

Contents

L. BIELIK: The Abductive Model of (Scientific) Explanation	5
M. DUŽI: Towards an Extensional Calculus of Hyperintensions	20
B. GEISTOVÁ ČAKOVSKÁ: Semantic and Pragmatic Meaning of Language Expressions.	46
V. HAVLÍK: Metamethodology and Naturalism.	56
M. KOSTEREC: The Weak Reading of Donkey Sentences in Dynamic Logics	78
L. KVASZ: Degrees of Inconsistency	95
J. MÁCHA: The Idea of Code in Contextualism and Minimalism	116
P. MATERNA: The Concept of Problem from the Viewpoint of the Theory of Constructions	137
M. NEŠTINA: Textualism and Legal Positivism in the Philosophy of Language	145
J. PEREGRIN – V. SVOBODA: Which Formula is the Right One? (Criteria of Adequacy of Logical Analysis)	163
M. POLÁK: Emotions: Raw and Cognitive Component	180
V. PUNČOCHÁŘ: Mathematical Realism and Penelope Maddy's Naturalism.	199
J. RAČLAVSKÝ: Contextualism vs. Minimalism and Methodological Principles	227
I. SEDLÁR: Infallibility and Logic	239
M. SCHMIDT: On the Realistic and Antirealistic Tendencies in the Object Ontology	255
M. TALIGA: Realism and the Principle of Empiricism	273
M. ZOUHAR: Two Kinds of Unarticulated Constituents	291

Obsah

L. BIELIK: Abduktívny model (vedeckého) vysvetlenia	5
M. DUŽI: Towards an Extensional Calculus of Hyperintensions	20
B. GEISTOVÁ ČAKOVSKÁ: Sémantický a pragmatický význam jazykových výrazov.	46
V. HAVLÍK: Metametodologie a naturalismus	56
M. KOSTEREC: Existenčné čítanie oslích viet v dynamických logikách	78
L. KVASZ: Stupne nekonzistentnosti	95
J. MÁCHA: The Idea of Code in Contextualism and Minimalism	116
P. MATERNA: Pojem problému z hlediska teorie konstrukcí	137
M. NEŠTINA: Textualizmus a právny pozitivizmus z hľadiska filozofie jazyka	145
J. PEREGRIN – V. SVOBODA: Která formule je ta pravá? (Kritéria adekvátnosti logické analýzy)	163
M. POLÁK: Emoce: primitivní a kognitivní složka	180
V. PUNČOCHÁŘ: Matematický realismus a naturalismus Penelope Maddy.	199
J. RAČLAVSKÝ: Kontextualismus vs. minimalismus a metodologické principy	227
I. SEDLÁR: Neomylnosť a logika	239
M. SCHMIDT: O realistických a antirealistických tendenciách v ontológii predmetov	255
M. TALIGA: Realizmus a princíp empirizmu	273
M. ZOUHAR: Dva druhy neartikulovaných zložiek	291

Preface

Česko-slovenské sympóziium k analytickej filozofii je tradičným vrcholom spolupráce medzi českými a slovenskými analytickými filozofmi. Už dávno prestalo byť fórom, na ktorom sa stretávajú predovšetkým pracovníci Filozofického ústavu AV ČR v Prahe a Filozofického ústavu SAV v Bratislave; početné zastúpenie na sympóziu mávajú aj filozofické pracoviská z univerzít v Prahe, Bratislave, Brne, Ostrave, Plzni, Banskej Bystrici, Hradci Králové, Ružomberku atď.

15. ročník sympózia sa konal v dňoch 13. – 16. 9. 2011 vo Vrútkach pri Martine. Na rozdiel od predchádzajúcich ročníkov sme sa rozhodli podobu sympózia upraviť tak, aby viac zohľadňoval skutočnosť, že analytická filozofia v Čechách a na Slovensku je pomerne široký filozofický prúd, ktorý sa zaoberá množstvom rozmanitých tém a problémov z filozofickej logiky, filozofie jazyka, filozofie mysle, filozofie vedy a ďalších oblastí. Preto sympóziium nemalo stanovenú jednotiacu rámcovú tému, ale namiesto toho sa počas neho usporiadalo niekoľko menších workshopov:

- workshop Realizmus, antirealizmus a filozofia vedy (koordinátor: Eugen Zeleňák; ostatní účastníci: Lukáš Bielik, Vladimír Havlík, Martin Paleček, Martin Schmidt, Miloš Taliga)
- workshop Filozofia exaktných vied (koordinátor: Ladislav Kvasz; ostatní účastníci: Vladimír Balek, Róbert Maco)
- workshop Emócie a intencionalita (koordinátor: Juraj Hvorecký; ostatní účastníci: Michal Polák, Jana Rončáková)
- workshop Hyperintenzionálna sémantika (koordinátorka: Marie Duží; ostatní účastníci: František Gahér, Jakub Macek, Pavel Materna)
- workshop Sémantický minimalizmus a kontextualizmus (koordinátor: Marián Zouhar; ostatní účastníci: Barbora Geistová Čakovská, Jakub Mácha, Jiří Raclavský)

Tematické zameranie workshopov dobre ilustruje, ktoré oblasti výskumu sú v našej analytickej filozofii dominantné.

Niektorí ďalší účastníci sympózia vystúpili so svojimi príspevkami vo všeobecnej časti, mimo spomínaných workshopov (Jindřich Černý, Miloš Kosterec, Marek Neština, Jaroslav Peregrin, Marek Picha, Vít Punčochář, Mark Risjord, Igor Sedlár, Prokop Sousedík, David Svoboda, Vladimír Svoboda). Prezentovali témy patriace do logiky, filozofie matematiky, epistemológie atď.

Toto mimoriadne číslo časopisu *Organon F* obsahuje väčšinu príspevkov, ktoré na sympóziu odzneli. Usporiadané sú abecedne podľa mien ich autorov, nie podľa workshopov, do ktorých boli pôvodne zaradené. Hlavný dôvod

spočíva v tom, že nie všetky príspevky z workshopov sa objavujú aj v tomto mimoriadnom čísle, a teda čitateľ by nezískal plastický obrázok o tom, akú podobu workshopy mali.

– mZ –

Abduktívny model (vedeckého) vysvetlenia

Lukáš Bielik

Univerzita Komenského v Bratislave

Abstract: The paper presents an attempt to elucidate Abductive model of (scientific) explanation. Some of the positive resemblances between Hempel's Deductive-Nomological model and Abductive model of explanation are discussed. Nevertheless, it is claimed that the process of hypotheses selection and final result of explanatory seeking is represented by abductive form of explanation. Several logical, semantic and methodological criteria are proposed to make abductive inference explanatory. Such an abductive explanatory model is supposed to repair some of the relevant defects of the D-N model. Moreover, the abductive model is presented as a methodologically fruitful model of scientific explanation.

Key words: models of explanation, D-N model of explanation, criteria of adequacy, abductive model of explanation.

Presvedčenie, že popri spoločných črtách existujú medzi vysvetľovaním v bežnej mimovedeckej komunikácii a vysvetľovaním v rámci vedeckého diskurzu viaceré dôležité rozdiely, bolo vo filozofii vedy (predovšetkým) druhej polovice 20. storočia motiváciou k predloženiu viacerých návrhov, ktoré mali zachytiť nevyhnutné a postačujúce kritériá, určujúce, čo možno považovať za adekvátne a korektné vedecké vysvetlenie.¹ Každý súbor takýchto kritérií, spojený s ich metodologickým zdôvodnením, možno nazvať *modelom vedeckého vysvetlenia*.

Pokiaľ predmet vysvetlenia obmedzíme na kategóriu udalostí (resp. výskytu udalostí), partikulárnych javov či faktov, opísaných singular-

¹ Medzi charakteristické rozdiely vo vysvetľovaní a vysvetleniach v bežnej komunikácii a vo vedeckom diskurze možno zaradiť odlišné nároky, ktoré kladieme na faktory, prostredníctvom ktorých vysvetľujeme, ako aj ich odlišnú mieru všeobecnosti.

ny mi alebo molekulárnymi (nekvantifikovanými) výrokmi, môžeme medzi už tradičné modely vedeckého vysvetlenia zaradiť *deduktívno-nomologický model* (D-N model), *induktívno-štatistický model* (I-Š model), *model štatistickej relevancie* (Š-R model), *model kauzálnych mechanizmov* (model KM), *unifikačný model* (U-model) či *pragmatické modely* (P-modely) (pozri napríklad Salmon 1989, Woodward 2009 či Zeleňák 2008). Každý z uvedených modelov má svoje silné i slabé stránky. Môžeme ich identifikovať tak, že porovnáme, nakoľko typické príklady vedeckých vysvetlení z určitej vednej disciplíny vyhovujú podmienkam, ktoré kladie príslušný explanačný model na adekvátne vedecké vysvetlenie. Napríklad výhodou D-N modelu, I-Š modelu i U-modelu je skutočnosť, že vyjadrujú explicitne spôsob, akým je vedecké vysvetlenie určitého javu prepojené s príslušnou vedeckou teóriou; inak povedané, ukazujú, že vysvetlenie je jednou z funkcií teórie. Napriek tomu je možné uviesť k týmto modelom príklady, ktoré by sme štandardne nepovažovali za vysvetlenia, hoci spĺňajú kritériá príslušného modelu. Uvedené modely by teda bolo potrebné doplniť či modifikovať spôsobom, ktorý by im zaručil imúnosť voči kritickým protipríkladom.

V tomto svojom príspevku sa nebudem venovať pokusom o korigovanie niektorého zo spomenutých modelov. Namiesto toho predstavím „model“ či presnejšie inferenčnú schému známu ako *abdukcia*, ktorú po doplnení viacerých metodologických kritérií, zdá sa, môžeme zaradiť medzi modely vedeckého vysvetlenia, a to aj s vedomím, že medzi štandardne uvádzanými modelmi explanácie zvykne chýbať. Skôr než sa k tomu dostaneme, pozrime sa na prvý z kľúčových modelov explanácie – D-N model. Pomôže nám priblížiť niektoré stránky vysvetlenia, ktoré budeme môcť využiť pri predstavení abduktívneho modelu vysvetlenia.

1 Hempelov D-N model

Carl G. Hempel spolu s Paulom Oppenheimom predstavili v práci *Studies in the Logic of Explanation* z roku 1948 (ktorá bola opakovane zverejnená v Hempel 1965 ako Hempel – Oppenheim 1965a) model vedeckého vysvetlenia, ktorý sa vzápätí stal východiskom rozsiahlych diskusií o všeobecných kritériách adekvátnosti vedeckých vysvetlení a voči ktorému sa začali vymedzovať aj návrhy ďalších konkurenčných explanačných modelov. Sám Hempel svoje pôvodné vymedzenie D-N

modelu ďalej rozpracoval a pokúšal sa ho spolu s I-Š modelom obhájiť voči viacerým námietkam (pozri najmä Hempel 1965b).

Podľa Hempela každé vedecké vysvetlenie je možné chápať ako úsudok deduktívnej alebo induktívnej formy, ktorý musí spĺňať niekoľko podmienok. Ak ponecháme teraz bokom vysvetlenia induktívnej formy, môžeme si D-N model priblížiť nasledovne:

Nech výrok E opisuje konkrétny jav, výskyt udalosti určitého druhu alebo nejaký fakt, ktorý chceme vysvetliť. To, čo tento výrok opisuje či konštatuje, môžeme nazvať *predmetom* vysvetlenia. V rámci štruktúry vysvetlenia je tento predmet, resp. príslušný výrok, ktorý ho opisuje, nazývaný *explanandum* vysvetlenia (explanácie). Východiskom vysvetlenia je podľa Hempela situácia, keď sa pýtame, prečo sa stalo to, čo opisuje výrok explananda. Otázky formy „Prečo E ?“, kde E je výrok konštatujúci, že došlo k určitej udalosti, resp. že nastal taký a taký stav vecí, sú podľa Hempela typickými otázkami, ktorými sa dožadujeme vysvetlenia. Vedecké vysvetlenie toho, prečo nastalo E , spočíva podľa neho v tom, že nájdeme také výroky L_1, \dots, L_n , kde $n \geq 1$, ktoré vyjadrujú určitý zákon všeobecnej formy, a také výroky C_1, \dots, C_m , kde $m \geq 1$, ktoré opisujú tzv. počiatočné (resp. hraničné) podmienky udalosti, ktorú chceme vysvetliť, pričom z uvedených výrokov možno zostrojiť deduktívny argument formy:

$$(DN1) \quad \frac{L_1, \dots, L_n}{C_1, \dots, C_m} \\ E$$

Výroky L_1, \dots, L_n a C_1, \dots, C_m predstavujú tzv. *explanans* príslušného vysvetlenia, teda všetky tie informácie, ktoré sú potrebné na vysvetlenie explananda. Na to, aby bol nejaký úsudok formy (DN1) aj explanačným úsudkom, musí spĺňať tieto štyri kritériá (porovnaj Hempel 1965a, 247-249):

1. Explanandum musí byť logicky odvoditeľné z explanansu.
2. Explanans musí obsahovať všeobecné zákony, ktoré sú nevyhnutné na odvodenie explananda.
3. Explanans musí obsahovať výroky, ktoré sú v princípe testovateľné prostredníctvom pozorovania alebo experimentu.
4. Výroky explanansu musia byť pravdivé.

Pokiaľ sú splnené prvé tri podmienky, ktoré Hempel považuje za kritériá logického charakteru, môžeme podľa neho hovoriť o *poten-*

ciálnom D-N vysvetlení. Ak je navyše splnená aj posledná podmienka, máme k dispozícii „pravdivé“ adekvátne vysvetlenie.²

Keďže zo schémy (DN1) nemusí byť dostatočne zrejmé, v čom vlastne vysvetlenie nejakej udalosti spočíva, vyjadríme si explanačnú úsudkovú schému (DN1) výrazovými prostriedkami predikátovej logiky. Ak pre jednoduchosť pripustíme, že v danom explanačnom úsudku operujeme len s jedným výrokom formy zákona a jedným výrokom opisujúcim počiatočné podmienky, tak základnú schému D-N modelu možno vyjadriť takto:

(DN2)	$(\forall x)(P(x) \rightarrow Q(x))$...	výrok formy zákona
	$P(a)$...	výrok o počiatočných podmienkach
	$Q(a)$...	výrok o stave vecí, ktorý chceme vysvetliť

Vysvetlenie toho, prečo má objekt a vlastnosť Q , teda spočíva podľa D-N modelu v tom, že zostrojíme úsudok formy (DN2), ktorý vlastne „hovoriť“, že keďže všetky objekty s vlastnosťou P majú aj vlastnosť Q a keďže objekt a má (alebo mal) vlastnosť P , má preto aj vlastnosť Q . Udalosť či jav, ktorý chceme vysvetliť, je tak potrebné „zastrešiť“ relevantným zákonom, ktorého je explanandum prípadom. Hempel doslova píše:

D-N vysvetlenie teda odpovedá na otázku „Prečo nastal jav explananda?“ tak, že ukáže, že tento jav je výsledkom určitých okolností, špecifikovaných v C_1, \dots, C_m v súlade so zákonmi L_1, \dots, L_n . [...] úsudok demonštruje, že za predpokladu určitých okolností a predmetných zákonov, *bolo treba* výskyt javu očakávať; a práve v tomto zmysle nám vysvetlenie umožňuje *pochopiť, prečo* tento jav nastal. (Hempel 1965b, 337)

Aby sme si uviedli jednoduchý názorný príklad explanačného úsudku, ktorého logickú formu možno zachytiť schémou (DN2), pýtajme sa, prečo sa nejaký konkrétny kus kovu rozťahol. Úsudok, ktorý možno uviesť ako (rekonštruované) vedecké vysvetlenie predmetného faktu, bude vyzeráť takto:

² Hempel hovorí v práci Hempel (1965b, 338) doslova o *pravdivom vysvetlení*, čo môže pôsobiť neprírodzene, keďže úsudky – a explanačia je podľa neho úsudkom určitého druhu – nezvykneme kvalifikovať ako „pravdivé“ (či nepravdivé).

(DN3) Všetky kovy, ktoré sú zahrievané (na teplotu T), sa rozťahujú.

Tento kus kovu bol zahrievaný (na teplotu T).

Tento kus kovu sa rozťahal.

Inak povedané, tento kus kovu sa rozťahal, pretože každý kov, ktorý je zahrievaný (na teplotu T), sa rozťahne; a tento kov bol zahrievaný (na teplotu T).

Viacerí filozofi voči D-N modelu vzniesli námietky a protipríklady, ktoré mali preukázať jeho defekty. Niektoré z nich, ako napríklad údajné vysvetlenie faktu, že XY mužského pohlavia neotehotnel, odvolaním sa na to, že žiadny muž, ktorý užíva antikoncepciu, neotehotnie a muž XY užíval antikoncepciu, nemožno považovať za pravé protipríklady k D-N modelu. Dôvod je jednoduchý: V prípade tohto úsudku nebola 2. podmienka adekvátnosti D-N modelu splnená. Aj keď je totiž pravda, že uvedený príklad poukazuje na problém rozlíšenia všeobecných výrokov, ktoré vyjadrujú len náhodné generalizácie od všeobecných výrokov, ktoré vyjadrujú zákony, D-N model nemá byť zároveň modelom demarkácie výrokov prvej kategórie od výrokov druhej kategórie. Jednoducho len predpokladá, že vhodný zákon, hypotézu či teóriu máme k dispozícii.

Na druhej strane existujú konštrukcie takých úsudkov, ktoré naozaj spĺňajú všetky štyri podmienky adekvátnosti D-N modelu vysvetlenia, no napriek tomu ich nie sme ochotní akceptovať ako adekvátne vedecké vysvetlenia. Uvediem dva príklady. Prvý príklad má reprezentovať vysvetlenie, prečo má určitý stĺp dĺžku d . Zo zákonov optiky a výrokov, ktoré špecifikujú jednak dĺžku tieňa, ktoré stĺp vrhá, a jednak informácie o tom, že tento stĺp stojí kolmo k zemi a dopadá naň svetlo pod uhlom 45° , možno odvodiť výrok explananda, že daný stĺp má dĺžku d . Zdá sa však, že úsudok pozostávajúci z týchto výrokov, by sme neakceptovali ako vedecké vysvetlenie. Viacerí pritom diagnostikujú problém naznačený týmto príkladom v tom, že D-N model nerešpektuje určitú asymetriu, ktorá je prítomná v zákonoch, vyjadrených formou funkčnej závislosti alebo ekvivalencie medzi viacerými veličinami (parametrami/premennými). Zdá sa však, že dôvody zamietnutia uvedeného úsudku ako vysvetlenia môžu byť aj iné. Prirodzenejšia je azda interpretácia, že neexistuje vedecká teória, ktorá by opisovala a vysvetľovala to, prečo majú jednotlivé stĺpy takú a takú dĺžku. Jednoducho povedané, nie každé explanandum môže byť predmetom vedeckého vysvetlenia. Zdá sa totiž prirodzené predpokladať, že na to, aby sme

určitý stav vecí mohli vedecky vysvetliť, musí byť konceptualizovaný pojmami určitej vhodnej vedeckej teórie. V uvedenom prípade možno teda príslušnú vlastnosť stĺpu – jeho dĺžku (resp. výšku) vysvetliť (podobne, ako je to aj v prípade ďalších artefaktov) odvolaním sa na dôvody či funkcie, ktoré zohrávali svoju úlohu pri jeho konštruovaní.

Druhý, svojou povahou odlišný príklad, ktorý zmieňuje americký filozof Peter Achinstein (pozri Achinstein 2010, 152n.), vychádza z nasledujúcich predpokladov. Predstavme si, že máme k dispozícii úsudok:

- (DN4) Ján zjedol libru arzénu v čase t .
Každý, kto zje libru arzénu, zomrie do 24 hodín.
 Ján zomrel do 24 hodín od t .

Predpokladajme, že premisy rovnako ako aj záver tohto úsudku sú pravdivé. Predstavme si však, že Ján zomrie do 24 hodín od času t v dôsledku autonehody, ktorej sa stal obeťou a ktorá nemá žiadny relevantný vzťah k jeho požitiu arzénu. Je zjavné, že úsudok (DN4) nepredstavuje vysvetlenie Jánovej smrti, aj keď tento úsudok spĺňa všetky štyri kritériá D-N modelu. Achinstein tvrdí, že jeden z hlavných problémov D-N modelu spočíva v tom, že jeho kritériá 1 – 3 robia vzťah medzi explanandom a explanansom *apriórny*, čo však nie je pre vysvetlenie kontingentných udalostí akceptovateľné, pretože *a priori* nevieme rozhodnúť, ktorá z možných príčin vysvetľovaného javu bola v danom prípade naozaj aktívna.

V ostatnej časti svojho príspevku sa pokúsim ukázať, že práve to-muto aspektu irelevantnosti sa môžeme vyhnúť akceptáciou explanačnej stratégie, ktorú nazývam *abduktívny model vysvetlenia*. Tento model si zároveň ponecháva pozitívne aspekty pôvodného D-N modelu, okrem iného aj skutočnosť, že predmet vysvetlenia je potrebné prepojiť s vhodnou vedeckou teóriou či s jej inferenčnými dôsledkami.

2 Abduktívne usudzovanie

Prv, než sa pozrieme na formalizáciu abduktívnych úsudkov, uvediem niekoľko príkladov, ktoré nám pripraví cestu k ďalšej analýze.

Príklad 1: Predstavme si, že sa vraciame z práce domov a blížime sa k dverám nášho bytu. Predpokladajme ďalej, že sme ráno pri odchode do práce mali informáciu o tom, že aspoň jeden rodinný príslušník bude doma v čase nášho príchodu z práce. Ako sa tak k dverám bytu

blížime, badáme, že sú pootvorené. Prečo sú však pootvorené? Jednou z možností (hypotéz), ktoré by mohli „dobre“ vysvetliť to, čo vidíme, je jednoducho predpoklad, že rodinný príslušník, ktorý do bytu vošiel ako posledný, dostatočne neprivil dvere a tie sa pootvorili (čo si však on sám nevšimol). Tento predpoklad zrejme najlepšie vysvetľuje pootvorené dvere nášho bytu vzhľadom na celý súbor poznatkov, ktoré máme v danom čase k dispozícii, resp. je lepším vysvetlením než ostatné možnosti, ktoré pripadajú do úvahy, napríklad možnosť, že sa nám do bytu vlámal zlodej.

Príklad 2: Podobne ako v predchádzajúcom prípade aj teraz si predstavme, že sa vraciame z práce domov, ale nemáme informáciu o tom, či niekto z rodinných príslušníkov je doma. Rovnako aj v tomto prípade vidíme, že dvere nášho bytu sú pootvorené, ale navyše si všimneme, že dvere sú v blízkosti zámky poškodené. Keď by sme mali teraz vysloviť domnienku, prečo sú dvere pootvorené (a zároveň poškodené), najlepším vysvetlením by bolo azda presvedčenie, že niekto sa nám vlámal do bytu (a zrejme v ňom už nie je, keďže nechal dvere pootvorené). Táto singulárna hypotéza by totiž evidenciu, ktorú pozorujeme, vysvetľovala lepšie než napríklad predpoklad, že rodinný príslušník nedostatočne privrel dvere bytu.

To, čo oba príklady zachytávajú, je spôsob, akým zvykneme v mnohých situáciách usudzovať, keď chceme niečo vysvetliť. V oboch prípadoch disponujeme informáciami, ktoré sa nejakým spôsobom vzťahujú k predmetu vysvetlenia. V oboch prípadoch môžeme zvažovať niekoľko konkurenčných hypotéz, ktoré by nám mohli vysvetliť to, na čo sa pýtame. Nie každá však za daných okolností môže predmet vysvetlenia aj naozaj úspešne vysvetliť. Rozdiel v dodatočných relevantných informáciách o okolnostiach, ktorých sa vysvetlenie týka, nás môže viesť k tomu, že si spomedzi viacerých potenciálnych hypotéz vyberieme práve tú, ktorá najviac alebo najpriamočiarejšie súvisí s daným súborom informácií.

3 Formy abdukcie

Abdukcia býva spolu s indukciou (v tej či onej forme) zaraďovaná medzi kategóriu úsudkov (inferencií), v ktorých vzťah medzi premisami a záverom nie je nevyhnutný (pozri Douven 2011), pričom však premisy abduktívneho úsudku udeľujú záveru určitú (v ideálnom prípade vysokú) mieru podpory či pravdepodobnosti. Zároveň spolu s induk-

ciou predstavuje triedu úsudkov, ktoré sú nemonotónne (por. Aliseda 2006, resp. Meheus 2011). Neexistuje však názorová zhoda, či abdukcia je len špecifickým druhom indukcie, alebo či naopak indukcia je len druhom abdukcie, resp. či abdukcia a indukcia tvoria dve samostatné nededuktívne úsudkové kategórie.³ Napríklad Charles S. Peirce (Peirce 1992) vymedzuje abduktívnu inferenciu (ktorú nazýva „hypotézou“) ako samostatnú podkategóriu tzv. syntetických úsudkov, medzi ktoré tiež patrí rovnako samostatná podkategória „indukcie“. ⁴ (My sa prikláňame k Peircovej tradícii.) Ako pôvodný príklad abdukcie uvádza úsudok (pozri Peirce 1992, 188):

- (AB1) Pravidlo: Všetky fazuľky z tohto vrecúška sú biele.
 Výsledok: Tieto fazuľky sú biele.
-
- Prípad: Tieto fazuľky sú z tohto vrecúška.

Peirce pritom charakterizuje abdukciu (hypotézu) ako prípad úsudku, keď „nachádzame veľmi zvláštne okolnosti, ktoré by boli vysvetlené predpokladom, že nastal prípad určitého všeobecného pravidla, a tento predpoklad preto prijmem“ (Peirce 1992, 189). Neskôr Peirce uvádza modifikovanú schému abdukcie, v ktorej je vyjadrený „hypotetický“ charakter vzťahu medzi abdukovanou hypotézou a tým, čo má vysvetliť; ide o úsudok:⁵

- (AB2) Pozorujeme zvláštny fakt C.
 Ak by však A bola pravda, fakt C by sa stal zrejmy.
-
- Preto je rozumné sa domnievať, že A je pravda.

³ Základný rozdiel medzi induktívnymi typmi inferencie (napríklad enumeratívnu indukciu či analógiu) a abdukciou je možné identifikovať na úrovni ich logickej formy (pozri ďalej). Pokiaľ však ide o niektoré logické vlastnosti, ako napríklad (ne)monotónnosť úsudkov, rozdiely medzi abdukciou a indukciou sa zmazávajú. Na druhej strane treba povedať, že rozdielnosť týchto dvoch kategórií nededuktívnych úsudkov sa môže komplexnejšie ukázať v metodologických funkciách ich aplikácie, a to či už v rovine explanácie, resp. predikcie, alebo v rovine konštruovania hypotéz, apod.

⁴ Na druhej strane G. Harman (pozri Harman 1965) považuje induktívne úsudky za podkategóriu abdukcie, resp. za podkategóriu úsudku - odvodnenia najlepšieho vysvetlenia („the inference to the best explanation“).

⁵ Peircovu druhú schému abdukcie možno nájsť v Peirce (1931 - 1958, CP 5.189); citované podľa Niiniluoto (2004, 63).

Pokiaľ túto úsudkovú schému vyjadríme prostriedkami predikátovej logiky a vypustíme atribúty, ktoré Peirce v úsudku (AB2) uvádza, môžeme si najjednoduchšiu formu abdukcie vyjadriť aj takto:

(AB3)	$(\forall x)(P(x) \rightarrow Q(x))$... hypotéza, teória, zákon, pravidlo
	$Q(a)$... explanandum
	$P(a)$... singulárna hypotéza, vlastný explanans

Obe schémy (AB2) i (AB3) sú len zjednodušenou formou abdukcie, kde medzi premisami figuruje len jedna základná hypotéza či teória.⁶ Vypustená je aj kvalifikácia, že A , resp. $P(a)$ je vysoko pravdepodobné alebo pravdepodobnejšie než ostatné uvažované alternatívy.

Len pre úplnosť dodajme, že okrem schém, ktoré pracujú s výlučne univerzálnymi hypotézami, môžeme uvažovať aj o pravdepodobnostnej abdukcii, teda o takom abduktívnom úsudku, ktorý má len pravdepodobnostnú alebo štatistickú hypotézu (pozri Niiniluoto 2004, 62):

(AB4)	Pravdepodobnosť, že ľubovoľný objekt s vlastnosťou F má aj vlastnosť $G = r$.
	$G(b)$
	$F(b)$

V ostatnej časti kvôli zjednodušeniu obmedzíme svoju pozornosť na tie abduktívne úsudky, ktorých hypotézy považujeme za univerzálne. Zároveň sa pokúsime načrtnúť komplexnejšiu schému abduktívneho úsudku, ktorá po špecifikácii ďalších kritérií a podmienok môže modelovať a rekonštruovať spôsob, akým možno podať vedecké vysvetlenie partikulárnej udalosti či stavu vecí.

⁶ V tejto súvislosti dodajme, že sa niekedy rozlišuje medzi abdukciou a schémou inferencie k najlepšiemu vysvetleniu (inference to the best explanation). O abdukcii sa pritom predpokladá, že pracuje len s jednou hypotézou, zatiaľ čo inferencia k najlepšiemu vysvetleniu pracuje s viacerými alternatívnymi hypotézami. My budeme používať termín „abdukcia“ a „abduktívna inferencia“ v širokom význame pokrývajúcom aj schému inferencie k najlepšiemu vysvetleniu.

4 Abduktívne vysvetlenie

Prv než uvedieme príklad na konkrétny abduktívny explanačný úsudok, predstavím všeobecnú explanačnú schému abdukcie spolu s viacerými logickými, sémantickými a metodologickými kritériami, ktoré by mali zabezpečiť adekvátnosť abduktívnej explanácie.

Nech H_1, H_2, \dots, H_n sú hypotézy logickej formy $(\forall x)(P_i(x) \rightarrow Q(x))$. Nech A_1, A_2, \dots, A_n sú výroky o potenciálnych počiatočných podmienkach (resp. príčinách) javu, ktorý chceme vysvetliť, a ktorých logickú formu vyjadruje formula $P_i(a)$. Výrok E (*explanandum*) nech opisuje jav, ktorý chceme vysvetliť, pričom nech jeho logickú formu vyjadruje formula $Q(a)$. Napokon nech množina poznatkov K , ktoré sa vzťahujú k empirickým okolnostiam (kontextu) predmetu vysvetlenia, nám umožňuje okrem iného aj rozhodnúť, ktoré z výrokov A_1, A_2, \dots, A_n nie sú pravdivé, prípadne ktoré z okolností vyjadrené týmito výroky pravdepodobne nenastali. Potom môžeme priebeh a výsledok procesu abduktívneho vysvetlenia vyjadriť schémou:

$$\begin{array}{l}
 \text{(AB5)} \quad E \\
 (H_1 \wedge A_1) \rightarrow E \\
 (H_2 \wedge A_2) \rightarrow E \\
 \dots \\
 (H_n \wedge A_n) \rightarrow E \\
 K = (\neg A_2 \wedge \dots \wedge \neg A_n) \\
 \hline
 A_1
 \end{array}$$

Namieste je jedno krátke upozornenie: V literatúre sa o abduktívnej inferencii zvykne uvažovať aj ako o metóde „objavovania“ hypotéz či ako o metóde výberu najpravdepodobnejšej alebo „najlepšej“ hypotézy z množiny alternatívnych hypotéz. To znamená, že v závere takéhoto abduktívneho úsudku sa objavuje jedna (najpravdepodobnejšia) z uvažovaných hypotéz. Uvedený postup môžeme reprezentovať schémou:

$$\begin{array}{l}
 \text{(AB6)} \quad E \\
 (H_1 \wedge A_1) \rightarrow E \\
 (H_2 \wedge A_2) \rightarrow E \\
 \dots \\
 (H_n \wedge A_n) \rightarrow E \\
 K = (\neg A_2 \wedge \dots \wedge \neg A_n) \\
 (\text{prípadne aj: } A_1) \\
 \hline
 H_1
 \end{array}$$

Schéma (AB6) sa od schémy (AB5) líši nielen formou výroku, ktorý vystupuje v jej závere (v schéme (AB6) má záver úsudku formu $(\forall x)(P_1(x) \rightarrow Q(x))$), zatiaľ čo v závere schémy (AB5) vystupuje výrok formy $P_1(a)$, ale aj svojou metodologickou funkciou. Zatiaľ čo schéma (AB6) má reprezentovať výber jednej spomedzi viacerých možných hypotéz, ktorá má vzhľadom na určité východiskové informácie v danom kontexte status najpravdepodobnejšej hypotézy, schéma (AB5) má reprezentovať explanačný výber najpravdepodobnejších okolností alebo stavu vecí, ktoré viedli k predmetnej udalosti. Naša explanačná schéma (AB5) nehovorí v závere nič o pravdivosti alebo nepravdivosti hypotéz H_1, H_2, \dots, H_n ; pravdivosť daných (vedeckých) hypotéz, ktoré môžu vyjadrovať viaceré odlišné príčiny vysvetľovaného javu, už predsa predpokladáme. Inak povedané, pri explanácii vychádzame z predpokladu, že jav, ktorý chceme vysvetliť, mohli spôsobiť viaceré odlišné okolnosti, ktorých kauzálny import na daný jav vyjadrujú viaceré konkurenčné (a nezlučiteľné) hypotézy. Avšak informácie z množiny K nám umožňujú vybrať spomedzi výrokov o počiatkových podmienkach (príčinách) javu ten, ktorý opisuje jeho najpravdepodobnejšie príčiny.

Aby bola schéma (AB5) vhodným modelom vedyckého vysvetlenia, musí byť doplnená o viaceré podmienky. Aj keď nemusí ísť o ich definitívnu podobu, domnievam sa, že navrhnuté podmienky sa približujú ideálu abduktívneho modelu vedyckého vysvetlenia udalosti:

- 1A. Každá hypotéza H_i musí byť empiricky testovateľná. (Ak má ísť o vedycké vysvetlenie, tak každá hypotéza H_i musí byť vedycká, t. j. musí byť prvkom určitého vedyckého teoretického systému.)
- 2A. Každá z uvažovaných hypotéz H_i musí byť vyjadrená v pojmoch, v ktorých je konceptualizovaný aj predmet vysvetlenia E .
- 3A. Zo samej H_i nie je možné deduktívne odvodiť E .
- 4A. Zo samej H_i nie je možné deduktívne odvodiť $\neg A_i$.
- 5A. Z E nie je možné deduktívne odvodiť A_i .
- 6A. Záver abduktívneho explanačného úsudku má obsahovať spomedzi výrokov A_i ten výrok, ktorého pravdivosť považujeme vzhľadom na empirickú bázu poznatkov K a predmet vysvetlenia opísaný výrokom E za najpravdepodobnejšiu.⁷

⁷ Podmienky 3A - 5A som prevzal a náležite prispôbil z prác Meheus (2011, 95), resp. z Aliseda (2006, 74).

Deklarované podmienky si azda vyžadujú stručný komentár:

Podmienka 1A vyjadruje podobne ako v Hempelovom D-N modeli akési minimálne kritérium, ktoré kladieme na explanačné hypotézy. Vysvetlenie empirického javu netestovateľnou hypotézou neprichádza v prípade vedeckého vysvetlenia do úvahy. Podmienku 1A však možno v prípade vedeckého vysvetlenia špecifikovať v tom zmysle, že explanačné hypotézy musia byť dobre testovanými prvkami určitej vedeckej teórie, resp. tvrdeniami odvodenými z určitej vedeckej teórie. Navyše by sme podmienku 1A mohli sprísniť a požadovať len výber *kauzálnych* hypotéz. Modifikácie a doplnenia v naznačenom smere sú teda prípustné.

Podmienka 2A má sémantický charakter. Jej požiadavka je jednoduchá: Hypotézy a teoretické výroky, ktoré berieme do úvahy pri vysvetľovaní určitého javu, musia byť vyjadrené v pojmoch, v ktorých je vyjadrený aj predmet explanácie; resp. predmet explanácie musí byť vyjadrený v pojmoch, ktoré sú vyjadrené príslušnými hypotézami.⁸ Ak k danému predmetu vysvetlenia nemáme takú (vedeckú) hypotézu alebo teoretický výrok k dispozícii, tak predmet vysvetlenia nie je (na teraz) vedecky vysvetliteľný. Ako som už predtým naznačil, nie každý predmet vysvetlenia možno aktuálne vysvetliť vedecky, t. j. na základe určitej vedeckej systematizácie reprezentovanej typicky teóriou.

Podmienky 3A až 5A zabezpečujú práve to, že úsudok má abduktívnu formu. Navyše, podmienka 5A zabezpečuje elimináciu samovysvetľujúcich úsudkov, teda úsudkov, ktoré v závere explanačného úsudku (teda vo vlastnom explananse) uvádzajú výrok explananda.

Podmienka 6A je zrejme najviac otvorená možnostiam ďalšej špecifikácie a formálnejšieho vyjadrenia. V metodológii abduktívneho usudzovania (modelovaného skôr schémou (AB6) než našou explanačnou schémou) sa za najlepšiu spomedzi uvažovaných hypotéz (vysvetlení) štandardne označuje tá, ktorá má najvyššiu pravdepodobnosť vzhľadom na dostupnú empirickú evidenciu. Niektorí autori, napríklad Lipton (2004), navrhujú zasa preferovať tú z potenciálnych hypotéz (vysvetlení), ktorá „prináša najviac pochopenia“. Nazdávam sa však, že pokiaľ má vysvetlenie uviesť príčinu alebo určitý determinujúci faktor

⁸ Technickejšiemu vymedzeniu tejto podmienky sa tu nebudem venovať. Upozorňujem len, že spresnenie tejto myšlienky je možné uskutočniť na pozadí dostatočne expresívnej sémantickej teórie, v ktorej by bolo možné hovoriť o ekvivalencii pojmov, resp. významov výrazov.

predmetného javu, tak vzhľadom na poznatky, ktoré sa týkajú predmetu vysvetlenia, je potrebné preferovať tie singulárne predpoklady, ktoré sú najpriamejšie, najjednoduchšie a najviac pravdepodobné.

Ako prípad abduktívnej inferencie, ktorá vyhovuje našim explanačným podmienkam, uvediem príklad prevzatý z práce Meheus (2011, 96n.), ktorý možno po modifikácii a doplnení rekonštruovať do nami navrhovanej schémy explanačnej abdukcie. Predstavme si, že náš priateľ Ján má teplotu a na tvári i na ostatnom tele sa mu objavil červený výsyp – malé červené škvrnky. Aktuálne máme k dispozícii dve (medicínske) hypotézy, pomocou ktorých by sme vedeli vysvetliť túto skutočnosť. Podľa prvej hypotézy každý, kto má osýpky (rubeolu), má aj teplotu a po tvári a tele sa mu robia malé červené škvrnky. Podľa druhej hypotézy každý, kto má ružienku (rubellu), má teplotu a po tvári a tele sa mu robia malé červené škvrnky. Predpokladajme zároveň aj dve pomocné hypotézy: Prvá hovorí, že nik nemá osýpky viac než raz. Druhá zasa to, že nik nemá ružienku viac než raz. Napokon, ak máme informáciu, že Ján pred rokom prekonal osýpky (a že nie je proti ružienke zaočkovaný, ...), môžeme jeho súčasné zdravotné symptómy vysvetliť vysoko pravdepodobným predpokladom, že Ján má ružienku.

Na záver sa spätne pozrime na Hempelov model a porovnajme ho s nami navrhovaným modelom explanačnej abdukcie. Môžeme si zaiste všimnúť značnú príbuznosť oboch modelov v zdôrazňovaní explanačnej funkcie hypotéz, zákonov a vedeckých teórií, ktoré zabezpečujú usúvzťažnenie predmetu explanácie s určitými okolnosťami, ktoré vysvetľovaný jav determinujú. No zatiaľ čo Hempelov model nepracuje s možnosťou vysvetlenia nejakej udalosti či faktu viacerými konkurénčnými hypotézami a predpokladá, že kontextu vysvetlenia a jeho aktérom sú informácie o počiatočných podmienkach vysvetľovaného javu epistemicky dostupné, abduktívny model vedeckej explanácie ponecháva priestor rozsiahlejšej báze explanačných hypotéz a umožňuje pracovať s predpokladom, že počiatočné podmienky vysvetľovaného javu, ktoré sú opísané výrokom A_i v závere abduktívneho úsudku, nemusia byť aktérom vysvetlenia v danom kontexte epistemicky dostupné. Na druhej strane, pokiaľ máme k dispozícii evidenciu, ktorá nám umožňuje o výroku A_i v závere úsudku konštatovať, že je pravdivý, a pritom tiež vieme, že ostatné okolnosti opisované alternatívnymi výrokmi o počiatočných podmienkach nie sú pravdivé, môžeme dané abduktívne vysvetlenie transformovať na explanačný úsudok dedukčnej formy.

Napokon problému, na ktorý Achinstein upozornil v súvislosti s apriórnosťou vzťahu medzi explanandom a explanansom D-N modelu vysvetlenia, sa abduktívny model vysvetlenia vyhýba. Výber najpravdepodobnejšej príčiny (počiatočných podmienok) spomedzi možných príčin (alternatívnych počiatočných podmienok) vysvetľovaného javu sa nedeje len na základe relevantných hypotéz, ale aj (a predovšetkým) na základe dostupných empirických informácií, ktoré sa vzťahujú k empirickým okolnostiam, v ktorých sa predmet vysvetlenia nachádza. A aj keď si jednotlivé podmienky abduktívneho modelu zrejme budú vyžadovať ďalšie špecifikácie, zdá sa, že sme poukázali na viaceré podstatné rysy, vďaka ktorým možno považovať abduktívnu inferenciu za samostatný *model vedeckého vysvetlenia*.

Táto štúdia vznikla v rámci riešenia projektu VEGA-V-11-093-00.

Katedra logiky a metodológie vied
Filozofická fakulta
Univerzita Komenského v Bratislave
Šafárikovo nám. 6
814 99 Bratislava
Slovenská republika
bielik@fphil.uniba.sk

Literatúra

- ACHINSTEIN, P. (2010): *Evidence, Explanation, and Realism. Essays in the Philosophy of Science*. Oxford University Press.
- ACHINSTEIN, P. (2010a): Can There Be a Model of Explanation? In: Achinstein, P. (2010): *Evidence, Explanation, and Realism. Essays in the Philosophy of Science*. Oxford University Press, 143-167.
- ALISEDA, A. (2006): *Abductive Reasoning. Logical Investigations into Discovery and Explanation*. Springer.
- DOUVEN, I. (2011): Abduction. In: Zalta, E. N. (ed.): *Stanford Encyclopedia of Philosophy*. Dostupné: <http://plato.stanford.edu/entries/abduction/>.
- HARMAN, G. (1965): The Inference to the Best Explanation. *Philosophical Review* 74, 88-95.
- HEMPEL, C. G. (1965): *Aspects of Scientific Explanation and other Essays in the Philosophy of Science*. New York: The Free Press.
- HEMPEL, C. G. – OPPENHEIM, P. (1965a): Studies in the Logic of Explanation. In: Hempel, C. G. (1965): *Aspects of Scientific Explanation and other Essays in the Philosophy of Science*. New York: The Free Press, 245-290.

- HEMPEL, C.G. (1965b): Aspects of Scientific Explanation. In: Hempel, C. G. (1965): *Aspects of Scientific Explanation and other Essays in the Philosophy of Science*. New York: The Free Press, 331-496.
- LIPTON, P. (2004): *Inference to the Best Explanation*. London – New York: Routledge.
- MEHEUS, J. (2011): A Formal Logic for the Abduction of Singular Hypotheses. In: Dieks, D. et al. (eds.): *Explanation, Prediction, and Confirmation*. Springer, 93-108.
- NIINILUOTO, I. (2004): Truth-Seeking by Abduction. In: Stadler, F. (ed.): *Induction and Deduction in the Sciences*. Dordrecht – Boston – London: Kluwer Academic Publishers, 57-82.
- PEIRCE, Ch. S. (1931 – 1958): *Collected Papers of Charles Sanders Peirce*. Eds. Hartshorne, C. et al. Cambridge: Harvard University Press.
- PEIRCE, Ch. S. (1992): Deduction, Induction, and Hypothesis. In: Houser, N. – Kloesel, Ch. (eds.): *The Essential Peirce. Selected Philosophical Writings*. Vol. 1. Indiana University Press, 186-199.
- SALMON, W. S. (1989): *Four Decades of Scientific Explanation*. Minneapolis: University of Minnesota Press.
- WOODWARD, J. (2009): Scientific Explanation. In: Zalta, E. N. (ed.): *Stanford Encyclopedia of Philosophy*. Dostupné: <http://plato.stanford.edu/entries/scientific-explanation/>.
- ZELEŇÁK, E. (2008): *Moderné teórie vysvetlenia a príčinnosti*. Katolícka univerzita v Ružomberku, Filozofická fakulta.

Towards an Extensional Calculus of Hyperintensions

Marie Duží

VSB-Technical University, Ostrava

Abstract: In this paper I describe an extensional logic of hyperintensions, viz. Tichý's Transparent Intensional Logic (TIL). TIL preserves transparency and compositionality in all kinds of context, and validates quantifying into all contexts, including intensional and hyperintensional ones. The received view is that an intensional (let alone hyperintensional) context is one that fails to validate transparency, compositionality, and quantifying-in; and vice versa, if a context fails to validate these extensional principles, then the context is 'opaque', that is non-extensional. We steer clear of this circle by defining extensionality for hyperintensions presenting functions, functions (including possible-world intensions), and functional values. The main features of our logic are that the senses of expressions remain invariant across contexts and that our ramified type theory enables quantification over any logical objects of any order into any context. The syntax of TIL is the typed lambda calculus; its semantics is based on a procedural redefinition of, inter alia, functional abstraction and application. The only two non-standard features of our logic are a hyperintension called Trivialization and a four-place substitution function (called Sub) defined over hyperintensions. Using this logical machinery I propose rules of existential generalization and substitution of identicals into the three kinds of context.

Keywords. Quantifying-in, extensional/intensional/hyperintensional context, transparency, ramified type theory, transparent intensional logic, extensional rules for three kinds of context.

1 Introduction

In this paper I introduce basic fundamentals of an extensional logic of hyperintensions developed within *procedural semantics* of Transparent Intensional Logic (TIL). Only an extensional logic will validate ex-

tensional principles like the rule of existential generalization or Leibniz's law of substitution of identicals. The cornerstone of TIL approach is that we avail ourselves of rich ontology organised into an infinite bi-dimensional hierarchy of types. We assign to terms and expressions occurring in hyperintensional contexts the very same meaning that we assign to those very same terms and expressions when occurring in intensional and extensional contexts. As a result of this top-down approach, the extensional logical rules apply indiscriminately to all contexts. The upside of our top-down approach is that referential transparency and compositionality of meaning are preserved throughout, together with semantic innocence, since we have no recourse to reference shift. At no point do we invoke contextualist epicycles to somehow create a second semantics for 'non-extensional' contexts. The perceived downside would be that we revise the prevalent extensionalist semantic theory of terms and expressions, in that we universalize Frege's semantics earmarked for *Sinn*-sensitive contexts to all contexts. Be that as it may, it is strength of our solution that it is emphatically not tailor-made specifically for validating extensional principles. Instead it is just yet another application of a large-scale background theory and our solutions are principled and not *ad hoc*.

The paper is organised as follows. In Section 2 I briefly summarise the history of the development of logical semantics from Frege *via* so-called syntactic and model-theoretic turn up to procedural or algorithmic turn. Section 3 introduces the core of TIL so that to describe logical machinery needed in the main Section 4 where basic principles of extensional logic of hyperintensions are introduced.

2 Historical Background

Going back to the history we encounter Frege who was (to the best of my knowledge) the first who were developing formal semantics. In (1892) Frege introduced the well-known semantic schema assigning to expressions their *sense* (*Sinn*) and *denotation* (*Bedeutung*). Wishing to save compositionality, Frege made the semantics of an expression depend on the linguistic context in which it is embedded. According to Frege an expression names its *Bedeutung* (extension) in ordinary contexts and *Sinn* (intension) in oblique contexts.¹ Frege, in an attempt

¹ There is another defect in Frege's semantics; extension of a sentence is its truth-value. Yet in case of empirical sentences, the truth-value depends on

to save compositionality, had recourse to contextualism. The price he paid is too high, indeed. No expression can denote an object, unless a particular kind of context is provided.² Yet such a solution is far from being natural. There are cases of real ambiguity, witness homonymous expressions. But would anybody say that ‘The author of *Waverley*’ were another such a case of homonymy? Hardly; unless, of course, their intuitions had been warped by Fregean contextualism. Furthermore, expressions can be embedded within other expressions to various degrees; consider the sentence

“Charles knows that Tom believes that the author of *Waverley* is a poet.”

The expression ‘The author of *Waverley*’ should now denote the ‘normal’ sense of the ‘normal sense’ of itself. Adding still further layers of embedding sets off an infinite hierarchy of senses, which is to say that ‘The author of *Waverley*’ has the potential of being infinitely ambiguous. This seems plain wrong, and is first and foremost an awkward artefact of Fregean semantics (see also Duží – Materna 2010, or Duží – Jespersen – Materna 2010, §1.5).

The second half of the last century can be characterized as a linguistic turn in semantics. We were developing systems of particular logics which are characterized by a language with a precisely defined syntax and a model set-theoretic semantics. The main goal of building such a system is to find a subset of sentences of the language, axioms of the theory, which characterize a given area under scrutiny and apply proper rules of inference in order to mechanically derive consequences of the axioms. If the system has a model, then it is consistent, and all we are interested in is manipulating symbols. Therefore, *linguistic or syntactic turn*.

Says David Kaplan:

During the Golden Age of Pure Semantics we were developing a nice homogenous theory, with language, meanings, and entities of

contingent facts, which is not a matter of logical semantics. See also Klement (2002).

² It is also remarkable that Frege in his effort to save compositionality postulated that even expressions lacking a standard denotation would have a conventional one. Thus he actually broke the principle of compositionality. I am grateful for this remark to Marian Zouhar.

the world each properly segregated and related one to another in rather smooth and comfortable ways. This development probably came to its peak in *Carnap's Meaning and Necessity* (1947). Each *designator* has both an intension and an extension. Sentences have truth-values as extensions and propositions as intensions, predicates have classes as extensions and properties as intensions, terms have individuals as extensions and individual concepts as intensions The intension of a compound is a function of the intensions of the parts and similarly the extension (except when intensional operators appear). There is great beauty and power in this theory. But there remained some nagging doubts: proper names, demonstratives, and quantification into intensional contexts. (Kaplan 1990, 13-14)

The mainstream in this direction was *Possible World Semantics* (PWS). Kripke characterizes this semantics as follows:

We define a proposition (...) as a mapping whose domain is K [a logical space of possible worlds] and whose range is the set $\{T, F\}$. (Intuitively, a proposition is something that can be true or false in each world; and (...) we identify propositions that are strictly equivalent, i.e. have the same truth-value in each world. (...) Notice that each proposition determines a unique set of worlds (the set of all worlds mapped into T), and that conversely each set of worlds determines a proposition (its 'characteristic function'). Thus a proposition could just as well have been defined simply as a subset of K . (Kripke 1963, §5.3)

Possible-world intensions are extensionally individuated and the PWS semantics is a *logic of intensions*, in particular the *model-theoretic* (hence set-theoretic) *theory of modalities*. Yet its individuation of meaning is too crude (up to equivalence), and thus it is not apt to solve the notoriously well-known problem of the analysis of belief sentences. Carnap in (1947) says that modal sentences like "It is necessary that P " are intensional with respect to the clause P . However, sentences about belief like "John believes that P " are *neither intensional nor extensional* with respect to P . He also criticised Frege's 'naming method' (now we would say the denotational semantics), because then we multiply the names *ad infinitum*, and we end up with the antinomy of naming. For Carnap, extension is not a matter of *logical semantics* because it is a matter of factual knowledge. Prior for the meaning is an *intension* independent of contingent facts that uniquely determines the extension (if any), but not *vice versa*.

In order to solve the problem of belief sentences, Carnap tried to define a stronger relation between expressions than L-equivalence that might rightly calibrate the identity of meaning (i.e. *synonymy*). He defined inductively the relation of *intensional isomorphism* on the set of sentences. Roughly, two sentences S and P are intensionally isomorphic if they are L-equivalent and each designator (either simple or composed) that is a constituent of S is L-equivalent to the respective designator of P . Thus sentences S and P have the same intensional structure if they are composed *in the same way* from designators with the same intensions. In my opinion, all these Carnap's tenets and philosophical desiderata are plausible and it might seem that he succeeded in defining the subject of beliefs, knowledge, convictions, etc. Moreover, his definition is independent of the language and syntactic structure in which this subject is encoded. So far so good; yet Carnap's method has been criticized by Alonzo Church (1954). Church's argument is based on two principles. First, it is Carnap's principle of tolerance (which itself is, of course, desirable), and second, which is less desirable, this principle makes it possible to introduce into a language *syntactically simple* expressions as definitional abbreviations of *semantically complex* expressions (for instance, in English 'fortnight' standing for 'a period of fourteen days').

Thus we can introduce into a language *primitive* symbols P and Q in this way: P is the set of natural numbers that are less than the number 3. Q is the set of natural numbers n for which there are natural numbers x, y, z such that $x^n + y^n = z^n$. But then P and Q are L-equivalent (because they denote the same set of numbers) and also intensionally isomorphic because they have no other constituent designators but themselves. Yet it is easy to believe that $\exists n (Qn \wedge \neg Pn)$ without believing that $\exists n (Pn \wedge \neg Pn)$.³ Church proposes a *synonymous isomorphism*: all the mutually corresponding designators must be not only L-equivalent but also synonymous, where the synonymy of syntactically simple designators must be *postulated* as a semantic base of a language. We can postulate any convention for introducing these synonymous abbreviations, but as soon as we postulate the meaning of a constant it becomes valid and cannot be changed by another convention.

Since the late 60-s of the last century many logicians strived for *hyperintensional semantics* and *structured meanings* (see also Fox – Lappin

³ For the proof of Fermat's theorem is difficult to discover (written in 1954).

2001). The structured character of meaning was urged by David Lewis in (1972), where non-structured intensions are generated by finite, ordered trees. This idea of ‘tree-like’ meanings obviously influenced George Bealer’s idea of ‘intensions of the second kind’ in his (1982). The idea of structured meanings was propagated also by M.J. Cresswell who defines structured meanings as ordered n -tuples (see Cresswell 1975; 1985). That this is far from being a satisfactory solution is shown in Tichý (1994), Jespersen (2003) and also Bealer (2004). In brief, these tuples are again set-theoretic entities structured at most from a mereological point of view, by having elements or parts (though one balks at calling elements ‘parts’, since sets, including tuples, are not complexes). Besides, tuples are of the wrong making to serve as truth-bearers and objects of attitudes, since a tuple cannot be true or be known, hoped, etc., to be true. Simply, tuples are ‘flat’ from the procedural or algorithmic point of view. The *way* of combining particular parts together is missing here.

In 1994 Moschovakis comes with an idea of *meaning as algorithm*. The meaning of a term A is “an (abstract, not necessarily implementable) algorithm which computes the denotation of A ” (2006, 27; see also 1994).⁴ Moschovakis outlines his conception thus:

The starting point ... [is] the insight that a correct understanding of programming languages should explain the relation between a program and the algorithm it expresses, so that the basic interpretation scheme for a programming language is of the form

$$\text{program } P \rightarrow \text{algorithm}(P) \rightarrow \text{den}(P).$$

It is not hard to work out the mathematical theory of a suitably abstract notion of algorithm which makes this work; and once this is done, then it is hard to miss the similarity of the above schema with the basic Fregean scheme for the interpretation of a natural language,

$$\text{term } A \rightarrow \text{meaning}(A) \rightarrow \text{den}(A).$$

This suggested at least a formal analogy between algorithms and meanings which seemed worth investigating, and proved after some work to be more than formal: when we view natural language with a programmer’s eye, it seems almost obvious that we can represent the meaning of

⁴ Moschovakis’ notion of algorithm borders on being too permissive, since algorithms are normally understood to be effective. (See Cleland 2002 for discussion.)

a term A by the algorithm which is expressed by A and which computes its denotation. (Moschovakis 2006, 42)

Yet much earlier, in (1968) and (1969) Pavel Tichý formulated the idea of *procedural semantics*. Thus, for instance, a sentence encodes an *instruction* how in any possible world at any time to execute the abstract *procedure* expressed by the sentence as its meaning, i.e., to evaluate the truth-conditions of the sentence. He developed a logical framework known today as *Transparent Intensional Logic* (TIL). In modern jargon, TIL belongs to the paradigm of *structured meaning*. However, Tichý does not reduce structure to set-theoretic sequences, as do Kaplan and Cresswell. Nor does Tichý fail to explain how the sense of a molecular term is determined by the senses of its atoms and their syntactic arrangement, as Moschovakis objects to ‘structural’ approaches in (2006, 27).

3 Foundations of TIL

From the formal point of view, TIL is a hyperintensional, partial typed λ -calculus.⁵ A main feature of the λ -calculus is its ability to systematically distinguish between functions and functional values. An additional feature of TIL is its ability to systematically distinguish between functions and modes of presentation of functions and modes of presentation of functional values. The TIL operation known as *Closure* is the very *procedure* of presenting or forming or obtaining or *constructing* a function; the TIL operation known as *Composition* is the very *procedure* of *constructing* the value (if any) of a function at an argument. Compositions and Closures are both multiple-step procedures, or *constructions*, that operate on input provided by two one-step constructions, which figure as sub-procedures of Compositions and Closures, namely *variables* and so-called *Trivializations*. Characters such as ‘ x ’, ‘ y ’ ‘ z ’ are words denoting variables, which construct the respective values that an assignment function has accorded to them. The linguistic counterpart of a Trivialization is a constant term always picking out the same object. In order to operate on X , X needs to be grabbed first. Trivialization is such a one-step grabbing mechanism. Similarly, in order to talk about China (in non-demonstrative and non-indexical English discourse) we need to name China, most simply by using the constant

⁵ For details on TIL see in particular Tichý (1988, 2004) and Duží et al. (2010).

‘China’. Trivialization is important in what follows, because in order to substitute one sub-construction for another inside a construction it is crucial to be able to grab those three individual constructions.

The logical core of TIL is its notion of *construction* and its *type hierarchy*, which divides into a ramified type theory and a simple type theory. The ramified type hierarchy organizes all higher-order objects, which are all constructions, as well as all functions with domain or range in constructions. The simple type hierarchy organizes first-order objects, which are non-constructions like extensions (individuals, numbers, sets, etc.), possible-world intensions (functions from possible worlds) and their arguments and values. The relevant definitions decompose into three parts. Firstly, simple types of order 1 are defined by Definition 1. Secondly, constructions of order n , and thirdly, types of order $n + 1$.

Definition 1 (types of order 1) Let B be a base, where a base is a collection of pair-wise disjoint, non-empty sets. Then:

- (i) Every member of B is an elementary *type of order 1 over B* .
- (ii) Let $\alpha, \beta_1, \dots, \beta_m$ ($m > 0$) be types of order 1 over B . Then the collection $(\alpha \beta_1 \dots \beta_m)$ of all m -ary partial mappings from $\beta_1 \times \dots \times \beta_m$ into α is a functional *type of order 1 over B* .
- (iii) Nothing is a *type of order 1 over B* unless it so follows from (i) and (ii).

Remark. For the purposes of natural-language analysis, we are currently assuming the following base of ground types, which is part of the ontological commitments of TIL:

- \circ : the set of truth-values $\{\mathbf{T}, \mathbf{F}\}$;
- ι : the set of individuals (constant universe of discourse);
- τ : the set of real numbers (doubling as temporal continuum);
- ω : the set of logically possible worlds (logical space).

Definition 2 (construction)

- (i) The *variable x* is a *construction* that constructs an object O of the respective type dependently on a valuation v : x v -constructs O .
- (ii) *Trivialization*: Where X is an object whatsoever (an extension, an intension or a *construction*), 0X is the *construction Trivialization*. It constructs X without any change.
- (iii) The *Composition* $[X Y_1 \dots Y_m]$ is the following *construction*. If X v -constructs a function f of type $(\alpha \beta_1 \dots \beta_m)$, and Y_1, \dots, Y_m v -construct entities B_1, \dots, B_m of types β_1, \dots, β_m , respectively, then the *Composition* $[X Y_1 \dots Y_m]$ v -constructs the value (an entity, if any,

of type α) of f on the tuple argument $\langle B_1, \dots, B_m \rangle$. Otherwise the *Composition* $[X Y_1 \dots Y_m]$ does not v -construct anything and so is *v -improper*.

- (iv) The *Closure* $[\lambda x_1 \dots x_m Y]$ is the following *construction*. Let x_1, x_2, \dots, x_m be pair-wise distinct variables v -constructing entities of types β_1, \dots, β_m and Y a construction v -constructing an α -entity. Then $[\lambda x_1 \dots x_m Y]$ is the *construction* λ -*Closure* (or *Closure*). It v -constructs the following function f of the type $(\alpha \beta_1 \dots \beta_m)$. Let $v(B_1/x_1, \dots, B_m/x_m)$ be a valuation identical with v at least up to assigning objects $B_1/\beta_1, \dots, B_m/\beta_m$ to variables x_1, \dots, x_m . If Y is $v(B_1/x_1, \dots, B_m/x_m)$ -improper (see iii), then f is undefined on $\langle B_1, \dots, B_m \rangle$. Otherwise the value of f on $\langle B_1, \dots, B_m \rangle$ is the α -entity $v(B_1/x_1, \dots, B_m/x_m)$ -constructed by Y .
- (v) The *Single Execution* 1X is the *construction* that either v -constructs the entity v -constructed by X or, if X is not itself a construction or X is v -improper, 1X is v -improper.
- (vi) The *Double Execution* 2X is the following *construction*. Where X is any entity, the *Double Execution* 2X is v -improper (yielding nothing relative to v) if X is not itself a construction, or if X does not v -construct a construction, or if X v -constructs a v -improper construction. Otherwise, let X v -construct a construction Y and Y v -construct an entity Z : then 2X v -constructs Z .
- (vii) Nothing is a *construction*, unless it so follows from (i) through (vi).

Definition 3 (*ramified hierarchy of types*)

T_1 (*types of order 1*). See Definition 1.

C_n (*constructions of order n*)

- (i) Let x be a variable ranging over a type of order n . Then x is a *construction of order n over B* .
- (ii) Let X be a member of a type of order n . Then ${}^0X, {}^1X, {}^2X$ are *constructions of order n over B* .
- (iii) Let X, X_1, \dots, X_m ($m > 0$) be constructions of order n over B . Then $[X X_1 \dots X_m]$ is a *construction of order n over B* .
- (iv) Let x_1, \dots, x_m, X ($m > 0$) be constructions of order n over B . Then $[\lambda x_1 \dots x_m X]$ is a *construction of order n over B* .
- (v) Nothing is a *construction of order n over B* unless it so follows from C_n (i)-(iv).

T_{n+1} (*types of order $n + 1$*) Let $*_n$ be the collection of all constructions of order n over B . Then

- (i) $*$ and every type of order n are *types of order $n + 1$* .
- (ii) If $0 < m$ and $\alpha, \beta_1, \dots, \beta_m$ are types of order $n + 1$ over B , then $(\alpha \beta_1 \dots \beta_m)$ (see T₁ ii) is a *type of order $n + 1$ over B* .
- (iii) Nothing is a *type of order $n + 1$ over B* unless it so follows from (i) and (ii).

Empirical languages incorporate an element of *contingency* that non-empirical ones lack. Empirical expressions denote *empirical conditions* that may, or may not, be satisfied at some empirical index of evaluation. Non-empirical languages have no need for an additional category of expressions for empirical conditions. We model these empirical conditions as *possible-world intensions*. Intensions are entities of type $(\beta\omega)$: mappings from possible worlds to an arbitrary type β . The type β is frequently the type of the *chronology* of α -objects, i.e., a mapping of type $(\alpha\tau)$. Thus α -intensions are frequently functions of type $((\alpha\tau)\omega)$, abbreviated as ' $\alpha_{\tau\omega}$ '. I will typically say that an index of evaluation is a world/time pair $\langle w, t \rangle$. *Extensional entities* are entities of a type α where $\alpha \neq (\beta\omega)$ for any type β .

Examples of frequently used intensions are: *propositions* of type $o_{\tau\omega}$ (denoted by non-indexical sentences), *properties* of individuals of type $(o\iota)_{\tau\omega}$ (denoted by predicates or nouns like 'being happy', 'being a mathematician'), binary *relations-in-intension* between individuals of type $(o\iota\iota)_{\tau\omega}$ (usually denoted by verbs like 'admire', 'kick'), individual *offices* or *roles* of type $\iota_{\tau\omega}$ (denoted by definite descriptions like 'pope', 'the president of CR', 'the first man to run 100 m under 9 s'), binary *relations-in-intension* between individuals and hyperintensions of type $(o\iota*_n)_{\tau\omega}$ (denoted by attitudinal verbs like 'believe', 'know', 'calculate', etc.).

The method of explicit intensionalization and temporalization encodes constructions of possible-world intensions directly in the logical syntax.⁶ Where w ranges over ω and t over τ , the following logical form essentially characterizes the logical syntax of empirical language:

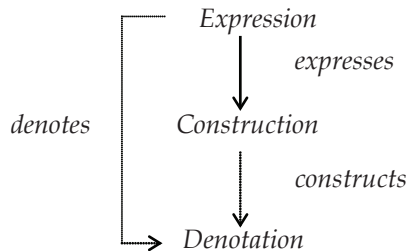
⁶ This is one of the issues in which TIL is superior compared to Montague's Intensional Logic (IL). Montague applies only implicit intensionalization; due to the lack of variables ranging over possible worlds (and times) IL does not validate the Church-Rosser 'diamond' property. Moreover, λ -abstraction, modalities and times must be imitated by special operators. As a result the law of universal instantiation, lambda conversion (β -rule) and Leibniz's Law do not generally hold in IL. For a critical comparison of TIL and Montague's IL see Duží et al. (2010, § 2.4.3).

$\lambda w \lambda t [\dots w \dots t \dots]$. If the Composition $[\dots w \dots t \dots]$ v -constructs an entity of type α , then the Closure itself constructs a function of type $((\alpha\tau)\omega)$, or $\alpha_{\tau\omega}$ for short, i.e. an α -intension.

Logical objects like *truth-functions* and *quantifiers* are extensional: \wedge (conjunction), \vee (disjunction) and \supset (implication) are of type (ooo) , and \neg (negation) of type (oo) . *Quantifiers* \forall^α , \exists^α are type-theoretically polymorphous total functions of type $(\text{o}(\text{o}\alpha))$, for an arbitrary type α , defined as follows. The *universal quantifier* \forall^α is a function that associates a class A of α -elements with **T** if A contains all elements of the type α , otherwise with **F**. The *existential quantifier* \exists^α is a function that associates a class A of α -elements with **T** if A is a non-empty class, otherwise with **F**.

Notational conventions: Below all type indications will be provided outside the formulae in order not to clutter the notation. Furthermore, ' X/α ' means that an object X is (a member) of type α . ' $X \rightarrow_v \alpha$ ' means that the type of the object v -constructed by X is α . We write ' $X \rightarrow \alpha$ ' if what is v -constructed does not depend on a valuation v . Throughout, it holds that the variables $w \rightarrow_v \omega$ and $t \rightarrow_v \tau$. If $C \rightarrow_v \alpha_{\tau\omega}$ then the frequently used Composition $[[C w] t]$, which is the extensionalization of the α -intension v -constructed by C , will be encoded as ' C_{wt} '. When applying truth-functions, identities $=^\alpha / (\text{o}\alpha\alpha)$, arithmetic operations and relations $>^\alpha, <^\alpha$, we will often use an infix notation without Trivialization and without indicating the type of a function. Instead of ' $[\text{}^0\exists^\alpha \lambda x B]$ ', ' $[\text{}^0\forall^\alpha \lambda x B]$ ' ($x \rightarrow_v \alpha; B \rightarrow_v \text{o}$) we will often write ' $\exists x B$ ', ' $\forall x B$ '. Thus, for instance, $[\text{}^0\forall^\tau \lambda x [\text{}^0\supset [\text{}^0=^\tau x \text{}^0\text{o}]] [\text{}^0\forall^\tau \lambda y [\text{}^0=^\tau [\text{}^0+ x y] y]]]$ will be encoded as $\forall x [[x = \text{}^0\text{o}] \supset \forall y [[x + y] = y]]$.

To summarize, our neo-Fregean semantic schema, which applies to all contexts, is this:



The most important relation in this schema is between an expression and its meaning (a construction). We can investigate *a priori* what (if anything) a construction constructs and what is entailed by it. Once a

construction is explicitly given as a result of logical analysis, the entity (if any) it constructs is already implicitly given, whereas it requires inquiry *a posteriori* to establish the reference of an empirical term at a given world/time pair. As a limiting case, the logical analysis may reveal that the construction fails to construct anything because it is improper. And if the construction is not improper, the denotation can be either a first-order object (i.e. a non-construction) or a lower-order construction. Intensional constructions (constructions of objects of type $(\beta\omega)$) are always proper, since they always construct an intension (including degenerate ones, which return no values at all or always the same value). In linguistic terms, every word whose sense is an intensional construction has a denotation, but will lack a reference at some or all $\langle w, t \rangle$ pairs, in case its denotation (a partial function) fails to return a value. This applies to, *inter alia*, ‘The pope’, ‘The first man to run 100 m under 9 s’, ‘The Evening Star’, or ‘John’s wife’.

Definition 4 (free and bound variables) Let C be a construction with at least one occurrence of a variable ξ .

- (i) Let C be ξ . Then the occurrence of ξ in C is *free*.
- (ii) Let C be X . Then every occurrence of ξ in C is ${}^0\text{bound}$ (‘Trivialization-bound’).
- (iii) Let C be $[\lambda x_1 \dots x_n Y]$. Any occurrence of ξ in Y that is one of x_i , $1 \leq i \leq n$, is λ -bound in C unless it is ${}^0\text{bound}$ in Y . Any occurrence of ξ in Y that is neither ${}^0\text{bound}$ nor λ -bound in Y is *free* in C .
- (iv) Let C be $[X X_1 \dots X_n]$. Any occurrence of ξ that is *free*, ${}^0\text{bound}$, λ -bound in one of X, X_1, \dots, X_n is, respectively, *free*, ${}^0\text{bound}$, λ -bound in C .
- (v) Let C be 1X . Then any occurrence of ξ that is *free*, ${}^0\text{bound}$, λ -bound in X is, respectively, *free*, ${}^0\text{bound}$, λ -bound in C .
- (vi) Let C be 2X . Then any occurrence of ξ that is *free*, λ -bound in a constituent of C is, respectively, *free*, λ -bound in C . If an occurrence of ξ is ${}^0\text{bound}$ in a constituent 0D of C and this occurrence of D is a constituent of X' v -constructed by X , then if the occurrence of ξ is *free*, λ -bound in D it is *free*, λ -bound in C . Otherwise, any other occurrence of ξ in C is ${}^0\text{bound}$ in C .
- (vii) An occurrence of ξ is *free*, λ -bound, ${}^0\text{bound}$ in C only due to (i)-(vi).

A construction with at least one occurrence of a free variable is an *open construction*. A construction without any free variables is a *closed construction*.

Definition 5 (*v*-congruent and equivalent constructions) Let $C, D / *_{\alpha} \rightarrow \alpha$ be constructions, and $\approx_v / (o *_{\alpha} *_{\alpha}), \approx / (o *_{\alpha} *_{\alpha})$ binary relations between constructions of order n . Using infix notation without Trivialization, $C \approx_v D, C \approx D$, we define: C, D are *v*-congruent, $C \approx_v D$, iff either C and D *v*-construct the same α -entity, or both C and D are *v*-improper; C, D are *equivalent*, $C \approx D$, iff C, D are *v*-congruent for all valuations v .

Corollaries. If C, D are *identical*, then C, D are *equivalent*, but not *vice versa*. If C, D are *equivalent*, then C, D are *v*-congruent, but not *vice versa*.

If meanings of expressions E_1, E_2 , that is the constructions expressed by them, are merely *v*-congruent, we will say that E_1, E_2 are *co-referential*. If meanings of expressions E_1, E_2 are *equivalent*, we will say that E_1, E_2 are *co-denotational* or *equivalent*.

The next notion we need to define is that of *synonymy*. Our notion of synonymy is defined in terms of *procedural isomorphism*. The term ‘procedural isomorphism’ is a nod to Carnap’s *intensional isomorphism* and Church’s *synonymous isomorphism*. Church’s Alternatives (0) and (1) leave room for additional Alternatives.⁷ One would be Alternative (1/2), another Alternative (3/4). The former includes α - and η -conversion while the latter adds a restricted β -conversion. If we must choose, we would prefer Alternative (3/4) to soak up those differences between β -transformations that concern only λ -bound variables and thus (at least appear to) lack natural-language counterparts.

One reason for excluding unrestricted β -conversion is the well-known fact that β -conversion is not an equivalent transformation in logics boasting *partial functions*, such as TIL. Another reason is that occasionally even β -equivalent constructions have different natural-language counterparts; witness the difference between attitude reports *de dicto* vs. *de re*. Thus for instance, if a, b are individuals, the difference between “ a believes that b is happy” and “ b is believed by a to be happy” is just the difference between β -equivalent meanings. If attitudes are construed as in possible-world semantics, i.e. as relations to intensions (rather than to hyperintensions), the former (*de dicto*) receives the analysis

$$\lambda w \lambda t [{}^0 \text{Believe}_{wt} {}^0 a \lambda w \lambda t [{}^0 \text{Happy}_{wt} {}^0 b]]$$

⁷ For Church’s Alternatives see Anderson (1998).

while the latter (*de re*) receives the analysis

$$\lambda w \lambda t [\lambda x [{}^0\text{Believe}_{wt} {}^0a \lambda w \lambda t [{}^0\text{Happy}_{wt} x]] {}^0b]$$

Types: *Happy* / ($\text{ot}_{\tau\text{ot}}$); $x \rightarrow_v \iota$; $a, b / \iota$; *Believe* / ($\text{ot}_{\text{ot}} \text{ot}_{\text{ot}}$).

The *de dicto* variant is the β -equivalent contractum of the *de re* variant. Both variants are equivalent because they construct one and the same proposition, the two sentences denoting the same proposition. Yet they denote this proposition in *different ways*, thus they are not synonymous. The equivalent β -reduction leads here to a *loss of analytic* information, namely loss of information about *which* of the two ways, or constructions, has been used to construct this proposition.⁸ In this case the loss seems to be harmless, though, because there is only one, unambiguous way to β -expand the *de dicto* version into its equivalent *de re* variant.⁹

However, unrestricted equivalent β -reduction sometimes yields a loss of analytic information that cannot be restored by β -expansion. The well-known example is the sentence “John loves his wife and so does Tom”. This sentence is ambiguous between two readings, sloppy and strict. On its sloppy reading John and Tom share the property of each loving their own wife (and both are exemplary husbands). On the strict reading they share the property of loving John’s wife (and there are troubles on the horizon). And these are two distinct properties. Thus there are two distinct analyses of “John loves his wife”:

$$\textit{Strict: } \lambda w \lambda t [\lambda x [{}^0\text{Love}_{wt} x [{}^0\text{Wife_of}_{wt} {}^0\text{John}] {}^0\text{John}]]$$

$$\textit{Sloppy: } \lambda w \lambda t [\lambda x [{}^0\text{Love}_{wt} x [{}^0\text{Wife_of}_{wt} x] {}^0\text{John}]]$$

But an unrestricted β -reduction turns these two redexes into one and the same contractum:

$$\lambda w \lambda t [{}^0\text{Love}_{wt} {}^0\text{John} [{}^0\text{Wife_of}_{wt} {}^0\text{John}]]$$

⁸ For the notion of analytic information, see Duží (2010) and Duží et al. (2010, §5.4).

⁹ In general, *de dicto* and *de re* attitudes are not equivalent, but logically independent. Consider “*a* believes that the pope is not the pope” and “*a* believes of the pope that he is not the pope”. The former, *de dicto*, variant makes *a* deeply irrational and most likely is not a true attribution, while the latter, *de re*, attribution is perfectly reasonable and most likely the right one to make. In TIL the *de dicto* variant is not an equivalent β -contractum of the *de re* variant due to the partiality of the office *Pope* / ι_{ot} .

A piece of analytic information has been lost and using the contractum one does not know which property should be applied to Tom.¹⁰

The *restricted* version of *equivalent* β -conversion we have in mind consists in substituting free variables for λ -bound variables of the same type, and will be called β_r -conversion. For instance, we see little reason to differentiate semantically or logically between “ b is believed by a to be happy” and “ b has the property of being believed by a to be happy”.¹¹ The latter sentence expresses

$$\lambda w \lambda t [\lambda w' \lambda t' \lambda x [{}^0\text{Believe}_{wt'} a \lambda w \lambda t [{}^0\text{Happy}_{wt} x]]_{wt} b]$$

This is merely a β_r -expanded form of

$$\lambda w \lambda t [\lambda x [{}^0\text{Believe}_{wt} a \lambda w \lambda t [{}^0\text{Happy}_{wt} x]] b]$$

Thus we define:

Definition 6 (*procedurally isomorphic constructions: Alternative (3/4)*)

Let C, D be constructions. Then C, D are a -equivalent iff they differ at most by deploying different λ -bound variables. C, D are η -equivalent iff one arises from the other by η -reduction or η -expansion. C, D are β_r -equivalent iff one arises from the other by β_r -reduction or β_r -expansion. C, D are *procedurally isomorphic*, denoted ${}^0C \approx {}^0D', \approx / (o^*_n^*)$, iff there are closed constructions $C_1, \dots, C_m, m \geq 1$, such that ${}^0C = {}^0C_1, {}^0D = {}^0C_m$ and all C_i, C_{i+1} ($1 \leq i < m$) are either α -, η - or β_r -equivalent.

Hence we advocate for the restricted β -conversion; yet β -conversion is the fundamental rule for computing the value of the function v -constructed by $[\lambda x Y]$ at an argument v -constructed by a construction C . Its (unrestricted) version ‘by name’ is this (where $Y(C/x)$ is the result of correct substitution of a construction C for x in Y):

$$[[\lambda x Y] C] \vdash Y(C/x)$$

¹⁰ For the solution of this problem see Duží – Jespersen (to appear).

¹¹ This is not to say we see no reason at all not to differentiate. For instance, it could be argued that one thing is to believe that a is happy and another is to believe that a has the property of being happy, because the latter at least appears to presuppose that the believer have the additional conceptual resources to master the notion of property. Thus a proper calibration of procedural isomorphism is still an open problem and it can depend on the area under scrutiny. More discussion on procedural isomorphism can be found in Jespersen (2010).

Due to compositionality, if C is v -improper the Composition $[[\lambda x Y] C]$ is v -improper as well. But if Y is itself a Closure then it is never v -improper.¹² Thus it may happen that the right-hand side is not equivalent to the left-hand side. For this reason we restrict the rule to C being a variable which is never v -improper.

But we do need a general rule of the λ -calculus for computing the value of a function. Fortunately, it turns out to be feasible to formulate a generally valid computational rule. A distinction familiar from programming languages based on the λ -calculus holds the key to the solution. The invalid rule above is moulded on the programming technique of calling a sub-procedure C by name: the sub-procedure itself is substituted for the ‘local variable’ x in the ‘procedure body’ Y . It is well-known among programmers that this technique can have undesirable side-effects, unlike the technique of calling a sub-procedure by value.¹³ The idea is simple: execute the sub-procedure C first, and then – *provided this execution does not fail* – substitute the construction of the result (‘pass by the value’) for x .¹⁴

The substitution method comes with two special functions.¹⁵ The polymorphous function Sub of type $(*_n *_n *_n *_n)$ operates on constructions as follows. When applied to constructions C_1, C_2, C_3 , Sub returns as its value the construction D that is the result of the correct (i.e. collisionless) substitution of C_1 for C_2 in C_3 . For instance, the result of the Composition $[{}^0Sub {}^{00}John {}^0x {}^{0[Wife_of_{wt} x]]}$ is the Composition $[{}^0Wife_of_{wt} {}^{00}John]$. The likewise polymorphous function Tr returns as its value the Trivialization of its argument. Thus the result of $[{}^0Tr {}^0John]$ is 0John . If the variable x ranges over ι , the Composition $[{}^0Tr x] v(John/x)$ -constructs 0John . Note one essential distinction between the function Tr and the construction Trivialization. Whereas the variable x is *free* in $[{}^0Tr x]$,

¹² See Definition 2, iii) and iv).

¹³ A recent reference to the distinction between ‘call by name’ and ‘call by value’ is Pierce (2002, 56-57). See also, for instance, Hyde (1996, Ch. 11) or Plotkin (1975).

¹⁴ For conversion by name, see Claim 2.5 and the subsequent proof in Duží et al. (2010, 267-268); for conversion by value, see Claim 2.6 and the subsequent proof in (ibid., 269-270). For the general strategy (inspired by programming languages) of distinguishing between succeeding, failing, and aborting with error, see also Van Eijck – Francez (1995).

¹⁵ Sub is introduced in Tichý (1988, 75) and Tr at (ibid., 68).

the Trivialization 0x binds the variable x by constructing just x independently of valuation.¹⁶

For simplicity's sake, we introduce the TIL translation of the rule of β -conversion by value in its simplified version for unary functions (generalization to n -ary functions is obvious):

$$[[\lambda x Y] C] \vdash {}^2[{}^0Sub [{}^0Tr C] {}^0x {}^0Y]$$

Note that the Composition on the right-hand side must undergo Double Execution. Provided C is v -proper, it v -constructs an entity, say e . Then the result of the first step (the substitution $[{}^0Sub [{}^0Tr C] {}^0x {}^0Y]$) is the construction $Y(e/x)$. The resulting construction must then be executed in order to obtain the value of the function v -constructed by $[\lambda x Y]$ at the argument e . Hence, *Double Execution*. Otherwise, if C is v -improper, the substitution fails to construct anything, because due to the compositionality constraint the whole Composition $[{}^0Sub [{}^0Tr C] {}^0x {}^0Y]$ is v -improper and so is ${}^2[{}^0Sub [{}^0Tr C] {}^0x {}^0Y]$ (see Definition 2, iii) and vi)). In this manner compositionality is preserved and the above rule of β -conversion by value is always valid even when C is v -improper.

Remark. In the project of a multi-agent system that our Laboratory of Intelligent Systems dealt with in 2004-2008 we use the computational variant of TIL, the *TIL-Script* functional programming language as the language of communication between agents.¹⁷ In the *TIL-Script* language we apply *only* this computational rule of *conversion by value*. The reason is this. Since the construction C can be v -improper, we need to implement a lazy evaluation mechanism in order to evaluate C only when needed. However, the properness of C can be checked only in the run time, because valuation v would supply values dependently on states-of-affairs.

¹⁶ Since TIL is a λ -calculus, all variable binding is λ -binding, except for Trivialization-binding. One area where Trivialization-binding plays a key role is existential quantification into hyperintensional contexts, where a quantifier is introduced with a view to binding a variable that occurs bound by Trivialization, because the variable occurs inside a Trivialized context. For discussion, see Duží et al. (2010, §5.3). For improved solutions, see Duží – Jespersen (to appear).

¹⁷ Project No. 1ET101940420: “Logic and Artificial Intelligence for multi-agent systems”; supported by the program “Information Society” of the Czech Academy of Sciences. For details see <http://labis.vsb.cz/>.

4 Rules for the Three Kinds of Context

At this point we have lined up everything we need in order to introduce the extensional logic of hyperintensions. Yet some may protest that extensional logic of hyperintensions sounds as an oxymoron like a roaring silence. For at least since the milestones Quine (1956) and Kaplan (1968) the validity of extensional principles, in particular of quantifying-in and existential generalization, has been fielded as a logical criterion for distinguishing

- a) extensional/transparent/'relational' contexts from
- b) non-extensional/opaque/'notional' contexts,

the idea being that extensional (etc.) contexts are those that validate quantifying-in.¹⁸ And conversely, if a context resists quantifying-in, it is caught in violation of one or more of the laws of extensional logic and as such eludes full logical analysis. What we are saying is that also intensional and hyperintensional contexts may be quantified into, but that the feasibility of doing so presupposes that it be done within an *extensional logic of hyperintensional contexts*. Deploying a non-extensional logic of hyperintensions to quantify into hyperintensional contexts would, indeed, be a non-starter, generating opacity and thereby making hyperintensional attitude contexts logically intractable. However, whether one accepts quantifying into (hyper-) intensional contexts or wants to restrict quantification to extensional contexts, like "Mary is happy", the logical question remains which contexts validate quantifying-in.

Tichý issues in (1986, 256; 2004, 654) a warning against a circular definition:

Q: When is a context extensional?

A: A context is extensional if it validates

- (i) *the rule of substitution of co-referential terms* and
- (ii) *the rule of existential generalization.*

Q: And when are (i), (ii) valid?

A: Those two rules are valid when applied to extensional contexts.

We steer clear of the circle by defining extensionality for

- a) *hyperintensions* presenting functions,

¹⁸ See Forbes (1996).

- b) *functions* (including possible-world intensions), and
 c) *functional values*.

These three levels are squared off with three kinds of context:¹⁹

1. *hyperintensional contexts*, in which a *construction* is not used to present an object, but is itself *mentioned* as functional argument (though a construction of one order higher needs to be used to mention this lower-order construction);
2. *intensional contexts*, in which a construction is *used* to present a *function* without presenting a particular value of the function (moreover, the construction does not occur within another hyperintensional context);
3. *extensional contexts*, in which a construction is *used* to produce a particular *value* of the function at a given argument (moreover, the construction does not occur within another intensional or hyperintensional context).

Example. Let the types of entities be: *Periodic* / ($o(\tau\tau)$); *Sin* / ($\tau\tau$); *Solve* / ($o(\iota^*)$) _{$\tau\omega$} ; the relation-in-intension between an individual and a construction the product of which the individual is seeking; π/τ ; *Tom* / ι .

- *Extensional context*: functional *value* is an object of predication (functional argument):

“*sin* $\pi = 0$ ”
 $[[{}^0\text{Sin } {}^0\pi] = {}^00]$

- *Intensional context*: function-in-extension is an object of predication:

“*Sine is a periodic function*”
 $[{}^0\text{Periodic } {}^0\text{Sin}]$

- *Hyperintensional context*: construction (“function-in-intension”) is an object of predication (a functional argument):

“*Tom is solving the equation sin x = 0*”
 $\lambda w \lambda t [{}^0\text{Solve}_{wt} {}^0\text{Tom } {}^0[\lambda x [{}^0\text{Sin } x] = {}^00]]$

A dual constraint of TIL has impact on the rules of inference. It is the constraint dictated by *properly partial functions*, which are undefined for some or all of their arguments, and *improper constructions*, which

¹⁹ For the definition see Duží et al. (2010, § 2.6 and 2.7).

fail to produce a product. Improperness stems from the procedure of applying a properly partial function f to an argument a , such that f returns no value at a . The procedure of functional application induces an extensional context. Thus when specifying the rules of quantifying-in, we must check whether particular constituent constructions occurring extensionally are improper. If none is, the particular rule of quantifying-in is valid.

The rules of *improperness* can be schematically summarized as follows. If a Composition is used in an extensional context as a procedure of application a properly partial function $F/(\beta\alpha)$ to an argument a/α and if F has no-value at a (value gap) then

$[{}^0F {}^0a]$ is *v-improper*

and so is any construction C occurring extensionally and containing $[{}^0F {}^0a]$ as a constituent; *partiality is strictly propagated up*:

$[\dots [\dots [{}^0F {}^0a] \dots] \dots]$ is *improper*

until the context is raised up to *hyper/intensional*:

intensional: $\lambda x \dots [\dots [\dots [{}^0F {}^0a] \dots] \dots]$ is *proper*
hyperintensional: ${}^0[\dots [\dots [{}^0F {}^0a] \dots] \dots]$ is *proper*

4.1 The Rules of Existential Generalization

Let $F/(\beta\alpha)$; a/α . Then the rules of existential generalization for particular contexts are as follows:

a) *extensional* context. Let $[\dots [{}^0F {}^0a] \dots]$ *v*-construct the truth-value **T**. Then the following rule is truth-preserving:

$[\dots [{}^0F {}^0a] \dots] \vdash \exists x [\dots [{}^0F x] \dots]; x \rightarrow_v \alpha$

Proof:

1. $[\dots [{}^0F {}^0a] \dots]$ assumption
2. $[\dots [{}^0F x] \dots]$ *v*(a/x)-constructs **T**
3. $\lambda x [\dots [{}^0F x] \dots]$ *v*-constructs a non-empty class
4. $[{}^0\exists \lambda x [\dots [{}^0F x] \dots]]$ EG, 3

Example: $\lambda w \lambda t [{}^0Wise_{wt} {}^0Pope_{wt}] \vdash \lambda w \lambda t \exists x [{}^0Wise_{wt} x]$;

Types: $Wise/(\alpha)_\tau; Pope/\iota_\tau; x \rightarrow_v \iota$.

b) *intensional* context. Let $[\dots \lambda y [\dots [{}^0F {}^0a] \dots]]$ *v*-construct **T**. Then the following rule is truth-preserving:

$$[\dots\lambda y [\dots [{}^0F {}^0a] \dots]] \vdash \exists f [\dots\lambda y [\dots [f {}^0a] \dots]]; f \rightarrow_v (\beta\alpha)$$

Proof:

1. $[\dots\lambda y [\dots [{}^0F {}^0a] \dots]]$ assumption
2. $[\dots\lambda y [\dots [f {}^0a] \dots]]$ $v(F/f)$ -constructs **T**
3. $\lambda f [\dots\lambda y [\dots [f {}^0a] \dots]]$ v -constructs a non-empty class
4. $[{}^0\exists\lambda f [\dots\lambda y [\dots [f {}^0a] \dots]]]$ EG, 3

Example: $\lambda w\lambda t [{}^0Believe_{wt} {}^0b \lambda w\lambda t [{}^0Wise_{wt} {}^0Pope_{wt}]] \models$
 $\lambda w\lambda t \exists f [{}^0Believe_{wt} {}^0b \lambda w\lambda t [{}^0Wise_{wt} f_{wt}]];$

Additional types:

$Believe/(\circ\iota_{\tau\omega})_{\tau\omega}$: an intensional attitude to a proposition; $f \rightarrow_v \iota_{\tau\omega}$

Gloss: If b believes that the Pope is wise then there is an office such that b believes that its holder is wise.

c) hyperintensional context. Let $[\dots [{}^0F {}^0a] \dots]$ v -construct **T**. Then the following rule is truth-preserving:²⁰

$$[\dots [{}^0F {}^0a] \dots] \vdash \exists c [{}^2Sub\ c {}^0F {}^0[\dots [{}^0F {}^0a] \dots]]; \\ c \rightarrow_v *_{n'}; {}^2c \rightarrow_v (\beta\alpha)$$

Proof:

1. $[\dots [{}^0F {}^0a] \dots]$ assumption
2. $[{}^2Sub\ c {}^0F {}^0[\dots [{}^0F {}^0a] \dots]]$ $v({}^0F/c)$ -constructs **T**
3. $\lambda c [{}^2Sub\ c {}^0F {}^0[\dots [{}^0F {}^0a] \dots]]$ v -constructs a non-empty class
4. $[{}^0\exists\lambda c [{}^2Sub\ c {}^0F {}^0[\dots [{}^0F {}^0a] \dots]]]$ EG, 3

The step 2 may be difficult to understand. Here is an additional explanation. The Composition $[{}^2Sub\ c {}^0F {}^0[\dots [{}^0F {}^0a] \dots]]$ $v({}^0F/c)$ -constructs the Composition $[\dots [{}^0F {}^0a] \dots]$. In order to v -construct **T**, this Composition must be executed again. Therefore, Double Execution.

Example: $\lambda w\lambda t [{}^0Believe^*_{wt} {}^0b [{}^0\lambda w\lambda t [{}^0Wise_{wt} {}^0Pope_{wt}]]] \models$
 $\lambda w\lambda t \exists c [{}^0Believe^*_{wt} {}^0b [{}^2Sub\ c {}^0Pope_{wt} [{}^0\lambda w\lambda t [{}^0Wise_{wt} {}^0Pope_{wt}]]]];$

Additional types:

$Believe^*/(\circ\iota^*_{n'\tau\omega})$: a hyperpropositional belief; $c \rightarrow_v *_{n'}; {}^2c \rightarrow_v \iota_{\tau\omega}$.

Gloss: If b (explicitly) believes* that the Pope is wise, then there is a concept of an office such that b believes* that the holder of the office is wise.

²⁰ I distinguish between a derivation rule, denoted ' \vdash ', and analytic entailment, denoted ' \models '. Of course, if a particular rule is truth-preserving then the corresponding entailment is valid.

Note: In this example the $[^0Sub\ c\ ^0Pope\ ^0[\lambda\omega\lambda\tau\ [^0Wise_{wt}\ ^0Pope_{wt}]]]$ is not Double executed, because b is related just to the Composition itself constructed by this substitution.

Hyperpropositional attitudes must be used if the attributer is reproducing faithfully b 's perspective. For instance, suppose that the office of Pope is identical with the office of the Bishop of Rome. Then it may be the case that though b believes that the Pope is wise, he may disbelieve that the Bishop of Rome is wise.

4.2 Leibniz's Rule of Substitution in the Three Kinds of Context

a) In an *extensional context* substitution *salva veritate* of v -congruent constructions is valid.

Example.

"The president of CR is the husband of Livia Klausova"

"The president of CR is an economist"

"The husband of Livia Klausova is an economist"

Proof:

1. $\lambda\omega\lambda\tau\ [^0President_of_{wt}\ ^0CR]_{wt} \approx \lambda\omega\lambda\tau\ [^0Husband_of_{wt}\ ^0Livia]_{wt}$ assumpt.
 2. $[^0Economist_{wt}\ \lambda\omega\lambda\tau\ [^0President_of_{wt}\ ^0CR]_{wt}]$ assumpt.
-
3. $[^0Economist_{wt}\ \lambda\omega\lambda\tau\ [^0Husband_of_{wt}\ ^0Livia]_{wt}]$ Leibniz 2)

Types: $President_of/(u)_{\tau\omega}$; CR/v ; $Husband_of/(u)_{\tau\omega}$; $Livia/v$; $Economist/(ot)_{\tau\omega}$;

b) In an *intensional context* substitution *salva veritate* of *equivalent* constructions is valid.

Example.

"The president of CR is the highest representative of CR"

"Tom wants to become the president of CR"

"Tom wants to become the highest representative of CR"

Proof:

1. $\lambda\omega\lambda\tau\ [^0President_of_{wt}\ ^0CR] \approx \lambda\omega\lambda\tau\ [^0Highest_Rep_of_{wt}\ ^0CR]$ assumpt.
 2. $[^0Want_be_{wt}\ ^0Tom\ \lambda\omega\lambda\tau\ [^0President_of_{wt}\ ^0CR]]$ assumpt.
-
3. $[^0Want_be_{wt}\ ^0Tom\ \lambda\omega\lambda\tau\ [^0Highest_Rep_of_{wt}\ ^0CR]]$ Leibniz 2)

Additional types. $Highest_Rep_of/(u)_{\tau\omega}$; $Want_be/(ou)_{\tau\omega}$; the relation-in-intension of an individual to an individual office; Tom/v .

c) In a *hyperintensional context* substitution *salva veritate* of *procedurally isomorphic* constructions is valid.

Example. Suppose that ‘azure’ and ‘sky-blue’ are synonymous.

“Tom believes* that Marie’s blouse is azure”

“Tom believes* that Marie’s blouse is sky-blue”

Proof.

1. [${}^0\text{Believe}_{wt}^* {}^0\text{Tom} [{}^0\lambda\omega\lambda t [{}^0\text{Azure}_{wt} [{}^0\text{Blouse_of}_{wt} {}^0\text{Marie}]]]]$] assumpt.
 2. ${}^{00}\text{Azure} \approx {}^{00}\text{Sky_Blue}$ assumpt.
-
3. [${}^0\text{Believe}_{wt}^* {}^0\text{Tom} [{}^0\lambda\omega\lambda t [{}^0\text{Sky_Blue}_{wt} [{}^0\text{Blouse_of}_{wt} {}^0\text{Marie}]]]]$] Leibniz

Some might object that this argument is invalid, because it is possible that Tom believes that Marie’s blouse is azure without believing that Marie’s blouse is sky-blue. We disagree and on this point we refer to Richard who says:

... It is impossible for a (normal, rational) person to understand expressions which have identical senses but not be aware that they have identical senses. (Richard 2001, 546-7)

Since the meanings of ‘azure’ and ‘sky-blue’ are identical by assumption, the two Trivializations ${}^0\text{Azure}$ and ${}^0\text{Sky_Blue}$ are not only equivalent but also *identical*, hence *procedurally isomorphic*. Hence the paradox of analysis is *not* a problem of hyperintensionality. Rather, it is a matter of linguistic incompetence (failure to recognize pairs of synonyms) and *not of logical incompetence* (failure to recognize pairs of procedurally isomorphic hyperintensions).

5 Conclusion

Once the three kinds of context, namely extensional, intensional and hyperintensional are defined, the extensional rules like the substitution of identicals and quantification into non-extensional contexts are trivially valid. There is no cogent reason for invalidity of Leibniz’s law. Only that we must substitute that object which is the object of predication in a given context.

Quantifying into hyperintensional contexts requires an *extensional logic of hyperintensions*. Much non-trivial footwork is required to lay out such a large-scale logical semantics. Once this is done, though, quanti-

fying into hyperintensional and intensional contexts turns out to be as trivially valid as quantifying into extensional contexts. However quantifying into hyperintensional contexts introduces a *technical complication* absent in quantifying into intensional and extensional contexts. In a hyperintensional context a construction occurs *mentioned* (as an argument of another function) rather than used (to construct a function). The complication is that, since every constituent of a *mentioned* construction itself occurs mentioned, the quantifier cannot bind any variables inside the hyperintensional context, thus rendering quantifying-in impossible. The solution consists in applying a *substitution technique* that makes the variables amenable to binding. Moreover, the substitution method made it possible to introduce the generally valid computational rule of a partial lambda calculus, *viz.* reduction ‘by value’.

*Department of Computer Science
VSB - Technical University
17. listopadu 15
708 33 Ostrava
Czech Republic
marie.duzi@vsb.cz*

Acknowledgements. This research was funded by the Grant Agency of the Czech Republic, project No. 401/10/0792, “Temporal aspects of knowledge and information” and also by the internal grant agency of VSB-TU Ostrava, project No. SP2012/26 „An utilization of artificial intelligence in knowledge mining from software processes“. Versions of this study were read by the author as an invited talk at the University of Western Australia, Perth, Australia, February 25th, 2011.

References

- ANDERSON, C. A. (1998): Alonzo Church’s contributions to philosophy and intensional logic. *The Bulletin of Symbolic Logic* 4, 129-171.
- BEALER, G. (1982): *Quality and Concept*. Oxford: Clarendon Press.
- BEALER, G. (2004): An inconsistency in direct reference theory. *The Journal of Philosophy* 111, 574-93.
- CARNAP, R. (1947): *Meaning and Necessity*. Chicago: Chicago University Press.
- CHURCH, A. (1954): Intensional isomorphism and identity of belief. *Philosophical Studies* 5, 65-73.
- CHURCH, A. (1993): A revised formulation of the logic of sense and denotation. *Alternative* (1). *Noûs* 27, 141-157.

- CLELAND, C. E. (2002): On effective procedures. *Minds and Machines* 12, 159-179.
- CRESSWELL, M. J. (1975): Hyperintensional logic. *Studia Logica*, No. 34, 25-38.
- CRESSWELL, M. J. (1985): *Structured Meanings*. Cambridge: MIT Press.
- DUŽÍ, M. (2010): The paradox of inference and the non-triviality of analytic information. *Journal of Philosophical Logic*, 39, 5, 473-510.
- DUŽÍ, M. – JESPERSEN, B. – MATERNA, P. (2010): *Procedural Semantics for Hyperintensional Logic. Foundations and Applications of Transparent Intensional Logic*. First edition. Berlin: Springer.
- DUŽÍ, M. – JESPERSEN, B. (to appear): Procedural isomorphism and restricted β -reduction. *Logic Journal of the IGPL*.
- DUŽÍ, M. – MATERNA, P. (2010): Can concepts be defined in terms of sets? *Logic and Logical Philosophy*, 19, 195-242.
- EIJCK, J. VAN – FRANCEZ, N. (1995): Verb-phrase ellipsis in dynamic semantics. In: Masuch, M. – Polos, L. (eds.): *Applied Logic: How, What and Why?* Kluwer, 29-60.
- FORBES, G. (1996): Substitutivity and the coherence of quantifying in. *The Philosophical Review*, 105, 337-371.
- FOX, C. – LAPPIN, S. (2001): A framework for the hyperintensional semantics of natural language with two implementations. *Lecture Notes in Computational Linguistics*, vol. 2009, 175-92.
- FREGE, G. (1892): Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik* 100, 25-50.
- HYDE, R. (1996): *The Art of Assembly Language Programming*. Available: <http://www.arl.wustl.edu/~lockwood/class/cs306/books/artofasm/toc.html>.
- JESPERSEN, B. (2003): Why the tuple theory of structured propositions isn't a theory of structured propositions. *Philosophia* 31, 171-183.
- JESPERSEN, B. (2010): How hyper are hyperpropositions? *Language and Linguistics Compass* 4, 96-106.
- KAPLAN, D. (1968): Quantifying in. *Synthese* 19, 178-214.
- KAPLAN, D. (1986): Opacity. In: Hahn, L. (ed.): *The Philosophy of W.V. Quine*. La Salle: Open Court, 229-289.
- KAPLAN, D. (1990): Dthat. In: Cole, P. (ed.): *Syntax and Semantics*. Vol. 9. New York: Academic Press. Reprinted in: Yourgrau, P. (ed.): *Demonstratives*. Oxford: Oxford University Press.
- KLEMENT, K. C. (2002): *Frege and the Logic of Sense and Reference*. New York: Routledge.
- KRIPKE, S. (1963): Semantical considerations on modal logic. *Acta Philosophica Fennica*, 16, 83-94.
- LEWIS, D. (1972): General semantics. In: Davidson, D. – Harman, G. (eds.): *Semantics of Natural Language*. Dordrecht: Reidel.
- MOSCHOVAKIS, Y. N. (1994): Sense and denotation as algorithm and value. In: Väänänen, J. – Oikkonen, J. (eds.): *Lecture Notes in Logic*. Vol. 2. Berlin: Springer, 210-249.
- MOSCHOVAKIS, Y. N. (2006): A logical calculus of meaning and synonymy. *Linguistics and Philosophy* 29, 27-89.
- PIERCE, C. B. (2002): *Types and Programming Languages*. London: MIT Press.

- PLOTKIN, G. D. (1975): Call-by-name, call-by-value and the λ -calculus. *Theoretical Computer Science*, 1, 125-159.
- QUINE, W. V. O. (1956): Quantifiers and propositional attitudes. *Journal of Philosophy* 53, 177-186.
- RICHARD, M. (2001): Analysis, synonymy and sense. In: Anderson, C. A. – Zelezný, M. (eds.): *Logic, Meaning and Computation: Essays in Memory of Alonzo Church*. Synthese Library, vol. 305, Dordrecht: Kluwer, 545-71.
- TICHÝ, P. (1968): Smysl a procedura. *Filosofický časopis* 16, 222-232. Translated as 'Sense and procedure' in Tichý (2004, 77-92).
- TICHÝ, P. (1969): Intensions in terms of Turing machines. *Studia Logica* 26, 7-25. Reprinted in Tichý (2004, 93-109).
- TICHÝ, P. (1988): *The Foundations of Frege's Logic*. Berlin – New York: De Gruyter.
- TICHÝ, P. (1994): The analysis of natural language. *From the Logical Point of View* 3, 42-80. Reprinted in Tichý (2004, 801-841).
- TICHÝ, P. (2004): *Collected Papers in Logic and Philosophy*. Edited by V. Svoboda, B. Jespersen, C. Cheyne. Prague: Filosofia, Dunedin: University of Otago Press.

Sémantický a pragmatický význam jazykových výrazov¹

Barbora Geistová Čakovská

Slovenská akadémia vied, Bratislava

Abstract: In this paper, I will focus on how to consider relevant pragmatic context parameters in understanding a sentence used in the context without introducing them into the field of semantics. In this sense, I will propose two options of analyzing the shift in the meaning of a used sentence, based on a use of the sentence, which work with propositions as determined, full-valued meanings of sentences independently of a use of the sentences in a pragmatic context.

Keywords: speaker, sentence, semantic meaning, pragmatic meaning, pragmatic context parameters, possible words and times, pragmatic reference, expression reference.

K téme prezentovanej v rámci ateliéru *Sémantický minimalizmus a kontextualizmus* pristupujem vo svojom príspevku z trochu všeobecnejšieho hľadiska a pracujem tu v istej miere s intuíciami týkajúcimi sa vybraných problémov. V literatúre o tejto téme si totiž možno všimnúť, že najmä intuície sú metódou využívanou zástancami kontextualizmu, proti ktorým sa usilujem touto staťou vystúpiť. Účinnou odpoveďou im tak môžu byť opäť najmä intuície, keďže argumentáciu čisto konkrétnymi modelmi by zrejme neprijali. Použitiu istých teoretických nástrojov sa však nevyhneme, pokladám za potrebné využiť ich, keďže je nutné vysvetliť pojmy, s ktorými budem v tejto stati pracovať. Za týmto účelom nám posluží najmä intenzionálna sémantika.

Vzhľadom na to, ako vágne alebo nejednoznačne sa používajú vo filozofii jazyka niektoré dôležité pojmy, vrátane základných pojmov

¹ Táto publikácia vznikla na Filozofickom ústave SAV v rámci grantového projektu VEGA č. 2/0019/12 *Jazyk a determinácia významu v komunikácii*.

ako jazykový výraz či jeho význam, spresním v úvode niekoľko bodov. Vetu prirodzeného jazyka chápem ako zložený výraz prirodzeného jazyka, ktorý má nejaký význam daný konvenčnými významami jednoduchých výrazov a istým spôsobom ich usporiadania, pričom táto veta je správne utvorená. Taký význam vety nazvem *doslovným* alebo *sémantickým významom*. Niektoré vety obsahujú indexické výrazy a ich význam vieme v plnej miere určiť až na základe istého kontextu použitia týchto viet, všeobecne však odlišujem indexikalizmus od kontextualizmu a vetám obsahujúcim indexické výrazy sa podrobnejšie venovať nebudem, keďže predstavujú iný problém. Niektoré vety však obsahujú výrazy citlivé na kontext iným spôsobom ako indexické výrazy. Príklady takých viet uvádzam nižšie. Tu sa význam vety použitej v nejakom kontexte modifikuje v závislosti od daného pragmatického kontextu, a to *modifikáciou* doslovného významu istým „spresnením významu“ kontextovo citlivých výrazov pomocou kontextu. Tento modifikovaný význam viet budem nazývať *vyjadreným významom*. To, ako k tomu dochádza, sa pokúsim vysvetliť neskôr. V prípade niektorých viet sú doslovný a vyjadrený význam totožné, ale v prípade množstva viet sa tieto dva významy odlišujú, keďže veľa viet obsahuje výrazy citlivé na kontext. Od týchto dvoch úrovní či druhov významu budeme odlišovať ďalšiu úroveň či druh významu, ktorý hovorca *komunikuje* použitím vety s doslovným či vyjadreným významom v nejakom pragmatickom kontexte. Nazvem ho *komunikovaným významom* a najlepšie je ho chápať podobne ako griceovskú mienenú implikatúru.

Pre spresnenie ešte doplním, že význam viet budeme chápať ako propozíciu, funkciu z možných svetov a časových okamihov (ďalej len svetamihov²) do pravdivostných hodnôt.

Pri použití nejakej vety hovorcom si musí adresát často nejakým spôsobom doplniť alebo modifikovať doslovný význam použitej vety (výpovede), aby identifikoval význam, ktorý hovorca danou výpoveďou vyjadroval či komunikoval. Ide nám o prípady vyjadrené príkladmi viet:

- (1) Všetky piva sú studené.
- (2) Jozef je vysoký.

a podobne.

² Svetamihy chápeme ako usporiadané dvojice $\langle w, t \rangle$.

Obidve z týchto viet obsahujú slová, ktorým konvenčne rozumíme a ktoré sú podstatné pre určenie doslovného významu viet. Konvenčný význam týchto slov a spôsob ich spojenia sa nemenia použitím v kontexte, preto vo všetkých kontextoch zostáva stabilný aj doslovný význam týchto viet. Doslovný význam takých viet explikujem ako plnohodnotnú propozíciu, ktorá si ďalej plní svoju funkciu pri určovaní pravdivostných hodnôt týchto viet.

Vety (1) a (2) však obsahujú kontextovo citlivé slová, napr. „všetky“ v prvom prípade a „vysoký“ v druhom prípade. Od kontextu použitia závisí to, že v prípade vety (1) hovorca napr. nechce vyjadriť, že všetky piva na svete sú studené, ale že studené sú piva v chladničke, prípadne v pivnici či komore. Má teda na mysli nejaké konkrétne univerzum pív, čomu adresát na základe vyhodnotenia pragmatického kontextu použitia tejto vety porozumie. Jozef z vety (2) môže byť vysoký napríklad vzhľadom na trpaslíkov, ale nie vzhľadom na výšku bežných ľudí. Ani v tejto vete sa toto doplnenie explicitne nevyjadruje, adresát však tomu na základe jej pragmatického kontextu použitia porozumie. Použitím v kontexte tak dochádza k určeniu vyjadreného významu daných viet na základe modifikácie ich doslovného významu kontextom. Vyjadrený význam budem nazývať aj *vyjadrenou propozíciou*, a práve na tento význam viet sa v tejto stati mienim sústrediť.

Napokon, hovorca použitím viet (1) a (2) môže komunikovať oveľa viac, napríklad aj to, že v prípade vety (1) si už môže adresát ísť pre pivo alebo v prípade vety (2) to, že Jozef určite dovidí na druhý koniec miestnosti ponad hlavy ostatných. Tejto úrovni, komunikovanému významu, sa ďalej nebudem venovať.³

Ako som už naznačila, v tejto stati sa sústredím práve na význam viet, ktorý som rozlíšila od ostatných ako vyjadrený význam použitých viet. Ak doslovný význam vety nazývame sémantickým významom, vyjadrený (aj komunikovaný) význam viet budeme nazývať *pragmatickým významom*, keďže k jeho určeniu potrebujeme poznať kontext použitia takej vety. Pragmatický význam vety sa v závislosti od pragmatického kontextu jej použitia mení, ale opačne to platí o jej sémantickom význame. Dalo by sa teda povedať, že jedna veta (použitá či nepoužitá) má jeden jediný, stabilný sémantický význam, ale môže

³ Viac k rozlíšeniu medzi doslovným významom vety, vyjadrenou propozíciou a mienenou propozíciou (komunikovaným významom) pozri v Zouhar (2011b).

viesť k množstvu pragmatických významov odlišných od vyjadreného významu aj navzájom, a to v závislosti od kontextov, v ktorých sa daná veta hovorcom použije a v ktorých sa jej pragmatický význam líši od sémantického významu. Tento názor sa nezdá byť problematický, bežná komunikácia je zasadená do nejakej situácie, pragmatického kontextu, v ktorom sa pragmatický význam vypovedaných alebo inak použitých viet stáva zrejým aj napriek tomu, že nie je explicitne vyjadrený v použitej vete. V prípade potreby dokáže hovorca pragmatický kontext doplniť tak, aby sa adresátovi tento význam stal zrejým. Bežná komunikácia funguje na báze istej hospodárnosti, kde nie je nutné doslova formulovať obsahy určené na komunikovanie inému na to, aby si účastníci komunikácie bez problémov rozumeli.

Problémy podľa môjho názoru začínajú v momente, keď niektorí autori tvrdia, že nejestvuje nič také ako ucelený či určený sémantický význam vety. Veta, či už použitá v nejakom pragmatickom kontexte, alebo nie, podľa týchto autorov nevyjadruje kompletnú, plnohodnotnú propozíciu, a preto nedokážeme určiť ani jej pravdivostné hodnoty. Až použitie vety v nejakom pragmatickom kontexte má poskytovať jej plnohodnotný význam, ktorý tu nazývam pragmatickým významom.⁴ Takto má až veta použitá v nejakom pragmatickom kontexte, na základe svojho nekompletného sémantického významu a pragmatického kontextu použitia, vyjadrovať kompletnú propozíciu, vďaka čomu vieme v závislosti od svetamihov vete priradiť konkrétne pravdivostné hodnoty.⁵ Pragmatický význam vety tak závisí od nejakého spolupôsobenia sémantického významu danej vety a istých pragmatických kontextových parametrov, ktoré sa pri danom použití vety aktivujú a sú relevantné pre určenie pragmatického významu danej použitej vety. Možnosti toho, ako a v akej rovine sa uskutočňuje spolupôsobenie sémantického významu s pragmatickými kontextovými parametrami pri určovaní pragmatického významu vety, sa čoskoro pokúsime analyzovať.

⁴ Znova upozorňujem, že tu sa venujem úrovni vyjadreného významu. Do opozície tak staviam sémantický (doslovný) význam viet a pragmatický (vyjadrený) význam viet. Pragmatický (komunikovaný) význam viet ma tu nezaujíma.

⁵ Vyjadrené na základe aplikácie sémantického modelu stručne prezentovaného v úvode.

Zástancovia kontextualizmu vo všeobecnosti tvrdia, že kontextové parametre nejakým spôsobom vstupujú do sémantického významu použitej vedy, teda vstupujú do nekompletnej propozície a vytvoria z nej kompletnú propozíciu.⁶ Keďže my chápeme propozíciu ako funkciu zo svetamihov do pravdivostných hodnôt, teda ako neštruktúrovaný objekt, toto vstupovanie a dopĺňanie propozícií je nemysliteľné. Bez ohľadu na to, nakoľko problematické sa nám to zdá, uplatnením zvolenej sémantickej koncepcie na toto hľadisko povieme, že táto kompletizovaná, a teda kompletná propozícia priraduje vetám v závislosti od svetamihov určité pravdivostné hodnoty. Svetamih tak v súlade s tým môžeme chápať ako istý druh kontextu, ktorý vstupuje do hry za účelom určenia pravdivostnej hodnoty vety či všeobecne referenta jazykového výrazu. Budem však rozlišovať medzi *pragmatickými kontextovými parametrami* použitia vety na jednej strane, napríklad časom jej výpovede, miestom, hovorcovými a adresátovými poznatkami o situácii, univerzom objektov, ktoré hovorca zohľadňuje, predchádzajúcou komunikáciou atď., pričom nie všetky identifikovateľné parametre musia byť relevantné pre určenie pragmatického významu použitej vety; a na druhej strane iným druhom kontextových parametrov, *svetamihom*. Zatiaľ čo pragmatické kontextové parametre spoluurčujú pragmatický význam použitej vety, svetamih spoluurčuje pravdivostnú hodnotu vety ako takej (použitej aj nepoužitej). Bez pragmatických parametrov nevieme určiť pragmatický význam použitej vety, na čo však nepotrebujeme svetamih. Tento argument v nami zvolenej koncepcii potrebujeme až pre určenie pravdivostnej hodnoty danej vety (použitej, alebo nie). Svetamih odlišujem od miesta a času použitia danej vety v pragmatickom kontexte, keďže moment výpovede sa môže líšiť od svetamihu vyhodnocovania jej pravdivostných hodnôt. Bližšie k tomu aj neskôr. Práve z dôvodu zdôraznenia tohto rozlíšenia používam spojenie *pragmatický kontext použitia vety*, kde čitateľovi mohlo slovo *pragmatický* pripadať nadbytočné.

V súvislosti so svetamihom nikto nepochybuje o tom, že ide o mimo-lingvistický, a teda aj mimosémantický parameter, ktorý nie je súčasťou významu výrazov. Zjednodušene povedané, ak propozíciu chápeme ako sémantický objekt a svetamih ako mimosémantický objekt, vďaka „spolupôsobeniu“ sémantického a nesémantického priradujeme vete vyjadrujúcej danú propozíciu pre daný svetamih určitú pravdivostnú

⁶ Napríklad Recanati.

hodnotu – nesémantický objekt. Pragmatické kontextové parametre výpovede vety by však podľa všetkého mali fungovať inak. Podľa kontextualistov totiž stoja už pri určovaní významu použitej vety, nie až pri určovaní jej pravdivostných hodnôt. Ako evidentne mimosémantické objekty, podobne ako svetamih, majú spolupôsobiť so sémantickým objektom tak, že ho nejakým spôsobom dopĺňajú, výsledkom čoho má byť opäť čisto sémantický objekt (pragmatický význam danej použitej vety). Pragmatické kontextové parametre vety by tak mali pôsobiť ešte v sémantickej rovine, stať sa nejakým spôsobom súčasťou čisto sémantického objektu – sémantického významu vety – a viesť k vytvoreniu ďalšieho čisto sémantického objektu! Vieme už, že problematyczny je spôsob, akým má k tomu dochádzať, či to, že takto sa určený význam použitej vety mení v závislosti od rôznych pragmatických kontextových parametrov, ktoré sa aktivujú pri rôznych použitíach tej istej vety. Vedie to ku konštatovaniu, že tá istá veta má v rôznych kontextoch použitia rôzne významy. Nehovoriac o tom, že nepoužitá veta zrejme nemá žiadny určený význam. O problémoch súvisiacich s naznačeným kontextualistickým pohľadom sa už veľa napísalo, nebudem sa im preto podrobnejšie venovať.⁷

Pokúsme sa teraz sústrediť na analýzu toho, čo je vlastne zmyslom komunikácie na základe použitia nejakej vety hovorcem v pragmatickom kontexte. Táto analýza má za cieľ vyhnúť sa problematickému zavádzaniu nesémantických objektov do sémantickej oblasti či tvrdeniu, že význam vety alebo akéhokoľvek iného jazykového výrazu bez jeho použitia v nejakom kontexte nie je určeným, kompletným významom výrazu.

Jedným zo spôsobov, ako to dosiahnuť, je chápať pragmatický význam použitej vety ako propozíciu vyjadrenú vetou, ktorá vznikne preformulovaním pôvodnej vety. Tak by sa už na jazykovej úrovni vyjadrili pragmatické kontextové parametre, ktoré sú relevantné pre vytvorenie pragmatického významu použitej vety odlišného od jej doslovného, sémantického významu. Napríklad, veta (1) by sa preformulovala na vetu (1*):

- (1*) Všetky pívá nachádzajúce sa o 16.00 hod v chladničke v byte č. 32 na Šancovej ulici č. 31 v Bratislave sú studené.

⁷ Napríklad Zouhar (2011), Zouhar (2011a), Geistová Čakovská (2011), pripravovaný článok Geistová Čakovská (2012), a i.

Veta (1*) vzniká preformulovaním vety (1) pri zohľadnení relevantných kontextových parametrov a explicitne vyjadruje to, čo sme nazvali pragmatickým významom vety (1) použitej v danom pragmatickom kontexte.⁸ Týmto pragmatickým významom je propozícia, ktorej v závislosti od iného druhu kontextu, svetamihu, priradujeme určitú pravdivostnú hodnotu. Pravdivostná hodnota vety (1*) je v nejakom svetamihu totožná s pravdivostnou hodnotou referovanou hovorcom použitím vety (1) v danom kontexte, pokiaľ sa tento hovorca v danom svetamihu nemýli, neklame ani inak adresáta nezavádza.

Kontextualisti by však mohli namietat', že sme vo vete (1*) nevyjadrili všetky podstatné kontextové parametre.⁹ V prípade, že by sme to uznali za potrebné pre danú komunikáciu, stále dokážeme vetu (1) preformulovať obsiahlejšie. Treba si však uvedomiť, že prípad vety (1) je v podstate jednoduchým príkladom. Určite by sme našli množstvo prípadov, kedy by preformulovanie použitej vety viedlo k vytvoreniu príliš dlhého výrazu, a niektorí autori tvrdia, že vyjadrenie všetkých relevantných pragmatických kontextových parametrov jazykovo nie je možné.¹⁰ Súhlasím minimálne s názorom o nepraktickom charaktere takej analýzy použitých viet. Mohlo by sa zdať, že sa táto analýza v prospech zachovania nedotknutosti sémantickej roviny vzdáva hospodárnosti komunikácie, ide však o teoretický nástroj, ktorý nehovorí o tom, ako majú hovorcovia medzi sebou komunikovať. Tento spôsob analýzy sa dá využiť aj pri analýze metaforických výrazov či výrazov s preneseným významom, čo však nie sú prípady, ktoré tu skúmame.

Pre uvedenie iného spôsobu riešenia sa vrátíme k jednej dôležitej poznámke. Myslím, že tu v skutočnosti nejde o významy výrazov, či už sémantický význam alebo pragmatický význam. V predchádzajúcom pokuse o analýzu problematických viet typu (1) a (2) som spomenula, že preformulovaná veta (1*) v nejakom svetamihu referuje pri splnení istých podmienok totožne s hovorcom vety (1). Týmto chcem odpútať pozornosť čitateľa od sémantickej roviny a upriamiť ju na rovinu referencie. Vlastne, v súvislosti s vetami, ani nie priamo na rovi-

⁸ V Zouhar (2011b) sa takto vzniknutá veta nazýva explicitnou vetou.

⁹ Pojem pragmatického kontextu a jeho relevantných parametrov je totiž v literatúre veľmi vágny a nie je zrejmé, čo všetko má taký kontext obsahovať a čo všetko má byť relevantné pri určení pragmatického významu.

¹⁰ Napríklad Searle.

nu referencie, ale skôr na rovinu stavu vecí vyjadreného¹¹ doslovným sémantickým významom použitej vety a modifikovaného relevantnými pragmatickými kontextovými parametrami. Pod stavom vecí rozumiem akýkoľvek stav vecí, ktorý sa dá prostredníctvom výrazov jazyka vyjadrovať, komunikovať. V prípade výpovede vety (1) je to napríklad to, že všetky pívá v danej chladničke v danom čase sú studené, v prípade výpovede vety (2) napríklad to, že Jozef je vysoký vzhľadom na ostatných ľudí v danej miestnosti. Kladieme si totiž otázku, čo je pre adresáta v rámci danej komunikácie dôležité: Poznať význam použitej vety?; uchopiť prostredníctvom nej nejaký komunikovaný stav vecí (nie nevyhnutne priamo empiricky), prípadne vedieť určiť, či touto vetou hovorca hovorí pravdu, alebo nepravdu?

Dôležité sú samozrejme oba. Ale adresát použitej vete rozumie vďaka jej doslovnému, sémantickému významu, lebo tento význam pokladáme za plnohodnotný a určený. Aby však adresát zistil, ako je to so stavom vecí, ktorý chce hovorca použitím tejto vety vyjadriť, potrebuje okrem jej významu zohľadniť aj relevantné pragmatické parametre kontextu, v ktorom bola daná veta použitá. A práve to pokladám za dôležité pri skúmaní toho, čo sme doteraz nazývali pragmatickým významom. Napríklad, veta (1) vďaka svojmu sémantickému významu vyjadruje taký stav vecí, že všetky pívá (v univerze diskurzu) sú studené.¹² Jej použitie hovorcom však vyjadruje stav vecí, že všetky pívá nachádzajúce sa teraz v chladničke sú studené. Adresát má uchopiť práve tento spresnený stav vecí. A adresát pozná relevantné pragmatické kontextové parametre, ktoré mu pomôžu uchopiť práve tento stav vecí, keďže je spolu s hovorcom situovaný v istom kontexte, v ktorom sa istým spôsobom vyzná, prípadne mu ho hovorca alebo niekto iný/niečo spresní.¹³ Tu je dôležité uvedomiť si, že pragmatické kontextové parametre pôsobia, podobne ako kontextový parameter svetamih, mimo sémantickej roviny vety, hoci sú ako nesémantické objekty iného druhu ako kontextový parameter svetamih a zohľadňujú

¹¹ Ide o iný vzťah vyjadrovania, ako je vzťah vyjadrovania medzi jazykovým výrazom a jeho významom.

¹² Teraz už budeme hovoriť iba o význame vety bez jeho rozlišovania na sémantický a pragmatický, keďže odteraz budeme rozlišovať iba jediný význam vety, ktorým je doslovný sémantický význam vety.

¹³ Samozrejme, riešime tu prípady korektnej bežnej komunikácie, kde nejde o hádanie, úmyselné zamlčovanie faktov atď.

sa ešte pred ním. Takto opäť nemusíme zavádzať pragmatické kontextové parametre ako cudzorodé objekty do roviny významu jazykového výrazu. Zohľadňujeme ich nie pre porozumenie vete vypovedanej hovorcom v hospodárnej forme vzhľadom na situáciu, v ktorej sa hovorca spolu s adresátom nachádza, ale preto, aby adresát vedel uchopiť stav vecí, ktorý veta sama osebe explicitne nevyjadruje, ale vyjadruje ho jej použitie v určitom pragmatickom kontexte. Adresát takto dokáže modifikovať, doplniť svoje uchopenie stavu vecí vyjadrených vetou a v prípade, že to pokladá za potrebné a je to pre neho možné, zistiť, či má hovorca pravdu, alebo nie. To je však už ďalší krok, kde sa až po uchopení stavu vecí, teda po zohľadnení pragmatických kontextových parametrov v súvislosti s použitou vetou, aplikuje kontext svetamihu. Ak je možné aj priamo empirické uchopenie komunikovaného stavu vecí a kontext svetamihu vie adresát aplikovať zároveň s pragmatickými kontextovými parametrami, vie zároveň zistiť, či hovorca hovorí pravdu, alebo nie.¹⁴ Druh pragmatickej referencie, o ktorej tu hovoríme, je vzťahom medzi hovorcom, vetou, relevantnými pragmatickými kontextovými parametrami, svetamihom a pravdivosťou hodnotou. Táto pravdivosťná hodnota sa môže líšiť od pravdivostnej hodnoty vety samej, čo je jedným zo známych rozdielov medzi výrazovou referenciou a pragmatickou referenciou.¹⁵ Ako nepotrebujeme vždy vedieť, či je nejaká veta pravdivá, alebo nepravdivá, aby sme získali nejaký poznatok a vedeli s ním ďalej pracovať, tak pre úspešnú komunikáciu nepotrebujeme vždy vedieť, či má hovorca pravdu, alebo nie. Vyjadrenému stavu sveta rozumieme vďaka tomu, že rozumieme významu vety a kontextu jej použitia. Tomuto kontextu však rozumieme iným spôsobom ako vete, keďže, podobne ako svetamih, nie je lingvistickým ani sémantickým objektom. Možnosť určenia tohto kontextu je oveľa väčšia, ako je možnosť overenia si pravdivosti toho, čo hovorca tvrdí, teda zistenia pravdivostnej hodnoty referovanej hovorcom použitou vetou v uvedených podmienkach, a to v takej miere, v akej je rozdiel medzi počtom úspešných komunikácií a počtom úspešností skúmania zistenia pravdivosti či nepravdivosti nejakej použitej vety. Relevantné pragmatické kontextové parametre sú nám bežne oveľa dostupnejšie ako kontextový parameter svetamih.

¹⁴ Ak je to pre neho možné.

¹⁵ Pozri Cmorej (2001), Zouhar (2004).

V prípade potreby analýzy v duchu uvedeného sémantického modelu sa dá vyjadrený stav vecí napokon vyjadriť istou vetou, kde sa vraciame k prvej možnosti analýzy. V tomto prípade však preformulovaním použitej vety nezačíname a nepredpokladáme ho pre uchopenie komunikovaného stavu vecí. Máme tu takto dve vety, jednu vyjadrujúcu význam bez zohľadnenia pragmatických kontextových parametrov, napr. vetu (1), a druhú formulovanú na základe významu prvej vety a so zohľadnením kontextových parametrov, napr. vetu (1*). Táto druhá veta je vetou, ktorú tretia osoba môže explicitne použiť na to, aby úplne vyjadrila to, čo hovorca komunikuje vetou (1), a to bez zmiešavania sémantickej a mimosémantickej roviny.

Filozofický ústav SAV
Klemensova 19
813 64 Bratislava
Slovenská republika
cakovskabarbor@gmail.com

Literatúra

- CMOREJ, P. (2001): Referencia, identifikácia a intencionálne akty. In: *Na pomedzí logiky a filozofie*. Bratislava: Veda, 245-263.
- GEISTOVÁ ČAKOVSKÁ, B. (2011): Problém identity jazykového výrazu a vzťahy synonymie z hľadiska logickej, lingvistickej a pragmatickej sémantiky. *Individuals and properties*. Špeciálne číslo *Organon F* 18, 115-125.
- GEISTOVÁ ČAKOVSKÁ, B. (2012): Identity of Linguistic Expressions and Lexical Synonymy in the Fields of Logical Semantics, Linguistic Semantics, and "Pragmatic Semantics". Pripravované vydanie vo vydavateľstve Onthos Verlag.
- RECANATI, F. (2008): *Philosophie du langage (et de l'esprit)*. Paris: Gallimard.
- ZOUHAR, M. (2004): *Podoby referencie*. Bratislava: IRIS.
- ZOUHAR, M. (2011): Kontext a sémantika predikátov. *Filozofický časopis* 59, č. 2, 187-206.
- ZOUHAR, M. (2011a): What is wrong with unarticulated constituents? *Human Affairs* 21, č. 2, 239-248.
- ZOUHAR, M. (2011b): *Význam v kontexte*. Bratislava: aleph.

Metametodologie a naturalismus

Vladimír Havlík

Akademie věd České republiky, Praha

Abstract: The article deals with the problem of consistency between metamethodology and naturalism in the philosophy of science. In the first part I clarify the concept of metamethodology in light of Lakatos' methodology of scientific research programs and his four theories of rationality. In the second part I analyze the naturalization of epistemology in Quine's perspective and the question whether metamethodology is compatible with his non-normative sort of naturalism. The final part focuses on the Laudan's version of naturalism which tries to retain normativity. I analyze Laudan's program for justification and appraisal of methodological rules and his attempt to formulate a metamethodological rule, which would be acceptable for all concurrent methodologies. From the naturalistic point of view it seems unacceptable that the naturalistic program would admit inviolable rule. I find such an objection to the metamethodological rule mistaken, but I formulate another objection as to why Laudan's rule is unacceptable. I propose my own version of metamethodological rules, which are inviolable but methodologically empty and trivial, and I try to show that the existence of such rules represent a worry for normative naturalism, but not a fatal flaw.

Keywords: philosophy of science, metamethodology, naturalism, methodological rules.

1 Co je metametodologie?

Metametodologie, obdobně jako jiné metatermíny, naznačuje, že je nadřazená v nějakém určitém smyslu jednotlivým metodologiím. Stejně tak jako v případech, kdy se předmětem zkoumání stává např. jazyk sám a jeho pravidla užití, uvědomujeme si, že toto zkoumání samo se odehrává také v nějakém jazyce a můžeme předpokládat, že zkoumání jazyka by mohlo probíhat v nějakém nadřazenějším jazyku – v metajazyku. Ponechme nyní stranou známé problémy, které takový

rozvrh provázejí, jako je nekonečný regres metajazyků nebo naopak pokus prozkoumat jazyk uvnitř jazyka samého. Chceme zmíněným příměrem pouze ilustrovat úroveň a vztahy, které metatermíny označují. Nadřazenost metatermínu tedy vyplývá z té skutečnosti, že existuje několik instancií nějakých entit, které se v jistých vlastnostech odlišují, ale v nějaké základní funkci shodují, a díky této shodnosti je pak možné uvažovat o zobecňujícím pohledu na tyto různé instance. Takové zobecnění zároveň ustavuje novou entitu, která ale nepatří do stejné úrovně entit, z kterých byla odvozena, ale do úrovně, která úroveň instancií entit překračuje – do metaúrovně. Z úrovně metajazyka tak zkoumáme jednotlivé instance jazyků a jejich pravidel, která se mohou stát opět instancemi generalizovaných metapravidel atp. v mnoha dalších případech.

V podobném smyslu je chápána i metametodologie. Jedním z těch, kdo o rozdílných metodologiích z hlediska metametodologie uvažoval (alespoň v našem smyslu), byl Imre Lakatos, když si uvědomil, že jednotlivé metodologie vědy může chápat jako výzkumné programy v duchu jeho originální metodologie vědeckých výzkumných programů. Rozdíl mezi metodologií a metametodologií je tak v tomto případě tříúrovňový. Jednotlivé metodologie se zabývají generalizací speciálních případů vědeckých postupů s cílem předložit normativní standardy, které by měly být při vědeckém bádání dodržovány. Úroveň metodologie je tedy úroveň generalizovaných pravidel vědecké činnosti. Jednotlivé instance postupů vědců v daných problémových situacích jsou zdrojem pro generalizace obecně závazných norem další vědecké činnosti. Jednotlivé případy postupů vědců jsou tedy obsahem první úrovně, zatímco metodologické generalizace těchto postupů jsou obsahem druhé úrovně. Třetí úroveň je pak úroveň metametodologická. V této úrovni je možné porovnat úspěšnost jednotlivých metodologií a zhodnotit je z hlediska vědecké praxe, ke které vedou. Lakatsovým záměrem je porovnat tyto možné metodologie či teorie racionality s ohledem na skutečnou historii. Metametodologie nám tak má poskytnout standardy, kterými můžeme zhodnocovat jednotlivé teorie racionality (tj. metodologie) a zvolit tu, která je nejvíce v souladu se skutečnou historií. Soupeřící metodologie tak mohou být zhodnoceny s pomocí normativně interpretované historie. Podle Lakatose lze chápat „všechny metodologie jako historiografické (nebo metahistorické) teorie (nebo výzkumné programy), jež lze podrobit kritice na základě posouzení racionálně historické rekonstrukce, ke které vedou.“ (Laka-

tos 1970, 109) Historie se tak stává testem racionálních rekonstrukcí, ke kterým vedou jednotlivé rivalitní metodologie. Nejlepší metodologií je pak ta, která maximalizuje roli vnitřních faktorů a minimalizuje roli vnějších faktorů, když generuje historickou rekonstrukci. Z metametodologického hlediska rozeznává Lakatos následující čtyři teorie racionality vědeckého vývoje (Lakatos 1970, 91):

- a) induktivismus
- b) konvencionalismus
- c) metodologický falzifikacionismus
- d) metodologii vědeckých výzkumných programů (MVVP).

Podle Lakatose musí být všechny tyto normativní rekonstrukce doplněny externími empirickými teoriemi tak, aby bylo možné vysvětlit zbylé neracionální faktory. Důvod je pochopitelný: „[H]istorie vědy je vždy bohatší než její racionální rekonstrukce“ (Lakatos 1970, 105). Přitom každá ze zmíněných čtyř racionálních rekonstrukcí určuje nějaký charakteristický vzor růstu vědeckých znalostí a s přihlédnutím k rozsahu externí historie, kterou je třeba doplnit interní historii racionální rekonstrukce, může být pak provedeno posouzení takové metodologie. Lakatos tak zaujímá metametodologické stanovisko výhradně z důvodů posouzení jednotlivých metodologií. Pro něho není rozhodující existence metametodologie jako takové, ale metametodologické hledisko je výhodné, protože umožňuje posouzení rivalitních metodologií. Konečně nejúspěšnější metodologií je právě Lakatosova MVVP, které se daří minimalizovat roli externí historie a maximalizovat roli interní historie. K nejpozoruhodnějším aspektům pak patří sebezduvodnění MVVP. Lakatos využije svou koncepci MVVP v historiografii tak, že ji aplikuje na metametodologické úrovni pro posouzení rekonstrukcí, ke kterým jednotlivé metodologie vedou, a tedy i pro posouzení rekonstrukce, ke které vede MVVP na metodologické úrovni. Díky tomu tedy metametodologická aplikace MVVP ospravedlňuje MVVP na metodologické úrovni jako plně legitimní a nejúspěšnější metodologii. Nás však nyní nezajímá tato pozoruhodná skutečnost, ale chceme se soustředit přímo na metametodologickou úroveň jako takovou. Je Lakatosova koncepce takové metametodologie, jež je založena na MVVP, jedinou možnou, nebo lze uvažovat o jejím jiném založení? V zásadě bychom našli odpověď již u samotného Lakatose, protože v zájmu prokázání schopností své historiografické metody nejprve aplikuje falzifikacionismus na normativní historiografické metaúrovni sám na sebe

a pak i na induktivismus a konvencionalismus, aby ukázal, že všechny metodologie jsou takto falsifikovány a že tedy falsifikacionismus nemůže být vhodným metakritériem pro posuzování jednotlivých metodologií (Lakatos 1970, 109). Přestože tedy falsifikacionismus nedává požadované výsledky na metametodologické úrovni, ukazuje na druhé straně, že je možné založit metametodologii aplikací jiné metodologie než MVVP. V souvislosti s tím vzniká i otázka, zda je nutné přijmout Lakatosovo kritérium historie pro posouzení racionálních rekonstrukcí a zda nemůže být empirické posouzení založeno jiným způsobem především v rámci silného naturalistického proudu ve filosofii vědy. Tolik pro násin Lakatosova přístupu k metametodologii. Podívejme se nyní na východiska epistemologického naturalismu.

2 Co je (epistemologický) naturalismus?

Kořeny naturalismu bychom měli sledovat pravděpodobně mnohem dále než ke Quinovi. Sám Quine se hlásí, pokud jde o silnou inspiraci a závazek k naturalismu, k Johnu Deweymu (Quine 1968, 26). V tomto ohledu je z hlediska naturalismu pak nejpodstatnější ta myšlenka, že „vědění, mysl a význam jsou částí toho samého světa, s kterým zachází (kterého se týkají), a mají být studovány v tom samém empirickém duchu, který oživuje přírodní vědu“ (tamtéž). Netriviálním důsledkem této na první pohled snad obecně přijatelné pozice je však radikální proměna epistemologie, s kterou Quine přichází. Není zde již žádné privilegované místo pro první filosofii (tamtéž), jež by v Descartově duchu předcházela empirickému poznávání vůbec. Epistemologie již tedy není v tradičním smyslu vědou zkoumající meze a možnosti našeho poznání světa a tím předcházející jakékoli vědění. Epistemologie již není fundamentálnější vědění o vědění, jež by zkoumalo skryté předpoklady vědění samého, ale stává se prostou součástí vědy, neboť ta je výrazem nejlepších a nejspolehlivějších metod poznávání světa.

Vzhledem k tématu, které v této stati sledujeme, je prvořadou otázkou slučitelnost naturalismu s metametodologií. Zdá se však, že Quinovy epistemologické požadavky koncipující teorii vědění, ne jako již nějaké svébytné filosofické předporozumění, ale spíše jako „kapitolu psychologie“, tomuto souladu přímo odporují. Jestliže epistemologie není již „první filosofii“ ani „metateorií poznání“, v jakém smyslu bychom mohli mluvit o metametodologii? Přestože se Quine otázkou

metametodologie přímo nezabývá, byl by takový závěr unáhlený. Naturalizace epistemologie neznamena pro Quina konec filosofie, či její případné nahrazení vědou, jak by se mohlo na první pohled zdát. V této otázce se přikláním k těm filosofům, kteří v Quinově snaze o naturalizaci epistemologie nevidí opuštění specifické role filosofie, ale např. tvrdí, že „Quinova teorie přirozeně vede k novému modelu vědění, v kterém filosofie hraje podstatnou roli: ne jako „první filosofie“ nebo „metavěda“, či jako „kapitola psychologie“, ale jako nezávislá disciplína sama o sobě, společně a komplementárně s vědou“ (Sher 1999, 491-492). V tomto smyslu se zdá být nerozlišování mezi „první filosofii“ a „metateorií“ neuvážené. Možná distinkce mezi „první filosofii“ a „metateorií“ je zobrazitelná jako rozdílné pojetí mezi topologií a chronologií. V případě metateorie je topologie následující: máme zde vertikální osu, s kterou spojujeme jednotlivé (meta)úrovně poznání a horizontální osu, s kterou spojujeme jednotlivé instancie poznání. Každá hypotetická „metateorie poznání“ pak odvozuje svoji nadřazenost z umístění na vertikální ose tím, že by měla být schopna zahrnout pod sebe všechny způsoby poznávání rozmístěné horizontálně v nižší úrovni. Na rozdíl od toho role „první filosofie“ nezapadá do takové topologické struktury. Její role je ještě výjimečnější a výsadnější, ale není zobrazitelná topologicky, ale spíše chronologicky nebo až genealogicky. Jejím úkolem je prozkoumání předpokladů možného poznání vůbec, což musí předcházet jakémukoli (i vědeckému) poznání. V tomto smyslu čerpá „první filosofie“ svou nadřazenost (či spíše předřazenost) z časové primárnosti a fundace – tj. ze zrozování vědění jako takového. Její role je tedy zakládající, zatímco role metateorie je unifikující a generalizující. Ruší tedy Quinův požadavek naturalizace epistemologie topologickou i chronologicko-genealogickou strukturu?

Úkol filosofie, tak jak ho formuluje Quine na mnoha místech (např. 1968, 1995), je zkoumat, „jak se může z nárazů částic a vln na naše smyslové receptory zrodit něco tak grandiózního, jako je naše teorie světa, vyjadřující naše poznání tohoto světa“ (Quine 1995). Quine si uvědomuje, že není jiné cesty k vědění o vědění než vědění samo a spíše než vědění samo jako hotové vědění, je to naopak právě proces jeho ustavování. Stimulace a dráždění smyslových receptorů na vstupu a sofistikované teoretické systémy odpovídající světu na výstupu. Pouze v tomto rozsahu se může odehrávat poznávací proces vedoucí k vědění a pochopitelně také i možné vědění o vědění (sebereflexivní poznávací vztah) je pak realizováno v tomto procesu samém, a nikoli

před ním nebo mimo něho. Zdůvodnění vědění nemůže být dosaženo mimo vědění samo, mimo samotný proces poznávání. Tím se epistemologie jako věda nárokuje si zdůvodnění vědění naturalizuje. Ztrácí tak své dosud významné postavení „první filosofie“, stojící mimo poznávací proces a stává se plně integrovanou do tohoto procesu samého. „Pro mne je epistemologie, nebo to co se jí nejvíce podobá, tím, jak my, živočichové, můžeme dosáhnout skutečné vědy, s pouze danými strohými neurálními vstupy“ (Quine 1981, 21). V tomto smyslu jde Quinovi při naturalizaci epistemologie jistě o proměnu dosavadního (a ne dosti uváženého) předpokladu, že adekvátnímu dosahování vědění musí předcházet analýza a stanovení podmínek takové činnosti vůbec, která je dosažitelná speciálními způsoby myšlení, jež jsou nějak významně odlišné od činnosti myšlení při běžném dosahování vědění. Quine tedy jistě ruší chronologicko-genealogickou vazbu epistemologie. V žádném případě však není naturalizace epistemologie spojena s topologickou strukturou možných generalizací. Epistemologie není „metateorií poznání“ pouze ve smyslu „první filosofie“, tj. „metateorií“ ve smyslu něčeho, co má zakládající význam pro poznávací procesy a stojí zároveň mimo tyto procesy samé. Epistemologie ale může být „metateorií“ ve smyslu unifikující a generalizující teorie, jež subsumuje jednotlivé případy a způsoby poznání. V tomto smyslu je metateorie plně závislá na poznávacím procesu samém a je možná pouze na základě jednotlivých instancí poznávací činnosti. V tomto smyslu je tak „metateorie“ *ex post* záležitostí a nijak nezakládá nebo nepředchází samotný poznávací proces. Jak říká sám Quine ve známé a často citované tezi:

Vztah mezi skrovným vstupem a plnohodnotným (torrential) výstupem je vztahem, který nás podněcuje k jeho studiu z těch samých důvodů, které vždy podněcovaly epistemologii: především pro pochopení toho, jak se evidence vztahuje k teorii a jakými způsoby teorie přírody transcenduje jakoukoli dostupnou evidenci. (Quine 1969, 83; kurzíva VH)

Vědění jako systém poznatků (teorie) je tedy nutně transcendující vzhledem k evidenci z jaké vychází. Může být tedy chápáno v tomto smyslu jako metateorie evidence a nemělo by vést ke snaze popřít možnost metaúrovňových generalizací.

Přestože Quine tvrdí, že je v pořádku, že jsme „přestali snít o dedukování vědy ze smyslových dat“ (Quine 1969, 84), je v určitém smyslu

stále platná snaha „staré“ epistemologie konstruovat přírodní vědu ze smyslových dat (Quine 1969, 83). Specifičnost nové (naturalizované) epistemologie ale spočívá podle Quina v tom, že musíme přihlídnout ke vzájemnému omezení ve dvou různých smyslech, ve smyslu epistemologie v přírodní vědě a ve smyslu přírodní vědy v epistemologii. Na jedné straně musíme pochopit, že epistemologie je kapitolou psychologie a na druhé straně, že přírodní věda jako celek obsahuje psychologii s mnoha částmi, a všechno to, že je naše vlastní konstrukce nebo projekce z našich stimulací smyslových receptorů (tamtéž). Quine v naturalizaci epistemologie vidí především jednotu, s jakou se vědění ustavuje. Vědě je třeba porozumět jako procesu ve světě tak, že toto porozumění nemůže být lepší než věda sama, jež je jeho předmětem (Quine 1969, 84).

V zásadě je tedy Quinův naturalismus slučitelný s metametodologií. Quine ovšem nikde neklade otázku po volbě metodologie jako Lakatos. Problém vědecké metodologie mu v kontextu naturalizované epistemologie nepřipadá jako rozhodující. Primárním úkolem je pro něho spíše pochopení toho jak se od jednotlivých stimulů můžeme dopracovat vědy v tom tvaru, jak ji známe. Quine tak věnuje spíše pozornost buď elementárním otázkám, jak dochází ke stimulaci exteroceptorů, další reflexi v jazyce v pozorovacích větách a následnému ustavování pozorovacích kategorikálů jako základu vědění, nebo pak zase komplexním otázkám vědy jako holistického celku apod. Volba vědecké metodologie ve smyslu normativního postupu při získávání nových poznatků zde nemá privilegovaný význam. Rozhodující jsou konceptuální otázky naturalizované epistemologie, ne zdánlivě pragmatické metody postupu k vědění. Možná i právě díky tomu čelí Quinova naturalizace epistemologie výtkám, že musí nutně svým zaměřením rezignovat na normativitu. Laudan například tvrdí, že „Quine, Putnam, Hacking a Rorty z odlišných důvodů tvrdí, že to nejlepší, co můžeme udělat, je popsat metody užívané přírodovědci, díky čemuž zde není prostor pro normativní metodologii, která by byla preskriptivního charakteru“ (Laudan 1987, 19). Další pak označují takový proud přímo za „deskriptivní naturalismus“ (Losee 2004, 135). Naturalizovaná epistemologie je primárně zaměřena především na to, jak získáváme vědění, což není identické s tím, jak bychom měli vědění získávat. Přestože se někteří autoři snaží Quina obhájit v této otázce, jako například P. Maddy (2007, 87), která cituje některá místa, kde se Quine brání ztrátě normativity v naturalizované epistemologii, jedná se v zásadě jen

o Quinova proklamativní tvrzení, že se normativní element, tak charakteristický pro epistemologii, neztrácí při její naturalizaci (Quine 1990, 1995). Poněkud konkrétnější je Quine, pokud jde o záchranu normativního náboje naturalizované epistemologie v následující citaci:

Naše spekulace o světě zůstávají předmětem norem a výstrah, které vyplývají z vědy samé při jejím osvojování. Takto je jedním z našich vědeckých zjištění skutečnost, že informace o světě se k nám dostávají pouze silami působícími na naše nervová zakončení; a toto zjištění má normativní váhu, varuje nás proti tvrzením telepatie a jasnovidectví. Normy se mohou měnit s tím, jak se věda vyvíjí. Například jsme mnohem obezřetnější na působení na dálku, než bychom byli v době Newtona. (Quine 1981, 181)

Pokud by měla tato Quinova představa uchování normativity epistemologie sloužit jako exemplární příklad, pak lze jen těžko vidět, jak by mohlo takového cíle být dosaženo. Quinem uváděný příklad je problematický sám o sobě. Není z něj především vůbec jasné, jak se můžeme odvolávat na normativitu zjištěných skutečností, neboť tradičním problémem axiologie je uvědomění si, že to, co má být, nijak neplyne z toho, co je. Jaká je pak normativní váha zjištěné skutečnosti, že informace o světě se k nám dostávají pouze drážděním našich nervových zakončení? Takové skutečnosti nás sice mohou ochraňovat od pseudovědeckých a nevědeckých vizí, ale jde v tomto případě opravdu o normativitu vzhledem k získávání vědění? V zásadě by norma měla vést k závaznému způsobu činnosti, jež by byla zárukou nabytí jistoty při získání či testování poznatků. V daném případě však stojí zjištěné existující skutečnosti proti v zásadě neexistujícím a smyšleným (věda proti nevědě) a nezdá se, že by nějakým způsobem mohly z toho, že *jsou*, získat status toho, že *by být měly*. Na druhé straně závazné způsoby činnosti při nabývání jistoty vědění musí čelit mnohem normativnějšímu požadavku. Musí získat svůj normativní status z mnoha různých *existujících* způsobů činnosti, a musí tedy hledat důvody pro to, že *by být měly*, jinde než v pouhé skutečnosti, že *jsou*. Běžné situace, v nichž získáváme vědění a v kterých se dovoláváme jejich normativity, jsou tedy mnohem odlišnější než Quinem uváděný příklad. Proto jsou zřejmě námitky vůči naturalizované epistemologii Quinova typu oprávněné. Na druhé straně se zdá, že existuje způsob, jak naturalizovat epistemologii a podržet její normativní status mnohem neproblematičtěji. Navíc taková forma epistemologie je pochopena jako metaepistemologická

koncepte, což je otázka, která nás nyní především zajímá. Metaepistemologická proto, že Laudan se nesnaží charakterizovat epistemický naturalismus ani tak jako epistemologii o sobě, ale spíše jako teorii „o“ filosofickém vědění, tj. metateorii, jež tvrdí, že filosofická tvrzení jsou rozhodována stejně tak, jako jsou rozhodována tvrzení v jiných oblastech života, jako je věda, přirozené myšlení a zákony (Laudan 1990, 44). Laudan chce říci, že z pozice metateorie, tj. v tomto případě z pozice epistemického naturalismu, se mezi jednotlivými teoriemi o světě nachází také teorie vědění, která obdobně jako ostatní vypovídá něco o tom, jak je svět konstituován. Filosofie z tohoto hlediska tedy není ani logicky prvotní vzhledem k těmto jiným formám zkoumání a ani nadřazenější jako způsob vědění (tamtéž). To plně odpovídá Quinovu pojetí vědy a filosofie, jež jsou z hlediska naturalismu jednoduše celkem, a není zde nic, co by stavělo filosofii v nějakém ohledu před vědu, nic v čem by byla primárnější nebo fundamentálnější. Laudan si však na rozdíl od Quina uvědomuje, že říkáme-li něco o vědě a filosofii, pak to říkáme nutně z nějaké metapozice, z metateorie. Topologická struktura je však to jediné, co tuto metateorii odlišuje od úrovně teorií, jež jsou předmětem jejich závěrů a poznatků. Metateorie si nečiní nárok na nijak výsadnější postavení, než má předmět, který zkoumá, a to ani z hlediska faktů, metod nebo cílů. Předpokladem je, že hodnotící standardy a procedury, jež metateorie používá, vznikají uvnitř praxe vědy a jsou hodnoceny tím samým způsobem, jak jsou hodnoceny vědecké teorie. Především odkazem k tvrzením o světě (Losee 2004, 130). Stejně tak není metateorie transcendentní disciplínou, v které by platily nehistorické (transhistorické), nezrušitelné principy.¹ Naopak i tyto principy podléhají vývoji a změně v závislosti na tom, jak se rozvíjí vědecké poznání a vědění vůbec. A konečně, ona metateorie, o které nyní hovoříme, má pro Laudana podobu filosofie vědy či metodologie. Laudan zdůrazňuje, že naturalistická epistemologie je teorií vědění, v rámci níž dochází k formulaci tvrzení (teorií nebo hypotézám) o vědeckém bádání, a tyto teorie či hypotézy jsou vystaveny stejným strategiím zhodnocení jako teorie či hypotézy v rámci vědy (Laudan 1990, 45). Obdobně také „epistemické doktríny nebo pravidla jsou omylné předpoklady či domněnky, stejně tak jako všechny ostatní elementy vědeckého poznání“ (Laudan 1990, 46). Laudan si podobně jako Lakatos tedy všimá

¹ Nezrušitelnost principů je odvozena od té skutečnosti, že neexistuje nic, vůči čemu by mohly být revidovány a jako principy či pravidla odvolány.

toho, že epistemologie nachází konkrétní podobu v metodologii vědy, jež je souborem principů a pravidel pro dosahování a zdůvodňování vědění. Avšak vzhledem k tomu, že existuje rozmanitost mezi jednotlivými metodologiemi, musí mít epistemologie nutně výšeúrovňový status. Laudan pak hovoří o „naturalistické metametodologii“ (Laudan 1990, 46), jež zachovává důležitou kritickou a preskriptivní roli filosofie vědy a která umožňuje volit mezi rivalitními metodologiemi a epistemologiemi vědy (Laudan 1987, 25). Jakým způsobem chce ale Laudan dosáhnout toho, aby naturalistická epistemologie byla normativním naturalismem?

3 Co je normativní naturalismus?

Normativnost epistemologie obecně a normativnost dané metodologie či filosofie vědy je běžně odvozena od toho, že standardy, s jejichž pomocí má být dosahováno vědění a jeho zdůvodnění, mají preskriptivní status. Předepisují činnosti, jež je nutné provádět v daných specifických stádiích vědeckého objevování nebo zdůvodňování. Normativní naturalismus předpokládá, že preskriptivní závěry vyplývají z deskriptivního popisu historie vědy. Jde sice o problematický, ale zároveň nijak převratný či originální požadavek. Předpoklad, že analýza historického vývoje vědy, případně vybraných exemplárních událostí, povede k formulaci standardů racionality, je běžným východiskem mnoha epistemologických koncepcí včetně Lakatosovy metodologie vědeckých výzkumných programů. Avšak na rozdíl od Lakatose, který doufal v odkrytí neměnných standardů racionality, počítá Laudanův normativní naturalismus s tím, že tyto „standardy, stejně tak jako vědecké teorie samy, mají pouze provizorní status. Jsou předmětem korekcí a mohou být opuštěny v průběhu další zkušenosti“ (Loose 2004, 130). Normativní naturalismus tedy předpokládá, že odhalení korigovatelných a historicky proměnlivých preskriptivních standardů, je možné na základě deskripce historie vědy prostřednictvím empirického zkoumání stejného typu, které běžně ve vědě užíváme. Takto dosažené standardy je pak nutné aplikovat ve vědě, aby zajistily její progresivní vývoj. Laudan chce tímto způsobem vybudovat naturalistickou teorii metodologie, která zachová preskriptivní roli filosofie vědy a která zároveň umožňuje volit mezi rivalitními metodologiemi a epistemologiemi vědy (Laudan 1987, 25).

Naturalismus i normativitu své koncepce obhajuje Laudan například následujícím způsobem:

Takový přístup je naturalistický, protože nahlíží na epistemologii a metodologii ve stejném rozsahu jako na vědy; je normativní, protože na rozdíl od mnoha forem naturalismu uchovává důležitou roli pro metodologická doporučení a zhodnocení. [...] Důležité je zdůraznit na počátku, že normativní naturalismus je *metaepistemologie* nebo mnohem přesněji *metametodologie*. Zamýšlí ukázat, jak pravidla zdůvodnění či pravidla metodologie mohou být sama zdůvodněna či oprávněna. (Laudan 1996, 173)

Laudanova *metametodologie* stojí podobně jako Lakatosova nad jednotlivými metodologiemi vědy, ale narozdíl od něho přichází s významně odlišnými požadavky, pokud jde o racionalitu a způsoby testování rivalitních metodologií. Navíc svou povahou není taková *metodologie apriorní* a *nekorigovatelná*, ale naopak je to podle Laudana *metodologie stejně nejistá jako věda*. „Nejde o to učinit metodologii mnohem spolehlivější (secure) než fyziku (tj. vědu), ale o to, ukázat, že je stejně spolehlivá jako fyzika“ (Laudan 1987, 29).

Na rozdíl od Lakatose (a všech dalších, kteří trvají na racionální rekonstruovatelnosti) nepovažuje Laudan racionální rekonstruovatelnost za vhodné *metakritérium* pro posouzení metodologických pravidel. Podle jeho názoru nelze metodologická pravidla posuzovat podle toho, zda se jeví jako racionální vzhledem k volbě teorie u významných vědců v signifikantních historických případech, protože zde neexistuje racionality sama o sobě nezávislá na progresu vědy. Laudanovi jde o to, že v historických případech zastávali vědci odlišné hodnoty (axiologies) a předpoklady (background knowledge) od našich a jejich volba díky tomu mohla být racionální, ale my již nemůžeme očekávat, že by to naše (rekonstruovaná) *metodologie* mohla odhalit (Laudan 1987, 29). To je důvod, jenž vede Laudana k domněnce, že metodologická pravidla nejsou nic jiného než vazby mezi cíli a prostředky. Z hlediska naturalistické epistemologie to je výhodné minimálně ze tří důvodů. Zaprvé vazba mezi cíli a prostředky je historicky rozmanitá a umožňuje tak posoudit různé signifikantní případy volby teorie vzhledem k držným hodnotám a předpokladům. Zadruhé lze tvrzení cíle-zprostředkujících vztahů (means-ends relationships) považovat za hypotetické imperativy, které tak získávají normativní náboj, a konečně zatřetí je možné úspěšnost těchto vazeb mezi cíli a prostředky podrobit empi-

rickému výzkumu. „Strategie výzkumu zahrnutá v metodologických pravidlech může být testována evidencí, že sledování pravidel nám umožní progres v realizaci našich kognitivních cílů“ (Laudan 1987, 29). To odpovídá naturalistickému požadavku „testovat metodologická pravidla stejným způsobem jako jiná deskriptivní nebo teoretická tvrzení“ (Laudan 1987, 25).

Laudan v rámci svého naturalismu navrhuje dvoufázový program pro posouzení hodnotících standardů a metodologických pravidel. V první fázi je nutné reformulovat standardy nebo pravidla do tvaru hypotetických imperativů: z tvaru pravidla ve formě kategorického imperativu (0) „mělo by se dělat y “ je třeba vytvořit pravidlo ve tvaru hypotetického imperativu typu (1) „Je-li hlavním kognitivním cílem x , mělo by se dělat y “, které má preskriptivní obsah a je vztaženo k požadovaným kognitivním cílům. Ve druhé fázi pak je třeba určit, zda uskutečnění y vede spíše k dosažení cíle x než jiné alternativy. Podle Laudana tato fáze vyžaduje historický výzkum, při kterém je třeba zjistit, jaké strategie zkoumání vedou k dosažení jakých kognitivních cílů. Pravdivost nebo zaručenost takto získaných metodologických pravidel potom vyplývá z tvrzení (2) „Dělání y vede k x mnohem pravděpodobněji než jiné alternativy“, které má formu statistického zákona (Laudan 1987, 25). Příklady transformace standardních metodologických pravidel do hypotetických imperativů jsou uvedeny a přeloženy z Losee (2004, 135):

Standardní pravidla	Hypotetické imperativy
„Vyvaruj se <i>ad hoc</i> hypotéz“ (Popper)	„Je-li cílem vyvinout riskantní hypotézu, pak bychom se měli vyvarovat <i>ad hoc</i> hypotéz.“
„Začleň koroborovaný převyšující obsah“ (Lakatos)	„Je-li progres cílem vědeckého výzkumného programu, pak hledej teorii T_{n+1} takovou, že: 1 T_{n+1} vysvětluje předcházející úspěch T_n , 2 T_{n+1} má vyšší empirický obsah než T_n , a 3 část převyšujícího obsahu T_{n+1} je koroborován.“

Reformulace metodologických standardů do hypotetických imperativů vede k tomu, že je možné poměřovat a volit mezi rivalitními me-

todologiemi, jež se tak stávají jen soubory hypotetických imperativů (tj. metodologických pravidel). Za završení úspěchu naturalistického programu lze pak považovat to, že volba metodologie není nyní problematičtější než volba teorie v libovolné jiné oblasti (Laudan 1987, 25).²

Zajímá-li nás nyní především kompatibilitnost metametodologie a naturalismu, pak se musíme soustředit na základní tenzi Laudanovy metametodologie. Jestliže jsme na jedné straně zdůraznili pozitivní roli Laudanovy strategie transformace metodologických pravidel do hypotetických imperativů, díky čemuž se objevuje možnost vybudovat metametodologický pohled na jednotlivé soupeřící metodologie a také základ jejich testování a případné volby, pak je nutné věnovat pozornost také dalšímu důsledku, jenž z této situace metametodologie vyplývá. Se získanou metametodologickou pozicí totiž získáváme i možnost formulovat určité metametodologické principy či standardy. Takovou možnost nelze obecně považovat za špatnou. Každá úroveň zobecnění připouští formulaci obecných souvislostí či zákonitostí. Problém však spočívá v tom, zda jsou takové principy či standardy kompatibilní s naturalismem? Než se pokusíme prozkoumat tuto otázku, podívejme se podrobněji na Laudanovo metametodologické pravidlo (rule) R_1 .

Laudanův požadavek testování metodologických pravidel vede k obtíži, které musí čelit rozličné metodologické koncepce – problému regresi či kruhu. Každý test metodologického pravidla může být proveden pouze na základě jiného či jiných již ustanovených pravidel. Takové zřetězení zdůvodňovaných pravidel v testech vede buď k nekonečnému regresi testů, nebo k uzavření takové řady do sebezduodňujícího se kruhu. Laudan si je vědom takového nebezpečí a předpokládá, že je možné přerušit kruh nebo zastavit regres, pokud se nám podaří nalézt ospravedlňující metaprincip (metapravidlo), který by společně sdílely všechny metodologie. Jeho původní naturalistické pravidlo propojující cíle s prostředky jejich dosažení ho nakonec

² Shoda v přístupu k vědě a metodologii, již má být dosaženo na empirickém základě, ale není bezproblémová. Laudan sice tvrdí, že rozhodování o volbě metodologie je založeno čistě empiricky, stejně tak jako při volbě teorie (např. ve vědě). Problémem je skutečnost, že ani volba teorie ve vědě neprobíhá na čistě empirickém základě a že pravděpodobně volba teorie není empiricky rozhodnutelná. (Laudan se tomu možná později brání, když vysvětluje, že neříká „že pouze empirie“, ale že říká, že „tak jako ve vědě“.)

vede k odhalení a formulaci takového metametodologického pravidla R_1 (Laudan 1987, 25):

Pokud činnosti určitého typu m v minulosti konzistentně poskytovaly určité kognitivní cíle e a rivalitní činnosti n selhávaly, pak předpokládáme, že budoucí činnosti řídicí se pravidlem „je-li tvůj cíl e , měl bys dělat m' “, mnohem pravděpodobněji poskytnou tyto cíle než činnosti založené na pravidle „je-li tvůj cíl e , měl bys dělat n .“

Přestože sám Laudan není plně spokojen s takovým pravidlem a chápe ho jako prozatímní (Laudan 1987, 26), nelze se ubránit pocitu z určité trivializace takového metametodologického pravidla. Pravidlo nám neříká v podstatě nic jiného než to, že je-li obsahem naší zkušenosti jistý postup vedoucí k danému cíli, zatímco jiný postup selhává, měli bychom dát přednost onomu postupu, jenž cíl poskytuje. Takové pravidlo se utváří ve všech možných zkušenostních kontextech našeho zacházení se světem a metametodologické je na něm pouze to, že je zobecněno pro všechny instanciaci takových pravidel a jako takové je transponováno do metametodologické úrovně. Přestože se Laudan snaží zdůvodnit své metapříklad následujícími „dobrymi důvody“:

- (1) „je prokazatelně obecně předpokládáno mezi filozofy vědy, a je tedy slibným kvaziarchimedovským bodem“
- (2) „zdá se být vhodným pravidlem pro učení se ze zkušenosti“ (Laudan 1987, 26),

vede přijetí takového metapříkladu k následujícím pochybnostem. (1) přestože je nepochybně vazba mezi cíli a prostředky naturalistického charakteru v rámci jednotlivých metodologií, nedochází při transpozici takového pravidla do metapříkladu na úrovni metametodologie ke ztrátě naturalistického přístupu? Nezavádí tím Laudan nezrušitelné metapříklad, které je tak obecné, že nemůže být opuštěno, a vymyká se tak naturalistickému požadavku empirického zdůvodňování? Nestává se takové metapříklad nezávislým na empirické vědě a nepatří tak určitým způsobem přeci jen do naturalismem odmítané *první filozofie*? Laudan si takové námitky nepřipouští. Přiznává však v souladu s problémem indukce, že antiinduktivisté budou tvrdit, že zde není záruka, že minulé vazby mezi strategiemi výzkumu a kognitivními cíli budou zaručeny i v budoucnu (Losee 2004, 134). Kromě této potíže s induktivní strategií je zde ale ještě další obtíž. Vědecká metodologie, tak jak ji Laudan představuje, by se musela omezovat pro dosažení po-

žadovaných cílů výhradně pouze na ty prostředky, které se již ukázaly jako úspěšné z hlediska jejich dosažení. Tím by se však metodologická role vědy stala z velké části konzervativní silou, která by vyžadovala setrvání u těch prostředků, jež byly schopné cíle s určitým úspěchem poskytnout, a nikoli progresivní silou, jež by nabádala k hledání efektivnějších prostředků, které by cíle poskytovaly lépe, tj. jednodušeji, efektivněji, rychleji, přesněji, jistěji atp. Tato námitka však není úplně opodstatněná z následujícího důvodu.

Losee považuje Laudanovo označení „normativního naturalismu“ za opodstatněné vzhledem k tomu, že metapříklad R_1 je plně normativní i plně naturalistické. Za (1) je natolik naturalistické, nakolik zhodnocení metodologických pravidel obsahuje empirický výzkum korelací cíle-zprostředkujících vztahů (means-end relationships), a za (2) je normativní, protože jde o aplikaci nezrušitelného (nerevidovatelného) metametodologického principu, který říká, že je třeba přijmout právě ta pravidla, která empirický výzkum odhalil jako efektivní korelaci mezi prostředky a cíli (Losee 2004, 135). Díky tomu bychom mohli předpokládat, že právě empirický výzkum korelací zajišťuje efektivitu doporučených prostředků pro získání požadovaných kognitivních cílů, a je tak zárukou progresivity vědy. V takovém případě by ovšem za progresivní rozvoj vědy zodpovídaly jednotlivé rivalitní metodologie, v jejichž rámci dochází k formulaci jednotlivých metodologických pravidel, a progres jako možný cíl vědy by nebyl zajištěn na úrovni metametodologie v rámci jejích metametodologických pravidel. Je však podivné si myslet, že se všechny metodologie shodnou na pravidlu R_1 , ale neshodnou se již na progresu vědy jako možném cíli vědy. Jinými slovy hypotetický imperativ „je-li tvým cílem správná věda, pak je nutné přijmout pravidla, jež povedou k progresu vědění“, by nebyl pravidlem přijatelným pro všechny soupeřící metodologie. Zdá se však, že takové pravidlo je naopak přijatelné všemi konkurenčními metodologiemi organizace vědeckého výzkumu, a z toho vyplývá, že Laudanovo metapříklad R_1 může být sice normativní, ale neodpovídá metametodologii, kterou bychom se měli řídit. Je to pravidlo, jež nás vede a zavazuje k přijetí již osvědčených prostředků k dosažení cílů, a nikoliv pravidlo, jež by nás vedlo a zavazovalo k hledání progresivnějších prostředků pro dosahování požadovaných cílů.

I v tomto případě ale musíme čelit ošidným a svůdným apriorním předpokladům. Jedním z nich je apriorní přesvědčení, že věda musí být nutně progresivní. V takovém případě bychom zmiňované závě-

ry zdůvodňovali mimoempirickými či mimovědeckými (metafyzickými) předpoklady, a unikali bychom tak naturalistické pozici. Udržet ji, znamená v tomto případě, stejně tak jako to činí Laudan, pokud jde o cíle-zprostředkující vztahy, obrátit se z metametodologické pozice k jednotlivým metodologiím a empirickým šetřením zjistit, zda neexistují metodologie, které by nepředpokládaly progresivitu poznání. Přijmeme-li dosavadní výsledky takových zkoumání, zjistíme, že rozlišíme-li mezi empirickým a teoretickým progresem, pak zřejmě není pochyb o „empirickém růstu (progresu)“ vědeckých poznatků v čase, ale jsou zde zásadní pochybnosti, pokud jde o odpovídající „teoretický růst (progres)“³ (Losee 2004, 2). Takové zjištění empirického charakteru nás již může vést v rámci naturalistické pozice k požadavku vybudovat takovou metodologii, která by reflektovala empirický růst poznatků, a nehledě na problematičnost teoretického růstu formulovala odpovídající metodologická pravidla pro dosahování růstu. Na metametodologické úrovni se pak nelze spokojit s Laudanovým pravidlem R_1 , které sice progresivitě přímo neodporuje, ale explicitně ji na metametodologické úrovni nevyžaduje. Výše zmíněnou tezi o progresu bychom mohli nyní transformovat do metametodologické úrovně a tvrdit, že není pochyb o ‚empirickém růstu (progresu)‘ *metodologických* poznatků v čase, ale jsou zde zásadní pochybnosti, pokud jde o odpovídající ‚teoretický růst (progres)‘ v rámci konkurenčních metodologií.

Normativní naturalismus je tedy snahou uchovat preskriptivní roli filosofie vědy i v rámci naturalistické pozice. Předpoklady i důsledky takové snahy jsme se snažili ukázat v souvislosti s tím, že normativní naturalismus znovu nastoluje otázku metametodologie jako vyšší metodologické úrovně, z níž je možné poměřovat jednotlivé konkurenční metodologie. V tomto smyslu je zásadní otázkou slučitelnost metametodologického konceptu a naturalismu. Přestože jsme našli určité

³ Losee rozlišuje mezi dvěma aspekty a cíli vědeckého poznání: 1) poznání vlastností a vztahů různých substancí. 2) formulace teorií, které vysvětlují, proč substance mají tyto vlastnosti a vztahy. Předpokládá dále, že není pochyb o tom, že progresivně vzrůstá vědění prvního typu a proto víme stále více, ale pochybuje již o tom, zda to platí také o teoriích, jež mají vysvětlovat tyto poznatky. Oprávněnou námitkou jistě je, že vazba mezi „empirickým“ a „teoretickým“ je mlhavá a často je „empirické“ těsně vázáno, či určeno „teoretickými předpoklady.“ Lze-li však prokázat vzrůst vědění prvního typu, a to za podmínek libovolně komplikovaného vztahu mezi empirickým a teoretickým, pak to postačuje ke stanovení úkolu vybudování takové metodologie, jež by tento růst reflektovala.

nedostatky Laudanova řešení především v souvislosti s metametodologickým pravidlem R_1 , které považujeme sice za podstatné, ale nepokládáme je za natolik fatální, aby vedly k destrukci normativně naturalistické pozice. V závěru naší úvahy o slučitelnosti metametodologie a naturalismu se věnujeme ještě několika dalším možným námitkám proti konzistentnosti normativního naturalismu.

Losee (2004) tak například upozorňuje na to, že Laudan znovu zavádí justifikační hierarchii, která není v souladu s jeho síťovým modelem, který zavedl k tomu, aby vyřešil problémy Kuhnova modelu paradigm. Kuhnův model neumožňoval racionální vysvětlení přechodu mezi paradigmaty, protože docházelo k radikální a náhlé proměně na všech úrovních od empirické přes teoretickou až k hodnotové. Laudanův síťový (retikulační) model naopak předpokládal, že při přechodu k novému paradigmatu zůstanou některé z úrovní fixní a proměna nastane pouze ve zbývající úrovni. To podle Laudana umožňuje racionální vysvětlení přechodu od paradigmatu k paradigmatu. Důsledkem takového předpokladu je, že žádná ze zmiňovaných úrovní není privilegiovaná a každá má vzájemné vazby s ostatními úrovněmi. Původní Kuhnova představa justifikační hierarchie, od výchozí empirické přes teoretickou až k axiologické, je tak zrušena. Neexistuje tedy hierarchie cílů, ale retikulační (síťový) model vazeb mezi cíli vědecké komunity, které se s časem mění. Losee v tomto smyslu tvrdí, že Laudanův metametodologický induktivní princip R_1 není v souladu s jeho síťovým modelem, ale že se dostává opět na vrchol justifikační hierarchie (Losee 2004, 135). Navíc se zdá, že vzhledem k jiným korelacím mezi prostředky a cíli, jež mohou selhat (např. byly v jisté době produktivními korelacemi a později přestaly být), se na rozdíl od nich zdá být tento princip neodvolatelný a zahrnutí takového nezrušitelného (neodvolatelného), transhistorického principu do filosofie vědy se zdá být porušením naturalistického programu (tamtéž). Pokud vezmeme takovou námitku vážně, pak se Laudanovi daří uniknout nekonečnému regresi nebo kruhovosti justifikace pouze za cenu porušení principu naturalismu. Jedná se však skutečně o porušení naturalistického programu?

Laudanova forma naturalismu předpokládá, že všechny hodnotící principy by měly být empiricky zdůvodněny a měly by být také odvolatelné vzhledem k nějaké empirické evidenci. Žádné principy by tedy neměly být zavedeny jako neodvolatelné principy, věčné nebo předzkušnostní pravdy. Je Laudanovo pravidlo R_1 skutečně takovým neodvolatelným principem? Domnívám se, že nikoli, protože výhrady

tohoto typu spočívají jednak na nedostatečném rozlišení mezi metodologickou a metametodologickou úrovní a jednak na určité neujasněnosti, pokud jde o neodvolatelnost zmiňovaného principu.

Losee například navrhuje určitou modifikaci principu R_1 , která by vedla k tomu, že by se takový princip stal nezrušitelným principem, jehož aplikace by určovala správnou hodnotící praxi. Bohužel takový princip by platil pouze ex post a nikdo by nemohl vědět, zda jeho aplikace v konkrétní situaci bude nebo nebude produktivní (Losee 2004, 136). Modifikace spočívá v následujícím dodatku: „aplikuj metametodologický princip na všechny případy kromě těch, v kterých dříve úspěšné cíle-zprostředkující korelace přestaly být úspěšné“ (Losee 2004, 135). Domnívám se, že v tomto případě nejde ani tak o to, že by se takový metametodologický princip stal nezrušitelným (nerevidovatelným) principem, ale spíše univerzálně platným principem. Univerzálně platných pravd ale existuje či může být zkonstruováno nekonečné množství (například tautologické tvrzení typu „A platí pouze v těch případech, kdy platí a neplatí ve všech ostatních“) a není vůbec jasné, proč by takové pravdy měly být narušením naturalistického programu. Navíc je pravděpodobné, že by takový princip byl zrušen ne pro svou univerzální platnost, ale pro svou univerzální nepoužitelnost. Jako princip neříká o aktuální použitelnosti principu nic a je tak metodologicky prázdný. Tolik k oné neujasněnosti neodvolatelnosti. Dostatečné nerozlišování mezi metodologickou a metametodologickou úrovní spočívá v něčem jiném. V metodologické úrovni jde o konkrétní úspěšné či neúspěšné naplňování cíle-zprostředkujících vztahů, ale v metametodologické úrovni již jen o obecnou formu cíle-zprostředkujícího vztahu jako takového. Jednotlivé metodologické principy se tedy liší jak v cílech, tak v činnostech, jež je poskytují v rámci jednotlivých konkurenčních metodologií, ale metametodologické principy jsou již většinou tak univerzální, že platí napříč jednotlivými metodologiemi a jsou již tak také formulovány. Není to tedy neodvolatelnost či nezrušitelnost takových principů, ale jejich univerzálnost, která je činí obtížněji falsifikovatelnými. Laudanovo pravidlo R_1 sdílí všechny metodologie, protože je zformulováno jako obecná souvislost mezi dosahováním cílů a činnostmi, jež je poskytují. V takto obecné formě je takový princip či pravidlo srovnatelný např. s vazbou mezi účinky a příčinami. Tvrzení, že „účinky nepředchází své příčiny“, má také univerzální formu a je přijatelné všemi metodologiemi. Není to však nezrušitelný princip či pravidlo. Pokud by empirický výzkum odhalil procesy, v kterých by

docházelo k narušení takového pravidla, pak by takové pravidlo muselo být revidováno.⁴ Konečně sám Laudan, jak jsem již uvedl, říká o cíle-zprostředkujících vztazích, a díky tomu zprostředkovaně i o pravidlu R_1 , že mají podobu statistického zákona (Laudan 1987, 25). Užití takových pravidel tedy podléhá plně všem kladům i záporům statistiky, a stejně tak i všem nedostatkům a nedokonalostem induktivní inference jako takové. Domnívám se tedy, že problém nespočívá v nezrušitelnosti takového pravidla, a že tedy takové pravidlo nenarušuje naturalistický program a není s ním neslučitelné. Na druhé straně považuji za mnohem palčivější problém jeho poněkud konzervativnější pojetí, jak jsem se snažil naznačit výše.

To však není jediná výhrada vůči Laudanovu normativně naturalistickému programu. Pro vhodnou ilustraci dalších možných obtíží, kterým by normativní naturalismus Laudanova typu musel čelit, si dovoluji zformulovat několik „svých“ hypotetických imperativů (striktně dle reformulačních pravidel, jaká Laudan narhuje). Vyjděme z Laudanova původního příkladu hypotetického imperativu reformulovaného z Popperova standardního pravidla o *ad hoc* hypotézách:

- (1) „Je-li cílem vyvinout riskantní hypotézu, pak bychom se měli vyvarovat *ad hoc* hypotéz.“

Uvažme nyní následující metodologické (hypotetické) imperativy:

- (2) „Je-li cílem vyvinout riskantní hypotézu, pak bychom měli připustit *ad hoc* hypotézy pouze v případě, že zvyšují její falsifikovatelnost.“
- (3) „Je-li cílem vyrovnat se s anomáliemi v rámci stávající teorie, je lepší uchýlit se k *ad hoc* hypotéze než ztratit dosud platnou teorii.“

(2) je také reformulované Popperovo pravidlo o *ad hoc* hypotézách, ke kterému dospěl v diskusi s Grünbaumem (Popper 1959, 50).⁵ (3) je reformulované a obecně přijatelné pravidlo s mnoha konkrétními his-

⁴ Některé interpretace kvantových jevů používají kauzální působení z budoucnosti a předpokládají tak, že účinek může předcházet svou příčinu.

⁵ Pro úplnost: Grünbaum následně ostře kritizuje Popperovo přesvědčení, že je možné uvažovat o *ad hoc* hypotézách ve škále, která je inverzně vztažena ke stupňům testovatelnosti (Grünbaum 1976, 6-9). Podrobněji viz Havlík (2008).

torickými případy. Všechny hypotetické imperativy nás normativně zavazují k činnostem a metodologickým postupům vzhledem k *ad hoc* hypotézám. (1) a (2) sledují stejné cíle prostřednictvím vzájemně výlučných strategií. (3) naopak vede k uchování stávající teorie za cenu připuštění *ad hoc* hypotézy. Všechny strategie ale mohou vést v daném historickém případě ke kladnému či zápornému výsledku. Jak v takovém případě chápat statistický zákon spojující činnosti s cíli? Řekněme, že bychom na základě historické analýzy zjistili, že úspěšnost hypotetických imperativů (1), (2) a (3) je pouze 33,33% pro každý z imperativů. Jaká strategie ohledně *ad hoc* hypotéz by měla být přijata do normativně naturalistické metodologie?

V závěru se ještě vrátím k otázce nezrušitelných principů. Snažil jsem se ukázat, že Laudanovo pravidlo R_1 není nezrušitelným principem, a nic tedy nebrání jeho přijetí do normativního naturalismu. Na druhé straně to neznamená, že nezrušitelná pravidla nemohou být zformulována. Opět pro vhodnou ilustraci zformuluji následující pravidla, která jsou již pravidly na úrovni metametodologie.

(Havlíkovy) metametodologické imperativy:

- (1) „Je-li cílem dosáhnout daných kognitivních cílů, pak bychom měli použít těch prostředků, které vedou k jejich dosažení, než jiných, které jejich dosažení nezaručují.“
- (2) „Je-li cílem formulace metodologických pravidel, pak bychom je měli formulovat tak, aby jejich aplikace nebyla v rozporu s minulou a stávající vědeckou praxí a aby proti jejich budoucí aplikaci nebyly známy žádné obtíže.“

(1) i (2) lze považovat za celkem dobré příklady nezrušitelných či neodvolatelných metametodologických pravidel. Obě pravidla mohou být jen těžko odvolána vzhledem ke způsobu jejich formulace a jejich zdůvodnění. Jsou nezpochybnitelná, ale banální a metodologicky prázdná. Taková situace je inspirací pro zformulování mého posledního metametodologického principu:

Havlíkův metametodologický princip:

- (HMP) „Je-li cílem formulace pravidel v rámci naturalistického programu, pak bychom je měli formulovat tak, aby neměla tvar nezrušitelných principů.“

(HMP) má tvar Laudanova hypotetického imperativu na metametodologické úrovni. Říká nám, jak postupovat, je-li cílem formulace pravidel přijatelných normativně naturalistickou metodologií. Avšak samo HMP je pravidlem, které zřejmě nějak patří do normativně naturalistické metodologie, čímž se dostává do rozporu samo se sebou, protože je formulováno nezrušitelným způsobem a je tak neodvolatelné. Dostáváme tak situaci obdobnou aplikaci verifikačního kritéria logických pozitivistů sama na sebe⁶, či některých dalších známých paradoxů založených na seberefenci. Domnívám se nicméně, že pro normativní naturalismus představují uvedená pravidla sice určité obtíže, nicméně nejsou pro samotný program nijak fatální. Vychází-li naturalistický projekt z jednoty vědy a epistemologie, pak heterogenita a nedokonalost vědy svědčí nějak i o heterogenitě a nedokonalosti epistemologie a metodologie. Jak poznamenává sám Laudan (1996, 221): „Zřejmá epistemická heterogenita činností a přesvědčení obvykle považovaných za vědecké by nás měly varovat, že je pravděpodobně marné hledat epistemickou verzi demarkačního kritéria.“ Možná se musíme vzdát nejen myšlenky ideální vědy, ale také myšlenky ideální metodologie.

Filosofický ústav AV ČR
 Jiřská 1
 Praha 1 110 00
 Česká republika
 havlik@flu.cas.cz

Literatura

- FREEDMAN, K. (1999): Laudan's Naturalistic Axiology. *Philosophy of Science* 66 (3), 537.
- GRÜNBAUM, A. (1976): Can a theory answer more questions than one of its rivals? *The British Journal of the Philosophy of Science*, No. 27.
- HAVLÍK, V. (2008): Status Ad hoc hypotéz ve vědě. In: Zouhar, M. (ed.): *Jednotliviny, všeobecniny, významy*. Bratislava: Filozofický ústav SAV, 94-110.
- LAKATOS, I. (1970): History of Science and Its Rational Reconstructions. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, Vol. 1970, 91-136.
- LAUDAN, L. (1987): Progress or Rationality? The Prospects for Normative Naturalism. *American Philosophical Quarterly*, Vol. 24, No. 1.

⁶ Známa kritika pozitivismu R. Ingardena.

- LAUDAN, L. (1990): Normative Naturalism. *Philosophy of Science*, Vol. 57, No. 1, 44-59.
- LAUDAN, L. (1996): *Beyond Positivism and Relativism, Theory, Method and Evidence*. Boulder, CO: Westview Press.
- LOSEE, J. (2004): *Theories of Scientific Progress. An introduction*. New York: Routledge.
- MADDY, P. (2007): *Second Philosophy. A Naturalistic Method*. Oxford: Oxford University Press.
- POPPER, K. R. (1959): Testability and ‚ad hocness‘ of the contraction hypotheses. *The British Journal of the Philosophy of Science*, No. 10.
- QUINE, W. V. O. (1969): *Ontological Relativity and Other Essays*. New York: Columbia University Press.
- QUINE, W. V. O. (1981): *Theories and Things*. Cambridge: The Belknap Press of Harvard University Press.
- QUINE, W. V. O. ([1990] 1994): *Hledání pravdy*. Praha: Herrmann a synové.
- QUINE, W. V. O. ([1995] 2002): *Od stimulu k vědě*. Praha: Filosofia.
- SHER, G. (1999): Is There a Place for Philosophy in Quine’s Theory? *The Journal of Philosophy* 96, No. 10, 491-524.

Existenčné čítanie oslích viet v dynamických logikách¹

Miloš Kosterec

Univerzita Komenského v Bratislave

Abstract: In this paper, two proposals for modification of two logical systems, namely Dynamic Predicate Calculus and Dynamic Semantics with Choice Functions, are introduced. These two systems do not grasp adequately the anaphoric bond in so-called donkey sentences. The aim of the proposal is to eliminate this limitation. First, the phenomenon of anaphor is introduced. Then, the author articulates the presuppositions for adequate grasping of this phenomenon for any semantic system. After that, the paper aims at particular problem of analysis of anaphor in donkey sentences. The analysis using two systems is demonstrated. The paper contains presentation of those parts of the two systems, which are relevant to this problem. After this, the author sheds light on problems connected with the weak reading and some limits of two systems in the issue. Conditions for adequate analysis of the weak reading are articulated. Then, the author demonstrates an analysis of the weak reading using the modified systems. The paper ends in discussion about the relation between the original and modified systems.

Keywords: anaphor, anaphoric bond, context, valuation, choice function, donkey sentences, weak reading.

1 Anafora

O anaforicky použitom výraze hovoríme, keď svoj význam nadobúda vzhľadom na predchádzajúci jazykový kontext. Anaforicky použitý

¹ Táto štúdia vznikla na Katedre logiky a metodológie vied Filozofickej fakulty UK v Bratislave v rámci projektu podporeného grantom VEGA č. 1/0046/11, *Sémantické modely, ich explanačná sila a aplikácia*. Ďakujem Mariánovi Zouharovi za kritické pripomienky k predchádzajúcim verziám článku.

výraz nadobúda svoj význam vďaka sémantickej väzbe na časť predchádzajúceho textu.² Líši sa v tom od tzv. deikticky použitého výrazu, pre ktorý sa to, na čo sa výraz vzťahuje, určuje až vzhľadom na situáciu prehovoru. Ako anaforické môžu vystupovať výrazy rôznych typov. Najčastejšie anaforicky používame zámená a deskripcie. Časť textu, vzhľadom na ktorú nadobúda anaforicky použitý výraz svoj význam, nazývame antecedent. Antecedentom môže byť rovnako množstvo rôznych typov výrazov ako individuové výrazy, všeobecné termíny a aj celé vety. Výsledkom sémantickej previazanosti medzi anaforicky použitým výrazom a jeho antecedentom je v prípade referujúcich výrazov koreferencia daných dvoch výrazov. Všetky tieto vlastnosti anaforicky použitého výrazu by mali byť zreteľné v adekvátnej sémantickej analýze tohto fenoménu.

Adekvátna sémantická analýza anafory musí uchopiť anaforickú väzbu³ medzi antecedentom a anaforicky použitým výrazom. V analýze významu anaforicky použitého výrazu musí byť zrejmé, že svoj význam nadobúda vzhľadom na predchádzajúci kontext.

2 Problematika oslích viet

Anaforická väzba sa vyskytuje v mnohých podobách, ktoré predstavujú široký obor pre aplikáciu prostriedkov sémantickej analýzy. V tomto článku sa obmedzím na problematiku sémantickej analýzy anaforickej väzby v tzv. oslích vetách (donkey sentences). Príklad, ktorý uviedol daný problém, je nasledovný:

(1) Každý farmár, ktorý vlastní nejakého oslíka, ho bije.

Oslia veta má charakter kvantifikovanej implikácie. Antecedent anaforicky použitého výrazu – *nejaký oslík* – sa nachádza v antecedente implikácie, avšak vo vedľajšej vete. Jeho dosah je obmedzený iba na určitú časť vety 1. Anaforicky použitý výraz – *ho* – sa nachádza v kon-

² Sémantickú povahu anafory obhaja napríklad Duží (2006). Jej článok nadviazal na diskusiu, ktorá prebehla v textoch Gahér (2002), Zouhar (2004). Podstatnú prácu v sémantickej analýze anafory predstavuje kapitola *Anaphora and meaning* v práci Duží – Jespersen – Materna (2010).

³ Z tejto pozície kritizuje Zouhar (2004) koreferenčnú teóriu.

zekvente danej implikácie. Nie je to však jediný prípad anaforickej väzby vo vete 1. Druhú dvojicu tvoria *každý farmár a on*⁴.

Oslie vety predstavujú problém pre tie sémantické teórie anaforickej väzby, ktoré sa pokúšajú reprezentovať význam anaforicke použitého výrazu pomocou viazanej premennej. Z hľadiska daných teórií je antecedentom kvantifikačný výraz a anaforicnú väzbu tento prístup umožňuje uchopiť nanajvýš pomocou výskytu premennej, ktorá by mala byť viazaná antecedentom anaforickeho výrazu. Kým pre dvojicu *každý farmár a on* takto môžeme anaforicnú väzbu opísať, pri druhej dvojici je to problém. Reprezentácia významu výrazu *ho* sa totiž nachádza mimo dosahu kvantifikátora, ktorý je súčasťou reprezentácie významu antecedenta. Z tohto dôvodu bola odmietnutá napríklad sémantická analýza daného javu pomocou predikátovej logiky.

3 Oslie vety a dynamické logiky

Problematika oslích viet bola jedným z podnetov pre aplikáciu dynamických logík na poli sémantickej analýzy prirodzeného jazyka. Spoločnou črtou dynamických logík je úsilie o opis vplyvu významu výrazov na jazykový kontext⁵. Vo všeobecnosti môžu nové tvrdenia kontext obohatiť, pozmeniť či zúžiť. Z tohto pohľadu je zrejmejšie, prečo sa extenzia významu jazykových výrazov⁶ v dynamických logikách často reprezentuje binárnou reláciou. Prvé členy v relácii zastupujú predchádzajúci jazykový kontext a druhé sú prejavom vplyvu významu aktuálneho tvrdenia na tento kontext. Opis významu formúl v dynamickej logike potom práve zahŕňa vysvetlenie, ako ovplyvňujú výrazy vo formulách výber prvkov v danej relácii⁷. Rôzne dynamické logiky sa líšia v tom, aké typy objektov môžu vystupovať ako prvky relácií reprezentujúcich zmenu kontextu, či spôsobom výberu prvkov do

⁴ V slovenčine je *on* v konzekvente danej implikácie zmlčané.

⁵ Dynamické logiky sa líšia v tom, ako modelujú jazykový kontext. Väčšinou má však podobu usporiadanej množiny významných prvkov, ktorými je daný kontext tvorený. Ovplyvňovanie jazykového kontextu v tomto prípade znamená ovplyvňovanie výberu prvkov do danej množiny. Nemení sa kontext v tom zmysle, že by sa menili slová v ňom použité.

⁶ Korektnejšie by som mal hovoriť o formulách. Nie každému výrazu prislúcha relácia.

⁷ Technickejšie povedané, význam výrazov je reprezentovaný pomocou výberového kritéria na členy relácie.

daných relácii, ktoré pripisujú jednotlivým formulám. V nasledujúcom texte sa venujem dvom dynamickým logikám: dynamickej predikátovej logike (ďalej len DPL) a dynamickej logike s výberovými funkciami (ďalej len DLV).

4 Oslie vety v DPL

DPL vznikla pri úsilí o rozšírenie predikátovej logiky s ambíciou uchopiť dynamickú stránku významu jazykových výrazov. Kým slovník DPL a PL je prakticky totožný, líšia sa v sémantickej interpretácii formúl z jazyka. Sémantická interpretácia formúl v DPL im neprisuduje pravdivostné hodnoty, ale relácie na valuáciach. Spomenul som, že sa dynamické logiky líšia o.i. výberom typov objektov do relácií, ktoré pripisujú formulám. Jazykový kontext je z pohľadu DPL opísaný množinou prípustných valuácií. Východiskom je jednoduchá intuícia. DPL je prvorádový systém a ako taký obsahuje iba individuové premenné. Význam individuových konštánt, či interpretácia predikátov sa samozrejme kontextom nemení. Čo sa však môže meniť, je ohodnotenie premenných. Niektoré ohodnotenia premenných sú vzhľadom na interpretáciu predikátov a individuových konštánt (ktorá funguje rovnako ako v predikátovej logike) adekvátne a niektoré nie. Ak sa teda v dynamickej logike môže meniť jazykový kontext, tak len vzhľadom na výber valuácií.

Konkrétny výber valuácií reprezentuje informáciu dodanú jazykovým kontextom. Formuly nadväzujúce na kontext ho môžu, ako som už spomenul, meniť rôznym spôsobom. Napríklad jednoduchá formula:

$$P_x$$

predstavuje výber tých valuácií, ktoré pripisujú premennej x individuum, ktoré má vlastnosť P . Výberové kritérium však nie je aplikované na všetky valuácie⁸, ale len na tie, ktoré boli dodané doteraz⁹. Väčšina typov formúl v DPL sa správa týmto spôsobom. Hovorí sa, že majú testový charakter, pretože testujú doterajší jazykový kontext pomocou

⁸ Okrem prípadu, keď je formula prvá v kontexte.

⁹ V situácii, že daná formula je v kontexte prvou, tak predstavuje výber zo všetkých valuácií. Predchádzajúcim kontextom neboli žiadne vylúčené a všetky sú takto prípustné.

nového kritéria. Pomocou testových formúl sa jazykový kontext môže iba zužovať. Na druhej strane formuly ako:

$$\exists x Px$$

fungujú ako obohatenie kontextu. Formula s existenčným kvantifikátorom vlastne iba tvrdí, že existuje nejaký prvok, ktorý má vlastnosť P. Existujú teda valuácie, ktoré pripisujú premennej x vhodné individuum. Existenčne kvantifikovaná formula nimi obohacuje kontext a výber neobmedzuje len na už predtým prípustné valuácie.

Pre problém oslích viet ešte dodám, ako sa v DPL narába s dosahom kvantifikátorov. Kým v predikátovej logike bol dosah určený použitím zátvoriek, v DPL je ovplyvnený použitím logických spojok. Niektoré (napríklad negácia) pre dosah kvantifikátora predstavujú bariéru a niektoré nie. Z hľadiska oslích viet mi postačí upozorniť, že implikácia nefunguje ako takáto bariéra. Existenčný kvantifikátor v antecedente môže viazať premennú nachádzajúcu sa v konzekvente. Viac o implikácii v DPL poviem neskôr.

Reprezentácia významu oslej vety 1 pomocou DPL vyzerá nasledovne:

$$(1a) \quad \forall x((Fx \wedge \exists y (Oy \wedge V(x,y))) \rightarrow B(x,y))^{10},$$

kde: F – byť farmárom, O – byť oslom, V- vlastniť, B – biť.

Anaforická väzba medzi výrazmi *nejaký oslík* a *ho* je opísaná pomocou viazaného výskytu premennej y pri predikáte B. Reprezentáciu významu anaforicky použitého výrazu *ho* predstavuje tento výskyt premennej y a anaforickú väzbu stelesňuje viazanosť tohto výskytu premennej y existenčným kvantifikátorom, ktorý je časťou reprezentácie významu výrazu *nejaký oslík*.

5 Oslie vety v DLV

DLV¹¹ predstavuje konkurenčný prístup k DPL v otázke sémantickej analýzy anafory. Kým v DPL je sémantika anafory založená na inštitúte viazanej premennej, DLV sa pokúša opísať anaforicky použitý výraz pomocou konkrétneho typu funkcie. Sémantická interpretácia

¹⁰ Tento typ prepisu podobných viet ponúka aj Groenendijk; pozri Groenendijk (1989, 11).

¹¹ Peregrin – von Heusinger (2003).

formúl v DLV rovnako ako v DPL pripisuje formulám relácie. Interpretácia predikátov a individuových konštánt je prakticky rovnaká. Spomeniem však dva podstatné rozdiely: prvkami relácií (a teda kontextom) nie sú valuácie premenných a kontext ovplyvňujú aj časti formúl, nielen formuly samé.

Prvý rozdiel pramení v odmietnutí použitia individuových premenných autormi systému DLV¹². Namiesto nich je zavedený inštitút výberovej funkcie. Konkrétne výberová funkcia jednoducho vyberá predikátu jedného predstaviteľa z množiny, ktorú mu prisúdila interpretačná funkcia. Výberové funkcie vyberajú individuum¹³ pre každý predikát. Navzájom sa líšia nanajvyš výberom individuí. Práve výberové funkcie sú však prvkami relácií, ktoré pripisuje sémantická interpretácia nielen formulám. Podľa tohto prístupu je kontext tvorený súborom v ňom spomenutých individuí. Výber významných individuí je zabezpečený výberovými funkciami. Kým v DPL bol predchádzajúci kontext reprezentovaný prípustnými ohodnoteniami premenných, v DLV ho stelesňujú prípustné výberové funkcie, ktoré predstavujú výber v kontexte významných individuí.

Druhý rozdiel sa zakladá na presvedčení, že kontext (teda výber individuí v ňom spomenutých, či podľa neho prípustných) ovplyvňujú už časti formúl. Pre mňa sú podstatné dva takéto prvky. Prvým je funkcia *a*. Má fungovať ako formálny náprotivok neurčitého člena¹⁴. Funkcia *a* sa aplikuje na predikáty. Výsledkom jej aplikácie je výber množiny výberových funkcií, ktoré sa od predchádzajúceho kontextu líšia nanajvyš tým, aké individuum vyberajú ako reprezentanta pre daný predikát. Funkcia *a* je podobná existenčnému kvantifikátoru DPL v tom, že môže obohatiť kontext aj o predtým nespomenuté možnosti. Druhou podstatnou funkciou je *the*. Ako náprotivok funkcie určitého člena (definite article) má za úlohu upriamiť pozornosť na už predtým spomenuté individua. Pri aplikácii funkcie *the* na predikát sa kontext nemení, skôr preberá. Zmena relácie, ktorú použitie funkcie *the* predstavuje, je vskutku triviálna. Funkcia *the* totiž kontext neobohacuje a ani nezužuje. Jej úlohou je naopak upriamiť pozornosť na už

¹² Pre možné dôvody pozri napríklad prácu Peregrin (2000b).

¹³ V podobe Peregrin - von Heusinger (2003) je DLV tiež prvorádovým systémom a preto môžem hovoriť o individuíach.

¹⁴ V angličtine, pričom v slovenskom jazyku sa často nahrádza výrazom *nejaký*.

predtým spomenuté indivíduá. Funkcia *the* nemení výber výberových funkcií. Vzhľadom na sémantickú analýzu anafory v systéme DLV je práve funkcia *the* najpodstatnejšia. Význam zámen je totiž opísaný práve pomocou nej a zámená často (ako aj v prípade vety 1) používame anaforicky.

Pre správne porozumenie formálnemu prepisu vety 1 v DLV dodám, ako je v tomto systéme reprezentovaný význam výrazu *každý*. V DLV mu zodpovedá funkcia *every*. Spája do formuly dva podvýrazy, pričom kontext nimi predstavovaný ovplyvňuje nasledovne. V kontexte ostane iba také indivíduá, pre ktoré platí, že ak sú časťami kontextu prvého výrazu, tak sú aj členmi kontextu druhého výrazu¹⁵.

Prepis vety 1 pomocou DLV:

(1b) *every*(vlastní(*a*(farmár),*a*(oslík)),*bije*(*he*,*it*))¹⁶

Význam antecedentného výrazu *nejaký oslík* je reprezentovaný pomocou *a*(oslík). Anaforickému výrazu *ho* zodpovedá *it*. Anaforická väzba je opísaná nasledovne. Antecedent kontext obohatí o také indivíduá, ktoré sú oslíkom (spomenutá funkcia *a* aplikovaná na predikát *oslík*) a zároveň ich vlastní *nejaký* farmár. Anaforicky použitý výraz potom vyberá z kontextu predtým spomenuté možnosti (zámeno *it* je prípadom funkcie *the*, pričom sa vzťahuje na veci, nie na osoby). Zámeno *it* sa takto vzťahuje na každé indivíduum, ktoré je oslíkom a je súčasťou predchádzajúceho kontextu.

6 Je všetko v poriadku?

Keby bol môj článok rekapituláciou, na tomto mieste by som skončil. Tvrdím však, že prepisy vety 1 ponúknuté v systémoch DPL a DLV nie sú *úplne* adekvátne. Veta 1 má totiž zrejme viacero možných čítaní. Vo všeobecnosti boli rozlíšené a pomenované dve rôzne čítania: všeobecné a existenčné (strong and weak reading). Všeobecné čítanie vety 1 vyjadruje parafráza:

(1c) Každý farmár, ktorý vlastní nejakého oslíka, bije každého oslíka, ktorého vlastní.

¹⁵ Aj autori DLV píšu, že funkcia *every*(A,B) je veľmi podobná $\forall x(Ax \rightarrow Bx)$. Pozri Peregrin - von Heusinger (2003, 9).

¹⁶ Peregrin - von Heusinger (2003, 16). Originál obsahuje anglické výrazy.

Existenčné čítanie vety 1 vyjadruje parafráza:

- (1d) Každý farmár, ktorý vlastní nejakého oslíka, bije aspoň jedného oslíka, ktorého vlastní.

Rozdiel medzi vetami 1c a 1d je v pravdivostných podmienkach. Vety, ktoré sa líšia pravdivostnými podmienkami, sa musia líšiť aj sémantickou reprezentáciou ich obsahu. Nieкто by mohol položiť otázku o oprávnenosti pripisovania spomenutých čítaní vete 1. Jazykovú evidenciu pre toto tvrdenie som získal z príkladov, ktoré sa objavili v diskusii o existencii možnosti viacerých čítaní oslích viet¹⁷. Príklady viet, ktoré čítame skôr existenčne:

- (2) Každý otec, ktorý má nejakého syna, ho nechá šoférovať do mesta.
 (3) Každý muž, ktorý má nejaký cent, ho dá do automatu na parkovanie.
 (4) Každý muž, ktorý má nejakú kreditnú kartu, ňou zaplatí účet.

Tieto vety sa na jazykovej úrovni veľmi podobajú vete 1. Formálna reprezentácia ich obsahu by však viedla rovnako ako v prípade vety 1 k neprirodzenému čítaniu daných viet. Tvrdím totiž, že prepisy 1a a 1b reprezentujú iba všeobecné čítanie vety 1. Vety 2, 3 a 4 však skôr čítame existenčne. Napriek tomu majú aj všeobecné čítanie. To je možno menej prirodzené¹⁸, ale napriek tomu je to logicky možné. Rovnako je možné existenčné čítanie vety 1. Adekvátna sémantická analýza viet 1 až 4 by však podľa toho mala poskytnúť všetkých 8 možností (2 pre každú z viet) a nielen 4.¹⁹ Je teda možné popísať existenčné čítanie vety 1 v systémoch DPL a DLV v ich súčasnej podobe? Tvrdím, že ak aj áno, tak len za cenu straty adekvátnej sémantickej reprezentácie anaforickeho výrazu a anaforickej väzby. V nasledujúcom texte podám dôvody svojich tvrdení.

¹⁷ Pozri napríklad Chierchia (1990), Schubert – Pelletier (1989).

¹⁸ Kanazawa (1992) sa pokúšal identifikovať čítania viet prostredníctvom obsiahnutých kvantifikátorov.

¹⁹ Netvrdím však, že by mala podať aj kritériá ich výskytu. To je skôr pragmatická otázka.

7 Jednoduchý model

Pre jednoduchšie porozumenie nasledujúcemu textu opíšem jednoduchý model, v ktorom budem demonštrovať svoje tvrdenia. Predstavme si veľmi jednoduchý svet, v ktorom sa nachádza len jeden farmár, ktorý vlastní iba dvoch oslíkov. Z času na čas zbije iba jedného z nich a vždy toho istého. Formálne môžem danú situáciu opísať nasledovne:

Univerzum = {a, b, c},
 a - Farmár,
 b - oslík,
 c - oslík,
 Vzťah vlastníctva: V = {<a,b>, <a,c>},
 Vzťah bitia: B = {<a,b>}.

Valuácie premenných (uvediem iba zaujímavé pre prípad vety 1):

$v1(x) = a, v1(y) = b$
 $v2(x) = a, v2(y) = c$

Výberové funkcie (uvediem iba zaujímavé pre prípad vety 1):

$e1(\text{farmár}) = a, e1(\text{oslík}) = b$
 $e2(\text{farmár}) = a, e2(\text{oslík}) = c$

Prepis 1a reprezentuje všeobecné čítanie kvôli sémantickej definícii implikácie:

$$[[\varphi \rightarrow \psi]] = \{ \langle g, h \rangle; h = g \ \& \ \forall k: \langle h, k \rangle \in [[\varphi]] \Rightarrow \exists j: \langle k, j \rangle \in [[\psi]] \}^{20}$$

kde g, h, k, j sú valuácie; $[[\varphi]]$ je sémantický obsah formuly φ .

Menej technicky povedané, význam implikácie v DPL je daný reláciou, do ktorej patria len také valuácie, pre ktoré zároveň platia tieto podmienky:

- a) $h = g$. Takúto podmienku obsahujú všetky formuly, ktoré majú testový charakter. Znamená to, že ide iba o výber adekvátnych možností z predchádzajúceho kontextu. Predchádzajúci kontext sa teda nanajvýš zúži. Každý valuácii, ktorá vyhoví podmienke obsiahnutej vo formule s testovým charakterom, sa priradí táto valuácia sama.

²⁰ Groenendijk (1989, 14).

- b) $\forall k: \langle h, k \rangle \in [[\varphi]] \Rightarrow \exists j: \langle k, j \rangle \in [[\psi]]$. Daná podmienka vlastne iba hovorí to, že na každú zmenu kontextu danú antecedentom implikácie ($\forall k: \langle h, k \rangle \in [[\varphi]]$, pričom k predstavuje možnú zmenu) musí byť schopný nadviazať konzekvent implikácie ($\Rightarrow \exists j: \langle k, j \rangle \in [[\psi]]$, k prepája antecedent a konzekvent). Inými slovami, v každom kontexte, ktorý je ovplyvnený antecedentom, musí byť interpretovateľný aj konzekvent.

Teraz ukážem, ako funguje prepis 1a pre môj model.

$$(1a) \quad \forall x((Fx \wedge \exists y (Oy \wedge V(x,y))) \rightarrow B(x,y))$$

pričom: F – byť farmárom, O – byť oslom, V – vlastniť, B – biť. Všeobecný kvantifikátor kladie podmienku, že v nasledujúcom kontexte budú iba tie valuácie, pre ktoré platí kvantifikovaná formula. Keďže premenná x sa môže vzťahovať na farmárov, predstavuje všeobecná kvantifikácia intuíciu, že podformula

$$(Fx \wedge \exists y (Oy \wedge V(x,y))) \rightarrow B(x,y)$$

musí platiť pre všetkých farmárov. Táto podformula má tvar implikácie. Hoci antecedentu vyhovujú obe valuácie $v1$ aj $v2$, význam konzekventu spĺňa iba $v1$. Podľa definície implikácie potom nebude valuácia $v2$ patriť do kontextu, ktorý bude výsledkom analýzy vety 1. Z pohľadu DPL je formula 1a v mojom modeli pri valuácii $v2$ nepravdivá. Formula 1a teda predstavuje iba všeobecné čítanie.

Prepis 1b taktiež reprezentuje iba všeobecné čítanie. Príčinou je sémantická definícia operátora *every*:

$$\begin{aligned} || \text{every}(S1,S2) || &= \{ \langle e, e' \rangle, e=e' \text{ \& pre každú } e1, \text{ ak } \langle e, e1 \rangle \in \\ || S1 ||, &\text{ tak potom existuje nejaká } e2 \text{ taká, že } \langle e1, e2 \rangle \in || S2 || \}^{21}; \\ || Q || &\text{ predstavuje sémantickú interpretáciu } Q; e, e' \text{ sú výberové} \\ &\text{ funkcie.} \end{aligned}$$

Táto definícia predstavuje dve podmienky, ktoré musia byť zároveň splnené:

- a) $e=e'$, použitie operátora *every* má teda testový charakter. Z predchádzajúceho kontextu vyberie iba vyhovujúce výberové funkcie. Vyhovujúcim výberovým funkciám priradí ich samé. Kontext teda použitie tohto operátora nanajvýš zúži.

²¹ Peregrin – von Heusinger (2003, 8).

- b) pre každú e_1 , ak $\langle e, e_1 \rangle \in \epsilon \mid \mid S_1 \mid \mid$, tak potom existuje nejaká e_2 taká, že $\langle e_1, e_2 \rangle \in \epsilon \mid \mid S_2 \mid \mid$. Pre každý kontext zodpovedajúci S_1 musí existovať interpretácia S_2 . Výberová funkcia e_1 prepája S_1 a S_2 . Ide prakticky o rovnakú podmienku ako pri dynamickej implikácii. Z popisu podmienok pre operátor *every* a dynamickejšiu implikáciu môžeme vidieť, že ich funkcia je prakticky rovnaká.

Ako funguje prepis 1b pre môj model?

(1b) *every*(*vlastní*(*a*(farmár),*a*(oslík)),*bije*(*he*,*it*))

Formula S_1 v tomto prípade: *vlastní*(*a*(farmár),*a*(oslík))

Pre zjednodušenie poviem, že daná formula ponechá v kontexte valúacie e_1 aj e_2 . Individuá nimi priradené predikátom *farmár* a *oslík* sú vo vzťahu *vlastní*. Vzhľadom na podmienku b) by mali obe výberové funkcie patriť do kontextu určeného frázou S_2 :

bije(*he*,*it*)

Výrazom *he* a *it* zodpovedá funkcia *the*. Tá nemení, ale iba preberá výberové funkcie z predošlého kontextu. Kým však výberová funkcia e_1 vyberá individuá, ktoré sú vo vzťahu *bije*, e_2 to nerobí. Výberová funkcia e_2 z daného dôvodu nebude patriť do reprezentácie významu priradenej vete 1. Hoci patrí do kontextu S_1 , nepatrí do kontextu S_2 . Kvôli použitiu operátora *every* teda nepatrí do kontextu obsahujúceho vetu 1. Formula 1b vylučuje výberovú funkciu e_2 ako prípustnú. Predstavuje teda všeobecné čítanie.

8 A čo existenčné čítanie?

Na tomto mieste sa ponúka otázka, či môžem prostriedkami DPL a DLV opísať existenčné čítanie vety 1. Áno, môžem. Napríklad nasledujúcimi prepismi:

(1e) $\forall x ((Fx \wedge \exists y (Oy \wedge V(x,y))) \rightarrow \exists y (Oy \wedge V(x,y) \wedge B(x,y)))$

pričom: F - byť farmárom, O - byť oslom, V- vlastniť, B - biť.

(1f) *every*(*vlastní*(*a*(farmár),*a*(oslík)),*vlastní*(*a*(farmár),*a*(oslík)) & *bije*(*he*,*it*))

Keď však zvážime použitie daných prepisov pri sémantickej analýze anaforicky použitých výrazov, narazíme na problém. V prepisoch 1a

a 1b bol jednoznačne identifikovateľný antecedent anaforického výrazu. V prepisoch 1e a 1f to je problém. Antecedentom anaforicky použitého výrazu je *ten*, ku ktorému má anaforickú väzbu. Viazaný výskyt premennej *y* pri predikáte *B* v prepise 1e je viazaný kvantifikátorom, ktorý sa nachádza v konzekvente implikácie. To je problém, pretože veta 1, ako podmienkové tvrdenie, zjavne obsahuje anaforický antecedent skôr. Rovnaká námietka stojí pri prepise 1f. Cenou za uchopenie existenčného čítania vety 1 pomocou daných prepisov bolo zmnoženie výskytu výrazov, ktoré by mali reprezentovať význam anaforického antecedentu. Zároveň anaforicky použitý výraz je podľa daných prepisov v anaforickej väzbe iba k jednej z dvoch reprezentácií významu anaforického antecedentu. Tá je navyše v sémanticky rozdielnej pozícii k tej, aká patrí pôvodnému významu výrazu vo vete 1. Za cenu uchopenia existenčného čítania vety 1 prestali byť prepisy 1e a 1f adekvátne z pohľadu sémantickej analýzy anafory.

Podobný problém čaká aj mnohé ďalšie pokusy. Dôvodom je podmienkový charakter vety 1. Anaforický antecedent sa nachádza v podmienke – výber farmárov vlastniacich nejaké oslíky. Anaforicky použitý výraz sa nachádza v dôsledku – bitie oslíkov. Problémom DPL, v ktorej sa anafora analyzuje pomocou viazanej premennej, je potreba viazať konkrétny výskyt premennej existenčným kvantifikátorom v implikácii. Ak je existenčný kvantifikátor v antecedente, prepis zastupuje všeobecné čítanie. Ak je existenčný kvantifikátor v konzekvente, nezohľadňuje daný prepis adekvátne sémantický fenomén anafory obsiahnutý vo vete 1. To je však problém, ak máme ambíciu použiť DPL pri analýze anafory. Zdá sa totiž, že sme vyčerpali všetky pozície, kde sa v implikácii môže existenčný kvantifikátor nachádzať. Podobné problémy má aj DLV. Význam *it* spočíva vo väzbe na posledné zmeny v kontexte. Ak sa tie zmeny stanú až v druhej fráze v operátore *every*, potom je už neskoro. Ak však pri použití systému DLV musíme použiť operátor *every* ako prirodzený náprotivok výrazu *každý* z vety 1, potom umiestnenie anaforického antecedentu do prvej frázy S1 predstavuje iba všeobecné čítanie. Máme tu teda problém.

9 Návrh riešenia

Riešenia 1e a 1f nie sú adekvátne z pohľadu uchopenia fenoménu anafory. Z hľadiska zmeny kontextu, ktorú existenčné čítanie vety 1

predstavuje, sú však adekvátnym opisom. Vzhľadom na môj model nevytlúči prepis 1e valuáciu v2 a prepis 1f výberovú funkciu e2. Ak je v kontexte splnené, že nejaký farmár bije aspoň jedného svojho oslíka, nevytlúčime z kontextu tie valuácie, ktoré priradia premennej x tohto farmára (prípadne také výberové funkcie, ktoré priradia predikátu *farmár* toto individuuum). Vlastne som tým formuloval podmienky na výber prvkov kontextu, ktoré má existenčné čítanie vety 1. Okrem opisu tohto čítania vety 1 však chcem aj analyzovať anaforu prostriedkami na to v daných systémoch určenými: viazaným výskytom premennej v DPL a aplikáciou funkcie *the* v DLV. Viem už, že všetky prostriedky na to potrebné sa nachádzajú vo vete 1 *na správnych pozíciách*. Ak ich chcem na týchto pozíciách zachovať, jediná možnosť nápravy tkvie vo funkcii implikácie či operátora *every*. Mohol by som upraviť ich funkciu tak, aby ich použitie opísalo aj existenčné čítanie *pri zachovaní pozícií* prostriedkov predstavujúcich analýzu anafory? Áno a v nasledujúcom texte predstavím návrh úpravy definície implikácie pre DPL a operátora *every* pre DLV, ktorých použitie splňa spomenuté podmienky.

Návrh zmeny definície implikácie pre DPL:

$$[[\varphi \rightarrow_v \psi]] = \{ \langle g, h \rangle; h = g \ \& \ \forall k: \langle h, k \rangle \in [[\varphi]] \Rightarrow \exists j: j[y]k \ \& \ \langle h, j \rangle \in [[\varphi]] \ \& \ \exists l: \langle j, l \rangle \in [[\psi]] \}$$

V čom je jadro zmeny? Táto definícia implikácie predstavuje, podobne ako pôvodná definícia implikácie, dve podmienky, ktoré musia zároveň platiť:

- a) $h = g$. Týmto som ponechal implikácii jej testový charakter.
- b) $\forall k: \langle h, k \rangle \in [[\varphi]] \Rightarrow \exists j: j[y]k \ \& \ \langle h, j \rangle \in [[\varphi]] \ \& \ \exists l: \langle j, l \rangle \in [[\psi]]$.

Druhá podmienka možno vyzerá na prvý pohľad odstrašujúco, jej funkcia je však jednoduchá. Táto podmienka iba znamená, že pre každú valuáciu k patriacu do kontextu obsahujúceho antecedent ($\forall k: \langle h, k \rangle \in [[\varphi]]$) musí existovať nejaká valuácia líšiaca sa nanajvýš ohodnotením premennej y ($\exists j: j[y]k$) splňajúca nasledujúce podmienky: musí patriť do kontextu obsahujúceho antecedent ($\exists j: j[y]k \ \& \ \langle h, j \rangle \in [[\varphi]]$) a zároveň v tomto kontexte musí byť interpretovateľný konzekvent ($\exists l: \langle j, l \rangle \in [[\psi]]$).

Použijem teraz pozmenenú implikáciu pri prepise vety 1:

$$(1g) \ \forall x ((F_x \wedge \exists y (O_y \wedge V(x, y))) \rightarrow_y B(x, y))$$

kde: F – byť farmárom, O – byť oslom, V – vlastniť, B – biť. Ako funguje tento prepis pre môj model? Zaujímavé sú pre mňa opäť iba valuácie v_1 a v_2 . Tie pripisujú premennej x farmára a . Obe valuácie zároveň spĺňajú podmienku určenú antecedentom. Obe totiž pripíšu premennej y oslíka, ktorého farmár a vlastní. Keďže valuácia v_1 sama spĺňa konzekvent (farmár a bije oslíka b), určite je splnená podmienka b) v mojej definícii implikácie: existuje valuácia, ktorá sa od valuácie v_1 líši nanajvyš hodnotou priradenou premennej y , ktorá spadá pod konzekvent. Jednoducho je to valuácia v_1 sama. Líši sa od seba samej nanajvyš hodnotou priradenou premennej y a spadá pod konzekvent. A čo valuácia v_2 ? Spĺňa antecedent a vieme, že existuje valuácia v_1 , ktorá sa od v_2 líši nanajvyš hodnotou priradenou premennej y , spĺňajúca konzekvent. Vďaka použitiu novej podoby implikácie nebude ani valuácia v_2 vylúčená z kontextu priradeného existenčnému čítaniu vety 1. Prepis 1g takto na jednej strane opisuje existenčné čítanie vety 1 a na druhej strane zároveň ponecháva prostriedky analyzujúce fenomén anafory na potrebných miestach. Prepis 1g obsahuje iba jeden kvantifikovaný výraz zodpovedajúci anaforickému antecedentu *nejaký oslík*, ktorý sa zároveň nachádza v antecedente implikácie.

Podobne sa dá upraviť aj definícia operátora *every*:

$$\begin{aligned} || \text{every}_p(S1, S2) || &= \{ \langle e, e' \rangle, e=e' \ \& \ \text{pre každú } e1, \text{ ak } \langle e, e1 \rangle \in \\ &|| S1 ||, \text{ tak existuje nejaká } e2 \text{ taká, že } e2=e1^{||P||} \ \& \ \langle e, e2 \rangle \in \\ &|| S1 || \ \& \ \text{existuje nejaká } e3 \text{ taká, že } \langle e2, e3 \rangle \in || S2 || \} \end{aligned}$$

Rozpíšem túto definíciu podrobnejšie. Opäť predstavuje niekoľko podmienok:

- a) $e=e'$, rovnako ako pri pôvodnej definícii som ponechal testový charakter operátora.
- b) pre každú $e1$, ak $\langle e, e1 \rangle \in || S1 ||$, tak existuje nejaká $e2$ taká, že $e2=e1^{||P||}$ & $\langle e, e2 \rangle \in || S1 ||$ & existuje nejaká $e3$ taká, že $\langle e2, e3 \rangle \in || S2 ||$.

Druhá časť predstavuje prakticky rovnaké podmienky ako pri novej definícii implikácie pre DPL. Najprv pre každú výberovú funkciu $e1$ musí existovať nejaká výberová funkcia $e2$, ktorá sa líši nanajvyš výberom individua predikátu P , ktorá patrí do kontextu frázy $S1$ (pre každú $e1$, ak $\langle e, e1 \rangle \in || S1 ||$, tak existuje nejaká $e2$ taká, že $e2=e1^{||P||}$ & $\langle e, e2 \rangle \in || S1 ||$). Zároveň potom musí byť splnená podmienka, že frá-

za S2 je interpretovateľná v kontexte končiacom $e2$ (& existuje nejaká $e3$ taká, že $\langle e2, e3 \rangle \in S2$).

Použijem pozmenený operátor pri prepise vety 1:

(1h) $every_{oslík}(vlastní(a(farmár), a(oslík)), bije(he, it))$

Obe výberové funkcie $e1$ aj $e2$ patria do kontextu frázy S1: $(vlastní(a(farmár), a(oslík)))$. Pre výberovú funkciu $e1$ existuje výberová funkcia líšiaca sa nanajvýš individuum priradeným predikátu *oslík*, ktorá vyhovuje fráze S2. Je ňou výberová funkcia $e1$ sama. Zároveň teda vieme, že existuje výberová funkcia, ktorá sa líši požadovaným spôsobom od $e2$. Preto bude aj $e2$ patriť do kontextu priradeného prepisom 1h vete 1. Prepis 1h predstavuje existenčné čítanie vety 1.

10 Diskusia o riešení

Tento článok uzatvorím diskusiou o mnou navrhovaných riešeniach. Zameriam sa pri tom na dve otázky:

- a) Nie je ponúkané riešenie *ad hoc*?
- b) V akom vzťahu sú systémy s pozmenenými definíciami k pôvodným systémom?

Pôvodné definície v DPL a DLV mali fungovať ako formálne náprotivky častí prirodzeného jazyka: *ak..., tak...* a *každý* alebo skôr *pre každého ak..., tak...*. Spojka *ak ..., tak ...* sama má veľké množstvo rôznych formalizácií v rôznych logických systémoch. Kým pôvodná definícia opisovala iba jedno z nich, nová definícia opisuje oveľa širšiu triedu rozličných možných porozumení danej spojke. Okrem vety 1 sme pomocou novo definovanej implikácie schopní pomerne elegantne analyzovať aj existenčné čítanie viet obsahujúcich viacnásobnú anaforu, napríklad:

Každý zákazník, ktorý má nejaké centy a nejaké bankovky, nimi zaplatí účet.

Existenčné čítanie sa môže vyskytovať v mnohých rôznych podobách. Stačí, keď veta obsahuje viaceré anaforicky použité výrazy, pričom existenčne čítať môžeme iba niektoré z nich. Preto som zväzil úpravu diskutovaných systémov. Môže sa teda zdať účelová, no napriek tomu umožní adekvátne opísať fenomén anafory vyskytujúci sa v mnohých existenčných čítaniach podmienkových viet. Ako som spomenul, nie-

kedy na správne opísanie nejakého fenoménu z prirodzeného jazyka vznikli celé nové logické systémy. Pokúsil som sa len upraviť dva už existujúce systémy. Navyše v prípade DPL ani sami pôvodní autori neváhali podobne upraviť definíciu implikácie pri analýze prísloviak²².

Pokiaľ ide o vzťah systémov s pozmenenými definíciami k pôvodným systémom, tvrdím, že pozmenené systémy sú schopné opísať všetko to, čo pôvodné systémy. Pôvodné definície sú totiž iba špeciálnymi prípadmi mnou zavedených definícií:

$$\begin{aligned} [[\varphi \rightarrow_0 \psi]] &= \{ \langle g, h \rangle; h = g \ \& \ \forall k: \langle h, k \rangle \in [[\varphi]] \Rightarrow \exists j: [0]k \ \& \ \langle h, j \rangle \in [[\varphi]] \ \& \\ &\exists l: \langle j, l \rangle \in [[\psi]] \} = \{ \langle g, h \rangle; h = g \ \& \ \forall k: \langle h, k \rangle \in [[\varphi]] \Rightarrow \exists j: j = k \ \& \\ &\langle h, j \rangle \in [[\varphi]] \ \& \ \exists l: \langle j, l \rangle \in [[\psi]] \} = \{ \langle g, h \rangle; h = g \ \& \ \forall k: \langle h, k \rangle \in [[\varphi]] \Rightarrow \exists j: j \\ &= k \ \& \ \langle h, k \rangle \in [[\varphi]] \ \& \ \exists l: \langle k, l \rangle \in [[\psi]] \} = \{ \langle g, h \rangle; h = g \ \& \ \forall k: \langle h, k \rangle \\ &\in [[\varphi]] \Rightarrow \exists l \langle k, l \rangle \in [[\psi]] \} \end{aligned}$$

|| every₀(S1,S2) || = {⟨e,e'⟩, e=e' a pre každú e1, ak ⟨e,e1⟩∈
|| S1 ||, tak existuje nejaká e2 taká, že e2=e1^{||0||} & ⟨e,e2⟩∈
|| S1 || & existuje nejaká e3 taká, že ⟨e2,e3⟩∈ || S2 || } = {⟨e,e'⟩,
e=e' a pre každú e1, ak ⟨e,e1⟩∈ || S1 ||, tak existuje nejaká e2 taká,
že e2=e1 & ⟨e,e2⟩∈ || S1 || & existuje nejaká e3 taká, že ⟨e2,e3⟩∈
|| S2 || } = {⟨e,e'⟩, e=e' a pre každú e1, ak ⟨e,e1⟩∈ || S1 ||, tak
existuje nejaká e3 taká, že ⟨e1,e3⟩∈ || S2 || }

*Katedra logiky a metodológie vied
Filozofická fakulta
Univerzita Komenského v Bratislave
Šafárikovo nám. 6
Slovenská republika
814 99 Bratislava
milos.kosterec@gmail.com*

Literatúra

- DUŽÍ, M. (2006): Anafora a význam. In: Zouhar, M. (ed.): *Jazyk z pohľadu sémantiky, pragmatiky a filozofie vedy*. Bratislava: Filozofický ústav SAV, 99-136.
- DUŽÍ, M. – JESPERSEN, B. – MATERNA, P. (2010): *Procedural Semantics for Hyperintensional Logic. Foundations and Applications of Transparent Intensional Logic*. Springer.

²² Pozri napríklad Groenendijk (1989, 40).

- GAHÉR, F. (2002): Anafora a pojmové postoje. In: Gáliková, S (ed.): *Filozofia Ludwiga Wittgensteina*. Bratislava: Veda, 130-155.
- GROENENDIJK, J. – STOCKHOF, M. (1989): *Dynamic Predicate Logic*. Available from <http://dare.uva.nl/document/3701>
- CHIERCHIA, G: (1990): *Anaphora and Dynamic Logic*. ITLI Prepublications for Logic, Semantics and Philosophy of Language.
- KANAZAWA, M. (1992): *Dynamic Generalized Quantifiers and Monotonicity*. Stanford University.
- PEREGRIN, J. – VON HEUSINGER (2003): Dynamic Semantics with Choice Functions. Available from http://www.ilg.uni-stuttgart.de/vonHeusinger/pub/inhalt/Choice_Functions.pdf
- PEREGRIN, J. (2000a): Reference and Inference: The Case of Anaphora. In: von Heusinger, K. – Egli, U. (eds.): *Reference and Anaphoric Relations*. Dordrecht: Kluwer (Studies in Linguistics and Philosophy), 268-286.
- PEREGRIN, J. (2000b) : Variables in natural language: where do they come from? In: Boettner, M. – Thummel, W. (eds.): *Variable-Free Semantics*. Osnabruck: Secolo, 46-65.
- SCHUBERT, L. K. – PELLETIER, F. J. (1989): Generically speaking, or, using discourse representation theory to interpret generics. In: Chierchia, G. – Partee, B. H. – Turner, R. (eds.): *Properties, Types and Meanings II*. Kluwer Academic Publishers, 193-268.
- ZOUHAR, M. (2004): Anafora a referencia. In: Zouhar, M. (ed.): *Používanie, interpretácia a význam jazykových výrazov*. Bratislava: Veda, 128-143.

Stupne nekonzistentnosti

Ladislav Kvasz

Akademie věd České Republiky, Praha

Abstract: Several mathematical theories, as for instance Newton's theory of fluxions and fluents, Frege's theory of the foundations of arithmetic, or Peano's theory of natural numbers were first formulated in a logically inconsistent form. Only after some period of time consistent formulations of these theories were found. The paper analyzes several historical cases of this "initial inconsistency". It suggests distinguishing three kinds of inconsistency according to the "distance" of the proposed inconsistent theory from its consistent variant. These three kinds correspond to whether re-formulations, relativizations or re-codings are needed for turning the inconsistent theory into a consistent one.

Keywords: inconsistency, theory of fluxions, non-standard analysis.

Nekonzistentnosť je logický fenomén, ku ktorému možno pristupovať z rôznych hľadísk. Vedľa snáh o vytvorenie parakonzistentných logík je možné jav nekonzistentnosti skúmať aj z historického hľadiska. Nové matematické teórie sa často rodia ako nekonzistentné, ale napriek nekonzistentnosti sa v ich rámci podarí odvodiť rad pozoruhodných výsledkov. Po kratšej či dlhšej dobe od odhalenia logickej nekonzistentnosti určitej teórie sa podarí vytvoriť jej logicky konzistentnú verziu, v rámci ktorej je možné odvodiť spravidla rovnaké tvrdenia ako v pôvodnej nekonzistentnej teórii. Klasickým príkladom je *diferenciálny a integrálny počet* vytvorený Newtonom a Leibnizom. Trvalo vyše sto rokov, kým sa Cauchy a Weierstrassovi podarilo predložiť logicky konzistentnú verziu tejto teórie, pričom väčšina viet dokázaných Newtonom a Leibnizom v rámci ich nekonzistentnej varianty diferenciálneho a integrálneho počtu sa ukázala ako správna aj z pohľadu konzistentnej teórie.

Keď logici komentujú túto epizódu, zdôrazňujú logickú nedostatočnosť pôvodnej, Newtonovej či Leibnizovej verzie. Napríklad A. W.

Moore v knihe *The Infinite* píše: „Napriek všetkej svojej hĺbke a kráse je uvedená argumentácia, ako sme videli, fundamentálne chybná. Zakladá sa na pojme nekonečne malého rozdielu (ktorý nie je celkom nič, ale ani celkom niečo) a tento pojem je zásadne nekoherentný“ (Moore 1990, 63). Takýto postoj má za následok, že za vedecky (a filozoficky) relevantné sa považujú až konzistentné formulácie diferenciálneho a integrálneho počtu pochádzajúce od Cauchyho a Weierstrassa. Rozboru základných pojmov Cauchyho a Weierstrassovej teórie sa venuje náležitá pozornosť, kým rozbor Newtonových a Leibnizových textov sa redukuje na úroveň *historických kuriozít*, ktoré si nezasluhujú hlbšiu analýzu. Tým, že sa ukázala ich nekonzistentnosť, sa zdá byť všetko povedané. Snažiť sa o filozofickú analýzu nekonzistentnej teórie nemá zmysel.

Som presvedčený, že tento názor je mylný. Samozrejme, v rámci nekonzistentnej teórie je (prostriedkami našej formálnej logiky) možné odvodiť prakticky čokoľvek, ale väčšina výsledkov, ku ktorým dospeli Newton a Leibniz pomocou svojich „logicky nekonzistentných“ teórií, sa ukázala byť správna. Keby sme vybrali náhodný systém nekonzistentných axióm, asi by sa nepodarilo zabezpečiť, aby sme systematicky odvodzovali iba „správne“ výsledky. Problém je, že odvodzovanie, ktoré Newton a Leibniz používali, nebolo čisto logickým odvodzovaním (a oni to ani netvrdili), ale išlo o kontextovo viazané odvodzovacie pravidlá. Ak je to naozaj tak, potom ani logická nekonzistentnosť nie je v ich prípade takým zásadným problémom, akým sa javí zo súčasného hľadiska. Mne nejde o obhajobu nekonzistentných teórií. Logická konzistentnosť je základom matematiky a nechcem ju spochybňovať. Snažím sa len porozumieť, ako možno v rámci logicky nekonzistentnej teórie odvodzovať takmer výlučne správne teoremy. Samozrejme, nie vždy sa to darilo, ale omyly boli zriedkavejšie, než by sme mohli očakávať. Nekonzistentné teórie, napríklad Newtonov a Leibnizov diferenciálny a integrálny počet či Eulerova analýza nekonečne malých veličín, mali napriek nekonzistentnosti veľký metodologický, logický a epistemologický význam. Preto je nutné podrobiť ich serióznej filozofickej analýze a nie ich odbaviť ako *historické kuriozity*.

Mojim cieľom je v rámci širokého pojmu logickej nekonzistentnosti zaviesť *tri užšie pojmy nekonzistentnosti*, ktoré vznikajú tak, že sa určia transformácie, ktoré sú potrebné na to, aby sa teória stala konzistentnou. V prípade týchto troch pojmov nekonzistentnosti pôjde teda o nekonzistentné teórie, ktoré sú prevoditeľné na konzistentné teórie pomocou špecifických transformácií. Chcem určiť mieru nekonzis-

tentnosti teórie tak, že opíšem, aké radikálne zmeny v jej jazyku treba urobiť na to, aby sa teória stala konzistentnou. Verím, že uvedené rozlíšenia prinesú nové argumenty do diskusie, ktorá sa vedie vo filozofii matematiky o Eulerovej variante diferenciálneho a integrálneho počtu, založenej na používaní nekonečne malých veličín. V priebehu 19. storočia boli nekonečne malé veličiny považované za nelegitímne objekty; Eulerova teória bola preto považovaná za nekonzistentnú teóriu a bola nahradená Cauchyho a Weierstrassovou teóriou limit.

Keď Abraham Robinson objavil neštandardnú analýzu, považoval ju za ospravedlnenie Eulerovej teórie. Robinson bol presvedčený, že ukázal, že Eulerova teória *nebola nekonzistentná*, ako si mysleli matematici 19. storočia, keďže pojmu nekonečne malej veličiny je možné dať presný zmysel. Niektorí filozofi však toto presvedčenie odmietajú. Tvrdia, že Robinson vo svojej konštrukcii nekonečne malých veličín použil nástroje, ktoré boli Eulerovi nedostupné, a teda jeho konštrukcia neukazuje oprávnenosť *Eulerových* úvah. Tak A. W. Moore v citovanej práci píše:

Nemecký logik Abraham Robinson (1918–1974) vynašiel to, čo je známe ako neštandardná analýza, a tým nakoniec dal zmysel pojmu nekonečne malej veličiny, väčšej ako 0, ale menšej ako ľubovoľné konečné číslo. Ale použil pritom logické postupy a techniky, ktoré siahajú ďaleko za možnosti matematikov sedemnásteho storočia. Bolo by preto anachronizmom chápať jeho prácu ako ospravedlnenie ich úsilia. Neukázala, že pojem nekonečne malej veličiny, ako jej rozumeli *oni*, bol koherentný. (Moore 1990, 69)

Moore tu robí pozoruhodný argumentačný ťah; ťah, ktorý je dostatočne typický.¹ Keď sa ukáže, že príslušná teória *nie je* nekonzistentná – za akú bola považovaná doposiaľ –, prechádza od objektívneho fak-

¹ Podobný ťah urobil Kuhn pri diskusii o einsteinovskej revolúcii vo fyzike, keď sa snažil diskvalifikovať argument, že newtonovská mechanika je limitným prípadom einsteinovskej mechaniky pri rýchlosti svetla rastúcej do nekonečna. Formálnu rekonštrukciu starej teórie (newtonovskej mechaniky či eulerovskej teórie nekonečne malých veličín) v rámci novej teórie (einsteinovskej mechaniky či robinsonovskej neštandardnej analýzy) uvedení autori odmietajú s poukazom na to, že pojmy použité pri formálnej rekonštrukcii nie sú pojmami starej teórie. Tento argument by bol platný, keby pojmy boli rozhodujúce pri budovaní teórií. Vedľa pojmov však existuje formálny rámec teórie, ktorý zachytáva dôležitejšie štruktúry, než je pojmová stavba.

tu nekonzistentnosti teórie do subjektívnej roviny toho, ako tejto teórii *rozumeli* či *nerozumeli* matematici určitého obdobia. To je ale z hľadiska filozofickej analýzy irelevantné. Spôsob, akým teórii rozumeli v určitom období, je vecou histórie vedy a nie filozofie. Takto Moore príslušnú teóriu odsúva z roviny filozofickej analýzy do roviny histórie. Síce už nie ako *historickú kuriozitu* (ako sa odsúvali nekonzistentné teórie), ale ako *psychologický aspekt* porozumenia teórii.

Aby sme sa mohli pustiť do analýzy nekonzistentných teórií, musíme si v prvom rade ujasniť, čo budeme na nich analyzovať. Domnievam sa, že okrem „*nezaujímavých*“ nekonzistentných teórií možno vyčleniť tri druhy nekonzistentných teórií, ktoré sú nekonzistentné „*zaujímavým spôsobom*“ (t. j. obsahujú konzistentné „*jadro*“, ktoré je nekonzistentným spôsobom uchopené).

1 Pojem re-formulačne nekonzistentnej teórie

Teóriu nazveme *re-formulačne nekonzistentnou* ak je logicky nekonzistentná, ale pomocou *re-formulácie* ju možno spraviť konzistentnou.² Re-formulačnú nekonzistentnosť tak možno považovať za omyl či chybu autora, keďže všetky prostriedky pre vytvorenie konzistentnej verzie príslušnej teórie už mal k dispozícii. Zdá sa, že autor re-formulačne nekonzistentnej teórie v dôsledku nešťastnej formulácie predpokladov, definícií či argumentov sa dostal do rozporu.³ Tento typ nekonzistentnosti nie je prekvapivý. Bolo ho nutné zaviesť iba preto,

² Pojem *re-formulácie* som zaviedol v knihe *Patterns of Change* (Kvasz 2008) ako takú zmenu teórie, pri ktorej sa nemení konceptuálny rámec, v ktorom je teória formulovaná, ale korigujú sa určité propozície. Príkladom *re-formulácie* bol objav planéty Neptún. Tento objav nezmenil pojem planéty, ale zmenil odpoveď na otázku, koľko planét má slnečná sústava. Preto je to *re-formulácia* písaná s pomlčkou. Pôvodná formulácia („*Slnečná sústava má 7 planét.*“) a nová formulácia („*Slnečná sústava má 8 planét.*“) sa logicky vylučujú (tým sa *re-formulácia* líši od *reformulácie* písanej bez pomlčky, keď sa ten istý obsah vyjadri pomocou iných slov).

³ Jav príbuzný re-formulačnej nekonzistentnosti sa vyskytol v euklidovskej geometrii pri „*dôkazoch*“ piateho postulátu. Poznáme vyše 6 rôznych „*dôkazov*“ tohto postulátu, o ktorých súčasníci, alebo príslušníci nasledujúcej generácie, zistili, že sú kruhové – ich autor skryto predpokladal to, čo tvrdil. Dôležité však je, že k odhaleniu cirkularity „*dôkazu*“ došlo v rovnakom konceptuálnom rámci ako bol ten, v ktorom bol „*dôkaz*“ sformulovaný (podobne ako konzistentná verzia re-formulačne nekonzistentnej teórie).

aby voči nemu vynikli dva ďalšie typy, ktorých zavedenie je vlastným cieľom state. V mnohých prípadoch, a verím, že Newtonov či Fregeho prípad patria medzi ne, v dobe formulácie teórie *nebolo možné vytvoriť logicky konzistentnú teóriu*, lebo konceptuálne rámce, v ktorých títo autori pracovali, neobsahovali prostriedky umožňujúce konzistentne vyjadriť obsahy, o ktorých vyjadrenie sa usilovali.

Príkladom re-formulačne nekonzistentnej (mini-) teórie je Cauchyho veta o spojitosti limity postupnosti spojitých funkcií. Cauchy, jeden z tvorcov striktných základov matematickej analýzy, dokázal chybné tvrdenie, že konvergentná postupnosť spojitých funkcií je spojitá. Pojomom re-formulačnej nekonzistentnosti chcem zvýrazniť skutočnosť, že Cauchy bol tvorcom konceptuálneho rámca, v ktorom bolo možné sformulovať korektnú verziu uvedeného tvrdenia. Od korektnej teórie tak Cauchyho delila re-formulácia. V tom sa Cauchyho teória líši od Newtonovej či od Eulerovej, ktoré od korektnej teórie oddeľuje prestavba celého konceptuálneho rámca matematickej analýzy. V prípade re-formulačne nekonzistentných teórií je potrebné analyzovať formulácie jednotlivých definícií, tvrdení a dôkazov, lebo je to práve táto rovina, v ktorej nastáva problém. Takáto epistemologická analýza by sa mohla pokúsiť odhaliť príčinu vzniku nekonzistentnosti.

a. Cauchyho veta o spojitosti limity postupnosti spojitých funkcií

Cauchy definuje spojitosť funkcie slovami: „*funkcia $f(x)$ je spojitá vzhľadom k x medzi danými hranicami, ak medzi týmito hranicami nekonečne malý prírastok $[a]$ neznámej $[x]$ bude mať vždy za následok nekonečne malý prírastok $[f(x + a) - f(x)]$ samotnej funkcie“ (Cauchy 1821, 26).⁴ Na tejto definícii je zaujímavé, že nedefinuje spojitosť v bode, ale hneď pre celý interval (pre x medzi danými hranicami). Prv ako uvediem vetu o spojitosti limity postupnosti spojitých funkcií, uvediem tri Cauchyho definície (Cauchy 1821, 6n.): definíciu premennej: „*Veličinu nazývame premennou, ak ju je možné uvažovať ako schopnú prijať postupne mnoho rôznych hodnôt. Takúto veličinu spravidla označujeme písmenom z konca abecede-**

rie bola podaná v tom istom konceptuálnom rámci ako bola sformulovaná nekonzistentná verzia).

⁴ Do textu definície som vložil v hranatých zátvorkách formálne vyjadrenie prírastku, veličiny a prírastku funkcie, ktoré Cauchy uvádza v texte tesne pred citovanou definíciou.

dy.“; definíciu limity: „Keď hodnoty postupne pripisované určitej premennej sa neobmedzene (indefinitely) blížia k určitej pevnej hodnote takým spôsobom, že nakoniec sa od nej líšia tak málo ako len chceme, táto pevná hodnota sa nazýva **limitou** všetkých ostatných hodnôt.“; a definíciu nekonečne malej veličiny: „Keď postupné číselné hodnoty premennej klesajú nekonečne (indefinitely) takým spôsobom, že klesnú pod ľubovoľné dané číslo, táto premenná sa stáva tým, čo nazývame **nekonečne malou veličinou**. Premenná tohto druhu má nulu za svoju limitu.“

Teraz už môžeme vysloviť príslušné tvrdenie: „Keď jednotlivé členy radu (1)

$$(1) u_0, u_1, u_2, \dots, u_n, u_{n+1}, \dots$$

sú funkciami tej istej premennej x , spojité vzhľadom k tejto premennej v okolí určitej hodnoty, pre ktorú rad konverguje, súčtom radu je opäť spojitá funkcia x v okolí tejto hodnoty“ (Ibid., 90).

Veta v tejto podobe neplatí. Prvý, kto Cauchyho vetu kritizoval, bol Niels Henrik Abel v práci Abel (1826). Ako protipríklad možno uviesť Fourierov rad funkcie, definovanej predpismi:

$$\begin{aligned} f(x) &= +1 \text{ pre } 0 < x < \pi, \\ f(x) &= 0 \text{ pre } x = 0, \\ f(x) &= -1 \text{ pre } -\pi < x < 0. \end{aligned}$$

Táto funkcia je zrejme nespojitá v bode $x = 0$, avšak jej Fourierov rad je⁵

$$f(x) = \frac{4}{\pi} \left\{ \sin(x) + \frac{1}{3} \sin(3x) + \frac{1}{5} \sin(5x) + \frac{1}{7} \sin(7x) + \dots \right\}$$

Máme tu príklad nekonečného radu tvoreného funkciami premennej x , ktoré sú spojité v okolí $x = 0$, avšak súčtom radu je funkcia, ktorá nie je spojitá.

b. Zavedenie pojmu rovnomernej spojitosti

V čom je problém s Cauchyho vetou možno nájsť vysvetlenie v učebniciach matematickej analýzy. V prípade funkcionálnych radov je potrebné rozlíšiť dva druhy konvergencie – bodovú a rovnomernú. Cauchyho definícia vystihuje konvergenciu bodovú, avšak k tomu, aby

⁵ Odvodenie tohto vzťahu možno nájsť v mnohých učebniciach, napríklad v Courant (1927, 517).

súčet radu spojitých funkcií bola opäť spojitá funkcia, bodová konvergencia nestačí, nutné je požadovať konvergenciu rovnomernú. Asi nemá zmysel uvádzať technické detaily teórie funkcionálnych radov; tie možno nájsť vo väčšine učebníc matematickej analýzy. Podobne nemá zmysel filozofický text zaťažovať históriou pojmu rovnomernej konvergencie; tú možno nájsť napríklad v stati Lützen (1999). Podstatné však je uvedomiť si, že oba pojmy – pojem bodovej ako aj pojem rovnomernej konvergencie je možné zaviesť v tom istom pojmovom rámci. Re-formuláciou Cauchyho definície limity je možné dostať definíciu rovnomernej konvergencie. Už v jazykovom rámci Cauchyho teórie **bolo možné** definovať pojem rovnomernej konvergencie a vysloviť Cauchyho vetu v konzistentnej forme. Teraz ma nezaujíma, prečo Cauchy nerozlišoval medzi bodovou a rovnomernou konvergenciou. To je (inak veľmi zaujímavá) otázka, ktorú prenechám historikom matematiky. Ide mi tu výhradne o skutočnosť, že Cauchyho nekonzistentnú teóriu delila od konzistentnej teórie obyčajná re-formulácia.⁶

c. Pojem re-formulačne nekonzistentnej teórie a jeho filozofické dôsledky

Krátko potom, ako Abraham Robinson objavil neštandardnú analýzu, Imre Lakatos napísal (ale nepublikoval) stať s názvom *Cauchy and the continuum: the significance of non-standard analysis for the history and philosophy of mathematics* (Lakatos 1966). V nej Lakatos vyslovil názor, že Cauchy sa nedopustil žiadnej chyby, ale používal iný pojem kontinua – nearchimedovské kontinuum, v ktorom existujú aj nekonečne malé veličiny. Aj keď sa tento názor nestal dominantným, existuje stály prúd literatúry argumentujúcej v prospech tejto alternatívy (napríklad Laugwitz 1987, Katz – Katz 2011). Podľa mňa je snaha vyložiť Cau-

⁶ Ide mi tu o možnosť formulácie korektnej teórie v jazykovom rámci, ktorý používal Cauchy. Uznávam, že keď máme jasne vymedzený pojem rovnomernej konvergencie, je ľahšie ho definovať v Cauchyho jazykovom rámci. Je možné, že na to, aby bolo možné tento pojem objaviť, bolo nevyhnutné Cauchyho rámec opustiť a že v Cauchyho jazykovom rámci je príslušná definícia neprirodená. Túto epistemologickú otázku však teraz nechávam bokom a k problému pristupujem z logického hľadiska. Z logického hľadiska je podstatné, že v príslušnom jazykovom rámci je možné sformulovať definíciu pojmu rovnomernej konvergencie. Či je možné k tejto definícii dospieť prostriedkami daného rámca, to je otázka, ktorou sa na tomto mieste nezaobieram.

chyho vetu prostriedkami neštandardnej analýzy neadekvátne, lebo Cauchyho teória je iba re-formulačne nekonzistentná, a preto je *jednoduchšie* (v zmysle, v akom pojem jednoduchosti používa Poincaré) interpretovať Cauchyho teóriu na pozadí archimedovského kontinua ako teóriu s jednou malou nekonzistentnosťou. Je nepravdepodobné, že by Cauchy skutočne používal nearchimedovské kontinuum a tento fakt by sa na takmer 400 stránkach jeho diela prejavil iba na jednom jedinom mieste – na dôkaze vety o spojitosti súčtu funkcionálneho radu.

2 Pojem objektačnej nekonzistentnosti

Teóriu nazvem *objektačne nekonzistentnou*, ak je logicky nekonzistentná, nie je možné ju urobiť konzistentnou pomocou re-formulácie, a vytvorenie konzistentnej varianty teórie vyžaduje nový konceptuálny rámec, t.j. *objektáciu*.⁷ Podobne ako re-formulačne nekonzistentnú teóriu možno spraviť konzistentnou pomocou re-formulácie, v prípade objektačne nekonzistentnej teórie tento proces vyžaduje objektáciu. Príkladom takejto teórie je Leibnizov a Newtonov diferenciálny a integrálny počet. Newton a Leibniz používali pri svojich dôkazoch postupy, ktoré sú logicky nekonzistentné. Keď tvrdím, že ich teórie sú *objektačne nekonzistentné*, chcem zvýrazniť, že nekonzistentnosť, ktorá sa v nich vyskytla, je síce rádovo väčšia ako re-formulačná nekonzistentnosť, ale na druhej strane je podstatne menšia než prínos týchto teórií. Newton a Leibniz sú tvorcami nového *reprezentačného nástroja*:⁸ objavili pojem

⁷ Pojmu *objektácie* (v anglickej verzii *relativization*) je venovaná druhá kapitola knihy *Patterns of Change*. Nie je možné stručne vysvetliť tento typ zmien, ale pri určitom zjednodušení si objektácie možno predstaviť ako vytvorenie *odstupu* (či *nadhládu*), teda schopnosť pozrieť sa na určitú oblasť zvonka. Tým sa niečo, v čom sme ponorení, mení v predmet, ktorý (v predstave) držíme pred sebou ako objekt. Príkladom objektácie je vznik perspektivistického maliarstva v renesancii alebo vznik substitúcie v algebre.

⁸ Kým objektácie sa viažu na konceptuálny rámec určitej teórie, re-rezentácie predstavujú ešte zásadnejšie zmeny jazyka matematiky. V *Patterns of Change* je reprezentáciám (v anglickej verzii *recoding*) venovaná prvá kapitola. Keď Newton a Leibniz vytvorili nový reprezentačný nástroj, vytvorili nové univerzum objektov – univerzum funkcií, funkcionálnych priestorov, atď. Toto univerzum je možné konceptualizovať rôznymi spôsobmi a objektácie predstavujú prechody medzi jednotlivými konceptualizáciami. Hlavné objektácie v dejinách matematickej analýzy sa spájajú s menami ako Euler, Lagrange, Cauchy, Riemann či Fréchet. Ale všetky tieto, ako aj

funkcie, zaviedli rozlíšenie funkcie a argumentu, zaviedli pojem derivácie a objavili vzťah medzi derivovaním a integrovaním. Tým zásadne zmenili celú matematiku. To, že pojmové ukotvenie týchto epochálnych inovácií bolo nekonzistentné, nie je dôležité. Od konzistentnej verzie diferenciálneho a integrálneho počtu Newtona a Leibniza delilo iba niekoľko objektácií.

a. Newtonova a Leibnizova varianta diferenciálneho a integrálneho počtu

Ako ilustráciu Newtonovej metódy výpočtu derivácie vezmeme pasáž z *Metódy postupností a fluxií*, (citovanú podľa Fauvel – Gray 1987, 385; dôraz LK):

Momenty fluentných veličín (to jest, ich nekonečne [indefinitely] malých častí, pridávaním ktorých narastajú v priebehu každého nekonečne [infinitely] malého úseku času) sú ako ich rýchlosti toku. Preto keď moment ľubovoľnej konkrétnej [fluenty], povedzme x , je vyjadrený ako súčin jej rýchlosti \dot{x} a nekonečne malého množstva o (teda ako $\dot{x}o$), tak momenty ostatných, v , y , z [...] budú vyjadrené ako $\dot{v}o$, $\dot{y}o$, $\dot{z}o$, [...] vidíme, že $\dot{v}o$, $\dot{x}o$, $\dot{y}o$ a $\dot{z}o$ sú k sebe navzájom [v rovnakých pomeroch] ako \dot{v} , \dot{x} , \dot{y} a \dot{z} .

Teraz, keďže momenty (povedzme $\dot{x}o$ a $\dot{y}o$) fluentných veličín (x a y) sú nekonečne [infinitely] malé prírastky, o ktoré tieto veličiny vzrastú počas každého nekonečne [infinitely] malého intervalu času, z toho plynie, že veličiny x a y počas ľubovoľného nekonečne malého intervalu času nadobudnú hodnoty $x + \dot{x}o$ a $y + \dot{y}o$. V dôsledku toho rovnica, ktorá vyjadruje vzťah fluentných veličín v každom okamihu, bude vyjadrovať tento vzťah rovnako medzi $x + \dot{x}o$ a $y + \dot{y}o$ ako medzi veličinami x a y ; a teda $x + \dot{x}o$ a $y + \dot{y}o$ môžu byť dosadené na miesta posledných veličín, x a y , v uvedenej rovnici.

Uvažujme teda rovnicu $x^3 - ax^2 + axy - y^3 = 0$ a dosadíme $x + \dot{x}o$ na miesto x a $y + \dot{y}o$ na miesto y : tým vznikne

$$(x^3 + 3\dot{x}ox^2 + 3\dot{x}^2o^2x + \dot{x}^3o^3) - (ax^2 + 2a\dot{x}ox + a\dot{x}^2o^2) + (axy + a\dot{x}oy + a\dot{y}ox + a\dot{x}\dot{y}o^2) - (y^3 + 3\dot{y}oy^2 + 3\dot{y}^2o^2y + \dot{y}^3o^3) = 0.$$

mnohé ďalšie konceptualizácie, sa týkajú tej istej re-prezentácie, toho istého univerza. Preto sú objektácie rádovo menšími zmenami jazyka ako je re-prezentácia a preto to, že teórie Newtona a Leibniza, sa „dajú dať do poriadku“ pomocou objektácie znamená, že ich chyba je rádovo menšia ako to, čo prinášajú.

Teraz, na základe predpokladu je $x^3 - ax^2 + axy - y^3 = 0$, a keď tieto členy odstránime a zvyšok vydělíme o , ostane

$$3 \dot{x} x^2 + 3 \dot{x}^2 ox + \dot{x}^3 o^2 - 2a \dot{x} x - a \dot{x}^2 o + a \dot{x} y + a \dot{y} x + a \dot{x} \dot{y} o - 3 \dot{y} y^2 - 3 \dot{y}^2 oy - \dot{y}^3 o^2 = 0.$$

Ale ďalej, keďže o je podľa predpokladu nekonečne malé, aby dokázalo vyjadriť momenty veličín, členy, ktoré ju obsahujú ako faktor, budú ekvivalentné s ničím vo vzťahu k ostatným. Preto ich vynechám a ostáva

$$3 \dot{x} x^2 - 2a \dot{x} x + a \dot{x} y + a \dot{y} x - 3 \dot{y} y^2 = 0.$$

Môžeme si všimnúť, že členy neobsahujúce o vždy vymiznú, rovnako ako tie, ktoré sú násobené o vo viac než prvej mocnine; a že zvyšné členy po vydelení o budú mať vždy tvar, aký majú mať podľa pravidla. To je, čo som chcel ukázať.

Z rovnice, ktorú uvádza Newton ako poslednú, možno ľahko odvodiť deriváciu veličiny y (t. j. implicitnej funkcie) v tvare, ako sme na to zvyknutí dnes: $dy/dx = \dot{y} / \dot{x} = (3x^2 - 2ax + ay)/(3y^2 - ax)$.

b. Berkeleyho kritika

V citáte z Newtona som zvýraznil dve pasáže, ktoré sú relevantné z hľadiska ďalšieho vývinu. Roku 1734, sedem rokov po Newtonovej smrti, vydáva George Berkeley svoj slávny spis *The Analyst or a Discourse Addressed to an Infidel Mathematician*, v ktorom predkladá prenikavú a vtipnú kritiku Newtonovej teórie fluxíí a fluent. Berkeleyho cieľom bolo ukázať, že matematická analýza, na ktorej spočíva celá moderná prírodoveda, nemá o nič pevnejší základ než náboženstvo so svojimi anjelmi a zázrakmi. Jadro Berkeleyho kritiky spočíva v tom, že s veličinou o Newton najprv pracuje, akoby bola rôzna od nuly (to je nutné k tomu, aby sa ňou dalo deliť), ale vzápätí ju položí rovnú nule (keď zanedbá diferenciály vyšších rádov). Podľa Berkeleyho, určitá veličina je buď rovná nule, ale potom je rovná nule počas celého výpočtu a teda ňou nemožno deliť, alebo sa nule nerovná, ale potom sa jej nerovná ani na konci výpočtu, a preto nemožno zanedbať jej vyššie mocniny. To, že matematici napriek týmto chybám dospeli k správnym výsledkom, vysvetľuje Berkeley vtipne ako *kompenzáciu chýb*: výpočty matematikov sú chybné, ale matematická analýza používa chyby vždy vo dvojiciach, takže sa v priebehu výpočtu kompenzujú a dajú správny výsledok. Správnosť výsledku nie je dôsledkom správneho postupu, ale len ná-

hodnej skutočnosti, že chyby sa navzájom zrušia. Preto výsledky matematickej analýzy nie sú o nič spoľahlivejšie ako zázraky či zjavenie.

c. Vytvorenie striktných základov diferenciálneho a integrálneho počtu

Je zaujímavé, že cesta, na ktorej matematici dospeli k logicky konzistentnej stavbe diferenciálneho a integrálneho počtu, v istom zmysle sleduje hlavnú myšlienku Berkeleyho (ironicky mieneného) výkladu správnosti výsledkov matematickej analýzy ako kompenzácie chýb. Matematici prestali deriváciu chápať ako podiel dvoch diferenciálov (či momentov) a pojali ju ako nedeliteľný výraz. Inak povedané, napevno spárovali výrazy presne tak, ako to opísal Berkeley, a pre tieto „zlepence“ vytvorili exaktný kalkul. Takže namiesto toho, aby (ako Newton) pracovali s nezávislými momentami veličín a robili s nimi operácie, „ktorých chyby sa po dvoch navzájom vykompenzujú“, pracovali iba s podielom dvoch momentov, teda s celkom, ktorý už mal chyby od začiatku „vykompenzované“. To bola základná myšlienka Lagrangea, ktorý výraz dy/dx považoval za kompaktný výraz, ktorého hodnota sa určuje metódami matematickej analýzy. Pritom Lagrange odmietol akýkoľvek limitný prechod a pre prácu s výrazmi typu dy/dx chcel vytvoriť pravidlá algebraického typu.

Lagrangeova koncepcia sa nakoniec nepresadila a ako úspešná sa ukázala až koncepcia Cauchyho, ktorá preberá Lagrangeovu myšlienku spojiť diferenciály do nedeliteľných výrazov, avšak pre určenie hodnoty týchto výrazov používa limitný prechod. Cauchyho verzia diferenciálneho a integrálneho počtu je tak vybudovaná z rovnakých zložiek ako Newtonova, iba sú tieto zložky skombinované iným spôsobom. Kým Newton najprv pre každú jednotlivú veličinu vytvoril pomocou limitného prechodu jej moment, a potom sa pokúša tieto momenty navzájom kombinovať, Cauchy najskôr (využívajúc Lagrangeovu myšlienku, siahajúcu svojimi koreňmi až k Berkeleymu) veličiny skombinuje do určitých pevných kombinácií a potom (na rozdiel od Lagrangea, ktorý limitný prechod odmietal) prejde s týmito kombináciami k limite. To je zásadná zmena. Newton chcel deriváciu pojať ako *podiel limitných hodnôt (t. j. momentov) veličín*, ale nebol schopný povedať, čo sa čím delí. Svoje predstavy zhrnul v *teórii prvých a posledných pomerov*, ktorú Berkeley parodoval, keď limitné hodnoty veličín, ktoré tvoria príslušný podiel, nazval duchmi zomrelých veličín. Cauchy vo

svojej teórii hľadá *limitu podielu konečných hodnôt* a chápe ju *nie* ako hodnotu, ktorú výraz nadobúda pre limitné hodnoty veličín, ale ako hodnotu, ktorú v limite nadobúda výraz ako celok.

Zdá sa, že rozdiely medzi Newtonovým, Lagrangeovým a Cauchyho pojatím matematickej analýzy spočívajú v *objektácii*, čiže sú to *rozdiely vo forme jazyka*.⁹ Newtonovo pojetie matematickej analýzy možno vyložiť ako pojetie založené na *perspektivistickej forme*, kým Lagrangeovo pojetie je založené na *kompozitívnej forme* a Cauchyho teória je založená na *interpretatívnej forme*. Bolo teda potrebné uskutočniť celý rad objektácií, aby vznikol konceptuálny rámec, v ktorom bolo možné Newtonove výpočty a argumenty konzistentne vyložiť.

d. Pojem objektačne nekonzistentnej teórie a jeho filozofické dôsledky

Keď prijmeme fakt, že existujú teórie, ako napríklad Newtonova teória fluxii a fluent, ktoré umožňujú pomocou heuristických schém usudzovania objaviť mnoho dôležitých poznatkov, ale z logického hľadiska sú nekonzistentné (t. j. protivník ako Berkeley je schopný pomocou tých istých pravidiel dospieť k paradoxným výsledkom), pričom v jazykovom rámci, v ktorom sú sformulované, ich nie je možné sformulovať v konzistentnej podobe, vrhá to nové svetlo na filozofické koncepcie vývinu vedy.

Na jednej strane to ukazuje neadekvátnosť *popperovského falzifikacionizmu*. Keby bol Newton falzifikacionistom, matematická analýza by asi nikdy nevznikla. Newton vedel, že to, čo robí, má mnohé nedostatky, ale našťastie sa nenechal odradiť a vytvoril teóriu, ktorá prešla postupnosťou štyroch objektácií, až vznikol konceptuálny rámec, v ktorom je možné Newtonove myšlienky konzistentne vyložiť. Teda to, že vedci často ignorujú fakty, ktoré odporujú ich teórii, nemusí byť prejavom samolúbosti či iracionality. Môžu si, rovnako ako Newton, na jednej strane uvedomovať, že to, čo sledujú (u Newtona to bola nová re-prezentácia), má zásadne väčší význam než to, čo ich teóriu vyvracia. Na druhej strane môžu cítiť, že vzhľadom na stupeň rozvoja discip-

⁹ Objektácie sú v *Patterns of Change* opísané ako zmeny formy jazyka. Existuje osem foriem: perspektivistická, projektívna, koordinatívna, kompozitívna, interpretatívna, integratívna, konštitutívna a konceptuálna.

líný nemá zmysel pokúšať sa rozpory riešiť, lebo v jazykovom rámci, ktorý je k dispozícii, to nie je možné (t. j. že ich teória je objektačne nekonzistentná, a teda pokusy riešiť problém pomocou re-formulácií nemajú šancu). Výrokov, ktoré dokladajú uvedomovanie si tejto situácie, je v dejinách mnoho, stačí spomenúť d' Alembertove slová „*Chod'te d'alej, viera sa dostaví!*“.

Na druhej strane prijatie vyššie uvedeného faktu spochybňuje predstavu vedeckej teórie ako súboru propozícií uzavrených vzhľadom k logickému vyplývaniu. Tento druh abstrakcie, ktorý je pre potreby matematickej logiky plne adekvátny, epistemológii a metodológii vedy nevyhovuje. Jav objektačnej nekonzistentnosti je významný tým, že takto nekonzistentnú teóriu nemožno ľahko urobiť konzistentnou (t. j. nejde o reformulačnú nekonzistentnosť). Nekonzistentnosť je súčasťou teórie počas dlhého obdobia (objektácie sú zmeny, ktorých stabilizácia si vyžaduje viaceré desaťročia). Ale aj napriek tejto nekonzistentnosti teória nesie v sebe jadro, ktoré sa vďaka *vnútornému rozvoju samotnej nekonzistentnej teórie* (t. j. bez zásahu zvonka) dá konzistentne sformulovať.

3 Pojem re-prezentačnej nekonzistentnosti

Teóriu nazvem *re-prezentačne nekonzistentnou*, ak je logicky nekonzistentná, nie je re-formulačne ani objektačne nekonzistentná, ale pomocou re-prezentácie ju možno spraviť logicky konzistentnou. Príkladom takejto teórie je Eulerova teória nekonečne malých veličín. Táto teória je omnoho „divokejšia“ než Newtonova teória fluxíí a fluent. Eulerovu teóriu nebolo možné spraviť konzistentnou pomocou objektácií. Cauchyho záchrana Newtonovej teórie spočívala práve v eliminovaní nekonečne malých veličín, čo pre Eulerovu teóriu nie je schodná cesta (povedané Berkeleyho slovami, Euler sa nestaral o to, aby sa chyby v jeho teórii vyskytovali v dvojiciach). Až keď Robinson, využívajúc nástroje úplne inej re-prezentácie (teórie množín), vytvoril model nekonečne malých veličín, ukázal, že Eulerovej teórii je možné dať logicky konzistentnú podobu. Teda na rozdiel od Cauchyho, ktorý sa pohyboval v rámci diferenciálneho a integrálneho počtu (teda v rámci tej istej re-prezentácie ako Newton), Robinson opúšťa Eulerov reprezentačný nástroj a konzistentnosť Eulerovej teórie dokazuje takpovediac zvonka, prostriedkami úplne iného jazyka.

a. Eulerova teória nekonečne malých

Na ilustráciu Eulerovho spôsobu práce s nekonečne malými a nekonečne veľkými číslami si ukážeme jeho odvodenie radu pre a^z , ktorú možno nájsť v Euler (1748, par. 114-116). Nech $a > 1$ a ω je „nekonečne malé číslo alebo zlomok taký malý, že je takmer rovný nule“. Potom

$$a^\omega = 1 + \psi$$

pre nejaké nekonečne malé ψ . Teraz položíme

$$\psi = k\omega.$$

Kde k závisí iba od a ; potom

$$a^\omega = 1 + k\omega.$$

Potom pre ľubovoľné reálne číslo i platí

$$a^{i\omega} = (1 + k\omega)^i.$$

Takže vďaka binomickej vete

$$a^{i\omega} = 1 + \frac{i}{1}k\omega + \frac{i(i-1)}{1.2}k^2\omega^2 + \frac{i(i-1)(i-2)}{1.2.3}k^3\omega^3 + \dots \quad (2)$$

Keď je teraz z nejaké konečné kladné číslo, potom $i = \frac{z}{\omega}$ je nekonečne

veľké, a dosadením $\omega = \frac{z}{i}$ do (2) dostávame

$$a^z = a^{i\omega} = 1 + \frac{1}{1}kz + \frac{1(i-1)}{1.2i}k^2z^2 + \frac{1(i-1)(i-2)}{1.2i.3i}k^3z^3 + \dots$$

Ale keď i je nekonečne veľké, sú $\frac{i-1}{i} = 1$, $\frac{i-2}{i} = 1$, etc., a tak dostávame

$$a^z = 1 + \frac{kz}{1} + \frac{k^2z^2}{1.2} + \frac{k^3z^3}{1.2.3} + \frac{k^4z^4}{1.2.3.4} + \dots$$

Prírodné logaritmy vzniknú, keď a zvolíme tak, aby bolo $k = 1$. Euler udáva hodnotu tohto a na 23 desatinných miest, zavádza pre toto číslo dodnes používané označenie e a píše

$$e^z = 1 + \frac{z}{1} + \frac{z^2}{1.2} + \frac{z^3}{1.2.3} + \frac{z^4}{1.2.3.4} + \dots$$

Na tomto a mnohých iných Eulerových odvozeniach vidno, že Euler nepoužíva nekonečne malé veličiny ako aproximácie, ako nástroj na výpočet hodnoty funkcie v blízkyh bodoch či okamihoch času, ako to robili Leibniz a Newton. Práve naopak, počíta s nimi ako s aritmetickými veličinami. Preto Berkeleyova stratégia „párovania“ chýb, na ktorej je založené Cauchyho vybudovanie základov analýzy, je tu, zdá sa, nepoužiteľná.

b. Základy Robinsonovej neštandardnej analýzy

O Eulerovej verzii matematickej analýzy nebolo nikdy ukázané, že by bola nekonzistentná. Koncom 18. storočia matematici dospeli k presvedčeniu, že pojmom nekonečne malého a nekonečne veľkého čísla nie je možné dať konzistentnú interpretáciu. Toto presvedčenie, zdieľané niektorými filozofmi aj dnes (pozri citát z Moora v úvode), stačilo k tomu, aby boli techniky Eulerovho kalkulu opustené a nahradené cauchyovskou analýzou. Preto bolo veľkým prekvapením, keď sa začiatkom 60-tych rokov minulého storočia Robinsonovi podarilo vytvoriť neštandardnú analýzu, ktorá priniesla tzv. hyperreálne čísla, t. j. systém, v ktorého rámci existujú nekonečne malé aj nekonečne veľké čísla. Tak sa vyše 200 rokov po tom, ako Euler publikoval *Introductio ad analysin infinitorum* (Euler 1748), ukázalo, že Eulerove techniky práce s nekonečne malými číslami sú plne konzistentné.

c. Pojem reprezentačne konzistentnej teórie a jeho filozofické dôkazy

Na vyššie uvedenej epizóde z dejín matematiky je pozoruhodné, že aj v matematike je možné stratiť dôveru na základe „ohovárania“. V podstate nik neukázal, že by na Eulerovom kalkule bolo niečo v neporiadku a aj napriek tomu sa presadilo presvedčenie, že pojem nekonečne malého a nekonečne veľkého čísla je vnútorne nekoherentný a teória bola opustená. Preto Eulerova teória asi nie je najvhodnejším kandidátom na re-rezentačne nekonzistentnú teóriu, lebo, prísne vzaté, nespĺňa prvú požiadavku definície tohto pojmu – aby príslušná teória *skutočne bola* nekonzistentná. Ona bola iba za takú *považovaná*. Ale na druhej strane to, že bola v skutočnosti konzistentná, jej nijako nepomohlo; bola odmietnutá, akoby nekonzistentnou bola.

Okrem pozoruhodného faktu, že v samom centre matematiky hrajú fenomény sociálneho presvedčenia takú dôležitú úlohu pri prijímaní či

odmietaní určitej teórie – teda aspoň krátkodobo, lebo nakoniec vyšla pravda najavo (i keď dvesto rokov zas nie je až taká krátka doba) – tento príklad je zaujímavý aj z iného hľadiska. Konzistentnosť eulerovskej analýzy bola totiž ukázaná prostriedkami teórie modelov za netriviálneho prispenia techník teórie množín, pričom obe tieto teórie patria k takým, ktorých *logická* a *expresívna* sila (v presnom, technickom zmysle slova, ako sú tieto pojmy zavedené v *Patterns of Change*) ďaleko prevyšujú logickú a expresívnu silu jazyka matematickej analýzy, v ktorého rámci bola Eulerova teória sformovaná. Preto keď sa vrátíme k druhému citátu z Mooreovej knihy, ktorý sme uviedli v úvode state, domnievam sa, že určite nie je anachronizmom chápať Robinsonovu prácu ako ospravedlnenie Eulerovho úsilia. Pri odmietaní Eulerovej teórie táto nebola odmietaná preto, že by sa ukázalo, že Eulerovo *porozumenie* nekonečne malej veličiny bolo nekoherentné. Matematikovia 19. storočia rovnako ako matematikovia dnes (na rozdiel od filozofov) nezaujíma, ako Euler *rozumel* pojmu nekonečne malej veličiny. Zaujíma ich, či pojem nekonečne malej veličiny je alebo nie je koherentný. Matematici 19. storočia si mysleli, že tento pojem nie je koherentný, preto Eulerovu teóriu zavrhlí (nezávisle od toho, čo si mysleli o tom, ako Euler pojmu nekonečne malej veličiny rozumel). Robinson, podľa môjho názoru, ukázal, že pojem nekonečne malej veličiny je koherentný, a tým ukázal, že Eulerova teória bola zavrhnutá neprávom. A ukázanie tohto faktu možno v plnej miere považovať za ospravedlnenie Eulerovej teórie.

4 Záver

Zdá sa, že najvážnejšou formou nekonzistentnosti je tá, ktorú nie je možné opraviť ani pomocou re-prezentácie. Tento druh nekonzistentnosti možno nazvať *absolútnou nekonzistentnosťou*. Teória, ktorá ju obsahuje, ostáva navždy mimo racionálneho diskurzu. Rozdiel v pohľade matematikov a filozofov na výskyt nekonzistentnosti v matematických teóriách je spôsobený zrejme tým, že filozofi nekonzistentnosť chápu automaticky ako nekonzistentnosť absolútnu. Matematici naproti tomu, vedení historickou skúsenosťou s (*re-formulačne*, *objektačne* a *re-prezentačne*) nekonzistentnými teóriami, sú ochotní sa s takýmito teóriami zaoberať. Štyri pojmy nekonzistentnosti, ktoré navrhujem rozlíšiť, asi nemajú veľký význam z hľadiska aktuálneho bádania. Dnes netušíme, aké *re-formulácie*, *objektácie* a *re-prezentácie* pri-

nesie budúcnosť. Majú však význam pri spätnom pohľade na vývin určitej disciplíny, a to ako pre historika tak aj pre filozofa.

a. Racionálnosť ignorovania nekonzistentností

Z pohľadu historika rozlíšenie rôznych typov nekonzistentnosti dáva racionálny obsah d'Alembertovej výzvy „*Chodzte ďalej, viera sa dostaví!*“. Možno ju parafrázovať ako: „*Netrápte sa re-formulačnou či objektačnou nekonzistentnosťou, keď kladiete základy novej re-prezentácie!*“ D'Alembert nabádal nestarať sa o nekonzistentnosť nižšej úrovne než je lingvistická úroveň, na rozvíjaní ktorej matematik pracuje. Ide o to, že v priebehu 18. storočia sa kladli základy novej re-prezentácie – diferenciálneho a integrálneho počtu. Preto paradoxy, ako je ten, na ktorý upozorňoval Berkeley, je rozumné *dočasne* ignorovať, a vrátiť sa k nim, až keď budú položené *základy novej re-prezentácie*.

Keď v Newtonovi a Leibnizovi budeme vidieť tvorcov nového *jazyka*, nemusíme klásť dôraz na logickú nekonzistentnosť ich *teórií* (ako A. W. Moore v citáte z úvodu tejto state) a môžeme si všímať to, ako pracujú s funkcionálnymi premennými (čo je zásadne nový typ premenných, ktoré diferenciálny a integrálny počet do matematiky priniesol), aké operácie na ne aplikujú (operácie derivovania a integrovania), To, že pre svoj kalkul nedokázali nájsť konzistentné zdôvodnenie, je vedľajšie. Najskôr musíme kalkul vytvoriť, aby sme ho mohli zdôvodňovať.

Tým nechcem obhajovať nekonzistentnosť. Chcem len povedať, že práca matematika prebieha na rôznych úrovniach, a *nekonzistentnosti nižších úrovní je možné dočasne ignorovať*¹⁰ a sústrediť sa na budovanie vyššej úrovne. Existovalo veľmi málo matematikov, ktorí sa stali zakladateľmi novej re-prezentácie – Newton a Leibniz, rovnako ako Frege či Cantor patrili medzi nich. Preto to, že sa v ich diele vyskytujú logické nekonzistentnosti, neznižuje jeho hodnotu. Uvedené nekonzistentnosti sa totiž dajú ošetriť, a nasledovnými generáciami aj ošetrené boli (v prípade Newtona Cauchym, v prípade Fregeho Russellom a v prípade Cantora Zermelom). Nechcem znižovať význam Cauchyho, Russellovej či Zermelovej práce, predsa však ho možno iba ťažko prirovnať k významu prác Newtona, Frege či Cantora. Myšlienka počítať obsahy

¹⁰ *Re-formulácia* je chápaná ako zmena nižšej úrovne než *objektácia*, rovnako ako *objektácia* je chápaná ako zmena nižšej úrovne než *re-prezentácia*.

a objemy útvarov invertovaním operácie derivovania; myšlienka považovať výroky za funkcie, či myšlienka pokračovať v generovaní číselnej postupnosti aj za nekonečnom – tieto myšlienky znamenajú prielom do úplne nových svetov. To, že sa pritom vyskytli určité logické problémy, je len vedľajšia nepríjemnosť.¹¹

b. Nevyhnutnosť ignorovania nekonzistentností

Klasifikácia nekonzistentností vysvetľuje opodstatnenosť ignorovania nekonzistentností aj v ďalšom zmysle. Newton a Euler *nemali šancu spraviť svoje teórie konzistentnými*. Jazykovými prostriedkami existujúcimi v ich dobe jednoducho nebolo možné vytvoriť konzistentné základy diferenciálneho a integrálneho počtu. Stáli teda pred možnosťou žiadny diferenciálny a integrálny počet nerozvíjať a držať sa prísnych štandardov logickej konzistentnosti. Potom by sme však nemali ani modernú fyziku ani modernú technológiu. Druhou alternatívou, práve tou, pre ktorú sa Newton, Euler a desiatky matematikov ich doby rozhodli, bolo rozvíjať diferenciálny a integrálny počet tak, ako sa dá.

To, že jazykovými prostriedkami dostupnými v dobe Newtona a Eulera skutočne nebolo možné vytvoriť logicky konzistentnú teóriu diferenciálneho a integrálneho počtu, alebo aspoň doposiaľ nik také niečo nepredložil, možno zdôvodniť pomocou teórie objektácií (v prípade Newtonovej teórie) a teórie re-prezentácií (v prípade Eulerovej teórie). Dôležité je uvedomiť si, že konzistentná verzia Newtonovej teórie *fluxii* a *fluent* bola vytvorená Cauchym za použitia *interpretatívnej formy* jazyka matematiky. Zdá sa však, že postupnosť foriem od perspektivistickej cez projektívnu, koordinatívnu, kompozitívnu až po interpretatívnu a integratívnu nie je náhodná, ale každá forma v tejto postupnosti je vybudovaná za využitia predošlej. Interpretatívna forma (prvá forma, ktorá je dosť bohatá, aby mohla sformulovať lo-

¹¹ Tento argument pripomína tézu I. Lakatosa, ktorú sformuloval v rámci svojej metodológie vedecko-výskumných programov. Tvrdil, že rodiaci sa program má spravidla horšie výsledky ako program etablovaný – často protirečí empirickým faktom, dáva horšie predpovede. Ale novému programu treba dať určitú dobu, kým sa dostatočne rozpracuje, a počas tejto doby nie je vhodné ho kriticky hodnotiť (a prípadne falzifikovať). Rozdiel oproti Lakatosovi je v tom, že v našom prípade ide o teórie matematické a nie vedecké, a pre matematiku Lakatos predložil úplne inú metodologickú analýzu – teóriu dôkazov a vyvrátení.

gicky konzistentnú teóriu diferenciálneho a integrálneho počtu) sa teda nemohla zrodiť prv, než boli vybudované formy ju predchádzajúce – *perspektivistická, projektívna, koordinatívna a kompozitívna*. Preto diferenciálny a integrálny počet, vybudovaný prostriedkami týchto chudobnejších foriem, je nutne logicky nekonzistentný.

Podľa môjho názoru nekonzistentnosť, na ktorú upozorňuje Berkeley, nie je nedostatkom Newtonovho myslenia. Padá skôr na vrub jazyka než konkrétneho matematika. Newtonova veľkosť spočíva v tom, že aj keď sa ako matematicky cítiaci duch logickej nekonzistentnosti štítil, akceptoval obmedzenia jazyka, a v jazyku s nedostatočným *abstrakčným, diferenciacným, explikačným, unifikáčným, relativizačným a regularizačným* potenciálom (pozri Kvasz 2010) vytvoril, čo bolo možné vytvoriť. Mojm návrhom je chápať Newtonovu teóriu ako priemet konzistentnej teórie diferenciálneho a integrálneho počtu do jazykového rámca, v ktorom sa mnohé rozlíšenia nevyhnutne stratia (teda jazyk, a nie Newton ich nie je schopný udržať) a produktom tohto procesu je potom metóda, ktorú napáda Berkeley. Ak je táto analýza správna, do značnej miery relativizuje význam logickej kritiky. Keby Newton poslúchol maximu logickej konzistentnosti, matematická analýza by nikdy nevznikla. Nevyhnutné bolo pestovať logicky nekonzistentnú teóriu až dotedy, kým vnútornou diferenciaciou jazyk nedospel k dostatočne bohatej forme, ktorá umožnila Cauchymu a Weierstrassovi predložiť logicky konzistentnú variantu tejto teórie.

Podakovanie

Príspevok vznikol v rámci programu *Fellowship Jana Evangelisty Purkyně* vo Filozofickom ústave AV ČR.

*Oddělení analytické filosofie
Filozofický ústav
Akademie věd České republiky
Jilská 1
110 00 Praha 1
Česká republika
ladislavkvasz@gmail.com*

Literatúra

ABEL, N. H. (1826): Untersuchungen über die Reihe *Journal für reine und angewandte Mathematik* **1**, 311-339.

- BERKELEY, G. (1734): *The Analyst; or, a Discourse Adressed to an Infidel Mathematician...* . . . In: Luce, A. A. – Jessup, T. E. (eds.) (1951): *The Works of George Berkeley, Bishop of Cloyne*, 4. London.
- BOSS, H. J. M. (1980): Newton, Leibniz and the Leibnizian tradition. In: Grattan-Guinness (ed.) (1980), 49-93.
- CAUCHY, A. L. (1821): *Cours d'analyse*. Citované podľa anglického prekladu: Bradley, R. E. – Sandifer, C. E. (2009): *Cauchy's Cours d'analyse, An Annotated Translation*. Springer.
- COURANT, R. (1927): *Vorlesungen über Differential- und Integralrechnung 1*. Berlin: Springer. Citované podľa ruského prekladu *Kurs differencialnogo i integralnogo isčislenija I*. Moskva: Nauka 1967.
- DAVIS, M. (1977): *Applied Nonstandard Analysis*. New York: John Wiley.
- EDWARDS, C. H. (1979): *The Historical Development of the Calculus*. New York: Springer.
- EULER, L. (1748): *Introductio ad analysin infinitorum*. Lausanne. Ruský preklad: *Vvedenie v analiz beskonečnych*. Moskva: GIFML 1961.
- FAUVEL, J. – GRAY, J. (eds.) (1987): *The History of Mathematics. A Reader*. London: Macmillan.
- GRATTAN-GUINNESS, I. (ed.) (1980): *From the Calculus to Set Theory*. London: Duckworth.
- GUICCIARDINI, N. (1999): Newtons *Methode* und Leibniz' *Kalkül*. In: Jahnke, H. N. (ed.) (1999), 89-130.
- JAHNKE, H. N. (ed.) (1999): *Geschichte der Analysis*. Heidelberg: Spektrum.
- JAHNKE, H. N. (1999): Die algebraische Analysis des 18. Jahrhunderts. In: Jahnke, H. N. (ed.) (1999), 131-170.
- KATZ, K. – KATZ, M. (2011): Cauchy's continuum. *Perspectives on Science* **19**.
- KUHN, T. S. (1962): *The Structure of Scientific Revolutions*. Chicago: Chicago University Press. Citované podľa slovenského prekladu *Štruktúra vedeckých revolúcií*. Bratislava: Pravda 1982.
- KVASZ, L. (2008): *Patterns of Change, Linguistic Innovations in the Development of Classical Mathematics*. Basel: Birkhäuser Verlag.
- KVASZ, L. (2010): Náčrt teórie potencialít jazyka matematiky. In: Kvasnička, V., Pospíchal, J., Návrat, P., Lacko P. a Trebatický P. (eds.): *Umelá inteligencia a kognitívna veda II*. STU v Bratislave, 263-290.
- LAKATOS, I. (1966): Cauchy and the Continuum: the Significance of Non-standard Analysis for the History of Mathematics. In: Lakatos, I. (1978), 43-60.
- LAKATOS, I. (1978): *Mathematics, Science, and Epistemology. Philosophical Papers of Imre Lakatos Volume II*. Cambridge: Cambridge University Press.
- LAUGWITZ, D. (1987): Infinitely small quantities in Cauchy's textbooks. *Historia Math.* **14**, 258-274.
- LEIBNIZ, G. W. (1684): Nová metóda maxím a miním, ako aj dotyčníc, ktorá nie je viazaná na zlomky ani na iracionálne čísla, a zvláštny spôsob ich počtu. In: Leibniz, G. W. (1956).

- LEIBNIZ, G. W. (1956): *O reforme vied*. Bratislava: Vydavateľstvo SAV.
- LÜTZEN, J. (1999): Grundlagen der Analysis im 19. Jahrhundert. In: Jahnke, H. N. (ed.) (1999), 191-244.
- MOORE, A. W. (1990): *The Infinite*. London: Routledge. Citované podľa druhého vydania z roku 2001.
- ROBINSON, A. (1966): *Non-Standard Analysis*. Amsterdam: North Holland.

The Idea of Code in Contextualism and Minimalism

Jakub Mácha

Masaryk University, Brno

Abstract: In this paper I discuss the idea of a semantic code in the contemporary debate between contextualism and minimalism. First, I identify historical sources of these positions in Grice's pragmatics and in Davidson's theory of meaning in order to sketch the role of a semantic code there. Then I argue that contextualism is committed to the idea of an *ad hoc* code, while minimalism involves a persistent code. However, the latter approach to a code requires disambiguation which must be carried out in the early stages of speech act processing. I raise a concern that primary pragmatic processes may be active here, especially in the case of disambiguating polysemous expressions, which could be problematic or even devastating for the minimalist program. At the end I evaluate a possible minimalist way out by examining the minimalist account of metaphor, which lies at the root of polysemy. If a code robust enough to deal with polysemy could be created, minimalist conceptions would present a new impetus to understand language as a code. Without such a code, very little would be left of the notion of a persistent code and hence of minimalism itself.

Keywords: code, contextualism, minimalism, disambiguation.

At the end of the twentieth century, it could have seemed that the idea of a semantic code had been discarded. Wittgenstein subsumed it under the Augustinian understanding of language, while Quine addressed it in his myth of a museum. In spite of its apparent rejection, the idea has recently inspired renewed interest. Curiously enough, this renewed interest has arisen thanks to thinkers who directly or indirectly follow in the footsteps of Wittgenstein and Quine.

The current debate focuses on the role of (linguistic or non-linguistic) context in communication and linguistics. If we admit the role of

context, how do we capture this role in theory? Given that the number of possible contexts is limitless, a language user cannot learn all the various impacts of the context. Others argue that, in communication and linguistics, we need to get by with finite means. This finite knowledge could be captured by a finite code coupled with finite recursive rules of composition, in which the understanding of meaning and the interaction of meaning with the context is, as it were, sublimated.

By semantic code I mean a form of mapping that assigns meaning to each element of language, usually a word.¹ This mapping captures language conventions. However, the terms “semantic code” and “language convention” are often used interchangeably. For the purposes of the present work, this indiscriminate use can be accepted, although the link between expressions and their meanings can be constituted in other ways, such as naturally (Cratylus’ hypothesis). If we define semantic code on the level of individual words, we obtain a finite mapping, as each natural language has only a finite number of words. This finite code can be extended into an infinite one using suitable rules of composition.

Two extreme cases of code are worth mentioning. At one extreme, a code is defined primarily on the level of sentences; at the other extreme, a code is defined on the level of letters. A code defined on the sentence level is infinite, as there is infinity of sentences in a natural language. A semantic code defined on the level of letters seems almost an absurdity (assuming we are concerned with a language that uses letters, such as English or Russian). However, strictly speaking, such a code is conceivable, if we systematically describe the semantic effect of each letter in the context of the letters that surround it. I mention this only because a similar problem pertains to lexical code.

What do we need a code for when studying language? This question can be answered only if we clarify the goals to be achieved by the study of language. We can study language in order to create a systematic theory of meaning and inference, where “systematic” has to be understood as “compositional”. Alternatively, we can study language in order to create a psychological theory of communication that involves such questions as how communication is possible, what our communication abilities are, and what things we are able to communicate. These

¹ I use “meaning” in a broad sense, which includes Fregean senses, rules, stereotypes, etc.

two aims are not in contradiction; under certain conditions, they can even be complementary.²

What is a semantic code good for, then, if we want to devise a systematic theory of meaning? The answer is obvious: a semantic code is suitable for modeling the semantic relation, because of the compositionality of the semantic code. A theory of meaning that was not based on the notion of a semantic code would not be systematic, and, thus, it would not meet the initial requirements. In a way, the idea of a semantic code is necessitated by the requirements.

If we aim to create a theory of communication, the usefulness of the semantic code can be clarified using the following transcendental consideration. People understand each other when using language, as they share a code. Or more precisely, in order for people to understand each other at all, they have to share a code. When a person utters “xsffg”, others will not understand the utterance, as the string is not an element of the domain of the shared code. By contrast, a person’s utterance “this is a writing desk” is readily understood by others. A semantic code is, here, a cognitive basis on which the psychological process of understanding is constructed. According to this consideration, knowledge of language amounts to knowledge of a semantic code, the rules of composition, and the rules governing how the code is applied in specific situations.

To sum up, a semantic code can be an idealized condition of the possibility of a semantic theory without the code’s having any psychological reality, and, conversely, a semantic code can have a psychological reality without being a condition of the possibility of communication.

For the sake of the following discussion, three problematic points concerning the idea of a semantic code need to be mentioned. The first problem is that natural languages undergo change, while a semantic code, as an abstract object, is unchanging. If we take a radical stance following Quine and Davidson, we can say that languages change from moment to moment, a fact that needs to be accounted for in the theoretical models of language. From this point of view, two notions of code can be distinguished: a code as an abstract *persistent* object, and a code as an object that is constructed in communication *ad hoc*. One could ob-

² With a little simplification, the first aim is one of minimalists, while the second aim is one of contextualists. Compare Borg (2004, 1-3), Jaszczolt (2005) or Bach (2011).

ject that the second notion of a code is so ephemeral that it does not even deserve to be called a “semantic code”. However, following the definition of a semantic code as a mapping between expressions and their meanings, such an ephemeral mapping is a semantic code per definition, no matter how ephemeral and *ad hoc* it may be. One problem of the notion of a shared *ad hoc* semantic code can be that it does not explain the possibility of communication. We cannot explain the possibility of communication by pointing to a shared code that has been created no sooner than in the very act of communication. The proponents of such a notion of a code have to explain how the communication partners arrive at their shared code. In the following, two methods of construction of a shared *ad hoc* code are mentioned. An advantage of *ad hoc* codes is that they work also on the level of sentences, as they need to specify the sentence meaning of only a small subset of sentences related to the situation at hand rather than the sentence meaning of all possible sentences.

The second problem of a semantic code may be its ambiguity. Natural languages contain polysemous expressions and homonyms, i.e. words that have more than one meaning. In order for a semantic code to be a mapping in the mathematical sense (each member of the domain is mapped to at most one member of the codomain) *disambiguation* needs to be done, which consists in assigning indices to terms in such a way that each term considered together with its index has exactly one meaning. For example, the word *hard* can be disambiguated as *hard*₁ (not easily penetrated, cut, or separated into parts) or as *hard*₂ (difficult to accomplish). The choice of the index of a term depends on the context. Capturing rules of disambiguation turns out to be a formidable task, as I will show later. The existence of polysemy and homonymy suggests that meaning does not depend exclusively on individual words but rather is spread throughout the context of sentences. This problem bears an analogy to a semantic code on the level of letters; a lexical code is possible if we are able to capture polysemy and homonymy with finite means.³ Homonymy and disambiguation that homonymy makes necessary present a problem only for the persistent view of a code. If

³ Wittgenstein (1974, 195) considers the hypothesis that the meaning of a particular word changes according to the day of the week. Such context dependence is systematic and can be captured in lexical code. However, if the meaning changed arbitrarily from day to day or from sentence to sentence, such context dependence could be captured in a code only using an infinite

a code is constructed ad hoc, homonymy does not emerge (except for when it is intentional), as it is resolved when the code is constructed.

1 Historical Background

Contemporary contextualism in the philosophy of language is the result of an effort to elaborate on, or solve, some problems of Gricean pragmatics. In fact, the influence of Gricean pragmatics on contextualism is explicitly acknowledged by contextualists themselves. By contrast, the intellectual sources of minimalist theories are not so clear. I will argue that what lies at the core of minimalism is Davidson's conception of language enriched with Grice's terminology and Grice's theory of communication. However, many elements of Davidson's conception of language can also be found in contextualism. Thus, a short overview of Grice's and Davidson's theories seems to be in order.

1.1 Gricean Pragmatics

Most contextualist authors adopt or refine Grice's basic terminology. To an extent, so do the semantic minimalists; semantic minimalism is also called Gricean minimalism. Grice introduces the technical terms "what is said" and "what is implicated". He distinguishes "what is said" from conventional meaning or "what is encoded", although he says the two are closely related. Roughly speaking, the transition from conventional meaning to what is said is a matter of semantics, while the transition from what is said to what is implicated is a matter of pragmatics. This close relation between conventional meaning and what is said is where contextualism picks up the threads of Grice's work by asserting that pragmatic phenomena are active also in this area. The notion of "what is (en)coded" captures the idea of code that is the subject of the present text.

Now I would like to sketch the relationship between conventional meaning and "what is said". In one of his most recent works, Grice came up with the notion of centrality of meaning, which has two independent properties: *dictiveness* and *formality*. Grice identifies the dictive content with "what is said"; formality is related to conventional

number of rules. To put it differently, it would be a code on the sentence level.

meaning.⁴ We can classify speech acts using these two properties. If a speech act has both properties, what is said is determined solely by the conventional meaning. The interesting cases are, however, those in which one of the two properties is missing. A speech act can lack its dictiveness; that is the case of a conventional implicature. Dictive speech acts that lack formality are especially interesting for contextualists. In these speech acts, what is said is partially or wholly independent of the conventional meaning of a sentence. This allows contextualism to be classified into a *moderate* form or a *radical* form, depending on whether it claims partial or complete independence of what is said on the conventional meaning. If “what is said” is at least partly independent of the conventional meaning, then it needs to be explained how this transition (from the conventional meaning to what is said), generally called *modulation*, is possible. Recanati calls this process “sense enrichment”; Sperber and Wilson (2008) speak of the process of a construction of meaning; Levinson goes so far as to allow “what is said” to be influenced by implicatures that were derived later. Here I have mentioned only the most influential theories.

1.2 Davidson’s Minimalist Conception of Meaning

Davidson aimed at providing an acceptable theory of meaning for natural languages. Such a theory should specify what each sentence of a language means, and thus capture semantic code on the sentence level. Furthermore, a theory of meaning has to be holistic in the following sense: “only in the context of the language does a sentence (and therefore a word) have meaning.” (2001b, 22) This has to be understood in such a way that a theory of meaning must contain all (true) sentences “*s* means *m*”. This intuitive requirement has led Davidson to use the well-known T-scheme: “*p*” is true if and only if *s*. Notice that giving the truth conditions of the sentence is, as a consequence, a way of giving its meaning. The meaning of sentence *p* is abstracted from *all* T-sentences that have “*p*” on the left side. But what is the scope of this universal quantifier?

⁴ Some authors identify Grice’s terms “formality” and “dictiveness” with Austin’s terms “locutionary” and “illocutionary”. See Bach (2004) or García-Carpintero (2006, 45-46).

A theory of meaning is, according to Davidson, an empirical theory that has to answer the question of what the meaning of sentence p is, rather than what meaning is *per se*. This theory has to be empirically testable on “sample cases”. In these cases, we do not deal merely with sentence p , but rather with its particular use, and thus utterance. However, Davidson—like later minimalists—explicitly insists that truth (and thus also truth conditions) is a property of sentences rather than utterances. A language user has available only a finite number of specific uses and the scope of the mentioned universal quantifier is—although getting broader—finite. Davidson calls such a theory of meaning a *prior theory*, while an extension (and reinterpretation) of the theory with the current language situation is called a *passing theory*. It can thus be said that language meaning is knowable in this restricted sense, and that if we are concerned with a rational (and psychological) reconstruction of human communication, this restriction of the scope of the universal operator needs to be taken seriously.

However, Davidson (and, later, also minimalism) offers a theory of meaning for sentences (type) of language. Thus, the scope of the mentioned quantifier is still left unspecified. Let us try to make the scope as broad as possible. If we are concerned with a theory of meaning for a particular language such as, say, English, we need to quantify over all possible uses of sentence p in English and, from them, abstract the meaning (i.e. truth conditions). Such a conception of meaning can be called minimal or minimalistic.

In the paper “Communication and Convention”, Davidson (2001a) questioned whether a shared conventional code is a necessary condition of communication. In order for communication to be successful, the speaker and her addressee need to assign the same meanings to the same words (which means that their passing theories have to match). Put differently, the addressee has to correctly determine “what is said”, i.e. the ad hoc code. In order for this condition to be met, it is not necessary that the speaker and the addressee share a persistent semantic code in advance. It does not matter what stands at the beginning of the communication process, during which the speaker performs a speech act and his addressee tries to figure out what is said; what matters is that there is an agreement at the end of the process. In order for this agreement to be possible (i.e. in order that the communication is successful), the language users have to share the same method of how to use their knowledge of language in a particular context. Davidson calls

this general process the “radical interpretation”, and later a transition from a prior theory to a passing theory. The conventions of language enter the process as a “practical crutch”, which sometimes helps and sometimes does not or need not. In the contemporary debate, a similar style of thought can be found in radical contextualism, as exposed in the following section.

2 Contextualism

I shall present a definition of contextualism that loosely follows Recanati (2010b). Recanati defines *methodological* contextualism as a position that admits that the meaning of each expression (rather than the meanings of only indexical expressions and demonstratives belonging to the *basic set*) can depend on context. This is followed by two varieties of contextualism which share the thesis that knowledge of contextual factors is necessary for the determination of the meaning of a sentence. Thus, the meaning of a sentence depends on the meanings of its parts, on the manner of their composition, and on context. Because of the specific contextual parameter, we have to speak of the meaning of a specific utterance (token) instead of the meaning of a sentence (type). *Moderate* (modulation based) contextualism is based on the assumption that the meaning of an expression is insufficient (and often unnecessary) for determining the meaning of an utterance that uses the expression. Conventional meaning undergoes primary pragmatic processes⁵ which Recanati calls modulation. Basic types of modulation include free (i.e. not linguistically mandated) enrichment, predicate transfer, and sense-extension. *Radical* contextualism goes even further by claiming that the idea of an association of expressions with meanings is untenable. This is called *generalized indexicality*. Every expression contributes to the meaning of a sentence by means of certain contextual factors. In my view, the difference between moderate and radical contextualism lies in their stance toward the idea of a persistent code. Radical contextualism rejects the idea, while moderate contextualism admits it in a constrained form. Radical contextualism (sometimes also called occasionalism) can also be seen as an effort to thoroughly re-interpret Wittgenstein’s thesis that meaning is use, which amounts to a rejection of Grice’s distinction

⁵ Secondary pragmatic processes pertain to a derivation of implicatures.

between what is semantically encoded and what is said (Baptista forthcoming, 3).

What are words associated with, then, if not with meanings? Recanati answers this question by offering a theory of *semantic potentials* inspired by Wittgenstein and Waismann. A word s is associated with a set of its legitimate uses $\{S_1, S_2, \dots, S_n\}$. If the word is newly used in a situation S_{n+1} , this new use is legitimate only if it is sufficiently *similar* to the uses already present in the set. However, similarity is a vague notion. To deal with this vagueness, Recanati uses the idea of the *contrast set* which makes it possible to abstract the property in respect to which the situations are deemed similar or dissimilar. New legitimate uses enrich the original set and thereby extend the semantic potential of the word.⁶

A code as understood by Recanati is not shared, as people cannot be expected to share the same semantic potential of every word. The sharing of semantic potential is possible only in a very limited manner; e.g. learning a mother tongue can be understood as passing on semantic potential in a shared situation. The condition of successful communication is a certain similarity between the semantic potentials of the speaker and her addressee (together with a shared method for determining this similarity). Two semantic potentials are similar if, in a given situation, they lead to the same judgment of their legitimacy.

Recanati's theory is strongly behavioristic and psychological. While it is inspired by Wittgenstein's conception of language, it is questionable that it faithfully captures Wittgenstein's ideas. However, Recanati is not concerned with an exegesis. More interestingly, his conception of language learning by association reminds one of Davidson's reconstruction of communication in the article "A Nice Derangement of Epitaphs" (2005). The enrichment of semantic potential with a new situation could be seen as a psychological reconstruction of the transition from prior theory to passing theory. Just like Davidson, Recanati admits the idea of code only as a set of temporary associations that dynamically change during communication. The association of expressions with meanings (or semantic potentials) is nothing persistent.

A common objection against all varieties of contextualism is that, if its main tenets were true, communication would be impossible. Although the theory of semantic potentials could seem to be an adequate

⁶ Recanati (2003), compare also Bezuidenhout (2002).

response to the objection, Recanati has come up with what he calls an *availability principle*.⁷ One of its many formulations is as follows:

What is said must be intuitively accessible to the conversational participants (unless something goes wrong and they do not count as ‘normal interpreters’). (Recanati 2004, 20)

This principle, which is a reflection of the psychological spirit of Recanati’s conception of language and communication, constrains the results of primary pragmatic processes to those that can be “normally” expected by the participants of communication. The trouble is with the word “normally”.⁸ The communicating partners must share the same intuitions or at least the same criteria of normality in order to have direct access to what is said. If normality were determined by introspection (as is suggested by the unfortunate term “intuitive”), the speaker could not tell whether he is normal or merely believes himself to be normal, which results in a case of private rule following.⁹ However, Recanati suggests that, in order to judge normality, it suffices to have epistemic access to the following factors: (a) the sentence spoken, (b) the meaning of the sentence (semantic code) and (c) relevant contextual factors (Recanati 2004, 20). This suffices, we are told, for the derivation of what is said. But then the availability principle is transcendently deduced rather than psychologically observed. It plays the role of a transcendental condition of successful communication, which places even harder demands on communication than the respective condition in the case of minimalism. Recanati’s availability principle requires both a shared semantic code and a shared relevant context, whereas minimalism demands a shared code only.

If the theory of semantic potentials is true, the participants of communication have no access to the shared meaning of a spoken sentence. Hence, the process of derivation of “what is said” is not shared and deterministic; we have no guarantee of successful communication. But in the fact, there can be no such guarantee in this psychological framework. As Recanati reminds us, communication is a matter of “negoti-

⁷ This principle distinguishes Recanati from other contextualists. The relevance theory comes with what it calls the *principle of relevance*. See e.g. Sperber – Wilson (2008, 177-181).

⁸ This objection is raised by Cappelen – Lepore (2005, 188f).

⁹ Wittgenstein (1958, §202).

ating meaning”, seeking and approaching the perfect understanding (Recanati 2010a, 8).

There is a sort of payoff here. Recanati includes contextual factors into his availability principle, which, on the one hand, allows him to keep the idea of code narrow, but, on the other hand, cannot guarantee successful communication. Cappelen and Lepore demand that the success of communication cannot depend on shared context (Cappelen – Lepore 2005, 182-189) and communication must be possible across contexts. This is to understand that they exclude contextual factors from conditions of communication. But then they have to demand a very robust semantic code. The rest of my paper is devoted to reflections upon the idea of such a robust code.

3 Minimalism

I admit that the goal of contemporary minimalists is unclear to me. While they claim they are concerned with creating a systematic semantic theory,¹⁰ they respond to objections that their minimalist conception of meaning has no psychological reality. (Cappelen – Lepore 2005, Ch. 12) If proponents of minimalism only aimed at a systematic semantic theory rather than the psychological reality of the minimal meaning, there would be little left to disagree about.¹¹

Minimalism can be defined in many ways. According to Cappelen and Lepore, the basic idea of minimalism is simple: The semantic content of a sentence *s* is the same for *all* utterances of the sentence *s*.¹² If a sentence depends on context in some way, this dependence is triggered semantically, as is the case with indexical expressions which belong to the *basic set*.¹³ Minimalism distinguishes between “what is semantically

¹⁰ E.g. in Borg (2004, 3).

¹¹ Recanati (2010, 12-14) claims that this form of minimalism is true by definition.

¹² Emma Borg explicitly rejects this definition: “I’d reject, then, the view of Soames (2002) and Cappelen – Lepore (2005) that minimal contents are those conveyed by *every* utterance of a sentence”. (2010b, 57)

¹³ This is a characterization by Cappelen – Lepore (2005, 143f); the (more radical) minimalism of Emma Borg (2009) adds the thesis that the dependence on context must be formally traceable.

expressed" (or the minimal content or meaning, minimal proposition) and "what is said", where the former is part of the latter. What is semantically expressed is a result of disambiguation in a broad sense and of the saturation of context dependent expressions. For minimalists, "what is said" is a pragmatic notion, whose content can significantly go beyond the content that is expressed semantically. The *speech act pluralism* of Cappelen and Lepore goes so far as to deny any systematic connection between what is semantically expressed and what is said.

Emma Borg's definition of minimalism is slightly different. Her definition of semantic content is this: "Semantic content for sentences is fully determined by syntactic structure and lexical content: the meaning of a sentence is exhausted by the meaning of its parts and their mode of composition" (2010b, 54). For Cappelen and Lepore, a process is pragmatic if it cannot be identified on the syntactic level of a sentence; for Borg (2009, §4), a process is pragmatic if it cannot be derived formally using lexical information only. Finally, let us mention a definition of minimalism from one of its opponents, François Recanati, who labels it I-minimalism:

What is said_{int} is affected by the bottom-up process of saturation but not by top-down processes such as free enrichment. (Recanati 2010, 14)

What is said_{int} is defined as the intuitive truth conditions of an utterance. Thus, this variety of minimalism is concerned with the psychological reality of what is said rather than merely with a systematic semantic theory. Crucial here is the division of processes into *top-down* and *bottom-up*. Roughly, a bottom-up process performs composition or synthesis from elements, whereas a top-down process performs decomposition and the analysis of a larger whole (sentence) into its elements. Recanati gives the following definition: "A 'top-down' or context-driven process is a pragmatic process which is not triggered by an expression in the sentence but takes place for purely pragmatic reasons". (Recanati 2010, 4)

What is the relationship of this terminology to Grice's above-mentioned classification? What is semantically expressed lies somewhere between Grice's conventional meaning and what is said. However, if we remove illocutionary force from Grice's definition of "what is said", we come close to minimal content:

The minimal content of utterance u of sentence s (in context c) is defined as the M-intended truth-conditional content of *sentence s (in context c)*, to whose truth the speaker is committed.

Furthermore, we need to clarify the relationship of minimal meaning to Davidson's conception of meaning. Here is an obvious connection. Both conceptions define meaning as an abstraction over all uses. All that I have written above about the problems of understanding the scope of such quantification applies also to minimalism. If a semantic minimalist tries to define minimal meaning, he arrives at something like Davidson's T-schema:

Utterance "A is red" semantically expresses the proposition that *A is red*. (Cappelen – Lepore 2005, 157)

There is a worry that what all the utterances of the sentence "A is red" have in common is only (or above all) that they contain the sentence itself. The minimal meaning, or – more specifically – minimal proposition, is for Cappelen and Lepore a metaphysical question in the sense of searching for common properties of objects that are included under a term (e.g. what is it that is common to all big things or all brave actions). Wittgenstein's objection would be that they may have nothing in common at all; they are connected merely by a network of similarities making up a family resemblance. Cappelen and Lepore call this position *metaphysical nihilism* and deem it absurd.

What is the notion of minimal meaning good for, then? Cappelen and Lepore give a clear answer: "In short: the proposition semantically expressed is our minimal defense against confusion, misunderstanding, mistakes and it is that which guarantees communication across contexts of utterance." (Cappelen – Lepore 2005, 185) Thus, minimal meaning has the role of a transcendental condition of communication that was attributed to the idea of a semantic code at the beginning. Contextualists object that the minimal meaning so construed is "an abstraction with no psychological reality" (Recanati 2001, 88). The problem lies in a too broad (or unconstrained) quantification over all uses of each sentence. Only a finite number of utterances can enter into an abstraction of minimal proposition if it is to keep its psychological reality.

Thus, a minimal proposition is what can be expected by an addressee who has no access to the context of an utterance. If there were no such propositions, this would not mean that communication was impossible, but merely that *some* sentences could not be understood with-

out context, which I take to be obvious. On the contrary, the thesis that *all* sentences can be at least minimally understood without the context of their utterance seems absurd to me.

What is the position of minimalism toward the idea of a semantic code, then? Semantic minimalism tries to be neutral: “[...] we want Semantic Minimalism to be neutral between the different accounts of how best to assign semantic values to linguistic expressions; e.g. objects, sets, properties, functions, conceptual roles, stereotypes, or whatever.” (Cappelen – Lepore 2005, 144) A further step on the way from semantic code to minimal proposition is the employment of rules for the composition of meanings; here, minimalism is also neutral. In spite of all that neutrality, it is necessary for these meanings and rules to be shared, as the derivation of the minimal proposition must be deterministic (according to E. Borg, even formal). In order for the minimal proposition to be able to serve as a condition of communication, we need a shared semantic code and shared rules of composition and saturation, even if minimalists are reluctant to describe the nature of the code. Thus it seems that semantic code and minimal proposition come rather close.

As outlined above, homonymy and polysemy present a problem for understanding language as a code. Contextualism deals with these phenomena at the level of negotiation of meaning. In minimalism, the derivation of minimal meaning requires disambiguation. However, there is the suspicion that disambiguation is a primary pragmatic process which enters the derivation of what is said before the determination of truth conditions.

Cappelen and Lepore describe the derivation of the minimal proposition as follows (Cappelen – Lepore 2005, 144f):

- (a) Specification of the meaning of every expression in the sentence.
- (b) Specification of the compositional rules.
- (c) Disambiguation of every ambiguous expression, where the ambiguity is due to homonymy or polysemy.
- (d) Clarification of vague expressions.
- (e) Saturation of semantic values of every context dependent expression.

Here, disambiguation is at the same level as saturation. Thus, it is a sub-propositional process. If saturation is the only process at this level that can interact with context, disambiguation must be independent of context.

The above-mentioned objection that it is unclear whether all utterances of a sentence have something in common is not only a matter of metaphysics, as it is claimed by Cappelen and Lepore. The objection that there is nothing that all utterances of the sentence “That is a hard thing” have in common can be answered by a semantic minimalist as follows: If you have found nothing in common, it may be because the term “hard” has multiple meanings. Get back and perform disambiguation thoroughly. In some cases, disambiguation can be done during the derivation of the logical form of the sentence;¹⁴ in other cases, disambiguation requires information from context such as the topic of conversation.

Emma Borg (2004, 142) deals with the problem in more detail and proposes the following solution: Homonymous expressions need to be disambiguated *before* semantic processing, if possible. If this is not possible, it is necessary to semantically process all the varieties of meaning, and pragmatically resolve them only later. Borg even allows the existence of homonymy that needs to be handled during semantic processing, if it is a case of *top-down feedback*, i.e. the gradual exclusion of contextually anomalous interpretations. This process differs from real pragmatic processes such as free enrichment in that no further information is added to what is said.

Disambiguation is a bottom-up process, as it is triggered by the presence of a homonymous expression. A top-down process is only the exclusion of anomalous interpretations. In this case we have here something like the above-mentioned Grice’s circle by Levinson, as later implicatures can influence what is said. It would be devastating for the minimalist program if a process of disambiguation would require information that was not lexically encoded. In the following, I will try to argue in this direction.

Disambiguation needs to proceed with respect to some code that maps each expression to a discrete set of meanings. But this is a simplification, as language meanings cannot really be understood as discrete sets. If meaning is holistic to an extent, criteria of identity for particular meanings cannot be exactly specified. So far I have spoken mainly

¹⁴ The sentences “They were armed with bows” and “He bows deeply” could be disambiguated in this manner, for the word “bows” must be a noun in the former sentence and a verb in the latter.

about the rather rare phenomenon of homonymy. By contrast, polysemy is a much more common phenomenon, one closely related to a holistic understanding of language. An expression is polysemous if it has multiple meanings that are in some way related, often one of the meanings being metonymically or metaphorically derived from another. Homonymous expressions are mostly considered distinct lexical units (and thus in need of indexing), whereas a polysemous expression has only one entry in the lexicon. Thus, it depends on the choice of the code whether, and to what extent polysemous expressions are considered related or quite different, i.e. whether they will be reduced to homonymy. If the code is fine-grained enough, disambiguation that uses the code requires finding a correct metonymic interpretation of many expressions.

Let us have a look at how the phenomenon of metaphor and metonymy is approached by minimalists, namely Emma Borg (2001). The interpretation of metaphor is based on a figurative interpretation function which assigns the set of (all) possible interpretations $p_1 \dots p_n$ to each literal meaning of metaphor p together with conceptual framework α :

$$f\langle p, \alpha \rangle = \{ p_1 \dots p_n \}$$

A conceptual framework is made up of relationships between concepts that are common in the given community or concepts of the particular speaker. Importantly, these relationships can depend on information that is not lexically encoded, i.e. on knowledge of the world. In the second step, a subset of those interpretations that are relevant in the given context needs to be chosen. The contextual factor c comes into play:

$$f^c\langle p, \alpha, c \rangle = \{ p_1^c \dots p_n^c \}$$

This second step, which involves the restriction of interpretations that are unsuitable in the given context, is analogical to the above-mentioned top-down feedback, in which anomalous interpretations of homonymous expressions are excluded. This process poses no problems for the minimalist derivation of meaning.

What is important for further consideration is the first step. If finding metaphorical interpretations requires extra-linguistic information, this information is probably in some way contained in the interpretations. If the same process of interpretation were active with polyse-

mous expressions, we would be dealing with a truly pragmatic process that precedes the determination of truth conditions, which would be fatal for minimalism.

Disambiguation of homonymous expressions is activated on the syntactic or semantic level and is mandatory (even if it were performed only after semantic processing). By contrast, disambiguation of polysemous expressions is not always mandatory. Thus, it can be governed only by pragmatic decisions. Consequently, the semantic information contained in a semantic code is insufficient for a complete determination of the meaning of an utterance, typically for truth conditions. But is this really the case?

This is exactly what contextualists claim. One of their types of modulation is “predicate transfer” or “metonymic transfer” (Recanati 2010, 166-168). It is a typically top-down process which is in many cases optional, i.e. it is triggered purely pragmatically. Recanati gives the following example:

There is a lion in the courtyard.

The expression “lion” can refer to a living lion or it can be metonymically transferred to a statue of a lion. A further metonymic transfer is possible, by which “lion” refers to a person born under the sign of Leo. Thus, the expression “lion” is polysemous in this way. Polysemy can even be defined as conventionalized modulation (Recanati 2010, 70f). This is where a way out for minimalism could be found. If it were possible to conventionalize all the information that is contained in polysemous expressions, there would be no need for pragmatic decisions. If we reduced polysemy to homonymy, the conceptual relations that, according to Emma Borg, are necessary for metaphorical interpretations would be lost. If minimalists succeed in including these relations into a semantic code, disambiguation of polysemous expressions would not be counted as a primary pragmatic process.

The last sentences of the previous paragraph are stated using conditionals, as I do not know of a conception of code robust enough to include all conceptual relations. It is an open question whether such a code can be constructed using finite means. Even if this were possible, it is unclear that such a code would be psychologically adequate (for instance, because of massive parallelism).

4 Minimalism Is (Radical) Contextualism

From the above discussion, it could seem that there are only a couple of varieties of contextualism and minimalism which cover all the possible semantic and pragmatic approaches to formulating a systematic theory of meaning and its interaction with context. This, of course, is not the case. Minimalism and contextualism are merely extreme positions in some sense, and there are several approaches to the unification of their various features.¹⁵

One group of approaches includes theories that extend the *basic set* of indexical expressions (as understood by minimalism) with further expressions or classes of expressions. Among these, let us mention Jason Stanley's theory of hidden indexical expressions, which places further indexical parameters on the sentence level, or the predicate indexicalism of Rothschild and Segal, which extends the set of indexical expressions with predicate expressions. In indexicalism, these further indexical expressions can be detected using syntactic methods and all semantic content has a propositional character (similarly as in minimalism).

Another group of approaches includes theories that extend the concept of proposition with a further parameter whose content needs to be gained from context. Thus, truth values depend not only on the possible world but also on a further parameter (we can call it a *generalized proposition*). If the proposition expressed in sentence *p* varies with context, it is possible to add this further parameter. The generalized proposition thereby becomes context invariant. It is a typically semantic approach consisting in an effort to capture context dependence by semantic means.¹⁶

An example of this approach is the non-indexical contextualism of John MacFarlane (2007). It extends proposition with a "count-as" parameter as a function from properties (resp. expressions that stand for properties) to intentions. For each property (e.g. big), the intention that

¹⁵ An overview of various positions can be found in Borg (2010a).

¹⁶ We can distinguish two kinds of context dependence: An expression or sentence is context dependent if it has different meanings in different contexts or if its truth depends on context. By adding a further parameter to proposition, we convert the former kind of context dependence to the latter one. Compare Cappelen – Lepore (2005, 146) or MacFarlane (2007).

is expressed by the property (e.g. big for a dog) is determined using context. This approach is reminiscent of Recanati's proposal to include a contrast set. As is admitted by MacFarlane himself, his proposal approaches radical contextualism, whose basic tenet is that a proposition is created only in context. We can see that a small modification of, or a small departure from semantic minimalism leads directly to radical contextualism.

The following line of argumentation is typical of proponents of minimalism (as an instance of *modus tollens*):¹⁷

Even a small departure from minimalism collapses into radical contextualism.

Radical contextualism is absurd and internally inconsistent.

No reasonable modification of minimalism is possible.

Proponents of radical contextualism can turn the argument the other way around while keeping the first premise:

Even a small departure from minimalism collapses into radical contextualism.

Semantic minimalism is inadequate and in need of modification.

Radical contextualism is the only acceptable position.

Both these arguments suggest that there is not much of a difference between minimalism and radical contextualism. If minimalists failed to solve the problems connected with the disambiguation of polysemous expressions, minimalism would collapse into radical contextualism without any modification. Thus, the notion of a code is at a crossroads. If a code robust enough to deal with polysemy could be created, minimalist conceptions would present a new impetus towards understanding language as a code. If such a code could not be created, very little would be left of the notion of a persistent code, other than perhaps a provisional code that could, for a certain time and in a certain language community, help in interpreting meaning or be active in the process of creating it. As Davidson has already claimed, it is a practical crutch.¹⁸

¹⁷ This is the main argument from Cappelen - Lepore (2005, 14) and also from Borg (2010a).

¹⁸ Supported by project GACR P401/11/P174.

Department of Philosophy
 Faculty of Arts
 Masaryk University Brno
 Arna Nováka 1
 602 00 Brno
 Czech Republic
 macha@mail.muni.cz

References

- BACH, K. (2004): Conversational Implicature. *Mind & Language* 9, 124-162.
- BACH, K. (2011): Review of Francois Recanati, Truth-Conditional Pragmatics. *Notre Dame Philosophical Reviews*. Available from: <http://ndpr.nd.edu/news/25657-truth-conditional-pragmatics-2/>
- BAPTISTA, L. (forthcoming): Say What? On Grice On What Is Said. *European Journal of Philosophy*.
- BEZUIDENHOUT, A. (2002): Truth-Conditional Pragmatics. *Noûs* 36, 105-134.
- BORG, E. (2001): An Expedition Abroad: Metaphor, Thought, and Reporting. *Midwest Studies in Philosophy* 25, 227-248.
- BORG, E. (2004): *Minimal Semantics*. Oxford: Oxford University Press.
- BORG, E. (2009): On Three Theories of Implicature: Default Theory, Relevance and Minimalism. *International Review of Pragmatics* 1, 1-21.
- BORG, E. (2010a): Meaning and Context: A Survey of a Contemporary Debate. In: Whiting, D. (ed.) *The Later Wittgenstein on Language*. Hampshire: Palgrave Macmillan, 96-113.
- BORG, E. (2010b): Minimalism and the Content of the Lexicon. In: Baptista, L. - Rast, E. (eds.) *Meaning and Context*. Bern: Peter Lang, 51-78.
- CAPPELEN, H. - LEPORE, E. (2005): *Insensitive Semantics: A Defense of Semantic Minimalism and Speech Act Pluralism*. Oxford: Blackwell.
- DAVIDSON, D. (2001a): Communication and Convention. In: Davidson, D.: *Inquiries into Truth and Interpretation* (2nd ed.). Oxford: Clarendon Press.
- DAVIDSON, D. (2001b): Truth and Meaning. In: Davidson, D.: *Inquiries into Truth and Interpretation* (2nd ed.). Oxford: Clarendon Press.
- DAVIDSON, D. (2005): A Nice Derangement of Epitaphs. In: Davidson, D.: *Truth, Language and History: Philosophical Essays*. Oxford: Clarendon Press.
- GARCÍA-CARPINTERO, M. (2006): Recanati on the Semantics/Pragmatics Distinction. *Crítica* 38, 35-68.
- JASZCZOLT, K. M. (2005): Review of E. Borg, *Minimal Semantics*. *Journal of Linguistics* 41, 637-642.
- MACFARLANE, J. (2007): Semantic Minimalism and Nonindexical Contextualism. In: Preyer, G. - Peter, G. (eds.) *Context-Sensitivity and Semantic Minimalism: New Essays on Semantics and Pragmatics*. Oxford: Oxford University Press, 240-250.

- RECANATI, F. (1993): *Direct Reference: From Language to Thought*. Oxford: Wiley-Blackwell.
- RECANATI, F. (2001): What is Said. *Synthese* 128, 75-91.
- RECANATI, F. (2002): Unarticulated Constituents. *Linguistics and Philosophy* 25, 299-345.
- RECANATI, F. (2003): The Limits of Expressibility. In: Smith, B. (ed.) *John Searle*. Cambridge: Cambridge University Press.
- RECANATI, F. (2004): *Literal Meaning*. Cambridge: Cambridge University Press.
- RECANATI, F. (2010a): *Truth-conditional Pragmatics*. Oxford: Oxford University Press.
- RECANATI, F. (2010b): Varieties of Contextualism. In: Achourioti, T. A. (ed.): *Proceedings of the Amsterdam Graduate Philosophy Conference - Meaning and Truth*. Amsterdam: ILLC Publications, 1-8.
- SOAMES, S. (2002): *Beyond Rigidity*. Oxford: Oxford University Press.
- SPERBER, D. – WILSON, D. (2008): A Deflationary Account of Metaphor. In: Gibbs, R. (ed.): *The Cambridge Handbook of Metaphor and Thought*. Cambridge: Cambridge University Press, 171-203.
- WITTGENSTEIN, L. (1958): *Philosophical Investigations*. Oxford: Blackwell.
- WITTGENSTEIN, L. (1974): *Philosophical Grammar*. Oxford: Blackwell.

Pojem *problému* z hlediska teorie konstrukcí

Pavel Materna

Akademie věd České republiky, Praha

Abstract: Transparent Intensional Logic (TIL) explicates objective abstract procedures as so-called *constructions*. Constructions that do not contain free variables and are in a well-defined sense ‘normalized’ are called *concepts* in TIL. An argument is given for the claim that *every concept defines a problem*. The paper treats just mathematical concepts, and so mathematical problems, and tries to show that this view makes it possible to take into account some links between conceptual systems and the ways how to replace a non-effective formulation of a problem by an effective one. To show this *in concreto* a well-known Kleene’s idea from his (1952) is exemplified and explained in terms of conceptual systems so that a threatening paradox is avoided.

Keywords: Transparent Intensional Logic (TIL), constructions, problem, concept, conceptual system.

1 Problém a pojem

Při čtení slavné Kolmogorovovy stati (1932) musí logického analytika přirozeného jazyka napadnout, že pojem *problém* si zaslouží explikaci z hlediska logické analýzy přirozeného jazyka, zejména když sám Kolmogorov problémy formuluje v přirozeném jazyce (obsahujícím ovšem matematické výrazy).

Zadání i řešení problému je nepochybně spjato s *procedurami*. Není divu, že mi jako nejvhodnější nástroj explikace připadá Tichého Transparentní Intenzionální Logika (TIL), která vyvinula *procedurální sémantiku*.¹

¹ Viz zejména Tichý (1988) a (systematický výklad současné podoby v) Duží – Jespersen – Materna (2010).

Podobně jako Montague (1974) vyšel i Tichý z předpokladu, že logická analýza přirozeného jazyka musí být po formální stránce inspirována nikoli predikátovou logikou (jakéhokoli řádu), nýbrž (typovaným) λ -kalkulem: ten totiž realizuje Churchovu geniální intuici, že logické operace lze v podstatě redukovat na *vytvoření funkce abstrakcí a aplikaci funkce na argumenty*.

TIL je spojena s typovou hierarchií, nejprve 1. řádu (kdy tzv. *konstrukce* jsou pouze užity, aby vytvořily množinové objekty, konkrétně funkce), posléze vyšších řádů, kdy konstrukce jsou také *zmiňovány*, takže se stávají objektem *sui generis*. Konstrukce jsou definovány a pojmenovány (proměnné, Trivializace, Provedení, Dvojí provedení, Kompozice, Uzávěr), což zde nemůžeme reprodukovat, takže odkazujeme hlavně na Duží – Jespersen – Materna (2010). Typy 1. řádu jsou množiny parciálních funkcí nad bází $\{0, \iota, \tau, \omega\}$, tj. po řadě *pravdivostní hodnoty, Univerzum, tj. individua, časové momenty, popř. reálná čísla, a možné světy*.

Rozvětvená hierarchie typů pak určuje typy vyšších řádů. Zhruba můžeme ideu této hierarchie vyjádřit takto: Konstrukce konstruuují objekty typu řádu n . Jsou to *konstrukce řádu n* . Množina všech konstrukcí řádu n tvoří nový typ: *typ řádu $n + 1$* , značený $*_n$. Tak například proměnná x v -konstruuující čísla, tedy objekty typu řádu 1, je konstrukce řádu 1 a je typu $*_1$. Je to tedy objekt typu řádu 2. Když v TIL zmiňujeme konstrukce, nemusíme kvůli tomu potřebovat metajazyk.

Mějme konstrukci C , která neobsahuje volnou proměnnou². C je *pojem*.

Příklad: S A. Churchem tvrdíme, že každý smysluplný (ne výhradně syntaktický) výraz vyjadřuje pojem (viz Church 1956), tj. zvolíme jakýkoli výraz (zde: češtiny), včetně věty. Například:

$3 > 2$ vyjadřuje pojem: $[^0 > ^0 3 ^0 2]$, kde [...] je instrukce k vytvoření Kompozice,⁰ vytváří Trivializaci a typy jsou $3, 2 / \tau, > / (\sigma\tau)$. Tento pojem konstruuje pravdivostní hodnotu T.

Čísla (reálná) větší než 0 vyjadřuje pojem $\lambda x [^0 > x ^0 0]$, konstruuje třídu kladných reálných čísel, apod.

Budeme se zabývat matematickými pojmy, tj. naše báze obsahuje jen 0 a τ , kde τ bude množina *přirozených čísel*. Následující tvrzení se týká i empirických pojmů, dále půjde už jen o pojmy matematické.

² Pro přesnost dodejme „a je jistým dobře definovaným způsobem normalizována“.

Tvrzení 1: Každý pojem definuje určitý problém. –

Argument: Pojem jakožto konstrukce je možno chápat jako návod na proceduru vedoucí k určitému výsledku. Když formulujeme problém, tak nás zajímá výsledek nějaké procedury.

Poznámka 1: Tichý při analýze tázacích vět (1978) rozlišuje sémantické jádro (*topic*) a pragmatický postoj. Tak jako můžeme mít k téže propozici postoj tvrzení, tázání nebo například chtění, tak i v našem případě můžeme mít k danému pojmu postoj spočívající v provádění dané konstrukce, ale i postoj navádění k proceduře. Ta procedura (pojem) je v obou případech stejná.

Poznámka 2: V případě otevřené konstrukce, tj. konstrukce obsahující aspoň jednu volnou proměnnou, je situace následující: Nechť konstrukce C obsahuje jednu volnou proměnnou, řekněme x . Pak pojem určený jednoznačně příslušným pojmem je konstrukce $\lambda x C$. Při $n > 1$ je více příslušných pojmů, a tedy i problémů. Například k otevřené konstrukci

$$[{}^0 > x y]$$

existují čtyři příslušné pojmy a tedy problémy:

$$\lambda x y [{}^0 > x y]$$

$$\lambda y x [{}^0 > x y]$$

$$\lambda x \lambda y [{}^0 > x y]$$

$$\lambda y \lambda x [{}^0 > x y]$$

Podle některých vykladačů pojmu *problém* by explikace měla zahrnovat jen případy, v nichž je obsažena nějaká netriviální obtíž. Např. právě uvedené konstrukce by pro ně jistě nereprezentovaly problémy. Je však velice vágní i nejednoznačné definovat stupeň obtížnosti, od něhož bychom teprve mluvili o problémech. Krom toho by takové kritérium znamenalo zavádění do explikace relativismu, resp. subjektivismu: Shora uvedené příklady představují jistě určitý stupeň obtížnosti (a vypočítat $3 + 5$, umí to každý v každém věku?).

2 Algoritmické a nealgoritmické pojmy

Pojem konstrukce v TIL není totožný s pojmem algoritmu. (Srovnej Tichý 1986.) Jednak může jít pouze o algoritmický výpočet jako např. $(3 + 5)$, jednak konstrukce jsou definovány i v těch případech, kdy není zaručena efektivní metoda (například se uznává jako legitimní aplika-

ce funkce na argument, aniž je vyžadován finitní charakter argumentu nebo efektivnost funkce). Protože pojmy jsou konstrukce, můžeme tedy rozlišit *pojmy algoritmické*, definující efektivní (rekurzivní, algoritmickou) proceduru, a *pojmy nealgoritmické*, kde příslušná procedura není efektivní.

Nealgoritmické pojmy jsou však dvojího druhu:

- a) pojmy, které jsou nealgoritmické a nejsou ekvivalentní žádnému algoritmickému pojmu, jako např. pojem množiny teorémů predikátové logiky 2. řádu, a
- b) pojmy, které jsou ekvivalentní určitému algoritmickému pojmu (a samy nejsou).

Jako příklad pojmu sub b) můžeme uvést (zapsáno matematickým jazykem)³:

$$\forall \lambda abc n (n > 2 \supset \neg [a^n + b^n = c^n]) \quad (\text{Fermatův poslední teorém})$$

Můžeme tedy definovat:

Definice 1

Problém indukovaný pojmem C: algoritmická konstrukce C' ekvivalentní C . –

Tvrzení 2: *Je-li pojem C nealgoritmický, pak problém logicky ekvivalentní je rovněž nealgoritmický.* –

Důkaz: Necht' konstrukce C' je logicky ekvivalentní konstrukci C . To znamená, že logická forma C je logicky ekvivalentní C' . Tyto logické formy se tedy liší pouze logickými konstantami. Je-li jedna z konstrukcí C , C' nealgoritmická, pak druhá bude rovněž nealgoritmická, protože logicky ekvivalentní transformací se z neefektivní kombinace logických konstant nestane kombinace efektivní.

Příklad: Je-li konstrukce $\forall \lambda x C$ nealgoritmická, pak logicky ekvivalentní konstrukce

$$\neg \exists \lambda x \neg C$$

je rovněž nealgoritmická.

³ Výrazy jazyka matematiky jsou blízké způsobu, jakým zapisujeme konstrukce. Např. srovnej výraz $(3 + 5)$ a zápis konstrukce $[^0 + ^03 ^05]$.

Vezměme konkrétně problém *prvočíselných čtveřic* (*prime quadruplets*): chceme zjistit, zda je nekonečně mnoho za sebou jdoucích čtveřic prvočísel s intervalem 0, 2, 6, 8. Nechť tedy *Kard(inalita)* je počet (v našem případě tříd o čtyřech prvcích), *c* je proměnná tříd čísel, *n, p* nechť jsou proměnné přirozených čísel. Náš problém lze napsat jako konstrukci (pojem), jejíž zápis bude opět proveden (kvůli jednoduchosti) běžným matematickým výrazem:

$$\exists n (n = \text{Kard } \lambda c (\exists p [[\text{Prvočíslo } p] \wedge [c = \lambda x [x = p] \vee [x = p + 2] \vee [x = p + 6] \vee [x = p + 8]]]])$$

(Jde o dosud nevyřešený problém. V kladném případě je těchto čtveřic konečně mnoho.)

Takto zadaný problém je nealgoritmický. Po provedení jakékoli logicky ekvivalentní transformace se na této skutečnosti nic nezmění. (Například $\neg \forall n \neg (\dots)$.)

Tvrzení 2 však může vést ke sporu! Lze totiž snadno dokázat:

Fakt: Třída těch pojmů *C*, které nejsou algoritmické a ke kterým byl nalezen problém *C'* tímto pojmem indukovaný není prázdná.

(Paradigmatickým příkladem je Fermatův poslední teorém.) Znamená to, že nalezený pojem/problém je ekvivalentní, ale nikoli logicky ekvivalentní.

Definice 2

Nechť = / ($o*_n*_n$) je relace mezi dvěma pojmy, které konstruují stejný objekt nebo jsou oba prázdné.

Problém *C* je ekvivalentní problému *D*, jestliže existuje konstrukce *E* tvaru [*C* = *D*] konstruující *T*. –

(Je-li konstrukce *E* algoritmická, pak je problémem indukovaným pojmem *C*. Viz **Definici 1**.)

3 Pojmové systémy

Kleene ve své klasické knize (1952, 317) upozorňuje na jeden vysoce důležitý fakt. Nechť *R* je efektivně vyčíslitelná relace. Definujme (schematicky) následující funkci:

$$\mu y R(x, y) \begin{cases} \text{Nejmenší číslo takové, že platí } R(x,y), \text{ jestliže } \exists y R(x,y) \\ 0 \text{ v opačném případě} \end{cases}$$

Tato definice neurčuje ovšem efektivní proceduru. Jak však ukazuje Kleene, existují takové instance relace R , pro něž efektivní procedura existuje, ale k níž se nedostaneme na základě shora uvedené definice.

Nyní si rozvedeme příslušný příklad, který Kleene zmiňuje. Necht' a, b, c jsou určitá celá čísla a R je definováno takto:

$$\lambda xy [ax + by = c]$$

Problém, který má Kleene na mysli, je⁴

$$\mathbf{P} \quad \lambda x [\mu \lambda y [ax + by = c]]$$

Tento problém je nealgoritmický: nemáme záruku, že $\exists y [ax + by = c]$ pro všechna x .

Mějme nyní gcd jako Euklidův algoritmus pro nalezení největšího společného dělitele dvou čísel. Problém

$$\mathbf{P}' \quad \lambda x [\mu \lambda y [gcd(a, b) = ax + by]]$$

je ovšem algoritmický. Lze nyní dokázat, že funkce, kterou konstruuje \mathbf{P}' , má pro

$$c = gcd(a, b)$$

stejně hodnoty jako funkce, kterou konstruuje \mathbf{P} , a pro

$$c \neq gcd(a, b)$$

nemají obě funkce hodnotu (resp. mají hodnotu 0 podle Kleeneovy definice). \mathbf{P} a \mathbf{P}' jsou tedy různé ekvivalentní konstrukce. (V TIL máme $\mathbf{P} = \mathbf{P}'$, ale ${}^0\mathbf{P} \neq {}^0\mathbf{P}'$.) \mathbf{P}' je tedy problém indukovaný problémem (pojmem) \mathbf{P} .

Předpokládejme, že chceme najít pojem indukovaný pojmem \mathbf{P} a ,nemáme k dispozici' pojem, který nazýváme Euklidovým algoritmem. Zřejmě nám nepomůže žádná logicky ekvivalentní úprava (**Tvrzení 2**) a k cíli nedojdeme. V okamžiku, kdy mezi použitelné pojmy zařadíme Euklidův algoritmus, je náš úkol splněn. Tento příklad ilustruje užitečnost zavedení dynamického měřítka naší pojmové kompetence: budeme definovat *pojmový systém*.⁵

⁴ Opět: příslušná konstrukce, která je analýzou uvedeného matematického zápisu.

⁵ Viz Materna (1998, 2004) a ovšem Duží – Jespersen – Materna (2010).

Definice 3

Nechť X je objekt, který není konstrukce. Pak 0X je *jednoduchý pojem*. Jednoduché pojmy nerozkládáme na složky a předpokládáme, že jsou dány.

Definice 4

Mějme konečnou množinu *Prim* $\{P_1, \dots, P_n\}$ jednoduchých pojmů a rozvětvenou hierarchii typů. Pojmy z *Prim* nazýváme *primitivní pojmy* a pojmy z nekonečné množiny *Der* pojmů odvozených z *Prim* na základě rozvětvené hierarchie typů pojmy *odvozenými (derived)*. Sjednocení $Prim \cup Der$ nazveme *pojmový systém*. –

Pojmový systém je jednoznačně určen množinou *Prim* a určuje, o kterých objektech můžeme mluvit. Je to ‚pojmová výbava‘, jakou máme díky tomuto systému k dispozici. Zde uvažujeme pouze matematické pojmy.

Je zřejmé, že nalézání efektivních postupů (algoritmických pojmů, které jsou indukovány daným pojmem) je závislé na tom, jaký pojmový systém máme k dispozici. Náš příklad ilustruje tento fakt. Zároveň můžeme fakt, že pojmové systémy se mění a celkově rozšiřují, chápat i jako podklad oprávněného matematického optimismu, který vyjádřil Gödel ve své filozofické kritice Turingova díla (1990, 306).⁶

To zřejmě neznamená, že vznikne pojmový systém, který umožní najít problém indukovaný pojmem *množina teorémů predikátové logiky 2. řádu* apod.⁷

Filozofický ústav
Akademie věd České republiky
Jilská 1
110 00 Praha 1
Česká republika
maternapavel@seznam.cz

⁶ „What Turing disregards completely is the fact that mind, in its use, is not static, but constantly developing, ...and that more and more abstract terms enter the sphere of our understanding.“

⁷ Tato práce vznikla s podporou Grantové agentury České republiky, projekt P401/10/0792.

Literatura

- CHURCH, A. (1956): *Introduction to Mathematical Logic*. Princeton: Princeton University Press.
- DUŽI, M. – JESPERSEN, B. – MATERNA, P. (2010): *Procedural Semantics for Hyperintensional Logic*. Berlin: Springer Verlag.
- GÖDEL, K. (1972): Some remarks on the Undecidability Results. In: Gödel, K. (2004): *Collected works, Vol. II*. New York – Oxford: Oxford University Press.
- KLEENE, S. C. (1952): *Introduction to Metamathematics*. New York, Toronto: D. van Nostrand Company.
- KOLMOGOROV, A. (1932): Zur Deutung der intuitionistischen Logik. *Mathematische Zeitschrift* 35, 58-65.
- MATERNA, P. (1998): *Concepts and Objects*. *Acta Philosophica Fennica* 63. Helsinki: Societas Philosophica Fennica.
- MATERNA, P. (2004): *Conceptual Systems*. Berlin: Logos Verlag.
- MONTAGUE, R. (1974): *Formal Philosophy: Selected Papers of Richard Montague*. Edited by R. Thomason. New Haven: Yale University Press.
- TICHÝ, P. (1988): *The Foundations of Frege's Logic*. Berlin – New York: De Gruyter.

Textualizmus a právny pozitivizmus z hľadiska filozofie jazyka

Marek Neština

Univerzita Komenského v Bratislave

Abstract: This paper addresses the problem which consists in the cases when the court properly makes substantive evaluative judgments in coming to a decision that determines what the law is. Since legal positivism says that one can determine what the law is without making any such judgments, there may seem to be a conflict here. The discussion of legal positivism throws light on contemporary disputes about the doctrine of interpretation known as “textualism.” S. Soames (textualism₂) states that the content of a legal text is what the lawmakers say, assert, or stipulate in adopting the text. A. Scalia (textualism₁) argues that the content of a legal text is the meaning – sometimes “the ordinary meaning” – of the text. The paper shows that legal positivism is consistent with the second definition of textualism.

Keywords: legal positivism, textualism, ordinary meaning, hard case, easy case, philosophy of language, philosophy of law, S. Soames, A. Scalia.

Tvrdenie, že právo je totožné s prijatými zákonmi, prípadne s pravidlami/normami¹, ktoré text zákonov vyjadruje, je súčasťou bežne prijímaných predstáv o právnom pozitivizme.² Spolu s opačným názorom, že platnosť práva závisí taktiež i od etických hodnôt a princípov, ktoré nie sú jeho bezprostrednou súčasťou, tvorí často obsah nášho

¹ Výraz „norma“ a „pravidlo“ budem v rámci tejto práce uvádzať synonymne.

² Právny pozitivizmus je filozoficko-právna koncepcia, ktorá sa týka identifikácie práva a odpovedá na otázku, čo je právo. Všetkým pozitivistickým koncepciám je spoločné, že odpoveď na otázku, čo je právo, závisí výlučne od toho, čo je stanovené a/alebo účinné. T. j., čo je stanovené prostredníctvom na to určenej autority a spoločensky účinné (pozri Alexy 2009, 26).

uvažovania o práve. Na prvý pohľad je však zrejmé, že takéto vymedzenie právneho pozitivizmu je nevyhovujúce, preto neprekvapuje, že ho mnohí autori odmietajú a ponúkajú miesto neho svoje vlastné definície. Napríklad v súčasnej anglosaskej filozofii práva je vymedzenie právneho pozitivizmu spravidla širšie, pričom za prevládajúci môžeme považovať názor, podľa ktorého je právo totožné s určitým druhom sociálneho poriadku, respektíve s určitým typom sociálnych faktov.³

Keďže mojím zámerom nie je formulovať správne vymedzenie právneho pozitivizmu, obídem obsahlu a komplikovanú polemiku, ktorá sa rozprúdila v súčasnej právnej filozofii, a poukážem iba na jeden zaujímavý filozoficko-teoretický problém, ktorý sa v nej vyskytol.⁴ Ak budeme totiž vychádzať zo spomenutej bežnej predstavy, ktorá sa v literatúre zvykne označovať aj priliehavejším výrazom „zákonný pozitivizmus“, môžeme začať uvažovať o teórii interpretácie známej ako textualizmus, podľa ktorej sa pri odhaľovaní obsahu textu právnych predpisov nemusíme a ani nemôžeme spoliehať na normatívne postoje jeho tvorcov, prípadne adresátov.⁵ Keď si uvedomíme, že zámerom právneho pozitivizmu je byť deskriptívnou teóriou práva, ktorá dôsledne rozlišuje medzi tým, čo je, respektíve nie je právo, a tým, čo má byť, respektíve nemá byť právo, bude nám zrejmé, že má s textualizmom mnoho spoločného. Inak povedané, ukáže sa, že medzi otáz-

³ Pod právnym pozitivizmom sa potom rozumie tvrdenie, že existencia a obsah práva závisí od určitých sociálnych faktov, nie od jeho hodnoty. Toto vymedzenie neimplikuje, že hodnoty, ktoré právo môže (ale nemusí) zosobňovať, sú nepoznateľné, bezvýznamné alebo periférne, naopak. Týka sa skôr skutočnosti, že právo je určitým druhom sociálneho poriadku (usporiadania), bez ohľadu na to, či je, alebo nie je v určitom stupni v zhode, napr. s ideálom spravodlivosti. Platnosť určitého konkrétneho práva v určitom štáte alebo spoločnosti je vecou zistiteľných sociálnych štandardov, ktoré jeho oficiálni predstavitelia uznávajú ako autoritatívne; napríklad môže ísť o legislatívne nariadenia, súdne rozhodnutia alebo spoločenské zvyky. Právo sa potom chápe ako sociálna konštrukcia svojho druhu, ktorá sa líši od iných sociálnych konštrukcií regulujúcich správanie jednotlivcov v rámci tej istej spoločnosti. Môžeme preto zjednodušene povedať, že právny pozitivizmus redukuje právo na sociálne fakty (pozri Raz 1979, 37-45).

⁴ Podrobnú a obsahlu diskusiu týkajúcu sa tejto problematiky výborne spracoval Tomáš Sobek (pozri Sobek 2010, 234-306).

⁵ Textualizmus patrí do teórie interpretácie práva; je jednou z jeho koncepcií. Týka sa výkladu textu právneho predpisu a odpovedá na otázku, čo je jeho obsahom (pozri Sobek 2008, 216-234).

kou, čo môžeme odvodiť z textu právnych predpisov, a otázkou, čo je právo, existuje vzájomná previazanosť.

Najznámejšia charakteristika textualizmu je, že obsah textu právneho predpisu je totožný s jeho doslovným sémantickým obsahom, respektíve významom, ktoré majú v ňom použité výrazy v prirodzenom jazyku, t. j. s tzv. bežným významom (the ordinary meaning; pozri Soames 2011, 42). Úlohou súdov je podľa textualistu Antonína Scaliu odhaliť obsah zákonov a iných právnych predpisov, ktorý je totožný s ich významom v prirodzenom jazyku, pričom ako súčasť interpretácie odmieta pátranie po intenciách zákonodarcu. Obáva sa totiž epistemologických problémov, ktoré vedú k tomu, že sudcovia majú pri vkladaní svojich vlastných politických a morálnych preferencií do interpretovaných textov voľnú ruku. Pretože sa chce tomu vyhnúť, považuje za nevyhnutné obmedziť sudcovskú tvorbu práva a zdôrazňuje pritom potrebu zachovať jej demokratické pravidlá (pozri Scalia 1997, 16-18). Takisto Tomáš Sobek v tejto súvislosti poznamenáva, že „textualizmus nie je primárne teóriou nejakého jazykového významu“ (Sobek 2008, 220). Jeho záujem o bežný význam slov je predovšetkým požiadavkou na transparentnú a v tom zmysle politicky korektnú zákonnú reguláciu, ktorá sa vyznačuje deľbou štátnej moci na zákonodarnú a súdnu, pričom inštitucionálna úloha sudcu spočíva hlavne v aplikácii, nie v tvorbe práva (pozri Sobek 2008, 220).

Scaliove obavy zdieľajú tí predstavitelia súčasných koncepcií právneho pozitivizmu, ktorí akceptujú tzv. *oddeľujúcu tézu*, ktorú môžeme zhrnúť do tvrdenia, že neexistuje nijaký pojmovo nevyhnutný súvis medzi právom a morálkou (pozri Alexy 2009, 43). Na základe tohto východiska argumentujú, že nie je úlohou súdov riešiť morálne nezгоды vznikajúce v spoločnosti, pretože tým je poverený legislatívny zbor ako reprezentant ľudu väčšinovým hlasovaním. Z toho pre nich vyplýva nedôvera voči súdnej diskreícii, ktorej chýba potrebná demokratická legitimita ohľadom tvorby práva a kladenie dôrazu na všeobecné pravidlá stanovené prostredníctvom demokratickej legislatívy (pozri Sobek 2010, 292-294). A pretože nie je správne, keď sa nerozlišuje medzi požiadavkami právnych noriem a morálnym názorom jednotlivca (typicky sudcu), parlament by mal formulovať zákony tak, aby ich normatívny význam bol dostatočne zrejмый z ich textov. „Interpreti zákona nemajú povinnosť (a niekedy ani možnosť) zisťovať, aký mali zákonodarcovia (resp. ich väčšina) úmysel alebo aký účel týmto zákonom sledovali“ (Sobek 2010, 296). Navyše jazyk ako spoločne zdieľaný ko-

munikačný kód je v určitom zmysle demokratická inštitúcia, a preto ho nik nemá „úplne pod kontrolou, ani zákonodarca“ (Sobek 2010, 296). V dôsledku toho zámer zákonodarcu je len politický konštrukt týkajúci sa inštitucionálneho postavenia zákonodarcu, ktorý „len odkazuje k tomu, že vôľou zákonodarcu bolo prijať tento text ako zákon“ (Sobek, 2010, 297).

Dá sa teda konštatovať, že právni pozitivistí uprednostňujú textualizmus ako interpretačnú techniku práve preto, lebo vyhovuje ich požiadavkám na nestranné a transparentné stanovenie obsahu práva. A keďže za jediný zdroj práva považujú právne predpisy vydané na to určenou autoritou (typicky parlamentom), pričom ich obsah pokladajú za totožný s obsahom práva, textualizmus súčasne odpovedá na otázku, čo je právo.

Zvažujme preto tento krátky argument:

1. Obsah práva je totožný s obsahom právnych predpisov.
2. Obsah právnych predpisov, pokiaľ v nich nie sú použité výrazy definované priamo v texte právnych predpisov, je totožný s významom, ktoré v nich majú použité výrazy v prirodzenom jazyku.
3. A teda obsah práva, pokiaľ nie je definovaný priamo v texte právneho predpisu, je totožný s významom, ktoré majú výrazy použité v právnych predpisoch v prirodzenom jazyku.

Námietka proti tomuto argumentu, ktorej sa budem ďalej venovať, bude nasledovná: Je ľahko overiteľným faktom, že sudcovia pri interpretácii právneho predpisu občas vytvárajú normatívne súdy, čím zavádzajú rozdiel medzi tým, čo je právo, a tým, čo má byť právom. Keďže (i) textualizmus nepripúšťa tvorbu normatívnych súdov a (ii) právny pozitivismus hovorí, že môžeme určiť, čo je právo bez toho, aby sme stanovili, čo má byť právom, zdá sa, že tento pohľad na súdnu interpretáciu nie je konzistentný ani s textualizmom, ani s právnym pozitivismom.

V mojom príspevku uvediem jednu z odpovedí na uvedenú námietku, ktorá spočíva v obhajobe právneho pozitivismu pomocou odlišného vymedzenia textualizmu. Niektorí autori si totiž v reakcii na uvedenú námietku všimli, že je chybou stotožňovať obsah právnych predpisov, t. j. význam výrazov v prirodzenom jazyku, s ich doslovným sémantickým obsahom a argumentujú, že je potrebné stotožniť ho

s tým, čo je jeho prostredníctvom komunikované, t. j. s jeho komunikovaným, respektíve tvrdeným obsahom. Ak totiž začneme uvažovať o faktickom fungovaní jazyka, je potrebné dôsledne rozlišovať medzi významom výroku ako takeého, významom konkrétnej výpovede, ktorú prostredníctvom tohto výroku robíme v určitom kontexte, a to opäť od toho, čo je touto výpoveďou v danej situácii tvrdené.⁶ V konkrétnej nadväznosti na súčasne podoby filozofie jazyka sa jej predstaviteľ Scott Soames domnieva, že textualizmus vlastne znamená, že obsah práva je totožný s tým, čo zákonodarcovia tvrdili, stanovili alebo inak nariadili prostredníctvom prijatia textu právneho predpisu (Soames 2011, 43).⁷ Tým sa dostávame k dvom odlišným vymedzeniam textualizmu, o ktorých budem viesť diskusiu na ďalších stránkach.

Textualizmus₁, podľa ktorého je obsah právneho predpisu totožný s doslovným sémantickým obsahom (významom) textu (vety, výroku), ktorý má text (veta, výrok) na základe jazykových konvencií pre jednotlivé výrazy v prirodzenom jazyku a spôsobu ich zloženia.

Textualizmus₂, podľa ktorého obsah právneho predpisu je totožný s doslovným sémantickým obsahom (významom) textu (vety, výroku); s ich sémantickým obsahom, ktorý je daný v určitom kontexte; a s asertívnym obsahom výpovede v konkrétnej situácii jej vyslovenia.

1 Zdôvodnenie výberu medzi textualizmom₁ a textualizmom₂

Všeobecné zdôvodnenie pre prijatie textualizmu₂ je, že právne predpisy sú vo svojej najvšeobecnejšej rovine aktmi komunikácie medzi suverénom a zvyškom spoločnosti. Andrei Marmor v tejto súvislosti pripomína, že hoci v tomto prípade, na rozdiel od bežnej komunikácie, ide o asymetrický vzťah, rovnako i tu sa môže prihodiť, že sme jednoducho neporozumeli tomu, čo hovorca oznámil, prípadne relevantnému kontextu, v ktorom to oznámil. Navrhuje preto, aby sme sa pokúsili oddeliť

⁶ Obsiahlejšiu diskusiu na tieto témy ponúka S. Soames v knihe *Philosophy of Language* (pozri Soames 2010, 145-173).

⁷ Podobne A. Marmor tvrdí, že komunikovaný obsah môže tvoriť: sémantický obsah, ktorý je totožný so sémantikou a syntaxou vyjadreného výrazu; asertívny obsah; implikovaný obsah (pozri Marmor 2011, 84).

sémantický a pragmatický aspekt jazyka právnych predpisov a analyzovali tak rôzne zložky, ktoré determinujú to, čo právo hovorí svojim adresátom v určitom konkrétnom kontexte (pozri Marmor 2007, 314).

Takisto S. Soames mieni, že problém vzťahu obsahu práva k jeho autoritatívnym zdrojom, t. j. právnym predpisom, môže byť pokladaný za špeciálny prípad všeobecnejšieho problému, a síce, čo determinuje obsah textov, s ktorými sme oboznámení v bežnom živote (pozri Soames 2009, 403). Tvrdí, že doslovný sémantický význam právneho textu je „príliš strohý a netransparentný na stanovenie jeho obsahu, takže najdôležitejšou úlohou interpretácie je zistiť, čo hovorí text právneho predpisu“ (Soames 2009, 422). Musíme odhaliť, čo relevantní autori povedali alebo k čomu sa iným spôsobom zaviazali tým, že ho prijali, lebo spravidla ten, kto vysloví nejakú vetu v určitom kontexte, sleduje jej vyslovením aj nejaký komunikačný zámer, súčasťou ktorého sú predpoklady o intenciách účastníkov komunikácie. Tieto pragmatické informácie spolupôsobia spolu so sémantickým obsahom viet na doplnení obsahu diskurzu, pričom sémantický obsah môže byť často len prostriedkom na to, aby sme sa dostali k pragmaticky obohatenému obsahu výpovede. Začíname síce doslovným sémantickým významom textu, ale nekončíme ním, keďže v bežných komunikačných situáciách je spravidla dôležitejšie to, čo hovorca mieni komunikovať pomocou svojich výpovedí, než sémantický význam použitej vety (pozri Soames 2009, 404).

2 Zložité a jednoduché prípady

V analytickej filozofii práva začiatok diskusií pojednávajúcich o úlohe jazyka pri určovaní obsahu práva podstatným spôsobom ovplyvnila 7. kapitola knihy *Pojem práva* od H. A. L. Harta. V nej obsiahnutý všeobecný záver, že jazykové aspekty právnych textov vo veľkej miere určujú obsah práva, pričom nechávajú niektoré právne problémy neurčité a otvorené interpretácii, pokladajú mnohí autori za rozumný a správny. V zhode s tým tvrdia, že ďalší filozofický prínos spočíva v rozpracovaní detailov, ktoré sú často nesmierne zložité a komplikované. Ich riešeniu by mohol napomôcť práve pokrok, ktorý bol dosiahnutý v posledných desaťročiach vo filozofii jazyka (pozri Marmor – Soames 2011, 3-4).

Vzťah Harta k filozofii prirodzeného jazyka je mimoriadne komplikovaný, pretože i keď bol ovplyvnený poprednými filozofmi pri-

rodzeného jazyka svojej doby, v samom jeho diele len veľmi zriedka explicitne nájdeme ten typ analýzy, ktorý je pre nich charakteristický. Súvisí to s tým, že jeho hlavným zámerom v uvedenej kapitole nebolo podať analýzu práva z jazykového hľadiska, ale chcel vyvrátiť „skepticizmus ohľadne pravidiel“ (rule skepticism) zastávaný americkými právnymi realistami. Tí podľa neho tvrdili, že právne pravidlá samy osebe len zriedka zaručia to, aký bude výsledok súdneho sporu, keďže môžu byť pochopené takmer akýmkoľvek spôsobom, takže rozhodnutie, ku ktorému nakoniec sudcovia dospeli, nie je spravidla stanovené na ich základe, ale do veľkej miery ho ovplyvňujú iné skutočnosti. Hart argumentoval, že tento „realistický“ obraz právnej praxe je prehnaný, pretože obsahuje jednak pomýlené chápanie jazyka a jednak sa zameriava len na prípady, v ktorých významy použitých viet neposkytujú definitívny návod pre rozhodnutie v takýchto situáciách (pozri Marmor – Soames 2011, 2). Sústredil sa preto na to, aby pomocou úvah o jazyku, ktoré môžeme stručne zhrnúť do nasledujúcich téz, mohol ukázať, v čom tkvie zásadný omyl právneho realizmu:

Formulácia právnych pravidiel v prirodzenom jazyku je závislá od významu všeobecných výrazov.

Význam všeobecných výrazov je daný *inter alia* ich použitím, preto existujú štandardné príklady, v ktorých je ich aplikácia na fakty neproblematická. Z hľadiska významu výrazu ide o jeho významové jadro (the core of meaning).

Väčšina všeobecných výrazov v našom jazyku je *de facto* vágna alebo má *otvorenú textúru*, takže pri ich aplikácii na fakty prípadu sa budú vyskytovať i tzv. hraničné prípady (borderline cases). Z hľadiska významu výrazu ide o jeho okraj, respektíve polotieň (pozri Marmor 2005, 96).

Ak zákon, respektíve právny predpis obsahuje všeobecné výrazy, ktoré sú vágne alebo majú *otvorenú textúru*, aplikácia právneho pravidla nie je v tých prípadoch, ktoré spadajú pod významový okraj výrazu, determinovaná v súlade s jazykovým pravidlom, ale explikáciou predpokladaného účelu právneho predpisu alebo iného relevantného hľadiska. O takomto type prípadu pojednáva Hart ako o zložitom prípade (hard case). Opačná situácia, keď fakty prípadu spadajú pod významové jadro príslušného výrazu, v dôsledku čoho je aplikácia právneho pravidla jasná a bezproblémová, predstavuje podľa Harta jednoduchý prípad (easy case).

Hartovo rozlíšenie medzi zložitými a jednoduchými prípadmi spolu s jeho jednoduchým ilustratívnym príkladom miestnej vyhlášky,

ktorá zakazuje vozidlám vstup do verejných parkov, často uplatňujú filozofi práva pri analýze jazykových aspektov právnych predpisov. Soames v nadväznosti naň formuluje odpoveď na otázku, čo tvorí obsah právneho predpisu. Všíma si pritom charakteristiku zložitých prípadov, ktoré na rozdiel do Harta ďalej delí na tzv. sémanticky zložené prípady (semantically hard cases) a na tzv. rýdzo zložené prípady (genuinely hard cases). Tvrdí, že prípad je sémanticky zložitý vtedy a len vtedy, keď doslovný sémantický význam jazyka právnych predpisov spolu s faktami prípadu nepostačuje na stanovenie (z hľadiska práva) správneho výsledku (pozri Soames 2009, 403). Keď takáto situácia nastane, úlohou interpretácie je zistiť, čo bolo pomocou neho tvrdené. Sémanticky zložené prípady sú podľa neho

- (i) tie, v ktorých tvrdené alebo stanovené intencie zákonodarcov vyplnia medzeru v sémantickom obsahu textu, ktorý nie je v nejakom zásadnom bode artikulovaný;
- (ii) tie, v ktorých zákonodarcovia čiastočne spresnili obsah vágného termínu jeho explicitným použitím (respektíve nepoužitím) na prípad, ktorý by bol za normálnych okolností považovaný za hraničný;
- (iii) tie, v ktorých zákonodarcovia nezámerné chybné opísali alebo stanovili obsah svojho vlastného (primárneho) tvrdenia, ak ide o ľahko vysvetliteľnú neznalosť extenzie použitých výrazov (pozri Soames 2011, 52).

Sémanticky zložené prípady charakterizuje názor, že prirodzený jazyk je veľakrát príliš strohý a netransparentný na to, aby sme len čisto pomocou doslovného významu stanovili obsah právneho textu, preto musíme vziať do úvahy aj zreteľné pozadie predpokladov, ktoré tvoria intencie zákonodarcu. Naopak, ak nastane situácia, že sudcovia i napriek identifikovaniu komunikovaného obsahu zákona nedokážu rozhodnúť súdny spor, ide o rýdzo zložitý prípad. To znamená prípad, v ktorom jeho výsledok nie je determinovaný ani faktami prípadu, ani obsahom právnych predpisov, takže k správne právnemu rozhodnutiu dospejeme len vtedy, keď pomocou inovovaného súdneho rozhodnutia vytvoríme nové právo (pozri Soames 2009, 403). Keďže rýdzo zložené prípady súvisia prevažne s problematikou vágnosti a vo svojom vymedzení sú blízke Hartovým zložitým prípadom, nebudem sa nimi, jednak pre krátkosť svojho príspevku a jednak preto, lebo mojím prvotným zámerom je skúmať úlohou intencií pri vymedzení textualizmu,

na tomto mieste ďalej zaoberať. Navyše urobím ešte jedno obmedzenie, a síce, že sa zameriam len na sémanticky zložité prípady (i).

Ohľadom súdnej interpretácie takisto pokladám za nevyhnutné zdôrazniť, že ak ide o jednoduché prípady, text právneho predpisu neinterpretujeme. Niekedy sa síce v tejto súvislosti píše o doslovnom výklade, ale táto formulácia je zavádzajúca. „Výraz „doslovný výklad“ je v tomto chápaní vlastne nezmyselný, pretože „doslovné“ čítanie zákona znamená jeho „nevýklad““ (Sobek 2008, 224). Navyše tvrdím, že rozlišovanie medzi zložitými a jednoduchými prípadmi je záležitosťou jazykových vlastností textu právneho predpisu, a nie spôsobu, akým sa dosahuje požadovaný výsledok z hľadiska práva. „Potreba interpretácie vzniká len vtedy, keď priame čítanie zákona nedáva z hľadiska potreby aplikácie práva uspokojivú odpoveď“ (Sobek 2008, 224). *In claris non fit interpretatio*, jasný text sa neinterpretuje. V jednoduchom prípade totiž z konvencií, ktorými sa riadi prirodzený jazyk (t. j. jazykových pravidiel), môžeme odvodiť právne pravidlo, nech už je jeho konkrétna explikácia akokoľvek náročná. Naopak, v zložitom prípade jazykové konvencie nepostačujú na to, aby bolo stanovené právne pravidlo.⁸ Čiže cieľom interpretácie je nájsť/vytvoriť právne pravidlo, čo však môže byť v mnohých ohľadoch pomerne ľahké. To znamená, že jednoduchý prípad z hľadiska jazyka právneho predpisu môže byť pomerne zložitý, pokiaľ ide o dosiahnutie konkrétneho riešenia súdneho sporu, ktoré si môže vyžadovať „veľmi sofistikované právne úvahy“ (Sobek 2008, 224). T. Sobek a A. Marmor (pozri Marmor 2005, 97) v tejto súvislosti spoločne uvádzajú názor Josepha Raza, ktorý tvrdil, že rozhodnúť niektoré regulované prípady je omnoho náročnejšie, ako je rozhodnutie vo väčšine neregulovaných prípadov (pozri Raz 1983, 182).⁹

Keď to zhrniem, je zrejmé, že v jednoduchých prípadoch sa právne pravidlo vyskytuje a nech už je jeho explikácia akokoľvek náročná, môžeme ho v princípe uchopiť a opísať. V zložitých prípadoch právne pravidlo chýba, preto ho musíme vytvoriť, pričom samotný akt jeho

⁸ Inak povedané: v jednoduchom prípade jazykové pravidlá determinujú právne pravidlá, avšak môže sa stať, že explikácia právneho pravidla je pomerne náročná. V zložitom prípade, pretože neexistujú jazykové pravidlá, neexistujú ani právne pravidlá, i keď vytvorenie nového právneho pravidla je pomerne jednoduché.

⁹ T. Sobek (pozri Sobek 2010, 26) kritizuje v tejto súvislosti Marmorov prístup. Avšak A. Marmor (pozri 2005, 95-98) podľa mňa dochádza k podobným záverom ako T. Sobek.

tvorby môže byť pomerne jednoduchý. V tejto situácii je však zrejmé, že ide o normatívne rozhodnutie, ktorým sa sudcovia v menšej alebo väčšej miere spolupodieľajú na normotvorbe. A keďže rozlíšenie medzi jednoduchými a zložitými prípadmi sa týka existencie, respektíve neexistencie jazykových pravidiel (t. j. konvencií), vzniká otázka, či v zložitých prípadoch nie sú v skutočnosti k dispozícii nejaké iné pravidlá, ktoré by umožňovali interpretáciu, ktorá by nebola normatívna. To znamená, či odvodenie právneho pravidla je v konečnom dôsledku závislé len od výskytu jazykového pravidla, alebo inak povedané, či obsah práva závisí výlučne len od doslovného sémantického obsahu zákona.

Scott Somes vidí práve prínos filozofie prirodzeného jazyka pre filozofiu práva v tom, že ukáže na pozadí komunikačnej praxe pravidiel, ktoré netvorí súčasť doslovného obsahu hovorcom vyslovených viet. To znamená, že okrem jazykových konvencií máme poruke aj iné pravidlá, ktoré zaručujú, že naša komunikácia prebehne úspešne. Konkrétne usudzuje, že doslovný obsah textu (vety) nebýva spravidla totožný s komunikovaným obsahom, preto to, čo je hovorcom komunikované, je nielen sémantický obsah textu (viet), ale niečo bohatšie, k čomu prispieva nielen jeho význam, ale aj zreteľné pozadie predpokladov. Napríklad to, čo hovorca explicitne vyjadří vyslovením vety (2), môže tiež povedať vetou (3), a to za predpokladu, že kontext urobí zreteľným to, čo hovorca mieni komunikovať výpoveďou, ktorú uskutoční jej použitím.

- (2) a Použil som kladivo ako nástroj na zatĺkanie klinecov.
- b Použil som kladivo ako zbraň na boj so zlodejom.
- c Použil som kladivo ako zarážku na dvere, aby som zabránil zatvoreniu dverí.
- (3) Použil som kladivo.

Slovné spojenie „použil som kladivo“ má svoj bežný význam (ordinary meaning), ktorý je v (2) doplnený o bežný význam frázy, ktorá nasleduje za ním. Ale na rozdiel od vety (2) v prípade vety (3) táto dodatočná informácia môže byť dodaná z kontextu, takže to, čo hovorca *mieni*, môže byť v oboch prípadoch totožné (pozri Soames 2009, 413-414). V nadväznosti na uvedené tvrdí, že o tzv. nenormatívnej interpretácii sa dá hovoriť práve vtedy, keď ide o sémanticky zložité prípady,¹⁰

¹⁰ Na rozdiel od normatívnej interpretácie, o ktorej hovorí vtedy, keď ide o rýdzo zložité prípady.

v ktorých nám umožňujú pravidlá riadiace našu komunikáciu identifikovať vopred existujúci obsah právneho predpisu (pozri Soames, 2011, 52). Kategória zložitých prípadov preto nevyhnutne nemusí viesť k tvorbe práva, ako to naznačoval pôvodný Hartov návrh, ale dovoľuje nám odhaliť podskupinu zložitých prípadov, ktorých riešenie je súčasťou obsahu textu zákona, pričom nie je závislé od jeho doslovného sémantického obsahu.

3 Smith v. The United States

Rozdiel medzi doslovným obsahom vety, „ktorý veta (výrok či veta iného druhu) má na základe jazykových konvencií pre jednoduché podvýrazy a spôsobu zloženia týchto jednoduchých výrazov do vety“ (Zouhar 2011, 725) a komunikovaným obsahom, ktorý pomocou vety hovorca tvrdí, tvorí podstatu súdneho sporu *Smith v. The United States*.¹¹ O čo v ňom išlo?

John Angus Smith ponúkol federálnemu tajnému agentovi automatickú zbraň výmenou za drogy. Následne ho zatkli, obvinili a odsúdili na základe federálneho zákona, ktorý ukladá obžalovanému okrem sankcie stanovenej za trestný čin obchodu s drogami aj ďalší trest, ak „počas a v rámci ... trestného činu obchodu s drogami ... používa ... zbraň“. Smith sa proti rozsudku odvolal, pričom sa jeho prípad dostal až k Najvyššiemu súdu Spojených štátov, ktorý potvrdil pôvodné rozhodnutie. Smith vo svojom odvolaní tvrdil, že dodatočný trest týkajúci sa použitia strelnej zbrane počas a v rámci obchodu s drogami sa vzťahuje len na situácie, v ktorých je strelná zbraň použitá ako prostriedok v boji, takže v zmysle uvedeného federálneho zákona jeho prípad, ktorý spočíva vo výmene strelnej zbrane za drogy, nepredstavuje prípad „použitia strelnej zbrane“.

Najvyšší súd Spojených štátov sa s týmto stanoviskom nestotožnil. Dospel totiž k záveru, že z textu zákona neplynie, že zbraň má byť použitá len ako prostriedok boja, ale týka sa akéhokoľvek použitia zbrane umožňujúceho v akejkolvek podobe spáchať trestný čin obchodu s drogami. Súd sa teda zaoberal otázkou, či opísané konanie predstavuje, respektíve nepredstavuje prípad použitia strelnej zbrane v súvislosti s trestným činom obchodu s drogami.¹² Presnejšie povedané,

¹¹ Viac o prípade je dostupné na http://www.oyez.org/cases/1990-1999/1992/1992_91_8674

¹² Dostupné na http://www.oyez.org/cases/1990-1999/1992/1992_91_8674

či môžeme hovoriť o obchodnej výmene strelnej zbrane za drogy ako o použití strelnej zbrane v rámci významu výrazu, ktorý je relevantný pre výklad právneho predpisu. Majorita súdu tvrdila, že keď slovo nie je definované zákonom, stanovujeme jeho význam spravidla v súlade s významom, ktorý má v prirodzenom jazyku, a preto extenziu výrazu „použitie strelnej zbrane“ tvoria nielen tí, ktorí používajú strelnú zbraň ako prostriedok boja, ale aj tí, ktorí ju používajú ako prostriedok výmeny. Keďže navrhovateľ použil svoju zbraň pri pokuse obstarat' si drogy jej výmenou za kokaín a doslovný význam slovného spojenia „použitie strelnej zbrane“ (v texte zákona) nevylúči toto použitie, môžeme tento jeho čin klasifikovať ako použitie strelnej zbrane, a teda navrhovateľ je subjektom dodatočného trestu.

Námietka druhej strany – Smith a dissent Najvyššieho súdu – rovnako vychádza z predpokladu, že obsah zákona, t. j. to, čo hovorí o pridaných trestoch, je totožný s významom výrazov v prirodzenom jazyku. Na rozdiel od majority súdu však dissent tvrdí, že význam výrazu „použitie strelnej zbrane“ je v prirodzenom jazyku odlišný, a síce použiť ju na účel, pre ktorý je normálne zamýšľaná, čiže ako prostriedok boja. Zdôvodňuje to tým, že použiť niečo spravidla znamená použiť to v súlade so zamýšľaným účelom. Napríklad bežný priemerne informovaný hovorca nebude spájať „použitie strelnej zbrane“ s obchodnou výmenou zbrane za drogy, ale skôr so známejším obrazom, v ktorom sa vyskytuje ako prostriedok boja. Hlavný predstaviteľ súdneho dissentu Antonín Scalia na margo toho poznamenáva, že „keď sa vás niekto opýta: „Používate paličku?“, neskúma, či máte striebornú vychádzkovú paličku svojho starého otca vystavenú v hale, ale chce vedieť, či sa s ňou prechádzate. Keď hovoríme o „použití zbrane“, znamená to, že hovoríme o jej použití na deštruktívne účely, ... ako o bojom prostriedku. A hoci „človek môže použiť strelnú zbraň mnohými spôsobmi“ ...vrátane artiklu výmeny..., nie je to bežný význam tohto výrazu“ (Scalia 1997, 23-24). Nakoľko teda význam slovného spojenia „použitie strelnej zbrane“ v prirodzenom jazyku vylúči obchodnú výmenu strelnej zbrane za drogy z rozsahu významu výrazu, ktorý má v právnom predpise, navrhovateľ nie je subjektom dodatočného trestu.

Soames vo svojom komentári k prípadu konštatuje, že majorita má pravdu, keď tvrdí, (i) že slovné spojenie „použitie strelnej zbrane“ sa vyskytuje v texte vo význame, ktorý má v prirodzenom jazyku, a (ii) že tento význam vopred nevylučuje iné použitie zbrane, ale mylí sa, ak si myslí, že to ukazuje, že obsah právnych predpisov – t. j. právo,

na uzákonenie ktorého bol použitý jazyk právnych predpisov – zahŕňa okrem použitia strelnej zbrane ako prostriedku boja aj iné použitie. Naopak Scalia podľa neho nemá pravdu, keď tvrdí, že bežný význam tohto slovného spojenia vopred vylúči použitie strelnej zbrane z predaja alebo výmeny. Ak sa totiž takýto výraz vyskytuje v nejakej vete, tak na to, aby sme dostali výsledné tvrdenie, musíme obsah výpovede skompletizovať, a to buď explicitným uvedením obsahu ďalšieho slovného spojenia (ako v (2)), alebo prostredníctvom pragmaticky dodaného obsahu z kontextu výpovede (ako v (3)). Keďže realizácia prvej možnosti prislúcha Kongresu Spojených štátov amerických, úlohou súdu bolo odvodiť, čo Kongres tvrdil z nekompletného sémantického obsahu poskytnutého prostredníctvom jazyka právneho predpisu. A teda v prípade, že sa v záznamoch Kongresu nevyskytuje nič, čo by explicitne potvrdzovalo názor majority, mali by sme vychádzať z predpokladu, že Kongres tvrdil, že dodatočný trest by sa mal pridať k trestnému činu, pri ktorom bola použitá strelná zbraň len vtedy, keď bola použitá ako prostriedok v boji (pozri Soames 2009, 414).

Najvyšší súd tento názor rozhodne nezdieľal, pretože tvrdil, že i keď pripustíme, že Kongres mohol mať na mysli bežnejšie použitie strelnej zbrane pri trestnom čine obchodu s drogami, jazyk právneho predpisu nie je nijako limitovaný a neposkytuje teda dôvod, aby jeho prostredníctvom Kongres nezohľadnil aj túto situáciu. Naopak, je podľa jeho názoru predsa evidentné, či už sú zbrane používané ako prostriedok výmeny, alebo pri ochrane nelegálnych obchodov s drogami, že ich samotné zavedenie na drogovú scénu výrazne zvyšuje nebezpečenstvo pre spoločnosť.

Soames zdôrazňuje, že toto stanovisko Najvyššieho súdu je nekonzistentné. Ak bol totiž (i) Kongres tej mienky, že zákaz sa týka len úzkeho rozsahu použitia strelnej zbrane, tak súd nemôže zároveň tvrdiť, že (ii) i keby platilo predchádzajúce tvrdenie, neznamená to, že sa to má i doceliť, keďže použitý jazyk explicitne nevylúčil širokú triedu prípadov, pričom (iii) máme dobrý dôvod na to, aby sme právny predpis aplikovali širšie (pozri Soames 2009, 414). Inými slovami, ak uvedieme, že Kongres tvrdil, že dodatočný trest sa pripája k tomuto úzkemu okruhu prípadov, potom skutočnosť, že jazyk právneho predpisu môže byť použitý (v inom kontexte) i na širší okruh prípadov, je irelevantná, pretože „obsah právnych predpisov je to, čo Kongres povedal prostredníctvom jazyka zákona, nie to, čo mohol a podľa mienky súdu mal povedať“ (Soames 2009, 415).

Súdny spor *Smith v. The United States* má slúžiť ako príklad komunikačnej situácie, ktorá vysvetľuje prijatie textualizmu.² Konkrétne ukazuje, prečo obsah (federálneho) zákona nie je totožný s jeho doslovným sémantickým obsahom (významom) textu (vety, výroku), a teda prečo na odhalenie obsahu, ktorý je prostredníctvom textu zákona tvrdený, potrebujeme zohľadniť intencie zákonodarcu. Spoločné východisko, a síce, že v prípade, ak výraz nie je definovaný v zákone, jeho význam je spravidla totožný s významom, ktorý má v prirodzenom jazyku, vedie k odlišným dôsledkom. Súdne rozhodnutie je založené na prijatí širšieho vymedzenia významu výrazu, takže z dvoch možných interpretácií právneho predpisu volí tú, ktorú pokladá za spoločenského hľadiska za najvhodnejšiu, pričom kritériom na určenie obsahu nie je komunikovaný zámer zákonodarcu, ale hľadisko adresátov právneho predpisu, ktorým je norma určená. Soames uvažuje podobne s tým rozdielom, že za relevantné kritérium pokladá intenciu zákonodarcu a trvá na tom, že interpretácia zákona musí byť s ňou v súlade. A teda hoci pripúšťa, že prirodzený jazyk nám ponúka aj inú možnosť, ako interpretovať zákon, robí tak iba preto, aby zdôraznil, že realizácia tejto možnosti nezodpovedá inštitucionálnemu postaveniu súdu.

Oba diskutované závery spája skutočnosť týkajúca sa vopred existujúceho obsahu právneho predpisu, ktorý nie je súčasťou doslovného sémantického obsahu textu zákona. O podobnej problematike píše Filip Melzer, ktorý v rámci pojmu interpretácia v práve rozlišuje medzi výkladom právnych predpisov a dotváraním práva (pozri Melzer 2010, 79; 116-118; 210-212). Výklad práva, ktorý má blízko k nenormatívnej interpretácii v sémanticky zložitých prípadoch, sa týka hlavne odstraňovania sémantických nejasností, pričom vychádza z pojmu právo, „ktorý zahŕňa vedľa právnych noriem i nepísané právne princípy a hodnoty, ktoré sú základom práva ako takého či konkrétneho právneho odvetvia, právneho inštitútu“ (Melzer 2010, 117). Kritizuje východisko stúpencov normatívnej teórie, ktorí právnu normu stotožňujú s jej jazykovým vyjadrením a pojem právo so súhrnom právnych noriem. Tvrdí, že „ak je legitimita odstraňovania sémantických nejasností popieraná, resp. jej právna povaha, je nutné vždy ju porovnať s alternatívnym riešením. Tým je len voľná úvaha sudcu“ (Melzer 2010, 117). Domnieva sa pritom, že s predvídateľnosťou práva a s princípom delby moci je skôr v zhode postup, ktorý je založený na objektívnych hľadiskách, ktoré majú, ak je to možné, pôvod v samotnom práve.

Výklad práva pomocou odstraňovania sémantických nejasností a nenormatívna interpretácia majú spoločný cieľ a tým je odhaliť vo pred existujúci obsah právneho predpisu. Avšak výklad práva sa netýka len zámerov zákonodarcu, či už reálneho, alebo hypotetického, ale jeho aplikácia je širšia, a preto na jeho základe môže byť metodologicky obhájiteľné aj riešenie, ku ktorému dospel súd. Konkrétne by v tejto situácii bolo možné interpretovať zákon v súlade s argumentmi objektívneho recentného výkladového cieľa (pozri Melzer 2010, 126). Takéto rozhodnutie preto nemusí nevyhnutne znamenať porušenie princípu deľby moci, lebo ak prijmeme predpoklad, že právo tvoria i nepísané právne princípy a hodnoty, ktoré sú základom práva ako takého či konkrétneho právneho odvetvia, právneho inštitútu atď., takáto interpretácia je v súlade s existujúcim právom. Toto porovnanie by nám teda nepomohlo rozhodnúť medzi záverom, ku ktorému dospel súd, respektíve Soames, keďže obe odlišné hľadiská (hovorca verzus adresát), ktoré riadia interpretáciu zákona, sú v princípe právne zdôvodniteľné.

Uvažujme teraz o Scaliovom riešení, ktoré sa zdá byť, nakoľko je jednoznačné, adekvátnejšie. Spočíva totiž v názore, že slovné spojenie „použitie zbrane“ má v prirodzenom jazyku len jeden relevantný význam, ktorý sa zhodou okolností zhoduje s tým, ktorý odhalil Soames pomocou intencií hovorcu, respektíve zákonodarcu. Hoci to na prvý pohľad vyzerá tak, že Scalia odmieta akékoľvek intencie, nie je to celkom tak. Svedčia o tom prinajmenšom príklady, ktoré používa na podporu svojej argumentácie. Odvoláva sa v nich na kontext použitia a na intencie bežných hovorcov v štandardných komunikačných situáciách, pričom na ne prihliada práve vtedy, keď vymedzuje význam výrazu „použitie zbrane“ v prirodzenom jazyku. Zrejme by nás to nemalo prekvapiť, pretože

jeden z úspechov moderného textualizmu bolo vykorenenie starej školy doslovného výkladu, ktorá sa zaujímala len o text, ale nebrala ohľad na kontext zákona. Nemali by sme sa nechať miasť tým, ako textualisti kritizujú svojich oponentov, že sa zaoberajú kontextom ešte po tom, čo už zistili jasný význam zákona. Textualisti totiž hovoria, že extratextuálne faktory máme zohľadňovať už pri zisťovaní toho, čo je jasný význam textu zákona. (Sobek 2008, 220)

Scalia teda nepopiera, že pri určovaní obsahu textu právneho predpisu hrajú úlohu intencie bežných hovorcov, ale odmieta pripísať intenciám zákonodarcu rozhodujúcu úlohu pri interpretácii právnych

predpisov (pozri Scalia 1997, 16-18). Ide totiž o dve odlišné veci. Soames sa totiž približuje skôr stanovisku intencionalistov, ktorí „sa zameriavajú sa na skúmanie legislatívneho zámeru zákonodarcu, keďže zákon chápú predovšetkým ako jeho autoritatívny akt. Zákonodarca má právomoc tvoriť právo, a preto musíme zákony čítať spôsobom, ktorý zodpovedá tomu, ako by ich čítal on sám ako ich autor Intencionalisti poukazujú na skutočnosť, že prirodzený jazyk je neurčitý, a preto potrebujeme skúmať zámer hovorca“ (Sobek 2008, 218). Scalia naopak tvrdí, že rozhodujúcim kritériom by mal zostať text zákona, ktorý by mal byť vykladaný v súlade s prirodzeným jazykom jeho adresátov. V skutočnosti tým neobhaja textualizmus₁, ale verziu textualizmu, ktorá sa v mnohých aspektoch podobá textualizmu₂, takže niektorí autori sa môžu oprávnene domnievať, že hoci svoj názor bráni pomocou textualizmu₁, prinajmenšom implicitne akceptuje textualizmus₂ (pozri Marmor – Soames 2011, pozn. 16).

4 Záver

Pri vymedzovaní textualizmu je nesmierne dôležité, akým spôsobom sa rozhodneme analyzovať prirodzený jazyk a na ktoré jeho aspekty budeme klásť dôraz. Domnievam sa v tejto súvislosti, že otázka dôležitosti úlohy intencií a celkového komunikačného pozadia pri určovaní obsahu právneho predpisu nespočíva v tom, či ich brať, respektíve nebrať do úvahy. Skôr je relevantné pýtať sa, koho intencie sú tie rozhodujúce, prípadne v ktorom časovom úseku a kto všetko v konečnom dôsledku patrí do príslušnej komunikačnej situácie. A teda musíme sa pozrieť na to, akým spôsobom máme modelovať vzťah medzi sémantickými a pragmatickými zložkami interpretácie právneho predpisu. Pretože ak budeme schopní tieto dve zložky oddeliť ostrou líniou, dokážeme tiež rozlíšiť medzi tým, čo tvorí nevyhnutnú súčasť obsahu právneho predpisu, a tým, čo je jeho zrušiteľnou zložkou. V opačnom prípade budeme musieť pripustiť, že právo obsahuje oveľa viac neistoty, než by sme boli ochotní pôvodne pripustiť.

Nemyslím si, že keď textualizmus vymedzíme takým spôsobom, ako to robí Scalia alebo Soames, bude nevyhnutne v rozpore s právnym pozitivizmom. Je to tak preto, lebo porozumenie textu právneho predpisu ostáva hlavným kritériom pri určovaní toho, čo je právo. Zároveň však budeme musieť prijať záver, že vedľa právnych pravidiel/noriem

ho tvoria i nepísané právne princípy a hodnoty, ktoré sú základom práva ako takého či konkrétneho právneho odvetvia, právneho inštitútu. A teda budeme musieť odmietnuť tzv. „zákonný pozitivizmus“ a nahradiť ho primeranejším vymedzením právneho pozitivizmu, ktorý zohľadní vopred existujúci právny obsah.

*Inštitút kontextuálnej teológie
Evanjelická bohoslovecká fakulta
Univerzita Komenského v Bratislave
Bartókova 8
811 02 Bratislava
nestina@fevth.uniba.sk*

Literatúra

- ALEXY, R. (2009): *Pojem a platnosť práva*. Bratislava: Kalligram.
- HART, H. L. A. (2004): *Pojem práva*. Praha: Prostor.
- MARMOR, A. (2011): Can the Law Imply More Than It Says? On Some Pragmatic Aspects of Strategic Speech. In: Marmor, A. – Soames, S. (eds.): *Philosophical Foundations of Language in the Law*. Oxford: Oxford University Press, 83-104.
- MARMOR, A. (2005): *Interpretation and Legal Theory: Revised Second Edition*. Oregon: Hart Publishing.
- MELZER, F. (2010): *Metodologie nalézání práva. Úvod do právní argumentace*. Praha: C. H. Beck.
- NEALE, S. (2007): On Location. In: O'Rourke, M. – Washington, C. (eds.): *Situating Semantics: Essays on the Philosophy of John Perry*. Cambridge: MIT Press, 251-395.
- NEŠTINA, M. (2011): Interpretačný obrat v právnej filozofii. *Organon F* 18, 144-164.
- RAZ, J. (1979): *The Authority of Law: Essays on Law and Morality*. Oxford: Oxford University Press.
- SCALIA, A. (1997): *A Matter of Interpretation*. Princeton: Princeton University Press.
- SOAMES, S. (2009): Interpreting Legal Text: What Is, and What Is Not, Special about the Law. In: Soames, S.: *Philosophical Essays, Volume 1: Natural Language: What it Means and How We Use It*. Princeton: Princeton University Press, 403-423.
- SOAMES, S. (2010): *Philosophy of Language*. Princeton: Princeton University Press.
- SOAMES, S. (2011): What Vagueness and Inconsistency Tell Us about Interpretation. In: Marmor, A. – Soames, S. (eds.): *Philosophical Foundations of Language in the Law*. Oxford: Oxford University Press, 31-57.

- SOBEK, T. (2008): *Argumenty teorie práva*. Praha: Ústav státu a práva.
- SOBEK, T. (2010): *Nemorální právo*. Praha: Ústav státu a práva.
- ZOUHAR, M. (2010): Kontextualizmus a jeho skryté pozadie. In: Zouhar, M. (ed.): *Kontext a význam*. Bratislava: Aleph.
- ZOUHAR, M. (2011): Neartikulované zložky sémantického obsahu a syntaktická elipsa. *Filozofia*, č. 8, 725-745.

Která formule je ta pravá? (Kritéria adekvátnosti logické analýzy)¹

Jaroslav Peregrin - Vladimír Svoboda

Akademie věd České republiky, Praha

Abstract: Transforming natural language sentences into formulas of a formal language (such as that of classical predicate logic) is a common practice that underlies most applications of logic to analysis of our reasoning/argumentation. Is this practice guided by any well established criteria? We argue that the answer is negative. The way from natural language to a formal one is much more tricky and much more arduous than it *prima facie* seems. We sketch a roadmap of this way and strive to explicate the criteria of adequacy of logical formalization that are implicit to the relevant practices. These considerations lead us to conceive logic as a project based on a search for a *reflective equilibrium*. Any formal system deserving the name *logic* must reach a balance between the authority of logical laws over individual arguments and their answerability to intuitive correctness of the bulk of such arguments.

Keywords: logic, argument, logical form, reflective equilibrium.

1 Formalizace úsudků

Jedna z nejběžnějších věcí, které se odehrávají v hodinách různých „úvodů do logiky“, je to, že učitel vyzve studenta, aby nějaký úsudek zapsal jazykem klasické výrokové logiky (KVL) a následně rozhodnul, zda je správný, nebo ne. Může to být třeba úsudek

(U1) Jestliže prší, je mokro
Není mokro

Neprší

¹ Děkujeme Grantové Agentuře ČR, která práci na tomto článku podpořila prostřednictvím financování projektu P401/10/1279.

V takovém případě je student pochválen, zapíše-li něco jako

$$(FU1) \frac{A \rightarrow B \quad \neg B}{\neg A}$$

a konstatuje-li, že jde o logicky správný úsudek.

Ale zatímco vyhodnocení druhé části studentovy odpovědi se řídí jasnými pravidly (KVL jednoznačně určuje, které úsudky zapsané jejím jazykem jsou správné, a které ne), pravidla pro vyhodnocení té první, to jest studentova přepisu úsudku vyjádřeného v přirozeném jazyce do jazyka KVL, nejsou vůbec jasná. Co nám dovoluje konstatovat, že je nějaký takový formální zápis *správný* a jiný *nesprávný* (nebo i jen že je jeden *lepší* než druhý)? V případě, který jsme uvedli, můžeme mít tendenci říkat, že je to tak nějak samozřejmé, ale nemá-li být logická formalizace jazyka – či jeho logická analýza – pouhým uměním, které by bylo lze poměřovat jedině subjektivním dojmem, který vyvolá v kritikovi, je zřejmě třeba se po nějakých takových pravidlech pít.

Abychom měli nějaký zajímavější případ, vezměme následující úsudek:

$$(U2) \frac{\text{Pokud ti nebylo 60 a nebyl jsi v Rusku, pak nemáš právo nám radit} \\ \text{Nebylo ti 60}}{\text{Nemáš právo nám radit}}$$

Přímočarou formalizací tohoto úsudku by se mohlo zdát být

$$(FU2) \frac{(\neg A \wedge \neg B) \rightarrow \neg C \quad \neg A}{\neg C}$$

Při takové formalizaci by ale tento úsudek vycházel jako nesprávný; přičemž intuitivně je (alespoň pro některé mluvčí češtiny, včetně autorů tohoto textu) správný. Alternativní formalizací by tedy mohlo být

$$(FU2') \frac{\neg(A \wedge B) \rightarrow \neg C \quad \neg A}{\neg C}$$

což už by ho vyhodnotilo jako správný. Jak máme obecně rozhodnout, zda větu

(V1) *Pokud ti nebylo 60 a nebyl jsi v Rusku, pak nemáš právo nám radit*
formalizovat jako

$$(FV1) (\neg A \wedge \neg B) \rightarrow \neg C$$

nebo jako

$$(FV1') \neg(A \wedge B) \rightarrow \neg C?$$

Na první pohled se může zdát, že formalizace je jakýmsi přímočarým ‚překladem‘, či dokonce výsledkem prosté náhrady jedněch jazykových symbolů jinými („transverbace“). Vezmeme nějaký symbol, třeba *A*, a nahradíme jím větu *Bylo ti 60*, jiným symbolem, řekněme *B*, nahradíme větu *Byl jsi v Rusku*, a dalším, *C*, větu *Máš právo nám radit*; a symboly \neg , \wedge a \rightarrow pak nahradíme po řadě negací (*ne-* předřazené slovesu věty), spojku *a* a spojení *pokud ... pak ...*.

Situace ale samozřejmě takhle jednoduchá není. I když pomineme negaci, která je v přirozeném jazyce velmi složitým jevem, není situace přímočará ani v případě dalších dvou spojek. Jak víme, že je \rightarrow rozumným protipólem českého spojení *pokud ... pak ...*? Protože se to prostě obecně bere za zřejmé? To asi není úplně dobrý důvod. Protože má výrok tvaru $A \rightarrow B$ stejné pravdivostní podmínky jako odpovídající věta tvaru *Pokud A, pak B*? Ani to nelze tvrdit – výroky tvaru $A \rightarrow B$ jsou v rámci KVL pravdivé právě tehdy, když je výrok *A* nepravdivý nebo je výrok *B* nepravdivý; avšak totéž lze stěží říci o českých větách tvaru *Pokud A, pak B*. K tomu, abychom takovou větu považovali za pravdivou, obvykle nestačí ani to, aby byl nepravdivý výrok *A* ani to, aby byl pravdivý výrok *B*, obvykle se předpokládá nějaká vnitřní (například příčinná) souvislost mezi skutečnostmi, o nichž tyto věty hovoří.

A situace není úplně přímočará ani tehdy, kdy uvažujeme o symbolu \wedge , jakožto protějšku spojky *a*, jakkoli v tomto případě se zdá, že pravdivostní podmínky formule $A \wedge B$ se s pravdivostními podmínkami české věty tvaru *A a B* shodují mnohem lépe. Jak víme, že *a* ve větě (V1) skutečně vyjadřuje konjunkci negací uvedených vět? Protože to tak „na povrchu“, tj. gramaticky, vypadá? Ale gramatika v tomto ohledu zřejmě nemůže být přímočarým vodítkem.² Není-li ale spolehlivým vodítkem

² Nikdo by asi neformalizoval větu *Nikdo není dokonalý* jazykem klasické predikátové logiky jako $\neg \exists x \neg D(x)$, jak by přímočaře odpovídalo její gramatické struktuře.

gramatika a nemůžeme-li formalizaci pojmout jako nějaký přímočarý překlad či ‚transverbací‘, jak vlastně máme formalizaci provádět a jaká jsou kritéria pro to, která formalizace je dobrá a která špatná?³

2 Logická analýza přirozeného jazyka

Klíčem k odpovědi na tyto otázky je, zdá se nám, úvaha o tom, *proč* se vlastně do něčeho takového, jako je formalizace (či logická analýza přirozeného jazyka), pouštíme. Je zřejmé, že tento podnik má co dělat s posuzováním toho, které úsudky jsou v pořádku a které ne. Co nám ale dovoluje něco takového posoudit? Vezměme úsudek

(U3) $\frac{\text{Všichni koně mají hlavu}}{\text{Nějaký kůň má hlavu}}$

Je to správný úsudek, nebo ne? Běžný zápis formy této věty v jazyce klasické predikátové logiky (KPL),

(FU3) $\frac{\forall x (F(x) \rightarrow G(x))}{\exists x (F(x) \wedge G(x))}$,

tento úsudek postihuje jako nesprávný. Přesto si myslíme, že většina mluvčích češtiny by ho měla spíše za správný.⁴ Co nám dává autoritu tvrdit, že je správný (a jeho obvyklá formalizace v KPL, či dokonce KPL sama, je vadná), nebo že je nesprávný (a intuice lidí, kteří si myslí opak, jsou pomýlené)?

Na tuto otázku lze odpovídat (přinejmenším) dvěma způsoby. První vychází z toho, že *logika* je něco objektivního, co je dáno pevně a nezávisle na nás lidech (může být například záležitostí nějakých neměnných abstraktních platónských idejí), a že je to tato transcendentní logika, čím musíme naše úsudky poměřovat. Druhý má naopak za to, že správnost úsudků je věcí toho, jak se kontingentně vyvinul příslušný jazyk, a že pro něj tedy není žádné transcendentní měřítko – jenom měřítko imanentní příslušnému jazyku, jeho inferenční struktura, která

³ Poněkud překvapivě v odborné literatuře nenajdeme příliš pokusů tyto otázky systematicky zodpovědět. Výjimku představují například práce Sainsbury (1991), Baumgartner – Lampert (2008) a zejména Brun (2004).

⁴ To je samozřejmě spekulace, jde o empirické tvrzení, které by muselo být potvrzeno příslušným průzkumem; s jistotou se ale dá říci, že by ho měli za správný někteří mluvčí češtiny.

je kontingentní a je daná tím, jak prostředky tohoto jazyka používají příslušní kompetentní mluvčí.

Na první pohled se může zdát, že je to ta první odpověď, která je jediná možná. Logika přece *není* empirická záležitost, zákony logiky přece nemůžeme ohýbat, jak chceme! Logika je nějak ‚nad‘ jazyky (případně vůbec nade vším), jazyky se jí musejí podřizovat, ne si ji podle libosti vytvářet!

Podívejme se ale na tuto odpověď více zblízka. Zdá se být například nepochybné, že logické pravidlo *modus ponens*, tj. pravidlo

$$(MP) \begin{array}{l} A \rightarrow B \\ A \\ \hline B \end{array}$$

platí zcela neotřesitelně a bez ohledu na to, jakými peripetemi se ubíral a ubírá vývoj češtiny či jiného faktického jazyka. Je jasné, že nakolik se implikace, tj. \rightarrow , řídí obvyklou pravdivostní tabulkou, je pravidlo (MP) nezvratné. Co nám to ale říká o úsudcích v češtině nebo v jiných jazycích? Samozřejmě, že pokud bychom jednoznačně věděli, že se sémantika nějaké spojky takového jazyka shoduje se sémantikou implikace KPL, mohli bychom zaručenou platnost (MP) přímo promítnout i na takový jazyk. Problém je v tom, jak tohle můžeme vědět. (Speciálně víme, že v běžných jazycích se spojka, která by se co do své sémantiky zcela shodovala s \rightarrow , jak je tato spojka vymezená v rámci KVL a KPL, zpravidla nevyskytuje.)

Podle čeho dokážeme posoudit, do jaké míry se nějaké spojení přirozeného jazyka sémanticky shoduje s nějakou spojkou definovanou v rámci logiky? Zdá se, že tu jsou dvě možnosti: nějak testovat shodu pravdivostních podmínek a nějak testovat shodu v chování v rámci úsudků (budeme též hovořit o *inferenční roli*). Jak však snadno nahlédneme, testovat to první v podstatě nemůžeme jinak než tak, že budeme testovat to druhé: abychom zkoumali, zda je nějaká věta pravdivá za takových a takových okolností, musíme tyto okolnosti nějak jazykem charakterizovat, a zkoumání pravdivosti věty za těchto okolností pak odpovídá zkoumání toho, zda je tato věta odvoditelná z oné charakterizace. To znamená, že fakticky se v podstatě musíme spolehnout na test chování v úsudcích. Tento test spočívá, zhruba řečeno, v tom, že zkoumáme, jak bude správnost úsudků ovlivněna tím, když do nich za nějakou spojku přirozeného jazyka dosadíme nějaký její protipól z logického jazyka.

To je ale opravdu jenom velmi zhruba řečeno. Ve skutečnosti samozřejmě nemůžeme v nějaké větě prostě nahradit třeba spojení *jestliže ... pak ...* spojkou \rightarrow , protože to, co takovou záměnou vznikne, bude přísně vzato jakýsi hybrid mezi větou přirozeného jazyka a výrokem jazyka logického kalkulu a jako takový nebude mít jasně daný význam, a tudíž ani pravdivostní podmínky. ‚Nepřísně vzato‘ si však jistě dovedeme představit, jaké pravdivostní podmínky můžeme takové větě přiřadit (podrobněji o tom níže).

Hlubším problémem se ale zdá být to, že takové zkoumání shody mezi inferenčními rolemi výrazů přirozeného jazyka a logických konstant předpokládá, že víme, které úsudky v přirozeném jazyce jsou správné (a to nezávisle na platnosti jakýchkoli úsudků v rámci logických kalkulů). Zdá se přirozené předpokládat, že přirozený jazyk, jako je čeština, má nějakou víceméně pevnou sémantiku (samozřejmě plus mínus variace dané odlišnostmi v jazykovém chování individuálních mluvčích nebo skupin mluvčích) a to také znamená, že nějaké úsudky v tomto jazyce jsou jednoznačně správné a některé jednoznačně nesprávné. Jak ale víme, které jsou ty správné? Inu, provizorně nám k tomu mohou stačit intuice nás samých, jsme-li rodilými mluvčími, ale poslední instancí by zřejmě musel být reprezentativní empirický výzkum mezi mluvčími češtiny.

Znamená to ale, že se logika stává *empirickou záležitostí*? V jistém smyslu se dá říct, že to tak je – rozhodně to ale není v onom smyslu, proti kterému bojovali otcové moderní logiky v čele s Fregem. Empirickou záležitostí rozhodně je, jaký význam je spjat s jednotlivými vyjádřeními přirozeného jazyka, a tedy i to jaká je inferenční role jednotlivých výrazů⁵.

⁵ Pavel Materna opakovaně argumentuje, že logická analýza přirozeného jazyka *není* empirická disciplína (viz například Materna 2005). Myslí to tak, že to, co je empirické, nepatří do *logické* analýzy (ale do lingvistiky). To, co bylo řečeno zde, se ale zdá naznačovat, že při takovém dělení by toho na logickou analýzu mnoho nezbylo. Zjištění, co který výraz přirozeného jazyka fakticky znamená, je nade vši pochybnost zjištění vyžadující empirické zkoumání (i když ti, kdo už mají za sebou to ‚empirické zkoumání‘, které stojí v základě prvotního osvojování si jazyka, snadno nabudou pocit, na tyto otázky dovedou odpovědět *a priori*) a protože inferenční role výrazu je věcí toho, co tento výraz znamená, platí totéž i pro tuto inferenční roli. Avšak stanovení inferenční role výrazu je stanovení toho, které úsudky obsahující tento výraz jsou správné a které ne – takže ani tohle by podle Maternova dělení nepatřilo do logiky. Logická analýza přirozeného jazyka by tak byla

Můžeme tedy například empiricky zjistit, že v nějakém jazyce neplatí *modus ponens*? Stěží; můžeme jistě zjistit, že nějaký úsudek, který jsme považovali za *modus ponens* ve skutečnosti není správný – závěrem ale pak nebude, že neplatí *modus ponens*, ale to, že něco, co se nám v přirozeném jazyce jevílo jako implikace, implikací ve skutečnosti není (kde implikací teď rozumíme jakoukoli spojku, pro kterou platí *modus ponens*)⁶. V tomto smyslu tedy logika *není* empirická a je transcendentní. Logikou ji ale dělá jediné to, že nějakým podstatným způsobem postihuje úsudky, které jsou fakticky (a tudíž v našich přirozených jazycích) správné⁷.

Měli bychom tedy logiku pojímat jako čistě popisnou disciplínu? Musíme rezignovat na to, že by nás byla schopna v našich úsudcích nějak korigovat? To ani zdaleka ne. Za prvé, to, co logika artikuluje, jsou *pravidla*. (Pravidla, která platí fakticky, pro nějaké jazykové společenství, a jejichž platnost je tak empirickým faktem; ale přes to všechno pravidla.) A za druhé, logika se neomezuje na to, že by jen vynášela na světlo hotová pravidla, ale svojí artikulací těchto pravidel je pomáhá dotvářet a fixovat; a navíc je pak nabízí k použití v nové, explicitní podobě: z mnohosti jazykových úzů napříč mluvčími a napříč jednotlivými úsudky extrahuje jakousi relativně jednoduchou normativní strukturu. Co se nám do ní nevejde, to prohlásíme za chyby, omyly či nepřesnosti. (Protože ale chceme stále tvrdit, že je tato struktura věcí toho, jak funguje faktický jazyk, dává tohle smysl jediné tehdy, když je těchto chyb, omylů ap. omezený počet – jinak by bylo rozumnější přehodnotit celou tu strukturu.)

Obrázek logiky, který nám z tohoto vychází, je tedy ten, že logická pravidla extrahujeme z empirických faktů týkajících se toho, jakými pravidly se náš jazyk řídí, že je nicméně extrahujeme způsobem, při kterém z vágních, nejasných a kolísavých pravidel implicitních užívání

omezena na zkoumání, které úsudky platí v nějakých umělých jazycích – a o tom, co platí v jakémkoli *přirozeném jazyce*, by nám nedokázala říci vůbec nic. Avšak závěr, že *logická analýza přirozeného jazyka* nám nedokáže říct nic o žádném přirozeném jazyce, nám připadá dokonale absurdní.

⁶ Viz o tom podrobněji Peregrin (2010).

⁷ Empiricky bychom tedy mohli zjistit, že v nějakém jazyce vůbec není nic jako implikace. Je však otázka, jestli bychom pak v souvislosti s pravidly takového jazyka vůbec hovořili o logice. A v důsledku vyvstává otázka, zda bychom v souvislosti se soustavou zvuků či inskripcí řídících se takovými pravidly vůbec hovořili o *jazyce*.

jazyka učiníme pravidla jasná, jednoznačná a explicitní, pravidla, která pak lze použít i k tomu, abychom do jisté míry hodnotili a opravovali to, co činí jednotliví mluvčí. Takto chápaná logika tedy odvozuje svou legitimitu z faktických struktur toho kterého jazyka, získává však vzhledem k těmto faktům určitou míru autonomie a s tím i určitý normativní rozměr. To je to, čemu budeme níže říkat *reflektivní ekvilibrium*: v jednom smyslu se platnost logického zákona odvozuje od toho, že jsou fakticky správné jeho instance, ale tyto instance jsou v jiném smyslu správné proto, že to stanoví tento zákon.

3 Co děláme, když formalizujeme úsudky?

Otázky, které musíme podle našeho názoru zodpovědět, abychom objasnili podstatu formalizace úsudků a logické formalizace přirozeného jazyka, a abychom pochopili možnosti a meze tohoto podniku, jsou nyní podle nás následující

- 1) Co vlastně děláme, když provádíme formalizaci věty přirozeného jazyka?
- 2) Jaká jsou (měla by být) kritéria, jež umožňují (měla by umožňovat) stanovit, která z alternativních formalizací věty přirozeného jazyka je správná?
- 3) Co dává logice autoritu rozhodovat o tom, které úsudky jsou (resp. nejsou) správné? Existuje vůbec něco takového?

Jak už jsme viděli, běžnou odpovědí na otázku 1) je to, že formalizace je něco jako překlad či ‚transverbase‘ (nahrazování znaků jinými znaky). V tomto bodě je ale třeba vzít v patrnost, že formalizací se běžně rozumí dvě poněkud různé věci. Podle prvního pojetí je výsledkem formalizace věty přirozeného jazyka nějaká smysluplná (pravdivá či nepravdivá) ‚věta‘ nějakého formalizovaného jazyka (skládající se výhradně ze symbolů s pevným významem). Podle druhého pojetí je jejím výsledkem *formule* formálního jazyka obsahující neinterpretované symboly, a tudíž sama o sobě nic neříkající, jenom zachycující *logickou formu* původní věty.

Vezmeme-li například větu

(V2) *Všichni koně mají hlavu,*

pak výsledkem její formalizace bude v prvním případě něco jako

(FV2) $\forall x (K(x) \rightarrow H(x)),$

kde K a H budou výrazy s významem víceméně odpovídajícím jejich českým protipólům *je kůň* a *má hlavu*, zatímco v tom druhém to bude

$$(FV2') \quad \forall x (F(x) \rightarrow G(x)),$$

kde F a G jsou bezobsažné symboly.

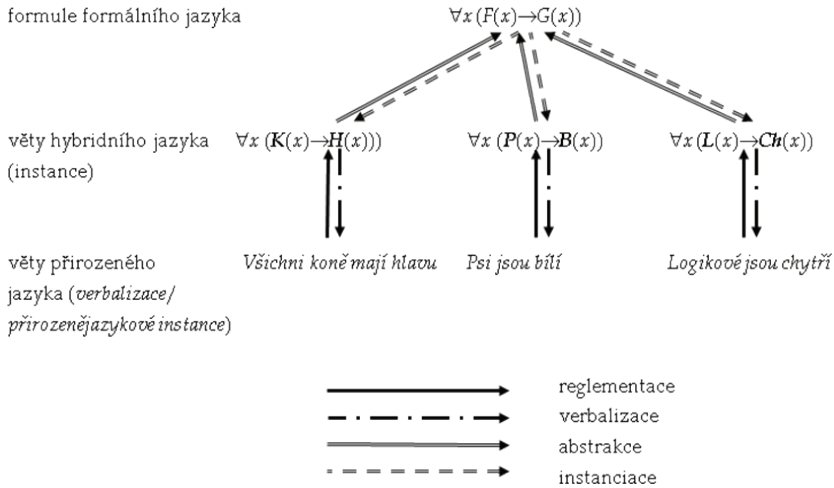
Zavedme terminologické rozlišení, podle kterého budeme o výroku, jako je (FV2), hovořit jako o *logické struktuře* věty (V2), a o formuli, jako je (FV2'), jako o její *logické formě*⁸. V prvním z těchto dvou případů by bylo možné uvažovat o tom, že formalizace je něco jako překlad; v tom druhém to rozhodně tak není.

Uvažujme ještě o té první možnosti: v jakém vztahu mohou být výrazy K a H , které jsou součástí (FV2), k českým výrazům *je kůň* a *má hlavu*? Jedna z možností je, že to jsou přímo zkratky za ony výrazy přirozeného jazyka – a že (FV2) je tedy takovým ‚hybridním‘ výrokem, poskládaným ze slov dvou různých jazyků, o jakých jsme uvažovali již výše. (V takovém případě samozřejmě musíme vyjasnit, jaký mají takové hybridní výroky význam, což ale jistě není nepřekonatelný problém: můžeme například prohlásit, že (FV2) je zkráceným zápisem českého výroku: *Pro každý objekt platí, že je-li koněm, má hlavu.*) Jinou možností je to, že K a H jsou součástí jazyka, do něž překládáme – pak to ovšem není čistě logický jazyk, tyto výrazy, protože mají pevný význam, jsou jeho *extralogickými konstantami*. Prostředky tohoto jazyka pak samozřejmě musíme zajistit, aby významy těchto konstant v nějakém smyslu odpovídaly významům příslušných českých výrazů a zároveň vhodně ‚zapadly‘ do sémantiky daného logického jazyka.

My se domníváme, že termín *formalizace* je rozumnější rezervovat pro onen druhý postup – totiž pro postup od věty přirozeného jazyka k formuli logického jazyka udávající logickou formu této věty. Tento postup si ovšem můžeme představit jako rozdělený do dvou kroků, z nichž první je krokem od věty přirozeného jazyka k její logické struktuře vyjádřené v ‚hybridním‘ jazyce, a ten druhý spočívá v tom, že se extralogické výrazy přirozeného jazyka přítomné ve výsledném ‚hybridním‘ výroku nahradí parametry. Tomu prvnímu kroku pak říkáme *reglementace*, zatímco tomu druhému *abstrakce*.

Můžeme si to znázornit na následujícím obrázku:

⁸ Tuto terminologii jsme zavedli a podrobněji rozebrali v Svoboda – Peregrin (2009).



4 Kritéria formalizace

Podívejme se nyní na otázku, kterou jsme výše označili číslicí 2: *Jaká jsou kritéria, jež umožňují stanovit, která z alternativních formalizací věty přirozeného jazyka je správná?* Uvažme znovu větu (V2) a tři alternativní návrhy na její formalizaci:

$$\begin{aligned} \text{(FV2')} \quad & \forall x (F(x) \rightarrow G(x)) \\ \text{(FV2'')} \quad & \forall x \forall y (F(x) \rightarrow G(y)) \\ \text{(FV2''')} \quad & \forall x (F(x) \wedge G(x)) \end{aligned}$$

(Ponechme stranou fakt, že se druhý i třetí z těchto návrhů mohou tomu, kdo má jistou praxi ve formalizaci v jazyce KPL, jevit poněkud absurdně.) Jak rozhodneme, který z těchto návrhů vybrat?

Vycházíme z toho, že některé úsudky, které obsahují větu (V2), jsou správné a jiné nesprávné. Konkrétně lze říci, že zatímco např. úsudek

$$\begin{array}{l} \text{(U4)} \quad \textit{Všichni koně mají hlavu} \\ \quad \quad \textit{Peruán má hlavu} \\ \hline \quad \quad \textit{Peruán je kůň} \end{array}$$

správný není, úsudek

$$\begin{array}{l} \text{(U5)} \quad \textit{Všichni koně mají hlavu} \\ \quad \quad \textit{Peruán je kůň} \\ \hline \quad \quad \textit{Peruán má hlavu} \end{array}$$

správný je. Jak budou tyto úsudky vypadat, když do nich dosadíme jednotlivé návrhy na formalizaci (V2)?

Všichni koně mají hlavu	$\forall x (F(x) \rightarrow G(x))$	$\forall x \forall y (F(x) \rightarrow G(y))$	$\forall x (F(x) \wedge G(x))$
Všichni koně mají hlavu <u>Peruán má hlavu</u> Peruán je kůň	$\frac{\forall x (F(x) \rightarrow G(x))}{G(a)}$ $F(a)$	$\frac{\forall x \forall y (F(x) \rightarrow G(y))}{G(a)}$ $F(a)$	$\frac{\forall x (F(x) \wedge G(x))}{G(a)}$ $F(a)$
Všichni koně mají hlavu <u>Peruán je kůň</u> Peruán má hlavu	$\frac{\forall x (F(x) \rightarrow G(x))}{F(a)}$ $G(a)$	$\frac{\forall x \forall y (F(x) \rightarrow G(y))}{F(a)}$ $G(a)$	$\frac{\forall x (F(x) \wedge G(x))}{F(a)}$ $G(a)$

Je zřejmé, že pokud jde o řádek, který odpovídá nesprávnému úsudku, dávají nám formalizace v prvním a druhém sloupci neplatné úsudkové formy, zatímco ta ve třetím sloupci vede k úsudkové formě platné. To nám dává důvod třetí z uvedených návrhů odmítnout.

Podíváme-li se pak na další řádek, vidíme, že tomuto intuitivně správnému úsudku odpovídá, v případě prvního sloupce platná, zatímco v případě druhého sloupce neplatná úsudková forma. To nám dává důvod odmítnout druhou z navržených formalizací, a z daných tří návrhů tedy upřednostnit (nepřekvapivě) ten první.

Pokusíme-li se formulovat jádro postupu, který jsme právě předvedli, ve formě principů, dostáváme následující dva:

(Princip spolehlivosti - PS)

Formule Φ je kandidátem na adekvátní formalizaci věty V v logickém systému S jen když platí: Pokud je úsudková forma, v níž se Φ objevuje jako premisa či závěr, platná v S , pak (téměř) všechny přirozenějazykové instance této úsudkové formy, v nichž se V vyskytuje jako instance Φ , jsou správné úsudky.

(Princip ambicióznosti - PA)

Formule Φ je tím lepším kandidátem na adekvátní formalizaci věty V v logickém systému S , čím více relevantních intuitivně správných úsudků⁹,

⁹ Předpokládáme, že každý z běžných logických systémů se zaměřuje na nějakou omezenou zamýšlenou oblast aplikace. Pohybujeme-li se například v rámci KVL, nebudeme jistě považovat za relevantní úsudky, jejichž správnost je věcí kvantifikace či nějakých modálních modifikátorů – za relevantní budeme brát jen ty, jejichž správnost je dána čistě výrazivem,

v nichž se V vyskytuje jako premisa nebo závěr, je instancemi (verbalizacemi) platných úsudkových forem S, v nichž Φ vystupuje jako formalizace V.

Tyto dva principy, zhruba řečeno, artikulují požadavek, že inferenční role formule, která představuje formalizaci přirozenějazykové věty (v rámci příslušného formálního jazyka), má co možná nejlépe aproximovat inferenční roli této věty v rámci daného přirozeného jazyka. *Princip spolehlivosti* přitom zhruba říká, že případ, kdy je verbalizace platné úsudkové formy (při dané formalizaci) intuitivně nesprávným úsudkem, příslušnou formalizaci diskvalifikuje; *princip ambicióznosti* naopak říká, že čím více relevantních intuitivně správných úsudků lze s využitím dané formalizace ‚odhalit‘ jako úsudky logicky správné (v daném systému S), tím je daná formalizace lepší.

S těmito principy jsme vystačili v případě našeho jednoduchého příkladu, obecně s nimi ale asi nevystačíme. Nepomohou nám například rozhodnout, zda máme za adekvátní formalizaci (V2) prohlásit už diskutovanou formuli $\forall x(F(x) \rightarrow G(x))$ (která vyšla vítězně z našeho mini-konkurzu) či $\forall x(F(x) \rightarrow G(x)) \wedge \forall x(F(x) \rightarrow G(x))$ (která v mini-konkurzu tohoto typu nutně obstojí stejně dobře, protože inferenční role obou těchto formulí jsou zjevně totožné.) Z tohoto důvodu formulujeme ještě dodatečná kritéria, která, jak se domníváme, používáme v procesu výběru „té pravé“ formule v případě, že ta předchodí nerozhodnou:

(Princip transparentnosti – PT)

Pokud máme dvě formalizace Φ_1 a Φ_2 věty V v jazyce logického systému S, které stejně dobře vyhovují kritériím (PS) a (PA), pak je adekvátnější ta, jejíž gramatická struktura je bližší gramatické struktuře věty V.

(Princip úspornosti – PU)

Pokud máme dvě formalizace Φ_1 a Φ_2 věty V v jazyce logického systému S, které stejně dobře vyhovují kritériím (PS) a (PA), pak je adekvátnější ta, která je úspornější co do výskytu symbolů jazyka logického systému S.¹⁰

které si KVL klade za cíl postihnout („a“, „nebo“, „jestliže ..., pak ...“ atd.). Některé správné úsudky, např. *Eva je slečna* tedy *Eva není vdaná*, pak asi nebudeme brát za relevantní z hlediska žádného logického systému – jejich správnost zjevně není záležitostí pouze takového výraziva, které bychom byli ochotni považovat za logické a které bychom měli ambice rekonstruovat pomocí logických systémů.

¹⁰ Podrobnější rozbor kritérií formalizace i kritiku alternativního přístupu lze nalézt v Peregrin – Svoboda (vyjde).

5 Reflektivní ekvilibrium

Když se ovšem nyní nad postupem formalizace, jak jsme ho zrekapitulovali, a nad kritérii, jak jsme je formulovali, zamyslíme, vyvstane před námi něco, co vypadá jako nezanedbatelný problém. Způsob testování, tak jak jsme ho popsali, i kritéria, která jsme z něj extrahovali, vycházejí z předpokladu, že chceme-li formalizovat nějakou větu, máme již formalizováno její ‚inferenční okolí‘, to jest, že můžeme pracovat s již danými formalizacemi spousty jiných vět, spolu se kterými tvoří daná věta úsudky. Když jsme testovali, která z navržených formalizací věty (V2) je nejlepší, tiše jsme předpokládali, že jsou již pevně dány formalizace vět *Peruán je kůň* a *Peruán má hlavu*. Něco takového ale samozřejmě obecně předpokládat nemůžeme, neboť bychom se tak dostali do bludného kruhu: k tomu, abychom mohli formalizovat jakoukoli větu, bychom již potřebovali mít hotovou formalizaci spousty dalších vět.

Z toho plyne, že proces určování logické formy se nemůže týkat jenom jednotlivé věty (to jest nemůže být ‚atomistický‘), ale vždy vyvolává otázky po logické formě mnoha dalších vět (tento proces musí být v principu ‚holistický‘). V praxi tomu je tak, že vyvstane-li otázka posouzení adekvátnosti formalizace jisté věty nebo jistého argumentu, postupujeme tak, že odhadneme logické formy některých vět (typicky těch nejjednodušších) a od tohoto ‚zkusného‘ přiřazení forem se propracováváme k formám dalších vět, přičemž stále testujeme, jak si výsledné formalizace vedou z hlediska výše uvedených kritérií, a případně zpětně korigujeme ty původní formalizace. Žádné formalizace nepovažujeme v rámci tohoto procesu za principiálně nerevidovatelné (i když někdy si třeba revizi dovedeme představit jen obtížně). Dokonce ani užitý logický aparát obecně nebereme jako pevně daný a nerevidovatelný.¹¹

Ilustrujme to na dvou jednoduchých příkladech. Řekněme, že před námi stojí úkol formalizovat větu

(V3) *Všichni Slováci mají společné geny,*

¹¹ Zejména v rámci logických systémů, které nejsou obdobně ‚usazené‘ jako systémy klasické logiky, vede nalezení paradoxů (neintuitivních důsledků) často k revizi daných systémů. Příklady takového postupu najdeme například v rámci dějin budování deontické logiky (viz například Hilpinen 1971).

a my ji postupem analogickým formalizaci (V2) zkusmo formalizujeme jako

$$(FV3) \quad \forall x (F(x) \rightarrow G(x))$$

Následně budeme adekvátnost této formalizace testovat pomocí úsudku:

$$(U6) \quad \frac{\begin{array}{l} \text{Všichni Slováci mají společné geny} \\ \text{Ján Slota je Slovák} \end{array}}{\text{Ján Slota má společné geny}}$$

Za předpokladu, že i zbývající dvě věty formalizujeme analogicky jako v případě úsudku (U5), dojdeme přímočaře k závěru, že naše zkusmá formalizace zřejmě není adekvátní. Zatímco úsudková forma

$$(FU6) \quad \frac{\forall x (F(x) \rightarrow G(x))}{\frac{F(a)}{G(a)}}$$

je v rámci KPL jasně platná, úsudek (U6) je zjevně nesprávný – je dokonce otázka, zda jeho závěr dává nějaký rozumný smysl. Ukázalo se tedy, že naše zkusmá formalizace je patrně v rozporu s principem **PS** a máme důvod ji zamítnout. V principu bychom ovšem mohli postupovat i jinak – mohli bychom trvat na tom, že (V2) a (V3) *mají* shodnou logickou formu a něco je v nepořádku s formalizací zbývajících vět v úsudku, nebo dokonce s logikou, která nám ‚neumožňuje‘ formalizovat dané – gramaticky velmi podobné – věty analogicky. Jako ‚přítěžující okolnost‘ bychom danému logickému systému, tj. KPL, mohli přičíst i to, že patrně neumožňuje formalizovat úsudek

$$(U7) \quad \frac{\begin{array}{l} \text{Všichni Slováci mají společné geny} \\ \text{Gen X je jedním z oněch společných genů} \end{array}}{\text{Všichni Slováci mají gen X}}$$

tak, aby vyšel jako logicky správný. Pokud totiž v rámci KPL nedokážeme tento úsudek takto adekvátně formalizovat, bude mít jakákoli formalizace v tomto systému mezery v naplňování kritéria **PA**.

Vezměme ještě jiný, jednodušší příklad. Řekněme, že máme za úkol formalizovat větu

$$(V4) \quad \text{Někteří Slováci jsou eurokomisaři.}$$

Zdá se, že zjevným kandidátem na adekvátní formalizaci bude formule

(FV4) $\exists x(F(x) \wedge G(x))$

Předložíme-li ovšem běžnému mluvčímu češtiny úsudek

(U8) *Maroš Šefčovič je eurokomisařem*
Maroš Šefčovič je Slovák
Někteří Slováci jsou eurokomisaři

bude, domníváme se, s velkou pravděpodobností minimálně pochybovat o tom, že je správný, i když jeho forma bude (při dané formalizaci (V4) a neproblematických formalizacích zbývajících dvou vět) zjevně platná. Z toho, že určitý člověk je Slovák a zároveň eurokomisař nplyne, že tuto funkci zastávají i další Slováci, což se věta v závěru úsudku zdá při svém běžném (intuitivním) čtení říkat.¹² Naznačená formalizace tak, zdá se, nesplňuje kritérium **PS** a měli bychom ji zřejmě odmítnout. To samozřejmě můžeme učinit a pokusit se o formalizaci, která tento problém odstraní, ale můžeme také říci, že rozdíl mezi významem vět *Někteří Slováci jsou eurokomisaři* a *Nějaký Slovák je eurokomisařem* je pro nás zanedbatelný, a můžeme tak trvat na tom, že je navržená formalizace adekvátní.

Z těchto příkladů je patrné, že v rámci logické analýzy do jisté míry svévolně stanovujeme, která věta je instancí dané formy, a která nikoli. Ale můžeme-li *jakýkoli* protipříklad k uplatňování logických pravidel, která navrhuje, interpretovat jako přijatelnou výjimku nebo jako problém, který lze řešit úpravami jinde v ‚systému‘, nemůže to vést k tomu, že všechna svá logická pravidla stanovíme nakonec zcela svévolně? Jistě ne. Veškerá legitimita pravidel, která v logice postulujeme, se totiž odvozuje od správnosti jejich instancí v přirozeném jazyce – můžeme tedy v zájmu jednoduchosti výsledných pravidel sem tam něco odsunout stranou jako výjimku, nemůžeme to ale dělat příliš často nebo příliš násilným způsobem.

Dostáváme se tak k obrázku, podle kterého je na jedné straně platnost obecných logických zákonů odvozena od faktické správnosti jejich instancí, zatímco na druhé straně je to, že je něco správná instance, stanoveno prostřednictvím logických zákonů. To samozřejmě vypadá trochu jako bludný kruh, ve skutečnosti je to ale podle nás nevyhnutelný osud logiky, se kterým se musí vypořádat tak, že bude směřovat k tomu, co Nelson Goodman nazval *reflektivním ekvilibriem*:

¹² Samozřejmě že pokud by v závěru daného úsudku byla věta *Nějaký Slovák je eurokomisařem*, neměl by se správností úsudku nikdo problém.

Deduktivní odvození jsou opodstatněna tím, že souhlasí v platných obecnými pravidly a [...] obecná pravidla jsou opodstatněna tím, že souhlasí s platnými odvozeními. Avšak tento kruh není bludný. Jde o to, že jak pravidla, tak jednotlivé inference jsou opodstatněny tím, že jsou uvedeny do vzájemného souladu. *Pravidlo upravíme, pokud z něj vyplyne odvození, které nejsme ochotni přijmout; odvození odmítneme, pokud porušuje pravidlo, které nejsme ochotni upravit.* (Goodman 1955, 67)

Tím se dostáváme k naší otázce 3): *Co dává logice autoritu rozhodovat o tom, které úsudky jsou, resp. nejsou správné?* Odpověď, ke které jsme takto dospěli, je ta, že úkolem logických formalismů je zachytit inferenční vlastnosti vět přirozeného jazyka, i když ne prostým popisem toho, co se ‚děje‘ v přirozeném jazyce, ale s využitím jisté míry zobecnování a systematizace. Proto je využití těchto formalismů vystaveno kritice obdobně, jako lze kritizovat jiné empirické generalizace. Avšak protože přirozené jazyky jsou neurčité, vágní a poskytují prostor k různým výkladům významu svých vyjádření, představuje formalizace způsob, jak pro jisté specifické účely učinit vyjadřování jednoznačnější a určitější, a imunizovat jej (alespoň do jisté míry) proti nekonzistentnosti. Z toho plyne, že se logika nemůže vyvíjet bez ohledu na empirická fakta týkající se faktických jazyků, současně však má i normativní rozměr, který jí dovoluje tyto faktické jazyky do jisté míry kritizovat. Jak formalizace, tak budování logických systémů bude ovšem vždy procesem hledání Goodmanovského reflektivního ekvilibria.

Oddělení logiky
Filosofický ústav
Akademie věd České republiky
Jilská 1
110 00 Praha 1
Česká republika
jaroslav.peregrin@uhk.cz
svoboda@site.cas.cz

Literatura

- BAUMGARTNER, M. – LAMPERT, T. (2008): Adequate formalization. *Synthese* 164, 93-115.
- BRUN, G. (2004): *Die richtige Formel*. Frankfurt: Ontos.
- GOODMAN, N. (1955): *Fact, Fiction, and Forecast*. Cambridge (Mass.): Harvard University Press.

- HILPINEN, R. (ed.) (1971): *Deontic Logic – Introductory and Systematic Readings*. Dordrecht: Reidel.
- MATERNA, P. (2005): Je „logická analýza přirozeného jazyka“ možná? *Filosofický časopis* 53, 267-276.
- PEREGRIN, J. (2010): Logic and Natural Selection. *Logica Universalis* 4, 207-223.
- PEREGRIN, J. – SVOBODA, V. (vyjde): Criteria for logical formalization. *Synthèse*.
- SAINSBURY, R. M. (1991): *Logical Forms. An Introduction to Philosophical Logic*. Oxford: Blackwell.
- SVOBODA, V. – PEREGRIN, J. (2009): *Od jazyka k logice*. Praha: Academia.

Emoce: primitivní a kognitivní složka

Michal Polák

Západočeská Univerzita v Plzni

Abstract: The paper deals with the problem of emotions. It begins with an analysis of emergence of the emotions according to three types of theories of the emotional process (cognitive, non-cognitive and somatic feedback theories) and shows that any of them has its own difficulties that are not acceptably solved within these theories. Therefore, the paper tries to propose a difference between a raw and cognitive part of an emotion. This difference should be a useful tool in our understanding of emotions.

Keywords: emotions, cognition, somatic feedback theory, two elements of an emotion.

Pocitová (afektivní) složka lidské zkušenosti tvoří důležitou část našeho zakoušení světa i nás samotných jako biologických nebo oduševněných bytostí. Pocitová složka lidské zkušenosti se projevuje v různých modech – např. v emocích, náladě, temperamentu nebo v tělesných počtících, jako je bolest. Existují různé teorie, jak vysvětlit tento aspekt lidské zkušenosti a jak jej objektivně zachytit. Ve stati se omezím na interní kontext vzniku subjektivní zkušenosti, přičemž opomímím neméně významný evoluční a sociální kontext. Interní přístup se pokouší porozumět vzniku subjektivní zkušenosti, ať už jde o zkušenost emocionální nebo percepční (např. zakoušení tělesné bolesti). Nejprve stručně představím základní teorie vzniku emocí – teorie kognitivní a nekognitivní, teorii somatické zpětné vazby – a ukáži některé jejich nevýhody. Poté představím vlastní návrh vzniku emocí, který se pokouší schematicky pokrýt jak proceduru vzniku základních emocí, jakými jsou strach, hněv, znechucení, překvapení, radost a smutek, tak i specifických tělesných prožitků, jakými jsou bolest, orgasmus, pocit hladu apod. Toto schéma je založené na rozlišování mezi primitivní

a kognitivní komponentou emocí. Cílem je zdůraznit a argumentačně podpořit tezi, že subjektivní prožívání emocí nelze doslova redukovat na kognitivní (čímž rozumím jazykovou a z ní odvozenou pojmovou) složku ani na složku somatickou (neurobiologickou), i když platí, že obě složky se mohou při konstituci emocí uplatňovat.

1 Kognitivní teorie emocí

Kognitivní teorie vzniku emočního procesu jsou sice v detailech různorodé, ale sdílejí jeden významný předpoklad. Emoce se utvářejí na základě aktivní manipulace s informacemi a to zejména v raném stádiu vzniku emočního procesu (tj. v době mezi percepcí vnějšího podnětu a spuštěním emoční tělesné reakce). Protože se jedná o proces, v němž hrají důležitou roli získané informace, hovoří se o kognitivním procesu. Na rozdíl od nekognitivních přístupů vzniku emočních stavů, kde se zdůrazňuje automatická odpověď na vnímaný stimul bez zapojení vyšších kognitivních procesů, je kognitivní přístup opřen o nutnost nejprve konceptualizovat lidskou zkušenost. To znamená, že nejprve dojde k ohodnocení percepční informace a až poté vznikne výsledná emoce. Emoce je tak utvořena s pomocí kontrastu mezi perceptuální informací (skutečně přijímanou smysly) a člověkem očekávanou informací. Různá míra napětí mezi těmito dvěma rovinami vede k vytvoření různých emocí.

První výhoda kognitivní pozice je v tom, že když různá individua odpovídají na tutéž událost různě, lze tento rozdíl v odpovědích spatřovat v rozdílu jejich kognitivního vybavení a ve způsobu, jak událost hodnotí (mají individuální preference a postoje). Například když fotbalový hráč obdrží nové angažmá, může cítit uspokojení (např. z vyššího platu), kdežto jeho schopný spoluhráč může tu samou situaci pociťovat se znepokojením (odešel mu zdatný partner). To samé platí i pro jednoho člověka, který dnes hodnotí rozchod s partnerem jako smutnou událost, ale kdyby se rozešel za půl roku a mezitím zjistil, že ho stávající partner podvádí, tak možná bude tu samou událost rozchodu pociťovat jako úlevu. Emoční reakce na týž podnět se tedy mohou měnit, takže individuální a časovou variabilitu prožitku lze stěžejně vysvětlit bez zapojení vyšších kognitivních procesů (Roseman – Smith 2001, 4).

Druhá výhoda kognitivní pozice spočívá v tom, že existuje řada událostí, které nejsou po fyzické stránce nijak podobné, ale přesto vy-

volávají stejné emoční reakce.¹ Například smutek lze vyvolat smrtí rodiče, narozením dítěte, rozvodem, nepřijetím na školu apod. Přestože tyto fyzické události nemají společného fyzického jmenovatele, takže emoci nelze vysvětlit perceptuální zkušeností, evoluční dispozicí reagovat na daný stimul nebo učením prostřednictvím generalizace, mají společného jmenovatele právě v emoci smutku. Schopnost přiřadit smutek k tak rozdílným událostem tedy musí být způsobena nějakými vnitřními kognitivními procesy (Roseman – Smith 2001, 4). Motivace, tendence k chování, cíle, přesvědčení, přání a úmysly jsou tím, co máme k dispozici ještě před registrací podnětu a co zformuje příslušný podnět do podoby určité emoce. V takto konstituované kognitivní teorii však vzniká otázka, jak vysvětlit emoční vyjádření tříměsíčního dítěte, které křičí, protože má pocit hladu, ale kognitivně (jazykově, resp. pojmově) by tuto situaci nedokázalo uchopit.

Třetí přednost kognitivní teorie je patrná v případě kognitivní teorie soudů. Jde o odnož kognitivní teorie, která vidí konstituci emoce v lidských soudech o světě. Souzení se zde chápe jako určitá mentální schopnost, která se uplatňuje při reflexi nás samotných i při reflexi vnějšího světa a jeho podmínek. Naše hodnoty, ideje a názory projikujeme do světa tím, že o něm něco soudíme (Solomon 2007, 204). Tyto soudy pak konstituují různé emoce. Například můj strach je utvořen ze soudů, že jsem napaden, že jsem fyzicky slabší než můj protivník, že neovládám příslušné obranné strategie apod. Marta Nussbaumová zdůrazňuje aktivní schopnost subjektu vyjadřovat o světě a situacích v něm obsažených určité soudy či přesvědčení, takže vznik emoce není ovlivňován pouze percepčním stimulem. Není to něco, co by se subjektu prostě čas od času stávalo pod vlivem vnějších okolností. Podle ní je důležitý typ souboru, který obsahuje příslušná přesvědčení, protože právě ten se vztahuje ke konkrétní emoci. Pokud se v takovém souboru změní nějaké přesvědčení (nebo nějaké vypadne), změní se i emoce, jež je důsledkem těchto přesvědčení, resp. soudů. Zároveň však platí, že ne všechny soudy konstituují nutně nějaké emoce. Například soud, že obloha je modrá, nemusí vyvolávat žádnou emoci.

¹ Problém „stejnosti“ emoční reakce je jeden z aspektů, na něž se zaměřuje tato stať. Stejnost či různost emoční reakce lze hodnotit na úrovni kognitivní (jazykově zformované) emoce a nikoli na úrovni primitivní emoce. Viz níže.

Jiný aspekt teorie soudů je, že nepředpokládá nutnost vědomých nebo svobodných soudů, které osoba musí vyjádřit o světě, aby vytvořila příslušnou emoci. Úplně postačuje, že tak činí jaksi spontánně a ne zcela záměrně a s rozvahou (Solomon 1977, 46). Například to, že mě někdo polil pivem, vyjádřím rychle (nevědomě) soustavou soudů, které způsobí příslušnou emoci zlosti. V tomto případě lze považovat za součást vlivů způsobujících emoci zlosti také pocit, který mám při dotyku mokrého oblečení na mém těle. Přitom ale ne vždy platí, že na emoci musí mít vliv tělesná reakce (Nussbaum 2004, 195). Zde se nabízí jedna námitka a jedna poznámka.

Především se nezdá, že bychom nemohli pociťovat strach nebo radost, aniž bychom měli nějaký takto komplexní kognitivní soud. Prostě pozorování novorozenců a kojenců naznačuje, že jejich radost nebo strach je skutečný, ačkoli se nedomníváme, že jejich kognice by byla na úrovni umožňující tvorbu složitých soudů. Je-li tomu tak, pak aspoň některé emoce musí vznikat nezávisle na kognitivním ohodnocení, anebo přinejmenším nezávisle na ohodnocení, které předpokládá teorie soudů. Neméně významná je skutečnost, že mnoho teorií emocí, a platí to i pro teorii soudů, pracuje s konceptem nevědomých kognitivních procesů. Zavedení nevědomých kognitivních procesů je často motivováno jednak empirickými zjištěními v kognitivní vědě, jednak explanačními důvody. Koncept nevědomých procesů se totiž lépe reaguje na námitku nekognitivních teorií, že emoce někdy vznikají tak rychle, že je těžké si představit nějakou ohodnocovací proceduru, která vzniku emoce předchází. Když však připustíme nevědomé ohodnocování, otevíráme tím složité otázky typu, co je substrátem nevědomých kognitivních procesů. Například, zda jsou to čistě formálně vymezitelné neuronální procesy, nebo nějaký jazyk myšlení.

Podobně jako teorie soudů se chovají psychologické teorie kognitivního vyhodnocení (*cognitive appraisal theories*). Zdůrazňují, že to, jak bude vypadat příslušná emoce, závisí především na způsobu, kterým individuum vyhodnotí a ocení situaci, v níž se nachází. Teorie kognitivního vyhodnocení se však nespolehají na pojmy lidové psychologie, jako je přesvědčení, přání, záměr apod., s nimiž počítá filosoficky motivovaná teorie soudů. Modely vyhodnocující příslušné percepční stimuly v rámci teorií kognitivního vyhodnocení jsou komplexnější a zahrnují řadu hodnotících komponent. Rozdíl mezi jednotlivými teoriemi jsou dány konkrétními hodnotícími prvky. Například Rosemanová uvádí sedm hodnotících komponent (neočekávanost situace, stav

situace, stav vzhledem k motivaci, pravděpodobnost, činitel, kontrolní potenciál, typ problému), přičemž každá komponenta se může vyskytovat v různých odpovídajících hodnotách (očekávaný/neočekávaný, chtěný/nechtěný, minimalizace trestu/maximalizace odměny, jistota/nejistota, příčina v okolnostech/příčina ve druhé osobě/příčina v první osobě, nízký/vysoký, instrumentální problém/vnitřní problém).² Klíčová idea tohoto modelu je v tom, že aby mohla nějaká emoce vzniknout, musí stimul projít systémem všech hodnotících komponent, a v každé z nich mu je přiznána některá z možných hodnot. Výsledná kombinace hodnot je pak určující pro typ emoční odpovědi.

Vůči tomuto přístupu lze namítat, že navrhuje relativně složitý systém hodnotících kritérií, a tudíž předpokládá vysokou časovou náročnost zpracování podnětů kognitivním systémem. Tuto námitku se pokouší řešit odkaz na aktivní roli nevědomí. Složitost hodnotících kritérií a procedurální a časová náročnost kognitivního ohodnocování je však předmětem kritiky nekognitivních teorií emocí.

2 Nekognitivní teorie emocí

Nekognitivní teorie jsou ty, které obhajují tvrzení, že soudy nebo kognitivní ohodnocení nejsou částí emočního procesu. Emoční reakce je podle nich často tak rychlá, že zde nevzniká prostor pro soudy, evaluace či vyšší kognici. Takto chápaný emoční proces je automatickou reakcí na relevantní podnět. Je to právě složitost a množství hodnotících kritérií, co klade vyšší náročnost na čas a formu zpracování kognitivním systémem. S tímto problémem se kognitivní teorie emocí vypořádávají odlišně od nekognitivních teorií emocí. Zatímco kognitivní teorie emocí připouští vysokou rychlost vzniku některých emocí, takřkajíc bez vědomého úsilí, a proto jsou nuceny uznat, že procedura ohodnocení vypadá automaticky proto, neboť se v těchto rychlých případech odehrává na nevědomé úrovni. Nekognitivní teorie emocí reagují na přítomnost rychlých emočních reakcí tvrzením, že aspoň některé (nebo všechny) emoce jsou nezatížené kognitivním ohodnocením.

Jejich úvaha oživuje běžný psychologický náhled na emoce, který počítá s tím, že emoce nejsou racionální a čistě logické komponenty naší psychiky, nýbrž iracionální a v širokém smyslu nekontrolovatelné

² Viz Roseman (2001, 68-69). Model je značně komplexní a nelze se mu zde podrobněji věnovat.

reakce na určité události. Přitom hlavní námitka nekognitivistů vůči kognitivnímu ohodnocení spočívá v tom, že malé děti a subhumánní živočichové jistě nemají kognitivní schopnosti popisované teorií soudů nebo teorií kognitivního ohodnocení, přesto jim nikdo neupírá schopnost produkovat kvalitní emoční reakce (srov. Nadel 2004, 115-116; Lewis 2008, 304). S takovou námitkou lze jen souhlasit, i když z hlediska vědeckého přístupu je stále předmětem sporů, nakolik jde pouze o automatickou reakci bez sémantické hodnoty a nakolik si je dítě plně vědomé toho, co prožívá. Podle mého soudu, a bez ohledu na empirická zjištění týkající se sémantické evaluace podnětu, děti, které reagují na úsměv úsměvem, mají přinejmenším to, co níže nazývám primitivní prožitek. Kdyby tento primitivní prožitek neměly, nedovedly by reagovat na podnět, kterým je úsměv dospělé osoby nebo staršího dítěte.³ Je však již diskutabilní, a to má být předmětem empirických výzkumů, v jakém věku již děti dovedou reflektovat sémantickou a intencionální stránku emočního prožitku.

Na nekognitivních teoriích je důležité, že kladou velký důraz na automaticnost emoční reakce na podnět. U Paula Ekmana a Paula Griffithse se tato skutečnost projevuje v deklaraci vrozenosti mechanismů, které zajišťují rychlost a optimální reakci na podnětovou situaci. Ekmanův model sice netvrdí, že všechny emoce jsou utvářeny nekognitivně, ale nezávislost na vyšších kognitivních funkcích platí aspoň pro tzv. primární nebo základní emoce. Ekmanův model vyhodnocující primární emoce pracuje se dvěma mechanismy: s mechanismem automatického ohodnocení a s afektivním programem. Primární emoce, mezi něž se podle Ekmana⁴ řadí strach, hněv, znechucení, překvapení, radost a smutek, vznikají díky činnosti mechanismu automatického ohodnocení (automatic appraisal mechanism). Podle Ekmana

musí existovat hodnotící mechanismus, který si selektivně všímá těch stimulů (externích nebo interních), které jsou příležitostí k aktivaci afektivního programu [...] Vzhledem k tomu, že interval mezi stimulem a emoční odpovědí bývá někdy zvlášť krátký, musí být hodnotící mechanismus schopný operovat s vysokou rychlostí. Často je ohodnocení

³ Tento názor podporují i habituační studie dětské schopnosti rozlišovat emoce. Dvou až tří měsíční děti jsou schopny rozlišit mezi obličejovými výrazy štěstí, smutku a strachu. Viz Nadel (2004, 115).

⁴ Viz Ekman – Friesen (1971). Někdy se sem řadí i opovržení.

nejen rychlé, ale odehrává se nevědomě. Proto musím postulovat automatickou činnost ohodnocovacího mechanismu. (Ekman 1977, 58)

Tento hodnotící mechanismus má podle Ekmana kontakt s pamětí týkající se předchozí zkušenosti, takže může porovnáním rychle vyhodnotit, jaký podnět se vztahuje k určité emoci. Na druhou stranu ale nejde o klasický kognitivní hodnotící systém, jak byl popsán výše. Jde spíše o evolučně vytvořený mechanismus, jakýsi modulární algoritmus, který efektivně reaguje na podněty bez zapojení vyšších kognitivních procesů. Avšak v tom, že hodnotící procesy mohou být nevědomé, se Ekmanův postoj blíží kognitivním teoriím. Oproti tomu v případě pomalého ohodnocování mohou být hodnotící procesy promyšlené a vědomé (Ekman 1977, 59). Když je činnost hodnotícího mechanismu u konce, nastoupí druhý mechanismus, afektivní program, který přiřadí příslušnou emoci k tělesné reakci.

Kromě teorií, které popisují proceduru vzniku některých emocí jako nekognitivní, zatímco jiných jako kognitivní, takže usilují o vyvážené stanovisko, existují i radikálnější teorie. K alternativním teoriím bývá řazen názor, že neexistují emoce, které by se zakládaly (primárně) na kognitivních procesech. Jádro příslušné emoce, například strachu, neobsahuje kognitivní složku vůbec. Pokud se nějaká emoce vyznačuje přítomností kognitivní složky, pak se jedná o sekundárně přidružený aspekt emoce, který se objevuje až tehdy, když dojde k vědomé reflexi příslušné podnětové události. Při vzniku samotné emoce se naopak uplatňuje výhradně mechanismus afektivního ohodnocení (*affective appraisal*), jak tvrdí Jenefer Robinsonová: „Můj návrh je takový, že existuje sada zabudovaných afektivních hodnotících mechanismů, které jsou u primitivnějších druhů a novorozenců automaticky vyladěny na specifický stimul. Avšak jak se lidé učí a vyvíjejí, mohou být tyto mechanismy obohaceny o komplexnější stimuly, zahrnující složité soudy nebo myšlenky“ (Robinson 2001, 41).

Robinsonová tedy tvrdí, že ohodnocovací mechanismus je jen jeden a primárně pracuje za účelem afektivního vyhodnocení příslušného podnětu. Kognitivní složka se k němu může přiřadit. Afektivní hodnocení se přitom odehrává v několika jednoduchých kategoriích typu „To je přítel“, „To je nepřítel“, „To se mi líbí“, „To se mi nelíbí“ apod. Toto jednoduché ohodnocení podnětu pak vyvolává fyziologickou aktivitu, tendenci k jednání, výrazové změny obličeje a společně tak konstituuje emoční odpověď. Domnívám se však, že abychom mohli tuto teorii

považovat za čistě nekognitivní, museli bychom tyto jednotlivé charakteristiky chápat čistě popisně. Měli bychom je brát jako něco, co při registraci a ohodnocení podnětu nevystupuje ve své jazykové, a tudíž ani kognitivní podobě. V opačném případě by tato teorie nemohla být kategorizována jako nekognitivní.

3 Teorie somatické zpětné vazby

Jak kognitivní tak i nekognitivní teorie vzniku emocí přiznávají určitý podíl tělesným pochodům na vzniku příslušné emoce. Liší se v tom, jak významný tento podíl je. Zcela převažující význam somatických pochodů pro vznik emocí zdůrazňuje teorie somatické zpětné vazby. Podle této teorie každá emoční kategorie má svůj specifický projev na tělesné úrovni, a právě díky tomuto zcela určitému vzorci somatické aktivity můžeme emoce diferencovat. Takže strach z hada má svůj nezaměnitelný tělesný projev, který je odlišný od pocitu štěstí, že had zmizel v trávě a neuštkl mě.⁵ To, co činí emoci skutečnou emoci, je zpětná vazba, která přichází z těla a je registrována myslí (mozkem). Mysl následně tyto tělesné údaje vyhodnotí a „přiřadí“ příslušnou emoci.

V klasické podobě je tato teorie známa od Williama Jamese a Carla Langea. Tento model nedávno revidoval Antonio Damasio a Jesse Prinz. V Jamesově modelu přísně vzato nejde o přiřazení nějaké emoce uschované v rezervoáru mysli k nějakému tělesnému stavu, ale samo vnímání tělesného stavu je onou emoci. Takže když se nám při pohledu na hada zvýší krevní tlak, tepová frekvence a změní výraz obličeje, pak subjektivní vědomé vnímání těchto tělesných změn je mentálním stavem, který se nazývá emoce strachu. Vzniklá emoce je tedy důsledkem a nikoli příčinou tělesného stavu. Neopodstatněné problémy s emocemi vznikají podle Jamese tehdy, když emoce chápeme jako zvláštní entity usazené kdesi v mysli a čekající, až je kognitivní systém vyzve, aby se ukázaly. Jamesova teorie je zcela opačná. „Tělesné změny následují bezprostředně po percepci vzrušujícího podnětu a náš pocit těchto tělesných změn, tak jak se objevují, je emoci“ (James 2007, 449). Pokud je

⁵ Zajímavé je, že tato teorie hodnotí emoce na kategoriální úrovni a nevyšímá si rozdílů mezi strachem z hada a strachem z výšky. Předpokládá tedy, že oba pocity jsou týž. To, co je odlišuje, je sémantická a intencionální stránka. Viz o tom níže.

tedy emoce vnímáním ne-volních fyziologických změn těla, pak ji můžeme chápat jako vedlejší produkt převážně vrozených behaviorálních a fyziologických reakcí. Důsledkem takového chápání vzniku emoce je i skutečnost, že bez tělesných změn, nemůže existovat žádná emoce. James se v tomto ohledu vyjadřuje přesně: „Když si představíme nějakou silnou emoci, a následně se pokusíme abstrahovat od našeho vědomí všech charakteristických tělesných pocitů a symptomů, pak zjistíme, že nám nezůstane nic. Nezůstane nám žádná 'látka myslí', z níž by mohla být emoce utvořena, nýbrž jen chladný a neutrální stav intelektuální percepce“ (James 2007, 451).

Pokud však přistoupíme na tvrzení, že bez tělesných změn (míněny jsou automatické tělesné změny v organismu, nikoli změny v mozku) nejsou ani emoce, pak lze jen stěží vysvětlit, jak by například herci mohli prožívat silné emoce. Jde o ty případy, kdy hercovy emoce nejsou způsobeny vnímáním vlastních tělesných změn, protože neexistuje žádný vnější podnět, který by tělesnou změnu dokázal přímo (bez činnosti vyšší kognice) vyvolat. Herci přece dobře ví, že když se na jevišti objeví had, nemusí se ho skutečně obávat, protože jde jen o atrapu. Moderní zobrazovací metody tento názor potvrzují. Ukazuje se, že v případech předstírané emoce se zapojují jiná centra v mozku než v případě emocí skutečných.⁶ To je další důvod, proč se níže pokouším rozlišit primitivní a kognitivní (jazykovou) stránku emocí. Toto rozlišení má ukázat, že emoce může být způsobena buď tělesnou zpětnou vazbou, nebo intelektuální aktivitou, nebo kombinací obojího.

Jesse Prinz vychází z James-Langeovy teorie a uznává roli percepce tělesného stavu při vzniku emoce. Na druhou stranu připouští námitky vůči této teorii, z nichž asi nejvýznamnější je ta, že takto úzce vymezené emoce nemohou nikdy vstřebat sémantickou a intencionální kognitivní složku. Emoce jsou v tomto případě takříkajíc uzavřené v tělesné percepci. Ve svém pozměňovacím návrhu Prinz přijímá roli zpětné vazby, která informuje mysl o tělesných změnách. Díky zpětné vazbě mysl tyto změny registruje (nominální obsah emoce), a to bez jakéhokoli kognitivního úsilí. Avšak vtělená emoce, která takto vznikne, má i svou složku reprezentační (reálný obsah emoce) (Prinz 2004b, 69). Emoce se

⁶ Viz Koukolík (2005, 158). Spontánní smích kontroluje systém amygdala-thalamus-hypothalamus-horní část mozkového kmene, kdežto volní herecký smích kontroluje premotorická oblast čelní kůry-motorická kůra-přední oblast mozkového kmene.

intencionálně vztahuje nikoli k tělesnému stavu, nýbrž k tomu, co reprezentuje. Tuto modifikovanou teorii nazývá vtělená ohodnocovací teorie (*embodied appraisal theory*). „Emoce jsou vtělené, stejně jako navrhol James a Lange. Jsou to percepce somatických změn. Ironicky však jsou též hodnotící“ (Prinz 2004a, 57). Ohodnocení Prinz definuje nikoli jako hodnotící soud, nýbrž jako reprezentaci vztahu organismu a jeho okolního prostředí. Jde tedy o intencionální vztah k vnějšmu stimulu, který vyvolal tělesné změny.

V této teorii se však musí počítat s tím, že vlastní konstituce intencionální reprezentace nemusí být nutně vázána na nějaké sémantické ohodnocovací procedury. Intencionální reprezentace může vzniknout i tehdy, pokud mysl nemá k dispozici mechanismus sémantického ohodnocení, který známe u dospělých jedinců. U nich většinou předpokládáme, že intencionalita je něco, co je pevně svázané s jazykem. Proto se intencionální vztah nemůže vytvářet výhradně na bázi jazyka, i když prostřednictvím jazyka může být popisován, a tudíž ani výsledná reprezentace nemá jazykový charakter. Kdyby byla veškerá intencionalita v tomto smyslu jazyková, pak bychom těžko vysvětlovali, jak se malé děti, které ještě neovládají žádný přirozený jazyk, mohou intencionálně vztahovat k určitému objektu.

4 Primitivní a konceptuální prožitek

V předchozí části jsme poukázali na sedm zajímavých rysů, které lze identifikovat u základních teorií emocí. První rys se týkal toho, že v rámci jakékoli teorie emocí musíme vysvětlit, jak je možné, že dva různí pozorovatelé mohou reagovat na tutéž událost různým emočním způsobem, anebo naopak, že jeden pozorovatel může reagovat na tutéž událost v různém čase odlišně. Zdá se, že kognitivní teorie emocí doveďou tyto aspekty vysvětlit lépe, než teorie nekognitivní.

Druhý rys upozorňuje na to, že existuje mnoho zcela odlišných fyzických událostí, které spolu nemají zdánlivě nic společného, přesto dokážou vyvolat stejnou emoci. V rámci kognitivní teorie se to vysvětluje tak, že subjekt v takových případech kognitivně (jazykově, potažmo konceptuálně) ohodnotí příslušnou fyzickou událost, čímž determinuje typ vzniklé emoce. Kdyby tak neučinil, nemohla by žádná emoce vůbec vzniknout. Problém však nastává v případech, kdy subjekt nedisponuje vyššími kognitivními funkcemi na úrovni, která takové ohodnocení umožňuje.

Třetí rys vychází z takového typu kognitivní teorie emocí, který využívá při ohodnocovací proceduře soudy, přičemž dále upozorňuje, že rychlost, s jakou některé emoce vznikají, naznačuje, že procedura souzení může v těchto případech probíhat nevědomě. Opět zde nastává potíže, a to s možnostmi kognitivního ohodnocování u malých dětí, jejichž schopnosti jsou v tomto ohledu zřejmě relativně malé. Podobně to platí i pro vysvětlení rychlosti, s níž emoce vznikají. Kdyby se ohodnocovací procedury realizovaly u dětí nevědomě, řekněme na neuronální úrovni, pak by tyto děti už musely mít od narození neuronální program, který takové ohodnocení umožňuje. Rozdíl mezi dětmi a dospělými by pak byl v tom, že dospělí lidé dovedou tyto procedury vědomě reflektovat díky přirozenému jazyku.

Čtvrtý rys zdůrazňuje, že teorie kognitivního ohodnocení může vypadat i tak, že obsahuje četná hodnotící kritéria, která společně tvoří relativně složitý systém. Tento kognitivní systém však může být výpočetně a časově vysoce náročný, takže by teoreticky vědomě nezvládl vyhodnotit podnět dřív, než dojde ke skutečné emoční reakci. V těchto případech se většinou poukazuje na roli nevědomých kognitivních procesů, které jsou rychlejší, protože nezapojují vědomou složku, která celý proces zpomaluje.

Pátý rys upozorňuje, že nekognitivní teorie emocí využívají automatický ohodnocovací mechanismus. Tento mechanismus efektivně reaguje na podněty bez zapojení vyšších kognitivních procesů a přiřadí k podnětu příslušnou emoci. Při rychlém ohodnocování jsou procesy nevědomé, při pomalém ohodnocování mohou být hodnotící procesy promyšlené a vědomé. Když je činnost hodnotícího mechanismu u konce, nastoupí druhý mechanismus, afektivní program, který přiřadí příslušnou emoci k tělesné reakci.

Šestý rys se týká somatických teorií zpětné vazby. Ty lépe vysvětlují rychlost, s níž osoba emočně reaguje na určitý podnět pomocí reflexe tělesných změn. Na druhou stranu tyto teorie otvírají otázku, jak odlišit pocit strachu z hada od pocitu strachu ze státní zkoušky, když se pocit vytváří vnímáním tělesných změn, jako je změna tlaku krve, tepové frekvence, pocení apod. V takových případech se nabízí návrh obohatit tyto procesy o procesy sémantické.

Šedmým rysem, který se také týká somatických teorií, je, že emoce jsou co do svého vzniku zcela závislé na somatických změnách. Toto tvrzení o závislosti emocí čistě na tělesných změnách je však neudr-

žitelné, neboť bychom nedokázali vysvětlit, jak vytvářejí emoce herci, kteří je předstírají.

Poslední rys reflektuje Prinzovu variantu somatické teorie, která vznikla evaluací James-Langeovy teorie a obohatila ji právě o onen sémantický aspekt. Prinz respektuje tělesnou složku emoce a zavádí navíc složku reprezentační, která se vyznačuje intencionalitou. Tím je schopen vysvětlit i některé jevy, které čistá somatická teorie vysvětlit nedokáže. Na druhou stranu vzniká opět otázka, jakým způsobem reprezentují podnět utvářející emoci malé děti. Ty děti, které ještě nedisponují jazykově zformulovanou intencionalitou, přestože je evidentní, že se velice brzy dovedou intencionálně vztahovat k různým objektům.

Jako reakci na výše uvedené rysy lze navrhnout koncept rozlišující primitivní a kognitivní složku emocí. Toto rozlišení má zapadat do představy, že emoce může být způsobena buď tělesnou zpětnou vazbou, nebo intelektuální aktivitou, nebo kombinací obojího. To, jakým z uvedených způsobů bude emoce konstituována, záleží především na tom, v jaké fázi ontogenetického vývoje se člověk právě nachází. Domnívám se, že v každé fázi jsou mu dostupné určité typy emočních reakcí a že s přibývajícím věkem se více uplatňuje kognitivní evaluace, což umožňuje diferencovat složitější a speciálnější typy emočních reakcí nebo dokonce emoční reakci vědomě zcela potlačit.

Primitivní prožitek je termín označující hlavní složku emocí, pomocí níž většinou emoce intuitivně identifikujeme. Vezmeme-li sadu základních emocí, jak je prezentuje Ekman a Friesen, což je strach, hněv, znechucení, překvapení, radost a smutek, pak tvrdím, že každá z těchto emocí může mít dvě složky: složku primitivní a složku kognitivní. Primitivní složka emoce představuje to, co prožíváme přímo a bezprostředně, když danou emoci zakoušíme. Ve filosofii myslí se nazývá termínem „kvále“ nebo „subjektivní kvalitativní zkušenost“. I když symbolický význam subjektivního prožitku může být značný (například důležitý pro přežití), je pojímán jako čistě performativní komponenta lidské emoce. Subjektivní zkušenost s nějakou emoci znamená, že se jedná o interní prožitek na fenomenální úrovni. Takové subjektivní zakoušení není podmíněno ohodnocovacími mechanismy na úrovni vyšších kognitivních funkcí, ani není spojeno s lidově-psychologickými teoriemi soudů. Tuto představu podporují možná poněkud paradoxně a po výtce nepřímou, i neurofyziologické empirické poznatky založené na studiu mozkových záznamů (Koenig 2004, 118). Tyto záznamy

ukazují, že v případě rychlých reakcí na emočně významné podněty, se zapojují spíše evolučně starší mozkové oblasti, jakými jsou limbický systém, talamus a amygdala.⁷

Podle teorie Josepha LeDoux (1998, 248) talamus odesílá informace o vnějších podnětech dvěma drahami. Kratší cesta vede do amygdaly, kdežto delší cesta do sensorických, především zrakových center v mozkové kůře. Kratší cesta zodpovídá za rychlou emoční reakci, která však často může být neadekvátní příslušné situaci. Například rychle emočně reagujeme na zkroucený klacek na cestě, protože se domníváme, že je to had. Klacek vyvolá prakticky okamžitou emoční reakci, již zde nazývám primitivní složkou emoce. V tomto případě se emoce vyskytuje pouze ve své primitivní podobě – jde čistě o prožívaný aspekt percepční zkušenosti. Tento primitivní prožitek však trvá většinou jen krátký časový okamžik, protože zpravidla následuje fokus na předmět a kognitivní ohodnocení. Kognitivní ohodnocení prokáže, že nejde o hada, nýbrž o zkroucený klacek, což vede k modulaci emoční reakce na adekvátní emoční hladinu. Zde se již objevuje komplexní emoce, tvořená jak nekognitivní (primitivní), tak i kognitivní (jazykovou, konceptuální) složkou.

Mechanismus vzniku primitivního prožitku se tedy z neurofyzilogického hlediska váže na evolučně starší centra v mozku, která byla přítomna již u prvních savců a zajišťovala funkci emocí a instinktů podílejících se na přežití (srov. Plhánková 2005, 392). Proti tomu vývojově nejmladší část mozku, již je neokortex, umožnil realizaci kognitivních ohodnocovacích procedur. Tomu odpovídá i náš model rozlišující u základních emocí primitivní a kognitivní složku. Tento rozdíl lze snadno naznačit, když si uvědomíme diferenci mezi pocitem překvapení a důvody, které k němu vedly. Je-li někdo překvapen, že dostal květinu bez zjevného důvodu, nebo pociťuje strach, když se dívá dolů z vysoké věže, pak je třeba oddělovat pocit překvapení, resp. strachu od jazykové a případně i konceptuálně zformovaných důvodů, které vedou k vyvolání takového pocitu, či pocit dále usměrňují. Květinu da-

⁷ Kromě představy, že máme primitivní prožitek na úrovni základních emocí, lze argumentovat ve prospěch primitivního prožitku jednak na případech somatické percepční zkušenosti, ale také na případech netělesných prožitků. Manifestace prožitekové komponenty na úrovni somatické percepcce jsou pocity lechtání, svědění, orgasmu, bolesti, hladu apod. Netělesné prožitky jsou např. pocity viny, krivdy apod. Těmto případům se zde nebudeme věnovat.

rovaná bez zjevného důvodu či pohled z vysoké věže jsou příčiny příslušného pocitu a netvoří jeho prožitkovou složku, nýbrž intencionální, příp. jazykově zformovanou složku.

Existují také kontroverzní případy, kdy není zcela jasné, zda se příslušná základní emoce vůbec může modulovat kognitivní komponentou. To je případ úzkosti. Emoce úzkosti je definována jako „neurčitý pocit obav či ohrožení, který se neváže k žádnému konkrétnímu objektu či události“ (Plháčková 2005, 405). Tato emoce tak postrádá minimálně jednu důležitou kognitivní komponentu, jíž je intencionalita. Takovou emoci bychom pak mohli vymezit čistě její prožitkovou performační komponentou, a nikoli konceptuálně. Je samozřejmě možné se pokusit přiblížit k povaze tohoto prožitku popisem událostí, které by mohly být jeho příčinou. Avšak takový popis je pouze opisem příčin, které ke vzniku příslušného prožitku vedly, a nikoli popisem prožitku samotného. Prožitková komponenta emoce tedy nic netvrdí a má pouze prezentační funkci nebo funkci habituační reakce.

Kognitivní (ve smyslu jazyková) komponenta je druhá složka emoce a uplatňuje se při kognitivním ohodnocování. Z evolučního hlediska se vyvinula později, a to u organismů s funkční mozkovou kůrou, která shromažďuje a zpracovává materiál ke kognitivnímu ohodnocení. V zásadě souhlasím s kognitivistickými teoriemi v tom, že při tvorbě emoce hrají důležitou roli ohodnocovací mechanismy, které pracují s informací na kognitivní úrovni. Nemyslím si však, že to platí pro všechny případy emoční zkušenosti. Ty případy, ve kterých se emoce vyznačuje konceptuální složkou, na sebe váží mnoho charakteristických rysů, které zde nelze detailně popisovat. Nicméně zpravidla platí, že prožitková a kognitivní komponenta jsou těsně spjaté v tom smyslu, že se vyskytují v příslušné emoci společně. Když máme například pocit strachu, ukazuje nám kognitivní jazyková komponenta, že se bojíme hada nebo zakroucené větve. Těsné sepětí obou komponent se zde projevuje ani ne tak v nevědomém ohodnocení percepční zkušenosti, jako spíše v uvědomělé reflexi intencionálního vztahu našeho subjektivního prožitku k vnější události. Abychom mohli rozlišit zakroucenou větev od hada, musíme vnější podnět nejen rozpoznat jako něco vnějšího a jako objekt, ale musíme mu také přiřadit příslušné vlastnosti. Toto přiřazení vlastností jako např. „je to dřevěné, zakroucené, neživé, nehýbající se“ se děje na kognitivní jazykové úrovni. Není přitom důležité, jestli se tak děje na nevědomé úrovni, nebo jestli proceduru ohodnocování vědomě reflektujeme.

Kognitivní jazyková komponenta emocí je důležitá i v případech, kdy usilujeme o diferenciaci základních emocí. Například u hněvu můžeme rozlišovat podrážděnost, vztek, spravedlivý hněv. Toto rozlišování není na primitivní úrovni vůbec možné, protože tam neexistuje mechanismus, který by to umožňoval. Proto se různé typy hněvu liší nikoli odlišnou prožitkovou komponentou, ale tím, že jim přiřadíme odlišné charakteristiky na kognitivní jazykové úrovni. Dokážeme například konceptuálně rozlišit různé intencionální objekty hněvu, dovedeme je reprezentovat, popsat situaci, díky níž se hněváme, což jsou případy, jimiž specifikujeme příslušnou základní emoci. Konceptualizované emoce jsou tedy strukturované, symbolické a vědomé. I když platí, že sama ohodnocovací procedura nemusí být vědomá.

Proti tomu primitivní komponenta emocí není intencionální. Nevztahuje se sama o sobě, tak jak je prožívána, k žádnému konkrétnímu objektu. Podle mého názoru totiž na primitivní úrovni máme jen jeden pocit strachu, jeden pocit radosti, jeden pocit překvapení atd. Sice lze hovořit o rozdílu mezi euforií a extází v případě štěstí, nebo mezi melancholií a zabloubaností v případě smutku, nebo mezi panikou a plachostí v případě strachu, ale tento rozdíl je jen na úrovni intenzity prožitku nikoli na úrovni typu prožitku (jako je radost, překvapení, strach apod.). To, co vzájemně odlišuje jednotlivé pocity strachu, je samozřejmě intencionální objekt, který ale vzniká až v momentě kognitivního ohodnocení na úrovni jazykového či konceptuálního zachycení.

Aby se emoce mohly k něčemu intencionálně vztahovat nebo aby mohly něco reprezentovat, potřebují kognitivní složku. Přitom nehraje roli, zda jde o intencionální vazbu k tělesným stavům, vnějším objektům a situacím nebo dokonce k jiným mentálním stavům v lidské mysli. Příklad vztahu k tělesným změnám může představovat kousnutí komárem. To pocítujeme jako bolestivý prožitek a intencionálně se vztahujeme k oblasti, v níž došlo ke kousnutí, případně i ke komárovi, který právě odlétá. Příklad intencionálního vztahu k mentálním stavům je také zjevný. Při čtení knihy se můžeme vžít do role hlavního hrdiny a soucítit s ním i na úrovni emočního prožitku (takovou zkušenost většina z nás má), přestože žádný intencionální objekt, jímž je onen smyšlený hrdina, reálně neexistuje. Takto konceptualizovanou intencionalitu dokazují i poněkud neobvyklé případy emocí, jako jsou pocity v amputovaných končetinách nebo pocity hladu u pacientů po úplné gastrektomii (viz Wangensteen – Carlson 1931).

Uvedl jsem, že je třeba rozlišovat primitivní a kognitivní komponentu základních emocí. Toto rozlišení nekoncepuji tak, že základní emoce (dle Ekmana) jsou zároveň primitivní emoce. Naopak se domnívám, že základní emoce se mohou vyskytovat v těchto podobách:

- 1) Emoce obsahující jen primitivní komponentu – to je případ subjektivního prožívání.
- 2) Emoce obsahující primitivní a kognitivní komponentu – to je případ většiny emocí trvajících delší dobu, kdy máme možnost si uvědomit a kognitivně (jazykově, příp. konceptuálně) zpracovat vnější i vnitřní (z paměti) podněty mající vliv na tvorbu emoce. Výjimku tvoří např. pocit úzkosti.
- 3) Emoce obsahující jen kognitivní komponentu.

Poměrně kontroverzní je poslední typ výskytu emocí, který naznačuje existenci kognitivní složky emocí při absenci primitivní (prožitkové) složky. Je také možné, že žádné pouze kognitivní emoce být nemohou, protože pak už by se nejednalo o emoce. Pokusím se díky příkladům spekulativně zamyslet, zda se v tomto případě skutečně jedná o *contradictio in adjecto* či nikoli. Předpokládám přitom, že vznik emoce je nějak spjat také s viscerální percepcí a percepcí hranic vlastního těla, která se projevuje ve způsobech chování a pohybech obecně.⁸

Prvním příkladem je výše uvedený případ s herci, kteří na jevišti předstírají emoce. Jejich umění spočívá ve věrném napodobení tělesné behaviorální a obličejové reakce někoho, kdo se kupříkladu právě směje vtipné situaci. Zde však možná dochází pouze k tělesné reakci, která vzniká jako důsledek konceptualizace emoce na kognitivní úrovni. Tuto představu podporují i zobrazovací metody snímající činnost lidského mozku. Při spontánním smíchu se aktivují starší části mozku (amygdala, část thalamu a hypothalamu, horní část mozkového kmene), kdežto při předstíraném smíchu se aktivují oblasti mozkové kůry (premotorická oblast čelní kůry, motorická kůra, přední oblasti mozkového kmene; viz Koukolík 2005 158). Na jiné případy umožňující rozlišovat primitivní komponentu emocí od komponenty kognitivní upozorňují neurologové. V patologické podobě se lidé smějí či pláčí, ale neprožívají příslušné emoce, tvrdí František Koukolík (tamtéž). Nebo

⁸ Tématu percepcce a rozdílu mezi primitivní a kognitivní složkou zkušenosti se budu věnovat v jiné stati.

si lze představit situaci předstíraného orgasmu, kdy se opět nezapojuje primitivní komponenta, i když se o to dotyčná žena může snažit. Emoční prožitek tedy musí být něco jiného než čistě tělesný behaviorální stav nebo jeho kognitivní konceptualizace.

Závěr

Cílem statě bylo jednak ukázat, že standardní teorie emocí, a to jak kognitivní, nekognitivní, tak i teorie somatické zpětné vazby, mají řadu významných rysů. Některé rysy zformulované v příslušných teoriích otvírají otázky, které jsou však srozumitelně řešitelné pouze v teoriích jiných. Jde například o to, jak je možné, že dva různí pozorovatelé mohou reagovat na tutéž událost stejnou/různou emoční reakcí. Nebo jak je možné, že stejná osoba může reagovat na tutéž událost v různém čase různě. Jde i o důležitou otázku, jak je to s možnostmi kognitivního ohodnocování u malých dětí, jejichž schopnosti jsou v tomto ohledu zřejmě relativně malé. Dále mne zajímalo, jak se vyrovnat s tím, že některé emoce nemusí být vázány na tělesné změny. V neposlední řadě jsem se věnoval otázce možnosti nevědomého kognitivního ohodnocování. Všechny tyto rysy se nějak dotýkají toho, co má vliv na vznik určité emoce a měly by být řešitelné v rámci naznačeného rozlišení. Nejednotné způsoby řešení těchto otázek jsem se aspoň částečně pokusil sjednotit díky rozlišení mezi primitivní a kognitivní (jazykovou) složkou emocí. Toto rozlišení by mělo umožnit vytvořit schéma zobrazující vznik různých emocí a také má ukázat, jaké vztahy mezi primitivní a konceptuální komponentou emoce mohou nastat. Pokud jde o rozdíly v emočních reakcích na tentýž (nebo naopak na různý) podnět, lze je vysvětlit právě zapojením kognitivního ohodnocení, které se realizuje na kognitivní úrovni u lidí, kteří disponují přirozeným jazykem. Například důležitá otázka emočního prožívání u kojenců naznačuje, že dítě nemá vyvinutou jazykovou složku, ale emoce zakoušet dokáže na úrovni primitivního prožitku. Skutečnost, že ne všechny emoce lze odvíjet od tělesných změn, jsem se pokusil ukázat na intencionální vazbě kognitivní složky emoce k jiným mentálním stavům. V souladu s nejnovějšími poznatky kognitivní vědy považuji nevědomé kognitivní ohodnocování za velmi pravděpodobné.

Hlavní teze této statě se týkala rozdílu mezi primitivní a kognitivní složkou základních emocí. Tento rozdíl jsem uplatnil na tzv. základní (primární) emoce. Tyto emoce, jakými jsou strach, hněv, znechucení,

překvapení, radost a smutek, vznikají buď jako automatické reakce na podnětovou událost, nebo se při jejich konstituci zapojuje i vyšší kognice. Všechny primární emoce mají primitivní (prožitkovou) komponentu, která je sama o sobě nekognitivní (nepojmová) a není ani vyjádřitelná jinak než opisně.

Katedra filozofie
Filozofická fakulta
ZČU v Plzni
Sedláčkova 19
306 14 Plzeň
miha@kfi.zcu.cz

Mezioborové aktivity
Výzkumné centrum Nové technologie (NTC)
Univerzitní 8
306 14 Plzeň
miha@ntc.zcu.cz

Literatura

- EKMAN, P. – FRIESEN, W. V. (1971): Constants across culture in the face and emotion. *Journal of Personality and Social Psychology* 17, 124-129.
- EKMAN, P. (1977): Biological and cultural contributions to body and facial movement. In: Blacking, J. (ed.): *A. S. A. Monograph 15, The anthropology of the body*. London: Academic Press, 39-84.
- JAMES, W. (2007): *The Principles of Psychology: Volume 2*. New York: Cosimo.
- KOENIG, O. (2004): *Emotion: neuroscience*. In: Houdé, O. (ed.): *Dictionary of Cognitive Science*. New York: Psychology Press, 118.
- KOUKOLÍK, F. (2005): *Mozek a jeho duše*. Praha: Galén.
- LEDoux, J. (1998): *Das Netz der Gefühle: Wie Emotionen entstehen*. München: Carl Hanser Verlag.
- LEWIS, M. (2008): The Emergence of Human Emotions. In: Lewis, M. – Haviland-Jones, J. M. – Barrett, L. F. (eds.): *Handbook of Emotions*. New York: The Guilford Press, 304-319.
- NADEL, J. (2004): Emotion: psychology. In: Houdé, O. (ed.): *Dictionary of Cognitive Science*. New York: Psychology Press, 115-116.
- NUSSBAUM, M. (2004): Emotions as Judgements of Value and Importance. In: Solomon, R. (ed.): *Thinking about Feeling: Contemporary Philosophers on Emotions*. New York: Oxford University Press, 183-199.
- PLHÁKOVÁ, A. (2005): *Učebnice obecné psychologie*. Praha: Academia.
- PRINZ, J. (2004a): Embodied Emotions. In: Solomon, R. (ed.): *Thinking about Feeling: Contemporary Philosophers on Emotions*. New York: Oxford University Press, 44-58.

- PRINZ, J. (2004b): *Gut Reactions: A Perceptual Theory of Emotions*. Oxford: Oxford University Press.
- ROBINSON, J. (2001): Emotion. Biological Fact or Social Construction? In: Solomon, R. (ed.): *Thinking about Feeling: Contemporary Philosophers on Emotions*. New York: Oxford University Press, 28-43.
- ROSEMAN, I. – SMITH, C. (2001): Appraisal Theory: Overview, Assumptions, Varieties, Controversies. In: Scherer, K. R. – Schorr, A. – Johnstone, T. (eds.): *Appraisal Processes in Emotion: Theory, Methods, Research*. New York: Oxford University Press, 3-19.
- ROSEMAN, I. (2001): A Model of Appraisal in the Emotion System. In: Scherer, K. R. – Schorr, A. – Johnstone, T. (eds.): *Appraisal Processes in Emotion: Theory, Methods, Research*. New York: Oxford University Press, 68-91.
- SOLOMON, R. (1977): The Logic of Emotion. *Noûs* 11, 41-49.
- SOLOMON, R. (2007): *True to Our Feelings: What Our Emotions Are Really Telling Us*. Oxford: Oxford University Press.
- WANGENSTEEN, O. H. – CARLSON, A. J. (1931): Hunger Sensations in a Patient after Total Gastrectomy. *Proceedings of the Society of Experimental Biology* 28, 545-547.

Matematický realismus a naturalismus Penelope Maddy¹

Vít Punčochář

*Akademie věd České republiky, Praha
Univerzita Karlova v Praze*

Abstract: This paper discusses different forms of mathematical realism and especially different kinds of advocating realistic position. Husserl's phenomenological approach, Gödel's (phenomenological and) methodological approach and Quine's naturalistic approach are presented. Afterwards, Maddy's realism is reconstructed as a synthesis of Gödel's and Quine's views, overcoming weaknesses of both. Also Maddy's later rejection of her own realism and its replacement by naturalistic project are discussed. At the end, we are offering our own interpretation of the problem of realism in mathematics.

Keywords: Mathematical realism, mathematical naturalism, Husserl, Gödel, Quine, Maddy.

1 Úvod

Matematický realismus je založen na představě, že matematické objekty existují nezávisle na našem jazyku, v němž o nich hovoříme, na našich definicích či konstrukcích těchto objektů, na naší schopnosti poznávat tyto objekty a celkově na naší matematické praxi. Pravdivost platných matematických vět je podle tohoto pohledu určena primárně charakterem objektivního světa matematických struktur a nikoli např. tím, že existuje důkaz těchto vět. Důkazy jsou chápány spíše jako posloupnosti kroků zachovávajících pravdivost než jako skutečné důvody pravdivosti.

¹ Práce na tomto článku byla podpořena z grantu GAČR P401/11/0371 Apriorní, syntetické a analytické od středověku po současnou filosofii.

Klíčová postava logiky 20. století a zastánce matematického realismu, Kurt Gödel, píše, že „objekty a teoremy matematiky jsou tak objektivní a nezávislé na naší svobodné volbě a naší kreativní činnosti, jako je fyzický svět“ (citát převzat z Maddy 1997, 89). Matematický realismus tedy může být chápán také jako teze, že matematika zkoumá matematické objekty ve stejném smyslu, v jakém fyzikové zkoumají objekty materiálního světa. Podobnou charakterizaci realismu najdeme také např. u Maddy (1990, 20) či u Kvasze (2010, 522).

Tato práce pojednává o různých způsobech zdůvodnění realismu, které ústí do různých podob realismu samotného. Rozlišíme přitom nejprve tři typy, které pracovním způsobem můžeme označit jako fenomenologický (Husserlův), metodologický (Gödelův) a naturalistický (Quinův) způsob zdůvodnění. Centrum práce pak tvoří výklad realismu Penelope Maddy, který je syntézou Gödelova a Quinova přístupu. Dalším tématem bude otázka, co vedlo Maddy k tomu, že svůj realismus opustila a proč ho nahradila svým naturalistickým postojem. V závěru uvedeme vlastní interpretaci realismu, která je z jedné strany motivovaná Husserlovým realismem a z druhé naturalismem Penelope Maddy, přestože ten byl původně zformulován spíše v opozici k realistické pozici. Vedle variace podob realismu samotného také význam slova „naturalismus“ bude v průběhu práce kolísat. Přes určité vnitřní souvislosti je třeba mít na paměti, že Quinův naturalistický způsob zdůvodnění realismu se odlišuje od naturalismu Penelope Maddy a na něm založené naší interpretaci realismu, kterou uvádíme v poslední části práce.

2 Husserlův realismus

Matematický realismus se jeví jako filosofie, která je v harmonickém souladu s běžnou matematickou zkušeností, s tím, co matematici prožívají, když provozují svoji vědu a co konec konců prožívá kdokoli, kdo řeší nějakou matematickou úlohu. Tak např. Moschovakis uvádí, že „hlavním bodem ve prospěch realistického přístupu k matematice je instinktivní jistota téměř každého, kdo vůbec někdy zkoušel řešit nějaký problém, že přemýšlí o ‚reálných objektech‘, ať už to jsou množiny, čísla nebo cokoli ...“ (citát převzat z Maddy 1990, 2).

Na specifickém charakteru matematické zkušenosti je založen Husserlův fenomenologický způsob zdůvodnění realismu. Ten je nejprehlednějším způsobem vyložen v Husserl (1913) a (1929).

Husserl vychází z perceptivní zkušenosti. Např. prožitek vidění chápe jako zvláštní stav vědomí, v němž nám jsou určitým způsobem dány materiální prostorové věci. Jinými slovy, ve specifickém charakteru prožitku smyslového vnímání je právě to, že se *v něm* před nás reálně staví určitý předmět (odlišný od prožitku samotného). A je celá sféra předmětů, které se nám takto dávají a mohou dávat, totiž viditelná fyzická příroda. Avšak předmět, který takto vnímáme, není pro Husserla „pouze něčím individuálním, nějakým ‚toto zde‘, něčím jednorázovým; je ‚sám v sobě‘ tak a tak ustrojený, a tudíž má svůj *svéráz*, svou skladbu *podstatných* predikabilí, které mu musejí příslušet... , aby mu pak mohla příslušet jiná určení, totiž určení sekundární, relativní“ (Husserl 1913, 23). V tomto smyslu má každá materiální věc svoji podstatu (eidos). K zachycení podstaty lze dojít tak, že předmět ve fantazii podrobíme svobodné variaci jeho možných podob a zachytíme invariant, kterým je právě ona podstata. Podle Husserla je dokonce možné vysledovat celou hierarchii podstat, protože z podstat můžeme získávat obecnější podstaty. Každý podstatný moment podstaty je sám podstatou. Podstaty získáváme nejen sledováním toho, co je pro jednotlivé předměty podstatné, ale také pozorováním toho, co má více různých předmětů společného. Podstata „materiální věc vůbec“ je podstatou všech materiálních věcí. Nejobecnější podstaty určují tzv. „regiony“ předmětů. Názorným příkladem je region „fyzická příroda“. Ale podstaty můžeme získávat i např. z vnitřního vnímání. Můžeme zachycovat podstaty vlastních světských duševních stavů (nás jakožto lidí ve světě). Tímto dospějeme k regionu „duše“. K regionům podstat se vztahují kategorie, které vyjadřují formu podstat (tato forma je sama podstatou odlišného typu). „Čistě logické základní pojmy... definujeme nyní jako *logické kategorie* či *kategorie logického regionu* ‚předmět vůbec‘. ... Patří sem ale i ‚významové kategorie‘... s ohledem na podstatné pravdy spojující ‚předmět-vůbec‘ a ‚význam-vůbec‘“ (Husserl 1913, 36). Mezi tyto čisté logické základní pojmy řadí Husserl i pojmy čisté matematiky.

Podle Husserla podstatu samu můžeme „vnímat“ analogickým způsobem, jako můžeme vnímat individuální přírodní objekty. Máme specifické prožitky, v nichž nám jsou dány podstaty. Způsob, jakým jsou nám dány, je rovnocenný s daností individuálních materiálních věcí v prožitcích vidění těchto věcí. „I nazření podstaty je ve vlastním smyslu názorem, stejně tak jako je eidetický předmět ve vlastním smyslu předmětem“ (Husserl 1913, 25). Dochází zde k rozšíření pojmů

„vnímání“, „zkušenost“, „nazírání“ a vzniká nový pojem tzv. „originárního“ uchopení předmětu, což představuje prožitek přímé danosti nějakého předmětu (předmětu v nejobecnějším, logickém smyslu).

Regiony se sice vyznačují vzájemnou komplikovanou provázaností (např. region „duše“ je fundován v regionu „materiální věc“, přestože se jedná o odlišné regiony), ale v jistém smyslu jsou rovnocenné. Husserl nevidí důvod tvrdit, že originární dávání se přírodního předmětu v prožitku vidění tohoto předmětu je více vnímáním než originární dávání se takového ideálního útvaru, jako je např. číslo v nazírání tohoto čísla. Dojem, že to tak je, je podle Husserla přirozeným předsudkem. „Slepota vůči idejím je jistým druhem duševní slepoty... Ve skutečnosti všichni a takřka neustále vidí „ideje“, „podstaty“, operují s nimi v myšlení, provádějí i soudy týkající se podstat“ (Husserl 1913, 53).

Husserlův „princip všech principů“ říká, že

každý originárně dávající názor je zdrojem oprávnění pro poznání a že vše, co se nám v „intuici“ originárně nabízí (takřkajíc ve své tělesné skutečnosti), je třeba brát jednoduše tak, jak se to dává, ale rovněž jen v mezích, v nichž se to zde dává. (Husserl 1913, 56)

Z tohoto hlediska má být filosoficky zdůvodněna např. aritmetika jako věda o číslech, protože předměty, o kterých pojednává, nám jsou originárně dány v odpovídajících názorech. Je přitom nutné v těchto vědách respektovat smysl originárně daných předmětů a chápat je skrze něj. To konkrétně znamená, že respektujeme, že čísla a jiné matematické objekty jsou objektivní, nezávisle existující předměty, stejně tak jako jimi jsou předměty přírodních věd.

Následující citace se opět vztahuje na předměty logiky pojaté v širokém smyslu, tj. zahrnující též matematické objekty:

Všechno toto objektivní nemá jen pomíjivou existenci toho, co vystupuje a zaniká v tematickém poli jako aktuální výtvar. Má také smysl bytí setrvalé platnosti, ba dokonce objektivní platnosti ve zvláštním smyslu, přesahující aktuální poznávající subjektivitu a její akty. Zůstává v opakování identickým, je znovu rozpoznáno ve způsobu jsoucího, které setrvává; má ve formě dokumentace objektivní existenci, stejně jako ostatní předmětnosti konkrétního světa: je tu v objektivním trvání pro každého, v tomtéž smyslu znovu pochopitelné, intersubjektivně identifikovatelné, existující, i když je nikdo nemyslí. (Husserl 1929, 50)

Je specifikem fenomenologie, že se u tohoto nezastavuje. Ať je jakkoli pro logiku a matematiku nutné takovéto pojetí svých objektů (založené na respektování jejich smyslu), filosofie těchto věd by si s tím podle Husserla vystačit neměla, protože věda, „která je přímo zaměřena na svou vlastní tematickou sféru a která rozvíjí svou aktivitu výlučně vzhledem k jejímu poznání, zůstává vězet v naivitě, která jí uzavírá filosofickou přednost radikálního sebepochopení a principiálního sebezduvodnění“ (Husserl 1929, 160).

Právě od této ideje se odvíjí Husserlův projekt transcendentální logiky, který je založen na přesunu intencionálního zaměření od objektů logiky k prožitkům, ve kterých se tyto objekty konstituují:

Opačný směr logické tematiky je směr subjektivní. Ten je zaměřen na hluboko skryté subjektivní formy, v nichž teoretický ‚rozum‘ uskutečňuje své výkony. Nejprve je zde předmětem tázání *rozum v jeho aktualitě*, totiž intencionalita, která probíhá v živoucím výkonu a v níž ony objektivní útvary mají svůj ‚původ‘. (Husserl 1929, 50)

Tím jsme narazili na paradoxně znějící aspekt celého myšlenkového postupu. To, co má smysl nezávislé existence, má zároveň původ v nějakém výkonu vědomí. K tomuto problému se ještě vrátíme v poslední části této práce.

3 Gödelův realismus

Husserlova fenomenologická filosofie logiky a matematiky zůstala bohužel více méně v podobě pouhého projektu. Tento projekt však velmi ovlivnil filosofické názory Kurta Gödela. Než vyložíme, jak se nechal Gödel fenomenologií inspirovat, musíme alespoň stručně nastínit problematiku základů matematiky.

Gödelův matematický realismus (podobně jako realismus Penelope Maddy) se vztahuje především na objekty teorie množin. To je motivováno tím, že existují dobré důvody považovat teorii množin (v poměrně jasném smyslu) za „ontologii matematiky“. Množiny nejsou žádné bezproblémové objekty. Russellův a další paradoxy ukázaly, že bychme-li udržet bezespornost, nevystačíme s jednoduchým intuitivním či „naivním“ pojetím množiny jako libovolného souboru objektů. Avšak ukázalo se, že předpokládáme-li pojem množiny jakožto daný, daří se redukovat na něj každý jiný matematický pojem. Jinými slovy, každý

matematický objekt lze modelovat jako určitou množinu. Pojem množiny je tedy natolik bohatý, že obsahuje veškerou rozmanitost matematických struktur. V důsledku lze každou matematickou větu přeložit do jazyka teorie množin.

Postupem času bylo navrženo pro teorii množin několik axiomatických systémů, z nichž nejvíce studovaná a používaná je tzv. Zermelo-Fraenkelova axiomatizace obohacená o axiom výběru² (značí se ZFC a bez axiomu výběru jen ZF). Úspěch a síla tohoto systému sestávajícího z několika málo přehledných formulí (resp. schémat) spočívá v tom, že v něm lze dokázat každou větu, která dosud byla v matematice dokázána.³ Existuje tedy jasný smysl, ve kterém lze tuto teorii považovat za věrnou formalizaci matematiky.⁴ Další výhodou je, že tento model (tj. ZFC) je sám matematickým objektem a lze o něm vynášet matematické výroky a dokazovat je. Např. lze o nějaké konkrétní formuli (která může být chápána jako překlad skutečného matematického tvrzení) ukázat, že není v tomto systému dokazatelná.

Zvláštní výzvu pro realismus představuje fakt, že existují určité matematické výroky, jejichž množinově teoretický korelát není v ZFC dokazatelný ani vyvratitelný (jinak řečeno, ZFC je neúplná teorie). A je to právě slavná Gödelova věta o neúplnosti, která říká, že situace nebude lepší pro žádný uzavřený axiomatický systém (je-li tento systém dostatečně silný a existuje-li algoritmus, na jehož základě lze rozhodnout, co je axiom tohoto systému a co ne). Jak lze očekávat, pro Gödela tento fakt neznamená, že nezávislé výroky nemají pravdivostní hodnotu, ani že tato hodnota je pro nás nedosažitelná. Podle Gödela význam pojmu množiny připouští otevřený proces přijímání nových a nových axiomů,

² Axiom výběru říká, že pro každou množinu x disjunktních množin existuje množina y , která obsahuje právě jeden prvek z každé množiny, která náleží x . Tento princip je „nekonstruktivní“ povahy, protože se velice často stává, že k dané množině x nejsme schopni množinu y zkonstruovat a na základě axiomu výběru přesto musíme předpokládat její existenci.

³ Toto tvrzení není samo matematickým tvrzením. Nelze tedy pro něj předvést striktní důkaz. V jeho prospěch však svědčí to, že je v souladu se zkušeností mnoha matematiků a nebyl k němu dosud nalezen žádný významný protipříklad. Můžeme tedy říct, že se jedná o tezi, která je stejně tak dobře podložena jako třeba darwinovská teorie evoluce.

⁴ Teorie množin nepředstavuje jediný možný způsob, jak vybudovat základy matematiky. Alternativou je především tzv. teorie kategorií. Viz např. Goldblatt (1979).

kteře vyjadřují další iterace operace „množina těch a těch prvků“ („set of ...“): „Axiomatický systém teorie množin, jak se dnes používá, je neúplný, ... může být doplněn bez libovůle novými axiomy, které jen rozvíjejí obsah pojmu množiny“ (Gödel 1947, 68-69).

Avšak na základě jaké metody máme tyto axiomy přijímat, když pojem množiny je základním pojmem, takže pro něj neexistuje žádná definice fixující jeho význam a ve stávajících axiomech je jeho význam uchopen jen částečně? Na základě čeho tedy lze objasňovat smysl pojmu množiny? Ve své době k tomu Gödel uvádí, že

dnes existují začátky jedné vědy, která tvrdí, že má systematickou metodu pro takové vysvětlování smyslu, totiž Husserlem založená fenomenologie. Vysvětlení smyslu zde spočívá v tom, že se ostřeji soustředíme na ty pojmy, jichž se to týká a to tím, že tuto pozornost jistým způsobem zaměřujeme, totiž na naše vlastní akty při používání těchto pojmů, na naše schopnosti tyto akty provádět atd. Přitom nám musí být jasné, že tato fenomenologie není vědou v témž smyslu jako vědy jiné. Spíše je postupem nebo technikou, která nás má uvést do nového stavu vědomí, v němž do podrobností zkoumáme námi používané základní pojmy našeho myšlení. (Gödel 1995, 209-210)

Níže uvidíme, že naturalismus Penelope Maddy představuje alternativní přístup k této problematice.

Gödelovo pojetí matematické zkušenosti se do značné míry podobá Husserlovu. Oba se domnívají, že tato zkušenost je analogická smyslové zkušenosti s materiálními objekty a oba na tom zakládají svůj realismus. Gödel říká, že

máme něco jako vnímání i objektů teorie množin, jak je vidět z faktu, že se nám axiomy vnucují samy jakožto pravdivé. Nevidím žádný důvod, proč bychom měli mít menší důvěru k tomuto druhu vjemů, tj. k matematické intuici, než k vnímání smyslovému, které nás vede k vytváření fyzikálních teorií a k očekávání, že další smyslové vjemy s nimi budou souhlasit, a navíc k víře, že otázka, která nyní není rozhodnutelná, má smysl a může být rozhodnuta v budoucnu. Paradoxy teorie množin jsou stěžejší překážkou pro matematiku, než smyslové klamy pro fyziku. (Gödel 1947, 80)

Fenomenologicky motivovaný realismus je však u Gödela doplněn něčím, čemu budeme říkat metodologicky motivovaný realismus. Ten analogii matematiky s přírodními vědami posouvá ještě dále. Podle

metodologicky motivovaného realismu matematické skutečnosti, které „vnímáme“, nemusí mít vždy povahu axiomů. My však podobně jako v přírodních vědách hledáme takovou teorii, jejíž důsledky jsou s těmito přímo ověřitelnými daty v souladu. Gödel naznačuje, že toto pojetí má zdroj u Russella:

Russell přirovnává axiomy logiky a matematiky k přírodním zákonům a logickou evidenci ke smyslovému vnímání, takže axiomy nepotřebují být nutně evidentní samy sebou, nýbrž jejich zdůvodnění spočívá spíše (právě tak jako ve fyzice) ve skutečnosti, že umožňují vyvození těchto „smyslových vjemů“; což by ovšem nevylučovalo, že také mají svého druhu vnitřní věrohodnost podobnou věrohodnosti fyziky. Myslím si ..., že následující vývoj dalekosáhle ospravedlnil tento názor. (Gödel 1944, 28)

Ilustrativním příkladem by zde mohla být třeba Bolzanova věta, podle které platí, že pokud funkce f , která je spojitá na intervalu $\langle a, b \rangle$, nabývá na a a b hodnoty s odlišnými znaménky, pak existuje bod c mezi body a a b , takový, že funkce f nabývá na c hodnotu 0. Toto tvrzení se zdá být zjevné, avšak v rámci teorie musí být odvozeno složitým důkazem z méně zřejmých principů. Bolzano sám se k této problematice vyjadřuje podobně jako Gödel (viz např. Bolzano 1810).

Ale Gödel jde ve svém „pragmatismu“ ještě dále. V matematice, podobně jako ve fyzice, lze někdy přijmout určité tvrzení čistě z toho důvodu, že činí teorii více elegantní:

I když odhlédneme od vnitřní nutnosti nějakého nového axiomu a dokonce i v případě, kdy nemá žádnou vnitřní nutnost, je možné pravděpodobné rozhodnutí o jeho pravdivosti i jiným způsobem, totiž induktivně studiem jeho „úspěšnosti“. Úspěšnost zde znamená plodnost na důsledky, zvláště na „verifikovatelné“ důsledky, tj. důsledky dokazatelné bez tohoto nového axiomu, jejichž důkazy pomocí nového axiomu jsou však podstatně jednodušší a jejichž nalezení je snadnější a které dovolují stáhnout do jednoho důkazu množství různých důkazů. ... Mohly by existovat axiomy s tak hojnými verifikovatelnými důsledky, vrhající tolik světla na celou oblast a dávající tak mocné metody na řešení problémů ..., že bez ohledu na to, zda jsou vnitřně nutné, musely by se přijmout aspoň v témže smyslu, jako jakákoli dobře zavedená fyzikální teorie. (Gödel 1947, 70)

Realismu gödelovského typu se často vytýká nejasnost nekauzálního vztahu mezi matematickými objekty a zkušeností, v níž je „vnímáme“. Na rozdíl od smyslového vnímání, nemáme k dispozici žádnou rozpracovanou teorii, která by popisovala, jak k takovému vnímání dochází, a která by byla včlenitelná do našeho celkového vědeckého obrazu světa. Tuto námitku zformuloval např. Benacerraf ve svém klasickém článku (1973) a jak uvidíme, právě v tomto ohledu se pokusí Penelope Maddy Gödelův realismus modifikovat.

4 Quinův realismus

Quinův realismus se velice odlišuje od Gödelova a my pouze stručně zformulujeme jeho hlavní myšlenku v té podobě, v jaké na ni navazuje Penelope Maddy v (1990). Podle Quinova holismu nekonfrontujeme se světem jednotlivé výroky, ale vždy jen celé teorie (viz Quine 1951). Naše ontologické závazky jsou určeny tím, přes jaké objekty musí probíhat proměnné těch teorií, které jsme přijali za platné (viz Quine 1948). Jinými slovy, měli bychom uzнат existenci právě těch objektů, které potřebujeme k formulaci našich nejefektivnějších vědeckých teorií představujících ty nejjednodušší pojmové rámce, do nichž jsme schopni vměstnat naši „syrovou“ zkušenost. Toto kritérium se vztahuje i na matematické objekty. Pokud k vyjádření některých vědeckých faktů (jako je například „molekula vody obsahuje tři atomy“) potřebujeme používat čísla a bez čísel je vyjádření těchto faktů nemožné, pak musíme uzнат, že čísla existují. Podobně je tomu s jinými matematickými objekty (jako např. s kontinuem), jichž používáme k vyjádření vědeckých faktů, i když tyto fakty mohou mít pochopitelně mnohem složitější podobu než v uvedeném příkladu.

Quinův naturalistický test pro existenci matematických objektů můžeme tedy formulovat jako ekvivalenci:

(QT) Daný matematický objekt existuje právě tehdy, když je nepostradatelný pro přírodní vědy.

Zvláštností Quinova realismu je, že je v kontrastu s běžným pojetím matematického vědění jakožto apriorního, nutného a nerevidovatelného. O tom, které matematické objekty existují a které nikoli, mohou rozhodovat čistě empirické důvody. Proto chápeme tuto argumentaci jako naturalistický způsob zdůvodnění realismu.

5 Realismus Penelope Maddy

Realismus Penelope Maddy kombinuje prvky Gödelova a Quineova přístupu ve snaze vyhnout se nedostatkům obou těchto variant realistické koncepce. Maddy vytýká Quinovu realismu, že není v souladu se skutečnou matematickou praxí: Matematikové v současné době nečekají na to, zda fyzika ospravedlní jejich aktivitu, neboť mají svoje kritéria a svoje metody ospravedlnění. Quinův realismus přiznává existenci pouze některým matematickým objektům – konkrétně těm, které jsou nezbytné pro přírodní vědy. Tak například transfinitní čísla – což jsou klíčové objekty teorie množin – by navrženým testem neprošla a celou vyšší teorii množin bychom pak museli považovat za pouhý „neinterpretovaný kalkul“, tj. pouhou hru s bezvýznamnými symboly. Jak si však potom vysvětlit úspěchy odborníků na teorii množin, nepochybnou smysluplnost jejich činnosti a především jejich neutuchající pocit nelibovůle v jejich disciplíně a toho, že skutečně hovoří „o něčem“ a „něco“ popisují?

Pro naše účely bude užitečné, když si Quinův test (QT) rozložíme na dvě implikace:

(QT1) Je-li daný matematický objekt nepostradatelný pro přírodní vědy, existuje.

(QT2) Je-li daný matematický objekt postradatelný pro přírodní vědy, neexistuje.

Maddy (1990) přijímá pouze polovinu Quinova testu, totiž tvrzení (QT1), které má zajistit existenci některých centrálních matematických objektů, např. kontinua. Avšak vzhledem k autonomii matematické praxe nemůže být fyzika konečným a jediným arbitrem existence matematických objektů. Maddy (1990) proto neakceptuje (QT2). Místo toho se domnívá, že aplikovatelnost částí matematiky ospravedlňuje realistický status matematické praxe jako celku. Pro obhajobu existence objektů neaplikované matematiky používá Maddy Gödelův metodologický způsob zdůvodnění.⁵

Gödelův realismus je v podstatě přesným opakem Quinova, neboť vychází důsledně z matematické praxe. Jeho nedostatkem podle Maddy (1990) je, že nezduvodňuje, proč bychom měli věty matematiky chápat tak, že se nějak vztahují ke skutečnosti.

⁵ Zdá se, že Quine dospívá ve své knize (1992) k podobnému stanovisku.

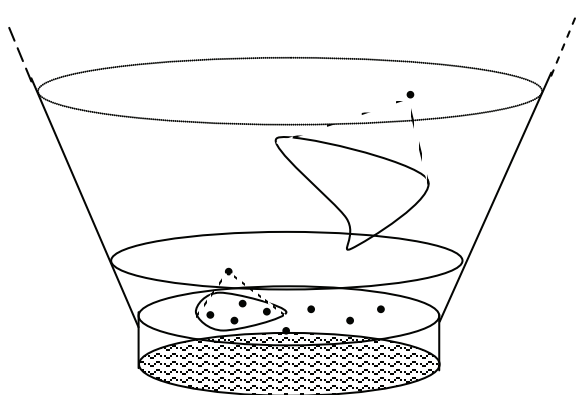
Pokud tedy shrneme a zjednodušíme syntézu Quinova a Gödelova realismu z Maddy (1990), dostaneme následující úvahu: Samotné metodologické úvahy nezajistí realistický status matematiky, můžeme však metodologicky zdůvodnit, že pokud existují některé centrální matematické objekty, řekněme kontinuum, pak existují i všechny ostatní matematické objekty, které naše nejlepší teorie kontinua vyžaduje, tedy např. i hierarchie transfinitních čísel. Tolik Gödel. (QT1) pak dodá antecedent této implikace, totiž tvrzení, že existují některé centrální matematické objekty, např. kontinuum. Důsledkem je, že musíme přijmout také existenci transfinitních čísel a jiných objektů neaplikované matematiky. Ve zkratce se dá tedy říct, že úsudek má formu: všechno nebo nic (Gödel), něco (Quine), tudíž všechno.

Fenomenologický způsob zdůvodnění realismu postuluje možnost nekauzálního „vnímání“ objektů, které jsou mimo prostor a čas a které tedy ani kauzálně působit nemohou. Maddy (1990) akceptuje Benacerraffovu kritiku, ale aby zachovala dvouvrstevnost vysvětlení matematické existence a tím udržela analogii matematiky s přírodními vědami, formuluje místo fenomenologického způsobu zdůvodnění něco, co bychom mohli nazvat psychologický způsob zdůvodnění. Podobně jako fenomenologický je i psychologický způsob zdůvodnění založen na jistém typu zkušenosti s matematickými objekty. Tato zkušenost však není pojata transcendentálně, jako u Husserla, a není to zkušenost objektů nějaké „třetí říše“ (viz Frege 1918). Místo toho se jedná o kauzálně způsobenou, fyziologicky fundovanou zkušenost s objekty tohoto světa.

Zvláštěností psychologického realismu tedy je, že matematické objekty v něm vystupují jako konkrétní entity v prostoru a čase, což je v rozporu s tradičním pojetím.

Maddy tvrdí, že doslova vnímáme vedle jednotlivých fyzických předmětů také množiny těchto předmětů, např. množinu tří vajíček v lednici (abychom užili její příklad, viz Maddy 1990, 58). Toto tvrzení podporuje některými psychologickými závěry Jeana Piageta a fyziologickou teorií Donalda Hebba. Přijmeme-li takovéto tvrzení, pak nemáme menší důvod věřit v existenci množin vajíček než v existenci vajíček samotných. Avšak množiny vajíček jsou matematické objekty. Přímou tedy vnímáme jisté matematické objekty jakožto součást světa.

Celý obraz fyzikálního světa, tak jak ho vidí Maddy v roce 1990, lze ilustrovat následovně:



Spodní vyplněný ovál zde reprezentuje „beztvarou hmotu“, nad ním je ovál obsahující hmotné předměty a z toho v jednotlivých vrstvách vyrůstá hierarchie množin „generovaná“ operací potence. Prvky množin tvoří na nejnižší úrovni hmotné předměty a na vyšších úrovních jiné množiny.

Na Chiharovu námitku, že perceptuálně neodlišíme hmotný předmět x od singletonu $\{x\}$ (tj. množiny obsahující pouze tento předmět), odpovídá Maddy, že nejlépe uděláme, když x a $\{x\}$ ztotožníme. Není to sice jediné možné řešení, je však obzvláště lákavé, protože poté se fyzikální objekty a matematické předměty dostanou do skutečné symbiózy: Platí-li pro materiální předměty $x=\{x\}$, pak vůbec každý konkrétní předmět je množinou, tedy matematickým předmětem, takže pokud ze světa odstraníme matematické předměty, zbude nám pouze „beztvará hmota“; a obráceně: protože každý reálný matematický předmět je fundován v hmotě, tak kdybychom odstranili ze světa hmotu, nezbyl by nám ani žádný matematický předmět. Hmotě nelze přiřadit číslo. Číslo lze přiřadit vždy jen množinám. Proto hmotu s množinami ztotožnit nelze. Avšak jednotlivým předmětům přiřadit číslo lze – a to vždy číslo jedna stejně jako singletonům těchto předmětů. Uvedenému ztotožnění tedy nic nebrání.⁶

⁶ Takovéto ztotožnění však nelze aplikovat na množiny, které nejsou materiálními předměty. Např. množinu $\{x, y\}$ nelze ztotožnit s množinou $\{\{x, y\}\}$, neboť první má dva prvky a druhá jeden.

Otázkou je, kde se v tomto obrazu nachází prázdná množina. Maddy argumentuje, že bez prázdné množiny se lze konec konců obejít a její existenci nemusíme předpokládat. Lze ji považovat za užitečnou fikci zjednodušující technickou práci. Ordinalní čísla, která jsou na prázdné množině vystavěna a na nichž zase závisí mnoho dalších matematických objektů, lze zkonstruovat i bez prázdné množiny. Množinově teoretický korelát posloupnosti přirozených čísel $(0, 1, 2, 3, 4, \dots)$ tvoří posloupnost

$$\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}, \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}\}, \dots$$

Tu lze nahradit např. řadou využívající materiální objekty x a y :

$$\{x\}, \{x, y\}, \{x, y, \{x, y\}\}, \{x, y, \{x, y\}, \{x, y, \{x, y\}\}\}, \{x, y, \{x, y\}, \{x, y, \{x, y\}\}, \{x, y, \{x, y\}, \{x, y, \{x, y\}\}\}\}, \dots$$

Maddy přijímá předpoklad běžně spojovaný s realismem, že pokud přijmeme, že existuje takto fixně daný svět matematických struktur (reprezentovaných množinami), pak každé tvrzení v jazyce teorie množin by mělo mít v této hierarchii množin jednoznačně určenou pravdivostní hodnotu. Tedy také každé tvrzení, které je nezávislé na ZFC, by mělo být buď pravdivé, nebo nepravdivé. Gödelovy metodologické úvahy naznačují, že existují způsoby argumentace zdůvodňující pravdivost resp. nepravdivost nezávislých matematických výroků. Tyto argumenty musí mít nedemonstrativní povahu v tom smyslu, že se nejedná o běžné matematické důkazy. Maddy (1990) považuje za největší výzvu pro svůj realismus popis, vysvětlení a ohodnocení takových možných argumentů.

Později, při snaze rozpracovat právě tuto klíčovou otázku, naráží Maddy na jisté problémy se svou koncepcí. Na jejich základě postupně realismus opustí a svůj nový přístup nazve naturalismem. Důvody, které Maddy vedou k odmítnutí svého realismu, si přiblížíme v následující části tohoto textu. K těmto důvodům se přidává důkladná kritika jejího realismu formulovaná v Lavine (1992), Balaguer (1994), Riskin (1994) a Carson (1996). Tato kritika, kterou Maddy v podstatě akceptuje, je shrnuta a rozebrána v Kvasz (2010).

6 Naturalismus Penelope Maddy

Maddy postupem času začala pochybovat o přijatelnosti principu (QT1), který tvořil významný pilíř jejího realismu. S odmítnutím tohoto principu realismus opouští.

Maddy ilustruje na příkladu atomu, proč nepostradatelnost nějakého konceptu nemusí nutně vést k přijetí jeho existence:

Do roku 1900 překypovala atomická teorie všemi pěti teoretickými přednostmi [které zformuloval Quine, jsou jimi jednoduchost, známost principů, rozsah, plodnost a testovatelnost důsledků]; její síla a užitečnost se stávala zřejmější a zřejmější s každým experimentálním a pojmovým krokem vpřed ... Navzdory těmto přednostem atomické teorie se vědci neshodovali ohledně reality atomů. (Maddy 1997, 138)

Přes nepostradatelnost atomů mnozí považovali atomy za užitečné fikce – a to až do té doby než Einstein uskutečnil jisté experimenty, v nichž Brownův pohyb částic v tekutině vedl k bezprostředním, mikroskopicky pozorovatelným efektům.

Zdá se tedy, že nepostradatelnost není postačující podmínkou existence. Na tomto základě začala Maddy pochybovat také o tom, zda se kontinuum v aplikacích používá v doslovném smyslu či jako idealizace – podobně jako se předpokládá při výpočtech trajektorií, že povrch daného úseku země je rovný či když vědci při analýze vln na povrchu oceánu předpokládají, že oceán je nekonečně hluboký (k tomu viz Maddy 1990, 143-152).

Maddy v (1990) odmítla (QT2), neboť není v souladu s matematickou praxí. Později v (1997) ale dospěla k závěru, že stejný problém platí i pro (QT1), i když v méně zřejmé podobě. Vraťme se k tématu nezávislých tvrzení teorie množin. Pokud přijmeme realismus založený na (QT1), pak otázka, zda mají nezávislá tvrzení teorie množin jednoznačně určenou pravdivostní hodnotu, závisí na dosud nerozhodnuté otázce týkající se doslovných aplikací kontinua ve fyzice. Pokud jsou všechny aplikace kontinua ve fyzice pouhé idealizace, nemáme dostatek důvodů přijmout do naší ontologie kontinuum spolu s celým množinově teoretickým rámcem, v němž je formulována jeho nejlepší teorie. Pak ale nemáme ani důvody domnívat se, že pravdivostní hodnota nerozhodnutelných tvrzení je jednoznačně určena. Pokud však potřebujeme doslovné aplikace kontinua, důsledkem (QT1) a na něm založeném realismu je, že hodnota těchto tvrzení jednoznačně určena je.

Pak by se ale zdálo, že odborníci na teorii množin budou s napětím sledovat vývoj některých fyzikálních teorií a v závislosti na něm se budou nebo nebudou pokoušet hledat nedemonstrativní argumenty ve prospěch té či oné pravdivostní hodnoty nezávislých tvrzení. To

se však neděje. K pokusům rozhodnout tato tvrzení dochází zcela bez ohledu na aktuální praxi fyziky. Role kontinua ve fyzikálních teoriích nemůže mít na status nezávislých tvrzení žádný vliv.

Dalším argumentem, který vedl Maddy k opuštění množinového realismu je, že při bližším pohledu na praxi teorie množin je významným způsobem zpochybněna její analogie s přírodními vědami. Můžeme říci, že pro přírodní vědy platí do jisté míry princip Ockhamovy britvy, tj. princip ekonomie teorie, podle něhož nezmnožujeme teoretické entity nad nezbytnou míru. Naopak teorie množin se snaží poskytnout co nejširší rámec, v němž lze při zachování konzistence rekonstruovat vše, co se dá smysluplně nazvat matematickou strukturou. V přírodních vědách se tedy snažíme o minimalizaci, kdežto v teorii množin naopak o maximalizaci ontologie. K tomuto problému se ještě vrátíme. Nyní však vyložíme hlavní ideu naturalismu Penelope Maddy z (1997).

Označení „naturalismus“ může být zprvu poněkud matoucí. Chápeme-li pod matematickým naturalismem „fyzikalizující“ a „biologizující“ tendence ve filosofii matematiky, pak je naturalismus Penelope Maddy méně naturalistický než její realismus. Je to právě realismus Penelope Maddy, který se snaží matematické objekty umístit do fyzikálního světa, jejich existenci obhájit s využitím (QT1) na základě podoby současné fyziky a zároveň podepřít fyziologickými a psychologickými teoriemi. Naturalismus Penelope Maddy naopak toto vše nechává stranou a obrací se pouze k matematické praxi, která se stává jedinou autoritou v takových otázkách, jako je problém pravdivostní hodnoty nezávislých tvrzení či sporných metod, např. uplatňování axiomu výběru či impredikativních definic⁷. Termín „naturalismus“ je převzat od Quina, jehož naturalistický názor spočívá v odmítnutí ideje Descartovské první filosofie, která má sloužit jako poslední arbitr poskytující vědě externí obhajobu, případně z externích pozic vědu korigovat. Podle naturalismu nám věda předkládá nejlepší obrázek reality, který máme k dispozici. Tento obrázek sice je korigovatelný, ale nikoli nějakými mimovědeckými prostředky. Naturalismus je tedy přesvědčení, že filosofické či metodologické otázky týkající se vědy musí být – pokud mají smysl – rozhodnuty ve vědě samotné a nikoli na základě jakýchkoli externích filosofických úvah. To je také smysl slavného Neurathova vý-

⁷ Impredikativní definice definují nějaký objekt pomocí odkazu na celek, v němž se tento objekt již nachází. Proto mají nekonstruktivní charakter.

roku, že věda je jako loď, kterou musíme přestavovat průběžně během plavby.

Tento respekt k vědecké praxi kontrastuje s Quinovým testem, neboť – jak jsme již uvedli – ten matematickou praxi nerespektuje. Matematický naturalismus formulovaný v Maddy (1997) chápe matematiku jako nezávislou na jakýchkoli externích filosofických či vědeckých principech a respektuje stejnou měrou tuto autonomii matematické praxe, jako Quinův naturalismus respektuje autonomii praxe přírodních věd. Je to tedy v jistém smyslu pravý opak „fyzikalizujících“ a „biologizujících“ tendencí ve filosofii matematiky.

Otázkou nyní je, do jaké míry zůstává realismus problémem, který lze v takto pojatém naturalistickém rámci řešit. Maddy je v tomto ohledu skeptická. Domnívá se, že zatímco fyzikální teorie nám říkájí mnoho o povaze existence fyzických objektů, např. že jsou to časoprostorové objektivně existující entity, matematika o těchto otázkách (vztahených na matematické objekty) mlčí.

Je zde potřeba vytyčit hranice mezi tím, co je naturalizovatelná filosofie a tím, co je – slovy Wittgensteina – pouhá próza čili první filosofie. Vytyčení této hranice je obtížný úkol, jehož řešení Maddy nepředkládá. Ukazuje pouze, že existují případy, u nichž je poměrně zřejmé, na jakou stranu hranice spadají. Jako příklad postoje, který spadá spíše do první filosofie než do vědy, uvádí Maddy verifikacionismus, tj. názor, že význam věty je dán podmínkami, za nichž lze tuto větu verifikovat. Do protikladu staví mechanismus, tj. názor, že všechny fenomény lze chápat jako výsledek působení mechanických sil. Mechanismus je filosofický postoj, který vznikl na základě vědecké praxe a který motivoval plodný vědecký projekt, jenž vedl až k nalezení důvodů pro zamítnutí tohoto postoje. Oproti tomu verifikacionismus povstal ve filosofickém kontextu a k žádnému plodnému projektu nevedl. Maddy se přiklání k názoru, že realismus je v tomto ohledu podobný spíše verifikacionismu než mechanismu.

Maddy však svůj naturalismus považuje spíše za metodu než za tezi (viz Maddy 1997, 200). Nechce tedy tvrdit, že realistické filosofické úvahy jsou zcela irelevantní. Spíše motivována jistými historickými důvody nechává tyto úvahy stranou a systematicky na ně nebere ohled při řešení metodologických otázek.

Co je tedy vlastním problémem, který Maddy pod hlavičkou naturalismu řeší? Jde především o naturalistické posouzení otázek teorie množin, které nemohou být řešeny čistě v rámci etablovaných matema-

tických teorií pomocí důkazu. Sem spadají především otázky posouzení nových kandidátů na axiomy, s jejichž pomocí lze rozhodnout některá klasická nezávislá tvrzení jako je např. hypotéza kontinua.⁸ Jedná se tedy o problém, který Maddy vytyčila již v (1990), ale nyní ho chce řešit čistě naturalistickými prostředky, tj. s vyloučením jakýchkoli externích motivací. Tyto prostředky spočívají ve vytvoření jednoduchého modelu matematické praxe. Model neobsahuje nic, co by bylo metodologicky irelevantní. To, zda je tento model adekvátní, je empirická otázka, která se testuje konfrontací modelu s reálnými historickými událostmi v matematice. Kandidáti na axiomy mají být posouzeni na základě tohoto modelu.

Model matematické praxe vychází z pozorování, že vždy když dochází k modifikaci stávající matematické praxe, především když dochází k přijetí nových prostředků, které byly dříve považovány za kontroverzní (např. axiom výběru), opakuje se jistý vzorec chování matematické komunity. Kontroverzní metody jsou postupně přijaty z toho důvodu, že vedou k takovým podobám matematiky, které představují efektivní prostředky k dosažení žádoucích cílů. Model matematické praxe (a v případě Penelope Maddy konkrétněji model množinově teoretické praxe) musí tedy sestávat z identifikace žádoucích cílů a z následného ohodnocení potenciálních metod a prostředků z hlediska jejich efektivity pro dosažení těchto cílů.

Jaké jsou tedy specifické cíle teorie množin, podle nichž bychom mohli poměřovat různé nové kandidáty na axiomy? Maddy identifikuje jeden základní cíl, jehož smysl jsme naznačili výše. Je jím vytvoření základů pro celek matematiky. (Přitom není řečeno, že nelze vybudovat základy matematiky ještě jiným způsobem, než jak se to praktikuje v teorii množin.) Z tohoto cíle vyplývají dvě dílčí maximy:

- (M1) Teorie množin by měla být jednotná.
- (M2) Množinově teoretická oblast by měla obsahovat maximální možnou rozmanitost matematických struktur.

⁸ Hypotézu kontinua lze formulovat jako tvrzení, že každá nekonečná podmnožina kontinua má buď mohutnost celého kontinua, nebo má mohutnost množiny přirozených čísel (ta má mohutnost menší než kontinuum). Jinými slovy, neexistuje žádná střední mohutnost mezi mohutností množiny přirozených čísel a mohutností kontinua. Otázka, zda je hypotéza kontinua pravdivé tvrzení, se vynořila hned v počátcích teorie množin. Později se však ukázalo, že se jedná o nezávislé tvrzení.

Mezi maximami (M1) a (M2) existuje určité pnutí. Řekněme, že posuzujeme dvě alternativní rozšíření ZFC. Nejjednodušším způsobem, jak postupovat ve shodě s (M2), by bylo přijmout obě rozšíření jako dvě alternativy, vymezující třeba dva různé pojmy množiny. Avšak (M1) koriguje tento postup a nutí nás k tomu, abychom mezi alternativami vybrali jen jednu. Z naturalistického hlediska je to tedy především maxima (M1), která stimuluje hledání a hodnocení nových kandidátů na axiomy a motivuje nás k řešení otázky rozhodnutelnosti nezávislých tvrzení.

Jak může naturalistické posuzování nových kandidátů vypadat? Penelope Maddy to ilustruje na příkladu axiomu konstruovatelnosti množin. Tento axiom je motivován russelskou představou, podle které množiny odpovídají propozičním funkcím, čili množina je něčím, co je určeno nějakou formulí s volnou proměnnou. Formulaci axiomu konstruovatelnosti předchází definice třídy L . Třída L se „konstruuje“ transfinitní rekurzí podle osy ordinálních čísel a každému ordinálu se v této konstrukci přiřadí množina definovatelných množin, jejichž prvky se nacházejí v předchozích úrovních. Konstrukce se tedy podobá standardní konstrukci „kumulativní hierarchie“ množin, avšak místo množiny všech podmnožin prvků předchozích úrovní sestává každá úroveň z množiny všech podmnožin (prvků z předchozích úrovní), které jsou definovatelné nějakou formulí. Axiom konstruovatelnosti pak říká, že každá množina je v tomto smyslu definovatelná, tj. že třída L tvoří univerzum všech množin označované standardně jako V . Stručně se toto tvrzení zapisuje jako $V=L$. Tento axiom zavedl Kurt Gödel a s jeho pomocí ukázal, že jak axiom výběru, tak hypotéza kontinua je relativně konzistentní vůči ZF, čili, jinými slovy, že negace těchto tvrzení nejsou v ZF dokazatelné (za předpokladu konzistence ZF). To se ukáže tak, že se sestrojí model teorie ZF obohacené o axiom konstruovatelnosti a ukáže se, že v něm platí obě tato tvrzení, tj. axiom výběru i hypotéza kontinua.⁹

⁹ Nechceme-li zabíhat do technických detailů, nemůžeme se vyjadřovat zcela přesně. Avšak mohlo by být užitečné alespoň stručně objasnit následující matoucí bod. Z druhé Gödelovy věty o neúplnosti plyne, že kdybychom předvedli přímou konstrukci modelu ZF, znamenalo by to, že ZF je nekonzistentní (konstrukci modelu by totiž šlo formalizovat v ZF samotné a ZF by tak dokazovala svoji konzistenci, což konzistentní teorie nemůže). Model ZF obohacené o axiom konstruovatelnosti se tedy nepopisuje přímo, ale na základě nezbytného předpokladu konzistence teorie ZF. Tento předpoklad zaručí existenci nějakého modelu (který však zkonstruovat

Sám axiom konstruovatelnosti je nezávislé tvrzení ZFC. Lze proti němu argumentovat z realistických pozic, že není důvod, aby univerzum všech množin, které existuje nezávisle na našem jazyku a naší matematické aktivitě, sestávalo jen z takových množin, které můžeme definovat. Avšak tento způsob argumentace by byl založen na externích důvodech, kterých se Maddy v (1997) snaží cíleně vyvarovat.

Maddy nabízí alternativní naturalistický způsob zdůvodnění, proč nepřijmout axiom konstruovatelnosti mezi základní množinové teoretické axiomy. Postup je takový, že ukážeme jasný smysl, v jakém je axiom konstruovatelnosti restriktivní a tedy v rozporu s maximou (M2). To se zprvu může jevit jako triviální tvrzení, neboť je zřejmé, že univerzum sestávající pouze z definovatelných množin je chudší než univerzum obsahující i nedefinovatelné množiny. Avšak jak Maddy v (1997) zdůrazňuje, situace není tak jednoduchá. Podobně by se mohlo zdát, že axiom fundovanosti, který zamezuje existenci množin, u nichž jsou možné nekonečné sestupy nejprve od celé množiny k jejím prvkům, poté k prvkům těchto prvků, poté k prvkům prvků prvků atd., podobným způsobem omezuje rozmanitost množinového univerza. Zdálo by se tedy, že Aczelova teorie množin AFA, která neobsahuje axiom fundovanosti a místo něj obsahuje axiom implikující existenci „nefundovaných“ množin, by měla mít z hlediska (M2) přednost před ZFC. Nefundované množiny sice představují nový typ množiny, nepředstavují však nový typ struktury, který by nebyl ničím nahraditelný v ZFC. Modely AFA jsou konstruovány uvnitř modelů ZFC a AFA je v ZFC interpretovatelná. AFA tedy není z hlediska (M2) přijatelnější než ZFC, neboť jejím přijetím nezískáme nic nového, co by nešlo v ZFC simulovat.

Avšak podobný pokus o obhajobu axiomu konstruovatelnosti není úspěšný. Konkrétně Maddy srovnává rozšíření ZFC o axiom konstruovatelnosti s rozšířením ZFC o formuli „existuje $0^{\#}$ “ tvrdící existenci

neumíme). V tomto modelu, ať už vypadá jakkoli, je fixována osa ordinálních čísel, podle které lze rekonstruovat třídu L . „Vyseparujeme-li“ pak z modelu takto získané L , dostáváme nový model ZF, v němž platí navíc $V=L$. Lze dokázat, že v tomto modelu dále platí hypotéza kontinua i axiom výběru. V praxi se lze vyhnout „realisticky“ znějící řeči o modelech tím, že ji nahradíme syntaktickou úvahou za pomoci pojmu interpretovatelnosti jedné teorie v druhé. Dalším bodem, který by si zasloužil objasnění, je fakt, že definovatelnost množin není v L tak důsledně aplikována, neboť není požadováno, aby sama osa ordinálních čísel obsahovala jen definovatelné ordinály. Na to upozorňuje i sama Maddy v článku (1993).

určitého „nekonstruovatelného“ objektu.¹⁰ Tato dvě rozšíření jsou vzájemně nekonzistentní, tj. nemůžeme přidat oba axiomy zároveň. (M1) nás motivuje k pokusům vybrat jednu z těchto dvou variant rozšíření, i když jde v obou případech pochopitelně o nezávislá tvrzení. Ukazuje se, že můžeme interpretovat teorii $ZFC+“V=L”$ v teorii $ZFC+“$ existuje $0^{##}$, což znamená, že veškerá matematika realizovatelná v $ZFC+“V=L”$ je v „ $ZFC+“$ existuje $0^{##}$ “ zachována, stačí pouze relativizovat kvantifikátory vzhledem k L . Avšak $ZFC+“$ existuje $0^{##}$ “ v $ZFC+“V=L”$ (ani v žádném jejím konzistentním rozšíření) interpretovatelná není. To znamená, že objekt $0^{##}$ představuje typ struktury, který nelze simulovat ničím v $ZFC+“V=L”$ a na druhou stranu v $ZFC+“V=L”$ neexistuje nic, co by nebylo nijak přítomno v $ZFC+“$ existuje $0^{##}$ “. Vzhledem k (M2) bychom tedy měli dát přednost teorii $ZFC+“$ existuje $0^{##}$ “.

Tímto příkladem jsme ilustrovali, jak vypadá naturalistický přístup k problému nezávislých vět. V poslední části této práce budeme argumentovat, že nejenže naturalismus Penelope Maddy není v rozporu s matematickým realismem, naopak představuje jeho silnou podporu.

7 Realismus motivovaný naturalismem Penelope Maddy

Jak jsme již uvedli, Maddy v (1997) nepovažuje otázku realismu za filosofický problém, který by byl v dostatečném kontaktu s aktuální matematickou praxí a rezignuje tedy na pokusy o jeho řešení. Mezi „filosofická“ tvrzení, která vyrostla z matematické praxe a dále na ni také působí, by zřejmě zařadila např. Churchovu tezi, podle níž každý z dosud navržených vzájemně ekvivalentních matematických pojmů algoritmu (tj. Turingovy stroje, rekurzivní funkce, vývojové diagramy atd.) je adekvátním zachycením přirozeného pojmu algoritmu či výše uvedené tezi, že každý matematický objekt lze modelovat jako množinu.

Avšak Maddy v (1997) netvrdí, že otázka realismu je nesmyslná a že bychom ji tedy vůbec neměli klást. Domnívám se, že pokud si ji položíme, je to právě naturalistický přístup, který představuje vhodnou metodu pro jedno z možných řešení. Uvádíme „jedno z možných řešení“, protože prvotní vymezení otázky (jaké jsme např. uvedli na začátku textu) je natolik vágní, že připouští natolik různá zpřesnění, že pro ně

¹⁰ Jedná se o určitou množinu přirozených čísel, jejíž existence není dokazatelná v ZFC.

dostáváme zcela opačné odpovědi. Dále se o jedno takové zpřesnění pokusíme a budeme zároveň argumentovat, že má jisté přednosti před alternativními způsoby zpřesnění. Vedle naturalismu Penelope Maddy se také necháme volně inspirovat Husserlovým přístupem s tím, že přeneseme problematiku ze zkušenostní do jazykové roviny.

Typicky lze vytýkat argumentům opírajícím se o fenomenologické aspekty naší zkušenosti, že pracují s něčím, co není intersubjektivně dostupné a je tudíž pro dialog a argumentaci nepoužitelné. Domnívám se však, že zdůvodnění toho, že prožitky nejsou intersubjektivně dostupné, závisí také na poukazu k určitému typu zkušenosti, neboť není pravda, že nelze hovořit o prožitku nějaké cizí osoby – a to o numericky tomtéž prožitku, o kterém hovoří tato osoba sama. Nelze však cizí prožitek vnímat, tj. nelze mít stejný typ zkušenosti s cizím prožitkem jako s nějakým materiálním objektem. Toto odmítnutí intersubjektivní nedostupnosti prožitků má tedy paradoxně fenomenologickou povahu – má v podstatě stejný charakter jako Husserlova a Gödelova argumentace ve prospěch realismu.

Nicméně a posteriori se zdá – srovnáme-li např. (ne)úspěšnost fenomenologického přístupu s takovým, který analyzuje matematiku na základě studia jejího jazyka –, že prožitky nejsou příliš vhodným typem objektů, na kterém by se dala bezpečně zakládat stabilní filosofie matematiky. Avšak je zajímavé povšimnout si, že fenomenologický přístup lze posunout do jazykové roviny. V té pak analogie mezi matematickou a perceptivní zkušeností začíná být problematická. V běžné řeči totiž často slouží perceptuální slovník k formulaci důvodů pro existenci nějakých materiálních objektů. Můžeme si např. představit dialog, v němž se osoba A ptá: „Proč si myslíte, že vrah měl komplice?“ Možnou odpovědí mající silnou důvodovou váhu by byla odpověď osoby B: „Svědék *zahlédl* utíkat dva muže z místa činu.“ Tím zdůvodňování může končit – dodatečné důvody často předložit nelze a také není potřeba. Je však pochybné tvrdit, že je podobný způsob zdůvodňování přípustný také v matematických textech. Sice se setkáváme s formulacemi jako „lze snadno nahlédnout, že ...“, či „to je zjevné“, ale tyto výrazy nikdy neslouží jako zdůvodnění daného tvrzení, jedná se pouze o strategickou zkratku v textu, která by měla být vždy nahraditelná rozvinutým vysvětlením. Přitom platnost axiomů, u nichž běžná argumentace končí, v současné době bývá málokdy dále zdůvodňována jejich „zjevností“, ale chápou se spíše jako „implicitní definice“, jejichž zavedení může být plně obhájeno jedině na základě komplexních teoretických důvodů (srov. Maddy 1988).

Avšak i kdybychom nepřijali, že matematická zkušenost je analogická smyslové zkušenosti (což na jazykové rovině odpovídá nepřijetí tvrzení, že slovník popisující matematickou zkušenost funguje stejně jako slovník popisující smyslovou zkušenost), stále ještě můžeme využít strukturu Husserlovy argumentace.

Základním východiskem fenomenologie je, že každý předmět má svůj „mysl“, který se musel vytvořit v prožitkovém proudu. Fenomenologie považuje za protismyslné předpokládat absolutní mysl, který vstoupil do vědomí zvenku. Vědomí je sféra, do které nemůže nic vstupovat. Naopak ale – ono samo jakoby ze sebe vystupuje tím, že konstituuje mysl vnějších předmětností, které ve vědomí nejsou. Tak např. i „svět sám má své celé bytí jako jistý „mysl“, který předpokládá absolutní vědomí jako pole udílení smyslu“ (Husserl 1913, 115). Co se např. čísel týče, i jejich mysl musel být konstituován ve vědomí – jak jinak bychom o nich mohli uvažovat než tak, že jsme si předtím sami, uvnitř svého vědomí jejich mysl zkonstitovali? „Cožpak ale nejsou, ptáme se, čísla tím, čím jsou, ať již je „vytváříme“, nebo nevytváříme?“ (Husserl 1913, 54). A čísla jsou podle svého smyslu právě něčím nezávislým, věčným, bezčasovým a bezprostorovým, transcendujícím prožitky, ve kterých byly stvořeny.

Shrňme ještě jednou, hlavní kroky Husserlovy argumentace:

Primárními objekty fenomenologie jsou prožitky. Ostatní objekty jsou zkoumány jen tak a natolik, jak a nakolik se nám „dávají v prožitcích“. Zajímá nás status objektů matematiky. Ty se nám dávají jakožto nezávislé na naší zkušenosti. Proto je na místě být realistou, neboť jinak bychom se provinili proti smyslu matematických objektů, jak jsme ho konstitovali ve svém vědomí.

Tím se fenomenologie dostává do zdánlivě paradoxní situace. Tvrdí, že mysl matematických objektů jsme konstitovali ve vědomí, avšak konstitovali jsme ho tak, že právě na základě tohoto smyslu je nutné považovat matematické objekty za nezávisle existující entity.

Domnívám se, že paradox v tomto postupu je pouze zdánlivý. Pouze se zde jeden problém nahlíží z různých perspektiv. Podobně se válec jeví z jedné perspektivy jako kruh a z jiné jako obdélník – na tom není nic paradoxního.

Pokusme se přetransformovat tuto fenomenologickou argumentaci do jazykové roviny:

Primárním objektem studia je jazyk. Ostatní objekty jsou zkoumány jen tak a natolik, jak a nakolik se nám „dávají v jazyce“. Zajímá nás status objektů matematiky. Ty v jazyce vystupují jakožto nezávislé na jazyce. Proto je na místě být realistou, neboť jinak bychom se provinili proti smyslu matematických objektů, jak je konstituován jazykem matematiky.

Tento modifikovaný argument má velkou výhodu. Lze mnohem lépe doložit, že matematické objekty vystupují v jazyce matematiky jako nezávislé na jazyce, než že matematické objekty vystupují ve vědomí jako nezávislé na vědomí. V protikladu k „apriorním“ argumentům pro fenomenologickou verzi jsou pro zdůvodnění jazykové verze vhodné právě naturalistické argumenty Penelope Maddy. Například její naturalistickou „demonstraci“, že V se nerovná L , můžeme pojmout jako evidenci ve prospěch tvrzení, že status množin v jazykové praxi teorie množin je „realistický“, tj. předpokládáme-li existenci množin, které nejsou definovatelné žádnou formulí, jednáme v souladu s matematickou praxí – a předpoklad opaku se prokazatelně dostává do konfliktu s cíli této praxe.

Tento způsob argumentace má naturalistickou povahu z toho důvodu, že zcela respektuje podobu matematické praxe a nemá ambice do ní nijak zasahovat. V tom se také liší od Husserlovy fenomenologie, která měla fungovat jako nezávislá apriorní disciplína čerpající své principy z vlastních zdrojů a poskytující základy a oprávnění všem ostatním vědám. To je přesně descartovský ideál první filosofie. Jak jsme viděli, naturalistický přístup – který Maddy ve své poslední knize nazývá druhou filosofií (viz Maddy 2007) – vznikl právě v opozici k tomuto ideálu. Naturalisticky motivovaný realismus tedy není rozhodnut apriori bez ohledu na faktickou podobu matematické praxe a nechce matematické předepisovat žádné principy (např. tak, že by nás realistické důvody vedly k přijetí principu vyloučeného třetího), nechce do matematiky nijak zasahovat, ale chce z ní čerpat pouze evidenci ve svůj prospěch. Realistická teze je sice zprvu vágní, ale je natolik smysluplná, že můžeme určit, které podoby matematiky ji podporují a které nikoli. Např. přijetí axiomu výběru, přijetí impredikativních definic a nekonstruktivních metod svědčí ve prospěch realismu a jejich nepřijetí svědčí ve prospěch pozic s realismem neslučitelných. To, zda máme tyto principy či metody přijmout, nemůže být rozhodnuto na základě nějaké filosofické pozice (např. realismu samotného), ale jsou-li tyto otázky na základě interních matematických důvodů rozhodnuty, můžeme to přijmout

jako evidenci svědčící ve prospěch či proti realismu. Jak sama Maddy podotýká, „mnoho těchto metodologických debat bylo rozhodnuto: impredikativní definice jsou povoleny, nepožaduje se, aby důkazy existence definovaly či konstruovaly entitu, jejíž existenci tvrdíme; axiom výběru byl přijat“ (Maddy 1997, 191). Avšak Maddy – v kontrastu s tím, co zde hájíme my – není ochotná na základě podoby matematické praxe činit realistické závěry. Říká, že „někteří mohou argumentovat, že úspěch těchto různých ‚platonisticky inspirovaných‘ metod lze pokládat za evidenci pro matematický realismus, ale naše naturalistická analýza tento názor nepodporuje“ (Maddy 1997, 193). Důvodem pro odmítavý postoj je, že podle Maddy realismus není teze, která by byla v dostatečném kontaktu s vědeckou praxí.

Hlavní závěr naší práce je v podstatě tento: Pokud lze problém realismu nějak smysluplně specifikovat, pak nejlépe jako otázku, jaký status mají matematické objekty v matematické praxi samotné. Tuto otázku umožňuje fakt, že matematika je schopna reflektovat svůj vlastní jazyk a tak se některé problémy, jako např. zda objekty matematiky existují nezávisle na jazyce matematiky a na matematických konstrukcích, dají formulovat jako interní matematické problémy. Z tohoto hlediska je ovšem velice relevantní, jak byly rozhodnuty výše zmíněné metodologické otázky. A byly rozhodnuty vesměs ve prospěch realismu.

Abychom otestovali alternativní stanoviska, můžeme zvážit následující tvrzení:

(T) Matematické objekty existují v závislosti na matematické praxi.

Co by se stalo, kdybychom integrovali tvrzení (T) do matematické praxe samotné jako její interní princip? Byly realizovány vážné pokusy toto učinit. Ty vedly k různým formám „konstruktivní matematiky“, která byla nejlépe rozpracována v díle L. E. J. Brouwera a jeho žáků, především A. Heytinga. Výsledná (tzv. intuicionistická) matematika založená na uvedeném principu se velmi odlišuje od klasické matematiky.¹¹

¹¹ Rozdíl je nejvíce patrný, když se podíváme, jaký charakter má intuicionistické kontinuum. Brouwer byl schopný ve své matematice dokázat, že každá totální funkce na uzavřeném intervalu je spojitá (viz např. Heyting 1966, 46 či Kolman 2008, 7. kapitola). To je v klasické matematice triviálně nepravdivé tvrzení. Snaha reformovat klasický pojem kontinua na filosofickém základě představovala pro intuicionismus jeden z hlavních cílů. Maddy tuto snahu komentuje již v roce 1990 takto: „Teorie reálných čísel tvoří základní součást kalkulu a vyšší analýzy a jako taková je mnohem

Může se zdát, že tento fakt podporuje konvencionalismus, tj. tvrzení, že existují různé podoby matematiky a je záležitostí konvence, jakou si vybereme. Na druhou stranu je potřeba vzít v úvahu, že klasická matematika nebyla otřesena konstruktivistickými pokusy a že nehrozí přímé nebezpečí, že by byla nahrazena intuicionistickou alternativou. Lze říci, že klasická matematika zvítězila na základě „přirozeného výběru“. Pro nás je zde podstatné pozorování, že integrujeme-li tvrzení (T) do matematické praxe, značně ji tím modifikujeme. Naopak integrujeme-li negaci tohoto tvrzení do matematické praxe, nic podstatného se nezmění. To chápeme jako argument ve prospěch realismu.

8 Závěr

V této práci jsme nastínili několik podob matematického realismu: Popsali jsme realismus Husserla, Gödela, Quina a Maddy a připojili jsme vlastní interpretaci této filosofické otázky. V závěru zvážíme některé námitky proti podobě realismu, kterou zde hájíme.

Nejprve se budeme zabývat námitkou, která je paralelní s možnou námitkou proti Husserlovu realismu. Když jsme transformovali strukturu fenomenologické argumentace na jazykovou úroveň, uvedli jsme, že *„je na místě být realistou, neboť jinak bychom se provinili proti smyslu matematických objektů, jak je konstituován jazykem matematiky“*. Není tím však zároveň vyjádřeno, že smysl matematických objektů je konstituován jazykem matematiky, a že matematické objekty v posledku přeci jen jsou na tomto jazyce závislé (stejně tak jako ve fenomenologii matematické objekty v posledku jsou závislé na prožitcích, v nichž se – dle fenomenologie – konstituují)? V jistém smyslu ano. Důležité však je, že pokud toto tvrdíme, změnili jsme perspektivu. Z této perspektivy psal např. Carnap některé své práce jako třeba (1934) či (1950). Při této interpretaci není realismus udržitelnou pozicí.

Lze však argumentovat, že existují metodologické důvody, proč problém realismu interpretovat spíše výše uvedeným způsobem než takto. Každé tvrzení může být vždy chápáno pouze jako tah v nějaké jazykové hře (ve smyslu Wittgensteina 1953) či jako pohyb uvnitř nějakého jazykového rámce (ve smyslu Carnapa 1950), případně jako pohyb, který nějaký nový jazykový rámec (spolu)konstituuje či starý přetváří. Tento úhel pohledu je vždy možný a v tomto smyslu je třeba souhlasit

pevněji zdůvodněna než jakákoli filosofická teorie matematické existence a matematického poznání“ (Maddy 1990, 23).

s Carnapem, že neexistují žádné absolutní externí tvrzení (viz Carnap 1950). Avšak tvrdit, že existence nějakého objektu je závislá na jazyce, a přitom se odvolávat na možnost tohoto úhlu pohledu, je podobné jako tvrdit nějakou tautologii a nejedná se tedy o příliš informativní teoretický výrok. Toto tvrzení neříká o moc více, než že pronesení existenčního tvrzení je vždy tvrzení nějakého jazyka a může být chápáno a analyzováno právě jen jako jazykové tvrzení. Být si vědom této samozřejmosti může být užitečné, podobně jako může být v nějaké matematické úvaze velmi užitečné uvědomit si, že platí p nebo $\neg p$, neboť tím můžeme chytrým způsobem rozdělit množinu možných případů na dvě disjunktní části, což může představovat důležitý krok při řešení nějakého problému. To nic nemění na faktu, že se jedná o bezobsažné tvrzení.

Naproti tomu interpretace matematického realismu inspirovaná naturalismem Penelope Maddy, jak jsme ji formulovali v předchozí části práce, vede k netriviálnímu problému a právě to zde prezentuji jako důvod pro její přijetí.

Mohli bychom dále namítat, že strategie, kterou jsme zde sledovali – totiž (zhruba) založení matematického realismu čistě na základě respektu k tomu, jak se historicky vyprofilovala řeč o matematických objektech –, je nepřijatelná, protože bychom takto mohli akceptovat třeba také existenci bájných bytostí pouze na základě povahy bájí, v nichž se o bájných bytostech mluví také realisticky.¹²

Avšak případ bájných bytostí se od případu matematických objektů podstatně liší. Máme zde velice stabilní praxi, která nám udává kritéria toho, kdy existuje či existoval nějaký materiální objekt a kdy nikoli. Existenci bájných bytostí odmítáme, právě protože je v rozporu s touto praxí. Matematická praxe s žádnou jinou stabilnější praxí v rozporu není.

Tuto práci uzavřeme poslední námitkou, která se vztahuje na celou naturalistickou metodu a kterou lze formulovat takto: Matematika sama může být vděčná velkým inovátorům, jakými byli např. Descartes či Newton, že nerespektovali stávající matematickou praxi a právě proto byli schopni modifikovat ji a něčím podstatně novým k ní při-

¹² Za příklad vděčím Igoru Sedlárovi, který podobnou námitku vznesl na Česko-slovenském sympoziu o analytické filozofii.

spět. Jejich motivy byly často filosofické – spadající do oné první filosofie, kterou naturalismus kritizuje.¹³

Domnívám se, že tento pohled je jen zdánlivě v rozporu s tím, co zde obhajujeme. Pokud v jednotlivém případě realistický (příp. antirealistický) pohled inspiruje nějakého matematika např. k zavedení nové metody, není tím tato metoda ještě obhájena. Metoda se musí etablovat na základě interních matematických důvodů a až poté zpětně můžeme říci (inspirováni naturalismem a respektem k matematické praxi), že efektivita metody v matematické praxi svědčí ve prospěch realistického (příp. antirealistického) názoru, který zavedení metody motivoval.

Filosofický ústav
Akademie věd ČR, v.v.i.
Jilská 1
110 00 Praha 1
Česká Republika
vit.puncochar@centrum.cz

Filozofická fakulta
Univerzita Karlova v Praze
Nám. Jana Palacha 2
116 38 Praha 1
Česká Republika
vit.puncochar@centrum.cz

Literatura

- BALAGUER, M. (1994): Against (Maddian) naturalized platonism. *Philosophia Mathematica* 2, 97-108.
- BENACERRAF, P. (1973): Mathematical truth. *The Journal of Philosophy* 70, 661-679.
- BOLZANO, B. (1810): Anhang über die Kantische Lehre von der Construction der Begriffe durch Anschauungen. In: *Beiträge zu einer begründeteren Darstellung der Mathematik*. Praha: Caspar Widtmann, 135-152.
- CARNAP, R. (1934): *Logische Syntax der Sprache*. Wien: Julius Springer.
- CARNAP, R. (1950): Empiricism, semantics, and ontology. *Revue Internationale de Philosophie* 4, 20-40.
- CARSON, E. (1996): On realism in set theory. *Philosophia Mathematica* 4, 3-17.
- FREGE, G. (1918): Der Gedanke. Eine logische Untersuchung. *Beiträge zur Philosophie des deutschen Idealismus* 2, 58-77.

¹³ Podobnou námitku vznesl na Česko-slovenském sympoziu o analytické filozofii Ladislav Kvasz.

- GÖDEL, K. (1944): Russell's mathematical logic. In: Schilpp, P.A. (ed.): *The philosophy of Bertrand Russell*. Evanston – Chicago: Northwestern University, 125-153. Citováno dle českého překladu: Gödel, K. (1999): *Filosofické eseje*. Praha: Oykoimenh.
- GÖDEL, K. (1947): What is Cantor's continuum problem? *The American Mathematical Monthly* 54, 515-525. Citováno dle českého překladu: Gödel, K. (1999): *Filosofické eseje*. Praha: Oykoimenh.
- GÖDEL, K. (1995): The modern development of the foundations of mathematics. In: Gödel, K.: *Collected Works III*. Edited by S. Feferman at al. Oxford: Oxford University Press. Citováno dle českého překladu: Gödel, K. (1999): *Filosofické eseje*. Praha: Oykoimenh.
- GOLDBLATT, R. (1979): *Topoi: The Categorical Analysis of Logic*. Amsterdam: North-Holland Publishing Co.
- HEYTING, A. (1966): *Intuitionism: An Introduction*. 2nd edition. Amsterdam: North-Holland publishing Company.
- HUSSERL, E. (1913): Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie. *Jahrbuch für Philosophie und phänomenologische Forschung* 1, Halle, 1-323. Citováno dle českého překladu: Husserl, E. (2004): *Ideje k čistě fenomenologii a fenomenologické filosofii I*. Praha: Oikoymenh.
- HUSSERL, E. (1929): Formale und transzendente Logik: Versuch einer Kritik der logischen Vernunft. *Jahrbuch für Philosophie und phänomenologische Forschung* 10, Halle, 1-298. Citováno dle českého překladu Husserl, E. (2007): *Formální a transcendentální logika*. Praha: Filosofia.
- KVASZ, L. (2010): Penelope Maddy medzi realizmom a naturalizmom. *Filozofia* 65, 522-537.
- KOLMAN, V. (2008): *Filosofie čísla*. Praha: Filosofia.
- LAVINE, S. (1992): Review of Maddy (1990). *The Journal of Philosophy* 89, 321-326.
- MADDY, P. (1988): Believing the axioms I. *The Journal of Symbolic Logic* 53, 481-511.
- MADDY, P. (1990): *Realism in Mathematics*. Oxford – New York: Oxford University Press.
- MADDY, P. (1993): Does V equal L? *The Journal of Symbolic Logic* 58, 15-41.
- MADDY, P. (1997): *Naturalism in Mathematics*. Oxford – New York: Oxford University Press.
- MADDY, P. (2007): *Second Philosophy*. Oxford – New York: Oxford University Press.
- QUINE, W.V.O. (1948): On what there is. *Review of Metaphysics* 2, 21-38.
- QUINE, W.V.O. (1951): Two dogmas of empiricism. *The Philosophical Review* 60, 20-43.
- QUINE, W.V.O. (1992): *Pursuit of Truth*. Cambridge: Harvard University Press.
- RISKIN, A. (1994): On the most open questions in the history of mathematics: a discussion of Maddy. *Philosophia Mathematica* 2, 109-121.
- WITTGENSTEIN, L. (1958): *Philosophische Untersuchungen*. Oxford: L. Blackwell Publishers.

Kontextualismus vs. minimalismus a metodologické principy

Jiří Raclavský

Masarykova univerzita, Brno

Abstract: The conflict between semantic minimalism and semantic contextualism seems to be irreconcilable. I suggest to view the controversy rather as a controversy concerning methodological principles which should govern investigation of language and meaning. It can be shown that some of the classically adopted principles are implausible, their a bit modified version would be more appropriate. Consequently, there will be no need to conform to any of the two extreme positions.

Keywords: contextualism; minimalism; semantics; methodological principles.

Jak známo, poslední více jak dekádu zuří v analytické filosofii jazyka konflikt mezi stoupenci sémantického kontextualismu a sémantického minimalismu. *Sémantičtí kontextualisté*¹ v zásadě říkají, že kontextuální jevy (okolnosti promluvy atp.) podstatně přispívají do sémantiky výrazů – více, než si klasičtí sémantikové vůbec kdy představovali. *Sémantičtí minimalisté*² jim oponují, přičemž říkají, že význam výrazu je jakoby minimální, není obohacován kontextuálním přispěním. (Níže si uvedeme pečlivější vystižení.)

¹ Mezi ně jsou obvykle počítáni Francois Récanati, Jason Stanley, Charles Travis, Robyn Carston, Dan Sperber, Deirdre Wilson.

² Zvláště Hermann Cappelen, Ernie Lepore, Emma Borg a snad i (neortodoxní) Kent Bach. V tuzemském prostředí se k minimalismu hlásí Marián Zouhar (srov. Zouhar 2011a).

Upřesním, že *formální sémantik*³ je sémantik, který modeluje jazyk, jeho významy, s pomocí formálních prostředků, tj. logických či matematických systémů. Tím se liší od *filosofického sémantika*, který formálních prostředků neužívá, byť o jejich existenci ví (měl by vědět). Spor kontextualismu a minimalismu se vede převážně mezi filosofickými sémantikami. Tímto se spor snad až příliš situuje do oblasti filosofie. Může pak vzniknout obava, že díky tomuto nakonec nedojde – podobně jako u jiných do filosofie příliš vsazených sporů – k jakémukoli plodnému ovlivnění praxe, totiž praktikované formální sémantiky, a to navzdory tomu, že nemálo praktikujících formálních sémantiků pochází z řad filosofů.

Nebude nevhodné, když se přiznám, že jsem praktikující formální sémantik. Více než rozvíjení filosofických pro et contra mi jde o to, jak vykonávat praxi. Proto se nebudu věnovat elaborování možných i nemožných forem minimalismu či kontextualismu nebo zase vypracování a dokládání deklarací minimalismu anebo kontextualismu, což obojí nalezneme v mnoha příspěvcích do celé diskuse.

Jde mi o to, jak provozovat sémantiku a proto chápu onen spor jako spor metodologický. Takže vlastně nepřekvapí, že na pozadí diskuse tematizují metodologické principy. Patřím k té možné většině praktikujících formálních sémantiků, jimž je bližší minimalismus spíše než kontextualismus, ovšem namísto ignorování kontextualistů se snažím v jejich celkovém návrhu nalézt poučení pro praxi. Nevyločil jsem přitom ani myšlenku, že je možná nutná paradigmatická změna v modelování významu. Domnívám se však nakonec, že žádná zásadní paradigmatická změna není nezbytná (ba snad ani žádoucí). Změnu potřebují jen dílčí *metodologické principy*, jimiž se praxe – ovšem jen do jisté míry – řídí.

V první části této stati vylíčím historicko-systémové pozadí této diskuse, což nám umožní překročit lokální perspektivu. Nahlédneme tak obecnější směrování dosavadního logicko-filosofického průzkumu jazyka, významu. V další části navážeme tím, že budeme ještě více zkoumat metodologické principy, které se v teorii jazyka vědomě či spíše nevědomě uplatňují. Ukážeme si, že vyznávání vyhraněné podoby nějakého principu vede k ocejchování jednou z nálepek ‚minimalismus‘ /

³ Známými příklady jsou např. Richard Montague, Pavel Tichý, Hans Kamp ad.

‚kontextualismus‘. Přitom ale vyhraněná podoba zdaleka není věcně přiměřená. Po adekvátnější formulaci daného principu pak ona nálepka aplikovatelná není. Jinými slovy, zastávání plausibilnějších podob daných principů vede k pozicím, které jsou zlatou střední cestou mezi extrémě skalního minimalismu a skalního kontextualismu.⁴

1 Paradigma klasické sémantiky a kořeny sporu

Účelem této sekce je si uvědomit, proč vlastně k diskutované kontroverzi došlo. Poukážeme též na dva rámcové principy, které směřují teorii jazyka: rozdílnost pragmatických a sémantických zkoumání a dále roli empirické evidence pro teorii významu.

Původní myšlenky té podoby sémantiky, která nás zde zajímá (nezajímá nás totiž sémantika, jak ji provozují lingvisté či informatici), lze bezpochyby nalézt u Gottloba Fregeho. Jeho hlavních návrhů bylo více, tím nejdůležitějším bylo modelovat jazykové významy formálními prostředky. Přispěl rovněž podezřením, že jazyková podoba nemusí vhodně korespondovat s úrovní významovou; isomorfie mezi syntaxí a sémantikou je nepravidelná. Uvědomme si však, že Fregemu nešlo zcela o jazykové významy jako takové, ale o jejich model z hlediska zájmů logiky, zejména determinování korektních úsudků. Toto vše od Fregeho převzal Bertrand Russell, ovšem významným hybatelem sémantiky byl díky svým sémantickým spisům Rudolf Carnap.

Odtud záhy dospějeme k postavám jako např. Richard Montague, kteří navrhnou formální významy dokonce už pro část jazyka (fragment angličtiny). Za poněkud ojedinělou figuru můžeme mít Pavla Tichého s jeho nedokončeným projektem *Meaning Driven Grammar*, jehož cílem byla počítačově prováděná sémantická analýza celého jazyka. V těchto snahách si nelze nevsimnout opouštění přílišné redukce vyvolané jen logickými zájmy a tedy podstatně větší materiální adekvátnost.

⁴ Jiným způsobem, jak se vyhnout nesmiřitelnosti kontroverze mezi minimalisty a kontextualisty, je poukázat na to, že mají při o dosti nejasný předmět nazývaný *what is said*. To klidně může být něco, co si mluvčí individuálně dospěkuje až v kontextu, zdaleka se nemusí jednat o (lingvistický) význam. Také se zdá, že kontextualistům jde spíš o významy promluv vět (pro mluvčí), kdežto minimalistům o významy vět. Čili se kontextualisté zřejmě zabývají (deskriptivní) pragmatikou, nikoli (deskriptivní) sémantikou.

Vývoj do počátku sedmdesátých let 20. století kdysi označil D. Kaplan za *zlatou éru formální sémantiky*. Zkusme se zamyslet, co pro ni bylo z hlediska sporu mezi minimalismem a kontextualismem příznačné. Uvádění principů sémantiky nalezneme v této éře zvláště u Carnapa, například v *Introduction to Semantics* (Carnap 1943) ilustroval studium celého jazyka na případu teoretika, který přichází mezi Eskymáky studovat jejich jazyk. Teoretikova zkoumání náleží do *pragmatiky*, když se při nich odvolává na *mluvčí* daného jazyka. Abstrahuje-li od *užití* výrazů mluvčími v *konkrétních komunikačních situacích*, náleží teoretikova zkoumání do *sémantiky*. Abstrahuje-li navíc od *významů*, teoretikova zkoumání náleží do *syntaxe*.⁵

Velké rozpaky při pohledu na současnou diskusi vzbuzuje, že se filosofové leckdy přou o *hranici mezi pragmatikou a sémantikou*, ale přitom často klasickou definici sémantiky a pragmatiky porušují, jak je z jejich polemik vidět. To by nebylo vážnou vadou, kdyby v plodné polemice s tou Carnapovou definicí uvedli, jakou definici vlastně užívají.

Klasickou definici sémantiky a pragmatiky jsme si ovšem připomněli spíše proto, abychom odhalili jeden význačný rys celého přístupu zlaté éry formální sémantiky k významu. Ze statí a knih sémantiků tohoto přístupu je totiž jasně patrné, že odvolávka na mluvčí je zcela potlačena, že od užití výrazů mluvčími, *od komunikačních kontextů* je *abstrahováno*.

Z filosofického hlediska toto může být důvodně napadáno. Při zastávání nějakých empiricistně orientovaných tendencí (Quine, Davidson, ...) lze namítat třeba následující. Spjatost výrazu s významem je přece vysledovatelná (kontrolovatelná, ...) vždy pouze v konkrétních jazykových komunikačních situacích. Od takovýchto kontextů, ale formální sémantik „abstrahuje“; tato abstrakce je tudíž přílišná.

Výhradou, z níž byly záhy vyvozeny metodologické důsledky, pak je, že sémantika zlaté éry se soustředila na významy asociované jen s *některými vybranými výrazy* a to jen v *několika málo vybraných* (informačních, oznamovacích, ...) komunikačních kontextech. Právě David Kaplan byl jedním z prvních, kdo upřeli svou pozornost ke skupině výrazů do té doby nestudovaných a navíc zřejmě pevně spjatých s užitím v komunikačních situacích (indexické výrazy ‚já‘, ‚zde‘, atd.). Začala tak být budována *formální pragmatika*.

⁵ Carnapovou definicí jsem se mj. zabýval v Raclavský (2007).

Z novějších zkoumání v tomto směru, který se mnohdy explicitně dovolává Kaplanových výzkumů, si nelze nezpomenout na badatele jako John Perry, Mark Crimmins či Mark Richard, kterými se v 80. a 90. letech 20. století ohlašují *proto-kontextualistické tendence*. Nazývám je tak proto, že relevantní kontextuální zakotvenost připisují i výrazům, které jsou tradičně pojímány jako kontextově neodvislé – pro ilustraci třeba vlastní jména. Svéráznou linií takovýchto tendencí pak představuje dvojdimenzionální sémantika (David Chalmers, Frank Jackson ad.), která v kontextu promluvy zjištěné skutečnosti vtěluje do tzv. sekundární intenze, tedy složky významu daného výrazu.

Zajímavou část formální pragmatiky tvoří již od první poloviny 70. let 20. st. formální traktování jazykové komunikace, příkladem je třeba teorie presupozice a komunikačních backgroundů spjatá zejména se jménem Roberta Stalnakera. Zde je zajímavé to, že tímto došlo aspoň do jisté míry k absorbování příspěvků P. F. Strawsona, Johna Langshaw Austina a zvláště Herberta Paula Grice i dalších. Je vhodné si připomenout, že tito se zabývali principy jazykové komunikace, nikoli čistě významem.

Strawsonovský příspěvek nabyl na významu v letech devadesátých. Strawson napadl Russellovu teorii deskripce poukazem na nevlastní deskripce, které přitom v kontextu komunikace nevedou k selhání sdělení, jak předpovídá Russellova teorie. Když v místnosti obsahující jeden stůl pronese mluvčí M_1 k mluvčímu M_2 :

„(The) stůl je pokryt knihami.“

pronesení této věty je pravdivé anebo nepravdivé pouze odvisle od toho, zda onen stůl v místnosti je pokryt knihami. V klasické sémantice je ale ona věta (a tedy jakákoli její promluva v té době) na každý pád nepravdivá (či chcete-li: bez pravdivostní hodnoty), poněvadž ona deskripce na nic nereferuje.

Tento problém má fundamentální význam pro metodologii sémantiky proto, že pravdivostní podmínky přisouzené té větě klasickou sémantikou se liší od pravdivostních podmínek, pro něž máme empirickou evidenci. Klasická sémantika, jež se zdála plausibilní, se zde tedy rozchází s empirickými daty, s tím, jak jazyk funguje.

Díky významu analýzy deskripce v teorii jazyka se mnozí sémantické 90. let 20. století (Peter Ludlow, Zoltan Szabó, aj.) snažili klasickou sémantiku přizpůsobit právě uváděné empirické evidenci. V této linii teoretiků opakovaně došlo k „ohnutím“ některých klasických princi-

pů. Například sémantik prohlásil, že logická forma (rozuměj explikans významu) dané věty obsahuje proměnnou pro pragmatické dourčení významu (v našem příkladu proměnnou pro „v této místnosti“). Už proti této vylepšené sémantice byly vzneseny námitky: sémantik přece nesmí do významu nějak „domontovávat“ kontextuální parametr, který ho napadl na základě vylíčení kontextu.

Právě připojování oněch na syntaktické rovině jen těžko odhalitelných proměnných pro kontextové parametry se stalo dosti rozšířené. Věta tak nemá za význam propozici, tedy věc plně určenou, ale tzv. propoziční radikál – propoziční torzo, které si musí uživatel jazyka v daném kontextu doplnit.⁶ Zvláště příznivci kontextualistických pozic pak říkají, že věta:

‚Prší.‘

znamená PRŠÍ NĚKDE, kde NĚKDE je chápáno jako (existenčním kvantifikátorem kvantifikovaná) *proměnná pro kontextuální parametr* místa, je to *neartikulovaná konstituenta* významu (,unarticulated constituent‘). Dalším notorickým příkladem je věta:

‚John je připraven [ready].‘

která – jak se můžeme dozvědět až z kontextu, kdy John aj. se připravují vyběhnout na běžecké dráze – znamená JOHN JE PŘIPRAVEN VYBĚHNOUT, neboli obecně je v logické formě té věty proměnná PRO NĚCO. (Jiné příklady vět: ‚Jablko je červené‘ – rozumí se, že na povrchu, ‚John otevřel dveře‘ – jak jinak než klíčem, ‚Jim je vysoký‘ – na Japonce bezpochyby, na NBA ne.)

Na přelomu 20. a 21. století se objevila řada textů filosofických sémantiků, kteří argumentovali ve prospěch větší a větší *kontextuální sensitivity* výrazů, a to i těch výrazů, které byly ve zlaté éře brány jako kontextuálně insensitivní.⁷ Často se přitom v těchto textech objevoval postup zpětně nazvaný *kontext posunující argument* (,context-shifting

⁶ Velmi dobře jsou tyto věci vyloženy např. v Récanati (2004). Tuzemským příkladem polemiky s takovýmto pojetím je Zouhar (2011).

⁷ Byť ne vždy byl tento starý přístup oprávněný. Nelze totiž nevidět, že ve zlaté éře byly traktovány jen některé výrazy – totiž ty, které šlo dobře analyzovat paradigmatickými metodami. Na ty neprozkoumávané výrazy se proto soustředily ony poněkud mimoparadigmatické linie (pro příklad DRT, Discourse Representation Theory, pro niž jsou třeba zájmena (indexické výrazy) mezi základními zkoumanými výrazy).

argument⁸): je uvážěn výraz E a jeho význam V_1 v kontextu K_1 ; na to je K_1 zaměněn pomocí jiného kontextu K_2 , v němž E očividně nemá význam V_1 , ale nějaký jiný význam V_2 . Pro konkrétní příklad,⁸ v ledničce je loužička mléka a mluvčí M_1 pronese větu ‚V lednici je mléko‘; v kontextu, kdy M_2 zjevně potřebuje mléko do své černé kávy, se ta věta (její promluva) jeví nepravdivá; v kontextu, kdy M_2 předtím uklízel ledničku, se ta věta (její promluva) jeví pravdivá.

Vyznavači právě popisovaných názorů začali hromadně zpochybňovat jasnost hranic mezi sémantikou a pragmatikou a hlavně propagovat masivní kontextualismus. Proto byli označeni za sémantické kontextualisty. Proti nim vznikla opozice teoretiků hlásících se k sémantickému minimalismu. Jejich hlavní tezí je, že vliv kontextu na sémantický obsah výrazů je mnohem menší (či spíše žádný), než si kontextualisté představují. Sémantika výrazu je tedy minimální v tom smyslu, že do ní nepatří ani kontextuální parametry, ani nějaké proměnné pro kontextuální parametry. Pro názorný příklad, věta ‚Prší‘ znamená PRŠÍ (minimální propozici) a nic jiného; věta ‚John je připraven‘ znamená JOHN JE PŘIPRAVEN, a nic jiného.

Zvláště kontextualisté obviňují minimalisty z nerealistického přístupu k jazyku. Je vskutku pochybné pomýšlet na propozici označovanou větou ‚John je připraven‘ či ‚Prší‘ mimo jakékoli kontexty. Není jasné, proč by mluvčí měl komunikovat takovéto propozice, které jsou prakticky setrvale pravdivé – vždyť vždy je něco, na co je John připraven (je připraven i nevystartovat, je připraven jít domů, je připraven biologicky zemřít, ...), snad vždy někde na Zemi prší.

Minimalisté věnovali určitou energii analýze postupů argumentace kontextualistů, aby tak dospěli ke zpochybnění všech jejich dílčích návrhů hromadně.⁹ Právě při tomto se začalo, a to spíše ze strany minimalistů, zakopávat o to, čemu zde říkáme metodologické principy, tj. postupy, jakými je žádoucí se řídit při analýze významu výrazů.

2 Metodologické principy v pozadí sporu

Jak už jsem naznačil, v poslední době začaly být některé z metodologických principů tu a tam zmiňovány a diskutovány. Třeba poukazem na tzv. naučitelnost jazyka se vícekrát argumentovalo v nepro-

⁸ Je přejet z Borg (2009, 96).

⁹ Nejznáměji Cappelen – Lepore (2005), Borg (2007).

spěch kontextualismu (jak se může mluvčí naučit význam výrazu, což intuitivně lze, když není s to dohlédnout všemožné kontexty?).¹⁰

Jsem přesvědčen, že ony principy by měly být diskutovány v mnohem větším rozsahu. To proto, že venkoncem jde o metodologické principy, které jsou pouze pomocné pro to, co je konečným *cílem* celého úsilí, totiž plausibilně explikovat jazyk a jazykovou komunikaci.

Průzkum metodologických principů, který níže ilustruji, by především měl: a) artikulovat běžně implicitně uplatňované metodologické principy (formální) sémantiky, b) popřípadě poukázat na některé jejich problematické stránky (např. ignorování šířky problematiky sémantiky), c) následně je reformulovat tak, aby byly plausibilnější,¹¹ d) prozkoumat důsledkové vztahy mezi principy.¹²

Následně by mohlo dojít – a níže si ukážeme, že skutečně dochází – k vyhnutí se příslušenství k jedné nebo druhé extrémní názorové pozici, kontextualismu či minimalismu.

(Některé metodologické principy takto k rozřešení této kontroverze samozřejmě přispět nemohou. Příkladem je princip, podle něhož jsou významy abstraktní entity (což se opírá o to, že komunikujeme shodné významy a ty významy proto nemohou být konkrétní, časoprostorově pomíjivé empirické předměty). Na základě tohoto principu je teoretik klasifikován jako realista, nicméně je neurčitelné, zda je kontextualistou či minimalistou.)

¹⁰ Ojedinělým příkladem zcela vědomého přístupu k metodologickým principům je text Zouhar (2010), jenž mi byl už v době rukopisu inspirací nejen pro zkoumání metodologických principů v Raclavský (2009), ale nakonec i pro tuto stať. Marián Zouhar totiž formuloval jistý princip (že do sémantiky nesmíme vsunovat znalosti z empirických či neempirických věd) a jeho užití ukazoval právě na příkladech vyňatých z nedávné doby, tj. též z prostředí kontroverze mezi minimalisty a kontextualisty.

¹¹ Snad netřeba říkat, že na základě plausibilních metodologických principů lze jistě více pokročit v řešení (formálně) sémantických úkolů.

¹² Také jsou tu souvislosti s proto-principy, předpoklady, jejichž příkladem je třeba předpoklad naučitelnosti: jazyk jsou s to ovládnout (naučit se) lidé s omezenými intelektuálními kapacitami. Všeobecně přijímaným důsledkem je, že do modelů jazyka je implementována konečnost prostředků (zejm. konečnost slovníku, gramatických pravidel), které přitom mají nekonečný komunikační potenciál (nekonečno vět, nekonečno významů, pro nekonečno komunikačních kontextů). Odtud zas vede cesta k principu kompozicionality významu. Atd.

Ty principy, o které mi jde, si ilustrujeme až na základě následujících (dvou) příkladů. Na první princip naráží porůznu se vyskytující formulace jako:

Z hlediska tradiční koncepce formální sémantiky ... něco je sémanticky relevantní, pouze pokud to může být dostopováno k syntaxi věty; v opačném případě je to pragmaticky relevantní.¹³

Čtenáři textu, z něhož byl citát vyjmut, může toto nijak neakcentované a neproblematizované tvrzení evokovat v sémantice dobře známý fakt. Tím je důraz na syntax v paradigmatickém pojetí formálního jazyka jakožto něčeho syntaktického, čemuž se přiřazují, a to zcela isomorfně, modelově-teoretické významy, přičemž tímto je zároveň explikován význam výrazů přirozeného jazyka.

Podle mého názoru se však jedná o mnohem závažnější skutečnost a právě ta musí být vytažena na světlo a přezkoumána. Skrývá se zde totiž jeden z fundamentálních principů klasické sémantiky. Ten byl navíc formulován v době, kdy jazyk byl zkoumán zcela bez ohledu na kontexty či mluvčí, poněvadž v té době bylo třeba řešit zcela jiné problémy. Jedná se jmenovitě o Carnapův *Princip tématu*:¹⁴

Význam (složeného) výrazu se skládá právě a pouze z významů, které jsou v tom výrazu zmíněny (vyjádřeny, pojmenovány).

Tento princip tedy říká, že např. ve významu deskripce ‚prezident USA‘ se nevyskytuje entita CLINTON, poněvadž v něm není zmíněna. Uplatnění tohoto principu v sémantice je možná názornější z jeho blízkce příbuzného *Principu adekvátní sémantické analýzy*: explikace významu složeného výrazu nesmí obsahovat svébytné významové komponenty, které nejsou v tom výrazu zjevně indikovány. Je tedy zjevné, kam princip míří – máme se vyhnout věcně nepřiměřeným explikačním významu.¹⁵

Lze též lehko uvidět, že tento princip (resp. principy) zcela ignoruje idiomy, jazykové zkratky (elipsy), metaforická užití výrazů. Zjednodušeně předpokládá pouze explicitní informativní oznamovací diskurz.¹⁶

¹³ Borg (2010, 584).

¹⁴ Přesný výrok ovšem viz v Carnap (1958, 98).

¹⁵ Principem tématu se blíže zabývám v Raclavský (2009, kap. I.1); sám zastávám jeho opatrnější formulaci.

¹⁶ Existuje tento zde uvažovaný „privilegovaný“, doslovný (informativní, věcný) kontext? Osobně jsem postupně dospěl k závěru, že ano: mimo

Zjednodušeně předpokládá neproblematickou vazbu výrazu s významem, ba vlastně isomorfii. Je tedy zjevně obecně neplatný.

V kontroverzi mezi kontextualisty a minimalisty jsou neartikulované podoby tohoto principu (viz např. citát výše) *kritériem* pro charakteristiku obou skupin. Minimalisté jsou ti, co si myslí, že syntax udává naprosto vše, co je sémanticky relevantní (předpoklad isomorfie), kdežto kontextualisté jsou ti, co s tím nesouhlasí.

Artikulovaný princip, tj. Princip tématu, a zamyšlení se nad ním, ovšem vede k jeho odmítnutí pro obecné cíle teorie významu (vysvětlení jazyka a jazykové komunikace). Vede tudíž i k opuštění skalní minimalistické pozice. Přitom je ale chybné být za nesouhlas s oním chybným principem oceňován jako kontextualista (kontextualistou je přece spíše ten, kdo činí součástí významu něco z kontextu promluvy). Jinými slovy, zastánce plausibilnější podoby Principu tématu¹⁷ se nachází ve středové pozici, která nespadá pod kterýkoli z obou extrémů. Protože plausibilitnost principu obnáší plausibilitnost pro cíle teorie jazyka a významu, středová pozice by měla být nejvíce žádoucí.

Druhý příklad principu, který prozkoumáme je *Princip jedinečného významu*:

Každý (tzv. jednoznačný) výraz má právě jeden, jedinečný význam.

Uplatňovat tento princip má např. tu praktickou výhodu, že při oscilaci mezi dvěma plausibilními kandidáty na explikans významu výrazu dochází k hledání argumentů ve prospěch jednoho z kandidátů (tyto postupy jdou samozřejmě počítačově implementovat).

Tento princip ovšem některé vede k poměrně vyhraněnému názoru, že jazyk je funkce z výrazů do významů, kdy jednomu výrazu odpovídá právě jeden význam.¹⁸ Tento princip může vést také k názo-

běžné kontexty (např. když za deště v našem bezprostředním okolí říkáme ‚Prší‘) je tu i komunikační kontext, kdy se ptáme – např. na sémantickém semináři – na význam slov (‚Co je významem ‚John je připraven‘?; tehdy je plně adekvátní odpovědí ‚JOHN JĚ PŘIPRAVEN‘). Pro linii mé následné argumentace však toto není podstatné.

¹⁷ Plausibilnější podoby by obsahovaly obraty jako ‚význam desambiguovaného výrazu v doslovném kontextu/informativním diskurzu‘. Poslání principu v této formulaci by bylo shodné – není např. důvod, proč by třeba v tzv. přímých (syntaktických) kontextech mělo být součástí významu nějaké deskripce individuum, na které náhodou ta deskripce referuje.

¹⁸ Srov. k tomuto např. Raclavský (2009, kap. IV.5).

ru, že výraz musí mít právě jeden význam „kontext nekontext“, tj. minimalismus. Každopádně zastávání tohoto principu v důsledku vede k neadekvátním vysvětlením významu všude tam, kde je intuice, že daný výraz má (popř. odvisle od kontextu) více jak jeden význam.

K dosažení plausibility sémantické analýzy přitom stačí přijatelná změna principu, resp. toho, co si z něj odvodíme. Jednou z možností je opustit názor, že jazyk je funkcí z výrazů do významů, a to ve prospěch názoru, že jazyk je funkce z výrazů do posloupností významů.¹⁹ Významy v dané posloupnosti jsou uplatnitelné v těch či oněch (syntaktických či pragmatických) kontextech.²⁰

Další možnost byla již dávno přijata praktikujícími sémantiky jako např. Montague či Tichý.²¹ Formální sémantik, resp. jeho počítačová implementace, neanalyzuje jazykové výrazy v podobě, jaké „najde na ulici“. Analyzuje až produkty lingvistické analýzy, určité stromové struktury (v lingvistickém prostředí známe též tzv. Logical Forms); jednomu běžnému výrazu pak přiléhá třeba třicítka desambiguovaných podob (tj. těch stromů). Každé takovéto syntaktické entitě pak přiléhá jedinečný význam, takže princip jedinečného významu je naplněn.

U obou možností přitom platí, že kontextuálnost není imputována do významu, jak činí kontextualisté. Přitom se ale nejedná o žádný známý minimalismus. Zajímavé na tomto „politicky nevyprofilovaném“ přístupu také je, že má nemálo praktikujících stoupenců, bez nichž by dokonce celý podnik analýzy jazyka byl jen nezávaznou zábavou.²²

¹⁹ Tento návrh jsem přejal z Lewis (1983).

²⁰ Tento model jazyka by šlo zajímavě a nosně rozpracovat. Setká-li se mluvčí M_2 s určitým výrazem E v kontextu K_1 , tak podle *pravidel* pro daný slovní kontext, či spíše typ komunikační situace, by si z dané posloupnosti vybral mluvčím M_1 v tom kontextu K_1 intendovaný význam. Myslím, že právě nějaké takovéto komunikační pravidlo bylo uplatněno v příkladu se (Strawsonovým) stolem (Raclavský 2009, 341, pozn. 11). Poznámávám, že typy komunikačních situací, resp. pravidla pro ně, jsou (na rozdíl od nekonečna komunikačních situací) naučitelné.

²¹ Viz k tomuto Raclavský (2006a), kde odkazují na oba teoretiky a hlavně ukazují, jak probíhá počítačová analýza lingvisty dodaných vstupů prostředky Tichého systému.

²² V korekturách dopisují poznámku, že tento rok v tuzemsku k diskutované tématice vyjde cenná kniha Zouhar (2011b).

Katedra filozofie
 Filozofická fakulta
 Masarykova univerzita
 Arne Nováka 1
 602 00 Brno
 Česká republika
 raclavsky@phil.muni.cz

Literatura

- BORG, E. (2007): *Minimal Semantics*. Oxford University Press.
- BORG, E. (2007a): Minimalism versus Contextualism in Semantics. In: Preyer, G. – Peter, G. (eds.): *Context-Sensitivity and Semantic Minimalism: New Essays on Semantics and Pragmatics*. Oxford University Press, 339-359.
- BORG, E. (2009): Meaning and Context: a Survey of a Contemporary Debate. In: Whiting, D. (ed.): *The Later Wittgenstein on Language*. Palgrave Macmillan, 96-113.
- BORG, E. (2010): Pragmatic Determinants of What is Said. In: Barber, A. – Stainton, R. J. (eds.): *Concise Encyclopedia of Philosophy of Language and Linguistics*. Elsevier, 582-585.
- CARNAP, R. (1942): *Introduction to Semantics*. Harvard University Press.
- CARNAP, R. (1958): *Meaning and Necessity*. (Second Edition). The University of Chicago Press (Phoenix Edition).
- LEWIS, D. (1983): Languages and Language. In: Lewis, D.: *Philosophical Papers. Vol. I*. Oxford University Press, 163-188.
- RACLAVSKÝ, J. (2006): Co je sémantika. *Filosofický časopis* 54, No. 2, 267-279.
- RACLAVSKÝ, J. (2006a): Počítačové zpracování přirozeného jazyka a Transparentní intenzionální logika. In: Dostálová, L. – Šebela, K. (eds.): *Organon V. (Kolik logiky za nás mohou učit počítače?)*. Praha: Univerzita Karlova, 127-135.
- RACLAVSKÝ, J. (2009): *Jména a deskripce: logicko-sémantická zkoumání*. Olomouc: Nakladatelství Olomouc.
- RÉCANATI, F. (2004): *Literal Meaning*. Cambridge University Press.
- ZOUHAR, M. (2010): Linguistic Knowledge, Semantics and Ideal Speakers (Some Remarks on the Methodology of Natural Language Semantics). *Epistemologia* 33, 41-64.²³
- ZOUHAR, M. (2010a): *Medzi sémantikou a epistemológiou jazyka*. Bratislava: Aleph.
- ZOUHAR, M. (2011): What is Wrong with Unarticulated Constituents? *Human Affairs* 21, No. 3, 249-261.
- ZOUHAR, M. (2011a): Kontext a sémantika predikátov. *Filosofický časopis* 59, No. 2, 187-205.
- ZOUHAR, M. (2011b): *Význam v kontextu*. Bratislava: aleph.

²³ Slovenskou verzi viz v Zouhar (2010a, 271-292).

Neomylnosť a logika¹

Igor Sedlár

Univerzita Komenského v Bratislave

Abstract: The paper presents preliminary results related to the logical analysis of the concept of infallibility. The concept is explicated using the resources of modal logic, namely in terms of the operator of possibility M and the operator of belief B . Infallibility with respect to p is explicated as impossibility of being wrong with respect to p . The infallibility operator is introduced: $I\phi$ is defined to mean the same as $\neg M(B\phi \wedge \neg\phi)$. We prove various theorems about validity or invalidity of certain formulas with I in special classes of models for the combined language.

Keywords: infallibility, belief, possibility, modal logic.

Začnime jednou samozrejmosťou:

(1) Ľudia sa môžu myliť.

Napriek tomu, že (1) znie skutočne triviálne, má v dejinách epistemológie dôležité miesto. Dejiny epistemológie môžeme totiž do určitej miery chápať ako rad pokusov ukázať, že (1) neplatí všeobecne. Inými slovami, ako rad pokusov dokázať nepravdivosť tvrdenia

(2) Ľudia sa môžu myliť vo všetkom.

Jadrom rôznych odpovedí na skeptickú výzvu je totiž úsilie ukázať, že existujú propozície, pri ktorých sa *nemôžeme myliť*. V Sedlár (2011) som načrtnol spôsob formálneho skúmania takýchto propozícií a uviedol som niekoľko príkladov. V tomto príspevku sa pozriem na samotný pojem neomylnosti.

¹ Tento príspevok vznikol na Katedre logiky a metodológie vied FiF UK v Bratislave ako súčasť výskumného projektu *Sémantické modely, ich explanačná sila a aplikácie*, podporeného grantom VEGA č. 1/0046/11. Účastníkom sympózia ďakujem za diskusiu a užitočné pripomienky.

1 Neomylnosť: syntax

Je jasné, že *mýliť sa* vzhľadom na nejakú propozíciu p znamená byť presvedčený o pravdivosti p , keď p je v skutočnosti nepravdivá. Ak rozšírime jazyk klasickej výrokovej logiky o unárny operátor B , pričom Bp znamená „Som presvedčený, že p “, tak skutočnosť, že sa vzhľadom na p mýlim, vieme zachytiť formulou

$$(3) Bp \wedge \neg p$$

Jazyk s B však na formalizáciu pojmu neomylnosti nestačí. Byť neomylný vzhľadom na nejaké p znamená, že je *nemožné* mýliť sa vzhľadom na p . Ak rozšírime jazyk s operátorom B o nový unárny operátor M , pričom Mp znamená „Je možné, že p “, tak skutočnosť, že som neomylný vzhľadom na p , vieme zachytiť formulou

$$(4) \neg M (Bp \wedge \neg p)$$

Ak v tomto rozšírenom jazyku známym spôsobom² definujeme operátor nevyhnutnosti L , tak formulu (4) vieme prepísať do ekvivalentnej a prehľadnejšej podoby

$$(5) L(Bp \rightarrow p)$$

Kvôli prehľadnosti môžeme v rozšírenom jazyku s B a M definovať operátor neomylnosti I : Nech $I\phi$ je skratkou za (5) a podobne pre ľubovoľnú formulu ϕ na mieste p . $I\phi$ teda znamená „Som neomylný vzhľadom na ϕ “, kde ϕ je ľubovoľná formula.

2 Neomylnosť: sémantika

V tejto časti sa pozrieme na sémantiku jazyka s B a M . Tú postavíme na známych východiskách: na relačnej sémantike pre jazyk s M a na minimálnej sémantike pre jazyk s B . Časti 2.1 a 2.2 majú úvodný charakter a stručne vysvetľujú relačnú a minimálnu sémantiku výrokovej modálnej logiky a ich interpretácie, ktoré sú pre tento článok podstatné: metafyzickú a doxastickú. V časti 2.3 definujeme modely pre jazyk s B , M . Tie sú kombináciou relačných a minimálnych modelov.

² Teda $L\phi$ nech znamená $\neg M \neg \phi$, pre ľubovoľnú formulu ϕ .

2.1 Možnosť

Modálna logika má v súčasnosti mnoho rôznych podôb,³ no jej zárodkom boli systémy, medzi ktorých nedefinované pojmy patrili aj pojem možnosti. Tento fakt môže byť vysvetlením skutočnosti, že modálna logika sa niekedy aj dnes pokladá primárne za „logiku možnosti a nevyhnutnosti“. To zďaleka nie je adekvátne, avšak toto prihliadnutie na korene modálnej logiky môže pomôcť pri vysvetlení jej relačnej sémantiky.

Relačná sémantika výrokovej modálnej logiky vychádza z leibnizovského chápania možnosti ako pravdivosti v nejakom možnom svete a nevyhnutnosti ako pravdivosti v každom možnom svete. Kvôli širšej aplikovateľnosti táto sémantika využíva aj reláciu „dosiahnuteľnosti“ medzi možnými svetmi, zvyčajne označovanú ako R . Východiskom tejto sémantiky je teda myšlienka, že výrok tvaru „Je možné, že p “ je pravdivý v nejakom svete x práve vtedy, keď existuje nejaký svet y , ktorý je dosiahnuteľný z x a zároveň p je pravdivé v y . Doplňme, že výrok tvaru „Je nevyhnutné, že p “ je pravdivý v x práve vtedy, keď p je pravdivé v každom y , ktorý je dosiahnuteľný z x .⁴

Vlastnosti R samozrejme vplývajú na platnosť výrokov. Napríklad výrok typu „Ak je nevyhnutné, že p , tak p “ bude platný, ak predpokladáme, že R je reflexívna (teda pre každý svet x platí Rxx). Trocha voľnejšie môžeme povedať, že rôznym chápaniam nevyhnutnosti vyhovujú rôzne „nastavenia“ relácie R . Pojem nevyhnutnosti sa však najčastejšie modeluje pomocou ekvivalenčnej relácie R (teda relácie, ktorá je reflexívna, symetrická a tranzitívna).

Sémantiku pre jazyk s M (ďalej „modálny jazyk“) teraz definujeme presnejšie. *Modálny model* je trojica $M = (W, R, V)$, kde W je neprázdna množina (množina možných svetov), R je reflexívna, symetrická a tranzitívna binárna relácia na W a V je funkcia, ktorá každej dvojici (ϕ, x) , kde ϕ je formula modálneho jazyka a $x \in W$, priradí nejakú pravdivosť-

³ Pozri Blackburn – van Benthem – Wolter (2006).

⁴ S modálnou logikou a jej sémantikou sa čitateľ môže bližšie zoznámiť napríklad v Peregrin (2004), Szomolányi (2003). Spomedzi zahraničných prác pozri napríklad Blackburn – de Rijke – Venema (2001), Hughes – Cresswell (1996), Chellas (1980) či van Benthem (2010).

nú hodnotu z množiny $\{0, 1\}$. V ďalej spĺňa nasledujúce požiadavky („vtt“ je skratka za „práve vtedy, keď“):

1. Pre každú výrokovú premennú p a každý svet $x \in W$ platí buď $V(p, x) = 1$, alebo $V(p, x) = 0$.
2. $V(\neg\phi, x) = 1$ vtt $V(\phi, x) = 0$
3. $V(\phi \wedge \psi, x) = 1$ vtt $V(\phi, x) = 1$ a $V(\psi, x) = 1$
4. $V(M\phi, x) = 1$ vtt $V(\phi, y) = 1$, pre nejaké také $y \in W$, že Rxy

Formula ϕ je *platná* v modálnom modeli $M = (W, R, V)$ práve vtedy, keď $V(\phi, x) = 1$ pre všetky $x \in W$. Keď definujeme $L\phi$ ako $\neg M\neg\phi$, tak $V(L\phi, x) = 1$ vtt $V(\phi, y) = 1$ pre všetky také $y \in W$, že Rxy . *Mod* nech je trieda všetkých modálnych modelov. $\text{Log}(Mod)$ nech je trieda všetkých formúl modálneho jazyka platných v každom $M \in Mod$.

Triedu $\text{Log}(Mod)$ môžeme chápať ako súbor tvrdení, ktoré v istom zmysle popisujú pojem možnosti (a teda aj pojem nevyhnutnosti). Bude užitočné uviesť niekoľko formúl, ktoré do $\text{Log}(Mod)$ patria:

- (6) $L(p \rightarrow q) \rightarrow (Lp \rightarrow Lq)$
- (7) $Lp \rightarrow p$
- (8) $Lp \rightarrow LLp$
- (9) $p \rightarrow MLp$
- (10) $Mp \rightarrow LMp$

Spomeňme, že každá spomedzi týchto formúl je platná v každej triede kombinovaných modelov, o ktorých budeme hovoriť nižšie. Rovnako je platná aj každá substitučná inštancia týchto formúl. Platnosť niektorých z nich bude dôležitým predpokladom v dôkazoch uvedených nižšie.

2.2 Logika presvedčenia

Slovné spojenia „Je možné, že ...“ a „Som presvedčený, že ...“ majú zaujímavé spoločné črty. Po prvé, oba môžeme chápať tak, že označujú výrokové operátory. Napríklad, ak za ... v „Som presvedčený, že ...“ dosadím nejaký výrok, tak dostanem znova nejaký výrok. Po druhé, oba spomenuté výrokové operátory sú neextenzionálne. Pravdivostná hodnota výrokov typu „Je možné, že p “ a „Som presvedčený, že p “ nie je určená výlučne pravdivostnou hodnotou výroku p .

Ihneď po formulovaní relačnej sémantiky pre modálnu logiku bolo zrejmé, že túto sémantiku je možné interpretovať doxasticky, a teda ju použiť ako formálny model pojmu presvedčenia; pozri klasickú prácu

Hintikka (1962). Takto vzniknuté formálne modely pojmu presvedčenia (a taktiež pojmu poznania) našli aplikácie v rôznych oblastiach.⁵

Z rôznych dôvodov⁶ sa však doxasticky interpretovaná relačná sémantika zdá byť príliš idealizujúcim modelom presvedčení. Ak opustíme niektoré diskutabilné predpoklady, dostávame sa od relačnej sémantiky k minimálnej sémantike.⁷

Minimálna sémantika pre jazyk s operátorom B (ďalej „doxastický jazyk“) je nasledovná. *Minimálny doxastický model* (ďalej iba „doxastický model“) je trojica $M = (W, N, V)$, kde W a V sú ako v relačnej sémantike, pričom N je funkcia, ktorá každému $x \in W$ priradí nejakú množinu podmnožín W . $N(x)$ je teda množina množín prvkov W . (Nemusí však ísť o množinu všetkých množín prvkov W !) V ďalej spĺňa nasledujúce požiadavky:

1. Pre každú výrokovú premennú p a každý svet $x \in W$ platí buď $V(p, x) = 1$, alebo $V(p, x) = 0$
2. $V(\neg\phi, x) = 1$ vtt $V(\phi, x) = 0$
3. $V(\phi \wedge \psi, x) = 1$ vtt $V(\phi, x) = 1$ a $V(\psi, x) = 1$
4. $V(B\phi, x) = 1$ vtt $|\phi| \in N(x)$, kde $|\phi| = \{y : V(\phi, y) = 1\}$

Formula ϕ je *platná* v doxastickom modeli $M = (W, R, V)$ práve vtedy, keď $V(\phi, x) = 1$ pre všetky $x \in W$.

Prečo by sme však mali súhlasiť s tým, že minimálna sémantika reprezentuje pojem presvedčenia? Po prvé, musíme si uvedomiť, že ak prvky W chápeme ako možné svety, tak množiny prvkov W reprezentujú propozície. Funkciu N potom môžeme chápať tak, že pre nejakého agenta stanovuje, o akých propozíciách je presvedčený v rôznych možných svetoch. $N(x)$ teda reprezentuje množinu propozícií, ktorým agent verí vo svete x .

Minimálna sémantika so sebou nesie iba jeden predpoklad. Ak totiž formulám ϕ a ψ zodpovedá v každom modeli tá istá propozícia (teda ak sú logicky ekvivalentné), tak v každom modeli platí aj to, že agent je presvedčený o ϕ práve vtedy, keď je presvedčený o ψ . Minimálna

⁵ K modálno-logickému modelovaniu pojmov presvedčenia a poznania pozri Fagin – Halpern – Moses – Vardi (1995), Meyer – van der Hoek (1995), kapitoly 12 a 13 práce van Benthem (2010) či prehľadový článok Meyer (2001).

⁶ Pozri najmä 9. kapitolu práce Fagin – Halpern – Moses – Vardi (1995).

⁷ K minimálnej sémantike pozri kapitoly 7 – 9 práce Chellas (1980).

sémantika je teda vhodným nástrojom modelovania presvedčení agentov, ktorí sú schopní zistiť, že dve propozície sú logicky ekvivalentné a prispôbiť tomu svoje presvedčenia.⁸ Ide skutočne o minimálnu požiadavku. Je totiž jasné, že sám projekt formálneho modelovania presvedčení vyžaduje, aby sa presvedčenia riadili aspoň nejakými všeobecnými pravidlami. (Ak by sa žiadnymi neriadili, neexistovali by žiadne pravidlá, ktoré by formálne modelovanie mohlo odhaliť.)

Presvedčeniam však môžeme pripisovať rôzne dodatočné vlastnosti. Nasledujúca tabuľka uvádza niektoré vlastnosti, ktoré sa v literatúre často objavujú. Pri každej vlastnosti uvádzame formulu doxastického jazyka, ktorá ju reprezentuje a zároveň vysvetľuje.

Vlastnosť presvedčenia	Formula
Konzistentnosť	$\neg(B\phi \wedge B\neg\phi)$
Uzavretosť na konjunkcie	$(B\phi \wedge B\psi) \rightarrow B(\phi \wedge \psi)$
Distribúcia do konjunkcie	$B(\phi \wedge \psi) \rightarrow (B\phi \wedge B\psi)$
Pozitívna introspekcia	$B\phi \rightarrow BB\phi$
Negatívna introspekcia	$\neg B\phi \rightarrow B\neg B\phi$
Logicky dokonalé presvedčenie	$B(p \vee \neg p)$

Možné dodatočné predpoklady týkajúce sa presvedčenia vieme v minimálnej sémantike zohľadniť špecifickými „nastaveniami“ funkcie N . V nasledujúcej tabuľke uvádzame formuly z predchádzajúcej tabuľky, pričom pri každej uvádzame vlastnosť N , ktorá podmieňuje platnosť danej formuly ($A, B \subseteq W$).

Formula	Vlastnosť N
$\neg(B\phi \wedge B\neg\phi)$	Ak $A \in N(x)$, tak $(W - A) \notin N(x)$
$(B\phi \wedge B\psi) \rightarrow B(\phi \wedge \psi)$	Ak $A, B \in N(x)$, tak $A \cap B \in N(x)$
$B(\phi \wedge \psi) \rightarrow (B\phi \wedge B\psi)$	Ak $A \in N(x)$ a $A \subseteq B$, tak $B \in N(x)$
$B\phi \rightarrow BB\phi$	Ak $A \in N(x)$, tak $\{y : A \in N(y)\} \in N(x)$
$\neg B\phi \rightarrow B\neg B\phi$	Ak $A \notin N(x)$, tak $\{y : A \notin N(y)\} \in N(x)$
$B(p \vee \neg p)$	$W \in N(x)$

⁸ Všimnime si, že minimálna sémantika nevyžaduje, aby bol agent schopný odhaliť všetky logické dôsledky svojich presvedčení. Požaduje sa iba „citlivosť na logickú ekvivalentnosť“.

2.3 Logika možnosti a presvedčenia

Sémantika pre jazyk, ktorý obsahuje M spolu s B (ďalej „dm-jazyk“) vznikne zlúčením relačných modelov (pre M) s minimálnymi modelmi (pre B). Pri minimálnych modeloch sme uvažovali o rôznych možných vlastnostiach N, takže aj pri kombinovaných modeloch bude nutné odlišiť viacero druhov.

Definícia 2.1. Základný dm-model (skrátene „základný model“) je $M = (W, R, N, V)$, kde W je neprázdna množina (množina možných svetov), R je reflexívna, symetrická a tranzitívna binárna relácia na W , N je funkcia z W do $P(P(W))$ (do množiny všetkých podmnožín W) a V je funkcia, ktorá každej dvojici (ϕ, x) , kde ϕ je formula dm-jazyka a $x \in W$, priradí nejakú pravdivostnú hodnotu z množiny $\{0, 1\}$. V ďalej spĺňa nasledujúce požiadavky:

1. Pre každú výrokovú premennú p a každý svet $x \in W$ platí buď $V(p, x) = 1$, alebo $V(p, x) = 0$
2. $V(\neg\phi, x) = 1$ vtt $V(\phi, x) = 0$
3. $V(\phi \wedge \psi, x) = 1$ vtt $V(\phi, x) = 1$ a $V(\psi, x) = 1$
4. $V(M\phi, x) = 1$ vtt $V(\phi, y) = 1$ pre nejaké také $y \in W$, že Rxy
5. $V(B\phi, x) = 1$ vtt $|\phi| \in N(x)$, kde tradične $|\phi| = \{y : V(\phi, y) = 1\}$

Formula ϕ je platná v základnom modeli $M = (W, R, N, V)$ práve vtedy, keď $V(\phi, x) = 1$ pre všetky $x \in W$. $MBel_E$ nech je množina základných modelov pre dm-jazyk. $Log(MBel_E)$ nech je trieda formúl dm-jazyka, ktoré sú platné v každom $M \in MBel_E$.

Definícia 2.2. Racionálny dm-model je taký základný model, v ktorom pre každú $A \subseteq W$ navyše platí: Ak $A \in N(x)$, tak $(W - A) \notin N(x)$. Triedu všetkých racionálnych modelov označíme $MBel_D$ a triedu všetkých formúl dm-jazyka platných v každom $M \in MBel_D$ označíme $Log(MBel_D)$.

Konjunktívny dm-model je taký základný model, v ktorom pre každú $A \subseteq W$ navyše platí: Ak $A, B \in N(x)$, tak $A \cap B \in N(x)$. Triedu všetkých konjunktívnych modelov označíme $MBel_C$ a triedu všetkých formúl dm-jazyka platných v každom $M \in MBel_C$ označíme $Log(MBel_C)$.

Regulárny dm-model je taký konjunktívny model, v ktorom pre každé $A, B \subseteq W$ navyše platí: Ak $A \in N(x)$ a $A \subseteq B$, tak $B \in N(x)$. Triedu všetkých regulárnych modelov označíme $MBel_R$ a triedu všetkých formúl dm-jazyka platných v každom $M \in MBel_R$ označíme $Log(MBel_R)$.

Extrémny dm-model je taký regulárny model, v ktorom pre každú $A \subseteq W$ navyše platí:

1. Ak $A \in N(x)$, tak $(W - A) \notin N(x)$
2. Ak $\text{ran}_R(x) \subseteq A$, tak $A \in N(x)$, kde $\text{ran}_R(x) = \{y : Rxy\}$

Triedu všetkých extrémnych modelov označíme $MBel_x$ a triedu všetkých formúl dm-jazyka platných v každom $M \in MBel_x$ označíme $\text{Log}(MBel_x)$.

3 Vlastnosti neomylnosti

V tejto časti dokážeme niekoľko tvrdení o platnosti respektíve neplatnosti rôznych formúl v rôznych druhoch dm-modelov. Pripomeňme, že $I\phi$ je skratkou za $L(B\phi \rightarrow \phi)$, teda za $\neg M(B\phi \wedge \neg\phi)$.

3.1 Základné modely

Základné modely reprezentujú kombináciu pojmu možnosti s minimálnym pojmom presvedčenia. Formuly obsahujúce operátor neomylnosti, ktoré patria do $MBel_E$ teda v istom zmysle predstavujú „základné zákony“ neomylnosti.

Tvrdenie 3.1. Nasledujúce formuly patria do $\text{Log}(MBel_E)$:

- (a) $(Bp \wedge Ip) \rightarrow p$
- (b) $Ip \rightarrow LIp$
- (c) $MIp \rightarrow Ip$
- (d) $(MIp \wedge Bp) \rightarrow p$

Bod a) môžeme chápať ako prvý test vhodnosti modálnej reprezentácie neomylnosti. Ide totiž o tvrdenie, ktoré by o neomylnosti určite malo platiť: Ak som presvedčený, že p je pravdivé a zároveň som neomylný vzhľadom na p , tak p .

Bod b) sa na prvý pohľad môže zdať trochu prekvapujúci: Ak som neomylný vzhľadom na p , tak som nevyhnutne neomylný vzhľadom na p . Podľa tohto tvrdenia neomylnosť nie je kontingentná. Ak som neomylný vzhľadom na nejakú propozíciu, tak nie je („metafyzicky“) možné, aby som taký nebol. Po zoznámení sa s dôkazom tohto tvrdenia (pozri dodatok) však bude jasné, že platnosť uvedenej formuly je triviálnym dôsledkom skutočnosti, že neomylnosť je v podstate modálna vlastnosť: Ip znamená $L(Bp \rightarrow p)$. Pre nevyhnutnosť však platí $Lq \rightarrow LLq$.

Podobne je možné komentovať aj bod c). Ten sa na prvý pohľad tiež môže zdať trochu prekvapivý: Ak je možné, že som neomylný vzhľadom na p , tak som skutočne neomylný vzhľadom na p . Tu si však treba uvedomiť, podobne ako pri bode b), že platnosť uvedenej formuly je triviálnym dôsledkom predpokladaných črt „metafyzickej“ možnosti. Stačí si uvedomiť, že každá formula tvaru $ML\phi \rightarrow L\phi$ je platná.

Podľa bodu d) platí, že ak je možné, že som neomylný vzhľadom na nejaké p , o ktorom som zároveň presvedčený, tak p je pravdivé. Toto silne znejúce tvrdenie je jednoduchým dôsledkom bodov a) a c) (pozri dodatok).

Prv, než pôjdeme ďalej, je dobré uvedomiť si, že neomylnosť neimplikuje pravdivosť:

Tvrdenie 3.2. Formula $I_p \rightarrow p$ nepatrí do $\text{Log}(MBel_E)$.

Uvedená formula nie je tiež platná v žiadnej spomedzi tried modelov, ktorým sa venujeme nižšie.

3.2 Racionálne modely

Podstatnou črtou racionálnych modelov je to, že ak $|\phi| \in N(x)$, tak $|\neg\phi| \notin N(x)$. Inými slovami, ak je niekto presvedčený o ϕ , tak nie je presvedčený o $\neg\phi$. Tieto modely teda reprezentujú iba *konzistentné* presvedčenia. Často sa predpokladá, že konzistentnosť presvedčení je nevyhnutnou podmienkou ich racionálnosti⁹ (odtiaľ názov tejto triedy modelov).

Je však zrejmé, že množina propozícií, o ktorých je niekto neomylný, nemusí byť sama konzistentná:

Tvrdenie 3.3. Formula $I_p \rightarrow \neg I\neg p$ nepatrí do $\text{Log}(MBel_D)$.

Platnosť tohto tvrdenia by čitateľa nemala prekvapiť. Predpoklad, že niekto je neomylný vzhľadom na „Bolí ma zub“, nevyklučuje možnosť, že by bol neomylný aj vzhľadom na „Nebolí ma zub“. Často sa dokonca predpokladá, že napríklad z neomylnosti „Bolí ma zub“ vyplýva neomylnosť vzhľadom na negáciu tohto výroku, teda vzhľadom na „Nebolí ma zub“. Dá sa však ukázať, že tak to vo všeobecnosti nie je:

Tvrdenie 3.4. Formula $I_p \rightarrow I\neg p$ nepatrí do $\text{Log}(MBel_D)$.

⁹ Avšak pozri Makinson (1965).

Toto tvrdenie si zaslúži určitú pozornosť. Ak sa vo filozofickej literatúre stretne s argumentmi v prospech názoru, že pri tvrdeniach určitého druhu sa nemôžeme mýliť, pričom daný argument pracuje výlučne s pozitívnymi tvrdeniami daného druhu,¹⁰ tak daný argument nie je platný. Stále sa môžeme mýliť pri negatívnych tvrdeniach daného druhu.¹¹

Je tiež zaujímavé si všimnúť, že formula $I p \rightarrow I \neg p$ nie je platná v žiadnej spomedzi tried modelov, o ktorých hovoríme nižšie.

3.3 Konjunktívne modely

Špeciálnou črtou konjunktívnych modelov je to, že v nich platí každá formula formy $(B\phi \wedge B\psi) \rightarrow B(\phi \wedge \psi)$. Tieto modely teda reprezentujú predpoklad, že presvedčenie je uzavreté na konjunkcie: Ak som presvedčený o ϕ , ψ , tak som presvedčený aj o ich konjunkcii, teda $\phi \wedge \psi$.

Zdá sa rozumné očakávať, že táto črta presvedčení sa preniesie aj na neomylnosť. Tak to však nie je:

Tvrdenie 3.5. Formula $(I p \wedge I q) \rightarrow I(p \wedge q)$ nepatrí do $\text{Log}(M\text{Bel}_c)$.

Toto tvrdenie ukazuje, že „uzavretosť na konjunkcie“ sa z presvedčenia neprenáša na neomylnosť. To, že som neomylný vzhľadom na ϕ , ψ , ešte neznamená, že som neomylný vzhľadom na $\phi \wedge \psi$.

3.4 Regulárne modely

Význačnou črtou regulárnych modelov je to, že v nich pre každé ϕ , ψ platí $(B\phi \wedge B\psi) \leftrightarrow B(\phi \wedge \psi)$. Reprezentujú teda predpoklad, že o konjunkcii dvoch tvrdení som presvedčený práve vtedy, keď som presvedčený o oboch daných tvrdeniach.

¹⁰ Tvrdeniami, ktoré nie sú ekvivalentné so žiadnym tvrdením, ktoré začína nepárnym počtom negácií.

¹¹ Zaujímavým príkladom je propozícia vyjadrená výrokom „Existuje propozícia, vzhľadom na ktorú sa mýlim“. Čitateľ si môže jednoducho overiť, že pri tejto propozícii sa nemôžem mýliť. Určite sa však môžem mýliť pri jej negácii, teda pri propozícii vyjadrenej výrokom „Nemýlim sa pri žiadnej propozícii“. Pozri aj Sedlár (2011). Ešte jednoduchším, ba až triviálnym príkladom sú tautológie klasickej výrokovej logiky. Pri žiadnej z nich sa nemôžem mýliť, no určite sa môžem mýliť pri ich negáciách.

Dalo by sa čakať, že sa táto črta prenesie na neomylnosť (podobne ako to bolo s uzavretosťou na konjunkcie pri konjunktívnych modeloch). Tak to však nie je.

Tvrdenie 3.6. Nasledujúce formuly nepatria do $\text{Log}(MBel_{\mathbb{R}})$:

- (a) $I(p \wedge q) \rightarrow (Ip \wedge Iq)$
- (b) $I(p \vee q) \rightarrow (Ip \vee Iq)$
- (c) $(Ip \vee Iq) \rightarrow I(p \vee q)$

Dá sa však ukázať, že platnosť formuly $B(p \wedge q) \rightarrow (Bp \wedge Bq)$ zaručuje skutočnosť, že neomylnosť je uzavretá na konjunkcie:

Tvrdenie 3.7. Formula $(Ip \wedge Iq) \rightarrow I(p \wedge q)$ patrí do $\text{Log}(MBel_{\mathbb{R}})$.

Zaujímavé je, že na platnosť tejto formuly nestačí predpoklad platnosti $(Bp \wedge Bq) \rightarrow B(p \wedge q)$. Musíme predpokladať opačnú implikáciu, teda $B(p \wedge q) \rightarrow (Bp \wedge Bq)$.

3.5 Extrémne modely

Ako už napovedá ich názov, extrémne modely pracujú s určitými veľmi silnými predpokladmi. Všimnime si, že pracujú s predpokladom, že modelované presvedčenia sú konzistentné. Ďalej pracujú so silnejším predpokladom, podľa ktorého platí, že ak $V(x, L\phi) = 1$, tak $|\phi| \in N(x)$. Inými slovami, ak je ϕ nevyhnutne pravdivá, tak je o nej agent presvedčený. Ešte jednoduchšie, agent je presvedčený o každej nevyhnutnej pravde.

Extrémnosť takýchto modelov však môžeme „zmierniť“ tým, že oslabíme interpretáciu operátora B. Predpokladajme, že pri práci s extrémnymi modelmi nemodelujeme „uvedomelé“ presvedčenia daného agenta, ale množiny tvrdení, ktoré z jeho „uvedomelých“ presvedčení vyplývajú.¹² Pri takomto čítaní nie je uvedená črta extrémnych modelov taká extrémna: Z presvedčení daného agenta samozrejme vyplýva každá nevyhnutná pravda. (Práve preto, že je nevyhnutná. Inými slovami, nemôže byť nepravdivá.)

Predpokladajme teda, že modelujeme agentov s konzistentnými presvedčeniami, pričom $B\phi$ chápeme ako „Z presvedčení daného agenta vyplýva ϕ “. Je jasné, že pri takomto čítaní sa zmení aj čítanie operátora I. Pri novom čítaní $I\phi$ znamená, že nie je možné, aby z presvedčení

¹² V literatúre sa v tejto súvislosti hovorí o „implicitných“ presvedčeniach.

daného agenta vyplývalo ϕ v prípade, že ϕ je nepravdivé. Inými slovami, $I\phi$ je pravdivé v x , ak v ľubovoľnom „dosiahnuteľnom“ svete y má daný agent také presvedčenia, z ktorých vyplýva ϕ iba v prípade, ak je ϕ v danom y skutočne pravdivé. Ide o slabšie chápanie neomylnosti, no iste ide o určitú formu neomylnosti. Označme ju neomylnosť*.

Tvrdenie 3.8. Nasledujúce formuly patria do $\text{Log}(MBel_x)$:

- (a) $BIp \rightarrow Ip$
- (b) $(BIp \wedge Bp) \rightarrow p$

Ide o prekvapivo silné tvrdenie. Podľa bodu a) sa nemôže stať, že by z mojich konzistentných presvedčení vyplývalo, že som neomylný* vzhľadom na p a zároveň by som nebol neomylný* vzhľadom na p . Ak z mojich presvedčení vyplýva, že som neomylný* vzhľadom na p , tak som skutočne neomylný* vzhľadom na p .

Podľa bodu b) je p pravdivé v prípade, že z mojich konzistentných presvedčení vyplýva, že som neomylný* vzhľadom na p a tiež samotné p . Ak z mojich presvedčení vyplýva p a tiež to, že som neomylný vzhľadom na p , tak p je pravdivé.

Predpokladajme teraz, že moje aktuálne presvedčenia sú konzistentné a obsahujú tvrdenie, že som neomylný* vzhľadom na p , spolu s tvrdením p . Ak toto tvrdenie obsahujú, tak z nich samozrejme vyplýva. Inými slovami, práve teraz verím, že z mojich presvedčení (nech by boli akékoľvek, pokiaľ sú konzistentné), nikdy nemôže vyplývať p v prípade, že by bolo nepravdivé, a zároveň verím aj tomu, že p je pravdivé. Potom p je pravdivé.

Na dokázané tvrdenie sa však môžeme pozrieť aj z trocha iného uhla. Bod a) napríklad naznačuje, že žiadna množina presvedčení, z ktorej vyplýva, že daný agent je neomylný* vzhľadom na p , nemôže byť konzistentná, pokiaľ daný agent skutočne nie je neomylný vzhľadom na p . Bod b) vraví okrem iného aj to, že ak sa mýlim vzhľadom na p a zároveň som presvedčený, že som neomylný* vzhľadom na p , tak moje presvedčenia nie sú konzistentné.

Tento dôsledok bodu b) je zaujímavý, pretože sa bežne nedomnievame, že aktuálny stav sveta ($\neg p$) by mohol ovplyvniť konzistentnosť našich presvedčení ($Bp \wedge BIp$).

4 Záver

Cieľom tohto príspevku bolo načrtnúť spôsob, akým je možné formálne skúmať pojem neomylnosti. Tento náčrt bol tiež doplnený o niekoľko výsledkov týkajúcich sa platnosti či neplatnosti rôznych formúl, ktoré reprezentujú tvrdenia o neomylnosti. Niektoré z uvedených výsledkov môžeme charakterizovať ako triviálne, no isto nie všetky.

Medzi netriviálne zistenia môžeme zaradiť napríklad tvrdenia o vzťahu medzi neomylnosťou a výrokovými spojkami (neomylnosť sa neprenáša na zložky konjunkcie, neomylnosť vzhľadom na p neimplikuje neomylnosť vzhľadom na $\neg p$). Tie poukazujú aj na jeden z možných prínosov tohto spôsobu skúmania pojmu neomylnosti: Bez dôkladnejšej formalizácie daného pojmu by možno našej pozornosti unikali.

Niektoré z prezentovaných výsledkov (najmä v časti 3.5) môžu do určitej miery budiť dojem kontroverznosti. Odporúčam ich chápať ako výsledky štúdia konzistentnosti množín presvedčení, ktoré môžu obsahovať presvedčenia o neomylnosti. Detailnejšia úvaha o predpokladaných vlastnostiach operátora možnosti M môže odhaliť, že nie sú také prekvapivé, ako sa na prvý pohľad zdá.

Uvedené výsledky sú však v každom prípade iba začiatkom a načrtnutý smer práce môže pokračovať rôznymi smermi. Zaujímavé by napríklad bolo skúmanie jazyka, v ktorom I vystupuje ako nedefinovaný operátor (i -jazyk). Keď sa zameriame na určitú triedu modelov, ako by vyzerala axiomatizácia množiny formúl (i -jazyka), platných v danej triede modelov?

Zaujímavé by bolo aj skúmanie operátora neomylnosti v kontexte iných druhov neklasických logík. Všimnime si totiž, že operátor I sme definovali v podstate pomocou pojmu striktnej implikácie¹³ ($I\phi$ vtt $B\phi$ striktné implikuje ϕ). Inú verziu tohto operátora však môžeme definovať aj pomocou iného druhu implikačnej spojky, napríklad pomocou relevantnej implikácie. Východiskom by mohla byť práca na epistemickej interpretácii relevantných logík, pozri Bílková – Majer – Peliš – Restall (2010).

¹³ Striktná implikácia sa definuje ako nevyhnutne pravdivá materiálna implikácia.

Dodatok: dôkazy tvrdení

V tomto dodatku uvádzame dôkazy tvrdení uvedených v hlavnom texte. Pri väčšine z nich ide o jednoduché cvičenia v modálnej logike.

Dôkaz tvrdenia 3.1. a) Vezmime si ľubovoľný základný model a nejaký svet x . Predpokladajme, že $V(Bp \wedge Ip, x) = 1$. To znamená, že $V(Bp, x) = V(Ip, x) = 1$. Podľa pravdivostných podmienok pre formulu Ip to však znamená, že $V(Bp \rightarrow p, y) = 1$ pre všetky také y , že Rxy . Z reflexívnosti relácie R však vyplýva, že $V(Bp \rightarrow p, x) = 1$. Platí však $V(Bp, x) = 1$, teda aj $V(p, x) = 1$.

b) Vyplýva z faktu, že každá formula tvaru $L\phi \rightarrow LL\phi$ patrí do $\text{Log}(MBel_E)$ (pozri formulu (8) a nasledujúci komentár).

c) Vyplýva z faktu, že každá formula tvaru $ML\phi \rightarrow L\phi$ patrí do $\text{Log}(MBel_E)$ (pozri formulu (10)).

d) Vyplýva z bodov a) a c).

Dôkaz tvrdenia 3.2. Stačí si vziať také $W = \{x\}$, v ktorom $V(p, x) = 0$ a $V(Bp, x) = 0$.

Dôkaz tvrdenia 3.3. Predpokladajme, že $W = \{x\}$, R je univerzálna na W , $V(p, x) = 1$ a $N(x) = \{W\}$. Jasne platí, že $V(Bp \rightarrow p, x) = 1$, aj $V(B\neg p \rightarrow \neg p)$. Teda $V(Ip, x) = V(I\neg p, x) = 1$.

Dôkaz tvrdenia 3.4. Majme $W = \{x, y\}$, R je univerzálna na W , $V(p, x) = 1$, $V(p, y) = 0$ a $N(x) = N(y) = \{\{y\}\}$. Jasne platí, že $V(B\neg p \wedge p, x) = 1$, aj $V(Bp \rightarrow p, x) = V(Bp \rightarrow p, y) = 1$. Teda $V(Ip, y) = V(\neg I\neg p, y) = 1$.

Dôkaz tvrdenia 3.5. Predpokladajme, že $V(Ip, x) = V(Iq, x) = 1$ pre nejaké x v nejakom M , no $V(I(p \wedge q), x) = 0$. To znamená, že $V(Bp \rightarrow p, y) = V(Bq \rightarrow q) = 1$ pre všetky také y , že Rxy . Na základe výrokovej logiky teda $V((Bp \wedge Bq) \rightarrow (p \wedge q), y)$ pre každé takéto y . Podľa druhého predpokladu však existuje také y , že Rxy a $V(B(p \wedge q), y) = V(\neg(p \wedge q), y) = 1$. To by však viedlo k sporu iba vtedy, ak by sme v rámci konjunktívnych modelov mohli z $V(B(p \wedge q), y) = 1$ odvodiť $V(Bp \wedge Bq, y) = 1$. To však nemôžeme.

Dôkaz tvrdenia 3.6. a) Z predpokladu nepravdivosti danej formuly v nejakom x vyplýva, že existuje také y , že Rxy a zároveň $V(Bp, y) = 1$ a $V(B(p \wedge q), y) = 0$, alebo existuje také z , že Rxz a $V(Bq, z) = 1$ a $V(B(p \wedge q), z) = 0$. Tieto dôsledky však nie sú sporné.

b) Z predpokladu nepravdivosti danej formuly v nejakom x vyplýva, že existujú také y, z , že Rxy a Rxz a zároveň $V(Bp \wedge \neg p, y) = V(B(p \vee q) \rightarrow (p \vee q), y) = 1$ a $V(Bq \wedge \neg q, z) = V(B(p \vee q) \rightarrow (p \vee q), z) = 1$. Z toho spor nevyplýva, stačí napríklad to, aby $V(q, y) = V(p, z) = 1$.

c) Z predpokladu nepravdivosti danej formuly v nejakom x vyplýva, že existuje také y , že Rxy a zároveň $V(B(p \vee q) \wedge \neg p \wedge \neg q, y) = 1$, pričom platí aj $V(Bp \rightarrow p, y) = 1$ alebo $V(Bq \rightarrow q, y) = 1$. Z týchto dôsledkov by vyplýval spor, iba ak by sme z predpokladu pravdivosti $B(p \vee q)$ v y mohli odvodiť $Bp \vee Bq$. To však jasne nemôžeme. Stačí model nastaviť tak, aby $V(Bp, y) = V(Bq, y) = 0$.

Dôkaz tvrdenia 3.7. Pozri dôkaz tvrdenia 3.5. V regulárnych modeloch platí predpoklad, ktorý potrebujeme na odvodenie sporu.

Dôkaz tvrdenia 3.8. a) Predpokladajme, že pre nejaké x máme $V(BIp, x) = 1$. To znamená, že $|Ip| \in N(x)$. Podľa vlastností extrémnych modelov teda $|\neg Ip| \notin N(x)$. Teda zároveň nie je pravda, že $\text{ran}_R(x) \subseteq |\neg Ip|$. To znamená, že existuje také y , že Rxy a $V(Ip, y) = 1$. Platí teda $V(MIp, x) = 1$. Na základe bodu c) tvrdenia 3.1 dostávame $V(Ip, x) = 1$.

b) Vyplýva z bodu a) tohto tvrdenia a z bodu a) tvrdenia 3.1.

*Katedra logiky a metodológie vied
Filozofická fakulta
Univerzita Komenského v Bratislave
Šafárikovo nám. 6
814 99 Bratislava
Slovenská republika
sedlar@fphil.uniba.sk*

Literatúra

- BÍLKOVÁ, M. – MAJER, O. – PELIŠ, M. – RESTALL, G. (2010): Relevant Agents. In: Beklemishev, L. – Goranko, V. – Shehtman, V. (eds.): *Advances in Modal Logic 2010*. London: College Publications, 22-38.
- BLACKBURN, P. – DE RIJKE, M. – VENEMA, Y. (2001): *Modal Logic*. Cambridge: Cambridge University Press.
- BLACKBURN, P. – VAN BENTHEM, J. – WOLTER, F. (eds.) (2006): *Handbook of Modal Logic*. Amsterdam: Elsevier.
- FAGIN, R. – HALPERN, J. Y. – MOSES, Y. – VARDI, M. Y. (1995): *Reasoning About Knowledge*. Cambridge (Mass.): MIT Press.
- FAGIN, R. – HALPERN, J. Y. (1988): Belief, Awareness, and Limited Reasoning. *Artificial Intelligence* 34, 39-76.
- HINTIKKA, J. (1962): *Knowledge and Belief. An Introduction to the Logic of the Two Notions*. Ithaca: Cornell University Press.
- HUGHES, G.E. – CRESSWELL, M.J. (1996): *A New Introduction to Modal Logic*. London – New York: Routledge.
- CHELLAS, B. (1980): *Modal Logic. An Introduction*. Cambridge: Cambridge University Press.

- MAKINSON, D. (1965): The Paradox of the Preface. *Analysis* 25, č. 6, 205-207.
- MEYER, J. -J. CH. - VAN DER HOEK, W. (1995): *Epistemic Logic for AI and Computer Science*. Cambridge: Cambridge University Press.
- MEYER, J. -J. CH. (2001): Epistemic Logic. In: Goble, L. (ed.): *The Blackwell Guide to Philosophical Logic*. Oxford: Blackwell Publishing, 183-202.
- PEREGRIN, J. (2004): *Logika a logiky*. Praha: Academia.
- SEDLÁR, I. (2011): V čom sa nemôžete mýliť? *Organon F* 18, č. 3, 351-362.
- SZOMOLÁNYI, J. (2003): Filozofické otázky logiky V. *Organon F* 10, č. 3, 328-340.
- VAN BENTHEM, J. (2010): *Modal Logic for Open Minds*. Stanford: CSLI Publications.

O realistických a antirealistických tendenciách v ontológii predmetov¹

Martin Schmidt

Univerzita Mateja Bela v Banskej Bystrici

Abstract: The paper identifies basic requirements that should be answered by an object ontology. It sheds light on the two major approaches to objects – reductive and non-reductive – in order to expose their difficulties in satisfying those requirements. The aim of the paper is to uncover realistic and antirealistic tendencies in both approaches.

Keywords: objects, ontology, substratum theory, bundle theory, realism, antirealism.

Úvod

V stati Schmidt (2010) som obhajoval tézu, podľa ktorej tradičná metafyzika nedokáže ponúknuť zástancom vedeckého realizmu zaujímavú koncepciu predmetu. Nasledovala kritika z pera P. Voleka (2011), ktorá diagnostikovala môj pesimizmus len ako dôsledok prehnaného fyzikalizmu. V niečom s P. Volekom súhlasím (Schmidt 2011), ale to teraz nie je podstatné. Postupne som si začal uvedomovať, že v diskusii sa nám prelínajú dve odlišné doktríny, pričom nie je jasné, voči ktorej/v prospech ktorej argumentujeme. Hovorím o ontológii predmetov a predmetovej metafyzike. *Pod ontológiou predmetov rozumieme skúmanie konštitúcie jednotlivín. V texte rozdelíme tieto koncepcie na redukčné (ide najmä o teóriu zhlukov) a nereduktívne (najmä teória sub-*

¹ This work has been funded by *On What There Is: Varieties of Realism and Their Influence on Science-Religion Dialog*, sponsored by the Metanexus Institute on Religion and Science, with the generous support of the John Templeton Foundation, and by KEGA 012UMB-4/2011.

strátov). Pod *predmetovou metafyzikou* však mienim doktrínu, ktorá považuje *nezávislú jednotlivinu* (nech už je jej ontologická konštitúcia akákoľvek) za *základný stavebný prvok reality*. Predmety sú podľa tejto metafyziky primárne, ostatné kategórie im len sekundujú: Vlastnosti sú vlastnosťami *predmetov*; vzťahy sú vzťahmi medzi *predmetmi*; udalosti sú zmenami *predmetov* a takto by sme mohli pokračovať.² Pritom ide o dve rozdielne doktríny: Ontológia predmetov odpovedá na otázku, čím predmet je, predmetová metafyzika hovorí o vzťahu predmetov voči iným ontologickým kategóriám (napr. vlastnostiam, vzťahom a udalostiam). V diskusii s P. Volekom sa však tento rozdiel prehliada. Lenže on sa prehliada bežne, resp. nepoznám literatúru, ktorá by s touto distinkciou explicitne pracovala. Vysvetlení je viacero. Tradičné prístupy sa snažia vysvetliť ontologickú konštitúciu predmetu izolovane od iných kategórií. Predpokladom je teda ontologicky nezávislý predmet, ktorý je však rovnako predpokladom predmetovej metafyziky. Zrejme platí aj opak. Dôvod, prečo sa tradičné koncepcie predmetov koncentrujú na ontologicky izolovaný predmet, môže byť v tom, že ich autori vopred prioritu predmetov predpokladajú (t. j. predpokladajú predmetovú metafyziku) a ich ontologickou konštitúciou chcú tento predpoklad len potvrdiť. Veď uvážme, že dejiny metafyziky sú z veľkej miery dejinami predmetovej metafyziky. Alternatívne projekty sa ešte stále považujú za metafyzických exotov (napríklad metafyzický monizmus, holizmus, ale aj štrukturalizmus).

V texte preskúmate potenciál ontológie predmetov a predmetovej metafyziky bez akýchkoľvek fyzikalistických predpokladov (s výnimkou príkladov). Preukáže sa, že ontológie predmetov dokážu výrazne zvýšiť svoj potenciál, ak sa vymania z dosahu predmetovej metafyziky, t. j. ak odmietnu prioritizovanie predmetov na úkor iných ontologických kategórií. Cieľom je dokázať, že v ontológii predmetov nejestvuje zhoda o miere ich objektívnej a nezávislej existencie. Jednotlivé teórie siahajú od predmetového realizmu až k miernemu antirealizmu. Pod nálepku ontológie predmetov sa dokonca pestujú aj antirealistické prístupy. Tieto skutočnosti významne napomáhajú kritikom, ku ktorým sa osobne hlásim, podľa ktorých predmety (ako objektívne a nezávislé entity) vôbec nejestvujú.

² Za prototyp takejto metafyziky sa považuje Aristotelovo učenie o substancii, ku ktorému sa ešte dostaneme.

1 Čo má koncepcia predmetov vysvetliť?

Náš postup bude jednoduchý. Najprv vyberieme aspekty, ktorých objasnenie od ontológie predmetov očakávame. Potom prejdeme k reprezentatívnym koncepciám a vo svetle už vybraných aspektov posúdime ich potenciál. Čitateľa však musím upozorniť, že výber aspektov a samotných koncepcií nebude predmetom osobitnej argumentácie. Snahou bolo vybrať tie momenty, ktoré sa v súvislosti s predmetmi najviac diskutujú, a to isté platí aj o koncepciách, od ktorých očakávame odpovede. Ak by existovali objektívne kritériá výberu, určite by som ich využil, ale o žiadnych neviem. Rovnako som si vedomý, že niektoré aspekty explikácie pojmu „predmetovosti“ v príspevku chýbajú, no neobávam sa, že by mohli do následnej argumentácie vniesť niečo radikálne odlišné. To sa v príspevku postupne objasní. Takže čo očakávame od ontologickej koncepcie predmetu?

Aspekt A. V prvom rade vysvetlenie toho, čo robí predmet predmetom, resp. čo odlišuje predmety od iných ontologických kategórií (najmä od vlastností). Je až zarážajúce, aký problém toto vysvetlenie obhajcom predmetov spôsobuje. Ak trochu prebehneme výklad, problém je v tom, že v koncepciách predmetov prevládajú tzv. reduktívne teórie, ktoré, obrazne povedané, redukovujú predmety na niečo iné (najčastejšie je to redukcia na zhluk vlastností). To iné však už predmetom nemôže byť (inak by redukcia nebola redukciou), a preto každá reduktívna teória predmetov nevyhnutne obsahuje prekážku, ktorá jej objektívne bráni identifikovať to špecifické, čo robí predmet predmetom. Nereduktívny prístup, na druhej strane, v tomto ohľade nezlyhá. Otázna je však cena, ktorú za túto výhodu platí, ale k tomu sa rovnako dostaneme až neskôr.

Aspekt B. Ďalším aspektom, s ktorým si musí teória predmetov poradiť, je identita a identifikácia. Klasika je všeobecne známa: Predmety (jednotliviny), na rozdiel od univerzálií, charakterizuje jedinečný výskyt. Následne platí, že správna teória musí viesť vysvetliť, čo garantuje jedinečnosť predmetov, t. j. v čom spočíva ich neopakovateľná identita. Situácia sa veľmi skomplikovala nástupom fyziky elementárnych častíc. Ontologický štatút častíc pripomína predmety, ale s tým rozdielom, že postulát jedinečnosti už v ich prípade neplatí. Presnejšie, existuje viacero (numericky odlišných) častíc, ktoré sú však kvalitatívne nerozlišiteľné. Treba priznať, že s takouto možnosťou filozofia vôbec nepočítala. Napríklad podľa zákona identity nerozlišiteľných platí,

že kvalitatívna identita implikuje numerickú identitu. Lenže napríklad elektróny predstavujú jasný protipríklad, lebo v ich prípade nastáva kvalitatívna identita, nie však numerická identita – kvalitatívne nerozlišiteľných elektrónov je veľa.³ Pod tlakom týchto skutočností sa začínajú rozlišovať dve samostatné kategórie: predmety-jednotliviny a predmety-individuá.⁴ Pod predmetom-jednotlivinou rozumieme entitu, ktorá má (na rozdiel od univerzálií) konkrétny výskyt: Jej existencia začala v určitom čase na určitom mieste, pretrváva a nakoniec, v konkrétnom čase a na konkrétnom mieste, aj zaniká. V prípade predmetov-individuá dodáme, že ich výskyt a existencia sú jedinečné a neopakovateľné. Presnejšie, ak je X individuom, musí byť jednotlivinou a zároveň musí platiť, že nejestvuje žiadna iná jednotlivina, ktorá by bola Xom , t. j. ktorá by mala presne tie isté monadické kvality ako X . Čo sa tým presne mieni? Uvažujme o dvojacom type vlastností – tie, ktoré má predmet nezávisle od svojho aktuálneho stavu, a tie, o ktorých to neplatí. Príkladom prvého typu vlastností je napríklad elektrický náboj elektrónu a príkladom druhého typu je jeho aktuálna poloha. V prvom prípade ide o monadickú vlastnosť, v druhom o polyadickú, t. j. relačnú (poloha elektrónu je určená jeho vzťahom k časopriestoru).⁵ Ak testujeme, či je X individuom, tak zisťujeme, či jestvuje ešte iná jednotlivina, ktorá by mala presne tie isté monadické vlastnosti ako X , t. j. tie isté vlastnosti prvého typu. Ak by takáto jednotlivina existovala, X by nemohlo byť neopakovateľným individuom. Takže každé individuum je jednotlivinou, opak však neplatí. Predpokladom je, aby dobrá koncepcia predmetu pracovala bez ohľadu na to, či máme do činenia s predmetom ako jedinečným individuom, alebo len s predmetom-jednotlivinou. Ak je predmet individuom, potom budeme od teórie požadovať vysvetlenie aj toho, čo ho odlišuje od iných individuá (aká je jeho identita). Ak bude predmet len jednotlivinou, potom toto vysvetlenie nepožadujeme. Len

³ Táto diskusia je zmapovaná vo štvrtej kapitole v French - Krause (2006).

⁴ Aby sme výklad zbytočne nekomplikovali, v prípade predmetov uvažujeme len o kontinuantoch, t. j. o takých predmetoch, ktoré pretrvávajú v čase. Rovnako sa obmedzíme len na predmety v aktuálnom svete, čím vylúčime komplikácie, ktoré by nám spôsobili predmety v iných možných svetoch (najmä v súvislosti s práve skúmaným aspektom).

⁵ Vo filozofii sa vlastnosti nezávislé od stavu predmetu označujú ako „vnútorné“, „inherentné“, „esenciálne“ a pod. Vonkajšie vlastnosti, relačné a akcidentálne majú blízko k tým vlastnostiam, ktoré fyzika považuje za závislé od aktuálneho stavu systému, v ktorom sa daný predmet vyskytuje.

zdôrazním, že splnenie tejto úlohy je motivované najmä tým, že bežné predmety sú individuá a my chceme vedieť, čo presne ich jedinečnosť konštituuje. Ak v tomto ohľade koncepcia principiálne nefunguje, je to jej vážny nedostatok.

Z tohto dôvodu sa v príspevku vyhneme riešeniam, ktoré považujú identitu za primitívny, ďalej neanalyzovateľný koncept. Hovoríme o prístupoch, ktoré sa opierajú buď o existenciu individuovej vlastnosti (tzv. *primitive thisness*), alebo individuovej podstaty (*haecceita*), ktoré identitu a identifikáciu trivializujú.⁶ Na ilustráciu, skutočné vysvetlenie identity a identifikácie predmetu *X* sa nemôže opierať o existenciu jedinečnej vlastnosti „byť predmetom *X*“. To je triviálne a rovnako triviálnu odpoveď by sme dostali aj od zástancov individuových podstat. Rovnako sa vyhneme, ale z iných dôvodov, identite a re-identifikácii jedného a toho istého individua v rôznych časových okamihoch. Ide o rozsiahlu a komplikovanú tému, ktorej adekvátny rozbor, v spojitosti s hlavnou témou tohto textu, prekračuje rozsah state, a preto si ju nemôžeme dovoliť.⁷ Tým však zároveň priznávam, že výsledky nášho skúmania nemôžu byť absolutizované a to aj s ohľadom na to, že skúmané doktríny neprejdú testom identity v čase.

Aspekt C. Predmety plnia dve základné funkcie, ktoré musí teória dostatočne objasniť. Hovoríme o úlohe predmetov ako subjektov predikácie a s tým súvisiacou funkciou nositeľov vlastností a vzťahov. Aj napriek zjavnej súvislosti nedokážeme obe funkcie zredukovať na jednu. Skúmané koncepcie, z dôvodu rozdielnych prístupov k ontologickej konštitúcii predmetov, ponúknu viaceré vysvetlenia ukotvenia atribútov v predmetoch, pričom niektoré z nich majú problémy v jazykovej rovine. Inak povedané, jestvujú také ontologické koncepcie, ktoré uspokojuvito objasnia rolu predmetov ako nositeľov atribútov, ale zlyhávajú vo vzťahu k predikácii. Z tohto dôvodu musíme obe funkcie posudzovať osobitne. V tejto súvislosti by sme mali spomenúť aj funkciu predmetov ako jednotiacich elementov, ktorých úlohou je konštituovať relatívne stabilné celky. Ak by sme z nášho ontologického repertoáru

⁶ Kritický prehľad týchto a podobných prístupov je v Lowe (2003, 87).

⁷ Adekvátny prístup by si vyžadoval explikáciu spôsobov pretrvávania individuí v čase (tých je niekoľko) a ich aplikáciu na ontológiu predmetov. Špecifický problém by spôsobovalo pretrvávajúce štruktúry (v prípade štrukturalistického chápania predmetov). To sú však témy na samostatné štúdie.

predmety vyškrtli, podľa kritikov by sme nedokázali adekvátne objasniť, prečo vo svete jestvujú relatívne stabilné celky a nie len napríklad rozptýlené vlastnosti. Je všeobecne známe, že najmä reduktívne koncepcie túto funkciu dostatočne nevysvetľujú.

Netvrdím, že ide o vyčerpávajúci a detailný súhrn očakávaných aspektov ontológie predmetov, no verím, že pre naše momentálne ciele je postačujúci. Ako bolo avizované, druhým krokom nášho výkladu bude testovanie vybraných ontologických koncepcií predmetov a to na základe ich potenciálu vysvetliť diskutované aspekty A, B a C. Predtým si však musíme v spleti ontológií predmetov urobiť aspoň čiastočný poriadok.

2 Reduktívne a nereduktívne prístupy k predmetom

Začnime kľúčovým delením ontológií predmetov na reduktívne a nereduktívne. Len zopakujeme, že reduktívne prístupy redukovujú predmety na inú ontologickú kategóriu, čo však predstavuje veľmi rizikovú stratégiu: Ide vôbec o teórie *predmetov*? Ontologická redukcia nie je zrejme nič iné ako ontologická eliminácia (Robinson 2009, časť 3.1). Ak redukovujeme X na Y, tak, ontologicky povedané, existuje len Y. Ak teda redukovujeme predmety na niečo iné, potom existuje to niečo iné (napr. vlastnosti), ale predmety strácame. To by reduktívne prístupy – ako teórie predmetov – úplne diskvalifikovalo, my však budeme od tohto rizika zatiaľ odhliadať. Aj tak je ale evidentné, že ideál ontológie predmetov je v nereduktívnych teóriách, lenže aj v ich prípade máme na výber. Na výber máme medzi nereduktívnymi teóriami, ktoré blokujú akúkoľvek ontologickú analýzu predmetu, a tie, ktoré analýzu umožňujú. Príkladom prvého typu je nominalizmus. Ten je voči predmetom nereduktívny (naopak, ostatné ontologické kategórie od predmetov závisia), no pojem predmetu považuje za primitívny, ďalej už neanalyzovateľný. V ontológii to znamená, že nominalistické predmety sú ontologicky jednoduché, neštruktúrované a teda diskusia o ich konštitúcii je bezpredmetná (Loux 1998, 94). Na strane druhej je tu napríklad teória substrátov. Tá vidí predmet dualisticky: kvalitatívne holý, neredukovateľný substrát (ten robí predmet predmetom) a vlastnosti. K detailom tejto koncepcie sa dostaneme, teraz je podstatné to, že jej konštituovanie predmetu hovorí o štruktúre, ktorú môžeme podrobiť – na rozdiel od nominalizmu – ďalšiemu skúmaniu. A teraz to podstatné: Nereduktívne objektové ontológie, ktoré neumožňujú ďalšiu

ontologickú analýzu predmetu, nás v texte zaujímať nebudú. Dôvod je v tom, že aspekty A, B a C, ktoré použijeme na testovanie úspešnosti jednotlivých ontológií, nedokážeme na tieto teórie korektne aplikovať. Podľa nich je kategória predmetu ontologicky neanalyzovateľná, lenže otázky ohľadom identity, vzťahu k vlastnostiam a pod. si vyžadujú pohľad do „vnútra“, čo napríklad nominalisti odmietnu. Netvrdím, že nominalizmus je z pohľadu objektivej ontológie stratený, tvrdím len to, že v našej diskusii nemá vzhľadom na jeho postoj k predmetom čo ponúknuť.

Najprv sa zameriame na nereduktívne teórie (samozrejme už len na tie, ktoré analýzu akceptujú), potom prejdeme k reduktívnym. V prvom prípade pôjde o rôzne verzie teórie substrátov, v druhom najmä o teórie zhlukov.

3 Nereduktívne ontológie

Prototypom nereduktívneho prístupu k predmetom je teória substrátov a prototypom teórie substrátov je lockeovská dualistická koncepcia:

Naša idea, ktorej dávame *všeobecné* meno *substancia*, je iba predpokladaným, ale neznámym nositeľom kvalít, o ktorých zisťujeme, že existujú. [...] Tohto nositeľa nazývame *substanciou*, čo v pravom zmysle slova v bežnej angličtine znamená to, „čo stojí pod“, alebo môže znamenať aj to, „čo podopiera“. (Locke 1983, 237-238)

Pod výrazom *substancia* sa mieni substrát, ktorý je kvalitatívne holý, a teda podľa Locka nepoznateľný. Spolu s ním je predmet konštituovaný aj vlastnosťami. Tie substrát buď exemplifikuje (ak sú vlastnosti univerzálne), alebo nesie (ak sú vlastnosti trópy⁸). Sám substrát však žiadne vnútorné (netriviálne) vlastnosti nemá. Na základe tejto charakteristiky môžeme do skupiny substrátov zaradiť aj také ontologické kategórie, akými sú holé jednotliviny a aj tzv. *thin particulars* (Robinson 2009, časť 3.2.2.). Úspechy a neúspechy celej koncepcie sú úspechmi a neúspechmi substrátov: Jedinečný substrát má zabezpečiť identitu

⁸ Trópy sú konkrétne časopriestorové výskyty vlastností, o ktorých platí kvalitatívna a aj numerická odlišnosť. Nepopierateľne pripomínajú viac jednotliviny ako univerzálne (Loux 1998, 79-80), čo môže byť zdrojom konceptuálnych nejasností (MacDonald 1998, 333). K trópom sa v texte ešte niekoľkokrát vrátime.

a identifikáciu jednotliviny/individua ako celku; substrát má byť nositeľom vlastností; vďaka substrátu majú jestvovať relatívne stabilné celky a najmä, substráty sú radikálne odlišné od iných ontologických kategórií, a tak predstavujú predmet *per se*. Aká je však realita? Napriek deklarovanej jedinečnosti jestvujú pochybnosti o tom, či substráty rolu identifikátorov naozaj zvládnu (aspekt B). Nejestvuje totiž nič, čo by odlišilo jeden substrát od druhého (Mertz 2001, 113-117). Všetky substráty majú presne tie isté formálne (triviálne) vlastnosti (napr. byť jednotlivinou) a čo sa týka netriviálnych vlastností, v ich prípade sú na tom takisto rovnako: Nemajú ani jednu. Inak povedané, pomocou substrátov dokážeme vysvetliť štatút predmetov ako jednotlivín (lebo substráty sú jednotlivinami), no nedokážeme objasniť štatút tých predmetov, ktoré sú aj neopakovateľnými individuami (lebo substráty také evidentne nie sú). Nepochybne ide o paradoxný záver, lebo hlavným motívom teórie substrátov je presvedčenie, že konštitúcia predmetu založená výlučne na vlastnostiach-univerzáliách nedokáže individuáciu vysvetliť. Z tohto dôvodu musí predmet obsahovať aj nekvalitatívnu zložku a tou je údajne neopakovateľný substrát (Loux 1998, 107-108). Ak máme pravdu, potom toto riešenie nie je dobré. V určitom zmysle je však problematické aj vysvetlenie toho, čo robí predmet predmetom (aspekt A). Táto teória to vysvetľuje prítomnosťou substrátu. O ňom však vieme len to, čo je dané v jeho pojme, t. j. v súhrne triviálnych, formálnych vlastností. Inak o substrátoch nič iné nevieme: Nevstupujú do žiadnych kauzálnych interakcií a nevyvolávajú zmeny, ktoré by nám umožňovali substráty, priamo alebo nepriamo, skúmať. Inertné substráty sú v predmete stratené, nepridávajú mu vôbec nič, pričom aj povaha predmetu je určená výlučne jeho vlastnosťami (nie substrátom). Preto aj Locke hovorí o substráte ako o principiálne nepoznateľnom komponente. Je teda na zváženie, či chce ontológia predmetov založiť svoju centrálnu kategóriu na tak problematickej entite. Súhrnom platí, že teória substrátov čiastočne zlyhá v prípade aspektu B a na aspekt A odpovedá problematcky. Aspekt C sa zdá byť obsiahnutý adekvátne. Nepoznám literatúru, ktorá by závažným spôsobom spochybnila úlohu substrátov pri stabilizácii predmetov, predikcii a v ich funkcii nositeľov vlastností. Tým sa však potenciál teórie nevyčerpáva.

Jestvujú minimálne dva spôsoby, ako ju vylepšiť, a oba prichádzajú z ne-predmetových metafyzík, t. j. od tých metafyzík, ktoré odmietajú prioritovať predmety na úkor iných ontologických kategórií. Prvou je použitie spinozovskej tézy o ontologickej neoddeliteľnosti atribútov

od substrátu, druhá možnosť má korene v štrukturalistickej metafyzike.⁹ Pilierom spinozovskej koncepcie je princíp ontologickej neoddeliteľnosti predmetov od modov (Heil 2003, 172-173), t. j. substrátov od vlastností a vzťahov. Oddeliť substrát od vlastností a vzťahov je možné, podľa tohto prístupu, len konceptuálne, a nie ontologicky. Teraz nechcem polemizovať s tým, čo si pod týmto princípom presne predstaviť, pouvažujme len o tom, čo by na teórii substrátov zmenil. V prvom rade by ju zbavil špekulatívnych entít a to preto, lebo základnou ontologickou jednotkou nového prístupu už nebude izolovaný substrát, ale substrát spolu s vlastnosťami. Takýto celok už má potenciál vstupovať do kauzálnych interakcií, vyvolávať zmeny a tým byť poznateľný. Problém je však v tom, že téza o neoddeliteľnosti modov automaticky vyúsťuje do extrémneho monizmu. Tomuto som sa venoval pri inej príležitosti (Schmidt 2010, 649) a preto len veľmi stručne. Ak pod modmi rozumieme aj vzťahy (čo podľa spinozovcov musíme), potom téza o ich ontologickej neoddeliteľnosti od substrátov automaticky zväzuje jeden substrát s tými substrátmi, s ktorými je bezprostredne spojený, a tak začnú tvoriť vzájomne prepojený celok. Lenže aj tieto ďalšie substráty sú spojené s inými substrátmi atď., až kým sa nezastavíme pri existencii jedného jediného celku, ktorý monisti označujú výrazom *Jedno*. Výsledkom je veľmi revizionistická metafyzika s viacerými nedostatkami (Schmidt 2010, 649-650). Oveľa prijateľnejšie vylepšenie spočíva v spojení teórie substrátov s niektorými prvkami štrukturalistickej metafyziky, konkrétne s tézou o kontextuálnom charaktere identity a identifikácie (Esfeld – Lam 2011, 152). Štrukturalisti odmietnu identitu na úrovni izolovaných predmetov, t. j. nezávisle od ich postavenia, úloh a vzťahov v systéme. Podľa nich nemá žiadny izolovaný predmet (ak čosi také vôbec jestvuje) presne určené kritériá identity a identifikácie, lebo tie nadobudne až výskytom v systéme. Identita je tak závislá od externých faktorov, čo však teórii substrátov prináša elegantné riešenie: Hoci sú substráty vnútorne nerozlíšiteľné, ich postavenie v systéme je jedinečné a práve tým sa zabezpečí ich individuácia. Ak by sme teda obetovali vnútornú identitu predmetu v prospech kontextuálnej,

⁹ Monizmus, vrátane Spinozovho monizmu, je asi najstaršou a pravdepodobne aj najdiskutovanejšou doktrínou (mierne) antirealistického prístupu k predmetom. Štrukturalizmus je radikálnejší, ten prioritizuje vzťahy (Esfeld – Lam, 2011, časť 8.2).

čo však predmetová metafyzika odmieta¹⁰, aj teória substrátov by dokázala úspešne vysvetliť aspekt B.¹¹

Pod pojem nereduktívneho prístupu k predmetom sme zaradili tie koncepcie, ktoré pracujú s kategóriou neredukovateľného substrátu. Spája ich neproblematické vysvetlenie C, jedna z verzií (štrukturalistická) dokonca odpovedá aj na B. Problém ostáva s A, ktorého objasnenie sa spolieha na postulát špekulatívnej entity (náprava v podobe spinovského princípu je vzhľadom na hroziaci monizmus riskantná).

4 Reduktívne teórie

Reduktívne prístupy k predmetom predstavujú ešte pestrejšiu plejádu koncepcií ako nereduktívne. Jedna z najstarších reductívnych koncepcií je aristotelovská koncepcia substancie. Substancia ako jednotlivina je hlavným stavebným prvkom reality, no zároveň podlieha redukcií na látku a formu.¹² Ak sa presunieme v čase, dnes sa vážne diskutuje o veľmi radikálnej redukcií, podľa ktorej predmet vzniká distribúciou (fundamentálnych) vlastností na časopriestorové body, resp. oblasti (Campbell 1990, 130). Podľa tohto prístupu je časopriestorový bod predmetom *per se*: Bod je pravým subjektom predikácie, on zabezpečuje individuáciu distribuovaných vlastností, on je garantom existencie stabilných celkov a pod. Ide o veľmi ekonomickú koncepciu, ktorá si trúfne odpovedať na naše aspekty B a C, no ako každý reprezentant reductívneho prístupu k predmetom, ani tento návrh nedokáže objasniť aspekt A (lebo predmety redukuje na iný druh entít). To isté platí aj o redukcií predmetov na vlastnosti, t. j. o teórii zhlukov. Jej sa budeme venovať podrobnejšie. Dôvody sú jednoduché. Hoci je redukcia predmetov na časopriestorové body veľmi zaujímavá a provokatívna, z pohľadu metafyziky ostane len na okraji záujmu (Simons 1994, 370). No a v prípade Aristotela je situácia veľmi komplikovaná. Na jednej strane je tu spomenutá redukcia na látku a formu a na strane

¹⁰ Lebo by sa tým narušilo výsadné postavenie predmetov, keďže by ich identita závisela od vzťahov.

¹¹ Ani štrukturalistický prístup k identite a identifikácii nie je ideálny (Ainsworth 2011). Zlyhá vtedy, ak postavenie predmetu nie je jedinečné, napr. v prípade, ak sa vyskytuje v symetrickej štruktúre.

¹² Kritická diskusia o Aristotelovej ontológii jednotlivín je v Robinson (2009, časť 2.2.2).

druhej nereduktívna ontológia tzv. fundamentálnych jednotlivín, akými sú napríklad živé bytosti a osoby (Loux 1998, 117-118; 129, pozn. 28). Riešenie tohto a ďalších protirečení v Aristotelovej metafyzike však prekračuje rámec nášho príspevku, a preto aj z tohto dôvodu ostávame len pri teórii zhlukov.

Podľa nej je predmet fúziou vlastností, t. j. vzťah vlastnosti k predmetu je vzťahom časti k celku. Žiadna iná zložka okrem vlastností predmet už nekonštituuje. Do radov zástancov patria D. Hume, A. J. Ayer, ale aj B. Russell (hoci len v neskoršom období). Teória je veľmi flexibilná a disponuje širokým manévrovacím priestorom. Zhluky môžu byť tvorené univerzáliami, trópami alebo ich kombináciou. Zhluk môže byť štruktúrovaný na pevné (esenciálne) jadro vlastností a na vlastnosti, ktoré sú v zhluku len náhodne, a rovnako jestvuje aj viacero riešení toho, čo spája vlastnosti do celku. Ide o tradičné témy teórie zhlukov, ktoré nás budú zaujímať len okrajovo.¹³ Teraz sa vráťme k našim aspektom A, B a C. O aspekte A sme už v kontexte reduktívnych prístupov hovorili. Ak redukuje predmety na ontologicky inú kategóriu, tak atribút predmetovosti sotva objasníme. Môžeme len doplniť, že azda jedinou možnou odpoveďou teórie zhlukov je tá, že v prípade predmetov máme do činenia so špecifickou skupinou vlastností a to tých, ktoré sú časťami zhlukov. Nie každá vlastnosť je časťou zhluku a to by malo odlišovať predmety (ako zhluky vlastností) od bežných vlastností. Najčastejšie sa uvažuje o kauzálnych interakciách medzi vlastnosťami, ktorých výsledkom by mali byť relatívne stabilné celky/predmety. Isteže, takéto odpovede nie sú uspokojujúce. Ved' atribút „byť časťou zhluku“ je opäť len vlastnosťou (vlastnosťou druhého rádu) a kauzálne interakcie sotva stotožníme s predmetmi *per se*. Situácia sa komplikuje aj tým, že zástancovia teórie nedávajú jednoznačnú odpoveď ohľadom tvorby zhlukov (Loux 1998, 99; Cleve 1985, 266-267; Simons 1994, 369) a aj z tohto dôvodu nedokážeme teóriu adekvátne posúdiť. Teórie zhlukov majú nedostatky aj v prípade aspektu B. Ak jednotliviny tvoria zhluky univerzálií, tak automaticky jestvuje možnosť existencie ich duplikátov – kvalitatívne identických, ale numericky odlišných predmetov.¹⁴ V prípade ne-individuá, akými sú napr. elektróny, je všetko v poriadku. Ak však zvažujeme individuá, situácia sa mení. Teória zhlukov sa môže oprieť len o kvalitatívnu odlišnosť (napr. o existenciu jedinečnej

¹³ Ich kritický výklad je v Cleve (1985, 266-268).

¹⁴ Substráty mali presne tomu zabrániť.

vlastnosti v danom zhluku/predmete), čo je však v prípade univerzálií problematická odpoveď: Ved' univerzálie sú preto univerzáliami, lebo ich môžu exemplifikovať *viaceré* predmety. Zástancovia však vidia záchranu v nahradení univerzálií trópami, t. j. jedinečnými výskytmi vlastností v čase a v priestore.¹⁵ Keďže trópy sú skôr jednotliviny-individuá ako univerzálie (Loux 1998, 80), ich výskyt v zhlukoch/predmetoch by mal zablokovať existenciu duplikátov, a tak umožniť ich identifikáciu a odlíšenie od iných (hoci aj veľmi podobných) zhlukov/predmetov. Tým by mohla aj táto teória odpovedať na aspekt B. To, či ide o uspokojivú odpoveď, je iná vec. Ved' problém ohľadom neopakovateľnej individuality predmetu sa len presunul na problém neopakovateľnosti trópov, čo je rovnako predmetom rozsiahlej diskusie (Loux 1998, 80). Situácia by sa zlepšila, ak by sme podobne ako v prípade substrátov neuvažovali o identite „zvnútra“ predmetu, ale z pohľadu jeho vzťahov, úloh a postavenia v systéme, t. j. štrukturalisticky. Po tejto zmene by bola možná dokonca aj individuácia elektrónov (a to bez pomoci kontroverzných trópov): Hoci v ich prípade ide o zhluky tých istých monadických vlastností (ten istý náboj, pokojová hmotnosť atď.), odlišujú sa vzájomnými vzťahmi. Ide o spomenutú dichotómiu tých vlastností, ktoré nezávisia od stavu predmetu, a tých, ktoré závisia (napr. relatívna poloha a hybnosť). A sú to vlastnosti druhého typu, ktoré nám pri identifikácii inak kvalitatívne neodlíšiteľných predmetov pomôžu, hoci neštrukturalistická metafyzika je tradične zahľadená do vlastností prvého typu (a preto nedokáže v prípade kvalitatívne identických entít adekvátne reagovať).

V prípade aspektu C sa vynárajú ďalšie problémy. Už sme uviedli, že predmet/zhluk je nositeľom vlastností v zmysle vzťahu časti k celku. Ako však vysvetliť úlohu zhluku v predikácii? Ide o často diskutovaný problém. Postavenie vlastností ako konštituentov zhluku automaticky spôsobí, že pripísanie vlastnosti zhluku musí byť buď triviálne pravdivé (ak je daná vlastnosť naozaj prvkom v zhluku), alebo triviálne nepravdivé (ak v zhluku nie je). To núti obhajcov hľadať pevné, esenciálne jadro zhluku a náhodnú perifériu, ktorá by umožňovala aj netriviálnu predikáciu (Cleve 1985, 267; Simons 1994, 376-378). Toto vylepšenie by eliminovalo aj ďalší nedostatok. Predmety vznikajú fúziou

¹⁵ Podľa teórie trópov sú dva výskyty akokoľvek podobných vlastností, napr. jedného a toho istého odtieňa červenej, nevyhnutne numericky odlišné (Loux 1998, 80), t. j. predstavujú dve rozdielne entity.

vlastností čo znamená, že akákoľvek zmena na úrovni vlastností spôsobuje zánik pôvodného zhľuku a vznik nového. Ide o jednu z foriem tzv. mereologického esencializmu (Loux 1998, 226), podľa ktorého je identita predmetu konštituovaná výlučne jeho pravými časťami. Ak sa ich nejaká zmena dotkne, predmet automaticky zaniká. Tento nešťastný dôsledok je možné blokovať tým, že zmena sa týka výlučne okrajových, akcidentálnych vlastností a nie esenciálneho jadra. Až zmena na úrovni jadra vlastností by vyvolala zánik pôvodného predmetu a vznik nového. V kontexte aspektu C je problematickým miestom aj objasnenie dôvodu, prečo niektoré vlastnosti tvoria stabilné zhľuky (predmety), ale tejto téme sme sa venovali v súvislosti s aspektom A.

Záverom môžeme konštatovať, že teória zhľukov, ako hlavný reprezentant redukcionistického prístupu k predmetom, dokáže čiastočne objasniť aspekty B a C, ale aj to len za určitých okolností: Zhľuky musia byť buď konštituované neopakovateľnými trópmi, alebo musíme prejsť k štrukturalistickému chápaniu identity. Osobne sa prikláňam k druhej možnosti. Rovnako je nevyhnutná (z viacerých dôvodov) štruktúracia zhľuku na pevné jadro a akcidentálnu perifériu. Samozrejme, aspekt A ostáva bez vysvetlenia.

5 Substráty, zhľuky a predmetová metafyzika

Bez príklonu k naturalistickej a fyzikalistickej metafyzike (s výnimkou príkladov) sme zistili, že reduktívne a aj nereduktívne prístupy k predmetom, v ich klasických podobách, majú vážne nedostatky. Tie však dokážu relatívne úspešne eliminovať, ak sa vzdajú predmetovej metafyziky, t. j. ak odmietnu prioritu predmetov na úkor iných ontologických kategórií. Je však prekvapením, že hlavný reprezentant nereduktívneho prístupu (teória substrátov) obstál lepšie než jeho reduktívny rival, hoci práve substráty bývajú terčom oveľa ostrejšej kritiky. Substráty bezproblémovo vysvetľujú stabilitu predmetov rovnako aj ich spojenie s vlastnosťami a väčší problém nevyvolávajú ani v súvislosti s predikáciou (aspekt C). V prípade identity stačí zavedenie štruktúrneho prístupu (aspekt B), pričom substráty majú dokonca aj potenciál vysvetliť odlišnosť predmetov od iných kategórií (aspekt A). V poslednom prípade však teória platí privysokú daň, ktorou je ontologický záväzok voči pasívnej, kauzálne neaktívnej entite. Ak však dokážeme akceptovať monizmus, ani toto nemusí byť problém (vďaka postulátu neoddeliteľnosti substrátov od atribútov), čím sa teória sub-

strátov stáva priam ideálnou ontológiou predmetov. Pravdou však je, že obe vylepšenia nemôžu fungovať v rámci jednej teórie a to preto, lebo monisti považujú vzťahy za zdanlivé (akceptujú jedine vzťahy *Jedného* k sebe samému), zatiaľ čo v štrukturalistickej metafyzike sú fundamentálne. Teória zhlukov ťahá evidentne za kratší koniec. V prvom rade nedokáže reagovať na aspekt A a s tým súvisí aj nejasná odpoveď ohľadom stability zhlukov/predmetov (aspekt C). Aby sa teória vyhla ďalším problémom v kontexte aspektu C (najmä s predikáciou a s úlohou predmetu ako nositeľa vlastností), zhluky by mali byť štruktúrované na pevné jadro a na náhodnú perifériu. Rovnako ako v prípade substrátov je potrebná aj akceptácia kontextuálnej identity (aspekt B), čím aj táto teória opúšťa rámec predmetovej metafyziky. Podstatné je, že žiadna iná teória predmetov nás už nemôže prekvapiť. Buď bude neredukcionistická, alebo bude redukcionistická. Prvý tábor je realistický, lebo verí v existenciu „predmetovosti“, ktorá je zásadne odlišná od iných ontologických kategórií. Skúmali sme substráty, ktorých hlavným problémom je kauzálna nečinnosť. Kauzálna interakcia je totiž, v našom svete, garantovaná vlastnosťami popri prípade vzťahmi, no tie nemôžu byť súčasťou nereduktívneho prístupu k predmetom. Preto sa zdá, že neredukcionistický tábor je odsúdený explikovať predmetovosť len pomocou neaktívnych a nepoznateľných entít, keďže kauzálne aktívne entity (vlastnosti a vzťahy) nemôžu byť jeho súčasťou. To je ale odstrašujúca perspektíva. No a redukcionisti predmety až tak vážne neberú (nahrádzajú ich, redukujú ich), a preto ich môžeme smelo označiť za, minimálne, miernych antirealistov.

Z predošlých riadkov je zrejme, že predmetová metafyzika je v porovnaní s ontológiou predmetov v ešte horšej pozícii: Nedokáže adekvátne objasniť svoju kľúčovú kategóriu. Priorizuje predmety, no prioritovanie sa negatívne odrazí v tom, že nedokáže ponúknuť ich vhodnú ontológiu. Paradoxne ju dokážu ponúknuť tie doktríny, ktoré ontologickú prioritu predmetov odmietajú. Takže si musíme vybrať: Buď chceme mať dobrú ontológiu predmetov, alebo predmetovú metafyziku. Nemôžeme mať oboje, čo potvrdzuje myšlienku z úvodu textu, že tradičné spájanie predmetovej ontológie s predmetovou metafyzikou je chybné. Môže byť však predmetová metafyzika úspešná, ak nepracuje s dobrou ontológiou predmetu? Odpoveď by mala byť negatívna, čo odhaľuje vážne napätie vo vnútri predmetovej metafyziky. Iný druh pochybností sprevádza aj ontológiu predmetov. Pokúsím sa to objasniť. V prípade reduktívnych a aj nereduktívnych teórií sme začali ich

klasickou formuláciou, identifikovali ich tradičné problémy a nakoniec sme ponúkli vylepšenia, a to aj za cenu spájania rôznych metafyzických pozícií (menovite štrukturalizmu a monizmu). Výsledkom boli viac-menej úspešne fungujúce koncepcie. Niečo v nich však nesedí. Je evidentné, že vylepšenia sme navrhovali *ad hoc*, len preto, aby teórie *konceptuálne* reflektovali ten alebo onen aspekt ontológie predmetov (aspekty A, B a C). Napríklad štrukturalistické kritériá identity a identifikácie sme uprednostnili len preto, že ich dokážeme aplikovať aj v prípade vnútorne neodlíšiteľných predmetov. Potom však dokážeme povedať len to, že náš návrh je konceptuálne prijateľný a nič viac. To je však veľmi málo: My nechceme vedieť, ktoré ontológie predmetov sú konceptuálne možné, *my chceme vedieť, akú ontológiu/ontológiu jednotliviny majú*. Na to je potrebné ontológiu otvoriť a vystaviť ju aj impulzom z externého prostredia, aby sme zistili, čo z konceptuálne možného je aj reálne. Analytická metafyzika, ako súčasť analytickej filozofie, stávla, aspoň vo svojich začiatkoch, na jazyk a jazykovú analýzu. V tomto duchu mi P. Volek (2011, 367) navrhol, aby som sa na predmety pozeral z pohľadu prirodzeného jazyka a to konkrétne v podobe Strawsonovho projektu deskriptívnej metafyziky. Tento návrh chápem tak, že prirodzený jazyk by mal rozhodnúť, ktoré z konceptuálne možných ontológií predmetov sú tie reálne. V Schmidt (2011) som dokazoval, že takáto zmena v porovnaní s jazykom vedy nič zásadné neprinesie, no teraz prichádzam s ešte radikálnejšou odpoveďou: Jazyková analýza ako arbiter ontologických koncepcií vedie k predmetovému antirealizmu. Pritom je jedno, či hovoríme o jazyku vedy alebo o prirodzenom jazyku. Zvážme Volekov návrh. Strawson vidí predmety najmä ako stabilizátory konceptuálnych schém (Strawson 1997, 66), čím sa však stávajú bytostne závislé od jazyka a konceptov. Toto nie je postoj predmetového realistu. Realista nemôže argumentovať v prospech existencie predmetov na základe toho, že ich potrebuje náš jazyk. Súčasťou realizmu je predsa presvedčenie o ich objektívnej a na jazyku nezávislej existencii. Ak sme zástancov reduktívnych teórií predmetov označili za miernych antirealistov, tak Strawsona musíme označiť za radikálneho antirealistu.¹⁶ Pritom jeho antirealistická pozícia nie je výsledkom voľ-

¹⁶ Tým naznačujem potenciálne protirečenie vo Volekovom texte: V argumentácii v prospech existencie predmetov predkladá ich antirealistické chápanie. Uvedomujem si však, že precízna formulácia by si vyžadovala oveľa detailnejšiu analýzu, do ktorej sa v tomto príspevku nebudeme púšťať.

by *konkrétneho* jazyka, ale jazykovej orientácie ako takej. Takže cesta pre serióznu ontológiu predmetov tadiaľto nevedie. Čitateľ už presne vie, kam mierim: Dobrá a realistická ontológia jednotlivín, ak je niečo také vôbec možné, nemôže ostať len pri apriórnych – konceptuálnych či jazykových – úvahách.

Záver

Bez náležitej dávky naturalizmu a fyzikalizmu vyústi ontológia predmetov buď do apriórnych špekulácií, alebo do antirealistického postoja. Ani jedna z týchto možností nie je z pohľadu predmetového realizmu akceptovateľná. Vráťme sa k identite a identifikácii v rámci teórie zhlukov. Zástancovia zhlukov hľadajú odpoveď aj v neopakovateľných trópoch. Tento návrh však ostane len konceptuálnym trikom, pokiaľ sa nepresvedčíme, že svet naozaj funguje tak, že vlastnosti a vzťahy sú trópy. Takýto dôkaz však zatiaľ nemáme. V úplne inom svetle je štrukturalistická individuácia. V jej prípade sa o dôkazoch už uvažovať dá (napr. holistické správanie tzv. silne korelovaných častíc), a preto nepredstavuje ani jazykový, a ani konceptuálny konštrukt. Samozrejme, môžeme sa mýliť. Ved' holizmus a na ňom založený štrukturalizmus môžu byť spochybnené a budúca metafyzika – podobne ako fyzika – sa bude musieť s touto zmenou vyrovnávať. Takže fyzikalizmus a naturalizmus nie sú v žiadnom prípade zárukou pravdy. Podobne je možné uvažovať aj o delení zhuku na esenciálne jadro a akcidentálnu perifériu, pričom jadro by tvorili vlastnosti nezávislé od stavu systému, perifériu vlastnosti, ktoré od stavu systému závislé sú. Fyzika toto delenie už úspešne používa, čo dáva štruktúracii predmetu silnú a hlavne nezávislú motiváciu. V podobnom duchu sa na teóriu zhlukov pozerá napr. P. Simons (1994, 378-382). On sa rovnako dovoľáva testovania konceptuálne možných ontológií (špeciálne teórie zhlukov) na základe miery ich korešpondencie s fyzikálnymi teóriami. Pre úplnosť treba dodať, že úloha testovania je skôr negatívna: Vylúčiť neadekvátne a neudržateľné ontológie. Druhou stranou mince je však skutočnosť, že fyzikalizmus nie je veľkým priateľom predmetov. Čím viac sa dostávame k fundamentálnej úrovni reality (napr. kvantová teória), tým máme väčší problém predmety nájsť. Takže ani tadiaľto cesta k predmetom nevedie a ja som ich – pre sympatie k fyzikalizmu – aj odmietol (2010). Dokážem však pochopiť aj menej radikálne riešenie. Ktoré? Malo by ísť o ontológiu, ktorá dokáže eliminovať bezbrehé (jazykové a konceptu-

álne) špekulácie, zároveň nemôže ašpirovať na fundamentálne vysvetlenia, no akceptuje predmety. Myslím si, že rozumné pozadie takýchto ontológií by sme našli napr. v biológii a v chémii, t. j. v pomerne silných disciplínach, ktoré predmety (zatiaľ) neeliminujú. To je však téma na iný príspevok.

Katedra filozofie
Fakulta humanitných vied
Univerzita Mateja Bela
Tajovského 40
974 01 Banská Bystrica
Slovenská republika
martin.schmidt@umb.sk

Literatúra

- AINSWORTH, P. (2011): Ontic Structural Realism and the Principle of the Identity of Indiscernibles. *Erkenntnis* 75, 67-84.
- CAMPBELL, C. (1990): *Abstract Particulars*. Oxford: Basil Blackwell.
- CLEVE, J. (1985): Three Versions of the Bundle Theory. In: Laurence, S. – Macdonald, C. (eds.) (1998): *Contemporary Readings in the Foundations of Metaphysics*. Oxford: Basil Blackwell, 264-274.
- ESFELD, M. – LAM, V. (2011): Ontic Structural Realism as a Metaphysics of Objects. In: Bokulich, P. – Bokulich, A. (eds.): *Scientific Structuralism*. Dordrecht: Springer, 143-159.
- FRENCH, S – KRAUSE, D. (2006): *Identity in Physics: A Historical, Philosophical, and Formal Analysis*. Oxford: Oxford University Press.
- HEIL, J. (2003): *From an Ontological Point of View*. Oxford: Oxford University Press.
- LOCKE, J. (1983): *Rozprava o ľudskom rozume*. Bratislava: Pravda.
- LOUX, J. M. (1998): *Metaphysics. A Contemporary Introduction*. London: Routledge.
- LOWE, E. J. (2003): Individuation. In: Loux, M. J. – Zimmerman, D. W. (eds.) (2003): *The Oxford Handbook of Metaphysics*. Oxford: Oxford University Press, 75-95.
- MACDONALD, C. (1998): Tropes and Other Things. In: Laurence, S. – Macdonald, C. (eds.) (1998): *Contemporary Readings in the Foundations of Metaphysics*. Oxford: Basil Blackwell, 365-384.
- MERTZ, D. W. (2001): Individuation and Instance Ontology. *Australasian Journal of Philosophy* 79, 45-61.
- ROBINSON, H. (2009): Substance. In: Zalta, E. N. (ed.): *Stanford Encyclopedia of Philosophy*. Dostupné: <http://plato.stanford.edu/entries/substance/>

- SCHMIDT, M. (2010): Predmety, kauzalita a vedecký realizmus. *Filozofia* 65, č. 7, 643-651.
- SCHMIDT, M. (2011): Je analýza jazyka metafyzikou? *Filozofia* 66, č. 4, 374-378.
- SIMONS, P. (1994): Particulars in Particular Clothing: Three Trope Theories of Substance. In: Laurence, S. – Macdonald, C. (eds.) (1998): *Contemporary Readings in the Foundations of Metaphysics*. Oxford: Basil Blackwell, 365-384.
- STRAWSON, P. F. (1997): *Individuá: esej o deskriptívnej metafyzike*. Bratislava: Iris.
- VOLEK, P. (2011): Predmety ako substancie. *Filozofia* 66, č. 4, 367-373.

Realizmus a princíp empirizmu¹

Miloš Taliga

Univerzita Mateja Bela v Banskej Bystrici

Abstract: There are many variants of scientific realism but few of them account for the role of testing by means of experience in empirical sciences. Moreover, if combined with a kind of empiricism, they usually lead to antirealism. The paper aims to expose causes as well as results of this perplexity. The solution is to revise empiricism quite radically, and to set its marriage with realism on methodological rather than epistemological level.

Keywords: scientific realism, empiricism, problem of induction, petitio principii, positivism, negativism, methodology.

Úvod

Ak chceme byť v oblasti filozofie prírodných (resp. empirických) vied realisti, a tak zastávať určitú verziu vedeckého realizmu, mali by sme si ujasniť, čo je cieľom empirickej vedy a ako je možné racionálne monitorovať jeho dosahovanie. Obhajoba empirickej vedy na rozvíjanie realizmu nestačí, o čom svedčí fakt, že niektoré podoby empirizmu môžu do viesť obhajcu niekam, kde sa možno ani nechcel ocitnúť: k antirealizmu. Hoci predostriem pár príkladov, jadro tohto článku bude postavené skôr na vecnom skúmaní vzťahu empirizmu a realizmu, ako na tézach kľúčových filozofov 20. a 21. storočia. V prvej časti sa stručne pozriem na trýzne P. Godfrey-Smitha, ktoré čitateľovi otvoria dvere do ostatných, dôležitejších častí. Ich úlohou nie je pritom ponúk-

¹ This work has been funded by *On What There Is: Varieties of Realism and Their Influence on Science-Religion Dialog*, sponsored by the Metanexus Institute on Religion and Science, with the generous support of the John Templeton Foundation. Práca bola podporená aj projektom KEGA 012UMB-4/2011. Ďakujem za recenzentské pripomienky.

nuť iba kritické argumenty proti konštruktívnym verziám empirizmu, ale aj konštruktívne načrtnúť negatívny (kritický) empirizmus. Druhá časť článku je venovaná problému indukcie, pretože viera v indukciu stále prežíva v podobe „pozitívnych doktrín“, podľa ktorých má metóda empirickej vedy konštruktívne poslanie. Na zbúranie tohto mýtu je potrebné preskúmať možnosti (a hranice) využitia empirickej evidencie a argumentácie. O to sa pokúšam v časti 3, ktorej závery nie sú pre tradičný empirizmus príliš lichotivé. Preto je v štvrtej časti načrtnutá stratégia ako empirizmus zachrániť a v piatej zasa vysvetlenie, prečo musí mať „negatívnu“ podobu. Šiesta časť vykresľuje pomocou kontrastu medzi „pozitivizmom“ a „negativizmom“ účel a zmysel metód používaných v empirických vedách. Úlohou záveru je navrhnúť stručnú odpoveď na otázku, ako súvisí takto zrevidovaný empirizmus s vedeckým realizmom.

1 Godfrey-Smith o realizme a empirizme

Podľa P. Godfrey-Smitha je jedným zo „skutočných a rozumných cieľov vedy poskytnúť presné opisy (a iné reprezentácie) toho, aká je realita“ (2003, 176). Godfrey-Smith zdôrazňuje, že uvedený názor, ktorý nazýva *vedeckým realizmom*, „zahŕňa aj poskytovanie presných reprezentácií tých aspektov reality, ktoré sú nepozorovateľné“ (tamtiež). Vedecký realista musí teda tvrdiť, že „veda sa môže racionálne snažiť o opísanie skutočnej štruktúry sveta, vrátane jej nepozorovateľnej štruktúry“ (2003, 220). Godfrey-Smith chce však okrem takto formulovaného vedeckého realizmu obhajovať aj určitý typ empirizmu a naturalizmu. A hoci je cieľom jeho knihy ich spojenie (2003, 219), jej čitateľ si všimne skôr problémy, ku ktorým zmienená snaha vedie, ako ich uspokojujúce riešenia. Jednou z príčin týchto ťažkostí je bezmocnosť Godfrey-Smitha podať vysvetlenie, ako súvisí naturalizmus s empirizmom. Odmieťa totiž tradičný empirizmus, podľa ktorého „je skúsenosť jediným zdrojom skutočného poznania sveta“ (2003, 8), čo ho vedie k otázke „Ako by sme mali opísať úlohu skúsenosti?“ (2003, 221). Vzápätí razí módný názor, že „správnym postupom je vykreslenie empirizmu v rámci naturalistického prístupu k filozofii“ (tamtiež), no jeho kresba zostáva bez konkrétnych kontúr. Vo svojej knihe uvádza iba všeobecné mottá typu:

Keď premýšľame o sebe samých z hľadiska biológie a psychológie, môžeme ustanoviť [establish] stále princípy reakcií našej vnemovej

sústavy na objekty a udalosti, ktoré sú od nás vzdialené. Môžeme rozlúštiť, ako nám náš systém vnímania pomáha zdolávať svet. Náš „prístup“ k svetu pomocou myslenia a teórií je veľmi komplikovaným druhom kauzálnej interakcie. (Godfrey-Smith 2003, 222)

Nijako však neukazuje, ako by si údajné ustanovenie takýchto „princípov“ poradilo s problémom indukcie, ktorý mu stojí v ceste (pozri časť 2 nižšie). Na druhej strane, ak sa má na mysli len toľko, že empirická veda je schopná formulovať vysvetlenia našich kognitívnych interakcií s prostredím, ťažko z toho usúdiť, ako jej pri tom pomáha skúsenosť. V oboch prípadoch je výsledok rovnaký: súvislosť medzi vedeckým realizmom a empirizmom je nejasná.

Godfrey-Smith si je vedomý problémov, ku ktorým vedie priamočiare chápanie vedeckého realizmu. Je presvedčený, že „ak by sme vyjadrili vedecký realizmus deklarovaním skutočnej existencie tých entít, ktoré uznáva dnešná veda, a ak by sa naše súčasné teórie ukázali byť nepravdivé, bol by nepravdivý aj vedecký realizmus“ (2003, 175). Preto tvrdí, že „pozícia vedeckého realizmu je zlučiteľná s rôznorodými názormi na spoľahlivosť našich súčasných teórií“ (tamtiež), neargumentuje však, v akom zmysle (ak vôbec) by mohli byť vedecké teórie spoľahlivé. Napokon iba upozorňuje na známy fakt, že „realisti sa musia vzdať myšlienky, podľa ktorej úspech vedy priamo či jednoznačne poukazuje na pravdivosť teórií“ (2003, 179). Namiesto očakávaného vysvetlenia súvislosti medzi vedeckým realizmom a empirizmom však suchopárne uzatvára:

Veda je taktikou vydávania aj tých najväčších teoretických ideí, otázok a sporov napospas testovaniu pomocou pozorovania. Vedeckou stratégiou je vyložiť myšlienky, zasadiť ich do okolitých rámcov a rozvinúť ich tak, aby sa aj v prípade tých najvšeobecnejších a najambicióznejších hypotéz o svete vyhľadávala ich konfrontácia so skúsenosťou. Takýto pohľad na vedu je odrodou empirizmu. (Godfrey-Smith 2003, 223-224)

Čo však empirická veda dosiahne, ak svoje teórie vystavuje skúške pomocou skúsenosti? Aká je úloha experimentov v empirickej vede? Na tieto otázky Godfrey-Smith v (2003) neodpovedá.² Ba čo viac, nezod-

² Ak mám byť presný, v kapitole 14 Godfrey-Smith uvádza určité ťažko dešifrovateľné náznaky, podľa ktorých (ako sa domnievam) by pozorovania mali potvrdzovať testované teórie. Ak by však aj moja domnienka bola

povedanú ponecháva aj otázku, ktorou svoj výskum začal: „Ako empirizmus pomáha filozofii vedy?“ (2003, 8). Nielenže nie je jasné, aký druh empirizmu je v hre, čitateľ nevie ani to, ako súvisí s vedeckým realizmom a s naturalizmom. Avšak práve ich spätosť chcel Godfrey-Smith obhájiť.

2 Empirizmus a problém indukcie

V nadväznosti na predošlú časť sa ponúka záver, že Godfrey-Smith nenašiel východisko zo svojich spleťtých úvah zrejme preto, že nečelil problému indukcie, ktorý sa stavia do cesty takmer každej verzii empirizmu. Táto skutočnosť sa stáva vypuklou najmä pri analýzach filozofických diel členov Viedenského kruhu. Keď napríklad Carnap (1966, 180) uvádza, že „Moritz Schlick raz poznamenal, že empirizmus môže byť definovaný ako názor, podľa ktorého neexistuje syntetické a priori“, netreba dodávať, že ide o dôsledok stotožnenia syntetických výrokov s výroky empirickými. Kritériá rozlišovania medzi analytickými a syntetickými pravdami, vlastné Viedenskému kruhu, viedli jeho členov totiž k názoru, že tie druhé *musia byť* poznávané pomocou skúsenosti, čo fakticky vytvorilo živnú pôdu pre problémy tradičného empirizmu. Filozofi Viedenského kruhu sa tak, možno nechtiac, ocitli aj zoči-voči problému indukcie, ktorý však nedokázali vyriešiť, a preto koketovali s rôznymi antirealistickými názormi. Ukážkou môže byť nasledujúce Carnapovo vyznanie:

Mnohí empiristi ... cítia, že empirista by nemal nikdy používať desivo nebezpečné slovo „pravdivý“. Napríklad Otto Neurath povedal, že hovoriť o zákonoch ako o pravdivých by bolo hriechom proti empirizmu. (Carnap 1966, 213)

Keďže zákony pokrývajú aj tie prípady, ktoré ešte neboli pozorované, je zrejmé, že ich pravdivosť nemôže byť legitímne vyvedená z prípadov už pozorovaných. Problém indukcie však neexistuje len v takto povrchné a nedôsledne sformulovanej podobe. Nejde len o to, že tzv. „zo-všeobecňovanie na základe faktov“ je diskutabilné. Naopak, problém indukcie je hlboko zakorenený v každej doktríne, podľa ktorej môže jeden súbor viet či presvedčení E (evidencia) zdôvodniť iný súbor H

správna, takáto úloha evidencie je z objektívnych príčin (uvedených v častiach 2 a 3 nižšie) neobhájiteľná.

(hypotéza), ktorý však svojím informačným obsahom prekračuje ten prvý. Pritom nezáleží na tom, čo sa presne má na mysli pod slovom „zdôvodniť“, či má ísť zo strany E o poskytnutie záruk alebo len náznakov v prospech či už pravdivosti alebo pravdepodobnosti H, alebo aspoň v prospech názoru, že H sa blíži k pravde, a pod. Dôležité nie je ani to, či symboly „E“ a „H“ zastupujú nejakú evidenciu a hypotézu alebo iné tvrdenia či teórie. Za predpokladu, že H prekračuje svojím obsahom E, je totiž v týchto hmlistých predstavách kľúčovou téza, podľa ktorej má mať E nejaký „pozitívny efekt“ na H, prípadne na vzťah relevantnej skupiny výskumníkov k H. Odtiaľ návrh pomenovať doktríny, osvojujúce si uvedenú tézu, *pozitívnymi smermi* v oblasti filozofie empirických vied.³

Argumentov proti pozitívnym smerom, a teda aj proti možnosti pozitívneho riešenia problému indukcie, je viac. Na tomto mieste spomeniem len jeden, pretože ukazuje, že práve spomínaný „efekt“, ktorý môže mať E na H, nemôže byť pozitívny a súčasne induktívny. K. R. Popper a D. W. Miller navrhli, aby sme tú časť obsahu, ktorým H prekračuje E, chápali ako kondicionál $E \rightarrow H$, pretože sú to práve dôsledky propozície $E \rightarrow H$, ktoré po pridaní k dôsledkom E tvoria obsah H, t.j. množinu všetkých deduktívnych dôsledkov H (pozri Popper – Miller 1987, 580). Otázka, ktorá nás zaujíma, znie, aký dopad má E na $E \rightarrow H$? Miller (1994, 73) odpovedá pomocou dvoch formúl pravdepodobnostného kalkulu. Prvá je teorémou odvodenou zo zákona násobenia a hovorí, že

$$(1) \text{ ak } H \text{ logicky implikuje } E, \text{ potom } p(H) \leq p(H | E),$$

teda, že H je vždy podporené svojimi logickými dôsledkami. Druhá je slabou verziou zákona doplnkov a hovorí, že

$$(2) \text{ ak } p(X | Z) < p(Y | W), \text{ potom } p(\sim Y | W) < p(\sim X | Z).$$

Keď do (1) dosadíme namiesto H propozíciu $E \& \sim H$, čiže negáciu $E \rightarrow H$, výsledkom bude:

$$(3) p(E \& \sim H) \leq p(E \& \sim H | E),$$

a keďže propozícia $E \& \sim H$ je doplnkom propozície $E \rightarrow H$ musí, podľa (2), platiť:

³ Pre stručnosť budem používať na ich označenie aj termín „pozitivismus“, a to aj napriek tomu, že sa tradične vyhradzuje pre doktríny Comta, Macha a logických empiristov (novopozitivistov).

$$(4) p(E \rightarrow H | E) \leq p(E \rightarrow H).$$

Veta (4) je hľadanou odpoveďou. Hovorí, že E má na $E \rightarrow H$ *negatívny* dopad, teda, že *evidencia E vždy⁴ oslabuje tú časť obsahu hypotézy H, ktorá prekračuje obsah E*. Ak sú axiómy pravdepodobnostného kalkulu pravdivé, musí byť pravdivá aj veta (4), ktorá sa dá z nich naznačeným spôsobom odvodiť.

Keďže propozícia $E \rightarrow H$ vyjadruje tú časť hypotézy H, ktorá prekračuje evidenciu E, vyvstáva otázka, aký dopad má evidencia E na ten obsah hypotézy H, ktorý má s H spoločný. Podľa Poppera a Millera (1987, 569) možno spoločný obsah H a E zachytiť disjunkciou $H \vee E$.⁵ Následne si stačí uvedomiť, že $H \vee E$ je jedným z dôsledkov E, a preto $p(H \vee E | E) = 1$. Evidencia E má teda na $H \vee E$ *vždy pozitívny dopad*. Ak napokon stupeň podpory (angl. support), ktorý H získava od E, rozpiseme spôsobom, ktorý navrhuje Popper (1997, 581):

$$(5) st(H | E) = st(H \vee E | E) + st(E \rightarrow H | E),$$

výsledkom bude záver, že ak H získava od E vôbec nejakú podporu, tá musí byť pripísaná výlučne *deduktívnym* vzťahom medzi H a E, pretože druhý sčítanec je vždy negatívny. Popper nepreháňal, keď o argumente, ktorý bol práve opísaný, povedal, že ním „končia dejiny indukcie“ (1997, 582).

3 Empirizmus, petitio principii a nekonečný regres

Takzvaná „pravdepodobnostná podpora“ medzi vedeckou hypotézou H a empirickou evidenciou E je tak výlučne záležitosťou ich deduktívnych vzťahov. Prípadnej premene deduktívnej *závislosti* medzi H a E na deduktívne *zdôvodnenie* H pomocou E stojí však v ceste nie-

⁴ Anonymný recenzent postrehol, že kvôli neostrej nerovnosti dovoľuje veta (4) aj „neutrálny dopad“ E na $E \rightarrow H$, teda situáciu, keď $p(E \rightarrow H | E) = p(E \rightarrow H)$. Takýto prípad však nastáva iba vtedy, keď $p(H | E) = 1$ alebo $p(E) = 1$, čo sú krajné možnosti, nezaujímavé vzhľadom k skúmanému problému. Bližšie pozri (Miller 1994, 73).

⁵ $H \vee E$ vyplýva totiž ako z H, tak aj z E. *Akékolvek* tvrdenie X možno teda „rozložiť“, vzhľadom na iné tvrdenie E, na dve zložky: na deduktívnu zložku $X \vee E$ a na nededuktívnu zložku $E \rightarrow X$. Dôkazom platnosti uvedeného „rozkladu“ je fakt, že konjunkcia $(X \vee E) \& (E \rightarrow X)$ je ekvivalentná tvrdeniu X.

koľko prekážok. Prvou je banálny fakt, že $p(H \vee E | E) = 1$ neznamená to isté, čo $p(H | E) = 1$. Ale ani skromnejšie chůtky, podľa ktorých by $p(H \vee E | E) = 1$ mohlo hovoriť niečo (čokoľvek) pozitívne v prospech H (keď už nie to, že H je pravdivé), nie sú namieste. Návrh, podľa ktorého evidencia E zdôvodňuje aspoň tú časť hypotézy H (menovite $H \vee E$), ktorá z E vyplýva, sa totiž dopúšťa chyby petitio principii. Ako hovorí Miller (2005a, 64):

Každý deduktívne platný argument je kruhový [is question-begging] v tom zmysle, že jeho záver je (implicitne alebo explicitne) obsiahnutý v jeho premisách, uvažovaných dohromady [taken together]. Preto je úplne nemožné, že by platný argument mohol zdôvodniť svoj záver, aspoň dovedty, kým výrok nemôže zdôvodniť sám seba.

Premisa E preto nie je dobrým dôvodom pre seba samú a ani pre žiadny zo svojich dôsledkov, čiže ani pre $H \vee E$. Táto prekérna situácia sa dá opísať aj tak, že sa pokúsime predstaviť si proces testovania H. Ak je výsledok experimentu vyjadrený vetou E, potom E *opakuje* časť toho, čo už vieme, menovite $H \vee E$. Bolo by teda nanajvýš hlúpe tváriť sa, že E je dobrým dôvodom v prospech $H \vee E$.

Niektó by však mohol namietnuť, že o H a ani o $H \vee E$ *nevieme*, či sú pravdivé (prípadne, či majú inú želanú vlastnosť) a že to, či takými sú, sa dozvedáme práve až vďaka výsledku testovania, t.j. vďaka E. Majme teda znova deduktívny argument z E na $H \vee E$. Vieme *vďaka nemu*, že $H \vee E$ je pravdivé? Keďže argument z E na $H \vee E$ nezdôvodňuje pravdivosť *svojich* premís, môžeme ich pravdivosť len predpokladať (alebo sa uchýliť k ďalšiemu argumentu, a tak rozbehnúť nekonečný regres). Pravdivosť (či iná želaná vlastnosť) toho, čo sme chceli argumentom z E na $H \vee E$ ešte len zdôvodniť (t.j. záveru), je tak v skutočnosti vopred predpokladaná spolu s predpokladom pravdivosti premisy E, v ktorej je záver obsiahnutý – typické petitio principii. Odpoveď na pôvodnú námietku by tak mohla znieť:

Ak premisy deduktívne platného argumentu ešte nie sú dokázané, argument nebude postačovať na dôkaz záveru (to je námietka klasického skepticizmu);⁶ ak však premisy dokázané už sú, potom je záver, ktorý je súčasťou premís, takisto dokázaný.⁷ ... Ani v jednom

⁶ Miller tu naráža na známy skeptický argument nekonečného regresu.

⁷ Na tomto mieste ide zasa o narážku na petitio principii.

prípade nie je argument, ktorý spája premisy so záverom, schopný asistovať pri zdôvodňovaní či dokazovaní záveru. (Miller 2006, 70)

Hoci je teda tzv. „pravdepodobnostná podpora“, ktorú hypotéza H získava od evidencie E, výsledkom deduktívnych vzťahov medzi H a E, v skutočnosti nejde o žiadnu podporu v prospech pravdivosti (či inej zelanej vlastnosti) hypotézy H.

A tak môže vzniknúť podozrenie, že ak by sme doterajšie argumenty spojili s tézou tradičného empirizmu, podľa ktorej je skúsenosť (E) zdrojom pravého poznania sveta (H), delil by nás už iba krok od upretia existencie nepozorovateľným entitám, o ktorých hovoria vedecké teórie (H) vyjadrujúce prírodné zákony. Niektorí (ako napr. Schlick) neváhali a vykročili smerom k antirealistickej téze inštrumentalizmu, podľa ktorej vedecké teórie nie sú pravými zmysluplnými výrokmi, ale iba nástrojmi (bez pravdivostnej hodnoty) pre predikciu pozorovateľných javov. V tejto súvislosti sú nasledujúce slová Donalda Davidsona vskutku výstižným mementom:

Antirealizmus je prejavom neskrotného nutkania západnej filozofie zaistiť, že čokoľvek, čo je skutočné, môže byť poznané: antirealizmus sa to pokúša dosiahnuť nepriznaním existencie tomu, čo podľa jeho verdiktu leží za hranicami ľudského poznania. ... Väčšinu reduktívnych -izmov treba chápať ako podoby antirealizmu: idealizmus, pragmatizmus, empirizmus, materializmus, behaviorizmus, verifikacionizmus. Každý z nich sa snaží osekať realitu tak, aby zapadla do jeho epistemológie. (Davidson 2001, 69)

Je preto najvyšší čas preskúmať otázku, či je vôbec možné zachrániť vedecký realizmus a spolu s ním, v podobe užitočnej pre objasnenie kontroly dosahovania cieľa empirických vied, aj empirizmus.

4 Nutnosť revízie princípu empirizmu

Vo svojom opus magnum (1959, oddiel 6) Popper poznamenáva, že jeho kritérium demarkácie nie je len odpoveďou na otázku, čo je charakteristickým znakom empirických vied, ale aj riešením Humovho problému indukcie:

Koreňom tohto problému je zdanlivý spor medzi tým, čo možno nazvať „fundamentálnou tézou empirizmu“ – tézou, že o pravdivosti či nepravdivosti vedeckých tvrdení môže rozhodovať iba skú-

senosť – a Humovym rozpoznaním neprijateľnosti induktívnych argumentov. Tento spor vzniká len vtedy, ak sa predpokladá, že všetky empirické vedecké tvrdenia musia byť „jednoznačne rozhodnuteľné“, t.j. že musí byť v princípe možná ako ich verifikácia, tak ich falzifikácia. Ak sa vzdáme tejto požiadavky a pripustíme ako empirické aj tvrdenia rozhodnuteľné iba v jednom smere – jednostranne rozhodnuteľné a, najmä, falzifikovateľné – ktoré môžu byť testované systematickými pokusmi o ich falzifikáciu, potom spor zmizne: metóda falzifikácie nepredpokladá žiadnu induktívnu inferenciu, ale iba tautologické transformácie deduktívnej logiky, ktorej platnosť je nesporná.

Popperovo riešenie problému indukcie je teda čisto *negatívne*: hoci skúsenosť (resp. evidencia vyjadrená vetou E) nemôže v žiadnom zmysle zdôvodniť hypotézu H, ktorá svojim informačným obsahom prekračuje E, E môže (za určitých podmienok)⁸ H vyvrátiť. O *riešenie* problému indukcie ide preto, že v pozadí stojí revidovaný princíp empirizmu. Predtým, ako sa naň pozriem bližšie, treba však zvážiť jednu zrejmú námietku, ktorá by mohla byť vznesená proti Popperovej teórii vedy. Mám na mysli názor, podľa ktorého je Popperova diagnóza sporu medzi tézou empirizmu a Humovými závermi zavádzajúca: Nie je snád' na to, aby skúsenosť (vyjadrená vetou E) rozhodla o nepravdivosti hypotézy H, potrebné najprv prijať vetu E ako pravdivú práve na základe skúsenosti? A ak áno, je pravdou, že E vyjadruje skúsenosť, ktorá je základom či už pre prijatie E alebo pre odmietnutie H?

Odpoveď na námietku pozostáva z troch krokov. Predovšetkým, a) neexistuje žiadna skúsenosť, ktorá by bola nositeľkou informácie a súčasne by nebola vyjadrená nejakou vetou. Ďalej, b) ak sú vôbec niektoré vety prijímané na základe skúsenosti, nezvyšuje to ich epistemický status: veta E, o ktorej sa tvrdí, že je akceptovaná ako pravdivá vďaka skúsenosti, je *hypotetická v rovnakej miere, v akej bola pred svojím prijatím*. A napokon, c) neexistuje žiadny epistemický základ či už pre prijímanie alebo pre odmietanie viet. Keďže a) iba opakuje známu tézu,

⁸ Na tomto mieste vznesol anonymný recenzent námietku, že „jednou z podmienok je aj to, že musí ísť o všeobecný výrok, nie o štatistickú hypotézu“. Nie je to pravda. H je možné falzifikovať aj ak je singulárnym tvrdením a po stanovení podmienok, za akých je konečná postupnosť náhodná, aj ak je pravdepodobnostnou (štatistickou) hypotézou. Bližšie pozri Popper (1959, oddiely 28 a 63 až 68) a Miller (1994, podkapitola 9.3).

podľa ktorej „neexistuje nič také, ako „čistá skúsenosť“, ale iba skúsenosť interpretovaná vo svetle očakávaní a teórií“ (Popper 1959, 446) – tézu dostatočne zargumentovanú viacerými autormi z viacerých uhlov pohľadu,⁹ – zameriam sa na b) a c). Krok b) vychádza z a): kvôli neexistencii „čistej skúsenosti“ by epistemickým dôvodom pre prijatie vety X mohla byť iba iná veta Y. Avšak, či už má väčší obsah veta X alebo veta Y, v oboch prípadoch platia námietky proti možnosti ich zdôvodnenia predostreté v častiach 2 a 3 vyššie. Posledný krok c) nie je teda iba odmietnutím fundacionalizmu, ale *akýchkoľvek* dobrých dôvodov: pre prijímanie viet (ako pravdivých), rovnako ako pre ich odmietanie (ako nepravdivých) *nie sú potrebné vôbec žiadne dobré dôvody*, ktoré by nás uistili o správnosti našich rozhodnutí alebo by aspoň načas uchlácholili naše prípadné pochybnosti. Preto ani *skúsenosť*, nech je to už čokoľvek, *nie je dobrým dôvodom pre prijímanie a ani pre odmietanie viet* (resp. teórií). V tomto okamihu nepochybne mnohí stratia trpezlivosť a prehlásia, že empirizmus bol definitívne hodený cez palubu. Nie je to tak: je zachovaný na metodologickej, nie epistemologickej úrovni. Pre vysvetlenie najprv opíšem Millerov argument o asymetrii medzi akceptáciou a odmietnutím, ktorý považujem za kľúčový. Potom sa vrátim k princípu empirizmu a pokúsim sa ho sformulovať v revidovanej podobe.

Aj povrchný znalec Popperovej filozofie si pamätá jeho časté zdôrazňovanie asymetrie medzi verifikáciou a falzifikáciou, ktorá „vychádza z faktu, že *vzhľadom k singulárnym tvrdeniam* sú univerzálne tvrdenia iba falzifikovateľné“ (1959, 50, pozn.*2). Miller však argumentuje, že pri hlbšom pohľade sa dá objaviť ešte základnejšia asymetria medzi akceptáciou a odmietnutím; asymetria, ktorá sa týka *všetkých* viet či hypotéz bez rozdielu, či sú už univerzálne alebo nie:

Hypotéza musí byť najprv prijatá, aby mohla byť odmietnutá, pretože iba ak je akceptovaná či braná v úvahu, teda sformulovaná a nastolená, môže byť podrobená logickému skúmaniu;¹⁰ aby mohla byť prediskutovaná, musí byť najprv zvažovaná. Nedáva však zmysel tvrdiť, že hypotéza musí byť najprv odmietnutá, aby mohla byť

⁹ Najznámejším z nich je azda Sellars (1956), kde je doktrína o bezprostrednej, čistej či neinterpretovanej skúsenosti podrobená nemilosrdnej kritike a prehlásená za „mýtus daného“.

¹⁰ Miller tu zámerne používa Popperove slová, ktoré možno nájsť v (1959, oddiel 2).

prijatá. To samozrejme znamená, že aspoň niektoré hypotézy musia byť prijímané bez argumentu a bez evidencie. (Miller 2006, 89)

Na pochopenie asymetrie medzi akceptáciou a odmietnutím si treba všimnúť dôsledky tejto na prvý pohľad jednoduchej úvahy. Jedným z nich je vysvetlenie, prečo sú niektoré formulácie princípu empirizmu neobhájateľné. Ak by sme napríklad chceli tvrdiť, že v empirických vedách môže iba pozorovanie a experiment rozhodnúť, či vedecké tvrdenia sú alebo nie sú *akceptované* či *odmietané*, uviazli by sme v nekonečnom regrese:

Žiadny empirický argument (vychádzajúci z pozorovania a experimentu) nemôže byť predložený za alebo proti akémukoľvek tvrdeniu Φ (vrátane tvrdenia, že hypotéza H by mala byť akceptovaná alebo vyradená), pokiaľ tvrdenie Φ nebolo akceptované. Na to, aby sme mali argument pre prijatie H, sme teda už museli prijať Φ . (Miller 2006, 89, notácia upravená)

Ďalším dôsledkom je skutočnosť, že „do právomoci princípu empirizmu môžu preto spadať iba rozhodnutia o tom, či hypotézy odmietnuť alebo nie“ (tamtiež). Sľubovaná asymetria medzi akceptáciou a odmietnutím je tak objasnená. Teraz si už stačí všimnúť iba to, že vedie k hľadanému princípu empirizmu, ktorý môže byť vyjadrený v nasledovnej podobe: „*Žiadna hypotéza, ktorá bola akceptovaná, nemôže byť odmietnutá bez pozorovania a experimentu*“ (Miller 2006, 90). V takto revidovanom empirizme niet miesta pre konštruktívnu úlohu skúsenosti: pozorovania, experimenty, výsledky testov, empirická evidencia – bez ohľadu na svoje pomenovanie môžu v hre vedy plniť *výlučne* „negatívnu“, t.j. kritickú, úlohu. Načrtnuté argumenty totiž objasňujú, prečo pozitívne smery v oblasti filozofie empirických vied prebývajú vo svete ilúzií.

5 Objektívne prázdne konanie

Zostáva už iba vysvetliť, ako a prečo je princíp empirizmu zachovaný na metodologickej, nie epistemologickej úrovni. Úlohou je teda odpovedať na otázky, ako je možné využiť argumenty a empirickú evidenciu pri odmietaní vedeckých hypotéz, a prečo nemôžu asistovať pri ich prijímaní. Takto vykreslený kontrast by mal vrhnúť viac svetla nielen na „negativizmus“, podľa ktorého je úlohou skúsenosti v empirických vedách pomáhať pri kritike, nie pri ustanovovaní vedeckých

hypotéz, ale aj na spôsob, akým súvisí práve zrevidovaný empirizmus s vedeckým realizmom.

Kontrast medzi pozitívnym a negatívnym prístupom vo filozofii empirických vied nespočíva len v tom, že ten prvý je ortodoxným a väčšinovým prúdom, zatiaľ čo ten druhý nadchýna menšinu filozofov. Kľúčový rozdiel je vecný. „Pozitivistí“ ochotne pripustia, že hypotézy sú akceptované vo vyššie naznačenom zmysle, t.j. že sú najprv zvažované a testované, no len preto, aby mohli byť následne, v prípade úspechu, vyzdvihnuté na piedestál. Za úspech sa pritom zvykne považovať nielen *konfirmácia* či *verifikácia* (čo sú pôvodne technické termíny z prostredia Viedenského kruhu), ale dokonca aj *koroborácia* (čo je termín, ktorý zaviedol Popper pôvodne preto, aby sa fakt, že teória prešla testom, nezamieňal s jej údajnou konfirmáciou – pozri Popper 1959, 248-249, pozn.*1). Spomínaný piedestál býva takisto označovaný rôzne: raz nálepkou „vedecké poznanie“, inokedy bonmotom „potvrdené presvedčenie“ či priamo nápisom „pravda“. Málokterých pritom trápi fakt, že koroborácia (t.j. skutočnosť, že vedecká hypotéza H prešla testom) nemôže hovoriť *vôbec nič* v prospech H (pozri Miller 1994, 120).¹¹ Jednoducho ho vytrvalo ignorujú. Podobne sú zľahčované aj argumenty proti možnosti akejkoľvek pozitívnej podpory v prospech H zo strany empirickej evidencie. A tak stále prežívajú reči, podľa ktorých je H akceptovaná ako pravdivá (alebo ako disponujúca inou želanou vlastnosťou) *až keď obstoí v testoch*. Podľa „pozitivismu“ je teda akceptácia vedeckých hypotéz dvojstupňovým procesom: z H musia byť v prvom rade vyvoditeľné dôsledky verifikovateľné skúsenosťou, resp. evidenciou E, rozhodujúcou je však fáza, keď E v značnej miere potvrdzuje H (pozri aj Miller 2006, 91). Na druhej strane, podľa „negativizmu“, „*neexistuje taká činnosť, ktorá by nám umožnila uznať za pravdivou už predtým prijatú hypotézu*“ (Miller 2006, 91), a to z objektívnych príčin uvedených v častiach 2 a 3 vyššie.

Miller (2006, 91) trefne poznamenáva, že dvojstupňový model prijímania hypotéz, typický pre „pozitivismus“, „dáva pramálo zmyslu, ak rast poznania považujeme za objektívnu záležitosť“, pretože „hypotéza, ktorá bola vedou akceptovaná ako hodná diskusie, zostáva akceptovaná, pokiaľ nie je odmietnutá: nič sa s ňou nedeje“. Ak by ale ktosi

¹¹ Otázka, na čo je dobrá koroborácia, bola v druhej polovici 20. storočia predmetom diskusie medzi D. Millerom, J. Watkinsonom a A. Musgravom. V našom prostredí ju približuje stať Jozefa Katinu (2012).

trval na tom, že sa hypotéza vďaka svojmu obstátiu v testoch môže stať prijateľnou, mal by si uvedomiť, že „v takýchto testoch sa nedozvedáme nič, čo by sme objektívne nevedeli už predtým: nedošlo k žiadnej objektívnej zmene; hypotéza bola akceptovaná, takou aj zostáva, a jej status je nezmenený“ (Miller 2006, 91). Z tohto hľadiska je údajná akceptácia hypotézy H po jej úspešnom otestovaní *objektívne prázdny konaním*,¹² a to bez ohľadu na to, či je H akceptovaná ako pravdivá alebo ako disponujúca inou želanou vlastnosťou.

Ako je to potom s odmietaním hypotéz? Nejde tiež o objektívne prázdne činy? Argument by mohol znieť tak, že predsa „aj keď je hypotéza na začiatku akceptovaná alebo na konci falzifikovaná a odmietnutá, jej objektívny status sa nemení“ (Miller 2006, 91). Prečo by teda malo byť počiatočné zvažovanie hypotézy H či jej prípadné odmietnutie v čomsi odlišné od akceptácie H na základe úspešných testov? Millerova odpoveď (2006, 91) je elegantná a jednoduchá: „Vzťah medzi hypotézou a množinou akceptovaných tvrdení sa mení vtedy, keď je táto hypotéza akceptovaná alebo odmietnutá. Nemení sa, keď je potvrdená alebo akceptovaná ako pravdivá.“

Je zjavné, že jadrom rozporu medzi „pozitivismom“ a „negativizmom“ je chápanie statusu vedeckého poznania: väčšinový prúd („pozitivismus“) považuje za poslanie empirického výskumu zmenu statusu skúmaných hypotéz (napr. z „hypotéza“ na „potvrdená teória“), vďaka čomu vedecké poznanie v konečnom dôsledku zväzuje s psychologickými faktormi (akým je napríklad rast dôvery výskumníkov v úspešne otestovanú hypotézu). Podľa menšiny („negativizmus“) nie je úlohou empirického výskumu potvrdzovať a ani podkopávať testované hypotézy, zámerom teda nie je vykonať čosi *hypotézam*, ale čosi *s hypotézami*.¹³ To, čo sa s hypotézami dá urobiť vďaka ich empirickému testovaniu, je pritom ich *odmietnutie*. Názor, podľa ktorého má empirické skúmanie hypotéz umožniť výskumníkom testované hypotézy nielen odmietañ, ale aj (či najmä) ich prijímať, bol zdiskreditovaný v častiach 2 a 3 vyššie: evidencia E nehovorí nič v prospech prijateľnosti hypotézy H. Ba čo viac, v časti 4 bolo vysvetlené, prečo je H fakticky akceptovaná ešte pred konfrontáciou s E.

¹² Výraz „objectively empty action“ pochádza od Davida Millera (2006, 91).

¹³ Miller (2006, 54) vykresľuje tento kontrast pomocou zvrátov „to do something to the theory“ („pozitivismus“) vs. „to do something with the theory“ („negativizmus“).

Teraz by už malo byť jasné, prečo nie je možné udržať empirizmus na epistemologickej, ale iba na metodologickej úrovni: skúsenosť (nech je to už čokoľvek) nemá žiadny vplyv na epistemický status hypotézy H,¹⁴ môže sa však stať súčasťou procesu (resp. metódy), ktorý skončí odmietnutím H. Zopakujem, že problém indukcie stojí v ceste každej doktríny D, podľa ktorej by spomínaný proces mohol skončiť prijatím H. Ak by aj napriek tomu D tvrdohlavo trvala na tom, že skúsenosť vedie k akceptácii H, prestávala by hovoriť o vedeckom poznaní (hypotézach) a poznávaní (vedeckých metódach) v objektívnom zmysle slova: možno sa nálada výskumníkov, ktorí postrehnú, že H obstála v empirických testoch, zlepší, ich poznanie to však neposunie nikam. Ako hovorí Miller:

Subjektivismus a induktivismus kráčajú ruka v ruke. Nemôžem dokázať, že sú chybné, a netvrdím, že som dokázal správnosť objektivistickéj teórie vedy, podľa ktorej môže byť hypotéza len empiricky testovaná – a to iba *potom*, ako bola predložená.¹⁵ Od samého začiatku sa priznalo, že ide o návrh, avšak o návrh, ktorý môže byť iba ťažko v nemilosti u tých, ktorí si myslia, že vedci sa pomocou svojich presvedčení o vedu starajú, no neustanovujú ju. (Miller 2006, 92)

6 Empirizmus na metodologickej úrovni

Ako pomáha argumentácia a empirická evidencia pri kritike vedec-kých hypotéz na metodologickej úrovni? Odpoveď bude zároveň zhrnutím kontrastu medzi „pozitivismom“ a „negativizmom“, ktorý bol naznačený v predošlej časti. Miller (2006, 127) tvrdí, že

kombinácia empirickej evidence a deduktívnej logiky nedáva žiadnu radu, čo by sme mali akceptovať alebo o čom by sme mali byť presvedčení, či čo by sme mali robiť, dokonca ani akú teóriu či postup konania by sme mali uprednostniť. Logika je zmierená s tým,

¹⁴ Čo znamená, že skúsenosť nemôže žiadnu hypotézu podporiť a ani oslabiť: hypotézy sa, jednoducho, vďaka svojej interakcii so skúsenosťou nemenia (pozri Miller 2006, 54).

¹⁵ Miller tu znova naráža na Popperove slová (pozri poznámku 10). Popper v (1959) zdôrazňoval, že jeho kritérium demarkácie, o ktoré v citáte ide, je *návrhom* ako riešiť kľúčové epistemologické a metodologické problémy.

že za prítomnosti evidencie naznačuje, čo by sme nemali akceptovať, čo by sme nemali robiť a čo by sme nemali preferovať.

Náznaky teda dáva (za určitých podmienok) logika, nie evidencia, a to aj napriek tomu, že empirická evidencia je prirodzenou súčasťou situácie, keď sa prírodní vedci venujú testovaniu svojich hypotéz. Ak sa skupine výskumníkov podarí za prítomnosti evidencie E odvodiť z premís PP pomocou platného „pozitívneho“ argumentu¹⁶ záver Z, môžu si myslieť, že sa *vdďaka tomuto postupu* niečo dozvedeli o prijateľnosti Z, *nebude to však pravda*, pretože Z bol v skutočnosti od počiatku zahrnutý v PP. Myslieť si, že uvedený spôsob argumentácie odhaľuje pravdivosť (či inú želanú vlastnosť) záveru Z teda neznamená nič iné, ako sa dopúšťať chyby známej ako *petitio principii* (pozri časť 3).

Keď sa však skupine výskumníkov podarí za prítomnosti E odvodiť z premís PP pomocou platného „negatívneho“ argumentu záver $Z \ \& \ \sim Z$, môžu *bez obáv tvrdiť*, že sa *vdďaka tomuto postupu* niečo dozvedeli, menovite, že PP nie sú pravdivými premisami. V takom prípade totiž nezáleží na tom, že záver $Z \ \& \ \sim Z$ musel byť zahrnutý v premisách PP od samého počiatku (inak by z nich nevyplýval), keďže uvedený spôsob argumentácie odhaľuje nepravdivosť (či inú neželanú vlastnosť) toho, čo bolo pôvodne zvažované ako nepoškvrnený kandidát na post poznania. Hoci výskumníci *vdďaka naznačenej metóde nevedia*, ktorá z premís má byť odmietnutá ako nepravdivá, *vedia*, že *niektorá z premís odmietnutá byť musí*. Tento záver je v súlade so známou Duhem-Quinovu tézou, ktorá sa však niekedy interpretuje v opačnom zmysle, a síce, že pomocou pozorovania nie je možné vyvrátiť žiadnu teóriu.¹⁷ Sám Quine sa vyjadril pomerne jasne:

Občas se také stane, že zkušenostní důsledky nějaké teorie nenanstanou; potom, nebude-li tomu nic bránit, vyhlásíme danou teorii za nepravdivou. Avšak ten neúspěch vyvrací jen teorii jako určitý celek, jako konjunkci mnoha tvrzení; poukazuje na to, že jedno či

¹⁶ Pod „platným pozitívnym argumentom“ mám na mysli akýkoľvek deduktívny argument, ktorého záverom nie je kontradikcia. Tzv. nededuktívne argumenty netreba brať do úvahy, pretože, hoci obsahujú vo svojich záveroch čosi, čo nie je obsiahnuté v ich premisách, tak len za tú cenu, že premisy to vždy oslabujú (pozri časť 2).

¹⁷ Pre rôzne, vzájomne nezlučiteľné, interpretácie Duhem-Quinovej tézy pozri napríklad Laudan (1990).

více těchto tvrzení není pravdivých, ale neukazuje, která to vlastně jsou. (Quine 2006, 128)

Existuje teda metóda, ktorá využíva ako argumenty, tak aj empirickú evidenciu, a to za účelom spochybnenia vedeckých hypotéz a ich odmietnutia ako nepravdivých. Neexistuje však žiadna metóda, ktorá by poukazovala na pravdivosť (či na inú želanú vlastnosť) hypotéz testovaných pomocou skúsenosti. Ak „skúsenostné dôsledky“ nejakej teórie „nastanú“, nehovorí to nič v jej prospech. Práve na tomto mieste vyniká kontrast medzi „negatívnymi“¹⁸ a „pozitívnymi“ metódami: *tie prvé pomáhajú výskumníkom zmeniť ich pôvodnú mienku, tie druhé im ju nemôžu pomôcť ustanoviť* (pozri časti 2 a 3) *a ani zmeniť* (čo ani nebolo ich úlohou). Vďaka tým prvým sa *môžu niečo odučiť*, pretože kontradikcia je jednoznačným znakom nepravdivosti, no tie druhé im nenaznačujú vôbec nič, a preto sa vďaka nim *ani nič neučia*. „Pozitívne metódy“ sú, jednoducho, sterilné.

Úlohou „negatívnych“ metód je teda pomáhať prírodným vedcom udržať svoj výskum pod kontrolou. To je možné preto, že sú súčasťou *procesov racionálneho rozhodovania*, ktorú z hypotéz odmietnuť a ktorú nie, a to aj napriek tomu, že *výsledné rozhodnutia*, ku ktorým sa vedci napokon odhodlajú, *nemôžu byť nikdy ospravedlnené* empirickou evidenciou a ani pravidlami logiky, ktorú vedci využívajú počas svojho výskumu. Samozrejme, ako poznamenáva Miller (2006, 127): „To neznamená, že by sme nemohli prekračovať odporúčania logiky. Logika dovoľuje to, čo nezakazuje. V tomto zmysle platí, že hoci existujú presvedčenia ..., ku ktorým sa dospieva pomocou ... kritických argumentov, neexistujú racionálne presvedčenia.“

Racionálnym presvedčením, napríklad o tom, ktorú z vedeckých hypotéz uprednostniť a ktorú zamietnuť, by totiž, podľa „pozitivistov“, bolo iba presvedčenie ospravedlnené výsledkami kritickej diskusie hypotéz. Ako sme však videli, kritická diskusia nemôže ospravedlniť *žiadne závery*, dokonca ani tie, ktoré sú z nej deduktívne vyvedené. Ak vedci akceptujú evidenciu E, ktorá je v spore s hypotézou H, môže sa zdať, že majú dobrý dôvod pre odmietnutie H. Tento dojem sa však rozplynie, keď si spomenieme na záver častí 2 a 3, podľa ktorého ne-

¹⁸ Pod „negatívnymi metódami“ nemám na mysli iba metódu falzifikácie, ktorú opísal Popper v (1959), ale aj využívanie argumentov *reductio ad absurdum* (RAA) za účelom kritiky, nie ustanovovania, tvrdení. Preskúmanie vzťahu falzifikácie a RAA by si zaslúžilo osobitnú pozornosť.

existujú dobré dôvody v prospech prijímania a ani odmietania tvrdení. Hoci teda „predostrieť vyvrátenie znamená poskytnúť argument ..., neznamená to argumentovať za (alebo proti) dôsledku; znamená to odvodiť dôsledok“ (Miller 2005b, 80). Proces rozhodovania o osude vedeckých hypotéz je tak racionálny preto, lebo v ňom vedci využívajú rozum, logiku a empirickú evidenciu za účelom vyriešenia „negatívneho“ verdiktu.

Záver

Výsledkom je empirizmus oslobodený od dogmy redukcionizmu, podľa ktorej je skúsenosť jediným zdrojom skutočného poznania sveta, ako aj od dogmy justifikacionizmu, podľa ktorej musí byť poznanie o svete zdôvodnené prostredníctvom skúsenosti. Otázka „pravých“ zdrojov poznania je irelevantná, pretože žiadny zdroj nedisponuje autoritou (pozri Popper 1963, 24): *skutočné poznanie nemôže byť zdôvodneným poznaním*. Úloha skúsenosti a zmysel empirického testovania vedeckých teórií, ktorý je v knihe Godfrey-Smitha (2003) zahalený rúškom tajomstva, sa však objasnil: *testovanie vedeckých hypotéz môže zrušiť ich akceptáciu, nemôže ju však vytvoriť* (pozri Miller 2006, 125). S týmto druhom empirizmu je plne zlučiteľný vedecký realizmus, podľa ktorého sa „veda môže racionálne snažiť o opísanie skutočnej štruktúry sveta, vrátane jej nepozorovateľnej štruktúry“ (Godfrey-Smith 2003, 220). Empirizmus opísaný v tomto článku totiž nepožaduje, aby všetko, čo je skutočné, bolo testovateľné. Empirické teórie z oblasti prírodných vied neopisujú len a iba pozorovateľné fenomény. Naopak, majú mnoho nefalzifikovateľných dôsledkov (pozri Miller 1994, 10). Po odmietnutí dogmy justifikacionizmu je však zrejmé, že ich prijímanie nie je podmienené existenciou žiadnych dobrých dôvodov, teda ani takých, ktoré by chceli pochádzať zo skúsenosti. Potrebné je iba to, aby teórie mali falzifikovateľné dôsledky a boli podrobené empirickým testom. Pokiaľ sú súčasťou vedy, dovtedy je všetko, o čom hovoria, chápané ako opis skutočnej štruktúry sveta, vrátane nepozorovateľných entít.¹⁹ Zmena tohto chápania nie je anomáliou, ale zmyslom metódy vedeckého bádania.

¹⁹ Preto je aj názor, že „zástancovia vedeckého realizmu veria v objektívnu existenciu tých entít, ktoré postulujú dobre fungujúce vedecké teórie“ (Schmidt 2011, 376) obhájitelný iba v naznačenom zmysle. Návrh, že až úspešné testovanie teórií poskytne dôvod veriť v existenciu postulovaných entít, by bol totiž reliktom antirealizmu.

Katedra filozofie
 Fakulta humanitných vied
 Univerzita Mateja Bela
 Tajovského 40,
 974 01 Banská Bystrica
 Slovenská republika
 milos.taliga@umb.sk

Literatúra

- CARNAP, R. (1966): *Philosophical Foundations of Physics. An Introduction to the Philosophy of Science*. New York – London: Basic Books, Inc.
- DAVIDSON, D. (2001): Indeterminism and Antirealism. In: Davidson, D.: *Subjective, Intersubjective, Objective*. Oxford: Clarendon Press, 69-84.
- GODFREY-SMITH, P. (2003): *Theory and Reality. An Introduction to the Philosophy of Science*. Chicago – London: The University of Chicago Press.
- KATINA, J. (2012): Analýza pojmu koroborácie v Popperovej teórii vedy. *Filozofia* 67, č. 2, 124–135.
- LAUDAN, L. (1990): Demystifying Underdetermination. In: Savage, C. W. (ed.): *Scientific Theories*. Minneapolis: University of Minnesota Press, 267-297.
- MILLER, D. W. (1994): *Critical Rationalism. A Restatement & Defence*. Chicago – La Salle: Open Court Publishing Company.
- MILLER, D. W. (2005a): Do We Reason When We Think We Reason, or Do We Think? *Learning for Democracy* 1, No. 3, 57-71.
- MILLER, D. W. (2005b): A Reply to Trudy Govier and Philip Adey. *Learning for Democracy* 1, No. 3, 78-80.
- MILLER, D. W. (2006): *Out of Error. Further Essays on Critical Rationalism*. Aldershot: Ashgate.
- POPPER, K. R. – MILLER, D. W. (1987): Why Probabilistic Support Is Not Inductive. *Philosophical Transactions of the Royal Society of London* 321, No. 1562, 569-591.
- POPPER, K. R. (1959): *Logic of Scientific Discovery*. London – New York: Routledge.
- POPPER, K. R. (1963): *Conjectures & Refutations*. London: Routledge.
- POPPER, K. R. (1997): Support a Countersupport. Indukce se stává kontraindukci, epagogé se vrací k elenchus. In: POPPER, K. R.: *Logika vědeckého bádání*. Preložil Jiří Fiala. Praha: Oikoyomenh, 578-586.
- QUINE, W. O. (2006): Naturalizace epistemologie. In: Marvan, T. – Dostálová, L. (eds.): *Willard van Orman Quine. Vybrané články k ontologii a epistemologii*. Preložil Tomáš Marvan. Plzeň: Západočeská univerzita v Plzni, 120-137.
- SELLARS, W. (1956): Empiricism and the Philosophy of Mind. In: Feigl, H. – Scriven, M. (eds.): *The Foundations of Science and the Concepts of Psychology and Psychoanalysis*, Minneapolis: University of Minnesota Press, 253–329.
- SCHMIDT, M. (2011): Je analýza jazyka metafyzikou? *Filozofia* 66, č. 4, 374-377.

Dva druhy neartikulovaných zložiek

Marián Zouhar¹

Slovenská akadémia vied, Bratislava

Abstract: It is often claimed that propositions expressed by certain kinds of sentence involve constituents that are unarticulated at the linguistic level. Two sorts of unarticulated constituents are sometimes distinguished in the literature – those that are explicitly unarticulated and those that are implicitly unarticulated. The main claim of the present paper is that, contrary to appearances, there are no implicitly unarticulated constituents and that all of them are in fact explicitly unarticulated ones. The argument runs as follows: If a propositional constituent is supposed to be implicitly unarticulated, it appears in the proposition on the basis of free enrichment (or expansion). On the other hand, if a propositional constituent is explicitly unarticulated, it appears in the proposition on the basis of another kind of process, namely saturation (or completion). However, it is shown that free enrichment (and expansion) leads to a serious problem concerning the unity of propositions. Consequently, if this problem is to be circumvented, free enrichment (and expansion) has to be rejected. As a result, implicitly unarticulated constituents have to be rejected as well.

Keywords: Expansion, explicitly unarticulated constituent, free enrichment, implicitly unarticulated constituent, literal meaning, proposition expressed, propositional unity.

1 Doslovný význam a vyjadrený obsah

Nech *doslovným významom* zloženého výrazu je obsah,² ktorý je určený výlučne sémantickými konvenciami pre jednoduché výrazy ob-

¹ Táto stať vznikla na Filozofickom ústave SAV v rámci grantového projektu VEGA č. 2/0019/12 *Jazyk a determinácia významu v komunikácii*.

² Doslovný význam jednoduchého výrazu je určený výlučne sémantickou konvenciou pre daný výraz.

siahnuté v danom zloženom výraze a spôsobom syntaktického usporiadania jednoduchých výrazov do zloženého výrazu.³ Ak zloženým výrazom je veta (výrok), doslovným významom môže byť *propozícia* (teda nositeľ pravdivostnej hodnoty). To však platí len v prípade niektorých špecifických druhov viet – doslovným významom viet viacerých druhov nemôže byť propozícia, pretože ich doslovný význam je určitým zásadným spôsobom neúplný, vďaka čomu sám osebe nestačí na determináciu žiadnej pravdivostnej hodnoty.⁴

Táto situácia nastáva napríklad v prípade vety, ktorá obsahuje *indexický výraz*, teda ukazovacie zámeno, osobné zámeno, príslovkové určenie času, resp. miesta, či iný výraz podobného druhu. Podľa viacerých renomovaných teórií by súčasťou vyjadrenej propozície mal byť *referent*, ktorý indexický výraz nadobúda vzhľadom na konkrétne

³ V tomto vymedzení ignorujem napríklad ustálené slovné spojenia či iné špecifické druhy výrazov, ktoré môžu byť zložené, ale ich doslovný význam je určený podobne, ako je to v prípade jednoduchých výrazov, t. j. výlučne na základe sémantickej konvencie. Alternatívne by sme mohli zložené výrazy tohto druhu považovať v skutočnosti za jednoduché. Úvahy v stati sú od rozhodnutia v tejto veci nezávislé.

⁴ Predpokladá sa tu, že propozície majú zložky, a teda sú *zloženými* entitami; zároveň sú tieto zložky nejakým spôsobom usporiadané, a teda propozície sú aj *štruktúrovanými* entitami. Táto koncepcia propozícií (pozri napríklad Kaplan 1975) je odlišná od toho, ako sa propozície chápu v intenzionálnej sémantike (kde sa považujú za funkcie z množiny možných svetov do množiny pravdivostných hodnôt, resp. z množiny možných svetov do chronológií pravdivostných hodnôt). Motiváciu v prospech koncepcie propozícií ako zložených štruktúrovaných entít uvádzať nebudem (pozri napríklad Soames 1987 alebo 2010). Dodávam však, že existujú rôzne argumenty proti tejto koncepcii (pozri napríklad Tichý 1994, Jespersen 2003 alebo Bealer 2004) – ktoré poukazujú napríklad na to, že nie je prípustné, aby tá istá entita bola zložená a štruktúrovaná, a zároveň bola nositeľom pravdivostnej hodnoty – a že jej problémy možno odstrániť, ak sa odliší (neštruktúrovaná) propozícia (v intenzionálnom zmysle) od konštrukcie takejto propozície, pričom zloženosť a štruktúrovanosť budú atribútmi propozičných konštrukcií. Túto cestu ponúka transparentná intenzionálna logika (pozri Tichý 1988, Raclavský 2009 alebo Duží – Jespersen – Materna 2010). To, čo hovorím v hlavnom texte v súvislosti s propozíciami, možno v mnohých prípadoch adekvátnejšie formulovať pre propozičné konštrukcie. Od týchto komplikácií však budem abstrahovať a štruktúrovanosť budem v texte zjednodušene pripisovať priamo propozíciám (korektnější spôsob vyjadrovania, keď som hovoril o propozičných konštrukciách, som zvolil v stati Zouhar 2011a).

použitie,⁵ a ten sa nedá určiť len vzhľadom na možný svet, ale až na základe relevantných parametrov kontextu použitia výrazu. Kontextom použitia indexického výrazu môže byť množina, resp. usporiadaná *n*-tica pozostávajúca z parametrov rôznych druhov, ako sú možný svet, čas použitia, miesto použitia, hovorca, adresát, množina určitých objektov atď.⁶ Doslovným významom vety s indexickým výrazom teda nie je propozícia, keďže takáto veta vyjadrí propozíciu až vzhľadom na relevantné parametre kontextu použitia. Vyjadrená propozícia v tomto prípade teda predstavuje inú vrstvu obsahu, ktorú treba odlišiť od vrstvy doslovného významu.

Analogickú situáciu môžeme zaznamenať aj v súvislosti s gramatickými tvarmi výrazov niektorých druhov. Napríklad gramatický prezent slovesa možno v niektorých prípadoch chápať tak, že sa vzťahuje na čas, v ktorom sa veta s takýmto slovesom použije; gramatické perfektum zase možno v určitých prípadoch chápať tak, že sa vzťahuje na čas, ktorý predchádza času použitia vety s takýmto slovesom. Gramatické tvary niektorých výrazov teda niekedy fungujú analogicky ako indexické výrazy – aj v tomto prípade musíme vziať do úvahy kontext použitia, aby sme určili všetky relevantné zložky vyjadrenej propozície, a teda takéto vety nebudú vyjadrovať propozície len v závislosti od sémantických konvencií a syntaktického usporiadania. Aj v tomto

⁵ Medzi renomované teórie tohto druhu patrí napríklad Kaplanova teória demonstratív (či všeobecnejšie indexických výrazov), podľa ktorej každý výraz (indexický aj neindexický) vyjadruje určitý *charakter* a označuje určitý *obsah* (pozri Kaplan 1989), či Perryho koncepcia indexických výrazov, v ktorej sa rozlišuje *rola* a *obsah* indexického výrazu (pozri Perry 1977). Charakter výrazu možno stotožniť s funkciou z kontextu použitia do obsahu; táto funkcia teda vyberá, čo je obsahom výrazu vzhľadom na kontext použitia. Obsahom výrazu (resp. jeho konkrétneho použitia) je referent či, presnejšie, funkcia, ktorá okolnostiam ohodnotenia (t. j. možným svetom či podobným entitám) priraduje referenty. Keďže v prípade indexických výrazov je táto funkcia konštantná, môžeme pre jednoduchosť priamo považovať referent indexického výrazu za jeho obsah. Napríklad, charakterom osobného zámena „ja“ je akési pravidlo, ktoré vyberá ako referent použitia zámena osobu, ktorá ho použila, a preto aktuálny používateľ zámena je obsahom jeho konkrétneho použitia.

⁶ V Kaplanovom chápaní je kontextom použitia usporiadaná štvorica, v ktorej sa vyskytuje hovorca, čas, miesto a možný svet. Z dôvodov, ktoré nebudem v tejto stati uvádzať, je však vhodné, ak doňho zaradíme aj ďalšie parametre.

prípade vyjadrená propozícia predstavuje inú vrstvu sémantického obsahu, ktorú treba odlíšiť od vrstvy doslovného významu.

Výrazy a gramatické tvary, ktoré sú v naznačenom zmysle kontextovo citlivé, budem súhrnne označovať termínom *kontextovo citlivé syntaktické jednotky*. V záujme čo najväčšieho zjednodušenia (a bez odkazu na akúkoľvek konkrétnu teóriu) predpokladajme, že v doslovnom význame viet obsahujúcich kontextovo citlivé syntaktické jednotky sa vyskytujú prázdne miesta, na ktoré treba doplniť hodnoty, ktoré kontextovo citlivé syntaktické jednotky nadobúdajú vzhľadom na kontext použitia. Zaplnením prázdnych miest hodnotami kontextovo citlivých syntaktických jednotiek vznikne propozícia. Doslovný význam viet s kontextovo citlivými syntaktickými jednotkami možno preto označiť Bachovým termínom *propozičný radikál* (pozri Bach 1994). Zaplnenie takýchto prázdnych miest považujem za *minimálnu modifikáciu* doslovného významu, propozičného radikálu. Vety, ktoré obsahujú kontextovo citlivé syntaktické jednotky, teda vyjadrujú propozície len vzhľadom na minimálne modifikácie ich doslovného významu.

Termínom *vyjadrený obsah* označme to, čo výraz vyjadruje vzhľadom na kontext jeho použitia. V prípade niektorých viet je vyjadreným obsahom ich doslovný význam; to sa týka takých viet, ktoré neobsahujú žiadne kontextovo citlivé syntaktické jednotky, a teda ich doslovným významom je propozícia. Vyjadreným obsahom viet, ktoré obsahujú aj kontextovo citlivé syntaktické jednotky, nie je ich doslovný význam, a teda v tomto prípade je netriviálny rozdiel medzi doslovným významom a vyjadreným obsahom. Vyjadrený obsah takejto vety sa získa z jej doslovného významu minimálnou modifikáciou.

2 Explicitne neartikulované zložky

V predchádzajúcej časti som prezentoval pomerne neproblematický obrázok interakcie medzi sémantickou rovinou jazyka a kontextom jeho použitia. Kontext použitia nemá vplyv na doslovný význam vety, ale má vplyv na jej vyjadrený obsah. Dôležité je pritom to, že každý zásah kontextu použitia do vyjadreného obsahu je vynútený syntaktickou rovinou vety. Mám tým na mysli skutočnosť, že do vyjadreného obsahu sa kontextovo podmienená zložka dostane iba v prípade, že v samej vete existuje taká syntaktická jednotka - výraz špecifického druhu, prípadne špecifický gramatický tvar nejakého výrazu - ktorej

táto kontextovo podmienená zložka zodpovedá. Môžeme povedať, že všetky zložky vyjadreného obsahu sú v nejakej podobe *artikulované* vo vete, ktorá daný obsah vyjadruje vzhľadom na kontext jej použitia.

Napriek – alebo, skôr, vďaka – jeho relatívnej jednoduchosti sa zdá, že tento obrázok interakcie medzi sémantickou rovinou jazyka a kontextom použitia celkom nezodpovedá tomu, čo poznáme z komunikačnej praxe. Zdá sa totiž, že keď hovorca použije vo svojej výpovedi určitú vetu, spravidla tak vyjadří obsah, ktorý nie je totožný nielen s doslovným významom vety, ale ani s vyjadreným obsahom vety (vzhľadom na kontext jej použitia) v uvedenom pomerne úzkom chápaní; často sa totiž stáva, že v hovorcovom použití veta vyjadruje obsah, ktorý je bohatší či nejako inak špecifickejší, ako je vyjadrený obsah vety v uvedenom chápaní. Preto sa považuje za viac či menej prijateľnú hypotéza, že niektoré zložky obsahu, ktorý veta vyjadruje vzhľadom na kontext jej použitia, sú v skutočnosti *neartikulované* v tom zmysle, že nezodpovedajú žiadnej syntaktickej jednotke vyskytujúcej sa vo vete.

Zopár príkladov môže ilustrovať túto hypotézu. Dobrým príkladom sú azda vety s tzv. privlastňovacími deskripciami ako „Albertovo auto“, resp. zloženými privlastňovacími demonstratívami ako „jeho auto“. Vety, ktoré obsahujú takéto výrazy, sú totiž v určitom zmysle neurčité. Vezmime si vety

- (1) Albertovo auto je nepojazdné.
- (2) Jeho auto je nepojazdné.

Veta (1) je neurčitá v tom zmysle, že sa v nej nešpecifikuje, aký je vzťah medzi Albertom a určitým autom; analogicky vo vete (2) sa nešpecifikuje, aký je vzťah referenta výrazu „jeho“ (vzhľadom na daný kontext použitia) a určitým autom. Albert (resp. referent výrazu „jeho“) môže dané auto vlastniť, môže ho mať požičané, môže na ňom jazdiť, môže v ňom sedieť, atď. Skutočnosť, že vo vetách (1) a (2) je zamľčaný vzťah medzi Albertom, resp. referentom výrazu „jeho“, a určitým autom, si môžeme uvedomiť vtedy, keď privlastňovaciu väzbu nahradíme zodpovedajúcou vedľajšou vetou; v tejto vedľajšej vete sa totiž musí špecifikovať daný vzťah, inak by sa nedala formulovať. Výraz „Albertovo auto“ môžeme teda nahradiť výrazmi „auto, ktoré Albert vlastní“, „auto, ktoré má Albert požičané“, „auto, v ktorom Albert jazdí“, „auto, v ktorom Albert sedí“, atď. V závislosti od toho, aký vzťah je relevant-

ný v danom kontexte použitia, možno použitím vety (1) vyjadriť rôzne propozície; napríklad:⁷

- (1a) »Auto, ktoré Albert vlastní, je nepojazdné«⁸
- (1b) »Auto, ktoré má Albert požičané, je nepojazdné«
- (1c) »Auto, v ktorom Albert jazdí, je nepojazdné«
- (1d) »Auto, v ktorom Albert sedí, je nepojazdné«

V záujme korektnosti treba dodať, že nie všetky privlastňovacie deskripcie, resp. zložené privlastňovacie demonstratíva, musia zamlčovať vzťah medzi subjektom (teda tým, čo patrí do definičného oboru vzťahu) a objektom (teda tým, čo patrí do oboru hodnôt vzťahu). To je napríklad prípad deskripcie „Albertova manželka“ alebo demonstratíva „jeho manželka“, ktoré sa vyskytujú vo vetách

- (3) Albertova manželka je neverná.
- (4) Jeho manželka je neverná.

Predpokladajme, že výraz „Albertova manželka“ v skutočnosti zamlčiava nejaký vzťah medzi Albertom a jeho manželkou. V takom prípade by sa táto deskripcia dala nahradiť zodpovedajúcou deskripciou s vedľajšou vetou, napríklad „manželka, s ktorou Albert žije“, „manželka, s ktorou Albert má deti“, „manželka, ktorú Albert bije“, atď. Keď to však urobíme, teda keď bližšie špecifikujeme, o ktorú Albertovu manželku ide, môžeme nechtiac spôsobiť komunikačný zmätok v tom zmysle, že adresát výpovede by mohol nadobudnúť dojem, že Albert má viacero manželiek, keďže dodatočnou informáciou spravidla mienime odlíšiť jednu z nich od ostatných manželiek.⁹ Na rozdiel od prípadu, s ktorým sme sa stretli v súvislosti s deskripciou „Albertovo auto“, nie je takéto doplnenie v prípade deskripcie „Albertova manželka“ potrebné. Dôvod spočíva v tom, že relevantný vzťah sa v danej deskripcii

⁷ Ak vzhľadom na kontext použitia vety (2) je referentom zámena „jeho“ Albert, propozície (1a) – (1d) možno vyjadriť aj vetou (2) (vzhľadom na kontext jej použitia).

⁸ Týmto spôsobom budem v stati reprezentovať propozície, resp. zložky propozícií. Uvedené zápisy teda možno považovať za mená propozícií, resp. zložiek propozícií.

⁹ Zostáva pritom ďalej nešpecifikované, či ide o manželky, ktoré má Albert v tom istom časovom období (a teda je bigamista), alebo ide o manželky, ktoré má v rôznych časových obdobiach (a teda sa napríklad rozviedol alebo ovdovel).

už predsa vyskytuje. Výraz „manželka“ totiž nemusí vystupovať len ako unárny predikát, ktorý utvorí vetu, ak sa spojí s jedným vhodným výrazom, ale môže vystupovať aj ako binárny predikát, ktorý utvára vetu s dvoma výrazmi vhodného druhu. V takom prípade sa aj špecifikuje vzťah medzi Albertom a ďalšou osobou. Veta (3) teda vyjadří ako svoj obsah propozíciu

(3a) »Albertova manželka je neverná«

Na rozdiel od výrazu „manželka“ nemôže výraz „auto“ vystupovať raz ako unárny, inokedy zase ako binárny predikát; vždy sa používa len ako unárny predikát, teda vzťahuje sa na vlastnosť, ktorú objekty môžu exemplifikovať. Keď sa však v propozícii – napríklad v propozícii (1a) – má vyskytovať Albert aj exemplifikát vlastnosti *byť auto*, musí sa v nej zároveň vyskytovať aj vzťah medzi týmito objektmi, a preto sa v propozícii vyskytuje aj niečo navyše, čo sa vo vete neartikuluje.¹⁰

Iným prípadom, ktorým možno ilustrovať výskyt neartikulovaných zložiek v propozíciách, sú vety, ktorých predikátom sú adjektíva. Z dôvodov, ktoré nás tu podrobnejšie nebudú zaujímať, sa prijíma ako štandardný názor ten, podľa ktorého vo vete, ktorej predikátom je adjektívum, je zamlčaný výskyt substantíva. Adjektíva totiž fungujú ako gramatické *modifikátory* substantív. Vezmime si ako príklad vetu

(5) Albert je ženatý.

Ak má štandardný názor pravdu, v tejto vete je zamlčané adjektívum, napríklad „človek“ alebo „muž“ (prípadne niečo podobné). Na sémantickej úrovni to znamená, že veta (5) vyjadří nejakú obohatenú propozíciu, napríklad:

(5a) »Albert je ženatý muž«

(5b) »Albert je ženatý človek«

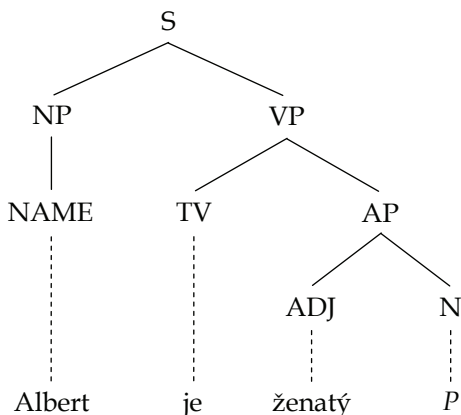
¹⁰ Dôvod, prečo je to tak, nebudem podrobnejšie špecifikovať, ale iba ho načrtnem. To, že propozície sú zložené a štruktúrované, naznačuje, že ide o jeden sémantický objekt s viacerými zložkami. To znamená, že propozície sú *jednotné* (keby to tak nebolo, nemohli by byť *jedným* objektom). Keby sa však stalo, že sa v propozícii majú vyskytovať dva objekty, ale bez vzťahu, ktorý by ich „spojil“, nemohli by tieto objekty tvoriť jednotu. Ak sa teda v propozícii má vyskytovať Albert aj exemplifikát vlastnosti *byť auto*, tak je potrebné, aby sa v nej vyskytoval aj vzťah medzi nimi. O probléme jednoty propozície budem v iných súvislostiach hovoriť v poslednej časti.

Entita »muž«, resp. »človek« je teda neartikulovanou zložkou vyjadrenej propozície.

Podobných prípadov možno nájsť viacero. Jednou zo záľub filozofov jazyka je hľadanie ďalších a ďalších prípadov, v ktorých sa vo vyjadrených propozíciách vyskytujú také zložky, ktoré nie sú vo vete artikulované. Domnievam sa však, že uvádzanie ďalších príkladov je zbytočné, keďže by malo byť jasné, že hypotéza, podľa ktorej nie všetky zložky obsahu, ktorý veta vyjadruje vzhľadom na kontext jej použitia, zodpovedajú nejakým syntaktickým jednotkám vyskytujúcim sa vo vete, vyzerá prijateľne.

Zdá sa však, že hoci sa v uvedených príkladoch môže ilustrovať výskyt takých zložiek, ktoré nie sú artikulované v *explicitnom* zmysle, predsa len v nejakom inom zmysle artikulované musia byť. Ide totiž o to, že ak adjektívum vystupuje ako modifikátor nejakého substantíva, tak v nejakej podobe by sa vhodný jazykový prostriedok, ktorý by mal reprezentovať zamlčané substantívum, vo vete (5) mal vyskytovať. Analogicky, genitívna väzba v privlastňovacích deskripciách, resp. zložených privlastňovacích demonstratívach by mala zastupovať vzťah subjektu k objektu, a teda vo vetách (1) a (2) by sa v nejakej podobe mal reprezentovať zamlčaný vzťahový výraz. Z tohto dôvodu sa zvyčajne uznáva, že aj keď tieto reprezentácie nie sú explicitné, teda viditeľné na úrovni povrchovej formy vety, predsa len sú vo vete implicitne prítomné a dajú sa reprezentovať na úrovni (skrytej) logickej formy vety. Presnejšie povedané, na úrovni logickej formy sa v daných výrazoch vyskytujú voľné premenné vhodných druhov. Vezmime si na ilustráciu vetu (5). Jej logickú formu môžeme zachytiť pomocou stromového diagramu v nasledujúcej podobe:¹¹

¹¹ Veta (výraz kategórie S) pozostáva z menného výrazu (NP) a slovesného výrazu (VP) (pokiaľ nejde o vetu, ktorá je priradovacím súvetím). Menný výraz možno tvoriť rôznymi spôsobmi; môže pozostávať napríklad z vlastného mena (výrazu kategórie NAME), ale môžu ho tvoriť aj iné výrazy (neuvádzam ich tu, keďže sú v našom príklade irelevantné). Aj slovesný výraz možno tvoriť rozličnými spôsobmi; môže napríklad pozostávať z tranzitívneho slovesa (výraz kategórie TV), ktorým môže byť napríklad tvar slovesa „byť“, a adjektívnej frázy (výrazu kategórie AP), ktorá pozostáva z adjektíva (výrazu kategórie ADJ) a podstatného mena (výrazu kategórie N). (Poznamenávam, že to, čo som označil ako výraz kategórie AP, by v jazykoch obsahujúcich určitý člen, mohol byť výraz kategórie NP.)



V uvedenom stromovom diagrame sa vyskytujú syntaktické reprezentácie výrazov „Albert“, „je“ a „ženatý“. Okrem toho sa tu vyskytuje premenná *P*, ktorá zastupuje výrazy kategórie N. Premenná *P* môže nadobúdať ako svoje hodnoty – ak sa pohybuje v extenzionálnom rámci – množiny objektov. V stromovom diagrame, ktorý by reprezentoval logickú formu vety (1), by sa zase vyskytovala premenná, ktorá by ako obor premennosti mala binárne relácie.

Neartikulované zložky, s ktorými sme sa stretli pri sémantických obsahoch viet ako (1) alebo (5), sú v skutočnosti reprezentované na úrovni logickej formy v podobe implicitných premenných (vhodného druhu), ktoré sa nevyskytujú na úrovni povrchovej formy vety. Ich neartikulovanosť teda spočíva v tom, že na úrovni povrchovej formy vety im nezodpovedá žiadna explicitná syntaktická jednotka z vety, no to neznamená, že im nezodpovedá žiadna syntaktická jednotka na nejakej inej syntaktickej úrovni – ako sme videli, na úrovni logickej formy možno nájsť vhodnú premennú. Ide teda o *explicitne neartikulované zložky* vyjadreného obsahu, ktoré sú však artikulované implicitne.

3 Implicitne neartikulované zložky

Niektorí filozofi jazyka – konkrétne prívrženci tzv. *kontextualizmu* – tvrdia, že okrem explicitne neartikulovaných zložiek musíme rozlíšiť aj *implicitne neartikulované zložky* vyjadreného obsahu. Presnejšie povedané, niektoré zložky, ktoré nie sú explicitne artikulované, nie sú artiku-

lované ani len na implicitnej úrovni logickej formy. F. Recanati rozlišuje dva druhy pragmatických procesov – *saturáciu* a *voľné obohatenie* (pozri napríklad Recanati 1993 alebo 2004). Saturácia priradzuje hodnoty premenným alebo kontextovo citlivým syntaktickým jednotkám, a teda ak sa do vyjadreného obsahu dostane nejaká neartikulovaná zložka (ako hodnota premennej) na základe saturácie, ide o explicitne neartikulovanú zložku. Na druhej strane voľné obohatenie je voľné v tom zmysle, že nezávisí od žiadnej syntactickej jednotky vyskytujúcej sa vo vete. Tie zložky vyjadreného obsahu, ktoré sa doňho dostali na základe voľného obohatenia, sú teda nielen explicitne neartikulovanými, ale aj implicitne neartikulovanými zložkami daného obsahu. Analogicky K. Bach – ktorý sa však napriek niektorým podobnostiam nedá začleniť ku kontextualistickému prúdu, ale patrí skôr k sémantickým minimalistom (pozri Zouhar 2011d, 195-196) – rozlišuje *kompletizáciu* od *expanzie* (pozri Bach 1994), pričom kompletizácia v podstate zodpovedá Recanatiho saturácii a expanzia voľnému obohateniu; expanziou sa do vyjadrenej propozície dostanú také zložky, ktoré nie sú artikulované na žiadnej úrovni syntactickej reprezentácie, no kompletizáciou sa fakticky priradzujú hodnoty explicitným či implicitným syntaktickým jednotkám vo vete.

V akých prípadoch možno vo všeobecnosti hovoriť o tom, že určitá zložka vyjadreného obsahu sa doňho dostala na základe voľného obohatenia, resp. expanzie? V princípe sa to stáva vtedy, keď veta vyjadruje nejakú kompletnú propozíciu (vzhľadom na kontext jej použitia) aj za predpokladu, že sa do nej nedoplní žiadna ďalšia zložka. To znamená, že tie zložky vyjadreného obsahu, ktoré sú implicitne neartikulované, nie sú *mandatórne* v tom zmysle, že sú nevyhnutné na to, aby sa vetou vyjadrila (vzhľadom na daný kontext použitia) propozícia, keďže veta vyjadrí propozíciu nezávisle od nich (pozri napríklad Recanati 2002, 323). Opäť uvediem zopár ilustrácií.

Keď hovorca použije vetu

(6) Albert vytiahol kľúč a otvoril dvere,

môže tak vyjadríť neobohatenú propozíciu

(6a) »Albert vytiahol kľúč a otvoril dvere«

ktorá je pravdivá, ak Albert vytiahol kľúč a otvoril dvere, pričom sa v nej nekonkretizuje nič iné. Táto propozícia nešpecifikuje mnohé veci, napríklad to, odkiaľ Albert vytiahol kľúč, ani to, čím otvoril dvere.

Propozícia (6a) bude pravdivá napríklad v prípade, že Albert vytiahol kľúč a otvoril dvere tak, že týmto kľúčom odomkol a ešte ním pootočil v zámke tak, aby sa dvere samy otvorili. Bola by však pravdivá aj v prípade, že Albert vytiahol kľúč a otvoril dvere rukou bez toho, aby na to použil tento kľúč. Nech je to akokoľvek, (6a) je propozícia, ktorá môže byť pravdivá, resp. nepravdivá, a teda použitím vety (6) možno vyjadriť propozíciu bez akéhokoľvek voľného obohatenia či expanzie.¹²

Lenže použitím tej istej vety môže hovorca vyjadriť aj obohatenú propozíciu, napríklad

(6b) *»Albert vytiahol kľúč a otvoril dvere kľúčom, ktorý vytiahol«*

(6c) *»Albert vytiahol kľúč z vrečka a otvoril dvere kľúčom, ktorý vytiahol«*

Propozícia (6b) by bola pravdivá len v prípade, že Albert použil na otvorenie dverí kľúč, ktorý vytiahol (pričom sa nešpecifikuje, odkiaľ ho vytiahol), a nie niečo iné. Zložka *»kľúč, ktorý vytiahol«* sa do vyjadrenej propozície dostala na základe voľného obohatenia, resp. expanzie, a teda by malo ísť o zložku, ktorá je neartikulovaná aj v implicitnom zmysle. V prípade propozície (6c) sú zase neartikulovanými zložkami *»z vrečka«* a *»kľúč, ktorý vytiahol«*. Na úrovni logickej formy im nezodpovedá žiadna implicitná premenná, a teda sa v propozícii (6c) vyskytujú na základe voľného obohatenia, resp. expanzie.

Iný príklad predstavujú podľa Recanatiho tzv. meteorologické predikáty (pozri okrem iného Recanati 2007). Podľa štandardného názoru vyjadrujú vety z meteorologickými predikátmi ako „pršať“ atď. propozície, ktoré musia zahŕňať aj miesto, na ktorých sa daný meteorologický jav odohráva (t. j. miesto je mandatórnou zložkou takej propozície; pozri napríklad Perry 1986 alebo Neale 2007). Recanati však tento názor nezdieľa a tvrdí, že miesto sa do takejto propozície dostane len na základe voľného obohatenia, a teda aj bez tohto miesta bude veta vyjadrovať kompletnú propozíciu (pozri napríklad Recanati 2002, 317). To znamená, že ak hovorca použije vetu

(7) Prší,

môže tak vyjadriť bližšie nešpecifikovanú propozíciu

¹² Pravda, gramatické perfektum naznačuje, že do vyjadrenej propozície budú patriť aj časové okamihy, no tie sa do nej dostanú na základe saturácie, resp. kompletizácie.

(7a) »Prší«,¹³

ale aj obohatené propozície, napríklad:

(7b) »V Bratislave prší«

(7c) »Niekde prší«

Zložky »v Bratislave« a »niekde« sú neartikulované nielen v explicitnom, ale aj implicitnom zmysle, keďže (7a) je plnohodnotná propozícia, do ktorej sa nemusí nič dopĺňať.¹⁴ To znamená, že na úrovni (skrytej) logickej formy vety (7) sa nevyskytuje žiadna syntaktická reprezentácia týchto zložiek propozície podobne, ako sa žiadna ich syntaktická reprezentácia nevyskytuje ani na („viditeľnej“) úrovni povrchovej štruktúry vety (7).

4 Dva druhy neartikulovaných zložiek?

Zdá sa, že rozdiel medzi explicitnými a implicitnými neartikulovanými zložkami vyjadreného obsahu je netriviálny a zároveň dobre zdôvodnený. Kontextualisti ako F. Recanati, ale aj nekontextualisti ako K. Bach považujú za potrebné rozlišovať tieto dva druhy neartikulovaných zložiek a odmietajú redukovať neartikulované zložky jedného druhu na neartikulované zložky druhého druhu. Naozaj však potrebujeme neartikulované zložky oboch druhov? Presnejšie povedané, naozaj dokážeme zdôvodniť existenciu neartikulovaných zložiek oboch druhov? Všimnime si, že neartikulované zložky jedného aj druhého druhu sa postulujú na to, aby sa vysvetlilo, ako možno použitím nejakej vety vyjadriť (vzhľadom na nejaký kontext jej použitia) propozíciu, ktorá je bohatšia, ako obsah, ktorý dostaneme z doslovného významu

¹³ Dodávam, že táto propozícia sa explicitne zmieňuje o čas, v ktorom sa daný meteorologický jav odohráva, keďže sloveso je v gramatickom prézente, ktorý sa vzťahuje na aktuálny čas výpovede.

¹⁴ To, samozrejme, neznamená, že dážď alebo nejaký iný meteorologický jav by sa mohol odohrať bez toho, aby sa odohrával na nejakom mieste. Vyskytovať sa na nejakom mieste je podstatná vlastnosť meteorologických javov, no to je metafyzická otázka. Ak je *metafyzicky nevyhnutné*, aby pršalo na nejakom mieste (a v nejakom čase), neznamená to, že je *sémanticky* (či *komunikačne*) *nevyhnutné*, aby sa zmienka o mieste, na ktorom prší, vyskytovala aj vo vyjadrenej propozícii. Recanati preto odlišuje metafyzické neartikulované zložky od komunikačných neartikulovaných zložiek (pozri Recanati 2002, 305 a ďalej).

vety na základe minimálnej modifikácie. Domnievam sa však, že to by sme dokázali aj v prípade, že si vystačíme len s neartikulovanými zložkami jedného druhu. Konceptia, ktorá túto možnosť ponúka, sa nazýva *minimálny indexikalizmus* a rozpracoval som ju v knihe Zouhar (2011d); preto ju nebudem opakovať na tomto mieste. Skôr sa zameriame na jeden kľúčový problém, ktorý by sme nemohli riešiť, keby sme akceptovali oba druhy neartikulovaných zložiek propozícií; skutočnosť, že tento problém vzniká, je dôležitým (negatívnym) argumentom proti koncepciám, ktoré tieto dva druhy neartikulovaných zložiek odlišujú.¹⁵

Kľúčové tvrdenie minimálneho indexikalizmu spočíva v tom, že všetky zložky vyjadrenej propozície zodpovedajú explicitným alebo aspoň implicitným syntaktickým jednotkám, ktoré sa vyskytujú vo vete. To znamená, že podľa minimálneho indexikalizmu žiadne voľné obohatenie alebo expanzia neexistuje a že ak sa vo vyjadrenej propozícii vyskytuje nejaká neartikulovaná zložka, tak sa do nej dostala na základe saturácie, resp. kompletizácie. Dôvod spočíva v tom, že každý zdanlivý prípad voľného obohatenia, resp. expanzie, sa dá odhaliť ako skrytý prípad saturácie, resp. kompletizácie. Ako som však naznačil, detaily tohto riešenia nebudem na tomto mieste opakovať. Z hľadiska našich terajších úvah je podstatné to, že ak sa nejaká neartikulovaná zložka dostane do propozície na základe saturácie, resp. kompletizácie, problém, na ktorý poukážem, nevzniká. Preto demaskovanie voľného obohatenia, resp. expanzie ako saturácie, resp. kompletizácie eliminuje akékoľvek nebezpečenstvo, že narazíme na spomínaný problém.

Tvrďím, že ak sa akceptuje konceptia propozícií ako zložených a štruktúrovaných entít, a zároveň sa pripúšťa voľné obohatenie, resp. expanzia ako legitímny spôsob obohacovania propozície o nové zložky, tak hrozí nebezpečenstvo, že takto obohatené „propozície“ v skutočnosti prestanú byť propozíciami. Propozícia totiž má byť *jednou*

¹⁵ Presnejšie povedané, tento problém vzniká primárne v súvislosti s Bachovou koncepciou, v ktorej sa odlišuje kompletizácia od expanzie, nie v súvislosti s Recanatiho koncepciou odlišujúcou saturáciu od voľného obohatenia. Dôvod spočíva v tom, že Recanati rozpracoval aj určitú formálnu teóriu, ktorá má eliminovať problém, na ktorý poukážem v hlavnom texte. Zdá sa však, že Recanatiho formálna teória vychádza z určitých nedorozumení, na ktoré som poukázal v prácach Zouhar (2011c) a (2011d, 300-306). To znamená, že Recanati sa v podstate podarilo ochrániť svoju teóriu neartikulovaných zložiek pred prípadnými námietkami len za cenu chybnéj formálnej teórie, a preto ju v skutočnosti neochránil dostatočne.

(hoci zloženou a štruktúrovanou) entitou a doplnenie ďalších zložiek neumožní zachovať túto jednotu. Pokúsim sa túto tézu obhájiť.

Začnime krátkou úvahou o Fregeho sémantickej koncepcii. Predpokladajme, že propozícii zodpovedá to, čo Frege nazýva *myšlienka*. Myšlienka je zložený a štruktúrovaný abstraktný objekt, ktorý možno rozložiť na jednotlivé zložky podľa určitých pravidiel. Pri akejkolvek analýze myšlienky dostaneme ako jej zložky zmysly rôznych druhov: V každej myšlienke sa vyskytuje aspoň jeden zmysel, ktorý obsahuje aspoň jedno prázdne miesto, na ktoré možno dopĺňať iné zmysly vhodných druhov, a určitý počet zmyslov, ktoré neobsahujú prázdne miesta. Počet prázdnych miest v zmysle prvého druhu je totožný s počtom zmyslov druhého druhu. Zmysel, ktorý obsahuje aspoň jedno prázdne miesto, je spôsobom danosti funkcie (pojmu) a v jazyku ho možno vyjadriť pomocou predikátového výrazu; zmysel, ktorý neobsahuje žiadne prázdne miesto, je spôsobom danosti predmetu a v jazyku ho možno vyjadriť tým, čo Frege nazýva *vlastné meno*. Z nášho hľadiska je teraz dôležité to, že jednotlivé zmysly obsiahnuté v myšlienke spolu tvoria určitú jednotu; jednotu získame práve vďaka tomu, že jednotlivé zmysly bez prázdnych miest zaplňajú prázdne miesta v tom zmysle, ktorý ich má. Ak »() je ženatý« je zmysel, ktorý obsahuje jedno prázdne miesto (ktoré je naznačené zátvorkami), a »Albert« je zmysel bez prázdnych miest, tak jeho dosadením do prázdneho miesta vznikne ďalší zmysel (myšlienka) »(Albert) je ženatý«. Keby sa zmysly bez prázdnych miest nedosádzali podobným spôsobom na prázdne miesta v zmysloch, ktoré ich obsahujú, nebolo by možné utvoriť zložený, no jednotný (jeden) objekt. Útvar »(Albert)«, »() je ženatý« je len čírym zoznamom dvoch zmyslov, ktoré spolu netvoria žiadnu jednotu. Myšlienkou je taký sémantický objekt, ktorý pozostáva s aspoň jedného zmyslu s prázdny-
mi miestami, ktoré sú však zaplnené inými vhodnými zmyslami tak, aby žiadne miesto nezostalo prázdne. Keby aspoň jedno miesto zostalo prázdne, nemohli by sme daný útvar považovať za myšlienku.¹⁶ Analogickú úvahu možno formulovať v súvislosti s propozíciami v našom chápaní.

Procesy saturácie, resp. kompletizácie si môžeme predstaviť tak, že zodpovedajú zaplňaniu akýchsi prázdnych miest v určitých sémantických entitách ďalšími vhodnými sémantickými entitami. Saturácia

¹⁶ Podrobnejšie sa týmto aspektom Fregeho koncepcie venujem v knihe Zouhar (2010, 89-122), resp. v stati Zouhar (2011b).

a kompletizácia teda umožňujú tvoriť z určitých sémantických objektov ďalšie (jednotné) sémantické objekty. Dôvod spočíva v tom, že z nepropozíčných sémantických entít vhodných druhov tvorí propozície. Na druhej strane voľné obohatenie, resp. expanzia má utvárať z propozícií ďalšie propozície doplnením nových zložiek. Problém je však v tom, že ak sa niečo klasifikuje ako propozícia, tak taký objekt nemôže obsahovať žiadne prázdne miesta, keďže by v takom prípade nemohlo ísť o propozíciu. Lenže ak neobsahuje žiadne prázdne miesta, nemožno ich ani ničím zaplniť, a teda doplnením ďalšej sémantickej entity nevznikne jeden jednotný útvar, ale len číry zoznam dvoch sémantických entít – pôvodnej propozície a nového sémantického objektu.

Vezmime si príklad. Predpokladajme, že z propozície (6a) (*»Albert vytiahol kľúč a otvoril dvere«*) chceme utvoriť propozíciu (6b) (*»Albert vytiahol kľúč a otvoril dvere kľúčom, ktorý vytiahol«*) doplnením zložky *»kľúč, ktorý vytiahol«*. Keďže (6a) je propozícia, neobsahuje žiadne prázdne miesto. To napríklad znamená, že sa v nej vyskytuje binárny vzťah *»() otvoril ()«*, pričom prvé prázdne miesto obsadzuje (zamlčaná) zložka *»Albert«* (t. j. určitý človek) a druhé prázdne miesto zložka *»dvere«* (t. j., zjednodušene, určitý exemplifikát vlastnosti *byť dvere*). Doplnenie zložky *»kľúč, ktorý vytiahol«* (t. j., zjednodušene, určitého exemplifikátu vlastnosti *byť kľúčom, ktorý Albert vytiahol*) predpokladá, že sa tu namiesto binárneho vzťahu *»() otvoril ()«* vyskytuje ternárny vzťah *»() otvoril () ()«*, a teda nová zložka zaplní zostávajúce prázdne miesto. Lenže v propozícii (6a) sa žiadne prázdne miesto nevyskytuje. Ak má voľné obohatenie, resp. expanzia utvárať z propozícií nové propozície, nemôže teda vyžadovať len doplnenie ďalšej zložky, ale by malo vyžadovať zásadnejšiu modifikáciu, ktorá spočíva v tom, že vzťah určitej árnosti sa nahradí vzťahom s vyššou árnosťou a následne sa zaplní nové prázdne miesto. Konkrétnejšie, z propozície (6a) možno získať propozíciu (6b) len tak, že sa binárny vzťah *»() otvoril ()«* nahradí ternárnym vzťahom *»() otvoril () ()«* a na nové prázdne miesto sa dodá zložka *»kľúč, ktorý vytiahol«*.

Lenže voľné obohatenie a expanzia neumožňujú nahrádzať jeden sémantický objekt iným sémantickým objektom, ale len dopĺňajú nové objekty. Keďže je to tak, doplnením zložky *»kľúč, ktorý vytiahol«* k propozícii (6a) nemôže vzniknúť propozícia (6b) – t. j. jeden jednotný objekt – ale zoznam dvoch sémantických entít, ktorý môžeme znázorniť takto: (*»Albert vytiahol kľúč a otvoril dvere«*, *»kľúč, ktorý vytiahol«*). To je principiálny a neprekonateľný problém akéhokoľvek procesu, ktorý má fungovať podobne ako voľné obohatenie a expanzia. Opakujem,

že tento problém nevzniká v súvislosti so saturáciou a kompletizáciou. Preto sa mu vyhneme, ak sa zbavíme voľného obohatenia a expanzie; zároveň všetky prípady, na ktoré sa pôvodne tieto procesy aplikovali, demaskujeme ako prípady saturácie alebo kompletizácie. Znamenalo by to okrem iného napríklad to, že doslovným významom vety (6) nebude propozícia (6a), ale zodpovedajúci propozičný radikál, ktorý obsahuje namiesto binárneho vzťahu »() otvoril ()« ternárny vzťah »() otvoril () ()«, pričom jedno miesto v ňom zostalo nezaplnené. Iste, tento návrh predpokladá podrobné rozpracovanie a zdôvodnenie, no tým sa nechcem zaoberať na tomto mieste.¹⁷

Dôsledkom týchto úvah je skutočnosť, že ak sa odmietne voľné obohatenie a expanzia, pričom sa nahradia saturáciou, resp. kompletizáciou, tak sa stratí aj motivácia rozlišovať dva druhy neartikulovaných zložiek. Explicitne neartikulované zložky sa stávajú súčasťou propozície na základe saturácie, resp. kompletizácie; implicitne neartikulované zložky sa zase do propozície majú dostať na základe voľného obohatenia, resp. expanzie. Tie propozičné zložky, ktoré sa pôvodne považovali za implicitne neartikulované, sa v skutočnosti demaskujú ako explicitne neartikulované, keďže predsa len im zodpovedá určitá syntaktická reprezentácia vo vete. Môžeme teda uzavrieť, že všetky neartikulované zložky propozície sú len explicitne neartikulované.

Filozofický ústav SAV
Klemensova 19
813 64 Bratislava
Slovenská republika
marian.zouhar@gmail.com

Literatúra

- BACH, K. (1994): Conversational Implicature. *Mind & Language* 9, No. 2, 124-162.
- BEALER, G. (2004): An Inconsistency in Direct Reference Theory. *Journal of Philosophy* 101, 574-593.
- DUŽÍ, M. – JESPERSEN, B. – MATERNA, P. (2010): *Procedural Semantics for Hyperintensional Logic. Foundations and Applications of Transparent Intensional Logic*. Dordrecht: Springer.

¹⁷ Koncepciu, ktorá túto vec rieši, možno nájsť v práci Zouhar (2011d, 3. kapitola).

- JESPERSEN, B. (2003): Why the Tuple Theory of Structured Propositions Isn't a Theory of Structured Propositions. *Philosophia* 31, 171-183.
- KAPLAN, D. (1975): How to Russell a Frege-Church. *Journal of Philosophy* 72, No. 19, 716-729.
- KAPLAN, D. (1989): Demonstratives. In: Almog, J. - Perry, J. - Wettstein, H. (eds.): *Themes from Kaplan*. Oxford: Oxford University Press, 482-563.
- NEALE, S. (2007): On Location. In: O'Rourke, M. - Washington, C. (eds.): *Situating Semantics: Essays on the Philosophy of John Perry*. Cambridge (Mass.): MIT Press, 251-393.
- PERRY, J. (1977): Frege on Demonstratives. *Philosophical Review* 86, No. 4, 474-497.
- PERRY, J. (1986): Thoughts without Representation. *Proceedings of the Aristotelian Society* 60, 137-152.
- RACLAVSKÝ, J. (2009): *Jména a deskripcie: logicko-sémantická zkoumání*. Olomouc: Nakladatelství Olomouc.
- RECANATI, F. (1993): *Direct Reference: From Language to Thought*. Oxford: Blackwell Publishing.
- RECANATI, F. (2002): Unarticulated Constituents. *Linguistics & Philosophy* 25, 299-345.
- RECANATI, F. (2004): *Literal Meaning*. Cambridge: Cambridge University Press.
- RECANATI, F. (2007): It is Raining (Somewhere). *Linguistics & Philosophy* 30, 123-146.
- SOAMES, S. (1987): Direct Reference, Propositional Attitudes, and Semantic Content. *Philosophical Topics* 15, No. 1, 44-87.
- SOAMES, S. (2010): *What Is Meaning?* Oxford: Princeton University Press.
- TICHÝ, P. (1988): *Foundations of Frege's Logic*. Berlin: Walter de Gruyter.
- TICHÝ, P. (1994): Cracking the Language Code. *From the Logical Point of View*, No. 2, 7-19.
- ZOUHAR, M. (2010): *Problémy jazyka a počiatky analytickej filozofie*. Bratislava: aleph.
- ZOUHAR, M. (2011a): Kontext a sémantika predikátov. *Filosofický časopis* 59, č. 2, 187-205.
- ZOUHAR, M. (2011b): The Structure of Frege's Thoughts. *History and Philosophy of Logic* 32, No. 3, 199-209.
- ZOUHAR, M. (2011c): What is Wrong with Unarticulated Constituents? *Human Affairs* 21, 239-248.
- ZOUHAR, M. (2011d): *Význam v kontexte*. Bratislava: aleph.