze softwaru MuTeP klade již na možnosti administrace psychologických testů výrazně větší důraz. Zajímavý příkladem protnutí psychologických úloh a kartografie je test „výškopis“ autorky I. Vidlákové (2011). Na mapě je zakresleno několik linií pomocí funkce „výběr linie“ a participant má za úkol zvolit tu linii, která odpovídá převýšení na schématu vpravo od mapy.

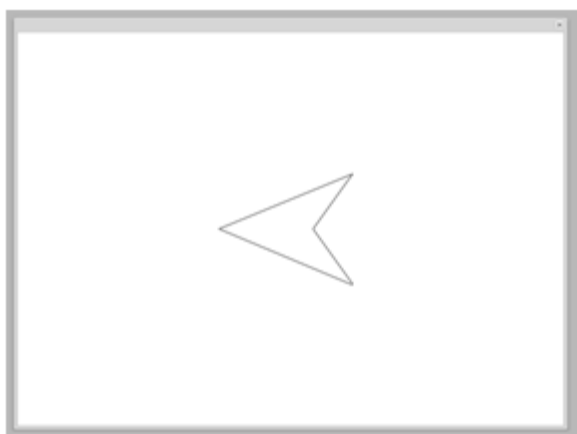


Obr. 5. Výběr řešení pomocí funkce „výběr linie“ (Vidláková, 2011)

Dnes již klasickým psychologickým testem je Witkinův Embedded figure test, který je založen na vyhledávání jednoduchých figur ve figurách komplexních. Vytvořili jsme elektronickou mutaci testu, která plně odpovídá i způsobem administrace původní Witkinově individuální papírové verzi. V individuální papírové formě trvá sběr dat za přítomnosti administrátora včetně vyhodnocení přibližně ¾ hodiny. Administrace na naší platformě dosahuje nejen větší přesnosti při měření, ale zároveň je umožněn skupinový sběr dat, což přináší nemalou časovou a finanční úsporu. Ve Witkinově případě je nejdříve na definovaný časový úsek exponována složitá figura, následně jednoduchá a poté znovu složitá (obr. 6), ve které již participant vyhledává jednoduchou figuru. Zároveň má možnost požádat o znovu předložení jednoduché figury (obr. 7). V daném případě je komplexní figura po dobu expozice skryta a administrátor zaznamenává i dobu těchto událostí. Platforma MuTeP umožňuje nejen odpovědi participanta resp. jejich správnost automaticky vyhodnotit, ale rovněž je k dispozici celkový protokol včetně všech událostí a jejich časů a trvání.



Obr. 6. Expozice komplexní figury po stanovení interval. Je povolena funkce „dialog window“, která umožňuje znovu předložit jednoduchou figuru



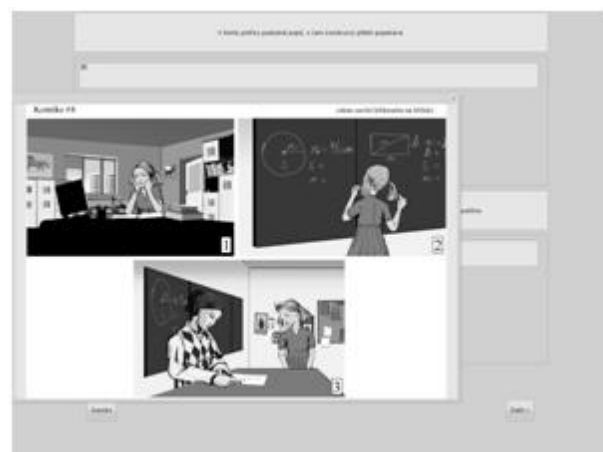
Obr. 7. Expozice jednoduché figury funkcí „dialog window“. Zároveň je zastřena komplexní figura

Oba výše zmíněné psychologické testy řadíme do kategorie performačních. Na platformě MuTeP byl již ale realizován mj. i původní projektivní test (TRS - Testu reakce na stres), který byl zároveň připraven i v papírové individuální a sešitové skupinové verzi (viz Šašinka et al., 2011). V testu reakcí na stres jsou vždy exponovány krátké komiksové příběhy a participant je následně písemně interpretuje (obr. 9) a popř. hodnotí formou dotazníku (obr. 8). Jak při dotazníkovém hodnocení, tak verbálním popisu má vždy možnost si vyvolat funkci „dialog window“ komiksový příběh. Zajímavostí, která zrcadlí jednu ze základních koncepčních idejí platformy, je formát výstupu dotazníku. Výzkumník má v databázi k dispozici nejen informace o finální volbě participanta, o

jeho odpovědích na jednotlivých škálách, ale je dostupná informace o procesu volby. V jakém pořadí otázek při volbě postupoval, zda volbu hodnoty realizoval okamžitě napoprvé, nebo „váhal“ a měnil odpovědi, zda nahlížel vícekrát na komiks atd. Tedy je postupováno cestou, raději uchovávat vše a nevyužít, než přijít o potenciálně zajímavá data.

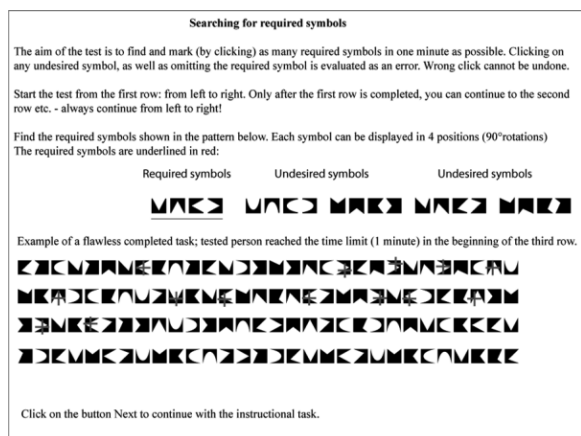


Obr. 8. Dotazník, který je součástí Testu reakce na stres



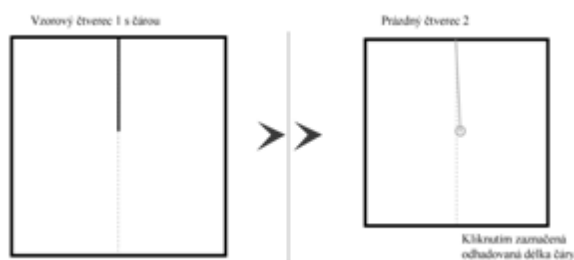
Obr. 9. Zobrazení komiksu funkcí „dialog window“ při verbální odpovědi

Příkladem užití funkce vícečetného bodového označování v rámci psychologické diagnostiky na platformě MuTeP je adaptace Švancarova diskriminačního testu (obr. 10), který je řazen mezi tzv. škrtačí testy (viz Svoboda, 2010) a je využíván při studiu pozornosti a psychomotorického tempa.



Obr. 10. Adaptace Švancarova diskriminačního testu

Dalším příkladem adaptovaného psychologického testu, který byl užíván mj. v kulturně-srovnávacích experimentech, je Frame-line test (viz Kitayama et al., 2003). Úlohou participantů je odhadovat absolutní resp. relativní délku čáry při změně velikosti čtverce. I zde v případě této adaptace je při zachování původního postupu administrace elektronická metoda na platformě výrazně efektivnější než původní papírová forma.



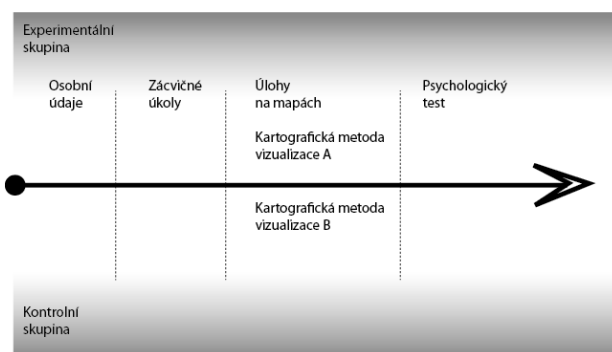
Obr. 11. Adaptace Frame-line testu

Na několika případech jsme si ukázali, jaké jsou možnosti platformy MuTeP již při užití základní funkcionality. Užití shodných principů manipulace při práci s různým podnětovým materiálem při realizaci kombinovaných experimentálních studií vede k vyšší racionalitě výzkumu. Participanté užívají jak v kartografických, tak psychologických subtestech stejné nástroje a tudíž se snižuje čas nutný pro zácvek. V následující kapitole si přiblížíme, proč je např. tento aspekt důležitý při realizovaných experimentech a jaké praktické výhody nabízí vysoká variabilita užívané platformy MuTeP.

### 3 Užívaný výzkumný design v oblasti interdisciplinárního zkoumání map

#### Název kapitoly

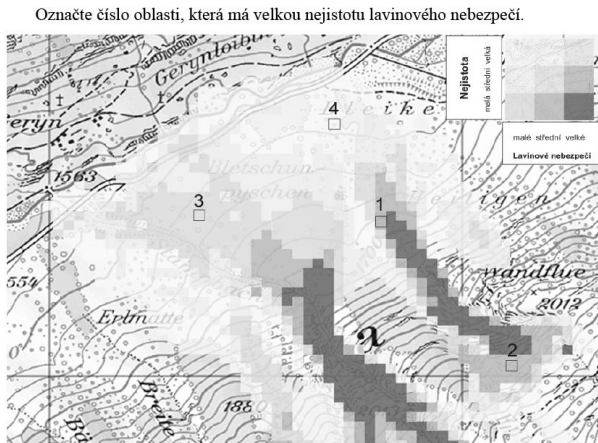
Typickým výzkumným designem užívaných v rámci psychologicko-kartografické spolupráce (viz. Konečný et. al, 2011, Štěrbá et al., 2011, Štěrbá a Šašinka, 2012) je experiment resp. následné pozorování na dvou skupinách v kombinaci s korelační studií (obr. 12). Bylo srovnáváno dvě či více kartografických metod vizualizace v kombinaci s psychologickým testem resp. úlohami. Zmíněným maticovým způsobem bylo možno zjistit nejen rozdíly v efektivitě a rychlosti zpracování alternativních kartografických metod vizualizace nebo kartografických děl obecně, ale zároveň bylo možné výkony dosažené v úlohách na mapě uvést do souvislosti s psychologickými koncepty např. inteligencí, kognitivním stylem a zjišťovat vliv interindividuálních rozdílů. Obdobný design výzkumů v rámci této oblasti není ničím zcela neobvyklým (viz Lobben, 2004, Lobben et al., n.d.), ovšem významným posunem dopředu je možnost realizovat úlohy z obou oblastí na jedné univerzální platformě.



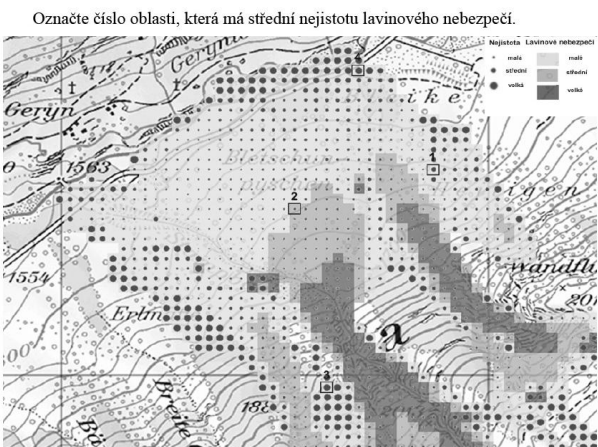
Obr. 12. Typický výzkumný design užívaný v rámci kartograficko-psychologické spolupráce

Jako příklad užití výše zmíněného výzkumného designu na platformě MuTeP krátce představíme studii zaměřenou na srovnání dvou alternativních metod vizualizace lavinového nebezpečí a jeho nejistoty (Šašinka, 2012). Stejná data byla zobrazena dvěma různými metodami – pomocí tzv. vnitřních (obr. 13) a vnějších (obr. 14) proměnných. Zároveň byla užitá mutace Navonova hierarchického testu (obr. 15), kterými lze odlišit analyticky resp. globálně zaměřené jedince. Díky maticové formě výzkumného designu jsme mohli porovnat nejen výkony na alternativních mapách mezi sebou, ale dospěli jsme ke zjištění, že jedinci, kteří dosahují nižších reakčních časů v subtestu lokálních číslic a jsou tedy spíše analytičtí, zároveň dosahují i nižších

reakčních časů při užití mapy s vnitřními proměnnými. Naopak jedinci více globálně orientovaní, kteří dosahují nižších časů u globálních číslic, dosahují zároveň nižších časů u mapy s vnějšími proměnnými.

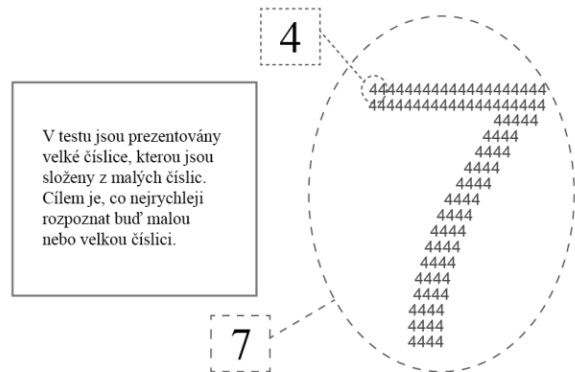


Obr. 13. Mapa lavinového nebezpečí a nejistoty – zobrazení pomocí vnitřních proměnných



Obr. 14. Mapa lavinového nebezpečí a nejistoty – zobrazení pomocí vnějších proměnných

### Test malých a velkých číslic



Obr. 15. Test velkých a malých číslic – mutace Navona hierarchického testu Takto se člení kapitoly vnitřně a na další podkapitoly.

## 4 Závěr

Možnosti uplatnění představené platformy MuTeP nevidíme pouze v oblasti kognitivní kartografie, ale přinejmenším stejný potenciál má mj. i v oblasti psychodiagnostiky. Nově vyvíjená verze se vydává cestou ještě větší interaktivity a adaptivity a rovněž je snahou zvýšit její šíři užití za pomoci modulární architektury.

## Poděkování

Príspevek vznikl v rámci projektu CZ.1.07/2.3.00/30.0037, který je spolufinancován Evropským sociálním fondem a státním rozpočtem České republiky.

## Literatura

- [1] Google web toolkit. <https://developers.google.com/web-toolkit/?hl=cs>
- [2] Kitayama, S., Duffy, S., Kawamura, T. & Larsen, J. T. (2003). Perceiving an object and its context in different cultures: A cultural look at new look. *Psychological Science*, 14 (3), 201–206.
- [3] Konečný, M., Kubíček, P., Šašík, Č. & Stachoň, Z. (2011). Usability of Selected Base Maps for Crises Management - Users Perspectives. *Applied geomatics*, 3(4), 189-198.
- [4] Kubíček, P. & Kozel, J. (2010). Cartographic Techniques for Adaptive Emergency Mapping. In

- H. Kremers, A. Susini (eds.). *Risk Models and Applications*. (pp. 49 – 62). Berlin: CODATA-Germany e.V.
- [5] Kubíček, P. (2011). Možnosti testování kartografické vizualizace polohové nejistoty. *Kartografické listy*, 19, 88 - 96.
- [6] Kubíček, P. (2012). Manuskript habilitační práce - Geoinformatizace v kartografii. Brno: GÚ PŘF Masarykova univerzita.
- [7] Lobben, A. K. (2004). Tasks, Strategies, and Cognitive Processes Associated With Navigational Map Reading: A Review Perspective. *The Professional Geographer*, 56(2), 270- 281.
- [8] Lobben, A.K., Olson, J.M. & Huang, J. (2005). Using fMRI in cartographic research. Proceedings of the International Cartographic Conference, A Coruna. Retrieved on April 4, 2012 from [http://geography.uoregon.edu/geocog/information/library/lobben\\_olson\\_05.pdf](http://geography.uoregon.edu/geocog/information/library/lobben_olson_05.pdf)
- [9] Mlejnková, K. (2011). Způsob orientace a sebelokalizace jedince v prostředí v závislosti na schopnosti používat mapu [diplomová práce]. Brno: GÚ PŘF Masarykova Univerzita.
- [10] Open layers. <http://openlayers.org/>
- [11] Sliviaková, A., Stachoň, Z., Šašinka, Č. & Zbořil, J. (2009). Posuzování uživatelských charakteristik kartografických produktů: interakce člověk a GIS. In *Kognice a umělý život IX* (pp. 297-304). Opava: Slezská univerzita v Opavě.
- [12] Svoboda, M. (2010). *Psychologická diagnostika dospělých*. Praha: Portál.
- [13] Šašinka, Č. (2012). Interindividuální rozdíly v percepci prostoru a map [dizertační práce]. Brno: PÚ FF Masarykova Univerzita.
- [14] Šašinka, Č., Kocábová, Z. & Jankovská, M. (2011). Test reakcí na stres: Vývoj nového grafického projektivního testu pro cílovou skupinu dětí a adolescentů. In *Psychologická diagnostika dětí a dospívajících: výzkum, prevence a školní poradenství*. Sborník abstraktů. Brno: Masarykova Univerzita.
- [15] Šašinka, Č. & Morong, K. (2012). Původní výzkumný nástroj pro oblast kartografie a psychologie – Multivariantní Testovací program (MuTeP). In P. Halama, R. Hanák, R. Masaryk (eds.). *Sociálne procesy a osobnosť 2012: Zborník príspevkov z 15. ročníka medzinárodnej konferencie* (pp. 188-194). Bratislava: Ústav experimentálnej psychológie.
- [16] Štěrba, Z., Stachoň, Z., Šašinka, Č., Zbořil, J., Březinová, Š. & Talhofer, V. (2011). Evaluace kartografických znakových sad v kontextu osobnosti uživatele. *Geodetický a kartografický obzor : odborný a vědecký časopis Českého úřadu geodetického a kartografického a Slovenského úřadu geodézie a kartografie*, 99/57(8), 173-182.
- [17] Štěrba, Z. & Šašinka, Č. (2012). Evaluation of Maps for Crisis Management in the Context of User's Cognitive Features. In H. Kremers & A. Susini (Eds.). *Risk Models and Applications* (pp. 79-88). Berlin: CODATA-Germany e.V.
- [18] Vidláková, I. (2011). Diagnostika předmětově – prostorové představivosti u dětí a dospívajících. In *Psychologická diagnostika dětí a dospívajících: výzkum, prevence a školní poradenství*. Sborník abstraktů. Brno: Masarykova Univerzita.
- [19] Wehrend, S. & Lewis, C. (1990). A Problem-oriented Classification of Visualization Techniques. In *Proceedings of the 1st conference on Visualization '90*, 139-143.
- [20] Zbořil, J. (2010). Kontextová kartografická vizualizace a její využití v krizovém managementu [dizertační práce]. Brno: GÚ PŘF Masarykova Univerzita.

# Kognitívny priestor

Ján Šefránek

Katedra aplikovanej informatiky,  
Fakulta matematiky, fyziky a informatiky,  
Univerzita Komenského, Bratislava  
sefranek@ii.fmph.uniba.sk

## Abstrakt

Tento text načrtáva pohľad na svet myšlienok. Predstavujem si všetky sformulované a ľuďom dostupné myšlienky – to je kognitívny priestor. Kognitívny priestor žije vlastným životom – pôsobí na ľudí, na spoločensvá, vyvoláva emócie, pribúdajú v ňom nové myšlienky, odohrávajú sa v ňom dialógy, spory a konflikty. Obsahuje veľmi rôznorodé vrstvy. V pozadí za touto predstavou sú osobnosti, ktoré dokážu pozitívne, ale aj veľmi negatívne, pôsobiť, vyžarovať myšlienkami i činmi do prostredia.

## 1 Úvod

Podnetom pre tento text boli verejné vystúpenia dvoch mužov. Zaujali ma tak neodbytné, že som začal rozmýšľať o sile myšlienok, o tom, ako na ľudí vplývajú myšlienky kedykoľvek a kdekoľvek vyslovené, napísané alebo demonštrované skutkami. Môžeme si predstaviť, že sídli v obrovskom kognitívnom priestore.

Zapôsobil na mňa predovšetkým kontrast k prevládajúcemu tónu, farbe a najmä obsahu súčasného slovenského prostredia, v ktorom sa šíria myšlienky alebo, presnejšie, reči. Kontrast k vyprázdnenému prostrediu smotánok, „módnych policajtov“, politických tárajov. Netreba však zabudnúť na nás, mlčiacu, nereagujúcu šedivú masu.

Mojou ambíciou nie je vyčerpávajúci systematický opis sveta myšlienok, jeho triedenie, charakterizácia dôležitých typov. Ponúkam iba niekoľko reflexií. Každopádne si však myslím, že ríša myšlienok si zaslúži pozornosť kognitívnej vedy.

Mal by som poznamenať, že reflexie, ktoré budú nasledovať, sú zmesou esejistických postrehov, pokusov o vhlad do problematiky, ale iste aj Cimrmanovských absurdít. Má to aj podobu politického pamfletu. Používam publicistický, nie vedecký jazyk. Domnievam sa, možno veľmi naivne, že pozvaná prednáška to znesie.

Na začiatok trocha nudných priblížení pojmov, aby bolo

jasnejšie, o čom bude reč. Pôjde iba o priblíženie pojmov, pokusy o ich presné vymedzenie v tomto kontexte by neboli produktívne. Reč bude o súhrne (priestore) všetkých myšlienok, ktoré sú dostupné nejakému (aspoň jednému) človeku. Tak chápem v tomto texte slovo kognitívny priestor. Za dostupnú myšlienku považujem takú, ktorá je sformulovaná v nejakom jazyku. Môže, ale nemusí to byť oznamovacia veta, o pravdivosti ktorej má zmysel uvažovať (výrok). Aj otázky alebo požiadavky sú schopné vyjadriť myšlienku. Na vyjadrenie zaujímavejších myšlienok zväčša treba viacero viet<sup>1</sup>.

Nebudem sa pokúšať o vylúčenie nejakých myšlienok z kognitívneho priestoru na základe nejakých kvalitatívnych požiadaviek.

Vyslovené a napísané myšlienky žijú vlastným životom, narážajú alebo víťazia, vstupujú do polemík, dialógov alebo zapadnú do zabudnutia.

Bude ma zaujímať emocionálne pôsobenie myšlienok a ich pôsobenie na spoločensvá, komunity. Nevyhnem sa ani tomu, aby som rozlišoval rôzne vrstvy v kognitívnom priestore. Možno ani veľmi nezjednoduším, keď napíšem, že sa pokúsím o pohľad na pozitívne a negatívne vrstvy.

## 2 Myšlienky a komunity

Značná časť priestoru myšlienok je rozdrobená medzi mnohé komunity. Teória šachových koncoviek zaujíma špecifickú skupinu ľudí, túto skupinu možno nazvať gravitačným poľom myšlienok, týkajúcich sa šachových koncoviek. Pravdepodobne gravitačné pole pikošiek o živote „celebrít“ bude oveľa väčšie.

Môžeme teda hovoriť o špecializovaných podpriestoroch kognitívneho priestoru. Každý člen komunity nemusí byť

---

<sup>1</sup>Nechcem tým poprieť zjavenie takých myšlienok ako “viem, že nič neviem”, “cogito ergo sum”, “o čom nemožno hovoriť, o tom treba mlčať”, atď.

oboznámený s každou myšlienkou daného podpriestoru, ale potenciál myšlienky z tohoto podpriestoru pritiahnúť záujem člena komunity tu je – závisí od sily myšlienky i od schopnosti človeka ju pochopiť, interpretovať, zaradiť do svojho názorového sveta. Názorový svet človeka je tu termín pre súhrn myšlienok, ktoré človek prijíma. Sila myšlienky (pre nejakú skupinu ľudí) je intuitívne zrejmy pojem. Môžeme rozlišovať pozitívnu (intenzita akceptovania myšlienky) a negatívnu (intenzitu odmietania).

Príkladom takýchto podpriestorov sú vedecké špecializácie. Pozornosť teraz sústredím na oblasť reprezentácie znalostí. Začnem o mojej práve prebiehajúcej práci. Pokúšam sa aktualizácie (updates) nemonotónnych báz znalostí sformalizovať pomocou tzv. assumption-based frameworks (ABF) [1]. ABF umožňujú predstaviť najznámejšie nemonotónne formalizmy ako dedukciu z nejakých odvolateľných, vyvrátených predpokladov. Preto opis aktualizácie ABF možno vidieť ako všeobecný rámec pre aktualizáciu báz znalostí, vyjadrených v rozličných nemonotónnych formalizmoch. Mňa najviac zaujíma prípad logických programov. Ak prácu úspešne ukončím a bude niekde prijatá, dozvie sa o nej niekoľko ľudí, možno bude mať pár citácií (jej špeciálnejší predchodca z roku 2012 má zatiaľ dve) a to je všetko. Pre mňa osobne je tam aj emocionálna motivácia. Netrúfam si odhadnúť, či (pozitívne :-)) zapôsobí na emócie niekoho iného. Skôr nie.

Obráťme pozornosť k vrcholom. Konferencie Principles of knowledge representation and reasoning sa usporadujú raz za dva roky a vytvorila sa tradícia ovenčiť prívlastkom Great moments in knowledge representation jednu pozvanú prednášku. V roku 2012 mal takúto prednášku jeden z popredných vedcov komunity Vladimir Lifschitz. Prednášku nazval Logic programming solution to the frame problem [2]. Problém rámca, frame problem, bol významným podnetom pre pestovanie nemonotónnych logík, lebo klasická logika nie je vhodná na jeho riešenie. Jeho obsah je, stručne, nasledovný. Pri reprezentácii znalostí o dynamických doménach by nebolo účelné vyjadriť, ktoré vlastnosti objektov sa nemenia po akejsi akcii. Problém rámca spočíva práve v tom, ako úsporne, ale korektne, bez neželaných dôsledkov, reprezentovať, že nejaká vlastnosť nejakého objektu sa nemení po danej akcii.

Základné riešenie ponúkol Ray Reiter v defaultovej logike. Napríklad defaultové pravidlo

$$\frac{\text{in}(\text{Alice}, s) : M(\text{in}(\text{Alice}, \text{next}(s)))}{\text{in}(\text{Alice}, \text{next}(s))}$$

interpretujeme tak, že Alice je v miestnosti aj v nasledujúcom stave, ak bola v predchádzajúcom a je možné konzistentne predpokladať, že v tej miestnosti ešte stále je (nie je dokázaný opak). Ak v báze znalostí, napríklad, máme, že v nasledujúcom stave vstúpil do miestnosti Bob, bolo by korektné odvodiť, že v nasledujúcom stave je v miestnosti nielen Bob, ale (ešte stále) aj Alice. Reiterovo riešenie problému rámca umožňuje takéto odvodenie.

Tento typ riešenia možno vyjadriť aj v logickom programovaní, založenom na sémantike stabilných modelov (ASP, answer set programming). Podstatným prínosom ASP je, že vďaka výkonným solverom umožnilo automatické usudzovanie, ktoré zahŕňa aj riešenie problému rámca. ASP je využívané aj v reálnych aplikáciách. Ich kód obsahuje aj riešenie problému rámca.

Riešenie problému rámca oceňujú špecialisti (stovky ľudí?), pre niektorých môže byť aj zdrojom emocionálnej reakcie. Pravdepodobne ho ocenia aj niektorí študenti alebo niekoľko ľudí, ktorí náhodne s ním prídu do styku. Na niektorých to možno zapôsobí aj emocionálne.

Navyše, treba poznamenať, že vyriešenie problému rámca a ďalšie prínosy, získané v nemonotónnych logikách posunuli logiku bližšie k postihnútiu spôsobov uvažovania, aké ľudia používajú v každodenných situáciách (common sense reasoning). Napriek tomu, nič to nemení na konštatovaní, že gravitačné pole tejto oblasti a intenzita jej pôsobenia na svet sú nepatrné.

Často sa stáva, že významný príspevok do nejakej vedy je vlastne drobnosťou z hľadiska celého kognitívneho priestoru, najmä z hľadiska vplyvu myšlienok na spoločenstvá, vrátane vplyvu emocionálneho. Kognitívny priestor je živý priestor, existuje iba vďaka tomu, že k nemu pristupujú živí ľudia. Prívlasztňujú si myšlienky, interpretujú ich, chápu ich, rozvíjajú ich, nechápu ich, prznia ich, udržiavajú ich pri živote, prinášajú nové myšlienky. Obdivujú ich, nenávidia ich. Niektoré vrstvy kognitívneho priestoru majú vo svojom gravitačnom poli množstvo ľudí. Práve pôsobenie myšlienok na veľké spoločenstvá je hlavnou témou tohoto textu. Samozrejme, nemyslím si, že takéto pôsobenie je cieľom vedeckých špecializácií.

Na tomto mieste treba poznamenať, že výsledky niektorých vedeckých komunit sa stávajú zložkou obrazu sveta, ovplyvňujú názorový kontext mnohých ľudí, spoluvytvárajú spoločný kognitívny priestor rozsiahleho spoločenstva, majú aj potenciál emotívneho pôsobenia na ľudí. Samozrejme, pri takomto všeobecnom rozšírení sú



dostupné hlavne ich interpretácie. Často sú zjednodušujúco interpretované, nedopochopené, neraz aj zle pochopené. Podobne možno vidieť vplyv myšlienok, ktoré viedli k významným technologickým inováciám.

Všimli sme si vrstvenie kognitívneho priestoru, spoločenský kontext týchto vrstiev, teda ich vzťahy s rôznymi komunitami. Uvedomili sme si, že dôležitým aspektom týchto vzťahov je emotívne pôsobenie. Domnievam sa, že základom tohoto emotívneho pôsobenia je dialogická až konfliktná povaha kognitívneho priestoru. Vrstvy, ktoré sú v spore, majú svoje fankluby. Fankluby sa vytvárajú podľa toho, ako príslušné myšlienky zodpovedajú záujmom, názorovému svetu a mentálnej výbave ich členov. Prejdeme teraz k dialogickej povahe kognitívneho priestoru, potom k jeho emotívnemu pôsobeniu.

### 3 Kognitívny priestor ako dialóg

Kognitívny priestor je poprepájaný. Niektoré myšlienky sú podporované, zdôvodniteľné inými. A naopak, niektoré myšlienky popierajú iné myšlienky. Kognitívny priestor je plný konfliktov, myšlienky ako keby boli vo vzájomnom dialógu.

Vstup človeka do kognitívneho priestoru je sčasti vecou náhody. Závisí od okolností, v ktorých sa k jednotlivým myšlienkam dostáva. Ponúkané myšlienky môže prijať nekriticky, automaticky. Ak sa neskôr dostane k myšlienkam, ktoré odporujú už prijatým, niekto ich automaticky odmieta, iný zvažuje argumenty pre a proti. Umenie argumentácie, zdôvodňovania je prostriedkom kultivácie kognitívneho priestoru.

Nemožno sa však domnievať, že argumentácia nám s definitívnou platnosťou odhalí, ktoré myšlienky sú platné, korektné a ktoré nie. Proces argumentácie nemôže bežať nekonečne dlho, zastavuje sa pri nejakých základných myšlienkach.

Preto môžeme dospieť nanajvýš k záveru, že nejaký argument je prijateľný vzhľadom na nejakú množinu argumentov. Ak k takej množine v dobrom prípade dospejeme, argumenty v tejto množine sa vzájomne podporujú a bránia proti kontraargumentom. Popri tej množine môže existovať iná, ktorá je v konflikte s prvou, ale jej argumenty sa vzájomne podporujú a bránia proti kontraargumentom.

Okrem toho, vyššie opísaný statický pohľad na argumentáciu nie je celkom adekvátny. Povaha reálnej

argumentácie je dynamická. Ľudia hľadajú nové relevantné argumenty a kontraargumenty. Často tento proces prebieha v mysli jediného človeka, ktorý hľadá, utvára svoj názor. Napokon, a predovšetkým, dialógy v reálnom kognitívnom priestore sa neriadia abstraktnými (logickými) pravidlami korektnej argumentácie. Reálna argumentácia v kognitívnom priestore sa výrazne líši od idealizovaného obrazu argumentácie, sformulovaného v [3]. Iba poznamenávam, že v predchádzajúcom odseku som načrtnol myšlienku prijateľného argumentu vzhľadom na množinu argumentov podľa [3].

### 4 Emotívna sila myšlienok a ich spoločenský dopad

Myšlienky nie sú oddelené od emócií. Naopak. Myšlienky sú živé a pôsobivé aj vďaka schopnosti nadchnúť alebo odpudzovať, navodiť príjemný pocit alebo odpor. Dokonca aj myšlienky, čo nezaujímú, robia tak často cez emocionálne pôsobenie – cez nudu až otrávenosť.

Tá istá myšlienka dokáže niekoho nadchnúť, v niekom vyvolať nepriateľské emócie. Niektoré myšlienky dokážu vyvolať davovú psychózu. Zosobniť to môžu Hitler alebo Stalin. Kognitívny priestor je plný nepráv, omylov, predsudkov, polopráv, ale i zámerných lží a demagógie. I tieto jeho vrstvy majú nemalú silu a gravitačné pole.

Posuňme sa od konštatovania emotívneho pôsobenia k identifikovaniu jeho príčin, alebo aspoň k poodhaleniu jeho črt. Na jednej strane, myšlienky vyvolávajú emocionálne reakcie. Na strane druhej, aj emócie, pozitívne či negatívne, záujmy, dobro i zlo vplývajú na vznik myšlienok. Týmto otázkam sa budem venovať v nasledujúcich sekciách – reč bude najprv o patológii kognitívneho priestoru, potom o tom, aké črty vtláča kognitívne priestoru skepsa, oslobodená od podriadenosti autorite spolu so slobodným hľadaním.

### 5 Patológia kognitívneho priestoru

Ak zamýšľame prejsť k tejto téme, netreba sa ani vracieť do histórie, k divnej spleti silného emotívneho pôsobenia myšlienok na množstvo ľudí s krvavými zločinmi štátneho aparátu pod diktátom milovaného vodcu a prvého proponenta tých istých myšlienok. Stačí sa pozrieť na niektoré internetové diskusné fóra, na bahno na Internete.

Jeden z miliónov príkladov: „Halík je nebezpečný havlistický Žid prosazujúci demoliberálny rozklad“.

Niektoré „myšlienky“ fungujú ako útočné zbrane alebo aspoň ako obyčajné nadávky. Netreba sa príliš zamýšľať nad tým, čo vlastne je havlistický Žid, stačí, že nenávisť namierená na Halíka, Havla a Židov je jedným dychom von. Netreba sa starať, čo to vlastne je, ten demoliberálny rozklad, iste je to niečo hrozného a odsúdeniahodného. Slová a „myšlienky“ dokážu pôsobiť na nízke pudy. Aj z nich býva štartovaná emocionálna sila myšlienky.

Pozrime sa na to detailnejšie, vychádzam z úvah v [4].

Patologické vrstvy kognitívneho priestoru sa často opierajú o nejaké mýty a predsudky. Zákernú silu majú nevedome prijímané predsudky, tie fungujú ako skryté predpoklady spôsobu myslenia, ktorý by sme mohli nazvať aj totalitným. Práve preto, že fungujú skryto, ťažko ich identifikovať a eliminovať.

Mýtus vyvolenosti: nejaká komunita je presvedčená – explicitne alebo podvedome – že je predurčená plniť mimoriadne, historické úlohy, že pravda je jej výsadou.

Mýtus vyvolenosti často v komunite zdôvodňuje čosi ako princíp poslušnosti. Keďže my sme tí, čo poznajú pravdu, čo majú jedinečné postavenie a vznešené ciele, preto je v záujme veci, aby všetci členovia komunity poslúchali vodcu alebo nejaké abstraktné normy.

Mýtus vyvolenosti spolu s princípom poslušnosti stavia bariéru medzi komunitou a všetkým, čo je mimo nej. Mýtus nepriateľov je extrémne vyhrotenie takéhoto postoja. Všetci tí, čo nepatria do mojej komunity, sú menejcenní a sú moji nepriatelia. Tento postoj generuje fanatizmus, vytvára atmosféru znemožňujúcu slobodné a kritické myslenie.

So spomínanými mýtmami a predsudkami súvisia falošné alternatívy – buď ideš s nami alebo proti nám. Buď si myslíš a hovoríš to, čo my alebo sú tvoje názory neprijateľné a odsúdeniahodné. Pozri sa na svet touto optikou a zistíš, že veci sú biele alebo čierne.

Ďalšou črtou je uzavretá myseľ, klapky na očiach – všímaj si iba to, čo podporuje bielo-čierne videnie sveta, niektoré fakty si možno všímať, iné nie; niektoré fakty je vhodné prapodivne interpretovať. Ak sú fakty v rozpore s ideológiou, tým horšie pre fakty. Diskusia s oponentmi nie je zaujímavá a prospešná. V politike opozícia nie je tolerovaná.

Živnou pôdou pre patologické vrstvy kognitívneho priestoru sú aj niektoré psychologické vlastnosti a sociálne determinanty ľudí. Sklon k „totalitnému“ spôsobu myslenia majú ľudia, ktorí sa potrebujú

podriadiť nejakej autorite. Pravdu poznajú podľa toho, čo hovorí autorita. Cítia sa dobre, ak sa prispôsobujú. Autorite, vodcovi, moci. Nepreverujú myšlienky, neanalyzujú ich kriticky, ak pochádzajú od autority. Ich podriadenie sa autorite je založené na iracionálnej viere. Názory „nepriateľov“ odsudzujú podľa naučených schém. Často sú to ľudia pokorní a zároveň agresívni, deštruktívni. Správne cítia, ako sa správať voči komu a čomu.

Aby tento obraz nebol bielo-čierny: v reálnych ľuďoch sa tieto črty namiešavajú v rôznej miere. Možno nik nie je úplne imúnny. Nepriaznivé spoločenské pomery dokážu takúto spojitú distribúciu spomínaných črt prevážiť smerom k totalitnému spôsobu myslenia. Alebo aspoň k trpkej koexistencii s ním.

Ďalším dôležitým faktorom, podporujúcim patologické vlastnosti kognitívneho priestoru, je jazyk. Niektoré defekty používaného jazyka podporujú defekty myslenia a myšlienok. A naopak, iné kvality jazyka bývajú dobrou obranou, predpokladom kritického, aj sebareflexívneho, myslenia.

Používanie „veľkých“, krásnych, úžasne znejúcich, ale prázdnych slov vytvára predpoklady pre posun myslenia od sféry kritickej racionality k emocionalite, zbavenej racionálnej kontroly. Vo vágnom jazyku sa lepšie darí lži a demagógii. Argumentácia vo vágnom jazyku je ťažšie kontrolovateľná, komplikuje riešenie konfliktov medzi argumentami a rozhodovanie o prijateľných argumentoch. Zato však uľahčuje lacné pseudovysvetlenia, predsudky, predčasné generalizácie, púhe dojmy, autoritatívne i doktrinárske tvrdenia.

Na rozdiel od racionálneho myslenia, ktoré hľadá fakty a argumenty, vyvracajúce vlastné názory, hypotézy, prijaté teórie, deformované (totalitné) myslenie nepochybuje o vlastných názoroch. So všetkými ostatnými však neznášanlivo polemizuje. Seba toleruje neobmedzene, ale nikoho iného netoleruje. Argumenty vôbec nepotrebuje.

## 6 Skepsa, hľadanie a sloboda

Skôr než sa dostanem k téme, mal by som objasniť, ako odlišujem pozitívny a negatívny vzťah, postoj k myšlienkam. Kritériom, ktoré ich (pre mňa) odlišuje, je otvorenosť (uzavretosť). Príznakom pozitívneho postoja je ochota hľadať podnety, poznatky, koncepcie v čo najrozličnejších spoločenských a myšlienkových prúdoch. Samozrejme, otvorenosť voči vplyvom nestotožňujem s eklektizmom. Uzavretosť pred vonkajšími vplyvmi a „cudzími“ myšlienkami považujem za negatívny postoj.

Orientácia v kognitívnom priestore je veľmi komplikovaná. Popri jeho konfliktosti, danej dialogickou povahou, si možno vybrať z rôznych možností, ktoré nie sú v priamom konflikte. Ako príklady možno uviesť rôzne paradigmy životného štýlu, umelecké smery, filozofické postoje, či školy. Rozdiely sú v jemnostiach, dôraze, hľadisku.

Aj ucelené sústavy myšlienok majú ďaleko k argumentačnému ideálu sformulovanému matematickou logikou. Konzistentnosť nie je zaručená. Ak o niektorých ich častiach možno povedať, že tvoria základy sústavy, ich pozícia sa líši od statusu matematických axiém.

Kognitívny priestor je neúplný priestor, je otvorený voči novým myšlienkam, voči revíziám a modifikáciám, je charakteristický veľkou mierou neurčitosti. Sú v ňom aj nepravdy, zámerné lži, myšlienky bez zreteľného významu.

Dôležitým orientačným prostriedkom je skeptický postoj k myšlienkam. V dejinách myslenia v tejto súvislosti žiaria napríklad Sokrates alebo Descartes. Každá myšlienka, aj vlastná, môže byť spochybnovaná, podrobená kritickému analýze. Jej pravdivosť sa testuje v konfrontácii s možnými kontraargumentami a falzifikujúcimi pozorovaniami. Na druhej strane, dôvera v jej pravdivosť sa zvyšuje vďaka dôkladnej argumentácii, ktorá vezme do úvahy čo najviac faktov a solídne potvrdených koncepcií (teórií).

Vecný jazyk umožňuje alebo uľahčuje kontrolu myslenia, kritickú analýzu, korektnú argumentáciu. Pod vecným jazykom nemám na mysli jazyk matematiky, ani hovorový jazyk, zneprehľadnený množstvom definícií. Veľmi často sa stáva, že kľúčovitá snaha o „presné“ definície vedie k tomu, že relatívne jasné pojmy sa definujú oveľa nejasnejšími pojmami. Treba sa zmieriť s tým, že o niektorých doménach je lepšie hovoriť vecným, ale bežným jazykom. Pokúsim sa povedať, čo mám na mysli pod vecným jazykom. Je to jazyk, ktorému sa usilujeme dať nejaký empirický zmysel. Keď ho používame, máme na mysli nejaké pozorovateľné situácie, sme schopní uviesť nejaké príklady a možné kontrapríklady, sme schopní postupného upresňovania, ak na tom záleží.

Prijaté, podopreté myšlienky neprinášajú definitívne uspokojenie, ale inšpirujú hľadanie. Do šírky i do hĺbky.

Cesta do šírky je menej zaujímavá, i keď dôležitá. Na nej sa objavujú nové fakty, nové detaily. Nachádzajú sa

lepšie vyjadrenia, upresnenia, drobné revízie.

Cestou do hĺbky sa hľadajú zdôvodnenia a vysvetlenia pre doposiaľ prijaté fundamenty, hľadajú sa alternatívne pohľady alebo analogické, či všeobecnejšie pohľady. Hľadajú sa revidované, spoľahlivejšie, presvedčivejšie základy.

Hĺbka sa dosahuje hľadaním súvislostí, kontraargumentov, robustnejších zdôvodnení, nadhľadom, kritickou (auto)reflexiou.

Pohybom do hĺbky je aj, možno predovšetkým, vybočenie z rámca nejakej paradigmy.

Základným postojom hľadajúceho skeptika je vedomie slobody. Slobodný a nezávislý človek nie je viazaný rešpektovaním žiadnej autority, pocitom podriadenosti alebo poslušnosti. Je presvedčený, že môže pochybovať o čomkoľvek. Vie, že popri pochybovaní je dôležité aj hľadanie. Hľadanie odpovedí na pochybnosti, otázky, hľadanie zaujímavých otázok, nových hypotéz a riešení. I hľadanie v známom kognitívnom priestore – hľadanie presvedčivých riešení, ktoré tam už sídlia.

Slobodne mysliaci človek nestavia hranicu medzi „my“ a „oni“. Nemyslí si, že iba my máme pravdu a oni sú naši nepriatelia. Naopak, hľadá podnety, inšpiráciu u odlišných koncepcií sveta a duchovna.

Hľadajúci, slobodný skeptik nie je bez emócií. Naopak, zaujímavé myšlienky ho emotívne motivujú a hľadá, ako sa dostať ďalej. To je pre neho zdrojom pozitívnych emócií. Potešenie má aj z korektnej a presvedčivej argumentácie. Je nespokojný, keď cíti vlastné nedostatky alebo nedostatky myšlienok, ktoré ho zaujali.

Najdôležitejšie však je emotívne zasadenie sa za myšlienky, ktoré majú potenciál pozitívneho pôsobenia na spoločnosť, nie iba na malé komunity. Pôsobenie na obraz sveta, šírený v spoločnosti. Na spôsob rozmýšľania, ale aj konania, pozitívny vplyv na spôsob a kvalitu života ľudí.

## 7 Zodpovednosť intelektuála

Na začiatku som spomínal, že tento text je motivovaný verejným vystúpením dvoch mužov. Verím, že ich pôsobenie a myslenie vyžaruje pozitívne do celého českého a slovenského spoločenstva. Napriek tomu, že majú dosť odporcov, a dosť hlasných.

Spomínaný zážitok pôsobil na mňa intenzívne asi aj

preto, že prebehol v čase vrcholiacich prvých priamych volieb českého prezidenta a súčasne v čase, keď som v duchu márne hľadal presvedčivého a dôstojného kandidáta na slovenského prezidenta. Ešte treba dodať, že obaja muži vstúpili do celospoločenskej diskusie.

Pokúsím sa načrtnúť povahu tej celospoločenskej diskusie a ich prínos do nej.

Samozrejme, naša celospoločenská diskusia je bohato štruktúrovaný podpriestor kognitívneho priestoru. Sú v ňom zaujímavé až pozoruhodné komunitné diskusie. Tu sa sústreďujem na jej – v istom zmysle slova – prikrývku alebo skôr prekryvku produkovanú politickými elitami, komerčnými televíziami, bulvárnymi médiami.

Politické elity prichádzajú s témami, ktorým by súdny človek vôbec nemal venovať pozornosť. Na druhej strane, keď s niečím prichádzajú „vodcovia“, ani solídnejšie média a premýšľajúci ľudia si to nemôžu nevnímať. A tak sa do celospoločenskej diskusie dostáva, či opozícia má alebo nemá spoločného kandidáta, i keď má iba jedného. Dojímavé starosti o jednomyselné hlasovanie opozície sa v inom kontexte prevetelia do rozhorčeného kriku toho istého krku o slobodnom mandáte poslanca. Okolo takýchto pseudoprotblémov sa to točí mesiace v zamorenom priestore verejnej diskusie.

Myšlienky, ktoré politické elity púšťajú do celospoločenskej diskusie, ostávajú zväčša na povrchu lacných hesiel, spoliehajú sa na nedostatočnú kritickosť tých, ktorým boli adresované a hlavne na to, že takéto reči adresáti chcú počuť. Sú prednesené s jediným úmyslom, aby zapôsobili na značnú časť populácie. Už tesne pod povrchom lacné heslá narážajú vzájomne na seba. Autori sa nestarajú o to, či sú konzistentné s tým, čo predniesli inokedy alebo aj práve v tom istom vystúpení, ani o to, aby boli podopreté serióznymi myšlienkami, aby zapadali do nich.

Tak sa stáva, že ľúbivé heslá súhrnne strácajú zmysel. Ak niekto chce byť hlasom desiatich nepriviligovaných miliónov, pýtame sa, či chce byť aj hlasom ostro napadnutých nacistických bojúvk, novinárov, závistlivých hlupákov? Aký je skutočný rečníkov vzťah k desiatim miliónom ľudí, medzi ktorými sú pozitívne iba nejaké ostrovy deviácie?

Od vrcholných predstaviteľov štátu by človek očakával, možno naivne, príklad hlbokej mysli, ktorá by pôsobila pozitívne na krajinu, na spôsob myslenia, ale aj konania ľudí.

Spomínaní dvaja muži na mňa zapôsobili slobodnými

postojmi, schopnosťou pochybovať a súčasne vnášať do verejného priestoru netradičné myšlienky a skutky. Každý z nich je výrazne iný, ale obaja obdivuhodne vyčnievajú. Vyžarujú kúzlo slobodných osobností, ktoré majú svoj názor, predstavy ako nezávisle konať, ktorí hľadajú a ktorí ponúkajú nádej.

Tomáš Halík a Robert Bezák sú tí dvaja muži.

Vôbec nie preto, že sú to duchovní. Naopak, obdivujem ako výrazne sa líšia od priemerného duchovného.

Vo všeobecnosti, kognitívny priestor obohacuje inakosť, schopnosť vymaniť sa z bežných klíšé a šablón, klásť si neobvyklé otázky a hľadať odpovede na ne. A súčasne ponúkať pozitívne emočné i racionálne podnety.

Tu mi nejde o to, aby som robil advokáta Bezákovi alebo Halíkovi, to nepotrebujú. Iba konštatujem – zapôsobili na mňa tak silne, že som sa rozhodol zamyslieť nad kognitívnym priestorom. Niektoré kľúčové slová tohoto textu – skepsa, hľadanie, hĺbka, nádej, sloboda – som odporoval ako podčiarknuté u Tomáša Halíka.

Sú aj vzorom pre intelektuálov, zahrabaných v milých a zaujímavých zábavkách ako ABF, aby sa hlasno vyjadrovali k defektom spoločnosti, v ktorej žijú a aby pred ňou vyzdvihovali pozitívne vízie.

## Pod'akovanie

Ďakujem programovému výboru za pozvanie na KUŽ 2013.

Ďakujem za podporu grantu VEGA 1/1333/12.

## Literatúra

- [1] A. Bondarenko, P.M. Dung, R. Kowalski, F. Toni: An Abstract, Argumentation-Theoretic Approach to Default Reasoning. *Artificial Intelligence* 93 (1997) 63-101
- [2] V. Lifschitz: Logic programming solution to the frame problem. <http://www.dis.uniroma1.it/~kr12/downloads/invited/lifschitzKR12.pdf>
- [3] P.M. Dung: On the Acceptability of Arguments and its Fundamental Role in Nonmonotonic Reasoning, Logic Programming and n-Person Games. *Artificial Intelligence* 77, 2 (1995) 321-358
- [4] J. Šeřránek: Duch totality. *Reflex* 3, 6.5.1969. Tiež aj *Filozofia* 46, 2 (1991) 212-216

## Dopravní chování jako úkol pro agentní simulaci

Michal Šimeček

Centrum dopravního výzkumu v. v. i.  
Líšeňská 33a  
michal.simecek@cdv.cz

### Abstrakt

Příspěvek se zabývá analýzou dopravního chování, která má sloužit pro vytváření modelu dopravní poptávky. Tradiční modely zohledňují jen některé z pravidelností, které se v lidském dopravním chování mohou najít. Tvorby dopravních modelů pomocí agentní simulace nabízí další možnosti, jak modelovat individuální účastníky dopravy ve větším detailu. Dostáváme se tak ale do oblastí, které dosud nebyly dostatečně prozkoumány a kde lidské chování nabývá na složitosti a komplexní provázanosti. Budou představeny vybrané aspekty lidského dopravního chování, které stojí za to modelovat. Zároveň budou nastíněny funkční moduly, pomocí kterých by měl agent toto chování napodobit. Tyto dílčí moduly budou zasazeny do nástinu celkového modelu agenta.

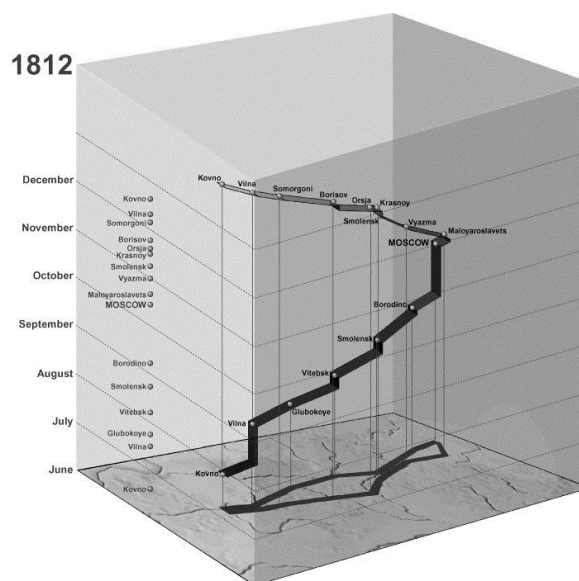
### 1 Úvod

Modelování dopravy se rozvíjí od 50. let dvacátého století v USA jako reakce na poválečnou ekonomickou konjunkturu, spojenou s nárůstem automobilové dopravy. Cílem těchto modelů je předpovídat budoucí stav dopravní poptávky a ukazovat dynamiku systému dopravy. Klasické modely předpokládaly rozdělení území do tzv. funkčních okrsků (například obytná nebo průmyslová zóna). Toto rozdělení málo odpovídá struktuře těch evropských měst, které nebyly zdevastovány socialistickým plánováním. Ve městech, která se rozvíjí organicky, jsou funkce území značně promíchány. Klasické modely jsou přezdívány čtyřstupňové, protože svoji funkčnost realizují ve čtyřech krocích. Prvním krokem je generování cest z jedné zóny do druhé. Ve druhém kroku se oddělí množství přepravy jednotlivým druhem (módem) dopravy. Třetí krok rozdělí vygenerované cesty do příslušných komunikací mezi zónami. Teprve ve čtvrtém kroku se simuluje zatížení dopravní sítě. Klasické dopravní modely neuvažují jednotlivé aktéry v dopravě, ale pracují spíše s množstvím přepravních požadavků, které je potřeba v tom kterém čase realizovat pomocí dopravní sítě. Hovoříme o agregovaných modelech. V poslední době se rozvíjí

modely, které simulují jednotlivé dopravní aktéry (tzv. disagregované modely). Tyto modely, mají-li vést k větší přesnosti a robustnosti modelů, ale potřebují nové zdroje dat o mobilitě [3].

### 2 Modely založené na dopravně-aktivitním přístupu

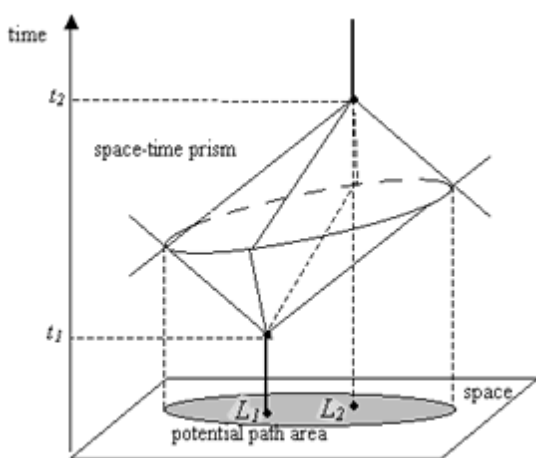
Dopravní chování lidí se odehrává v časoprostorovém kontinuu. Dráze jednotlivce v tomto kontinuu se říká cesta či trasa (path). Tyto cesty mají charakter kontinuálního procesu ve vymezeném časovém intervalu, lze si je představit jako navazující sekvenci aktivit, které mohou mít nejrůznější charakter, význam a dobu trvání [7][8].



Obr. 1. Prostorově-časová krychle (Kraak, 2003)

Stejně, jako lidské dopravní chování nelze oddělit od prostoru a času, ve kterém se odehrává, je potřeba uvažovat také o tom, že člověk produkuje dopravní chování za nějakým cílem. Hodlá na příslušném místě a

v příslušném čase vykonávat určitou aktivitu. Právě časoprostorové omezení aktivity jej nutí k pohybu. Problémem se zabýval švédský kartograf Torsten Hägerstrand, který mluvil o geografii času. S geografii času souvisí časoprostorová omezení, která Hägerstrand definoval. V určitém časovém úseku se může člověk dostat z posledního známého bodu určitým prostředkem pouze do určité vzdálenosti. Jak časový úsek narůstá, roste i hraniční vzdálenost. V časoprostorovém kontinuu vzniká trychtýřovitý útvar omezující možný výskyt sledovaného člověka. Známe-li počáteční pozici ( $L_1$ ) a budoucí pozici ( $L_2$ ), vznikne v časovém úseku mezi  $t_1$  a  $t_2$  útvar, složený ze dvou spojených kuželů (space-time prism). V takto definovaném časoprostoru se mohl sledovaný člověk pohybovat (Obr. 2).



Obr. 2. Prostorovo-časové znázornění oblastí možného pobytu mezi současnou lokací ( $L_1$ ), budoucí známou lokací ( $L_2$ ) a mezi současným a budoucím známým časem ( $t_1$  a  $t_2$ ).

Uvažování časoprostorových omezení je důležité pro realističnost disagregovaných modelů. V případě modelů agregovaných, kde nemodelujeme jednotlivé individuální cesty, nehrají časoprostorové limity podstatnou roli.

### 3 Dopraveně-aktivitní vzorce

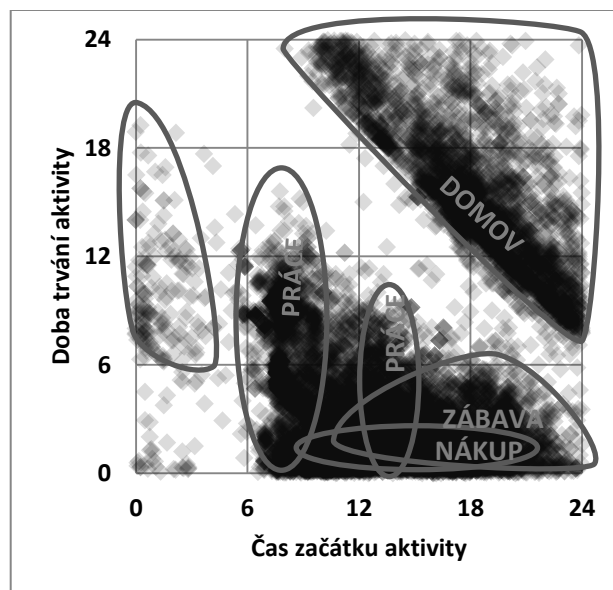
Časoprostorové vzorce, které lidé produkují, nemají náhodný charakter, ale jsou určitým způsobem organizované. Nenáhodnost bude více zřejmá, pokud poznáme, jaké aktivity jsou právě realizovány. Na obrázku 3 je mapa tras, které obvykle realizují odpoledne: dříve to byla trasa mezi body 1-2-4 nebo 1-3-4, dnes jsou to trasy 5-6-4 nebo 5-3-4. Podobnost mezi trasami začne být zřejmá, když budeme vědět, že v místě 4 je domov, v místě 1 bylo staré a 5 je nové zaměstnání. V místech 2 a 6 jsou velké samoobsluhy a v místě 3 je malá samoobsluha.

Jde tedy ve všech případech o cestu z práce domů, při které je realizován běžný nebo poněkud větší nákup.



Obr. 3. Trasy po Brně.

Na úrovni realizovaných aktivit se tedy vzorec příliš nemění. Aktivitní přístup na rozdíl od klasických modelů dokáže popisovat takové jevy, jako je změna zaměstnání. Pokud si zobrazíme aktivity lidí podle jejich začátku a délky trvání (viz obrázek 4), zjistíme, že existují vzorce, které se v menší nebo větší míře opakují prakticky u všech.

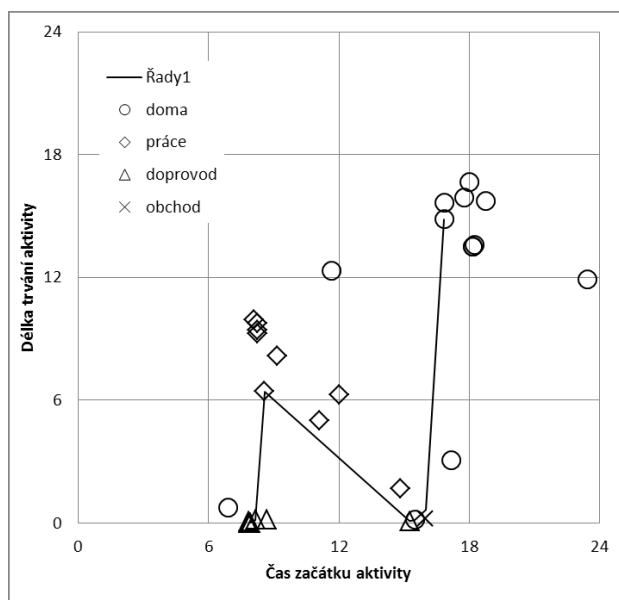


Obr. 4. Závislost druhu aktivity na čase jejího začátku a době trvání (data pochází z dopravního průzkumu v Pardubicích).

Lidé obvykle neopouští domov před šestou hodinou. Často pak míří do zaměstnání, které někteří opouští kolem poledne (pauza na oběd). Nákupní aktivity se

realizují zhruba od osmé hodiny do večera a netrvalí dlouho. Za zábavou začínají lidé chodit před polednem, ale poslední aktivity zábavy začínají až dlouho do noci. Nejdelsí aktivity, které se týkají zábavy, lidé obvykle zahajují kolem sedmé hodiny večerní a stráví na nich až šest hodin. Domů se snaží dostat co nejdříve, aby zde mohli strávit co nejvíce času. Jednotlivé skupiny obyvatel (např. zaměstnaní, v domácnosti, studenti, lidé dle výše příjmů) mohou vykazovat vzorce, které si jsou vzájemně podobné. Podobné výsledky, jako jsou zobrazeny na obrázku 4, můžeme získat i z dopravních průzkumů z jiných oblastí. Podobný graf lze získat i z dat z průzkumu Mobidrive [2], který byl uskutečněn ve Švýcarsku.

Dopravně-aktivitní vzorce jednotlivých lidí jsou ovšem ještě mnohem jednodušší. Uvedme příklad dvou týdnů muže, který ráno obvykle vodí dítě do školy a pak pokračuje do práce (obrázek 5). Z práce někdy vede dítě ze školy, ale většinou jde rovnou domů. Toho dne vedl dítě ze školky a stávil se v obchodě.

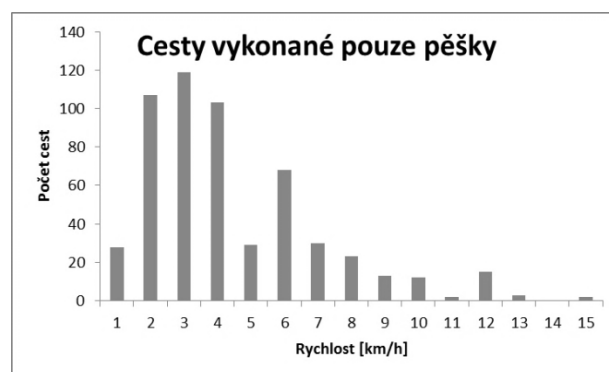


Obr. 5. Dopravně-aktivitní vzorec jednoho člověka. Čárou jsou vyznačeny aktivity v jednom dni.

### 3 Kognitivní výbava agenta

Pokud redukuje lidskou mysl pouze na ty funkce, které umožňují vznik dopravního chování, co musí umět? Především musí být schopen si podle potřeb naplánovat aktivity (které mají tyto potřeby saturovat). K tomu musí mít nějakou představu o území, kterým se má pohybovat a musí mít představu o dopravních vlastnostech různých

tras. Území tedy musí být reprezentováno jakousi kognitivní mapou, která umožní identifikovat vhodný cíl cesty a zvolit k ní vhodnou trasu a cestovní prostředek. Tato kognitivní mapa není úplná. Člověk a ani agent nemůže znát veškeré detaily o světě, ve kterém se pohybuje a brát je do úvahy. Bylo by obtížné, aby agent při každém rozhodování nad mapou procházel všechny možné alternativy. Realističtější a také uživanější je využití tezí omezené racionality a znalost o světě agentovi omezit na obvyklá místa a trasy. V praxi jsou tyto habituální prvky kognitivní mapy proti ostatním značně zvýhodněny [6]. Mapa rovněž nemůže být spolehlivá, aktualizovat by se měla teprve na základě zkušenosti, kterou agent udělá při realizaci svých denních plánů. Data z Pardubického dopravního průzkumu pouze pro cesty učiněné pěšky ukazují značný rozptyl v cestovní rychlosti, která byla vypočítána z odhadů časové a geografické délky cesty. Tyto odhady dělali sami respondenti průzkumu a vyplývalo by z nich, že cirká čtvrtinu cest lidé z Pardubic v podstatě běhají (Obr. 6.). Lepším vysvětlením může být, že lidé mívají poměrně špatnou představu o prostorovém uspořádání (vzdálenostech) v okolí svého bydliště, na rozdíl od představy časové (např. za jak dlouho se někam dostanou pěšky).



Obr. 6. Rychlost je počítána z časového a prostorového údaje délky cesty. Tato data byla získána dotazováním od respondentů v dopravním průzkumu v Pardubicích.

Časový údaj trvání dopravy mezi dvěma body je pro plánování cest důležitější, než prostorová vzdálenost. Zahrnutím časových délek cest do reprezentace vzniká poněkud odlišná mapa od mapy geografické. Proti geografické mapě je tato mapa pokrivená. Vzdálenosti na časové mapě jsou doby přepravy, které mohou být v nesouladu s geografickými vzdálenostmi.

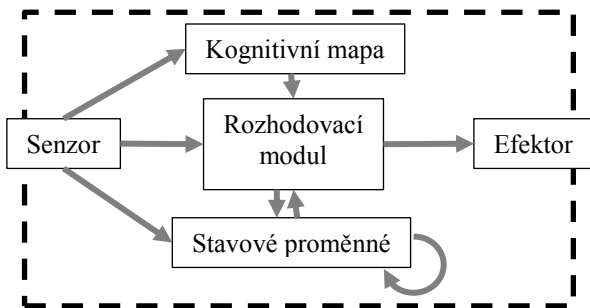
Naše představa je, že by agent měl mít k dispozici mapu území, složenou z uzlů a hran, jak bývá definováno území v modelech dopravy. V uzlech budou umístěny

příležitosti k příslušným typům aktivit, jejich časová omezení (otvírací hodiny obchodu) a indexy užitečnosti provozování aktivity právě na tomto místě. Hrany budou mít atributy očekávané doby průjezdu v příslušném dopravním módu a s tím spojeným komfortem.

Aby agent mohl plnit svoji kognitivní mapu aktuálními informacemi, schopnost přijímat potřebné informace ze svého okolí pomocí senzoru.

Dále by měl agent mít soubor stavových proměnných, které obsahují momentální naléhavost uspokojení určitých potřeb, ale také informaci o tom, jestli má právě v tomto okamžiku k dispozici (u sebe) například automobil nebo permanentní jízdenku MHD. Zároveň musí existovat pravidla pro zacházení s těmito prostředky. Bylo by třeba nesmyslné, aby agent přijel autem k zastávce MHD, strčil si své auto pod paži a nastoupil do tramvaje.

Situaci poněkud komplikuje skutečnost, že lidé nežijí odděleně od druhých a i své dopravní chování koordinují se svým sociálním okolím. Mohou se vzájemně zastupovat při naplňování potřeb (např. nákup pro společnou domácnost), mohou sdílet dopravní prostředky (např. společný automobil nebo přenosnou jízdenku MHD), mohou sdílet své přepravy nebo aktivity (např. doprava taxislužbou). Bude tedy nutné umožnit některým agentům, aby mohli své dopravní plány společně koordinovat. Aby to mohli dělat podobně, jako to dělají lidé v dnešní realitě mobilních telefonů, musí jim (některým) být umožněna vzájemná komunikace na dálku. Tuto komunikaci je potřeba zahrnout do senzoru a efektoru agenta v podobě možnosti posílat a obdržet zprávy podle předem stanoveného komunikačního protokolu.

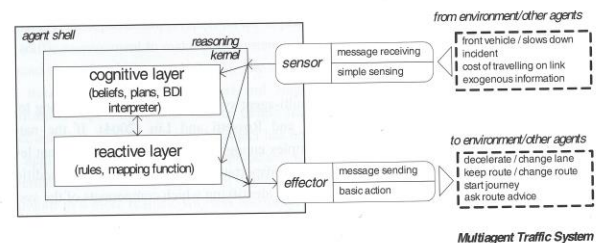


Obr. 7. Nástin architektury agenta simulujícího dopravní chování.

Všechny tyto moduly slouží k rozhodování o jednotlivých cestách agentního modelu jedince. Samotné rozhodování má na starost rozhodovací modul, který bývá v agentních simulacích dopravně-aktivitních vzorců

modelován na základě mikroekonomické teorie užitku [1][4]. Rozhodování se týká toho, kdy a jaká aktivita bude uskutečněna, na kterém místě, jakým dopravním prostředkem a kudy se na příslušné místo agent vypraví. Vznikne sada možných alternativ, pro které se vypočítá užitková funkce a po té se zvolí nejvýhodnější varianta. Výhodou využití teorie užitku je možnost díle části rozhodovacího procesu ověřit na základě studia preferencí pomocí analýzy diskretních voleb [1].

Naše současná představa o základní architektuře agenta je zachycena na obrázku 7. Šipky znázorňují tok informace mezi moduly. Informace ze senzorů vedou k aktualizaci kognitivní mapy. Komunikační zprávy mezi agenty pronikají přímo do rozhodovacího modulu. Přímé spojení mezi senzorem a modulem stavových proměnných aktualizuje stav potřeb. Pokud sensor zjistí, že nějaká potřeba je saturována aktivitou, modul stavových proměnných situaci aktualizuje. Některé stavové proměnné mohou nabývat na významu s postupem času (únava, hlad). Rozhodovací stroj využívá informaci z kognitivní mapy k rozhodování o naplnění potřeb v modulu stavových proměnných. Efektor je oslovován pouze rozhodovacím modulem, řeší základní akce a odesílá zprávy příslušným agentům. Sensor a efektor tedy lze rozdělit na složku vnímání/akce a na komunikační část, jako je tomu u modelu na obrázku 8.



Obr. 8. Dvourvrstvá architektura agenta modelujícího chování řidiče (Rossetti, Liu, 2005).

## 4 Závěr

Jsme nyní na začátku tvorby architektury agenta. Identifikovali jsme základní požadavky, které by měl agent splňovat, aby dokázal věrohodně simulovat dopravní chování lidí. Z požadavků vyplývají vlastnosti vstupů a výstupů modelu, které kromě jednoduché senzory a jednoduchých akcí musí zvládat také příjem a odesílání zpráv.



Plánovaný model sice nepůjde až na úroveň mikrosimulací dopravního chování (simulaci dopravního proudu), měl by však být schopen simulovat řadu dílčích úkolů, jako je hledání parkování, reakce na objíždnou trasu. Některé z těchto úkolů bude možná lepší simulovat místo uživatelských funkcí přímým mapováním senzorických vstupů na jednotlivé základní akce, jako je tomu u modelu Rossettiho a Liu [6].

Agentní architektura autorů Rossettiho a Liu [6] je rozdělena do dvou vrstev. Vyšší (kognitivní) vrstva zajišťuje komplexní rozhodování jako je volba příští aktivity a volba plánu, který se bude následovat. Nižší (reaktivní) vrstva obsahuje sadu jednoduchých pravidel, které přímo propojují percepci s akcí. Autorům slouží tato úroveň k simulaci takových dějů, jako je jízda v koloně nebo změna silničního pruhu.

## Poděkování

Tento příspěvek vznikl za podpory grantové agentury Technologické agentury České republiky v rámci grantového úkolu vav-rodos TE01020155.

## Literatura

- [1] Ben-Akiva M., Lerman S. R. (1985). *Discrete Choice Analysis: Theory and Application to Travel Demand*. The MIT Press, ISBN-13: 978-0-262-02217-0.
- [2] Chalasani, V.S. and Axhausen, K.W. (2004) *Mobidrive: A six week travel diary*, Travel Survey Metadata Series, 2, Institute for Transport Planning and Systems (IVT); ETH Zürich, Zürich.
- [3] Jovicic G. (2001): *Activity based travel demand modelling – a literature study*. Denmark TransportForskung. ISBN 87-7327-055-5.
- [4] Karlström a. (2005): *A Dynamic Programming Approach for the Activity Generation and Scheduling Problem*. 25 – 42. In Timmermans H. [ed.]: *Progress in activity-based analysis*. Elsevier.
- [5] Kraak M. (2003): *Geovisualization illustrated*. *Journal of Photogrammetry and Remote Sensing*, Volume 57, Issues 5–6, April 2003, Pages 390–399.
- [6] Rossetti R. J. F., Liu R. L. (2005): *Activity-based Analysis of Travel Demand Using Cognitive Agents*, 139-160. In Timmermans H. [ed.]: *Progress in activity-based analysis*. Elsevier.
- [7] Thrift, Nigel J (1977): *An introduction to time-geography. Concepts and techniques in modern geography*, 13. Norwich: Geo Abstracts, University of East Anglia.
- [8] Yu, Hongbo; Shaw, Shih-Lung (2008): "Exploring potential human activities in physical and virtual spaces: a spatio-temporal GIS approach". *International Journal of Geographical Information Science* 22 (4): 409–430.



## Prostor ve fotografii a v obrázku

Michal Šimeček

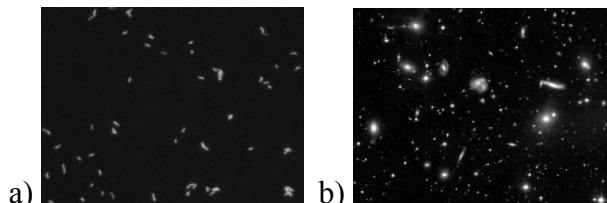
Psychologický ústav AV ČR v. v. i.  
 Veveří 967/97, Brno  
 michal.simecek@volny.cz

### Abstrakt

Vnímání prostoru na obrázku nebo fotografii se v mnoha rysech liší od přímého zrakového vnímání prostoru. Rozdíly jsou dosti značné, ačkoli si je většinou neuvědomujeme. Kromě naší kulturní zkušenosti s realistickými obrázky a fotografiemi, která takové vnímání prostoru umožňuje, jsou ve hře další faktory. Patří sem jednak vztah mezi zobrazovaným prostorem a tím, kdo jej zobrazuje (kamerou), ale také relace mezi výsledným obrazem (fotografií) a pozorovatelem (například velikost obrazu a pozorovatelova vzdálenost od něj). Tato složitost neumožňuje například pozorovateli z fotografie spolehlivě odhadovat egocentrické vzdálenosti jinak, než na pořadové úrovni, a i při dalších odhadech jej nutí zapojovat vlastní percepční zkušenost víc, než ve skutečnosti situace dovoluje.

### 1 Úvod

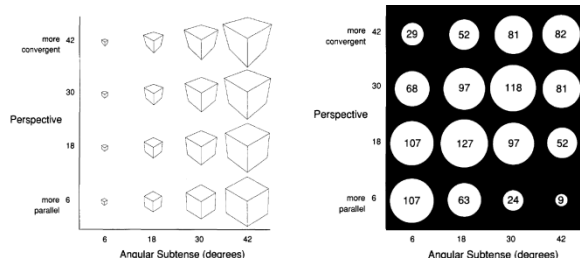
Pozorovatel vnímá scénu vyobrazenou na ploše obrazu, scéna však, pokud jde o její prostorové vztahy, není na obraze určena bez možnosti alternativních výkladů. Pozorovatel má k dispozici řadu pomůcek, které mu umožní identifikovat v obraze jeho prostorovou hloubku. Jedná se o takzvané obrázkové nápovědi, jako jsou perspektivní zkreslení, gradienty textur, okluze a další. Tyto nápovědi však podávají neúplné a nepřesné informace. Některé z vlastností zobrazení scény souvisí s vlastnostmi zobrazovacího zařízení (ať je jím optické fotografické zařízení nebo oko, mysl a ruka výtvarníka), které obraz pozorovateli zprostředkovává. Tohoto prostředníka nemá ovšem pozorovatel pod kontrolou. Extrémní případ je na obrázku 1, kdy musí pozorovatel rozhodnout o zcela zásadně odlišných měřítcích obou scén. Obvykle jsou interpretativní rozhodnutí pozorovatele mnohem jemnější a také případně chybné závěry, ke kterým může dojít, se od reality tolik neodchylují. Pokud zobrazení není extrémní a představuje spíše všednodenní pohled, zastupuje kamera v mysli pozorovatele jeho vlastní zrak. To může vést k některým předčasným interpretacím vnímaného obrazu, a k pozorovatelným chybám, které z toho vyplývají.



Obr. 1. Obrázek nemusí informovat o měřítku zobrazené scény: mikroskopické zobrazení bakterií tuberkulózy (a), fotografie kupy galaxií v souhvězdí Panny (b).

### 2 Velikost obrazu

Stejně, jako pozorovatel nezná přesně výškové umístění kamery, nezná také parametry objektivu, respektive parametry části obrazu, který je mu prezentován. Bude nás nyní zajímat vztah úhlové velikosti záběru a úhlové velikosti zobrazení scény v oku pozorovatele. Yang a Kubovy [13] prokázali, že percipovaná adekvátní míra perspektivního zkreslení zobrazené scény je závislá na její úhlové velikosti. V jejich experimentu měly pokusné osoby rozhodovat, který ze dvou prezentovaných objektů vypadá spíše jako krychle. Objekty přitom variovaly v míře perspektivního zkreslení a v úhlové velikosti. Obrázek 2 ukazuje podnětový materiál a počty, kolikrát byl pokusnými osobami ten který objekt považován za adekvátní krychli. Z výsledků vyplývá, že čím je pozorovaný objekt (scéna) úhlově větší, tím větší perspektivní zkreslení je považováno za adekvátní.



Obr. 2. Vztah mezi úhlovou velikostí a perspektivním zkreslením (upraveno podle Yang a Kubovy, 1999).

V Ohnakově a Matsudově experimentu [6] autoři neshledali vztah mezi impresí hloubky fotografie a jejím zvětšením, ale při dotazu na absolutní vzdálenost místa ve fotografické scéně od pozorovatele hrála velikost fotografie roli. Vnímaná absolutní vzdálenost místa ve scéně klesala se zvětšujícím se formátem fotografie. Stejný efekt nastává při zvětšující se ohniskové vzdálenosti objektivu [5], což rovněž vede ke zvětšení obrazu.

### 3 Vzdálenost obrazu od pozorovatele

Ohnaka a Matsuda [6] při experimentování s percipovanou absolutní vzdáleností na fotografii zjistili, že vnímání těchto vzdáleností závisí kromě velikosti obrázku také na vzdálenosti pozorovatele od něj. S rostoucí vzdáleností obrazu od pozorovatele se vnímaná absolutní vzdálenost objektu ve scéně zmenšovala. Podobné zjištění již dříve učinil Smith [8][9].

Nabízí se zde myšlenka, že klíčovou roli sehrává úhlová velikost obrázku na sítnici pozorovatele, která je funkcí jak skutečné velikosti obrazu, tak jeho vzdálenosti od pozorovatele.

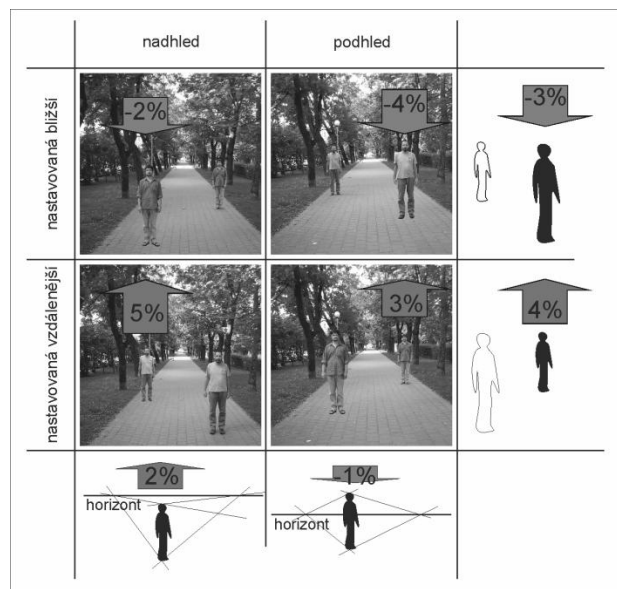
### 4 Zakřivení vizuálního prostoru

Zrakem vnímaný prostor není stejný ve všech směrech. Efekt hloubkové komprese je studován již dlouhou dobu (např. [2][1]). Hloubková komprese znamená, že percipovaná vzdálenost v hloubkovém směru od pozorovatele se jeví pozorovateli jako kratší, než je tomu ve skutečnosti. S rostoucí distancí tento rozdíl proporčně narůstá. Ve frontálních směrech k tomuto jevu nedochází (Loomis et al., 1992). Mění se však vnímané orientace, což ukazuje na zakřivení vnímaného prostoru [11]. Watanabe [12] dokládá, že vnímání fotografické scény podléhá těmto zkreslením mnohem výrazněji, než vnímání scény reálné nebo vnímání fotografické scény viděné stereoskopicky.

### 5 Výška záběru

Zastupuje-li kamera oko pozorovatele, bude zobrazená scéna vnímána tak, jako by byla nazírána z výšky jeho vlastních očí. V našem experimentu jsme do fotografické scény bohaté na prostorové nápovědi umístili dvě stejné postavy do různých vzdáleností od kamery. Tato scéna se promítala na monitoru počítače. Probandův úkol byl nastavit pomocí klávesnice velikost jedné z postav tak, aby odpovídala svojí velikostí scéně. Měřitko scény bylo

určené druhou postavou. Všechny položky bylo možné rozdělit do čtyř skupin podle dvou dichotomických charakteristik: zda byla nastavovaná postava blíž nebo dál a zda byla scéna zabírána nadhledu nebo podhledu (Obrázek 3).



Obr. 3. Jednotlivé typy položek experimentu, kde měli proband za úkol nastavovat velikost jedné ze dvou stejných postav ve scéně. Čísla představují mediány procentuálních rozdílů proti požadované hodnotě.

Rozdíly mezi položkami, kde se nastavovala bližší nebo vzdálenější postava, ukazují na hloubkovou kompresi vnímaného prostoru na fotografii. Pokusné osoby nastavovaly postavy tak, aby si jejich úhlová velikost byla podobnější, než by odpovídalo perspektivnímu zřeslení scény [7]

Při porovnání scén zabíraných z nadhledu nebo podhledu byly opět nalezeny významné rozdíly. V obou případech pokusné osoby nastavovaly velikost adjustované postavy tak, aby scéna více odpovídala pohledu z výšky očí [7].

Kromě toho se ukázalo, že jsou výsledky tohoto experimentu závislé na skutečné výšce pokusných osob. Zatímco pozorovatelé malého vzrůstu neměli potíž nastavovat výšky (průměrně) vysokých osob zabíraných z podhledu, potíže jim dělaly scény z nadhledu. U vysokých pokusných osob byla sice opačná tendence, ale neprojevila se tak výrazně.

Domníváme se, že zatímco lidé malého vzrůstu mají zkušenost s vnímáním druhých lidí z podhledu, zkušenost s nadhledem mají omezenou. Naproti tomu vysocí lidé mají zkušenost jak s nadhledem (jsou vyšší,

než většina ostatních), tak i s podhledem, pokud například sledují stojící osoby, když sami sedí.

Pokud scéna kromě osob neposkytuje další jednoznačné a přesné informace o měřítku (relevantní objekty známé velikosti), může do výsledků vstupovat také pozorovatelův odhad skutečné velikosti zobrazených postav na obrázku, se kterým experimentovali Kato a Higashiyama [4].

## 6 Velikost objektů

Prostorové odhady ve fotografické scéně jsou závislé na druhu objektů, které k těmto odhadům slouží. Pokud se odhady dělají za pomoci objektů známé velikosti, budou výsledky podobnější odhadům v reálném prostoru. Matsuda a Takezawa [6] to zjistili experimentem, kdy pozorovatel měl odhadovat egocentrické vzdálenosti stojící dospělé mužské postavy na fotografii. Když byl místo člověka do fotografické scény jako target umístěn arteficiální disk s neznámou fyzickou velikostí, zvýšila se nejistota odhadů pozorovatelů.

## 7 Závěr

Vnímání prostorových charakteristik fotografických scén je závislé na fyzikálních vlastnostech kamery (ohnisková vzdálenost objektivu, výška záběru). Na druhou stranu do podoby vjemu vstupují také vlastnosti sledovaného obrazu (velikost, vzdálenost od pozorovatele). Všechny tyto parametry ovlivňují konečnou podobu vjemu společně s dalšími faktory. Například řada experimentů dokládá, že viditelné orámování zobrazené scény a další kontext potlačují její hloubkový účinek [2].

Celý proces je ovlivněn zkušeností pozorovatele, který se pokouší interpretovat scénu pomocí svých zvykových předpokladů. Má například tendenci interpretovat scénu tak, jako by ji viděl vlastníma očima. Kromě toho existují doklady, že odhadování prostorových vlastností scény závisí na objektech ve scéně, skrze které se prostorové odhady dělají.

Velká nejistota se týká zejména egocentrických odhadů vzdáleností objektů ve scéně, kdy je obtížné pro pozorovatele určit absolutní velikost těchto vzdáleností. Proto někteří autoři místo egocentrických distancí pracují při hodnocení percipované hloubky scény se vzájemnými vzdálenostmi a orientacemi více objektů ve scéně [10].

## Poděkování

Tento příspěvek vznikl za podpory Grantové agentury České republiky v rámci grantového projektu „Prostorové vnímání fotografických scén“ 13-28709S.

## Literatura

- [1] Baird J. C. & Biersdorf W. R. (1967). Quantitative functions for size and distance judgments. *Perception & Psychophysics* 2, 161-166.
- [2] Deregowski J. B. & Parker D. M. (1996). The depiction of distance: A Bartellian analysis. *Perception*, 25, 177-185.
- [3] Gilinsky, A. S. (1951). Perceived size and distance in visual space. *Psychological Review* 58, 460-482.
- [4] Kato, K. & Higashiyama, A. (1998). Estimation of height persons in pictures. *Perception & Psychophysics*, 60 (80), 1318-1328.
- [5] Matsuda, T. (2002). Depth perception of persons in pictures photographed with various focal lengths. *Ritsumeikan Journal of Human Sciences*, 2, 47-54.
- [6] Ohnaka Y. & Matsuda T. (2008). Effects of Aspect-ratio and Size of Photographs upon the Depth Impression and the Depth Perception on their Scenes. *Ritsumeikan Journal of Human Sciences*, 17. 15-24.
- [7] Šimeček M., Šikl R., Lukavský J. (2011). Gulliver na Marsu. Sborník z konference Kognice a umělý život XI. FPF Slezská univerzita, 269 – 273. ISBN 978-80-7248-644-1.
- [8] Smith, O. W. (1958). Comparison of apparent depth in a photograph viewed from two distances. *Perceptual and Motor Skills*, 8, 78-81.
- [9] Smith, O. W. & Gruber, H. (1958). Perception of depth in photographs. *Perceptual and Motor Skills*, 8, 307-313.
- [10] Wagemans J., van Doorn A. J., Koenderink J. J. (2011). Measuring 3D point configurations in pictorial space. *I-Perception* 2, 77-111.
- [11] Wagner, M. (2006). *The Geometries of Visual Space*. Mahwah, NJ: Lawrence Erlbaum Associates.
- [12] Watanabe T. (2006). Geometrical Structures of Photographic and Stereoscopic Spaces. *The Spanish Journal of Psychology* 9/2, 263-272.
- [13] Yang T. & Kubovy M. (1999). Weakening the robustness of perspective: Evidence for a modified theory of compensation in picture perception. *Perception & Psychophysics*, 61 (3), 456-467.



## Socio-technická interakce: Virtuální prostředí jako důsledek informatizace společnosti

Zdeněk Smutný

Fakulta informatiky a statistiky, VŠE v Praze  
nám. W. Churchilla 4, 130 67 Praha 3  
zdenek.smutny@vse.cz

### Abstrakt

Příspěvek předkládá nový pohled na důsledky masivního procesu informatizace se současným akcentem na internetové služby. Téma této socio-technické interakce bude představeno v perspektivě sociální informatiky s cílem prezentovat její rozšířené paradigma pro výzkum ve virtuálním prostředí. Pod pojmem virtuální prostředí chápeme jednotlivé služby (např. Google Search, Facebook nebo Twitter) poskytované přes internet nejen lidem, ale také softwarovým robotům, tzv. softbotům. Tyto služby si lze představit jako možné světy s vlastními pravidly a možnostmi, přičemž zde dochází k interakcím nejen v rámci jedné služby, ale také napříč službami. Můžeme tedy mluvit také o virtualizaci (společnosti), ovšem v jiném pohledu, než jaký nám nabízejí autoři jako P. LÉVY nebo C. W. HORROCKS v souvislosti s virtuální realitou na počátku 21. století. V příspěvku se zaměříme také na kognitivní procesy člověka a především softbota, které probíhají v nově formovaném virtuálním prostoru. Budeme argumentovat směrem k virtuálnímu prostředí jako platformě, která je jednotná pro člověka a softboty (stroje), kdy toto prostředí je jednodušší/přístupnější pro softboty na jedné straně a na druhé straně dochází k ořezání (kognitivních) možností člověka v tomto prostředí. V příspěvku budou dále diskutovány související implikace.

### 1 Úvod do problematiky

Jaké konsekvence přináší současná technika společnosti, respektive jaké má dopady na socio-ekonomické prostředí? – to je otázka, na niž musejí být v průběhu času neustále pozměňovány a upřesňovány odpovědi. Abychom byly schopni ptát se na důsledky, je vhodné nejdříve porozumět pozadí tvorby techniky, neboli tomu, kde se tato mimopřirozená činnost člověka bere. Při odpovědích na tuto otázku nám mohou pomoci známá díla autorů J. O. GASSETA,<sup>1</sup> M. HEIDEGGERA<sup>2</sup> či

L. MUMFORDA.<sup>3</sup> V jejich pracích nalezneme podstatu a souvislosti vzniku této činnosti, které vrhají člověka vpřed, aby překonával okolostojičnost, dnes stejně jako před sto lety.

Vedle filosofů, historiků a sociologů nalezneme i další, především obchodníky, kteří se ptají na důsledky techniky v socio-ekonomickém prostředí ve specifických oblastech tak, aby své závěry přetavili kupříkladu v zisk nebo společensky prospěšný výsledek. V této souvislosti se nejčastěji v poslední několika dekadách opakují otázky po důsledcích pronikání informačních a komunikačních technologií (ICT) do všech oblastí fungování společnosti (informatizace). Bývá také zmiňována tzv. socio-technická interakce, která dostává novou kvalitativní úroveň, pokud je tím míněna především interakce člověka (na úrovni personální či organizační) s ICT.<sup>4</sup>

V kontextu v úvodu položené otázky s akcentem na problematiku informatizace je ze všech aplikovaných oblastí nejbliže – dnes již velmi široká oblast – sociální informatika. V dalším textu ji představíme podrobněji, přičemž naším neskromným cílem bude rozšířit její paradigma ve světle současného „pokroku“.

V tomto příspěvku se zaměříme na změnu prostředí, ve kterém člověk dnes žije a které jej zásadním způsobem určuje. Nakonec, determinace prostředím se odráží jak u živých organismů, tak u strojů člověkem vytvořených, a to ať už se jedná o fyzický nebo virtuální stroj.<sup>5</sup> Daný subjekt/objekt, který existuje v určitém prostředí, jej musí vnímat a dále být schopen v něm interagovat, aby přežival. Jednoduše lze říci, že každý takový subjekt/objekt (člověk živý organismus stejně jako stroj)

<sup>2</sup> M. Heidegger: *Věda, technika a zamýšlení*. Oikoymenh, Praha, 2004.

<sup>3</sup> L. Mumford: *Technics and civilization*. The University of Chicago Press, London, 2010.

<sup>4</sup> Oborů, které se touto problematikou zabývají z různých pohledů, nalezneme hned několik, například computer-aided technologies, ambient intelligence, ubiquitous computing nebo v příspěvku rozebíraná sociální informatika.

<sup>5</sup> Například soft robot, respektive softbot, viz [5, str. 38]

<sup>1</sup> J. O. Gasset: *Úvaha o technice*. Oikoymenh, Praha, 2011.

vnímá jiné prostředí, respektive různě velkou část světa, jež je schopen zpracovat a ve kterém je schopen existence.<sup>6</sup> Z tohoto důvodu je vzhledem k naší úvodní otázce zásadní nový fenomén spojený s informatizací, jenž vytváří kvalitativně nové virtuální prostředí, přinášející nové implikace jednak pro člověka na úrovni personální i organizační, jednak strojů, a dopomáhá k prolínání jejich prostředí.

## 2 Sociální informatika

Současná a velmi mladá koncepce z konce 90. let 20. století zabývající se socio-technickou interakcí je sociální informatika v podání R. KLINGA. Ten ji definoval jako „interdisciplinární studium designu (návrhu tvorby), užití a důsledků informačních a komunikačních technologií, které zohledňují jejich interakci s institucionálním a kulturním kontextem.“ [7] Jak se však tento obor začal ve světě prosazovat, jeho kontext se postupem času rozšiřoval, aby obsáhl do té doby izolované přístupy podobného zaměření nebo názvu.<sup>7</sup>

Na evropské půdě najdeme první zmínky o podobných oborech ze začátku 80. let 20. století na University of Oslo, kde se problematikou komunikace (včetně člověka a počítače) zabýval tým kolem sociologa a později také kognitivního psychologa S. L. BRÁTENA. Právě v tomto období se setkává R. KLING, pozdější zakladatel sociální informatiky v USA, s BRÁTEMEM, který ho informuje o oblasti „socialinformatikk“ (původně nazývané jako socio-informatics dle [13]). Sociální informatika byla později uznána roku 1984 jako nový obor norskou vládou [1].

Ačkoli se takto koncipovaná sociální informatika z pohledu sociologie a komunikace uchytila, rozvíjela (např. díky I. ROGGENOVI) a vyučovala se na několika norských univerzitách, tak se pod tlakem nových fenoménů transformuje například na „samfunnsinformatikk“ (Community Informatics<sup>8</sup>) na Nesna University College a omezuje se na jednotlivé předměty. V Dánsku roku 1985 vzniká na University of Aalborg program „Humanistisk Informatik“ (Humanistic

<sup>6</sup> Jinak vyjádřeno: subjekt/objekt vnímá prostředí, jež postačuje pro jeho existenci. Neboli, nepotřebuje vnímat celý svět ve své plné složitosti, neboť nutnost interpretace takového objemu dat na jakékoli úrovni je nadbytečná vzhledem k jeho existenci.

<sup>7</sup> Jako například na Ljubljanské univerzitě, kde se takto nazvaný obor (družboslovna informatika) vyučuje od roku 1985.

<sup>8</sup> Do češtiny hůře přeložitelné jako „komunitní informatika“, která se zaměřuje na problematiku využívání informačních a komunikačních technologií různými komunitami. Zahrnuje do sebe zejména sociální informatiku (v pojetí R. KLINGA), a dále počítačovou vědu, informační vědu a knihovnictví a další oblasti spadající do sociálních věd. Na této univerzitě jsou součástí studia digitální kultury.

informatics) právě pod vlivem „norské školy“, kde prošla také dlouhým vývojem, přičemž se dnes zaměřuje na mediální komunikaci, informační systémy a informační vědu, včetně důrazu na problematiku využívání informačních technologií.<sup>9</sup> „Humanistická i komunitní informatika“ našla odezvu i ve zbytku světa a dále se rozvíjí na dalších univerzitách. Jinak orientovaná sociální informatika směrem k využití informatiky jako nástroje pro sociologický výzkum se dodnes rozvíjí na Ljubljanské univerzitě. Vedle zmíněného západoevropského přístupu existuje také jinak koncipovaný ruský přístup nazývaný sociální informatika a zaměřený především na udržitelný rozvoj společnosti v podmínkách informatizace, jehož tradice sahá do dob Sovětského svazu a jehož současnými představiteli jsou A. D. URSUL a K. K. KOLIN.

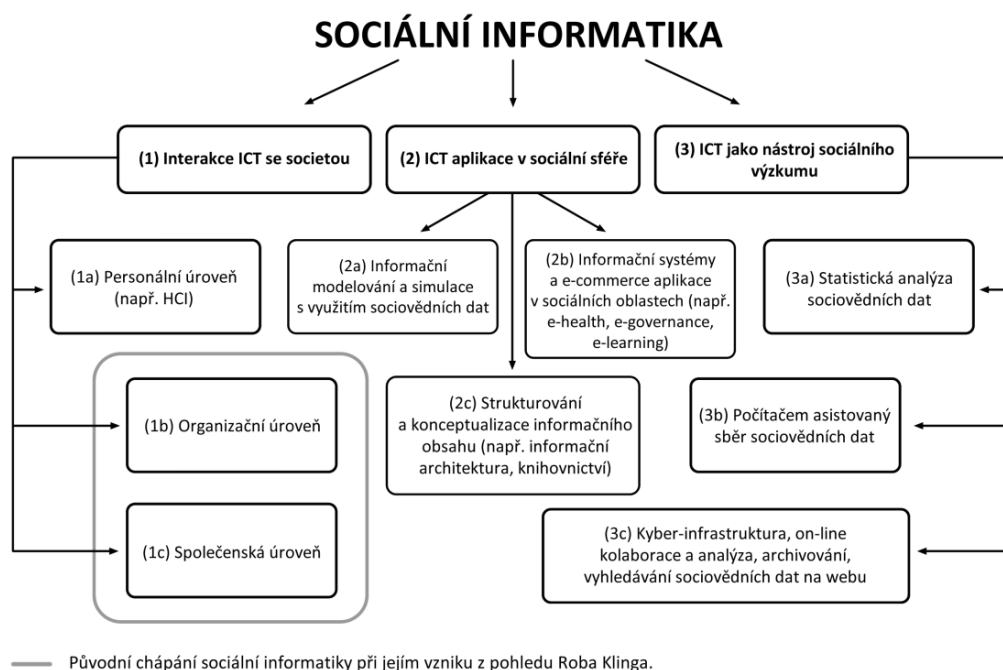
Na Obr. 1 je znázorněn aktuální pohled na sociální informatiku vystavěnou na základech položených R. KLINGEM a rozšířenou především o evropské koncepce. V následujících podsekcích představíme jádro myšlenek, které stály za vznikem tohoto oboru v KLINGOVĚ pojetí a zároveň představíme současné problémy s touto koncepcí spojené. Přičemž budeme argumentovat směrem ke změně pohledu tak, aby byl tento obor jako celek schopen dalšího kontinuálního rozvoje. Tato změna pohledu vychází z aktuálního technologického vývoje a změn, které do života člověka přináší ICT.

### 2.1 Výzkum socio-technické interakce

Jak jsme poukázali v předchozí části, problematika interakce mezi člověkem (společností) a technikou zprvu inspirovala především sociology k návrhu nových oblastí výzkumu, respektive tvorbě nových oborů. V 80. letech 20. století sociologové vymýšlejí také teorie socio-technické interakce na obecné úrovni (nejenom ICT). Z tohoto období upozorníme na dvě: SCOT (Social Construction of Technology) a ANT (Actor Network Theory). SCOT vychází ze sociálního konstruktivismu a jeho autory jsou W. BIKER a T. PINCH [2]. Základy ANT položili M. CALLON, B. LATOUR a J. LAW, přičemž autoři obou přístupů se znali a obě teorie vznikaly na evropské půdě. Těmito koncepty se o přibližně deset let později inspiroval také R. KLING, ovšem není s nimi plně spokojen ve světle bouřlivého vývoje ICT v 90. letech 20. století a pokládá vlastní fundamenty sociální informatiky. Jako stěžejní část představuje vlastní přístup k výzkumu socio-technické interakce – metodiku/model STIN (Socio-Technical Interaction Networks).

<sup>9</sup> Více informací: <http://www.huminf.aau.dk>





Obr. 1. Pojetí sociální informatiky, tak jak je chápáno a prezentováno v současné době (2013) na Ljubljanské univerzitě. Převzato z [15].

Ve svém pojetí se snaží vyvážit jak hledisko sociální, tak technické, zaměřuje se především na informatizaci a mění se také úhel pohledu, kdy on sám je profesorem počítačové vědy, informačních systémů a informační vědy, kdežto doteď v této oblasti dominovaly přístupy sociologů. KLINGOVY pohnutky k vytvoření nového přístupu můžeme vysledovat v bouřlivém vývoji oblastí souvisejících s počítači, zejména umělé inteligence, informačního managementu či podnikových informačních systémů v 70. a 80. letech 20. století v USA. KLING se snažil již v této době záhy po získání doktorátu kritizovat „pozitivistický“ přístup k informačním a komunikačním technologiím. Tuto dobu v USA můžeme ilustrovat optimismem M. MINSKEHO, který roku 1970 prorokoval, že během tří až osmi let bude umělá inteligence na úrovni lidské, a bude tedy schopna například číst Shakespeara nebo říkat vtipy. Obdobný pohled nabízela i aplikace počítačových technologií v praxi, kde vládl „naivní determinismus“ ve smyslu: čím více počítačů, tím více se ve firmě zefektivní práce a ušetří finanční prostředky. [6] Přístup KLINGA k tomuto technokratickému pohledu byl velmi kritický, neboť také u empirického pozorování v praxi se ukazovala řada problémů (především problém lidského faktoru), které s sebou přinášely často přesný opak. Tento paradox stál ve vší jednoduchosti také u zrodu sociální informatiky na akademické půdě, která měla nabídnout východisko a zlepšit či zjednodušit pohled na složité interakce, ke

kterým dochází mezi lidmi a technikou (ICT) na úrovni určité organizace, a její společenské dopady.

Klíčovou myšlenkou sociální informatiky v podání R. KLINGA je, že ICT neexistuje v izolaci, a to ať už sociální nebo technologické. Na jedné straně člověk či společnost vytváří dané technologie pro svou potřebu a na druhé straně je potřeba vhodně tyto nové technologie zasadit do sociálního rámce organizace či společnosti. Takto pojatá sociální informatika spadá pod aplikovanou informatiku (jako např. softwarové inženýrství), neboť se zabývá spíše praktickým řešením problémů než tvorbou explanačních teorií. [8, str. 15]

Cílem je tvorba empirické základny, na které sociální informatika staví svůj výzkum a nabízí hodnotný vhled do současných problémů a otázek v souvislosti s informatizací společnosti. Výsledkem je i tvorba dvou modelů, respektive metodik, pro odhalení „konceptního rámce“ světa – STIN [10] a teorie uživatelů jako sociálních aktérů [9].

## 2.2 Strasti a propasti sociální informatiky

Takto načrtnutá oblast s vlastními metodami a nástroji výzkumu měla své stoupence i odpůrce. Odpůrcům vadila určitá nejednoznačnost a vágnost v rétorice KLINGA, která se zrcadlila i do STINu, kdy postup tvorby modelu závisel do značné míry na zkušenosti tvůrce a jeho

interpretačních schopnostech. Přesto tato oblast zaznamenala rozkvět v letech 1999-2007. Zásadním problémem byla také Klingova předčasná smrt v roce 2003, kdy tyto přístupy ztratily svého nejdůležitějšího obhájce.

Navíc mnozí z pokračovatelů R. Klinga, držící se myšlenky predikce dopadu IT, expandovali do nových oblastí směrem k Big Data nebo sociální dynamice na internetu. Důvodem byl neuvěřitelný rozmach a konsekvence internetu, jehož dopad R. Kling do své smrti silně podceňoval [6]. Lze říci, že právě podcenění fenoménu internetu mělo za následek postupný úpadek zavedených přístupů a přinutilo jeho následovníky k průzkumu nových metod pro výzkum dopadu internetových technologií na společnost. Tímto novým směrem se vydali i jeho žáci a pokračovatelé jeho myšlenek (S. SAWYER, H. ROSENBAUM, E. T. MEYER). V současné době nám však chybí holistický pohled, jednotící linie (přístup), neboť výzkum se značně tříští do izolovaných oblastí, jakoby se sociální informatika rozpadla na jinak samostatně žijící oblasti.<sup>10</sup> Sociální informatika tak získává spíše nádech transdisciplinární platformy [14] s tím, že jednotlivé části spolu obtížně spolupracují, neboť každá vidí jinou část světa pro své zkoumání. Postrádáme tedy vymezení základních principů současné interakce, použitelné pro tvorbu informačních systémů v podmínkách virtuálního prostoru a pro jejich následný výzkum a hodnocení.

V druhé části této práce se budeme zabývat východiskem pro sociální informatiku, které spatřujeme v novém pohledu na problematiku důsledků, jež přineslo masivní rozšíření přístupu k internetovým službám po roce 2000, a následných implikací, které směřují do života člověka.

### 3 Internet a virtuální prostředí

Nejdříve vysvětleme výchozí pojmy (internet, kyberprostor a virtuální prostředí), abychom je dostali do vzájemných vztahů vzhledem k této práci. Internet je počítačová síť, která sama o sobě nic nevytváří, je třeba software a komunikace v rámci sítě. Teprve tak může vzniknout například prostředí webů a jejich interakce s uživateli. Právě tato nejdostupnější oblast webů bývá ztotožňována s pojmem internet. Dalším užívaným pojmem, který byl přejet ze sci-fi literatury, je kyberprostor neboli doslova „řízený prostor“, jež byl vytvořen člověkem. Principiálně opravdu veškerý software a hardware je výtvořem člověka a člověk se ho snaží užít a řídit ke svému prospěchu. V našem pojetí však používáme raději pojem virtuální svět, jenž je tvořen

<sup>10</sup> Tyto trendy dokumentují témata mezinárodní konference o sociální informatice (SocInfo), viz [www.socinfo2012.com](http://www.socinfo2012.com)

dílčími virtuálními prostředími, což jsou jednotlivé internetové služby (Facebook, Google Search aj.) mající své vlastní možnosti.<sup>11</sup> V tomto pojetí se odráží chápání výrazu virtuální jako možný, navazuje na myšlenku možných světů, kdy v tomto prostředí můžeme vytvořit jakoukoli službu ~ ohraničené prostředí. Jsme limitováni pouze možnostmi hardwaru, softwaru a lidskou představitostí. Vedle takto kontinuálně budovaného virtuálního světa existuje původní fyzický svět, bez něhož by neexistoval ten virtuální.

Fyzický svět postihuje oblast naší přímé percepce, kterou nabízí naše tělo. Je to prostředí, ve kterém přirozeně interagujeme s okolím včetně sociální interakce mezi lidmi. Primárně byla fyzická těla určena právě ke komunikaci skrze fyzický svět, přičemž každé tělo má své omezení, ať už se jedná o receptory nebo možnosti zpracování velkého množství dat. S rozvojem centrální nervové soustavy a vyšší mozkové činnosti u člověka se objevují další dva paralelní světy – svět snů a svět představ, jež můžeme pokládat za předchůdce virtuálního světa. Ze světa snů a představ se stává fundamentální pilíř našeho současného abstraktního myšlení a tvořivosti, která umožnila rozvoj společnosti.

Z hlediska fylogeneze mají sny úkol tříditi a zpracovávat percepci vytvořené přes den; tato schopnost dává určitému druhu konkurenční výhodu. Z ontologického pohledu snění zajišťuje syntézu percepce s již obsaženými znalostmi v endoceptu, což je model světa v mysli jednotlivce, zlepšuje adekvátnost reakcí v prostředí a vytváří zkušenosti jednotlivce. „Lidské myšlení pak při dalších projekcích také propojuje dřívější zážitky s aktuálními, aby zjistilo, zda již něco podobného nebylo zakoušeno, zda nový stimul už něčemu neodpovídá, případně aby tomu také přizpůsobilo reakci na aktuální stimul.“ [12, str. 108] Při tomto procesu dochází také k abstraktním a tvořivým procesům, kdy jsou předkládány i scénáře a kombinace, které by v reálném případě nenastaly. Nové koncepty se tak mohou vynořit při interakci existujících myšlenkových konceptů. [3, str. 192] Od snů je tak jen kousek k vědomým představám, kdy nám mysl představuje různé kombinace řešení problému, jež vyvěrá z nevědomí – například ve formě tzv. aha efektu ~ uvědomění si řešení problému.

<sup>11</sup> Stejně jako fyzický svět má svá pravidla (fyzikální zákony), která ohraničují a určují jednotlivá prostředí v něm obsažená, tak obdobně aktuální možnosti hardwaru a softwaru ohraničují virtuální svět. Přičemž v každém světě mohou existovat dílčí prostředí či systémy (černá díra, planeta / Facebook, Twitter) s jinými možnostmi/funkcemi. Výhodou virtuálního prostoru je, že zde nejsou pevné zákony, které by člověk nemohl změnit (škálování hardwaru, změna protokolu, architektury atd.). Snad až na ty, které opět klade fyzické prostředí na hardware.

## 3.2 Společné prostředí pro člověka a stroje

Je nasnadě, že právě tato schopnost člověka pracovat s abstraktními představami<sup>12</sup> stála u zrodu dnešního virtuálního světa. Dělení na svět fyzický a virtuální je totiž společné jak člověku, tak pokročilým strojům, které člověk vytvořil, právě díky abstraktním konstruktům, se kterými je naše mysl schopna pracovat a se kterými pracují i stroje. Lze si představit jednoduchou analogii: člověk si představí jablko – vznikne tak v mysli určitý konstrukt, který charakterizuje (má referenci na) jablko; obdobně si stroj vytvoří objekt jablko při běhu své instance a může s ním dále pracovat (za předpokladu, že oba již jablko znají). Je třeba zdůraznit, že se jedná o analogický proces, nikoli však stejný. Najdeme zde i známý analogický dualismus u člověka a stroje: mysl a tělo vs. software a hardware. Člověk díky své invenci dokáže přetvářet na softwarové úrovni své představy (odrazit je ve virtualitě), a co je nejdůležitější, pomocí ICT je aktivně sdílet s dalšími lidmi. Vznikla tak nová kvalita života člověka na základě tvorby nového světa (prostředí). A právě úvahy o tomto člověkem vytvořeném světě, který sjednocuje lidi a stroje, chceme dále rozvíjet. Virtuální svět chápeme jako společnou platformu pro aktivní interakci strojů (softbotů) a člověka. Stejně jako v případě strojů (fyzických i virtuálních) se jedná o prostor, který vytvořil člověk. Jedná se tedy o prostší prostředí než komplexní fyzický svět, člověk tedy může snáze uzpůsobit možnosti a schopnosti stroje právě takto omezenému prostředí než například u fyzického robota, který by se měl pohybovat pouze ve fyzickém světě a vykonávat rozličné činnosti.<sup>13</sup> Zároveň dochází k ořezání (kognitivních) možností člověka v tomto světě, v jednotlivých virtuálních prostředích. Výsledkem je určitá rovnocennost interakcí člověka a stroje ve virtuálním prostředí.

Do teď byly srovnány pouze stroje, působící buď jenom ve fyzickém nebo virtuálním prostředí. Nelze opomenout fyzické stroje,<sup>14</sup> jejichž činnost se odráží jak

ve fyzickém, tak virtuálním prostředí, obdobně jako dnešní člověk, který zrcadlí sebe do internetových služeb nebo je využívá k další interakci. Nedochází však k odrazu skutečného člověka, nýbrž pouze jeho alternativy, kterou sám vytváří. V praxi pak dochází k masivní virtualizaci společnosti.

Při pohledu na virtuální svět a jeho „namapování“ na fyzický můžeme spatřit další analogii, protože to, co se nám zde rýsuje, silně připomíná to, co víme o procesech v mozku (mysleny jsou pouze procesy, nikoli mozek samotný). Existuje zde další dualismus, kdy jinak vypadá fyzická úroveň a něco zcela jiného se děje na virtuální úrovni (a to i z hlediska přenášených dat/informací). Stejně tak i reakce na podnět vypadá obdobně, jako když se podíváme na snímek mozku pořízený pozitronovou emisní tomografií. Při určitém podnětu reagují jen některé části mozku. Nalezneme zde moduly, které reagují různě intenzivně vždy při daném podnětu, a pak zde najdeme jiné části, které se zapojují nepravidelně a které zpravidla při neurovědním výzkumu ignorujeme. Velmi to připomíná reakci ve virtuálním prostředí, kdy určité sdělení (ale i nová služba či produkt) se šíří podle vnějších podnětů.<sup>15</sup> Ve virtuálním světě tedy dochází k paralelnímu souběhu služeb, ze kterých emergují celospolečensky zásadní témata, která ji neustále přetváří (ovlivňují), obdobně jako se myšlenky dostávají do vědomí a ovlivňují aktuální chování člověka ve fyzickém světě.

## 3.1 Virtualita

S pojmem virtuality a virtuální reality je spjata i její dělení z pohledu člověka a současných technologií na tzv. silnou a slabou virtualitu. Silná virtualita je technologicky podmíněna (pro pohyb ve virtuálním světě – helma, speciální rukavice aj.): „Pocit vnoření přichází ze zařízení, která izolují naše smysly natolik, že se člověk cítí přenesen na jiné místo.“ [4, str. 36] Slabá virtualita zasahuje i náš fyzický svět, typickými příklady jsou komunikace přes mobilní telefon, domácí počítač či bankovní terminál. Trendem mělo být postupné přibližování se k silné virtualitě, tak byla načrtnuta budoucnost na konci 90. let 20. století. Ovšem realita současnosti je naprosto jiná – slabá realita nám postačuje. Obdobně si to představoval jak C. W. Horrocks, tak i P. Lévy a předpokládali nutnost vytvoření hardwarových

---

pak může dotazováním na server získat informace o aktuálním zpoždění jeho spoje.

<sup>15</sup> Zpráva se šíří jinak rychle na Twitteru a jinak na Facebooku. V některých státech je odezva a sdílení silnější, jinde slabší. Přes některá sociální média se daná zpráva nemusí šířit vůbec. Zajímavou ukázkou v rámci jedné služby zobrazující množství tweetů o Fukushima v jednotlivých státech vs. šíření radioaktivního mraku naleznete zde: [www.socialintensity.org](http://www.socialintensity.org)

---

<sup>12</sup> Abstraktní konstrukt vytváří mysl člověka na základě smyslových dat, která jsou jí dodávána z vnějšku, přičemž to, co je pro vzniklý obraz skutečně podstatné, jsou apriorní formy mysli, nikoliv vnější data.

<sup>13</sup> V tomto ohledu může přijít námitka, že je obdobně jednoduché vytvořit robota na výrobní lince, která je jeho omezeným prostředím, ve kterém pracuje; dané prostředí je také jednoduché. Nicméně, ve virtuálním prostředí se snáze ztrácí ponětí, zda daná entita pracující v daném prostředí je člověk nebo stroj – argumentace by tedy směřovala k Turingově testu.

<sup>14</sup> Například autobus MHD napojený přes internet na fleet management system, kde se odráží jeho fyzická poloha na mapě a je možné dále upravovat jeho cestu za účelem optimálního řízení vzhledem k aktuální dopravní situaci. Návštěvník města,

prostředků, které by nějakým způsobem více izolovaly naše smysly a vrhaly je do alternativního světa. [11, str. 64] Lévy při svých úvahách předpokládal, že se člověk bez takových pomůcek nedokáže dostatečně oprostit od fyzického světa a plně vnímat ten virtuální. Dle našeho soudu tuto činnost nyní dělá mozek člověka, který je onou interakcí (nikoli jen smysly) připoután k virtuálnímu prostředí, stejně jako při percepci a reakcích ve fyzickém světě. Děje se tak i přesto, že využíváme pouze některé vybrané exteroceptory. Při projekci do alternativního světa náš mozek „dokresluje“ situaci díky tomu, že zpracování percepcí se děje právě v mozku, stejně jako při imaginaci příběhu z knihy, s tím rozdílem, že zde dochází k interakci s virtuálním prostředím. V takovém světě pak sebezhotovuje člověk své alternativy sebe sama. Obdobnou analogii nalezneme také u interakce na sociálních médiích, kdy se nejedná o naplnění (role, obsahu), nýbrž vyplnění, v tomto případě vedoucí k uspokojení sociálních potřeb jednotlivců.

#### 4 Závěr

V příspěvku byly představeny důsledky masivní socio-technické interakce na život člověka a společnosti na úrovni virtuálního prostředí. Implikace, které s sebou přináší prolínání fyzického a virtuálního světa, se promítají do každodenní činnosti člověka a přispívají také k rozvoji nového druhu socio-ekonomických vztahů ve virtuálním prostředí.

V souvislosti se sociální informatikou se jedná o opomíjený sjednocující pohled na prostředí internetových služeb, které v tomto oboru chybí. Sociální informatika nabízí obecně koncipované metody a nástroje pro výzkum socio-technické interakce jako je STIN, které se však zaměřuje pouze na důsledky ve fyzickém světě. Problémem je neschopnost jejího paradigmatu uchopit zásadní fenomény, které přináší masivní dostupnost připojení k internetu – virtuální svět. Zaměřuje se na dílčí výzkumy v oblasti internetu bez širšího kontextu. Pro ten zde nabízíme nový pohled, jenž nabízí platformu pro další systematický výzkum a umožňuje holistický náhled na interakce, které zde probíhají. Toto téma si však zaslouží širší rozpracování.

#### Poděkování

Tento příspěvek vznikl za podpory interní grantové agentury na Fakultě informatiky a statistiky, VŠE v Praze v rámci grantu IG409012.

#### Literatura

- [1] About Social Informatics. *Rob Kling Center for Social Informatics*. 2010 [cit. 2013-04-02]. Dostupné z: <http://rkcsi.indiana.edu/About/SI.shtml>
- [2] W. E. Bijker, T. P. Hughes, T. J. Pinch: *The Social construction of technological systems: new directions in the sociology and history of technology*. The MIT Press, Cambridge, Mass., 1987.
- [3] M. R. Byrne: *The Rational Imagination: How People Create Alternatives to Reality*. The MIT Press, Cambridge, Mass., 2005.
- [4] C. W. Horrocks: *Marshall McLuhan a virtualita*. Triton, Praha, 2002.
- [5] J. Kelemen: *Strojovia a agenty*. Archa, Bratislava, 1994.
- [6] J. L. King, S. Iacono a J. Grudin: Going Critical: Perspective and Proportion in the Epistemology of Rob Kling. *Information Society* 4 (2007).
- [7] R. Kling: What is Social Informatics and Why Does it Matter? *D-Lib Magazine* 1 (1999).
- [8] R. Kling, H. Crawford, H. Rosenbaum, S. Sawyer a S. Weisband: *Learning from Social Informatics: Information and Communication Technologies in Human Contexts*. The Center for Social Informatics, Bloomington, 2000.
- [9] R. Kling, R. Lamb: Reconceptualizing users as social actors in information systems research. *MIS Quarterly* 2 (2003).
- [10] R. Kling, G. McKim, A. King: A Bit More to IT: Scholarly Communication Forums as Socio-Technical Interaction Networks. *Journal of the American Society for Information Science and Technology* (2003).
- [11] P. Lévy: *Kyberkultura*. Karolinum, Praha, 2001.
- [12] K. Pstružina: *Pojednání o lidském myšlení III*. Ekopress, Praha, 2010.
- [13] I. Roggen: Specialization Course in Web Sociology and Social Informatics. *Universitetet i Oslo*. 1998 [cit. 2013-03-28]. Dostupné z: <http://folk.uio.no/iroggen/WEBSociologyINFOeng.html>
- [14] Z. Smutný, V. Řezníček: Příspěvek k současné filosofii a metodologii vědy se zaměřením na inter- a transdisciplinární výzkum. *E-Logos – Electronic Journal for Philosophy* 3 (2012).
- [15] Social Informatics. *FDVinfo.net*. 2006 [cit. 2013-03-29]. Dostupné z: <http://www.social-informatics.org/uploadi/editor/1158776249plakat%20ENG1.pdf>

## Konekcionistický model epizodickej pracovnej pamäti

Martin Takáč

Centrum kognitívnych vied KAI FMFI UK  
Mlynská dolina, 842 48 Bratislava  
Email: takac@ii.fmph.uniba.sk

Alistair Knott

Dept. of Computer Science, University of Otago  
PO Box 56, Dunedin 9054, New Zealand  
Email: alik@cs.otago.ac.nz

### Abstrakt

V príspevku predstavíme nový výpočtový model uchovávaní epizód v pracovnej pamäti implementujúci Baddeleyho predstavu epizodického zásobníka a konzistentný s myšlienkou, že epizódy sa uchovávajú ako pripravené sekvencie pozornostných a motorických operácií (Knott). Model pozostáva z viacerých prepojených samoorganizujúcich sa máp (SOM). Technickou inováciou modelu je, že umožňuje transformácie medzi časovou a priestorovou reprezentáciou sekvencií a obsahuje médium, kde môžu byť jednotlivé sekvencie reprezentované viaceré naraz, takže môžu medzi sebou súťažiť, resp. je možné reprezentovať ich rozdelenie pravdepodobnosti. To sa využíva pri predikcii pokračovania epizódy (tendencia zhora nadol pri pozorovaní epizódy) alebo na rozhodovanie o ďalšom kroku pri aktívnom vykonávaní epizódy. Model taktiež reprodukuje niektoré experimentálne zistenia o reprezentácii pripravených sekvencií v prefrontálnej kôre.

### 1 Úvod

Klasický model pracovnej pamäti [4] pozostáva z fonologickej slučky, vizuálno-priestorového náčrtníka a centrálnej riadiacej zložky. Alan Baddeley [3] neskôr rozšíril tento model o médium na uchovávanie sémantických reprezentácií epizód resp. udalostí – tzv. epizodický zásobník. V tomto článku prezentujeme výpočtový model epizodického zásobníka.

V našom prístupe predpokladáme, že epizódy sú prezentované v pracovnej pamäti ako pripravené senzomotorické rutiny resp. plány, t.j. ako sekvencie pozornostných a motorických operácií. Motivácií pre takúto sekvenčnú organizáciu je viacero. Podľa súčasných simulačných teórií porozumenia jazyka (napr. [7]) reprezentácia epizódy by mala odrážať proces, ktorým skúsenosť epizódy vznikla, a umožniť jeho neskoršie znovuprehratie resp. simuláciu. Senzomotorické procesy prítomné pri prežívaní epizódy majú pri vhodne zvolenej časovej škále sekvenčnú štruktúru. Podľa [5] senzomotorické spracovanie prebieha pomocou sekvenčne štruktúrovaných *deiktických rutín*, ktorých atomickými elementami sú senzorké alebo moto-

rické akcie, tzv. *deiktické operácie*. Podrobnejšia argumentácia o sekvenčnej štruktúre senzomotorických procesov participujúcich na prežívaní konkrétnych epizód je uvedená v [9]. Ďalší argument sa týka mechanizmu, pomocou ktorého sú epizódy ukladané v dlhodobej pamäti. Na tomto procese sa zúčastňuje hipokampus. Podľa [11] úloha hipokampu spočíva v aktívnom znovuprehrávaní sekvencií reprezentácií evokovaných počas pôvodnej senzomotorickej skúsenosti. Ešte ďalší argument pochádza z produkcie reči: predpokladá sa, že počas generovania vety je význam, ktorý má byť vyjadrený, uložený/aktivovaný v pracovnej pamäti. V [9] zdôvodňujeme sekvenčnú reprezentáciu vetných významov a v [14] uvádzame výpočtový model generovania viet založený na takejto reprezentácii.

V nasledujúcej časti uvádzame požiadavky, ktoré by dizajn nášho modelu mal spĺňať, potom nasleduje popis architektúry modelu a simulačných experimentov zameraných na okamžité vybavovanie z pamäti, predikciu pokračovania sekvencie a riadenie rozhodovania.

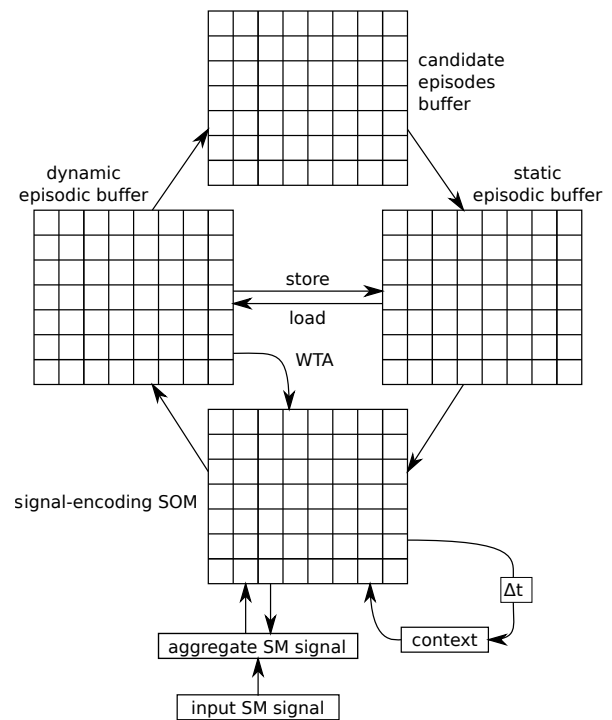
### 2 Požiadavky na model

V prvom rade, model by mal byť schopný registrovať, ukladať a neskôr prehrať v senzomotorických oblastiach sekvenciu signálov v čase. Ďalej by sa mal byť schopný naučiť často sa vyskytujúce sekvencie a vedieť predikovať pokračovanie sekvencie. Ten istý model epizodickej pracovnej pamäti by mal fungovať pre prežívanie epizódy z perspektívy jej aktívneho vykonávateľa, ako aj pasívneho pozorovateľa. V prvom prípade agent aktívne riadi vykonávanie epizódy, preto potrebuje predikovať, ktoré akcie sú výhodné, t.j. ktoré akcie participujú v zapamätaných epizódach asociovaných s kladnou odmenou. V druhom prípade anticipuje, ako sa bude pozorovaná epizóda ďalej vyvíjať. V oboch prípadoch si potrebuje vybrať z konkurenčných alternatív. Preto potrebujeme dva druhy konkurencie: jednu na úrovni viacerých alternatív jednotlivých akcií v každom kroku časovej sekvencie, druhú na úrovni celých zapamätaných sekvencií/epizód. To však predstavuje technický problém: potrebujeme médium, v ktorom by mohlo byť reprezentovaných viacero časových sekvencií naraz, s rôznou úrovňou

aktivity a tak umožniť ich vzájomnú konkurenciu (rôzne úrovne aktivity môžu byť interpretované ako rozdelenie pravdepodobnosti možných víťazov). Tento problém sme vyriešili transformáciou časových sekvencií na lokalistické/bodové reprezentácie.

Kľúčovým rozhodnutím v návrhu modelu bolo použitie samoorganizujúcich sa máp (SOM) [10]. SOM je (obvykle dvojrozmerná) mapa neurónov s topografickým usporiadaním, ktoré sú plne prepojené s vrstvou vstupných neurónov. Po prezentácii tréningového vzoru na vstupnej vrstve každý neurón mapy reaguje priamo úmerne vzdialenosti (napr. euklidovskej) medzi vstupným vektorom a vektorom svojich vlastných váh. Váhy neurónu s najsilnejšou odpoveďou a jeho susedov na mape sa následne adaptívne posunú smerom k vstupnému vektoru. SOM majú niekoľko veľmi dobrých vlastností, vďaka čomu sú vhodné na modelovanie pamäťových funkcií [12]:

1. Sú kombináciou distribuovanej a lokalistickej reprezentácie: distribuovaná informácia je uchovávaná vo váhach spojení medzi vstupnou vrstvou a mapou, lokalistická informácia je reprezentovaná pozíciou neurónu v mape.
2. Kvôli veľkému množstvu parametrov je reprezentácia v SOM robustná voči neúplnému a zašumenému vstupu. SOM funguje ako (odšumovací) autoasociátor: po rozpoznaní zašumeného vstupu je možné zrekonštruovať jeho zapamätanú podobu z váh víťazného (najsilnejšie odpovedajúceho) neurónu.
3. Mapa obsahuje aj tzv. neobsadené neuróny, ktoré nereprezentujú žiaden jeden konkrétny vstup, ale skôr kombináciu viacerých, čo umožňuje modelovať zmiešanie spomienok.
4. V mape môže byť aktívnych viacerých neurónov súčasne, čo umožňuje reprezentovať viaceré konkurujúce si alternatívy a ich rozdelenie pravdepodobnosti (vyjadrené aktivitou neurónov reprezentujúcich jednotlivé alternatívy).
5. Mapa má vlastnosť „magnifikácie“, t.j. najčastejšie vstupy sú reprezentované väčšími oblasťami mapy a preto sú medzi nimi možné jemnejšie rozlíšenia ako medzi zriedkavými vstupmi.
6. Mapa zachováva vzťahy podobnosti: podobné vstupy sú reprezentované neurónmi, ktoré sú blízko seba.
7. Tréningový proces – samoorganizácia je biologicky plauzibilná a nepotrebuje učiteľa.



Obr. 1: Architektúra modelu

### 3 Architektúra modelu

Náš model pozostáva z viacerých navzájom prepojených samoorganizujúcich sa máp. Najprv popíšeme ich funkciu, potom technické detaily.

Architektúra modelu je na obr. 1. Sieť dostáva na vstupe sekvenciu senzomotorických (SM) signálov postupne tak, že v každom okamihu je vo vstupnej SM oblasti aktívny signál zodpovedajúci jednej sensorickej alebo motorickej operácii (*input SM signal*). Vstupné SM signály reprezentujú alebo agentove vlastné akcie (pozornostné alebo motorické), alebo vonkajšie podnety (objekty, alebo pozorované akcie). Vstupné SM signály prechádzajú cez oblasť agregovaného SM signálu (*aggregate SM signal*), kde sa kombinujú s očakávaniami/predikciami šírenými zhora z vyšších oblastí, do oblasti kódovania signálu (*signal-encoding SOM*). Táto oblasť je implementovaná architektúrou typu MSOM [13], čo je variant SOM s rekurentnými spojeniami, ktoré reprezentujú časový kontext vstupu: táto sieť má ako ďalší vstup tzv. kontextový vektor, ktorý je kombináciou váh a kontextového vektora víťazného neurónu z minulého časového kroku. Oblasť kódovania signálu je médium, kde nastáva konkurencia medzi kandidátmi na reprezentáciu fragmentov sekvencie: táto sieť sa tréningom samoorganizuje tak, že jej jednotlivé neuróny kódujú vstupné signály v rôznych sekvenciálnych kontextoch,

podobne, ako sa to deje v prefrontálnej kôre [6].

Vítazný neurón sa v každom kroku skopíruje do oblasti izomorfnnej s oblasťou kódovania signálu – tzv. dynamického epizodického zásobníka (*dynamic episodic buffer*), kde sa temporálna sekvencia transformuje na distribuovaný priestorový vzor: v každom kroku sekvencie víťaz z oblasti kódovania signálu zanechá stopu v dynamickom epizodickom zásobníku tým, že aktivuje príslušný izomorfný neurón. Prvý signál v sekvencii má najvyššiu aktivitu, ostatné sú postupne zapamätané s klesajúcou aktiváciou, takže na konci sekvencie dokážeme podľa veľkosti aktivít zrekonštruovať poradie, v akom boli neuróny aktivované. Ako ukázal Averbeck v štúdií makakov [1], podobne sa to deje aj v prefrontálnej kôre. Vstupnú sekvenciu zakódovanú v dynamickom epizodickom zásobníku možno prehrať, teda spätne transformovať na časovú sekvenciu mechanizmom podobným na kompetitívne zaraďovanie (*competitive queuing*, napr. [8]): najaktívnejší neurón v dynamickom epizodickom zásobníku vyšle signál do oblasti kódovania signálu (cez spojenie označené na obr. 1 ako WTA), potom svoju aktivitu inhibuje a celý proces sa iteratívne opakuje, pokiaľ v dynamickom epizodickom zásobníku existujú nadprahovo aktívne neuróny. Signál z dynamického epizodického zásobníka aktivuje príslušný neurón v oblasti kódovania signálu, ktorý môže premietnuť svoj váhový vektor do oblasti agregovaného SM signálu.

Aby bolo možné takúto sekvenciu vykonať opakovane, uchováva sa v tzv. statickom epizodickom zásobníku (*static episodic buffer*), ktorý má rovnakú štruktúru ako dynamický, avšak neaktualizuje sa deštruktívne počas prehrávania sekvencie. Priestorový vzor uchovávaný v statickom epizodickom zásobníku sa skopíruje do dynamického, odkiaľ sa potom vyššie opísaným mechanizmom prehráva.

Teraz zameriame pozornosť na médium, v ktorom môžu byť súčasne aktívne reprezentácie viacerých celých sekvencií. Statický ani dynamický epizodický zásobník nemajú túto vlastnosť, pretože distribuované priestorové vzory by mohli pri súčasnej aktivácii navzájom interferovať. V tzv. epizodickom zásobníku kandidátov (*candidate episodes buffer*) je to možné, pretože priestorové vzory z dynamického zásobníka sa tu transformujú na lokalistické reprezentácie. Epizodický zásobník kandidátov je implementovaný ako klasická SOM, ktorej neuróny sú plne prepojené na vstupy prichádzajúce z dynamického epizodického zásobníka. Počas tréningu sa táto SOM naučí reprezentovať epizódy s podobným priestorovým vzorom blízko seba. Tu využívame robustnosť SOM voči neúplnému vstupu: v každom okamihu počas prezentácie sekvencie (kým sa prvok po prvku postupne vytvára distribuovaná priestorová reprezentácia v dynamickom epizodickom zásobníku) vyjadruje profil aktivít v epizodickom zásobníku kandidátov rozdelenie pravdepodobnosti úplných epizód zapamätaných vo

váhach jeho neurónov. Toto rozdelenie možno interpretovať ako predikciu možného pokračovania momentálne pozorovanej epizódy, alebo ako tendencia (bias), ktorá ovplyvňuje výber ďalšej operácie v prospech tých, ktoré boli v minulosti odmenené (pozri časť 4.4).

Po určení víťazného neurónu v epizodickom zásobníku kandidátov sa aktivita spätne šíri do nižších vrstiev procesom rekonštrukcie zhora nadol. Počas rekonštrukcie sa váhy víťazného neurónu z epizodického zásobníka kandidátov skopírujú do statického a následne aj dynamického epizodického zásobníka. Deštruktívna iteratívna aktualizácia v dynamickom epizodickom zásobníku spôsobí sekvenciu aktivácií neurónov v oblasti kódovania signálu: projekcia váh týchto neurónov nadol do oblasti agregovaného SM signálu vytvorí očakávaná zhora nadol, ktoré možno skombinovať so vstupným signálom prichádzajúcim zdola nahor.

Teraz predstavíme niektoré technické detaily navrhnuť architektúry. Oblasť vstupného SM signálu používa lokalistické (one hot) kódovanie, t.j. každá možná SM operácia je reprezentovaná iným neurónom.<sup>1</sup> Oblasť agregovaného SM signálu je izomorfná s oblasťou vstupného SM signálu. Oblasť kódovania signálu je dvojrozmerná MSOM [13] so 400 neurónmi ( $\alpha = 0,4$ ,  $\beta = 0,5$ , rýchlosť učenia 0,1 a gausovské okolie znižujúce sa lineárne z 10 na 0,5). Pôvodný algoritmus MSOM sme zmodifikovali tak, že pri hľadaní víťaza v oblasti kódovania signálu sú zo súťaže vylúčené neuróny, ktoré už boli víťazmi v predchádzajúcich krokoch tejto sekvencie. Aj keby časová sekvencia prehrávaná v oblasti vstupného SM signálu obsahovala opakujúce sa prvky, ich lokalistická reprezentácia v oblasti kódovania signálu bude rôzna. To je nutné preto, že víťazné neuróny postupne vytvárajú priestorovú stopu v dynamickom epizodickom zásobníku: aktivity rovnakých víťazných neurónov by sa naskladali na seba a informácia o opakovaní by sa stratila.

Statický aj dynamický epizodický zásobník sú izomorfné s oblasťou kódovania signálu, teda každý z nich má 400 neurónov (akurát bez váhových spojení). Časová postupnosť víťazov v oblasti kódovania signálu sa zaznamenáva v dynamickom epizodickom zásobníku ako priestorová stopa izomorfnie prislúchajúcich neurónov s exponenciálne klesajúcou aktivitou:  $n$ -tý neurón v postupnosti má aktivitu  $\delta^{n-1}$ , kde  $\delta = 0,8$  a všetky nepoužité neuróny majú nulovú aktivitu. 400-rozmerný vektor reprezentujúci úplný priestorový vzor v dynamickom epizodickom zásobníku sa pošle ako tréningový vstup do epizodického zásobníka kandidátov až po skončení celej epizódy. Epizodický zásobník kandidátov je klasická SOM s 900 neurónmi, konštantnou rýchlosťou učenia 0,9 a Gausovským okolím znižujúcim

<sup>1</sup> V experimente opísanom ďalej používame 35 SM operácií a teda 35 vstupných neurónov.

sa lineárne z 10 na 0,5.

## 4 Experimenty a výsledky

Experimenty sme zamerali na overenie, či má model želanú funkcionálnosť, konkrétne či umožňuje okamžité vybavenie práve prežitej epizódy, predikciu možného pokračovania pozorovanej sekvencie, a riadenie rozhodovania počas aktívneho vykonávania epizódy. Ďalej sme skúmali podobu vytvorených reprezentácií a kapacitu epizodického zásobníka kandidátov.

### 4.1 Trénovanie

Model sme trénovali na sekvenciách predstavujúcich päť rôznych typov senzomotorických rutín, pomocou ktorých sú epizódy prežívané. Konkrétna sekvenčná štruktúra jednotlivých typov rutín má kognitívnu motiváciu a je podrobne zdôvodnená v [9]. Sú to rutiny pre netranzitívne akcie (napr. MUŽ KÝCHNUŤ<sup>2</sup>), tranzitívne akcie (napr. MUŽ POHÁR UCHOPÍŤ), lokomočné akcie s lokatívnym cieľom (napr. MUŽ KRÁČAŤ DOM DO), jednoduché kauzatívne akcie (napr. MUŽ POHÁR SPÔSOBIŤ ROZBIŤ) a kauzatívne akcie, ktorých výsledkom je pohyb (napr. PES KOSŤ SPÔSOBIL SKOTÚLAŤ STÔL POD). Všetky SM sekvencie boli vystavané z 35 SM operácií.

Každú simuláciu sme opakovali 10 krát s rôznymi náhodne inicializovanými hodnotami váh v sieti a rôznymi tréningovými množinami stochasticky vygenerovanými tou istou sadou prepisovacích pravidiel. Výsledky 10 behov boli spriemerované. Každá tréningová množina pozostávala z 500 sekvencií, z ktorých v priemere 13 malo dĺžku dva, 86 dĺžku tri, 126 dĺžku štyri a 274 dĺžku šesť. Sekvencie mohli obsahovať viac rovnakých prvkov – 19,1% sekvencií obsahovalo dva rovnaké signály a 0,9% obsahovalo tri. Trénovanie trvalo 200 epoch – v každej epoche boli tréningové sekvencie prezentované sieti v náhodnom poradí a kontext MSOM bol po každej sekvencii resetovaný.

### 4.2 Okamžité vybavenie z pamäti

Najprv sme testovali schopnosť modelu uložiť a okamžite prehrať SM sekvencie. Táto schopnosť spočíva v interakcii medzi oblasťou kódovania signálu a dynamickým epizodickým zásobníkom. Časová postupnosť SM operácií vo vstupnej SM oblasti vytvorí priestorový vzor v dynamickom epizodickom zásobníku, potom sa zresetuje kontext MSOM v oblasti kódovania signálu a víťazný (najaktívnejší) neurón v dynamickom epizodickom zásobníku iteratívne aktivuje zodpovedajúci neurón v oblasti kódovania

<sup>2</sup>Zápis rutín pozostáva z označení SM operácií resp. „významov“, nie slov nejakého jazyka.

signálu a následne sa inhibuje. Váhy aktivovaného neurónu sa skopírujú do oblasti agregovaného SM signálu a výsledná časová sekvencia v oblasti agregovaného signálu sa porovná s pôvodnou vstupnou. Natrénovanej sieti sme prezentovali 200 testovacích sekvencií: 100 náhodne vybraných z tréningovej množiny a 100 nových, ktoré sa v tréningovej množine nevyskytovali. Model bol schopný korektné prehrať 99,4% (SD=0,49%) sekvencií z tréningovej množiny a 98,6% (SD=0,92%) nových sekvencií.

### 4.3 Predikcia pokračovania sekvencií

Ďalej sme testovali schopnosť modelu generovať očakávania/predikcie o sekvenciách pomocou aktivít v epizodickom zásobníku kandidátov. Predikcia je vlastne vyvolanie najpodobnejšej minulej epizódy zapamätanej vo váhach neurónov tejto SOM. Váhy víťazného neurónu sa skopírujú do statického epizodického zásobníka a pomocou dynamického epizodického zásobníka sa prehrávajú ako sekvencia v oblasti kódovania signálu, odkiaľ generujú očakávania zhora nadol pre SM operácie. Natrénovanej sieti sme prezentovali 500 sekvencií z tréningovej množiny a počas prehrávania sme po každom prvku skúmali očakávania generované sieťou. Doplňenie epizódy na základe jej fragmentu je nejednoznačná úloha: ten istý fragment môže mať viacero pokračovaní, ktoré sieť videla počas tréningovania. Preto sme vyhodnocovali iba to, nakoľko je sekvencia vyvolaná z pamäti kompatibilná s aktuálnym fragmentom vstupnej sekvencie. Napríklad s fragmentom PES MAČKA . . . je kompatibilná epizóda PES MAČKA NAHÁŇAŤ, ale nie epizóda PES MYŠ NAHÁŇAŤ. Z 2936 predikcií (jedna po každom fragmente) bolo kompatibilných 97,1% (SD=1,4%). Aby sme overili schopnosť siete robiť spätnú rekonštrukciu, vyhodnotili sme aj kompatibilitu pre 500 predikcií z úplných epizód. Na týchto sekvenciách sme dosiahli 95,2% (SD=2,7%) korektnú rekonštrukciu.

### 4.4 Riadenie rozhodovania

Taktiež sme skúmali, ako je možné použiť náš model pre riadenie rozhodovania počas aktívne generovaného správania. V tomto prípade sekvencia SM signálov vo vstupnej SM oblasti zodpovedá vykonávanej (nie iba pozorovanej) behaviorálnej sekvencii. Top-down očakávania generované v oblasti agregovaného SM signálu možno interpretovať ako ovplyvňovanie rozhodovania o ďalšom kroku. Počas vykonávania sekvencie dynamický epizodický zásobník obsahuje čiastočný, resp. postupne vytváraný priestorový vzor, ktorý sa porovnáva so zapamätanými sekvenciami v epizodickom zásobníku kandidátov rovnako ako pri predikcii pokračovania pozorovanej epizódy. Rozdiel je v tom, že teraz je súťaženie zapamätaných sekvencií v epizodickom



**Tabuľka 1:** Postupné dopĺňanie epizódy MAČKA SEDIETĚ MUŽ ZA z fragmentov, ovplyvnené zapamätanými odmenami ( $\rho = 0,5$ ). Epizóda vyvolaná v kroku 1 je zmiešanina, ktorá je výsledkom projekcie váh neobsadeného neurónu v epizodickom zásobníku kandidátov. Ostatné vyvolané epizódy sú korektné. Vo všetkých krokoch okrem posledného môžeme pozorovať dominanciu správania obsahujúcich interakcie s kreslom. V poslednom kroku bola aktuálna situácia silnejšia ako odmenou ovplyvnené spomienky a bola zrekonštruovaná korektné. Epizódy vyvolané v krokoch 3 a 4 nie sú kompatibilné s aktuálne vnímaným fragmentom. Bodka (.) označuje signál konca epizódy.

Por. č.	Vstupný fragment	Vyvolaná epizóda
1	MAČKA . . .	MAČKA DRŽAĚ . KRESLO KRESLO
2	MAČKA SEDIETĚ . . .	MAČKA SEDIETĚ KRESLO BLÍZKO .
3	MAČKA SEDIETĚ MUŽ . . .	MAČKA SEDIETĚ KRESLO BLÍZKO .
4	MAČKA SEDIETĚ MUŽ ZA . . .	MAČKA KRESLO SPŔSOBIĚ ÍŠĚ LOPTA ZA .
5	MAČKA SEDIETĚ MUŽ ZA .	MAČKA SEDIETĚ MUŽ ZA .

zásobníku kandidátov ovplyvnené odmenami/úspešnosťou asociovanou s minulými epizódami tak, že v minulosti úspešné epizódy majú vyššiu pravdepodobnosť stať sa víťazom, čo dosiahneme modifikáciou epizodického zásobníka kandidátov nasledujúcim spôsobom.

**Trénovanie.** Každému neurónu v mape epizodického zásobníka kandidátov pridáme ďalšiu (skalárnu) váhu, tzv. *odmenovú váhu*, ktorá slúži na zapamätanie odmeny asociovej s epizódou (pôvodné váhy, ktoré slúžia na zapamätanie epizódy, budeme nazývať *epizodické váhy*). Zopakujme, že epizodický zásobník kandidátov je trénovaný iba kompletnými priestorovými vzormi z dynamického epizodického zásobníka, teda po ukončení epizódy. V prípade, že po ukončení epizódy dostal agent nejakú odmenu, táto slúži ako učiaci signál pre odmenovú váhu. Najprv sa štandardným spôsobom nájde neurón, ktorý má epizodické váhy najbližšie k vstupnému vektoru. Potom sa štandardným spôsobom aktualizujú jeho epizodické váhy a epizodické váhy jeho susedov, a tiež sa aktualizuje jeho odmenová váha pomocou odmenového signálu (ak je k dispozícii) tak, aby odmenová váha tvorila kľavý priemer všetkých minulých odmien asociovaných s týmto neurónom.

**Výber víťazného kandidáta.** Aktivity neurónov pre výber kandidáta sa počítajú ako

$$a_i = (1 - \rho) \exp(-d_i) + \rho r_i, \quad (1)$$

kde  $a_i$  je aktivita  $i$ -teho neurónu,  $d_i$  je euklidovská vzdialenosť medzi vektorom epizodických váh  $i$ -teho neurónu a vstupným vektorom,  $r_i$  je odmenová váha  $i$ -teho neurónu, a  $\rho$  je zmiešavací koeficient odmeny. Experimentovali sme s hodnotami  $\rho = 0,25$  a  $\rho = 0,5$ . Ak nastavíme  $\rho = 0$ , modifikovaný model sa zredukuje

na pôvodný, takže jeho dva operačné režimy (pozorovanie, riadenie správania) môžeme chápať ako prepínanie medzi nulovou a nenulovou hodnotou  $\rho$ .

Riadenie správania sme testovali tak, že sme nechali zbehnúť znovu pôvodné experimenty s tým, že sme poskytli ako trénovací signál spolu s epizódami aj odmenu: všetky sekvencie/epizódy, ktoré obsahovali SM signál KRESLO boli asociované s odmenou 1 a všetky ostatné epizódy s odmenou 0 (žiadna odmena). To zodpovedá hypotetickému scenáriu, v ktorom sú všetky interakcie s kreslami pre agenta nejakým spôsobom výhodné. Počas testovania sme zrekonštruovali sekvencie z epizodického zásobníka kandidátov rovnakým spôsobom ako pri predikcii pokračovania epizód. Potom sme porovnali počty zrekonštruovaných epizód obsahujúcich SM signál KRESLO v pôvodných a súčasných experimentoch. V prípade  $\rho = 0,25$  počet stúpol v priemere o 14,2% (SD=6,3%) a v prípade  $\rho = 0,5$  dokonca o 61% (SD=5,0%), čo sa dá interpretovať ako jasná preferencia správania obsahujúcich interakcie s kreslami (pozri Tab. 1). Ďalej sme vyhodnotili kompatibilitu epizód vyvolaných zo zásobníka kandidátov s dosiaľ vygenerovaným fragmentom epizódy. Kompatibilita bola 94,5% (SD=1,2%) pre  $\rho = 0,25$ , ale klesla na 63,7% (SD=3,2%) pre  $\rho = 0,5$ . Tieto výsledky napovedajú, že hodnota  $\rho$  by mala vyjadrovať rozumný kompromis medzi atraktivnosťou zapamätaných epizód a ich podobnosťou aktuálne vnímanej/vykonávanej epizóde.

#### 4.5 Súvislosť s neurálnymi dátami

V tejto časti popíšeme ako vlastnosti reprezentácií, ktoré vytvára náš model pri trénovaní, korešpondujú s reprezentáciami identifikovanými v prefrontálnej kôre (PFC) makakov.

Niektoré PFC neuróny kódujú jednotlivé operácie v plánovanej sekvencii berúc do úvahy sekvenčný kon-

		MUŽ3			
MUŽ1					
	MUŽ5				
MUŽ4		MUŽ2	MUŽ6		

**Obr. 2:** Pozície víťazných neurónov v mape natrénovanej oblasti kódovania signálu pre výskyty SM signálu MUŽ v rôznych sekvenčných kontextoch: MUŽ1 LOPTA SPÔSOBIŤ SCHOVAŤ MUŽ2 POD, MUŽ1 PES OBJAŤ, PES MUŽ3 SPÔSOBIŤ ÍŠŤ MUŽ4 POD, MUŽ1 SEDIET MUŽ5 BLÍZKO, PES PES SPÔSOBIŤ SCHOVAŤ MUŽ6 BLÍZKO. Na obrázku je zobrazená iba (ľavá dolná) časť oblasti kódovania signálu s 20x20 neurónmi.

text, v ktorom sa vyskytujú [6]. V našom modeli majú túto vlastnosť neuróny v oblasti kódovania signálu. Obr. 2 znázorňuje pozíciu víťazných neurónov v mape natrénovanej oblasti kódovania signálu pre výskyty SM signálu MUŽ v rôznych sekvenčných kontextoch.

Iné neuróny v PFC kódujú jednotlivé operácie v plánovanej sekvencii tak, že relatívne úrovne aktivácie určujú poradie vykonania operácií [1]. Aktivita niektorých z týchto neurónov sa dynamicky mení počas vykonávania sekvencie tak, že je maximálna pred vykonaním príslušnej operácie a inhibovaná po nej (v našom modeli im zodpovedajú neuróny v dynamickom epizodickom zásobníku). Iné neuróny majú aktivitu invariantnú počas vykonávania celej sekvencie (v našom modeli im zodpovedajú neuróny v statickom epizodickom zásobníku).

Niektoré oblasti PFC sú médiom, kde rôzne konkurenčné alternatívy sekvencií medzi sebou súťažajú ([2]). V našom modeli je takýmto médiom epizodický zásobník kandidátov.

#### 4.6 Kapacita epizodického zásobníka kandidátov

Schopnosť správnej spätnej rekonštrukcie závisí na dostatočnej kapacite epizodického zásobníka kandidátov (Tab. 2). Pri príliš malej kapacite sú viaceré epizódy (najmä tie, ktoré sú si podobné) reprezentované vo váhach toho istého neurónu, a preto pri rekonštrukcii dochádza k zmiešaniu spomienok, ktoré sa vyskytuje aj v ľudskej pamäti.

## 5 Záver

V tomto článku sme navrhli konekcionistickú architektúru schopnú uchovávať a prehrávať časové sekvencie. Zamerali sme sa najmä na zaujímavé technické vlastnosti ako sú transformácie medzi časovými a priestorovými reprezentáciami a schopnosť reprezentovať viacero súťažiacich

**Tabuľka 2:** Závislosť obsadenosti pamäti a rekonštrukčnej schopnosti od veľkosti epizodického zásobníka kandidátov. Čísla v zátvorkách vyjadrujú smerodajnú odchýlku (SD). Položka „Zmiešané neuróny“ označuje podiel neurónov v mape, ktoré reprezentujú viac ako jednu epizódu, položka „Rekonštrukcia“ označuje úspešnosť spätnej rekonštrukcie (kompatibilitu predikcie z úplných epizód, pozri časť 4.3).

Veľkosť SOM	Zmiešané neuróny	Rekonštrukcia
10 x 10	83,0% (2,4%)	15,5% (1,5%)
20 x 20	16,6% (1,4%)	74,2% (2,2%)
30 x 30	1,3% (0,5%)	95,2% (2,7%)

alternatív súčasne. Náš model je implementáciou Baddeleyho epizodického zásobníka. Prezentovaná sieť je len prvým krokom k úplnému modelu epizodickej reprezentácie v pracovnej pamäti, avšak už teraz možno vidieť niekoľko jej zaujímavých aplikácií. V [14] sme opísali detailný model generovania viet, ktorý je založený na myšlienke sekvenčnej reprezentácie sémantiky viet, t.j. epizód. Model prezentovaný v tomto článku rozpracúva túto myšlienku v omnoho väčšej úrovni detailu a je ho možné priamo zapracovať do modelu generovania viet. Takisto vidíme zaujímavú oblasť aplikácie prezentovaného modelu v interpretácii viet. Pravdepodobnostné modely interpretácie viet majú na vstupe sekvencie slov a na výstupe produkujú rozdelenie pravdepodobnosti možných významov – sémantických reprezentácií. Neurálna implementácia takéhoto modelu by mohla generovať epizodické reprezentácie priamo v epizodickom zásobníku kandidátov, ktorý explicitne kóduje rozdelenie pravdepodobnosti možných kandidátskych epizód. Porozumenie významu vety by potom zahŕňalo prehranie SM sekvencie reprezentovanej víťazným kandidátom, v súlade s modernými simulačnými teóriami významu.

## PodĎakovanie

Ďakujeme Lubici Beňušovej za podnetné diskusie o architektúre modelu. Tento výskum bol financovaný grantami VEGA 1/0439/11 a KEGA 076UK-4/2013.

## Literatúra

- [1] Averbeck, B., Chafee, M., Crowe, D., Georgopoulos, A.: Parallel processing of serial movements in prefrontal cortex. PNAS 99(20), 13172–13177 (2002)
- [2] Averbeck, B., Sohn, J., Lee, D.: Activity in prefrontal cortex during dynamic selection of action sequences. Nature Neuroscience 9(2), 276–282 (2006)

- [3] Baddeley, A.: The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences* 4(11), 417–423 (2000)
- [4] Baddeley, A., Hitch, G.: Working memory. In: Bower, G. (ed.) *The psychology of Learning and Motivation*, pp. 48–79. Academic Press (1974)
- [5] Ballard, D., Hayhoe, M., Pook, P., Rao, R.: Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences* 20(4), 723–767 (1997)
- [6] Barone, P., Joseph, J.P.: Prefrontal cortex and spatial sequencing in macaque monkey. *Experimental Brain Research* 78, 447–464 (1989)
- [7] Barsalou, L.: Grounded cognition. *Annual Review of Psychology* 59, 617–645 (2008)
- [8] Bullock, D., Rhodes, B.: Competitive queuing for serial planning and performance. In: Arbib, M. (ed.) *Handbook of brain theory and neural networks*, 2ed., pp. 241–244. MIT Press, Cambridge, MA (1996)
- [9] Knott, A.: *Sensorimotor Cognition and Natural Language Syntax*. MIT Press, Cambridge, MA (2012)
- [10] Kohonen, T.: Self-organized formation of topologically correct feature maps. *Biological Cybernetics* 43, 59–69 (1982)
- [11] Lee, A., Wilson, M.: Memory of sequential experience in the hippocampus during slow wave sleep. *Neuron* 36, 1183–1194 (2002)
- [12] Miikkulainen, R.: Trace feature map: A model of episodic associative memory. *Biological Cybernetics* 67, 273–282 (1992)
- [13] Strickert, M., Hammer, B.: Merge SOM for temporal data. *Neurocomputing* 64, 39–71 (2005)
- [14] Takac, M., Benuskova, L., Knott, A.: Mapping sensorimotor sequences to word sequences: A connectionist model of language acquisition and sentence generation. *Cognition* 125, 288–308 (2012)



## Mobilní monitorovací systém pro seniory

Petr Tučník, Petr Šuba

Fakulta informatiky a managementu  
Univerzita Hradec Králové  
Rokitanského 62, 500 03 Hradec Králové  
petr.tucnik@uhk.cz, petr.suba@uhk.cz

### Abstrakt

Príspevok popisuje návrh a aktuálně probíhající technické řešení mobilního snímacího zařízení pro oblast ambient assisted living (AAL) zaměřené na péči o seniory. Cílem je vytvořit cenově dostupné zařízení na bázi smartphonů kombinující přenos signálu z více typů senzorů monitorujících zdravotní stav a další relevantní údaje o uživateli. Tato data jsou bezdrátově přenášena na server, kde jsou vyhodnocena a po zpracování jsou dále přenášena zdravotnickému personálu, a podle potřeby je obsluha upozorňována na případné vzniklé kritické situace. Přenosová cesta signálu je zde klíčovým faktorem při volbě nosné technologie a následného kódování, především pak z hlediska elektromagnetické sluchitelnosti systémů, mezi nimiž má komunikace probíhat. Řešení je určeno pro seniory se zdravotními komplikacemi mírnějšího typu, zejména pak takových, které vyžadují kontinuální monitoring.

### 1 Úvod

V oblasti inteligentních prostředí je aktuálně stále větší pozornost věnována oblasti péče o nemocné osoby či osoby vyššího věku. Význam takto zaměřených technických řešení se zvyšuje – krom jiných vlivů – přirozeně také spolu s tím, jak dochází ke stárnutí populace, což je trend prakticky ve všech ekonomicky vyspělých zemích. Dostupnost zařízení, které dokáží tento typ funkcionality poskytovat, je vyšší a situace se spolu s pokračujícím vývojem komunikační a výpočetní techniky postupně zlepšuje, což dovoluje implementovat hardwarově náročnější a komplexnější systémy.

Tato práce tematicky spadá do oblasti tzv. ambientní inteligence (dále AmI) a inteligentních prostředí. Tato prostředí si obecně kladou za cíl nerušivým způsobem napomáhat uživateli či obyvateli v realizaci jeho plánů a v dosahování jeho cílů. Klíčovými atributy takových prostředí jsou diskretnost, důvěrnost a ochrana dat a předvídatelnost chování. Pokud se ovšem jedná o projekty domácí péče - a obecně o systémy určené pro péči o seniory – zvyšuje se důraz na efektivní asistenci,

kvalitní monitoring a vyhodnocování (rozpoznávání) situací, které mohou potenciálně představovat zdravotní či bezpečnostní riziko. Alternativní technická řešení zaměřená na zvýšení kvality života starších osob lze nalézt v [2] nebo [8]. Tímto směrem je zaměřen i tento příspěvek.

### 2 Předpoklady využívání systému

Inteligentní prostředí mají celou řadu možných využití. Pravděpodobně komerčně nejzajímavější jsou využití pro zábavu a volnočasové aktivity, úspornější chod objektu (tzv. inteligentní domy) nebo bezpečnost. Daly by se jistě vyjmenovat i další možnosti využití, platí však, že společným atributem tohoto pojetí AmI je to, že v těchto případech je očekávána nejvyšší míra diskretnosti, („nerušivosti“) systému, a za ideální lze považovat situaci, kdy uživatel o chodu systému téměř neví, viz také práce Wicherta [11].

Tento pohled se ovšem liší v kontextu zdravotní péče a asistence, což jsou oblasti, které v bydlení seniorů mají primární důležitost. Požadavky na takové systémy uvádí např. Larizza a kol. [6]. V tomto kontextu lze také zmínit např. španělský projekt AmIVital, který se zaměřuje na komplexní zlepšení služeb a kvalitu života seniorů formou služeb zvaných e-Health, e-Information, e-Learning, e-Leisure, e-Assistance [10], jedná se tedy o poměrně komplexní řešení, oproti kterému je navrhovaný monitorovací systém jednodušší. Diskretnost systému při využití v oblasti zdravotní péče a asistence ztrácí na významu a důležitější roli zde hraje kvalita asistenčních služeb a přesnost případných diagnostických postupů, pokud má mít systém i medicínskou aplikaci, viz např. systém pro detekci respiračních obtíží [3], pádů [5], nebo se zaměřením na oblast rehabilitace [7]. Tato oblast už spadá spíše než do klasické ambientní inteligence do oblasti anglicky nazývané „ambient assisted living“ (obvykle označované akronymem AAL) nebo do oblastí s přízvisky „homecare“ (domácí péče) či „assisted healthcare“ (asistovaná zdravotní péče).

Kromě standardních požadavků na systém pro správu inteligentního prostředí, jako je např. zabezpečení dat a dostatečná kvalita jejich přenosu, hraje v případě monitorovacího systému pro seniory velkou roli i otázka soukromí. Výsledky šetření, které uvádí Kanis a kol. [4], poměrně jednoznačně ukazují, že některá technická řešení, jako je pořizování video a audio záznamů při monitorování uživatelů, jsou považovány za příliš velký zásad do soukromí. Lze předpokládat, že by to představovalo zásadní překážku pro masovější využívání systému. Detailněji je problematika soukromí řešena v části 4.2.

Zmíněný problém představuje pouze jeden z mnoha příkladů, které by bylo možno v této souvislosti uvést. Obecně lze říci, že téma soukromí a dostatečného zabezpečení systémů na bázi Aml představuje aktuálně významné a dosud nepřilíší zpracované téma vyžadující další pozornost, viz [12]. Tento pohled sdílí i Rothenpieler a kol. [9], přičemž navrhuje, aby data pořízená při monitorování byla plně pod kontrolou uživatele/obyvatele. Technicky dostupným řešením, které nabízí Rothenpieler, je umístění vypínače, tzv. „privacy switch“, při jehož zapnutí systém přestane pořizovat záznam. To je užitečné především tehdy, když má uživatel hosty nebo provádí činnosti, při nichž vyžaduje soukromí. Toto řešení je technicky snadno proveditelné, levné a přístupné i uživatelům, kteří jsou technicky méně zdatní.

Panuje obecná shoda na tom, že objem dat, která monitorovací systém pořizuje, by měl být co možná minimální. Rothenpieler přímo uvádí [9]: „*Objem dat pořizovaných o uživateli by měl být minimalizován na takovou úroveň, která ještě umožní efektivně plnit monitorovací úlohu... data by měla být udržována anonymní tak dlouho, jak jen je to možné, a měla by obsahovat pouze odkaz na místo pořízení (adresu domácnosti), pro případ, že by takovou informaci potřebovala např. záchranná služba.*“

Zohlednění požadavků na dodržování soukromí je základní předpokladem pro úspěšnou aplikaci systému. Studie provedená Kanisem [4] ukazuje, že uživatelé jsou ochotni sdílet podrobnosti o své každodenní rutině prakticky pouze se svým lékařem. Pregnantně je to vyjádřeno jedním z účastníků tohoto průzkumu [4]: „*Svou sousedku pravidelně zdravím, ale nepřejí si, aby věděla cokoli o mém soukromém životě.*“ Neochota sdílet veřejně svá data je také zohledňována Rothenpielerem [9], který pro sběr dat navrhuje využít servisního portálu, kde jsou shromažďována data soukromého i obecného charakteru, přičemž část je přímo zadávána uživatelem. Příklady těchto dat ukazuje Tab. 1.

Tab. 1. Příklady informací shromažďovaných systémem (volně podle [9]).

Zadáno uživatelem	Snímáno senzory
Jméno	Spotřeba vody
Věk	Spotřeba energie
Adresa	Teplota
Koníčky	Sledování pohybu osob
Přátelé	Vlhkost vzduchu
Aktivity	

Všechny tyto informace by měly být soustředěny do zmiňovaného servisního portálu. Ten by měl být dostatečně zabezpečen, ale způsobem, který nebude vyžadovat přílišnou participaci uživatelů. Předpokládá se jejich malá technická zdatnost. Všechna (bezdrátově) přenášená data musí být zabezpečena šifrováním minimálně na úrovni přenosových protokolů. V případě průniku do zabezpečené sítě mohou být informace zneužity např. za účelem vloupání nebo zveřejnění citlivých osobních údajů. Z důvodu zabezpečení proti zneužití systém může generovat datový tok i v případě, že data nepřenáší, aby nebylo možno jen na základě informace o datovém toku rozhodnout o přítomnosti a/nebo kondici monitorované osoby.

Požadavky na soukromí jsou do jisté míry v protikladu s tím, že určitý objem anonymizovaných dat by mohl být využit pro poskytovatele externích služeb (dovážka jídla a nákupů, zdravotní servis, úklid, apod.) a v podstatě zajímavý pro využití pro komerční účely. To v současné době představuje pro navrhovaný systém stále ještě otevřený problém.

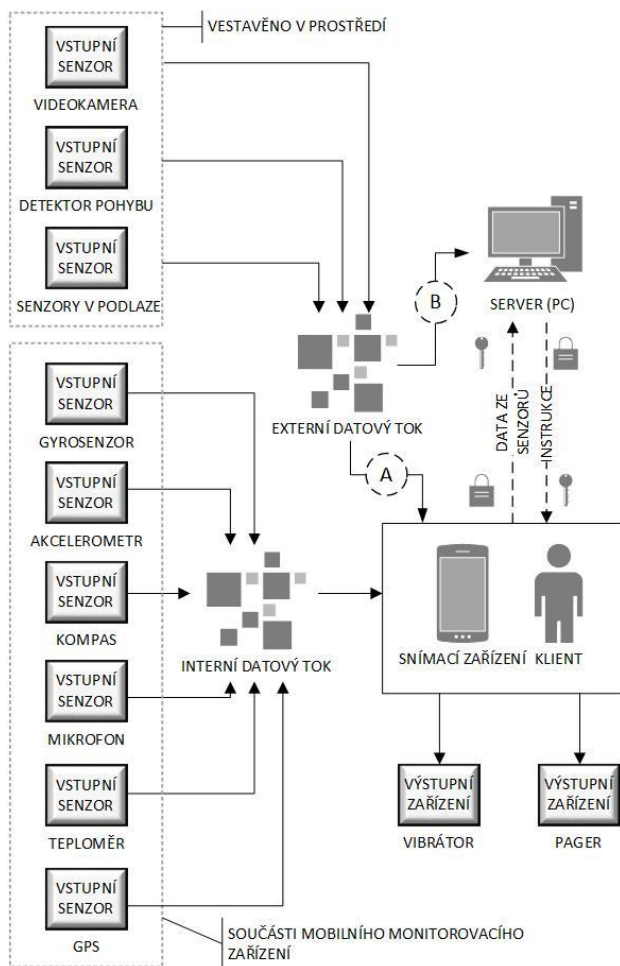
Shrnutí výše uvedených požadavků na monitorovací systém je následující:

- Systém musí v maximální možné míře respektovat soukromí uživatele.
- Data budou snímána jen v takovém rozsahu, který bude dostačující k plnění monitorovací úlohy.
- Uživatel bude mít možnost část monitorovacích prostředků vypnout dle svého uvážení.
- Veškerá přenášená data budou zabezpečena po celé cestě signálu i zpracovaných informací.

### 3 Architektura systému

Inteligentní prostředí, jak bude dále uvažováno, je z perspektivy popisovaného systému tvořeno sítí samostatných čidel a komponent, které jsou schopné autonomní činnosti, tzn., že fungují nezávisle na samotném mobilním měřicím zařízení. Tato zařízení tvoří

část systému, která je chápána jako externí. Jako příklad takové externí soustavy senzorů lze zmínit např. cenově příznivé technické řešení pokrytí většího prostoru čidly, které popisuje Baumgartner [1]. Poměrně dobrý přehled nositelných senzorů (ať již ve formě speciálních oděvů nebo jen dílčích technických komponent) pak nabízí Patel [7].



Obr. 1. Vstupní data pro monitorovací systém.

Interní část mobilního monitorovacího zařízení (dále MMZ) pak tvoří senzory funkčně integrované do mobilního zařízení, které uživatel nosí u sebe. Sem patří především gyrosenzor, akcelerometr, kompas, mikrofon, teploměr a GPS.

Data, která jsou snímána senzory, viz Obr. 1, jsou dále zaslána na server, kde jsou zpracována a vyhodnocena. U externích zdrojů dat, jako jsou videokamery nebo detektory pohybu, může být přenos směřován přímo do serveru k vyhodnocení (tzn., že nemusí nutně jít skrze mobilní monitorovací zařízení). Z technických důvodů

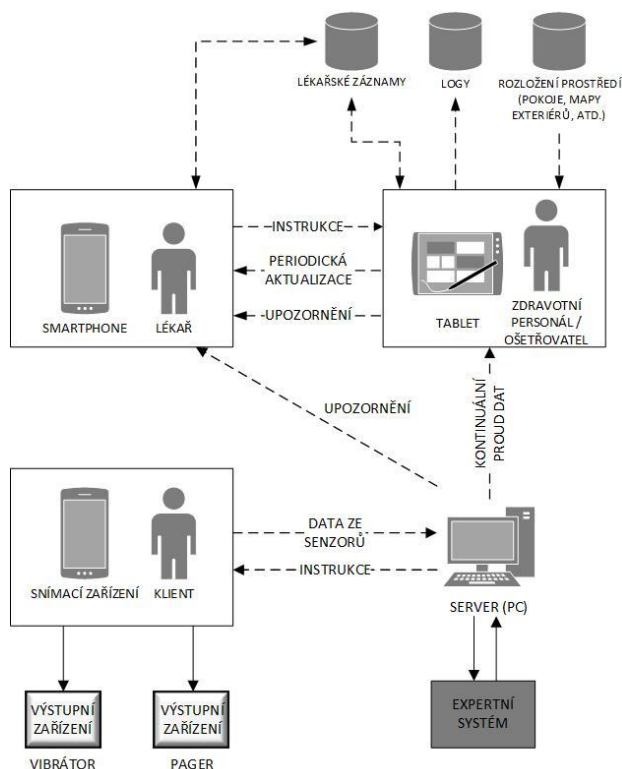
ale může být v některých případech výhodnější využít mobilního zařízení pro zprostředkování přenosu (zejména u bezdrátových přenosů). Proto Obr. 1 uvádí obě varianty, značené jako „A“ a „B“. Přenos je kódován.

Systém je navržen tak, že primárně počítá s vestavěnými senzory, které dovolují realizovat primární požadovanou funkcionalitu, tzn. diagnostiku a monitorování zdravotního stavu. Přítomnost externích čidel je předpokládána a počítá se s jejich využitím, představuje však jen doplňující zdroj informací. Díky modularitě návrhu je možno libovolnou sadu senzorů přidat, či odebrat (případně deaktivovat).

Snímací zařízení disponuje aktuátory (vibrátor, pager, případně i reproduktor, sluchátko, tlakové signalizační zařízení), které mohou uživatele upozornit např. na dávkování léků, zobrazit zprávy od lékaře či ošetřujícího personálu, systémové zprávy apod.

Z hlediska poskytování informací o zdravotním stavu upozorňuje na zajímavý aspekt Rothenpieler [9], a to skutečnost, že uživatel systému má „právo na nevědomost“ (v anglickém originále ozn. jako right of nescience (unknowingness)). Např. v případě, že automatický diagnostický systém diagnostikuje vážný stav nebo onemocnění, měl by o tom být nejprve informován ošetřující lékař. Sdělení takové informace přes rozhraní systému by bylo jednak necitlivé a především je nutné verifikovat rozpoznanou diagnózu. Lékař je také schopen kvalifikovaně vysvětlit důsledky takové diagnózy a ošetřit další případné souvislosti (jako úprava medikace apod.). Tento aspekt navrhovaný systém respektuje, jelikož zdravotní stav vyhodnocuje zdravotní/ošetřující personál, resp. lékař. V případě využití pro účely domácí péče pak tato informace bude dostupná členu domácnosti, který o uživatele pečuje. Je pak na jeho uvážení, jak s ní případně naloží.

Komunikaci v rámci systému zachycuje Obr. 2. Kromě klientského snímacího zařízení, které má uživatel u sebe, a serveru, na něž jsou data posílána, se předpokládá přenos dat zdravotnímu personálu (např. zdravotní sestra vykonávající dohled nebo ošetřovatel) a přes klientské zařízení v odlišné roli (také například smartphone) jsou informace volitelně přístupné i ošetřujícímu lékaři. Lékař dostává periodicky aktualizovaná a sumarizovaná data (push technika), periodu lze upravit dle potřeby. Kontinuální přenos je energeticky náročný a není efektivní s ohledem na předpokládanou skutečnost, že lékař vykonává i jinou činnost. Současně díky využití push techniky je možno vyvolat událost i v zařízení monitorujícího personálu (alarm).



Obr. 2. Komunikace v rámci systému.

Řešení zachycené na Obr. 2 je vhodné pro větší počet klientů, např. domovy důchodců a podobná zařízení. V případě domácí péče postačuje komunikace pouze přes klientské zařízení ošetřovatele (v Obr. 2 znázorněn jako tablet), na které je možno odeslat informaci o nastalé události i periodické sumarizované informace. Komponenta klientské aplikace pro lékaře v tomto případě není nezbytná. Pro diagnostiku základních stavů je na serveru implementován expertní systém, jehož úkolem je rozpoznávat a vyhodnocovat vznikající události (např. dýchací obtíže, srdeční obtíže, pády) a s patřičným zdůvodněním, je-li vyžádáno obsluhou, upozorňovat na tyto situace ošetřující osobu.

Expertní systém má k dispozici informace i o běžných provozních událostech, jako je stav baterií jednotlivých komponent a upozorňuje na jejich stav. Tedy v případě selhání klientského zařízení, nebo v případě ztráty konektivity upozorní technický dohled a aktivuje případnou záložní signálovou cestu, namísto toho aby detekoval kompletní selhání životních funkcí. Současně, pokud není do zvoleného času sjednána náprava, je schopen vyhlásit „všeobecný“ poplach a informovat všechny zúčastněné subjekty.

## 4 Dílčí technická řešení

Jak bylo zmíněno v první kapitole, je třeba se v průběhu technického řešení MMZ vyrovnat s řadou překážek převážně etického, ale i právního charakteru.

### 4.1 Bezpečnostní hlediska

Tak jako u každého systému, v případě předkládaného je třeba vyřešit otázku bezpečnosti. Požadavek na bezpečnost je v tomto případě dán nejen citlivostí získaných, ukládaných i zpracovávaných dat, ale i nutnou spolehlivostí systému. Schopnost detekovat vlastní chyby je nutným a základním požadavkem.

#### 4.1.1 Zabezpečení v rámci sledovaného objektu

Pod pojmem signálová cesta je v kontextu navrhovaného řešení rozuměna celá cesta všech podob získaných dat. Signálová cesta začíná ve snímači, kde je měřená veličina převedena na elektrický signál. Z toho důvodu jsou při výběru komponent preferovány ty snímače, které signál před opuštěním zařízení kódují a šifrují.

Dalším krokem na cestě signálu je průchod prostředím před opuštěním sledovaného prostoru (např. bytem monitorované osoby). I tato část musí být zabezpečena minimálně na fyzické úrovni, ale v případě bezdrátového přenosu je šifrování zcela nezbytné.

Následuje dekodování, předzpracování, filtrování a případně další úpravy signálu před jeho zpracováním a multiplexací. Zařízení se zodpovědností za tento krok musí krom předpokládané fyzické bezpečnosti poskytovat i ochranu před neautorizovaným zásahem. Krom úmyslného zásahu do signálové cesty (zvědavost, úmysl poškodit) je nutno připravit klientské zařízení i na neúmyslné zásahy do systému (děti, neznalost...). Zásahy do klientského zařízení operujícího v signálové cestě proto musí podléhat autorizaci minimálně proti lokální autoritě (před zásahem vyžadováno heslo).

#### 4.1.2 Signálová cesta mezi klientem a serverem

Zde je odpovědnost zcela odkázána na standardní síťové technologie. Z toho důvodu je třeba zajistit šifrování na obou stranách komunikačního kanálu, neboť po opuštění fyzických hranic sledovaného objektu není možno zajistit fyzickou bezpečnost a je třeba počítat s obecnou zranitelností MAN i WAN technologií.

#### 4.1.3 Zabezpečení úložiště



Úložištěm dat je chápána taková část systému, která je na konci šifrovaného komunikačního kanálu a odpovídá za zpracování a vyhodnocení dat i za jejich perzistenci. Z obecného pohledu není podstatné, zda se jedná o distribuovaný či kompaktní systém. Z podstaty musí být zajištěny všechny součásti proti vnitřnímu i vnějšímu útoku na úrovni fyzické i logické. Stav, kdy signál opouští zařízení v nešifrované podobě je nepřijatelný.

#### 4.1.4 Dostupnost informací

Dostupnost informací, tedy přesněji dostupnost jen potřebných informací autorizovanému subjektu je restrikcí vlastností uvedených výše v bodu 4.1. Zabezpečení signálové cesty bude částečně narušeno, pokud bude jakýmkoli způsobem zveřejněna identita sledované osoby a proto je třeba i s touto informací nakládat velmi citlivě. Rovněž úložiště je omezeno ve své bezpečnosti úlohou v případě potřeby část informací zpřístupnit libovolnému subjektu. Z požadavku na dostupnost informací vyplývá závěr, že systém bude jen tak bezpečný, jak bezpečný bude přístup k jeho výstupům. Kritické je tedy zabezpečení cesty směrem k entitám typu *Lékař* a *Ošetřovatel* dle Obr. 2.

Zmíněné entity budou do systému přistupovat po autentizaci a autorizaci. Současně bude využito standardních technologií pro přenos informací. Nelze se spolehnout na zajištění přenosových technologií zvláště z důvodu předpokládané mobility koncových zařízení. Jako vhodné řešení se jeví autentizace prostřednictvím SSL certifikátů a šifrování přenosového kanálu odpovídajícím algoritmem (HTTPS).

#### 4.2 Etická hlediska

Ve shodě s [9] jsou etická hlediska klíčovým atributem celého projektu. Pokud se nepodaří získat důvěru klientů, nebude systém možno nasadit v zamýšlených aplikacích.

##### 4.2.1 Transparentnost

Celý systém musí být postaven a zdokumentován tak, aby bylo možno jej představit i člověku s limitovanými technickými schopnostmi. Součástí tohoto přístupu je kromě dokumentace i komunikace s uživateli pro ně přijatelnou formou (např. tištěné materiály s převahou kvalitních grafických symbolů).

V dalším zkoumání bude řešena otázka vhodnosti zveřejnění zdrojových kódů, i s přihlédnutím k zajištění bezpečnosti systému.

##### 4.2.2 Získávání dat v přijatelné podobě

Podstatným argumentem proti kontinuálnímu poskytování citlivých informací je požadavek na soukromí, alespoň v určitých časových úsecích. Pokud je tento požadavek vznesen z důvodu návštěvy, či soukromého rozhovoru – je zcela na místě a těžko jej lze vyvracet potřebou monitorování stavu, když klient je ve společnosti dalších osob.

Jiná situace ale nastává v případě, kdy je monitoring odmítán z důvodu studu v choulostivých situacích, jakou je například návštěva toalety. V takovém případě dochází ke sporu mezi potřebami klienta a je potřeba je uspokojit obě. Klient se současně dostává do mnohem nebezpečnější situace, než když leží na lůžku a současně je pro něj nepřijatelná asistence dalších osob, či dokonce sledování neznámými osobami.

Zde je spatřován prostor pro technické řešení na bázi pravidlového systému přímo v klientském zařízení a v přenosu jen omezené sady dat, ze kterých není možno sestavit kompletní obraz a zvuk popisující choulostivou situaci.

Vyšší míry soukromí má být dosaženo průchodem nasnímaných dat přes softwarový low-pass filtr přímo ve snímacím zařízení. Klient by s touto vlastností měl být seznámen předem a měl by mít možnost si dopředu vyzkoušet, jaká data jsou monitorujícímu personálu k dispozici. Teprve v případě oboustranné shody (např. klient – lékař) na potřebě odeslat krátkodobě kompletní data bude mít klient možnost filtr deaktivovat a zpřístupnit tak na omezenou dobu dozorujícímu subjektu kompletní data včetně plného zvukového signálu.

#### 4.3 Technika snímání a vyhodnocování dat

Strmost uvažovaného filtru pro zvukový signál by měla být na základě odhadnutých hodnot 20dB na dekádu a prahová frekvence bude stanovena pro jednotlivé veličiny experimentálně. Přesnější vlastnosti low-pass filtru budou konzultovány s lékaři i dobrovolnými klienty a budou porovnány s výstupy medicínálních přístrojů.

Omezení kvality přenášených dat se rovněž pozitivně projeví na množství kontinuálně přenášených a ukládaných dat.

Ospravedlněním pro takový přístup je skutečnost, že frekvence tělesných činností, které je potřeba monitorovat jsou hluboko pod akustickým pásmem. Zde je míněna zejména frekvence dýchání a srdeční činnost, kde jsou obvyklé hodnoty ve zlomcích případně v jednotkách Hz.

Právě z těchto hodnot a z hodnot získaných o změně polohy senzoru na hrudi klienta je možno určit případné známé znaky nežádoucích stavů a predikovat výpadky těchto funkcí.

Porovnávané veličiny jsou do značné míry dány sensorovou výbavou a tedy je možno při zachování modulárního přístupu veličiny přidávat, či odebrat podle konkrétního nasazení systému. Základní varianta by však již měla umožnit detekci pádu, dlouhodobé nehybnosti, potíží s dechem, srdeční činnosti a místa události.

## 5 Závěr

V současné době probíhají měření na zdravých dobrovolnících. Získaná data nejsou zatím nijak upravována ani deformována, takže bude možno na jejich základě stanovit prahové hodnoty (min i max) jednotlivých veličin. Odlišnosti těchto veličin od cílových klientů budou později porovnány a bude možno sestavit algoritmy pro předběžné zpracování dat a vyvolání události. V ideálním případě i pro vyvolání odpovídající reakce na událost.

V nejbližší době se předpokládá postupně dořešení technické konstrukce MMZ, otestování na modelových situacích při monitorování zdravého jedince (figuranta) a teprve po dokončení těchto přípravných fází jeho nasazení pro reálné testování na uživatelích zamýšlené cílové skupiny. Pro tyto účely je již předběžně dohodnuta spolupráce s lékařským zařízením odpovídajícího typu.

## Poděkování

Tento příspěvek vznikl za podpory projektu „Informační, kognitivní a interdisciplinární podpora výzkumu – INKOV, registrační číslo CZ.1.07/2.3.00/20.0001, který je spolufinancován Evropským sociálním fondem a státním rozpočtem České republiky.

## Literatura

- [1] T. Baumgartner et al.: Hallway Monitoring: Distributed Data Processing with Wireless Sensor Networks. In: *REALWSN 2010, LNCS 6511* (P. J. Marron et al., eds.), Springer-Verlag, Berlin Heidelberg, 2010, pp. 94-105
- [2] A. N. Belbachir et al.: Ambient Assisted Living for Ageing Well – an Overview. *Elektrotechnik & Informationstechnik* (2010) 127/7-8: 200-205.
- [3] A. J. Jara et al.: An Ambient Assisted Living Platform to Integrate Biometric Sensors to Detect Respiratory Failures for Patients with Serious Breathing Problems. In: *IWAAL 2011, LNCS 6693* (J. Bravo, R. Hervás, V. Villarreal, eds.), Springer-Verlag, Berlin, 2011, pp. 122-130
- [4] M. Kanis et al.: Ambient Monitoring from an Elderly-Centered Design Perspective: What, Who and How. In: *Aml 2011, LNCS 7040* (Keyson et al. eds.), Springer-Verlag Berlin, 2011: 330-334.
- [5] M. A. Laguna, J. Finat: Remote Monitoring and Fall Detection: Multiplatform Java Based Mobile Applications. In: *IWAAL 2011, LNCS 6693* (J. Bravo, R. Hervás, V. Villarreal, eds.), Springer-Verlag, Berlin Heidelberg, 2011, pp. 1-8.
- [6] M. Larizza et al.: Studies to Determine User Requirements Regarding In-Home Monitoring Systems. In: *UMAP 2012, LNCS 7379* (J. Masthoff et al., eds.), Springer-Verlag, Berlin, 2012, pp. 139-150
- [7] S. Patel et al.: A Review of Wearable Sensors and Systems with Application in Rehabilitation. *Journal of NeuroEngineering and Rehabilitation* (2012) 9:21
- [8] R. Rogers et al.: Living Longer Independently – a Healthcare Interoperability Perspective. *Elektrotechnik & Informationstechnik* (2010) 127/7-8: 206-211
- [9] P. Rothenpieler et al.: Privacy Concerns in a Remote Monitoring and Social Networking Platform for Assisted Living. In: *Privacy and Identity 2010, IFIP AICT 352* (S. Fischer-Hübner et al., eds.), IFIP International Federation for Information Processing, 2011, pp. 219-230
- [10] Z. Valero et al.: AmlVital: Digital Personal Environment for Health and Well-Being. In: *MobiHealth 2010, LNCS 55* (J. Lin, K. S. Nikita, eds.), Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, 2011, pp. 160-167
- [11] R. Wichert: Challenges and Limitations of Intelligent Assisted Living Environments. In: *Aml 2010, LNCS 6439* (B. de Ruyter et al. Eds), Springer-Verlag, Berlin Heidelberg, 2010, pp. 304-309
- [12] Ziefle et al.: When Your Living Space Knows What You Do: Acceptance of Medical Home Monitoring by Different Technologies. In: *USAB 2011, LNCS 7058* (A. Holzinger, K. M. Simonis, eds.) Springer-Verlag, Berlin Heidelberg, 2011, pp. 607-624

## Vliv kontextu na orientaci v prostoru

Michal Vavrečka, Jakub Kužilek, Lenka Lhotská

Fakulta elektrotechnická ČVUT  
Karlovo náměstí 13, Praha 2  
vavrecka@fel.cvut.cz, lhotska@fel.cvut.cz

### Abstrakt

V prezentovaném experimentu jsme se zabývali faktory, které ovlivňují preferovanou navigační strategii u skupiny 32 dobrovolníků. Zkoumali jsme vliv allothetických (environmentálních) a idiothetických (individuálních) faktorů. Allothetická informace byla modifikována pomocí stoupající složitosti průchodu virtuálním tunelem, která způsobila větší chybovost a přesnost navigace u tunelů se větší délkou a počtem zatáček oproti jednodušším tunelům. Zkoumaným idiothetickým faktorem byl způsob prezentace cílů v rámci sensorického horizontu při pohybech tunelem vzad a mimo sensorický horizont účastníka při pohybech tunelem vpřed. Výsledky prokázaly vliv na preferovaný způsob navigace, jelikož poměr allocentrické proti egocentrické navigace se změnil z 65:35 u pohybu vpřed na 3:97 při pohybu vzad.

### 1 Úvod

Přestože se může z laického hlediska jevit orientace a navigace v prostoru jako triviální kognitivní schopnost, při bližším zkoumání zjistíme, že je situovaná v prostředí velmi komplexním fenoménem. Dokazují to stovky výzkumů v oblasti kognitivní psychologie a neuropsychologie, které identifikují faktory, které mají vliv na tuto schopnost. Také v oblasti neurověd existuje velké množství studií, které mapují mozkové oblasti zodpovědné za reprezentaci v prostoru z mnoha různých perspektiv. Ukazuje se, že během orientace v prostoru jsou současně aktivována mozková centra v okcipitálních, parietálních, temporálních i frontálních lobech, přičemž jejich součinnost je pro efektivní pohyb v prostředí nezbytná. Cílem výzkumů je kromě mapování mozkových center také získávání znalostí umožňujících tvorbu umělých kognitivních systémů, které budou vykazovat podobné schopnosti jako lidé. Požadavek situovanosti, jako základní schopnosti virtuálních či robotických agentů se objevuje již v základních učebnicích umělé inteligence (Pfeifer, Scheier, 1999)

Současné výzkumy prostorové navigace v oblasti psychologie a neurověd se nejčastěji soustřeďují

na výzkum orientace pomocí rozdílných referenčních rámců. Jedná se o lidskou schopnost reprezentovat prostor z různých perspektiv a popřípadě popsat vztahy objektů v prostoru. V literatuře zabývající se prostorovou kognicí je často referenční systém reprezentován jako soustava ortogonálních os, jejichž centrem je retina, hlava, tělo nebo další body, objekty či pole v prostoru (Behrmann, 2000; Colby a Goldberg, 1999; McCloskey, 2001).

Studie věnované neurálním korelátům navigace v prostoru poukazují na rozdíl mezi užitím egocentrického (prostorové vztahy jsou vztahovány k pozorovateli) a allocentrického (prostorové vztahy jsou popisovány podle stabilních objektů v prostředí) referenčního rámce (Fink et al., 2003). Byly lokalizovány fronto-parietální oblasti včetně posteriorního parietálního kortexu a premotorické kůry nacházející se v pravé hemisféře, související s egocentrickým referenčním rámcem. V případě allocentrického rámce byla aktivní pouze část těchto center (Galati et al., 2000). Problematickým místem zmíněných studií je způsob, jakým byla ZO prezentována úloha. Jelikož se jednalo o užití statických stimulů, při kterém je způsob prezentace velmi odlišný od normálního vnímání a navigace v prostoru, nastávají problémy s ekologickou validitou. Odlišným způsobem prezentace se zabýval ve svých výzkumech Gramann (2005,2006,2010), který použil jako stimuly průchody virtuálním tunelem. Ty obsahují pouze vizuální informaci (světelný tok), bez přítomnosti objektů, které by mohly změnit způsob orientace a navigace subjektu v prostoru. Subjektu je navíc prezentovaná dynamická scéna, která je shodná se stimuly z prostředí. Výzkumy za použití virtuálního prostředí vedly k identifikaci rozdílů v aktivitě Broadmannových oblastí 7 a 32. V dalších studiích (Gramann et al. 2010, Lin et al. 2009) byl výsledek potvrzen a pomocí přesnějších metod analýzy byl identifikován pokles v EEG alfa aktivitě během otočení jednoho rámce vůči druhému v zahnutých částech virtuálního tunelu.

V těchto pracích se díky použité metodologii podařilo identifikovat také fenomén spontánní dispozice k používání allocentrického či egocentrického referenčního rámce. Jelikož není během navigace v

tunelu participant instruován, aby užíval specifický referenční rámec, je možné měřit převládající (preferovaný) způsob navigace. U většiny účastníků se objevuje setrvalá tendence používat jeden referenční rámec. Analýza výsledky prokázala, že použití specifického referenčního rámce nemá vliv na navigační chybu, měřenou jako rozdíl odhadované a skutečné polohy orientačního cíle (Gramann et al., 2005,2010). Studie administrované v 3D prostoru (Vavrečka, 2009, Vavrečka et al.,2012) však identifikovaly, některé faktory, které dokáží systematicky měnit preferovaný referenční rámec. V případě vertikální navigace dochází u skupiny participantů používajících egocentrický rámec k přepnutí do allocentrického. Příčinou může být konflikt s vnímání vestibulární informace oproti vizuální. Efekt byl potvrzen na vzorku 300 lidí (Gramann et al., 2012).

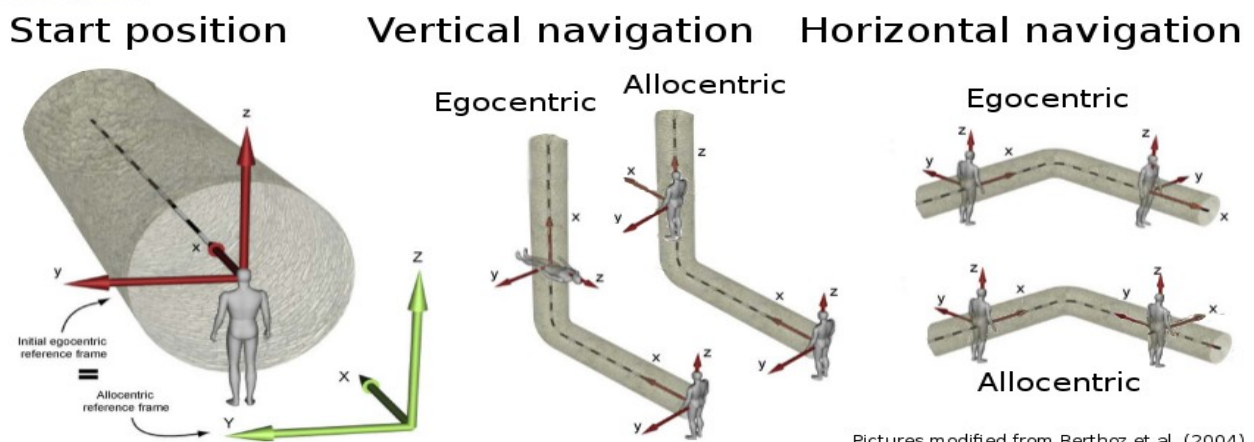
Cílem prezentovaného výzkumu je identifikace dalších faktorů, které by mohly způsobit změnu ve spontánním užívání referenčních rámců. Jelikož byla dosud úloha ve všech studiích využívajících navigaci ve virtuálním tunelu prezentována způsobem, během kterého se odhadovala poloha mimo sensorický horizont účastníka (odhadované místo leželo vždy za zády), rozhodli jsme se modifikovat úlohu a prezentovat také průchody tunelem s odpovědí ležící v rámci sensorického horizontu. Zajímalo nás, zda jednotlivci změni svou preferovanou navigační strategii díky změně kontextu úlohy. Také nás zajímalo, zda má vliv komplexita prostředí (v našem případě simulovaná pomocí délky virtuálního tunelu a počtu zatáček) na změnu použitého referenčního rámce. Zajímalo nás, zda změna kontextu zadání úlohy, ovlivní preferenci referenčního rámce. Dle naší předběžné hypotézy by měla navigace v rámci sensorického horizontu vést k užívání egocentrického referenčního rámce oproti navigaci mimo sensorický horizont. Naopak se vzrůstající komplexitou navigace očekáváme přechod z egocentrického rámce k allocentrickému, jelikož neustálé přepočítávání polohy subjektu vůči výchozímu bodu není při více zatáčkách efektivní.

## 2 Metoda

Experimentální vzorek se skládal z 32 účastníků (5 žen a 27 mužů). Průměrný věk byl 29,4 rok. Malá velikost výzkumného vzorku je dána hlavním záměrem experimentu, spočívající v měření EEG aktivity a její následné analýze. Všichni účastníci měli normální zrak nebo příslušnou zrakovou korekci a neužívali dlouhodobě žádné léky ovlivňující psychiku. Pět účastníků byli leváci. Jedna ZO byla z následné analýzy vyjmuta, jelikož měla anamnézu organického poškození mozku. U dvou účastníků selhalo zařízení pro záznam odpovědí.

Jednotlivé stimuly (tunely) sestávaly s rovného, zahnutého a následně opět rovného úseku, přičemž délka průchodu jednotlivými úseky byla konstantní (3s rovný úsek a 3s úsek zahnutý). Zahnutí tunelu se lišilo v rozmezí od 90 do 150 stupňů s intervalem 30 stupňů. Pro každý ze 4 směrů (vlevo, vpravo, nahoru, dolů) tedy subjekt procházel 3 tunely s různou mírou zahnutí. Z pohledu subjektu sestávala úloha z následujících částí. Nejprve 1 vteřinu sledoval černou obrazovku se záměrným křížem. Poté se na 1s objevil statický pohled do tunelu nebo z tunelu ven (podle typu pohybu popsaném níže). Následně subjekt projížděl virtuálním tunelem konstantní rychlostí.. Celková délka tunelu závisí na počtu segmentů (9s pro tunely s jednou zatáčkou, 15s pro tunely s dvěma zatáčkami a 21s pro tři zatáčky v tunelu). Po prezentaci průjezdu se uprostřed obrazovky objevila šipka a souřadné osy, které mohl subjekt ovládat pomocí trackballu. Jeho úkolem bylo nastavit šipku směrem, kde podle něj ležel počátek tunelu.

Experiment sestával z 6 bloků, během kterých byl subjektu opakovaně prezentovány průchody virtuálním tunelem. Bloky se odlišovaly mírou komplexity a také způsobem prezentace průchodu



Pictures modified from Berthoz et al. (2004)

Obr.1 Rozdíl mezi egocentrickou a allocentrickou navigací

tunelem. První tři bloky obsahovaly průjezd virtuálními tunelů dopředu, tzn. počátek tunelu ležel vždy mimo senzorický horizont účastníka. Bloky se odlišovaly mírou komplexity prostředí. V prvním bloku procházel subjekt 15 tunelů, které obsahovaly pouze jeden zahnutý segment. Ve druhém a třetím bloku stoupal počet zatáček a tím i míra komplexity, konkrétně 2 zatáčky ve druhém bloku a 3 zatáčky ve třetím bloku. Bloky 4 až 6 byly koncipovány podobně. Vznášel počet zatáček (od 1 do 3), změněn byl pouze směr pohybu tunelu. Subjekt se pohyboval v tunelu v opačném směru a počátek tunelu byl v tomto případě v jeho senzorickém horizontu. Pořadí tunelu bylo náhodně namícháno tak, aby se neopakoval dvakrát po sobě stejný směr zatáčky. V každém bloku byly obsaženy 3 tunely, které obsahovaly pouze rovný segment. Sloužily jako kontrolní úlohy, které měly ověřit, že subjekt úlohu chápe a že nedává náhodné odpovědi. Tyto kontrolní tunely zároveň slouží k vyhodnocení EEG aktivity, jelikož obsahují signál během úseků, ve kterém je egocentrická a allocentrická navigace identická. Během průchodu všemi tunelů byla zaznamenávána aktivita EEG pomocí 62 elektrodového systému a vzorkovací frekvencí 500 Hz. Jako referenční elektroda slouží Fcz. Výsledky EEG analýzy nejsou v tomto článku prezentovány.

Experiment byl prezentován v klidné místnosti na speciálně upraveném křesle, umožňujícím měření v horizontální i vertikální poloze. V našem případě byl subjekt měřen pouze vsedě. Stimuly byly promítány na 24" monitor, který byl umístěn 50 cm před očima subjektu. Úroveň očí byla nastavena do středu monitoru. Stimuly byly promítány v rozlišení 1080x1080 bodů na displej počítače, přičemž zabíraly 34x34 stupňů zorného pole subjektu. Speciální zábrany kolem křesla znemožňovaly vnímání okolí. Subjekt měl u pravé ruky umístěn trackball pro zadávání správných odpovědí. Ten byl zvolen aby neměli ovládání zjednodušeno lidé, kteří často pracují s myši.

Po prezentaci byly vyhodnoceny odpovědi subjektu a podle nich určeno, který z referenčních rámců používá. Byly vyřazeny pokusy, během kterých překračovala chyba odhadu 45 stupňů v obou směrech horizontální i vertikální roviny. Tyto odpovědi jsme považovali jsme je jako nesplnění zadání. Zbývající odpovědi byly vyhodnoceny podle algoritmu, který určil, zda se subjekt orientoval v egocentrickém nebo allocentrickém rámci. Pokud subjekt umístil svou odpověď tak, že nešlo rozlišit, zda použil egocentrický nebo allocentrický rámec (šipka umístěna mezi dva rámce) s rozlišením menším než 3 stupně, byla odpověď vyhodnocena jako chybná a nebyla použita v další analýze.

Výsledky pro jednotlivé bloky byly vyhodnocovány také z hlediska použití preferovaného

referenčního rámce. Měřítkem hodnocení je počet odpovědí konsistentních s užitím jediného referenčního rámce. Jestliže se odpovědi v daném bloku shodovaly alespoň v 70 procentech, byl subjekt posuzován jako reprezentativní uživatel jednoho referenčního rámce. Každý účastník navíc podstoupil rozhovor, ve kterém byla upřesněna jeho strategie při navigaci.

### 3 Výsledky

V úvodní fázi jsme se zaměřili na analýzu strategií pro jednotlivé bloky. Zajímá nás poměr mezi účastníky, kteří používali allocentrický a egocentrický rámec. Nejprve jsme spočetli relativní podíly vzhledem k počtu účastníků, kteří měli konzistenci odpovědi nad 70 procent pro daný blok. Pro první blok (1 zahnutý segment při pohybu dopředu) preferovalo allocentrický rámec 65% ZO a egocentrický rámec 35%. Pro druhý blok (dva zahnuté segmenty po předu) byl poměr mezi allocentrickým a egocentrickým rámcem 87% oproti 13%. V případě třetího bloku jsme zjistili že 75% participantů používá allocentrický a 25% egocentrický. Překvapením byly výsledky pro identické tunely, které ZO projížděli pozadu (bloky 4-6). Při čtvrtém bloku (jeden zahnutý segment pozadu) se zcela změnila použitá strategie u účastníku, neboť pouze 3% preferovaly allocentrický rámec, přičemž 97% používalo egocentrický rámec. Podobný efekt se ještě více projevil v pátém a šestém bloku, žádný účastník nepreferoval allocentrickou strategii.

Jelikož je užití relativních podílů zkráceno nestejným počtem ZO, které preferovaly jeden referenční rámec v daném bloku, vypočetli jsem také absolutní podíly (viz. Tab. 1). Důvodem je zvyšující se obtížnost v tunelech s více zahnutými segmenty, které působily účastníkům problémy (bloky 2,3,5,6). Počet osob pro s nativní strategií byl pro bloky 1-6 roven 20,15,4,26,13 a 2 lidem. Právě malé počtu pro obtížné bloky znemožňují jejich přesnější interpretaci.

Blok/ <i>popis</i>	Strategie	Rel. podíl	Abs. podíl
Blok 1	Allo	65%	45%
<i>1zat/vpřed</i>	Ego	35%	24%
Blok 2	Allo	87%	45%
<i>2zat/vpřed</i>	Ego	13%	7%
Blok 3	Allo	75%	10%
<i>3zat/vpřed</i>	Ego	25%	3%
Blok 4	Allo	3%	3%
<i>1zat/vzad</i>	Ego	97%	86%
Blok 5	Allo	0%	0%
<i>2zat/vzad</i>	Ego	100%	45%
Blok 6	Allo	0%	0%
<i>3zat/vzad</i>	Ego	100%	7%

Tab.1. Přehled absolutních a relativních podílů nativních strategií pro jednotlivé bloky.

Druhá fáze se týkala analýzy chyb během orientace. Byly vyhodnoceny všechny odpovědi, které nepřesahovaly povolenou mez 45 stupňů (viz. výše). Jelikož distribuce chyby pro jednotlivé strategie nemělo normální rozložení, použili jsme pro testování statistické významnosti Wilcoxonův test.

Nejprve jsme vypočetli průměry allocentrických a egocentrických odpovědí pro všechny bloky. Průměrná úhlová chyba v horizontální rovině pro allocentrickou strategii byla -0.99 stupně (SD 14.84). Ve vertikální rovině byla průměrná úhlová chyba -0.47 stupně (SD 14.69). Pro egocentrický rámec byla horizontální průměrná úhlová chyba -0.12 (SD 15.86) a vertikální -0.23 (SD). Statistický test neprokázal signifikantní rozdíl v horizontální ( $p=0.25$ ) ani vertikální rovině ( $p=0.67$ ). Podrobnější analýza pro jednotlivé bloky prokázala signifikantní rozdíly pouze ve vertikální rovině pro třetí blok ( $p=0.02$ ) a pátý blok ( $p=0.03$ ). Přehled výsledků pro jednotlivé bloky je uveden v Tab.2 a 3.

Blok/ <i>popis</i>	Strategie	Prům. chyba	SD
Blok 1	Allo	-0.54	10.86
<i>1zat/vpřed</i>	Ego	0.01	11.81
Blok 2	Allo	-2.27	17.71
<i>2zat/vpřed</i>	Ego	-2.80	20.51
Blok 3	Allo	-0.11	17.21
<i>3zat/vpřed</i>	Ego	-0.71	16.31
Blok 4	Allo	-0.40	8.81
<i>1zat/vzad</i>	Ego	0.01	11.86
Blok 5	Allo	2.18	14.75
<i>2zat/vzad</i>	Ego	-0.19	17.51
Blok 6	Allo	-2.75	14.20
<i>3zat/vzad</i>	Ego	1.18	19.29

Tab.1. Přehled úhlových chyb v horizontální rovině pro jednotlivé bloky.

Blok/ <i>popis</i>	Strategie	Prům. chyba	SD
Blok 1	Allo	-0.38	11.68
<i>1zat/vpřed</i>	Ego	-0.19	8.17
Blok 2	Allo	-0.82	16.00
<i>2zat/vpřed</i>	Ego	-0.35	17.32
Blok 3	Allo	-1.44	16.32
<i>3zat/vpřed</i>	Ego	3.76	18.05
Blok 4	Allo	2.11	15.39
<i>1zat/vzad</i>	Ego	0.29	9.28
Blok 5	Allo	4.75	16.32
<i>2zat/vzad</i>	Ego	-0.07	14.99
Blok 6	Allo	-2.58	13.07
<i>3zat/vzad</i>	Ego	-3.62	18.08

Tab.1. Přehled úhlových chyb v horizontální rovině pro jednotlivé bloky.

## 4 Diskuze

Ve výsledcích se prokázalo, že schopnost orientace v prostoru nelze vnímat jako kognitivní schopnost, která je nezávislá na kontextu zadání úlohy či způsobu jejího plnění. Předchozí studie byly založeny pouze na zjišťování rozdílů mezi egocentrickou a allocentrickou

navigací, nestudovaly však detailně, které další faktory se na navigaci podílí. Současné výzkumy (Gramann et al., 2010, 2012) poukazují na fakt, že navigace v prostoru je ovlivněna specifickými faktory, ovlivňujícími výsledky. Až v posledních letech začíná být studován vlivu přirozené tendence jednotlivce (do češtiny obtížně přeložitelný výraz *proclivity*) řešit úlohu v kontextu jejího zadání. V prvotních studiích (Gramann et al., 2005) byl studována tendence používat odlišné navigační strategie při jednotném zadání úlohy. Následující studie prokázaly, že tento fenomén se řídí specifickými pravidly. Navigační tendence je ovlivňována allothetickými a idiothetickými faktory. Allothetické (environmentální) faktory jsou spojeny se strukturou a komplexitou prostředí. V předchozích studiích bylo prokázáno, že neroste chybovost při orientaci v jednotlivých referenčních rámcích se stoupající složitosti navigace (Gramann et al., 2005). V našem případě se však vliv tohoto faktoru projevil jak klesajícím počtem správných odpovědí v tunelech s více zatáčkami, tak ve vzrůstající průměrné chybě při rostoucí komplexitě navigace. Rozdíly ve výsledcích můžeme hledat v manipulaci s parametrem komplexity v obou studiích. Přestože jsou oba výzkumy založené na průchodu virtuálními tunely, vzrůstal se složitostí tunelů v této studii také čas potřebný na průchod. V případě Gramannovy studie stoupal počet zatáček, ale čas průchodu byl shodný. Výsledky nám tedy naznačují, že faktor času stráveného během navigace má vliv na chybovost navigace. Podle výsledků se ovšem neprojevil vliv komplexity na změnu preferované strategie, neboť poměry strategií pro jsou podobné pro bloky 1-3 stejně jako pro bloky 4-6. Dalším allothetickým faktorem, který se v této oblasti vyskytuje je založen na rozdílu mezi navigací v horizontální a vertikální rovině. V předchozích studiích jsme prokázali (Vavrečka et al., 2009, 2012), že při navigaci ve vertikální rovině dochází u části vzorku k přepnutí z egocentrického do allocentrického rámce. V této studii se přesnějšímu rozboru rozdílu mezi horizontální a vertikální navigací nevěnujeme, přesto hodláme tyto data v budoucnu analyzovat. Kromě allothetických faktorů je možné přikládat vliv na změnu strategie také idiothetickým (individuálním) faktorům. Během poslední studie (Vavrečka, 2012) jsme dospěli k hypotéze, že by se na preferenci strategie mohl podílet vliv sensorického horizontu, jelikož ve všech předchozích studiích byl počátek tunelu umístěn za ZO. Jak potvrdily výsledky, jestliže jsme umístili počátek tunelu v rámci sensorického horizontu subjektu (pomocí průchodů tunelem vzad), změnila se poměr allocentrické proti egocentrické navigaci z 65:35 pro cíle ležící mimo sensorický horizont pro pohyb vpřed na poměr 3:97 pro cíle v rámci sensorického horizontu při pohybu vzad.

## 5 Závěr

Tyto výsledky nás vedou k závěru, že je potřeba allothetické a idiothetické faktory dále podrobněji zkoumat, jelikož výsledky prokazují, že navigace v prostoru je v závislosti na kontextu. Jednotlivé vlivy je potřeba systematicky zmapovat, abychom byli schopni zachytit jejich vzájemné vztahy a interakce. V praktickém životě nám to třeba umožní vysvětlit, proč používají rozliční lidé jiné způsoby reprezentace prostoru a jakým způsobem můžeme zlepšovat nástroje usnadňující navigaci, jestliže máme k dispozici znalosti o odlišnostech při užívání specifických referenčních rámců.

## Poděkování

Tento příspěvek vznikl za podpory grantové agentury GAČR v rámci grantového úkolu číslo P407/11/P696.

## Literatura

- [1] Behrmann, M. (2000). Spatial reference frames and hemispatial neglect. In *New cognitive science*.
- [2] Colby, C.L., Goldberg, M.E. (1999). Space and attention in parietal cortex. *Annual Review of Neuroscience*, 22:319-49.
- [3] Fink, G.R., Marshall, J.C., Weiss, P.H., Stephan, T., Grefkes, C., Shah, N.J., Zilles, K., Dietrich, M., (2003). Performing allocentric visuospatial judgments with induced distortion of the geocentric reference frame: an fMRI study with clinical implications. *NeuroImage* 20, 1505–1517.
- [4] Galati, G., Lobel, E., Vallar, G., Berthoz, A., Pizzamiglio, L., Le Bihan, D., (2000). The neural basis of egocentric and allocentric coding of space in humans: a functional magnetic resonance study. *Exp. Brain Res.* 133, 156–164.
- [5] Gramann, K., Müller, H.J., Eick, E., Schönebeck, B., (2005). Empirical evidence for separable spatial representations in a virtual navigation task. *J Exp Psychol Hum Percept Perform* 31, 1199–1223.
- [6] Gramann, K., Müller, H., Schönebeck, B. & Debus, G. (2006). The neural basis of egocentric and allocentric reference frames in spatial navigation: Evidence from spatio-coupled current density reconstruction. *Brain Research*, 1118, 116-129.
- [7] Gramann, K., El Sharkawy, J., & Deubel, H. (2009). Eye-Movements during Navigation in a Virtual Tunnel. *International Journal of Neuroscience*, 119(10), 1755–1778.
- [8] Gramann, K., Onton, J., Riccobon, D., Müller, H.J., Bardins, S., Makeig, S. (2010). Human brain dynamics accompanying use of egocentric and allocentric reference frames during navigation. *Journal of Cognitive Neuroscience*, 22(12), 2836-2849

- [9] Gramann, K. (2012). Embodiment of spatial reference frames and individual differences in reference frame proclivity. [Review]. *Spatial Cognition and Computation*.
- [10] Lin, T.C., Chiou, T.C., Ko, L.W., Duann, J.R., Gramann, K., (2009). EEG-based spatial navigation estimation in a virtual reality driving environment. In: *Proceedings of the Ninth IEEE International Conference on Bioinformatics and Bioengineering*, 435-438.
- [11] McCloskey, M. (2001). Spatial representation in mind and brain. In B. Rapp (Ed.), *What deficits reveal about the human mind/brain: A handbook of cognitive neuropsychology* (pp.101-132). Philadelphia: Psychology Press.
- [12] Pfeifer, R., Scheier, C.(1999). *Understanding Intelligence*. MIT Press, Cambridge, USA
- [13] Vavrečka, M., Gerla, V., Lhotská, L., Brunovský, M., (2012). Frames of reference and their neural correlates within navigation in a 3D environment, *Visual Neuroscience* 29 (03), 183-191
- [14] Vavrečka, M., (2009). The neural correlates of spatial reference frames processing, *Cognitive processing*, 10:2, Springer Berlin, p. 342-345.



## Brána věděni otevřena: nový pohled na výpočty

Jiří Wiedermann

Ústav informatiky AV ČR  
Pod Vodárenskou věží 2, 182 07 Praha 8, Česká republika  
jiri.wiedermann@cs.cas.cz

### Abstrakt

V přednášce představíme nový pohled na výpočty – totiž jako na procesy generující znalosti. Z tohoto zorného úhlu výpočty realizují různé více či méně úplné a v různé míře formalizované teorie s podobnou vnitřní strukturou popisující mono- či heterogenní znalosti. Nové vidění výpočtů umožňuje chápat dosavadní vývoj počítačů a informačních technologií ve zcela nové perspektivě. Dovoluje extrapolaci vzhledem k budoucnosti počítání, které bude zaměřené na získávání, shromažďování a tvořivé využívání znalostí, a to vše převážně v přirozeném jazyce. Evoluce a efektivita výpočtů bude vnímána prostřednictvím jejich schopností vyvářet nové poznatky. Tento přístup má široké konotace v oblasti umělé inteligence, v kognitivních vědách, ve filozofii, epistemologii a metodologii vědy.

### 1 Úvod

Výpočet již není tím, co odjakživa býval. Ještě tak do konce osmdesátých let dvacátého století se otázkou, co je to výpočet, nikdo netrápil. Odpověď byla jasná: výpočet je to, co dělá Turingův stroj [10], [16], anebo ekvivalentní výpočetní zařízení. S příchodem moderních výpočetních technologií a počítačových sítí, společně s novými metodologiemi a s novými poznatky fyziky a biologie bylo stále jasnější, že výpočet je daleko širší, omnoho běžnější a podstatně komplexnější fenomén než ten, který lze přímočaře modelovat pomocí Turingova stroje. Příklady takových „problematických“ výpočtů jsou biologicky motivované výpočty neuronových sítí, DNA výpočty, sebe-sestavovací struktury, davové (swarm) výpočty, amorfní výpočty, atd., anebo fyzikou inspirované výpočty, jako jsou kvantové výpočty, relativistické výpočty a či hyper-Turingovské výpočty. V neposlední řadě jsou zde

„technologicky umožněné“ výpočty realizované v Internetu a pomocí mobilních sítí.

Jak na tyto nové trendy reagovala vědecká komunita z příslušných vědních oblastí, zejména informatiky, fyziky a filozofie? (Ne)překvapivě, nesjednotila se na společném názoru, nýbrž se rozpadla na několik názorových skupin. Např. profesor informatiky D. Frailey zastává radikální názor, že jakýkoliv proces realizuje výpočet [11]. Další informatici, např. P.S. Rosenbloom [13] anebo R. Bajcsy [3] se domnívají, že výpočet je proces transformace informace. Jiní význační badatelé k tomu dodávají další podmínku: proces transformace informace musí být řízen přechodovou funkcí a v podstatě jde o manipulaci symbolů (např. Fortnow [10], Denning [7], Connery [6] anebo filozof Searle [2]). Ještě další – např. A.V. Aho [1] anebo Searle [15] požadují, aby výpočetní proces odpovídal nějakému výpočetnímu modelu. Fyzik Ed Fredkin [12] to shrnul lakonicky: „*Věc se má takto: výpočet je normálně přehršle bitů, která se vyvíjí v čase, a dále zde máme jeho pohon – počítač*“. Další fyzik, D. Deutsch [9] přitvrzuje ještě více – výpočetní model musí být fyzikálně realizovatelný. Zdá se, že kolik badatelů, tolik názorů.

Bohužel, žádný se shora uvedených názorů není bez chyby - nezachycují totiž některé aspekty výpočtů, které se zdají být důležitými. Chápání výpočtů jako procesy vede k příliš širokému okruhu objektů a jevů, které jsou nadány výpočetní schopnosti – např. kámen anebo počasí. Podobně je tomu tak i v případě nahlížení na výpočty jako na procesy zpracovávající informace. Každou změnu lze totiž chápat jako součást takového procesu, a tudíž jakákoliv změna ve vesmíru by byla součástí nějakého výpočetního procesu. Pokud požadujeme, aby změny byly řízeny přechodovou funkcí (v rámci nějakého formálního modelu výpočtu), tak se

dostáváme k cirkulární definici, protože výpočet pak je cokoliv, co konají počítače, a naopak, cokoliv, co činí počítače, je výpočet. Navíc, lpění na formálním modelu výpočtu, který musí odpovídat fyzikálně realizovatelnému procesu sráží pojem výpočtu na konkrétní fyzikální úroveň, která může být velmi vzdálena od rozumné úrovně abstrakce, na které je výpočet uvažován (např. výpočet nad reálnými čísly).

Je zajímavé, že všechny předchozí pokusy o definici pojmu výpočet mají přece jenom několik věcí společných. První věcí je konsensus, že výpočet je proces. Druhou věcí je skutečnost, že všechny definice se soustřeďují na to, JAK je výpočetní proces realizován, tj. na to, JAK funguje příslušný „hardware“. To ovšem neumožňuje dostatečný vzhled to pojmu výpočtu a nutí nás vidět jako výpočet jakoukoliv nesmyslnou operaci s daty. Na druhé straně, veškerá naše zkušenost s výpočty ukazuje přesně opačný směr: nás vždy především zajímá, CO dělá výpočet, co činí pro nás návrháře výpočtů, uživatele, pozorovatele. Co dělá výpočet je vyjádřeno v jeho návrhu (a v případě programovatelných systému v návrhu jeho software). Znalost toho, jakým způsobem dělá výpočet to, co dělá, je méně zajímavá, protože opět – veškerá naše zkušenost (i teorie) ukazuje na to, že to, co dělá výpočet lze implementovat mnoha různými způsoby pomocí rozličného hardware. Závěr je, že pro pochopení pojmu výpočtu bychom se měli soustředit nikoliv na to, JAK výpočet probíhá, nýbrž na to, CO výpočet dělá, jaký je jeho smysl a užitek.

Nuže, co to je, co výpočet dělá? Naše odpověď je prostá: *výpočet generuje znalosti*. Výpočet generuje znalosti nad znalostní doménou, pro kterou byl systém navržen, pokud byl navržen, anebo v rámci, ve kterém se „sám“ vyvinul, pokud se jedná o evoluční systém.

Pokud je toto řečeno, musíme umět odpovědět na otázku, co je to znalost. To je ovšem ošemetná otázka, která se zkoumá zejména ve filosofii tisíce let. Z hlediska vyčerpávající definice je to ještě hůře uchopitelný pojem jako pojem výpočtu. Do jisté míry je to však výhoda, protože to činí tezi, že výpočet je generování znalostí, robustnou. Pokusy o

její vyvrácení tak mohou vést jak k lepšímu pochopení pojmu výpočtu, tak i pojmu znalosti.

Pro naše další účely bude plně postačující definice převzatá z wikipedie.

*Znalost je obeznámenost s někým anebo s něčím a zahrnuje fakta, informace, chování, popisy anebo schopnosti získané zkušeností anebo výchovou. Vztahuje se k praktickému anebo teoretickému porozumění subjektu. Může být implicitní (např. v případě schopností anebo odborné znalosti) anebo explicitní (např. v případě teoretického porozumění subjektu). Může být více anebo méně formální a systematická.*

Dle předchozí definice je znalost záležitost výsostně závislá na pozorovateli (observer-dependent) a takovým je i pojem výpočtu. To je jeden z důležitých rozdílů mezi chápáním výpočtu jako zpracování informací, které na jedné straně činí tento pojem méně závislým na pozorovateli, ale na straně druhé vede k příliš obecnému chápání tohoto pojmu, jak bylo vysvětleno výše.

Předností naší definice výpočtu je více. Poprvé, umožňuje zřetelně oddělit objekty, které počítají (tj. generují znalost) od těch ostatních. Například, podle naší definice můžeme kameni přiřadit schopnost počítat (viz např. [5]) pouze tehdy, bude-li dokazatelně generovat znalost. Podruhé, nová definice dovoluje klasifikaci výpočtů z hlediska kvality a kvantity znalostí, které produkují. Pořetí, nezávislost definice na výpočetním mechanismu je vítána protože pokrývá řadu známých a ještě neznámých výskytů výpočtů. Počtvrté, definice podporuje myšlení o výpočtech na vysoké abstraktní úrovni, která je důležitá při návrhu umělých systémů a pro pochopení činnosti jiných přirozených systémů které vznikly evolucí. Popáté, definice řeší některé problémy, které se vztahovaly ke „klasické“ definici výpočtu. Např. v kognitivní vědě se široce diskutuje otázka, co je kognice, když ne výpočet [18]. Pokud vidíme kognici jako schopnost získávat, shromažďovat a využívat znalosti, pak tato definice kognice odpovídá naší definici výpočtu v jeho nejrozvinutější podobě. Konečně, naše definice má velký metodologický a filozofický význam, protože

<b>Současné výpočetní systémy</b>		
<b>Výpočetní systém</b>	<b>Odpovídající znalostní doména</b>	<b>Typ vytvářené znalosti</b>
Akceptory Rozeznávače  Translátory Vědecko-technické výpočty Dokazovače vět Operační systémy Databázové a informační systémy Řídicí systémy Vyhledávače  Umělé znalostní systémy	Formální jazyky Formální jazyky  Funkce, relace Matematika Logika Počítačové komponenty a periferie Relace nad strukturovanými konečnými doménami Vybrané oblasti lidské činnosti Relace nad nestrukturovanými potenciálně neohrazenými doménami Reálný svět, věda	Příslušnost k jazyku Příslušnost/nepříslušnost k jazyku Funkční hodnota Řešení výpočetních problémů Důkazy Správa počítačových aktivit Odpovědi na formalizované dotazy Monitorování a řízení Odpovědi na dotazy v přirozeném jazyce Hypotézy, nové znalosti, vysvětlení
<b>Přirozené výpočetní systémy</b>		
Organizmy, buňky Mozek, mysl, sociální síť Vesmír	Životní prostředí Poznaný svět Vesmír, vědy, život,...	Život, chování, inteligence Znalosti o světě Život
<b>Ne-Turingovské výpočty</b>		
Kružítka a pravítka BSS počítač [4] Orákula  Superturingovské výpočty	Euklidovská geometrie Teorie reálných čísel Podmnožina A množiny přirozených čísel  Formální jazyky třídy $\Sigma_2$	Euklidovské konstrukce Hodnoty reálných funkcí Charakteristická funkce množiny A Příslušnost k jazyku

soustředí pozornost na smysl výpočtů, kterým je vytváření znalostí a tím staví výpočty do centrální pozice nástroje, který je v pozadí veškerého pokroku.

Argumenty, dokumentující a podporující naše chápání výpočtu na příkladech dosavadního vývoje výpočetních technologií uvedeme ve druhé části tohoto článku. Ve třetí části se zaměříme na vnitřní strukturu výpočtů z hlediska generování znalostí. Čtvrtá část se zabývá problematikou generování

nových znalostí. Závěrečná pátá část je shrnutím Článek vznikl na základě původního článku [19] a jeho doplněním dalším materiálem.

## 2 Výpočet jako generování znalostí

Cílem této části je podat argumenty podporující následující tezi:

**Teze:** Výpočet je proces generování znalosti.

Tabulka na předchozí straně uvádí přehled různých výpočetních systémů. U každého z nich definuje odpovídající znalostní doménu a typ vytvářené znalosti.

V této tabulce je důležité si všimnout několik věcí. Předně, první část tabulky je uspořádána více-méně chronologicky podle pořadí, ve kterém se jednotlivé systémy objevily, resp. byly uvažovány v teorii. Z této části tabulky je dále zřetelný nárůst požadavků na generování znalostí: čím dále v tabulce, tím více znalostí výpočty generují, tím obecnější a méně formální jsou znalostní domény, zachycují stále větší výseky reálného světa, zvyšuje se jejich interaktivita a tím více se vkrádá do komunikace se systémy přirozený jazyk. Z tohoto hlediska jsou klasické systémy uvažované v teorii vypočítatelnosti velmi primitivními a v podstatě degenerovanými systémy pro generování znalostí. Trend ke zvyšování kvality a kvantity produkovaných znalostí je nicméně velmi zřetelný.

Druhá část tabulky zachycuje tzv. *přirozené výpočetní systémy*. To jsou systémy, které nejsou navrhované lidmi. Vzhledem ke svým schopnostem se přirozeně řadí mezi výpočty, což by podle „klasických“ definic výpočtu nebylo možné, protože neznáme odpovídající výpočetní mechanismy.

V poslední části je naše definice natahována na skřípec. Původně se jednalo o pokus falsifikovat tezi pomocí případů, které jsou obecně pokládány za výpočty, byť poněkud exotické. Jak vidno, teze přežila i tento útok. Tyto případy současně ukazují, proč není dobrý nápad požadovat primárně od výpočtů, aby byly fyzikálně realizovatelné. To je v dobrém souladu s praxí – v geometrii, matematice

nejdůležitějších přínosů nového chápání počítání. a teorii vypočítatelnosti jsou tyto procesy bez problému považovány za výpočetní procesy.

## 3 Struktura znalostí

Je jisté, že pokud se soustředíme na to, CO dělají výpočty místo toho, JAK to dělají, ztratíme možnost zkoumat jemnější detaily tohoto procesu např. tak, jak se to děje v teorii vypočítatelnosti anebo v teorii výpočetní složitosti. Na druhé straně, náš přístup umožní pohled na jiné, doposud přehlížené aspekty výpočtů. Náš přístup totiž umožňuje vhléd do struktury znalostí, které využívají výpočty pro generování nových znalostí. Pro dosažení svých cílů využívá výpočet znalost odpovídající znalostní domény. V dalším budeme předpokládat, že odpovídající znalostní doména je dána ve formě teorie.

V tomto kontextu teorii nebude chápat pouze v jejím úzkém smyslu, tak jak je používána zejména v matematice anebo v logice, nýbrž ji budeme chápat jako analytický nástroj pro porozumění, vysvětlení a zodpovězení dotazů, poskytování řešení a predikcí v různých oblastech života, anebo pro generování chování. Teorie má obvykle podobu faktů, tvrzení a principů pro odvozování dalších tvrzení, ale může vypadat i jinak – např. jako sémantická síť, soubor omezení a podmínek, platných pro daný výpočet, atd. V rámci takové teorie výpočet produkuje znalost ve formě odpovídající uvažované znalostní doméně. Takové znalosti udržované ve formě znalostní báze mohou být součástí uvažované teorie.

Následující tabulka udává přibližný a orientační pohled na strukturu znalostí v různých znalostních doménách. Příklady pokrývají znalostní domény od formálních teorií až po neformální (resp. těžce formalizovatelné) teorie zachycující chování v běžném životě.

Z tabulky je vidět různá úroveň formálnosti, úplnosti a pravdivosti odpovídajících teorií. Ve složitějších případech je charakteristická heterogenní znalost. V těchto případech je důležitým prostředníkem mezi teoriemi přirozený jazyk. Pro přirozený jazyk je důležitá nejen jeho syntaxe, ale

	Logika a matematika	Přírodní vědy a filozofie	Mysl a humanoidní kognitivní systémy
<b>Oblast zájmu</b>	Abstraktní entity	Empirická data, ideje	Percepce, kognice
<b>Základní prvky znalostí</b>	Axiomy, definice,	Pozorování, fakta	Stimuly, multi-modální koncepty, domněnky (beliefs), epizodická paměť
<b>Odvozovací pravidla</b>	Deduktivní systém	Racionální myšlení, logika	Pravidla a asociace formované procesem statistického učení
<b>Finální forma znalostí</b>	Výroky, teorémy, důkazy	Tvrzení, teorémy, hypotézy, vysvětlení, vědecké zákony, predikce	Konceptualizace, chování, komunikace, přirozený jazyk, myšlení, znalost světa formulovaná převážně v přirozeném jazyce a ve formě teorií

zejména jeho sémantika, která „přiřazuje“ význam jednotlivým slovům a větám jazyka. Sémantika je opět znalost a jako taková může být také reprezentována teorií. Z tohoto hlediska nabývá každý výpočet, i ten, který generuje znalosti na základě porozumění přirozenému jazyku, homogenní strukturu. Znalosti takového výpočtu mají podobu systému spolupracujících teorií. Tento systém je neuvěřitelně složitý, protože v zásadě pro každé slovo jazyka musí obsahovat teorii správného použití tohoto slova. Taková teorie obecně závisí nejen na slově samotném, ale i na kontextu, ve kterém je slovo použito. V případě vtělených kognitivních systémů kontext neznamená pouze gramatický kontext, ale obecně celkovou percepční situaci. To vše vede ke značné provázanosti příslušných teorií. Obecně lze říci, že o takových teoriích a jejich spolupráci víme velice málo. Zde je však vidět přínos pohledu na výpočty jako na procesy generující znalosti – jakkoliv hluboký a detailní (klasický) pohled na mechanismy realizující výpočty nemůže nijak přispět k osvětlení sémantiky výpočtů.

Právě prezentovaný vhléd do struktury znalostí má ještě jeden principiální význam. Ilustruje, že *výpočty generují novou znalost ze znalosti*, která je v pozadí každého výpočtu. S jistou nadsázkou lze říci, že *výpočet je znalost v akci*.

#### 4 Struktura výpočtů

Ve shodě s naší tezí, že výpočet je generování znalostí, můžeme předpokládat, že za každým výpočtem existuje teorie, alespoň v principu (tj. v praxi nemusí být explicitně či do detailu známá). V rámci této teorie příslušný výpočetní proces pro daný vstup generuje znalost, vyjádřenou opět prostředky této teorie (viz předchozí tabulku). Samozřejmě, aby to platilo, musí existovat důkaz, vysvětlení, že výpočetní proces funguje tak, že v rámci dané teorie pro daný vstup skutečně bude generovat požadovanou znalost.

Nechť  $C$  je výpočet, jenž v rámci teorie  $T$  generuje pro daný vstup  $I$  znalost  $K$  prostřednictvím výpočetního procesu  $P_M$ , který pro vstup  $I$  realizuje výpočetní mechanismus  $M$ , a necht'  $E$  je důkaz či vysvětlení, že tomu tak skutečně je. Výpočet je pak jednoznačně určen předchozí pětici:  $C = (T, I, K, P_M, E)$ .

Vysvětlení  $E$ , že proces  $P_M$  generuje  $K$  v rámci teorie  $T$  pro daný vstup  $I$  závisí na vlastnostech procesu  $P_M$  (a tedy zprostředkovaně na  $M$ ). Ze zkušenosti s klasickými výpočty víme, že stejný problém lze řešit na stejném typu počítače různými způsoby, a také jej lze řešit na různých typech počítačů (deterministicky, nedeterministicky, sekvenčně, paralelně, atd.), a to platí samozřejmě obecně, nejen pro klasické výpočty. Takže  $E$  závisí na  $P_M$  a  $I$ , ale také, samozřejmě, na  $T$  a  $K$ . Nicméně,  $P_M$  nemá přístup k  $T$  ani  $K$ , takže pokud  $P_M$

potřebuje znát teorii T anebo její část, musí příslušný poznatek obsahovat ve svých datech. To znamená, že pro K vlastně existují v obecnosti dva různé důkazy. Jeden důkaz je veden prostředky teorie T, a druhý důkaz E musí také brát zřetel na vlastnosti procesu  $P_M$ .

Všimněme si, že výpočet v našem pojetí je definován pouze pro (jeden) daný vstup a vlastně představuje „jeden běh“ zařízení M na tomto vstupu. Takže předchozí důkazy jsou důkazy právě pro jeden vstup. Důkaz korektnosti výpočtu pro daný vstup není totéž jako důkaz korektnosti algoritmu, který platí pro všechny přípustné vstupy.

Již v současné praxi programování je situace taková, že tvůrce programu by si měl být vědom, alespoň v principu, teorie T, požadovaného výsledku K, a má k dispozici výpočetní model M, v rámci kterého navrhne výpočetní proces  $P_M$  tak, aby generoval požadovanou znalost K. Důkaz, že tomu tak je, by také měl být schopen dodat, protože jinak si nemůže být jistý, že proces  $P_M$  skutečně dělá to, co se po něm chce.

U programů specializovaných na generování znalostí může být situace taková, že výpočetní proces P využívá přímo a pouze prostředky teorie T. To je například možné, pokud teorie T je tzv.  $\lambda$ -kalkul a výpočetní proces používá pouze odvozovací pravidla tohoto kalkulu.

Poslední věc, kterou je dobré si uvědomit, je skutečnost, že „klasické“ definice výpočtu zmiňované v úvodu tohoto článku, zhruba odpovídají v naší notaci zápisu  $C = (P_M)$ , tj. jakémukoliv procesu generovanému zařízením M.

Vyjadřovací a vysvětlovací síla teorie, která je v pozadí každého výpočtu, podmiňují schopnost příslušného výpočtu generovat novou znalost. Jak je však možné, aby výpočet generoval znalosti, které by již nebyly nějakým způsobem obsaženy ve znalostech (rozuměj: ve znalostní bázi) resp. datech, které má výpočet k dispozici?

Odpověď je tato: na základě dotazu či příkazu generovaného interně výpočtem anebo dodaného externě uživatelem výpočet může odhalit nové

vztahy mezi poznatky, které má k dispozici ve znalostní bázi anebo najít takové poznatky, které splňují požadovaný vztah. Pojmem „odhalit“ anebo „objevit“ myslíme doslova učinit něco poprvé zjevným, tj. poskytnout informaci o něčem, co až do té doby nebylo známé.

Pokud tedy mluvíme o tvořivosti ve smyslu generování znalostí, musíme mít pořad na mysli, že znalost může být generována pouze ze znalosti. Tudíž musí existovat dva protichůdné procesy: proces získávání znalostí, a proces generování znalostí.

Existuje mnoho způsobů, jak získávat znalost: pomocí usuzování a logiky, vědeckou metodou, metodou pokusu a omylu, pomocí algoritmu, zkušenosti, intuicí, od autority, čtením, poslechem, konverzací, pozorováním, atd. Účelem *získávání znalostí* není pouze naplnit systém znalostmi, ale především je uspořádat, vpravit – opět pomocí výpočtu – do existujících teorií, které se vztahují k dané znalostní doméně. To znamená, že proces získávání znalostí je současně procesem budování, doplňování anebo aktualizace stávajících teorií. V jistém smyslu je tento proces i procesem generování znalostí uvnitř výpočtu.

Účelem procesu *generování znalostí* – tj. výpočtu – je produkovat znalost v odpovědi na interní anebo externí podněty. Budeme rozlišovat dvě základní metody tvorby znalostí: syntaktické a sémantické dolování znalostí. Obě metody používají specifické inferenční mechanismy, jejichž účelem je objevit skryté souvislosti v datech.

*Syntaktické dolování znalostí* bere v potaz pouze syntaxi odpovídajících dat a nikoliv jejich význam (sémantiku), a také inferenční mechanismus základní teorie. Tento proces hledá skryté vzory zejména v rozsáhlých datech kombinací metod umělé inteligence, strojového učení, statistiky a databázových systémů.

*Sémantické dolování dat* je hlavním nástrojem kreativity. Také hledá skryté souvislosti v datech, avšak nikoliv syntaktické, nýbrž sémantické souvislosti. Obyčejně se vychází ze sémantiky základního vzoru a hledá se jiný vzor, často v jině

sémantické oblasti, který má podobnou sémantickou strukturu jako původní vzor a navíc splňuje další sémantická omezení zadaná v dotaze.

Úkol sémantického dolování je často zadán v přirozeném jazyce. To znamená, že hledané prvky a vztahy, které mají splňovat, jsou popsány v lingvistických termínech. To neobyčejně komplikuje úlohu, protože význam lingvistických termínů musí být známý. Výpočet musí „rozumět“ přirozenému jazyku. Jak jsme zmínili v předchozí části, význam každého prvku přirozeného jazyka je popsán odpovídající teorií. To znamená, že sémantické dolování vlastně vyžaduje odhalení podobností mezi různými teoriemi.

Nalezení v jistém smyslu „paralelní“ teorie k dané teorii v jiné oblasti znalostí zásadně přispívá k lepšímu pochopení každé ze zúčastněných teorií. Navíc umožňuje očekávat, že vztahy, které platí v jedné teorii, budou platit i mezi odpovídajícími prvky teorie druhé. To jsou důležité předpoklady pro vzhled do daných teorií a jejich vysvětlení a pochopení.

Bohužel, v současné době víme velmi málo o obecných mechanismech odhalování sémantických znalostí. Skoro nic nevíme o odhalování sémantické nekonzistence mezi různými teoriemi, týkajícími se stejné znalostní domény. Odstranění takových nekonzistencí může vést k průlomovým poznatkům [8].

Jednou ze základních metod sémantického dolování, která má za sebou tisíciletou historii (aniž by to bylo považováno za metodu sémantického dolování), jsou *analogie* a *metafory*.

Jako příklad uveďme následující úkol, jehož cílem je nalezení analogie. Hledáme pojem, který má k pojmu *město* podobný vztah, jako má pojem *řeka* k pojmu *krajina*. Pokud takový pojem najdete (není to *kanalizace*) a dovedete ho logicky zdůvodnit, tak máte výbornou ilustraci toho, jak „v sobě“ můžete nalézt znalost, o které jste nevěděli, a také o složitosti tohoto procesu.

Problematika sémantického dolování znalostí však již přesahuje záměr tohoto článku. Zájemce

odkazujeme např. na práci [14], [17] a [20] a mnoho dalších prací z oboru dolování znalostí.

Tato část opět ukázala význam chápání výpočtů jako procesů generujících znalosti: přirozeným způsobem upřelo naši pozornost na problematiku kreativity, protože znalosti nelze generovat jinak než kreativně. Vidění výpočtů jako zpracování informace takový potenciál nemá, protože se nezajímá o smysl výpočtů.

## 5 Závěr

V práci jsme navrhli nový, alternativní pohled na výpočty, který odhlíží od jejich realizace a naopak se soustřeďuje na to, co je úkolem výpočtů: generování znalostí. Ukázali jsme, že tento pohled vychází z dosavadních trendů v oblasti informačních technologií a má mnohem větší motivační a vysvětlovací potenciál než tradiční chápání výpočtu jako procesu transformace informace. Epistemologický přístup k výpočtům staví do centra pozornosti jejich smysl – *proč vlastně počítáme*. Umožňuje soustředit se na problematiku generování znalostí, která bude ve stále větší míře dominovat nasazení informačních (a robotických) technologií. Výpočet se tak stává univerzálním nástrojem pro rozvoj vědění.

## Literatura

- [1] Aho, A.V.: Computation and Computational Thinking. Magazine Ubiquity, Volume 2011. Issue January, January 2011, Article no. 1
- [2] Almond, P.: Machines Like Us, an interview by Paul Almond with John Searle. Machines Like Us, March 2009, <http://machineslikeus.com/interviews/machine-s-us-interviews-john-searle-0>
- [3] Bajcsy, R.: Computation and Information. Comput. J. 55(7): 825 (2012)
- [4] Blum, L., Shub, M., Smale, S.: On a Theory of Computation and Complexity over the Real Numbers: NP-completeness, Recursive Functions and Universal Machines. Bulletin of the American Mathematical Society 21 (1), 1989

- [5] Chalmers, D. J.: Does a Rock Implement Every Finite-State-Automaton? *Synthese*, 1996, Vol.8 No.3, pp.309-333, Kluwer
- [6] Conery, J. S.: Computation is Symbol Manipulation. *Comput. J.* 55(7): 814-816 (2012)
- [7] Denning, P. J.: What is Computation? (opening statement), *Magazine Ubiquity*, Volume 2010, Issue October, October 2010, Article No.~1
- [8] D. Deutsch: Creative Blocks, *AEON Magazine*, 02 October 2012
- [9] Deutsch, D.: What is Computation? (How) Does Nature Compute? In: Zenil, H. (Editor): *A Computable Universe: Understanding and Exploring Nature as Computation*, World Scientific Publishing Company, 2012, pp. 551-566
- [10] Fortnow, L.: The Enduring Legacy of the Turing Machine. *Comput. J.* 55(7): 830-831 (2012)
- [11] Frailey, D. J.: Computation is Process. *Magazine Ubiquity*, Volume 2010, Issue November, November 2010 Article No.~5
- [12] Fredkin, E.: What is Computation? (How) Does Nature Compute? (Transcription of a live panel discussion, with participants C.S. Calude, G., J. Chaitin, Ed Fredkin, T .J. Leggett, R. de Ruyter, T. Toffoli, S. Wolfram). In: Zenil, H. (Editor): *A Computable Universe: Understanding and Exploring Nature as Computation*, World Scientific Publishing Company, 2012, pp. 673-726
- [13] Rosenbloom, P. S. Computing and Computation. *Comput. J.* 55(7): 820-824 (2012)
- [14] G. Lakoff, and M. Johnson: *Metaphors We Live By*. Chicago, IL: The University of Chicago Press, 1980
- [15] Searle, J.: Is the Brain a Digital Computer? *Proceedings and Addresses of the American Philosophical Association* 64 (1990) 21-37
- [16] A.M. Turing, On computable numbers, with an application to the Entscheidungsproblem, *Proc. London Math. Soc. Series 2*, 42 (1936) 230-265
- [17] P. D. Turney: Similarity of semantic relations. *Computational Linguistics*, 32 (3), 2006, pp. 379–416.
- [18] van Gelder, T.: What Might Cognition Be, If Not Computation? *The Journal of Philosophy*, Vol. 92, No. 7, 1995, pp. 345-38
- [19] Wiedermann, J., van Leeuwen. J.: Rethinking Computation. *Proc. 6<sup>th</sup> AISB Symposium on Computing and Philosophy, The Scandal of Computing: What Is Computation*. The University of Exeter, April 2013
- [20] Wiedermann, J.: The Creativity Mechanisms in Embodied Agents: An Explanatory Model. To appear in: *Proc. 2013 IEEE Symposium Series on Computational Intelligence (SSCI 2013)*, April 2013



# Register autorov

## A

Andrejková, Gabriela ..... 5

## B

Bačová, Viera ..... 141

Ballová Mikušková, Eva ..... 53

Bečev, Ondřej ..... 16

Blšťáková, Petra ..... 37

Blahuta, Jiří ..... 26

Burian, Jan ..... 43

## C

Chalupa, David ..... 58

Cienciala, Luděk ..... 65

Ciencialová, Lucie ..... 65

Cimrová, Barbora ..... 79

Clementis, Ladislav ..... 87

## Č

Čavojová, Vladimíra ..... 53

Čermák, Petr ..... 26

Číhalová, Martina ..... 73

## D

Dillinger, Viliam ..... 93

Dudeková, Katarína ..... 101, 141

## F

Farkaš, Igor ..... 107, 223

## G

Gerlich, Jakub ..... 180

## H

Hanák, Róbert ..... 53

Horváth, Marek ..... 166

Husáková, Martina ..... 115

## J

Jagla, Fedor ..... 79

Janikova, Anna ..... 123

## K

Kamhal, Dezider ..... 128

Knott, Alistair ..... 265

Konečný, Matúš ..... 133

Kostovičová, Lenka ..... 101, 141

Kužilek, Jakub ..... 279

Kuzma, Miron ..... 147

## L

Lhotská, Lenka ..... 279

Lúčný, Andrej ..... 153

## M

Malinovský, Ľudovít ..... 157

Marček, Dušan ..... 166

Markošová, Mária ..... 173

Matula, Kamil ..... 123

Menšík, Marek ..... 180

Mikulecký, Peter ..... 188

## N

Nahodil, Pavel ..... 197

## O

Olajec, Libor ..... 205

Oravec, Jozef ..... 5

## **P**

Pecháč, Matej ..... 211

Pstružina, Karel ..... 216

## **R**

Rebrová, Kristína ..... 223

Riečanský, Igor ..... 79

Rybár, Ján ..... 231

## **S**

Sirota, Miroslav ..... 141

Smutný, Zdeněk ..... 258

Stachoň, Zdeněk ..... 235

## **Š**

Šašinka, Čeněk ..... 235

Šefránek, Ján ..... 243

Šimeček, Michal ..... 249, 254

Štěpán, Jan ..... 73

Šuba, Petr ..... 272

## **T**

Takáč, Martin ..... 265

Tučník, Petr ..... 272

Tůmová, Štěpánka ..... 123

## **V**

Vavrečka, Michal ..... 279

Vician, Michal ..... 153

Vítků, Jaroslav ..... 197

## **W**

Wiedermann, Jiří ..... 285