V čom sa nemôžete mýliť?

Igor Sedlár[1]

Univerzita Komenského v Bratislave

Abstract: The paper sketches an analysis of the notion of a self-fulfilling belief in terms of doxastic modal logic. We point out a connection between self-fulfilling beliefs and Moore's paradox. Then we look at self-fulfilling beliefs in the context of neighborhood semantics. We argue that the analysis of several interesting self-fulfilling beliefs has to make essential use of propositional quantification.

Keywords: Self-fulfilling belief, doxastic logic, neighborhood semantics, propositional quantification, Moore's paradox.

Dôležitým rozdielom medzi pojmami *poznania* a *presvedčenia* je to, že poznanie na rozdiel od presvedčenia *implikuje pravdivosť*. Ak viete, že p, tak p musí byť pravda. Ekvivalentne, ak p nie je pravda, tak nemôžete vedieť, že p.[2] Na druhej strane, ľahko sa môže stať, že budete presvedčení o niečom, čo v skutočnosti nie je tak.

Inými slovami, „Ak x vie, že p, tak p" je pravda pre ľubovoľnú osobu x a propozíciu p, zatiaľ čo pri „Ak je x presvedčený, že p, tak p" to tak nie je. V našich presvedčeniach sa môžeme mýliť. Môžeme sa však mýliť vo všetkom? Neexistujú propozície, ktoré nemôžu byť nepravdivé, ak je o nich niekto presvedčený? Cieľom tejto štúdie je pouvažovať o tejto problematike na pozadí doxastickej logiky, teda formálno-logic-

[1] Tento text vznikol v rámci projektu *Sémantické modely, ich explanačná sila a aplikácia* podporeného grantom VEGA č. 1/0046/11. Práca na tomto článku bola podporená tiež Grantom UK pre mladých vedeckých pracovníkov č. UK/510/2011 a Štipendiom Vzdelávacej nadácie Jana Husa a Nadácie Univerzity P. J. Šafárika v Košiciach.

[2] Podmienka pravdivosti je pri analýzach pojmu poznania prijímaná všeobecne, pozri 8. kapitolu práce Huemer (2002). K analýze poznania pozri napríklad aj Audi (2003) a Lehrer (1999).

kého prístupu k pojmu presvedčenia.[3] Do veľkej miery pritom pôjde o vymedzenie základných pojmov a náčrt určitého okruhu problémov. Hlbšie skúmanie týchto problémov bude obsahom iných štúdií.

Našu otázku však môžeme chápať dvoma spôsobmi. Po prvé, môžeme sa pýtať na to, či existuje nejaká taká propozícia *p*, že ak je o jej pravdivosti presvedčený ktokoľvek, tak skutočne musí byť pravdivá. Po druhé, môžeme sa pýtať, či pre každého existuje nejaká taká propozícia, že ak je daný človek presvedčený o pravdivosti danej propozície, tak táto propozícia nemôže byť nepravdivá. Samozrejme, toto nie sú ekvivalentné otázky, aj keď z kladnej odpovede na prvú vyplýva kladná odpoveď na druhú otázku. My sa budeme venovať druhej, teda slabšej, otázke.

Ďalej, naša otázka sa nezaoberá možnosťou zdôvodnenia pravdivosti našich presvedčení, a teda nie je možné povedať, že by bola motivovaná primárne problémom skepticizmu. Je však zrejmé, že s možnosťami zdôvodnenia pravdivosti určitých presvedčení súvisí.

1 Moorov paradox

Náš problém je teda nasledovný. Nech si vezmeme ľubovoľnú osobu x, existuje pre x taká propozícia p, že nie je možné, aby bol x presvedčený o pravdivosti p, a zároveň by p bola nepravdivá? Ekvivalente, existuje také p, že „Ak je x presvedčený, že p, tak p“ nemôže byť nepravdivá? Takéto propozície by isto boli zaujímavými prípadmi, v ktorých sa hranica medzi poznaním a presvedčením začína strácať.

Nastal čas zaviesť formálny zápis, ktorý naše úvahy sprehľadní. Nech p je výroková premenná zastupujúca ľubovoľný výrok, ďalej nech a je označenie nejakej konkrétnej osoby a B_a nech znamená „a je presvedčený/á, že…“. Našou otázkou teda je, či existujú také hodnoty premennej p, že

(1) $B_a p \to p$

je nevyhnutne pravdivá.

Prvé priblíženie k riešeniu tohto problému prichádza z trocha nečakanej strany. Je ňou známy epistemologický problém, *Moorov para-*

---

3 Klasickou prácou v tejto oblasti je Hintikka (1962). K formálno-logickému prístupu k epistemologickým pojmom pozri napríklad aj Fagin – Halpern – Moses – Vardi (1995).

dox.[4] Jeho podstata je nasledovná. Je jasné, že o každom, kto by sa hlásil k protirečivému tvrdeniu typu „2 + 2 = 4 a zároveň 2 + 2 ≠ 4“, by sme povedali, že si protirečí. Nikto nemôže byť racionálne presvedčený o pravdivosti takéhoto tvrdenia. Zdá sa však, že rovnako nemôžeme byť racionálne presvedčení ani o pravdivosti niektorých tvrdení, *ktoré sami osebe nie sú protirečivé*. Príkladom takéhoto tvrdenia je „2 + 2 = 4, ale ja neverím, že 2 + 2 = 4“. Toto tvrdenie nie je protirečivé, pretože môže byť pravdivé. Ak by som bol trestuhodne matematicky negramotný, mohol by som neveriť, že 2 + 2 = 4. Skúsme si však predstaviť, že by som *veril* v pravdivosť tvrdenia „2 + 2 = 4, ale ja neverím, že 2 + 2 = 4“. Môžeme si to vôbec predstaviť? Takéto presvedčenie je neprípustné rovnako ako presvedčenie o pravdivosti protirečivého tvrdenia, avšak samotné tvrdenie „2 + 2 = 4, ale ja neverím, že 2 + 2 = 4“ protirečivé nie je!

Je pritom zrejmé, že táto situácia nastáva, nech namiesto „2 + 2 = 4“ použijeme *akýkoľvek* výrok. To znamená, že

(2) $\neg B_a (p \wedge \neg B_a p)$

je nevyhnutne pravdivá. Neexistuje taká propozícia p, že a by mohol byť presvedčený o „Je pravda, že p, no ja nie som presvedčený o tom, že p.“

Ako to však súvisí s našim pôvodným problémom? Jednoducho. Dá sa totiž ukázať, že (2) je ekvivalentná s určitým variantom (1). Predpokladajme, že $\neg B_a (p \wedge \neg B_a p)$. Dôležité je uvedomiť si, že presvedčený o dvoch tvrdeniach X, Y by som mal byť práve vtedy, ak som presvedčený aj o ich konjunkcii $X \wedge Y$. Z nášho predpokladu teda vyplýva $\neg(B_a p \wedge B_a \neg B_a p)$. To však na základe výrokovej logiky znamená $\neg B_a \neg B_a p \vee \neg B_a p$, čiže

(3) $B_a \neg B_a p \rightarrow \neg B_a p$

Táto formula je to, čo sme hľadali. Uvádza totiž príklad tvrdenia, ktoré nemôže byť nepravdivé v prípade, že a je presvedčený o jeho pravdivosti! Takýmto tvrdením je akákoľvek substitučná inštancia formuly $\neg B_a p$, teda napríklad tvrdenie „a neverí, že 2 + 2 = 4“. Ak teda súhlasíme s (2), tak a sa pri $\neg B_a p$ nemôže mýliť.

---

[4] Zaujímavú kolekciu súčasných článkov o Moorovom paradoxe nájdete v Green – Williams (2007).

2 Minimálne modely

Je tu však jeden problém. Platí skutočne princíp, ktorý sme použili v prvom kroku nášho zdôvodnenia? Prehľadne ho môžeme vyjadriť ako

(4) $B_a (p \wedge q) \leftrightarrow (B_a p \wedge B_a q)$

Pravdou je, že v doxastickej logike sa jeho platnosť často predpokladá. Platnosť (4) priamo vyplýva zo spôsobu, akým sa v sémantike tzv. *normálnych* doxastických logík interpretuje operátor presvedčenia B_a.

Aby sme to jasne uvideli, stručne si túto sémantiku predstavme. *Normálne doxastické modely* sú trojice $M = (W,R,V)$, kde W je neprázdna množina (neformálne ju môžeme chápať ako množinu „možných situácií" alebo „možných svetov"), R je binárna relácia na W a V je ohodnotenie, teda funkcia, ktorá každej výrokovej premennej priradí nejakú množinu svetov (neformálne, $V(p)$ je množina svetov, v ktorých je p pravdivé). Základom doxastickej interpretácie týchto modelov je relácia R.[5] Fakt, že dva svety s, t sú v relácii R (čo zapíšeme ako Rst), totiž neformálne zodpovedá skutočnosti, že vo svete s pokladá a svet t za možný. Inými slovami, a nemá v s dosť silné presvedčenia na to, aby mohol na základe nich vylúčiť možnosť, že skutočnému svetu zodpovedá t. Ak je však a vo svete s presvedčený o pravdivosti nejakej propozície, napríklad p, tak p musí byť pravdivá v každom svete, ktorý a pokladá za možný.

Technicky, $B_a p$ platí vo svete s práve vtedy, keď p platí v každom takom s', že Rss'. Nejaká formula je potom *logicky pravdivá* práve vtedy, keď je pravdivá v každom bode každého normálneho doxastického modelu.[6] Teraz môžeme ukázať, že (4) je logicky pravdivá. Predpokladajme, že vo svete s platí $B_a (p \wedge q)$. To znamená, že $(p \wedge q)$ platí v každom takom s', že Rss'. To však na základe vlastností konjunkcie platí práve vtedy, keď v každom takom s' platí p a zároveň q. V pôvodnom s teda musí platiť ako $B_a p$, tak aj $B_a q$. To znamená, že v s platí $B_a p \wedge B_a q$.

---

5 Predpokladajme, že chceme modelovať presvedčenia iba jednej osoby, povedzme a. V takomto prípade nám postačí jedna relácia. V prípade, že chceme modelovať presvedčenia viacerých osôb naraz, musíme do modelu zahrnúť pre každú osobu jednu reláciu.

6 Stručný úvod do epistemickej a doxastickej logiky a relačnej sémantiky nájdete v Meyer (2001). K relačnej sémantike a k modálnej logike všeobecne pozri napríklad Hughes – Cresswell (1996), z novších prác Blackburn – de Rijke – Venema (2001) či van Benthem (2010).

Skutočná platnosť (4) je však problematická a v epistemológii sa o nej diskutuje.[7] Doxastická logika by nemala nahradzovať epistemologické teórie, mala by skôr slúžiť ako nástroj, ktorý umožní niektoré diskusie a problémy sprehľadniť. Ak teda s princípom (4) nesúhlasíte, mali by sme opustiť pôdu normálnych doxastických logík a využiť služby menej „teóriou zaťažených" logík.

Toto nie je situácia, ktorá by bola v epistemickej a doxastickej logike novinkou. Normálne doxastické (a epistemické) logiky sú často kritizované za to, že v nich platia princípy, o ktorých by sme len ťažko mohli povedať, že opisujú reálne osoby. Môžeme si napríklad overiť, že keď si vezmeme akúkoľvek logicky pravdivú formulu X, tak aj formula $B_a X$ je logicky pravdivá. To však znamená, že podľa normálnych doxastických logík by a mal byť presvedčený o každej logickej pravde.[8] Východiskom je opustenie spôsobu interpretácie operátora presvedčenia, ktorý je súčasťou relačnej sémantiky. Inými slovami, musíme vytvoriť novú interpretáciu, čiže konštruovať *slabšie* doxastické logiky.

Jednou rodinou takýchto logík sú tzv. *klasické* doxastické logiky, založené na *minimálnych modeloch*. Sémantika týchto logík (budeme o nej hovoriť ako o minimálnej sémantike) vychádza z chápania propozície ako množiny možných svetov.[9] Relácia R, známa z relačnej sémantiky, sa v minimálnej sémantike nahrádza funkciou, ktorá každému svetu priraďuje množinu propozícií, ktoré majú vzhľadom na daný svet nejaké význačné postavenie.[10] Teda namiesto toho, aby bol napríklad svet s v rámci modelu v relácii R so svetmi s', s'',… , je svetu s priradená množina, ktorej prvkami sú nejaké propozície, teda množiny svetov.

Formálne, *minimálny doxastický model* je trojica $M = (W,N,V)$, kde W je neprázdna množina, N je funkcia, ktorá každému prvku W priradí nejakú množinu množín prvkov W a V je ohodnotenie.[11] Doxastická

---

7 Pozri napríklad Kyburg (1970) a (1997).

8 Viac k tomuto problému a možným riešeniam nájdete napríklad v 9. kapitole Fagin – Halpern – Moses – Vardi (1995).

9 K tomuto chápaniu propozícií pozri napríklad Stalnaker (1976), jeho kritiku je možné nájsť napríklad v Soames (1987).

10 Ku klasickým logikám a minimálnej sémantike pozri Segerberg (1971) a Chellas (1980). Minimálnu sémantiku nezávisle od seba formulovali D. Scott a R. Montague; pozri Montague (1970) a Scott (1970).

11 Minimálnu sémantiku môžeme chápať ako zovšeobecnenie relačnej sémantiky. Relačná sémantika „spája" svety so svetmi, minimálna sémantika spája svety s množinami svetov.

interpretácia minimálnej sémantiky je nasledovná. Predpokladajme, že nás zaujíma iba jedna osoba, napríklad a. Funkcia N špecifikuje, o ktorých propozíciách je a v jednotlivých svetoch presvedčený. V súlade s doxastickou interpretáciou budeme funkciu N označovať N_a.[12] Majme napríklad $W = \{s_1, s_2, s_3\}$ a nech $N_a(s_1) = \{\{s_1, s_3\}, \{s_1, s_2\}\}$, pričom $V(p) = \{s_1, s_3\}$ a $V(q) = \{s_1, s_2\}$. Takto je v rámci minimálnej sémantiky vyjadrený fakt, že a je vo svete s_1 presvedčený o p a q.

Interpretácia operátora B_a tomu zodpovedá. Vo svete s platí $B_a\ X$ práve vtedy, keď množina svetov, v ktorých X platí („propozícia, vyjadrená formulou X"), patrí do $N_a(s)$ Ľahko si môžeme uvedomiť, že keď na funkciu N_a nekladieme žiadne obmedzenia, vyššie uvedené problematické princípy neplatia. Minimálne modely nás napríklad nenútia pripustiť, že sme presvedčení o každej X, pravdivej v každom svete každého modelu, teda o každej logicky pravdivej X. Takáto X v rámci každého modelu vyjadruje „maximálnu propozíciu", totožnú s množinou svetov daného modelu W (logicky pravdivá formula je pravdivá všade, v každom svete). Funkciu N_a ale nič nenúti k tomu, aby množinu W priradila každému svetu každého modelu. Inými slovami, môžeme mať model a v ňom svet s tak, že W nepatrí do množiny propozícií $N_a(s)$. Formula X môže byť platná, no $B_a\ X$ platná byť nemusí. Podobne si viete overiť, že v rámci minimálnych modelov nie je platný ani princíp (4).

Dôležitou otázkou teraz je, či existujú také p, že (1) je logicky pravdivá vzhľadom na minimálne modely. Existujú propozície, pri ktorých sa nemôžeme mýliť, aj keď nepredpokladáme silné princípy typu (4)? Dá sa ukázať, že existujú. Vezmime si napríklad nasledujúcu propozíciu:

(5) $\{s \mid N_a(s) \neq \emptyset\}$

Tvoria ju také svety, ktorým funkcia N_a nepriraďuje prázdnu množinu. Sú to teda svety, v ktorých je a o niečom presvedčený. Môžeme mať taký svet t, že a je v ňom o tejto propozícii presvedčený, no zároveň v ňom nie je pravdivá? Kedy však povieme, že nejaká propozícia ako množina svetov je v nejakom svete pravdivá? Odpoveď je jednoduchá. Nejaká propozícia je v určitom svete pravdivá práve vtedy, keď je daný svet jej prvkom. Môže t nebyť prvkom (5) v prípade, že (5) patrí do $N_a(t)$? Je jasné, že nemôže. Ak by totiž nebol jej prvkom, tak

---

12 Ak by sme chceli modelovať presvedčenia viacerých osôb $a,b,c,\ldots$, súčasťou modelu by boli viaceré funkcie $N_a, N_b, N_c, \ldots$.

by bol prvkom jej doplnku, teda množiny všetkých svetov, ktoré do danej propozície nepatria. Doplnok propozície (5) je však množina $\{s \mid N_a(s) = \emptyset\}$. Ak by t patril do *tejto* množiny, tak by jasne platilo $N_a(t) = \emptyset$. Lenže potom by a nemal v t žiadne presvedčenia, čiže by nemohol byť presvedčený ani o (5)!

Ďalším príkladom je propozícia

(6) $\{s \mid s \notin \bigcap (N_a(s))\}$

Túto propozíciu tvoria také svety s, ktoré nepatria aspoň do jednej propozície, ktorej v s náš a verí. Propozícia (6) teda zodpovedá tvrdeniu, že a má aspoň jedno nepravdivé presvedčenie. Môže však byť a presvedčený o tomto tvrdení bez toho, aby bolo zároveň pravdivé? Je jasné, že nemôže. Ak by totiž bolo nepravdivé, tak by a nemal nepravdivé presvedčenie, čiže všetky jeho presvedčenia by boli pravdivé. Ak by bol zároveň presvedčený o tom, že aspoň jedno jeho presvedčenie je nepravdivé, tak by skutočne aspoň jedno jeho presvedčenie muselo byť nepravdivé. Tým sa dostávame do sporu.

3 Sebapotvrdzujúce presvedčenia

Tu sa dostávame k všeobecnej otázke, ktoré propozície okrem (5) a (6) nemôžu byť nepravdivé v prípade, že im a verí. Pre ktoré hodnoty p je (1) platná v rámci minimálnych modelov? Je jasné, že na to, aby sme na túto otázku mohli dať uspokojivú odpoveď, musíme ju položiť v kontexte jazyka doxastickej logiky. Pýtame sa teda, pre aké formuly X je $B_a X \to X$ platná v rámci minimálnych modelov.

Najprv si však musíme vyjasniť, *aký* doxastický jazyk nás zaujíma. Samozrejme, musí to byť jazyk, ktorý obsahuje doxastický operátor B_a a známe výrokové spojky. Stačí však takýto jednoduchý jazyk na opísanie propozícií, ktoré nás zaujímajú? Dá sa ukázať že nestačí. V prvom rade, takýto jazyk nestačí ani na popísanie propozícií (5) a (6). Presnejšie, neexistuje formula X obsahujúca iba výrokové premenné spojky a operátor B_a, ktorá by v ľubovoľnom minimálnom modeli M a svete s platila práve vtedy, keď s patrí do (5), a podobne je to aj v prípade propozície (6).

Ukážme si to na prípade propozície (5). Ako sme zdôraznili, táto propozícia v podstate tvrdí, že existuje nejaká propozícia, o ktorej je a presvedčený. Samozrejme, ak v s platí napríklad $B_a\, p$, tak existuje propozícia, o ktorej je a v s presvedčený – je ňou práve propozícia, vyjadrená v rámci modelu premennou p. Opačná implikácia však neplatí vo všeobecnosti. Ak je $N_a(s)$ neprázdna, tak to ešte neznamená, že medzi jej prvky patrí práve propozícia vyjadrená v rámci modelu premennou p! Ľahko si môžeme uvedomiť, že niečo podobné môžeme povedať o propozícii vyjadrenej *ľubovoľnou* formulou minimálneho doxastického jazyka. To, že a verí v *nejakú* propozíciu, ešte neznamená, že verí v *práve túto* propozíciu.

Aké rozšírenie minimálneho doxastického jazyka nám umožní uviesť formulu, ktorá platí v (M,s) práve vtedy, keď s patrí do (5)? Sú tu dve možnosti. Po prvé, takúto formulu môžeme zhotoviť pomocou *propozičnej kvantifikácie*.[13] Je to prirodzená voľba: Ak (5) hovorí o existencii nejakého presvedčenia bez toho, aby sa špecifikovalo, o aké konkrétne presvedčenie ide, tak na jej vyjadrenie potrebujeme formulu, ktorá dokáže niečo podobné. Definícia formuly doxastického jazyka s propozičnou kvantifikáciou obsahuje dodatočnú klauzulu, ktorá je presne tým, čo potrebujeme: Ak je X formula, tak aj $\exists pX$ je formula. Potom sa dá jednoducho ukázať, že formula $\exists pB_ap$ platí v (M,s) práve vtedy, keď s patrí do (5). Vo svetle vyššie uvedenej diskusie to však znamená, že formula

(7) $B_a\exists pB_ap \rightarrow \exists pB_ap$

je platná v rámci minimálnych modelov. Ak verím, že v niečo verím, tak skutočne v niečo verím. Formula $\exists pB_ap$ je teda príkladom tvrdenia, ktoré nemôže byť nepravdivé v prípade, že a je presvedčený o jeho pravdivosti. Propozícia (6) je podobným spôsobom vyjadrená formulou $\exists p(B_ap \wedge \neg p)$. Formula $\exists p(B_ap \wedge \neg p)$ je teda ďalším príkladom tvrdenia, o ktorom a nemôže byť nepravdivo presvedčený. Všimnime si však, že tieto fakty platia iba o osobe a, explicitne spomenutej v uvedených formulách. Pre iné osoby platiť nemusia.

Po druhé, podobný účinok bude mať rozšírenie minimálneho doxastického jazyka o možnosť formulovať nekonečné disjunkcie. Efekt existenčnej kvantifikácie v $\exists pB_ap$ totiž môže simulovať nekonečná disjunkcia $B_aX_1 \vee B_aX_2 \vee \ldots$, pre všetky formuly X_i nášho jazyka.

Tu si však musíme uvedomiť, že naša pôvodná otázka sa musí spresniť. Ak sa totiž pýtame na formuly (doxastického jazyka s pro-

---

[13] Klasickými prácami na tému propozičnej kvantifikácie v modálnej logike sú Bull (1969) a Fine (1970). K propozičnej kvantifikácii v kontexte minimálnych modelov pozri Gabbay (1971).

pozičnou kvantifikáciou) X, pre ktoré je $B_a X \rightarrow X$ platná v rámci minimálnych modelov, tak naša otázka má určité triviálne odpovede, ktoré by sme mali eliminovať. Po prvé, $B_a X \rightarrow X$ je platná pre všetky platné X. Po druhé, uvedená implikácia je platná pre všetky také X, že $\neg B_a X$ je platná. Je jasné, že sa nemôžem mýliť, ak verím v propozíciu, ktorá nemôže byť nepravdivá. Je tiež jasné (aj keď toto znie možno trocha neintuitívne), že nemôžem mať nepravdivé presvedčenie o propozícii, o ktorej nemôžem byť vôbec presvedčený.[14] Takéto triviálne prípady však isto nie sú uspokojivými odpoveďami na našu pôvodnú otázku, ktorú preto musíme spresniť.

Povieme, že formula doxastického jazyka s propozičnou kvantifikáciou X *vyjadruje sebapotvrdzujúce presvedčenie pre a* práve vtedy, keď sú splnené nasledujúce tri podmienky:

1. $B_a X \rightarrow X$ je platná vzhľadom na minimálne modely.
2. X nie je platná vzhľadom na minimálne modely (existuje teda model M a svet s tak, že X nie je pravdivá v (M,s)).
3. $B_a X$ je v minimálnych modeloch splniteľná (existuje teda (M,s), v ktorom je $B_a X$ pravdivá).

Našu otázku teraz môžeme položiť presnejšie: *Ktoré formuly spĺňajú tieto podmienky*?

Ideálne by určite bolo, ak by existovala mechanická procedúra (algoritmus), ktorá by nám pre ľubovoľnú formulu X povedala, či spĺňa uvedené podmienky, alebo ich nespĺňa. Ak by nás zaujímalo, či nejaká formula vyjadruje sebapotvrdzujúce presvedčenie, stačilo by aplikovať takúto procedúru a v konečnom čase by sme sa dozvedeli výsledok. Problém je však v tom, že žiadna takáto procedúra neexistuje.

Dá sa totiž ukázať, že modálne logiky s propozičnou kvantifikáciou, ktoré sú dostatočne slabé na to, aby pripúšťali doxastickú interpretáciu, sú nerozhodnuteľné. K tomu pozri napríklad Fine (1970). Fine ukazuje, že najslabšia normálna modálna logika K obohatená o propozičnú

[14] V súvislosti s formulou $\exists p(B_a p \wedge \neg p)$ sa vynára dôležitá otázka. Zdá sa, že až príliš pripomína formulu $B_a p \wedge \neg p$, o ktorej platí niečo podobné ako o $p \wedge \neg B_a p$: Nezdá sa byť možné, aby o nej bol *a* racionálne presvedčený (v súvislosti s touto formulou sa hovorí o *komisívnej* forme Moorovho paradoxu, zatiaľ čo v súvislosti s $p \wedge \neg B_a p$ sa hovorí o jeho *omisívnej* forme). Platí teda $\neg B_a (p \wedge \neg B_a p)$, ba dokonca všeobecne $\neg \exists p B_a (p \wedge \neg B_a p)$. Je však zrejmé, že $\neg B_a \exists p (p \wedge \neg B_a p)$ nie je dôsledkom $\neg \exists p B_a (p \wedge \neg B_a p)$. Teda ak aj predpokladáme platnosť $\neg \exists p B_a (p \wedge \neg B_a p)$, $\neg B_a \exists p (p \wedge \neg B_a p)$ platná byť nemusí.

kvantifikáciu je nerozhodnuteľná. Množina formúl doxastického jazyka s propozičnou kvantifikáciou platných vo všetkých minimálnych modeloch je podteóriou takto obohatenej *K*, postavenou na tom istom jazyku. To však znamená, že táto množina je tiež nerozhodnuteľná.[15] Inými slovami, neexistuje algoritmus, ktorý by pre ľubovoľnú formulu *X* rozhodol, či je platná vo všetkých minimálnych modeloch, alebo nie je. Je však zrejmé, že hypotetický algoritmus, ktorý by rozhodoval, či ľubovoľná *X* spĺňa podmienky 1 – 3, by musel rozhodnúť, či formuly $B_a X \rightarrow X$, *X* a $\neg B_a X$ sú univerzálne platné, alebo nie sú.

4 Záver

Neexistencia algoritmu, ktorý by pre ľubovoľnú formulu *X* rozhodol, či vyjadruje sebapotvrdzujúce presvedčenie, však nie je dôvodom na paniku. Existujú mnohé teórie, ktoré sa často využívajú aj napriek svojej nerozhodnuteľnosti. Dobrým príkladom je logika prvého rádu. Ako by sme však mohli využiť predstavenú teóriu sebapotvrdzujúcich presvedčení, založenú na minimálnych modeloch?

V prvom rade nám umožňuje skúmať sebapotvrdzujúce presvedčenia formálnymi prostriedkami. Aké vlastnosti má množina formúl, vyjadrujúcich sebapotvrdzujúce presvedčenia? Aké vlastnosti majú formuly, ktoré do nej patria? Podobné otázky môžu viesť k zaujímavým zisteniam.

Po druhé, veľkou prednosťou tejto teórie je jej všeobecnosť. Všimnime si, že minimálne modely tvorí akákoľvek neprázdna množina, akákoľvek funkcia, ktorá každému prvku tejto množiny priraďuje určitý počet množín prvkov danej množiny, a ohodnotenie, teda funkcia, ktorá určitým základným symbolom nejakého jazyka priradí množiny prvkov danej množiny. Fenomény, ktorými sa táto stať zaoberala, je teda možné nájsť v akejkoľvek predmetnej oblasti, ktorú môžeme reprezentovať minimálnymi modelmi. Je pritom jedno, ako minimálne modely neformálne interpretujeme. Skúmané fenomény sú totiž nezávislé od našej pôvodnej motivácie – *W* nemusí predstavovať „možné svety“ a množinu $N_a(s)$ nemusíme interpretovať ako množinu propozícií, ktorým *a* verí v *s*.

Aké zaujímavé vlastnosti majú sebapotvrdzujúce presvedčenia? V akých oblastiach môžeme nájsť podobné fenomény? Tieto problémy čakajú na hlbšie preskúmanie.

---

15 Pozri aj Tarski – Mostowski – Robinson (1953).

Katedra logiky a metodológie vied
Filozofická fakulta UK
Šafárikovo nám. 6
814 99 Bratislava
Slovenská republika
sedlar@fphil.uniba.sk

LITERATÚRA

Audi, R. (2003): *Epistemology. A Contemporary Introduction to the Theory of Knowledge*. (2. vydanie). Londýn: Routledge.

Blackburn, P. – de Rijke, M. – Venema, Y. (2001): *Modal Logic*. Cambridge: Cambridge University Press.

Bull, R. (1969): On modal logic with propositional quantifiers. *The Journal of Symbolic Logic* 34, No. 2, 257-263.

Chellas, B. (1980): *Modal Logic. An Introduction*. Cambridge: Cambridge University Press.

Fagin, R. – Halpern, J. – Moses, Y. – Vardi, M. (1995): *Reasoning about Knowledge*. Cambridge, Massachusetts: MIT Press.

Fine, K. (1970): Propositional quantifiers in modal logic. *Theoria* 36, No. 3, 336-346.

Gabbay, D. M. (1971): Montague type semantics for modal logics with propositional quantification. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik* 17, No. 3, 245-249.

Green, M. – Williams, J. N. (eds.) (2007): *Moore's Paradox. New Essays on Belief, Rationality, and the First Person.* Oxford: Oxford University Press.

Hintikka, J. (1962): *Knowledge and Belief*. Ithaca: Cornell University Press.

Huemer, M. (ed.) (2002): *Epistemology. Contemporary Readings*. Londýn: Routledge.

Hughes, G. E. – Cresswell, M. J. (1996): *A New Introduction to Modal Logic*. Londýn: Routledge.

Kyburg, H. E. (1970): Conjunctivitis. In: Swain, M. (ed.): *Induction, Acceptance, and Rational Belief*. Dordrecht: D. Reidel.

Kyburg, H. E. (1997): The rule of adjunction and reasonable inference. *The Journal of Philosophy* 94, No. 3, 109-125.

Lehrer, K. (1999): *Teória poznania*. Bratislava: Infopress.

Meyer, J.-J. C. (2001): Epistemic logic. In: Goble, L. (ed.): *The Blackwell Guide to Philosophical Logic*. Oxford: Blackwell Publishers.

Montague, R. (1970): Universal grammar. *Theoria* 36, 373-398.

Scott, D. (1970): Advice in modal logic. In: Lambert, K. (ed.): *Philosophical Problems in Logic*. Dordrecht: D. Reidel, 143-173.

Segerberg, K. (1971): *An Essay in Classical Modal Logic*. Uppsala: The Philosophical Society.

Soames, S. (1987): Direct reference, propositional attitudes, and semantic content. *Philosophical Topics* 15, 47-87.

Stalnaker, R. (1976): Propositions. In: MacKay, A. – Merrill, D. (eds.): *Issues in the Philosophy of Language*. New Heaven: Yale University Press, 79-91.

Tarski, A. – Mostowski, A. – Robinson, R. (1953): *Undecidable Theories*. Amsterdam: North-Holland Publishing Company.

van Benthem, J. (2010): *Modal Logic for Open Minds*. Stanford: CSLI Publications.