

# FIKCJONALIZM MODALNY

Piotr Warzoszczak



OŚRODEK  
BADAŃ  
FILOZOFICZNYCH

# **BIBLIOTEKA OŚRODKA BADAŃ FILOZOFICZNYCH**

**Komitety redakcyjne serii:** Marta Bucholc, Tadeusz Ciecierski,  
Paweł Grabarczyk, Witold Hensel, Sebastian Szymański

Copyright©2014

Piotr Warzoszczak and *Ośrodek Badań Filozoficznych*

**ISBN: 978-83-931671-8-0**

**Recenzja wydawnicza:** prof. UŁ. dr hab. Janusz Kaczmarek

**Redakcja i korekta:** Jagna Świdarska

**Projekt okładki:** Karolina Pluta

**Skład komputerowy:** Juliusz Doboszewski

Książka wydana ze środków subsydium profesorskiego Fundacji na rzecz Nauki Polskiej, przyznanego prof. drowi hab. T. Szubce na realizację tematu „Filozofia analityczna: historia i najnowsze wyzwania”.

**Ośrodek Badań Filozoficznych**

**Warszawa 2014**

# Spis treści

<b>Wstęp</b>	<b>3</b>
<b>I Dyskurs niezobowiązujący ontologicznie</b>	<b>14</b>
1 Teoria uprawdziwiaczy . . . . .	16
2 Superweniencja prawdziwości na bycie . . . . .	37
3 Ufundowanie prawdziwości na bycie . . . . .	48
<b>II Dwa typy fikcjonalizmu</b>	<b>57</b>
1 Wspólne cechy różnych odmian fikcjonalizmu . . . . .	59
2 Metafikcjonalizm <i>versus</i> fikcjonalizm przedmiotowy .	75
2.1 Metafikcjonalizm . . . . .	84
2.2 Fikcjonalizm przedmiotowy . . . . .	86
3 Spór w fikcjonalistycznej rodzinie . . . . .	98
<b>III Fikcjonalizm modalny</b>	<b>106</b>
1 Metafikcjonalizm modalny . . . . .	108
1.1 Fikcja realisty modalnego . . . . .	110
1.1.1 Lewisowska koncepcja światów możli- wych . . . . .	110
1.1.2 Teoria odpowiedników . . . . .	127
1.2 Metafikcjonalistyczne zastosowania fikcji reali- sty modalnego . . . . .	141
2 Przedmiotowy fikcjonalizm modalny . . . . .	143

3	Fikcjonalizm modalny — rewolucyjny projekt czy zastana praktyka? . . . . .	146
4	Fikcjonalizm modalny D. Armstronga . . . . .	148
<b>IV Fikcjonalizm modalny pod lupą</b>		<b>150</b>
1	Wiedza o światach możliwych . . . . .	150
2	Obiekcja Brocka-Rosena . . . . .	153
3	Problemy z funktorem opowieści . . . . .	171
4	Problem drobnoziarnistości treści fikcji . . . . .	204
4.1	Treść możliwościwiatowych twierdzeń jednostkowych . . . . .	215
4.2	Treść ogólnych twierdzeń możliwościwiatowych	245
4.3	Co z ontologią reprezentacji mentalnych? . . .	251
5	Inne problemy z treścią fikcji . . . . .	253
5.1	Problem niekompletności treści fikcji . . . . .	255
5.2	Problem sztuczności treści fikcji . . . . .	265
6	Trudności umiarkowanego fikcjonalizmu modalnego .	271
6.1	Fakty modalne — pierwsze podejście . . . . .	276
6.2	Rola fikcji światów możliwych . . . . .	280
6.3	Problem ze wzorcami egzemplifikacji . . . . .	289
<b>V Stanowiska konkurencyjne</b>		<b>297</b>
1	Instumentalizm <i>versus</i> fikcjonalizm . . . . .	297
2	Modalizm <i>versus</i> fikcjonalizm modalny . . . . .	308
<b>Zakończenie</b>		<b>324</b>
<b>Indeks</b>		<b>329</b>
<b>Bibliografia</b>		<b>331</b>

# Wstęp

Przed kilkoma dekadami światy możliwe stały się stałym elementem filozoficznego krajobrazu. Swoją ugruntowaną pozycję zawdzięczają głównie zastosowaniu ich do analizy treści twierdzeń<sup>1</sup> modalnych, tj. twierdzeń, w których występują funktory modalne takie jak „jest konieczne, że...”, „jest możliwe, że...”, „jest przygodne, że...”. Standardowo analizuje się je według następujących schematów:

- (N) jest konieczne, że  $\alpha$  wtedy i tylko wtedy, gdy (wtw) dla każdego świata możliwego  $w$ , jeśli  $w$  jest osiągalne ze świata aktualnego @, to  $\alpha$  w  $w$ ;
- (P) jest możliwe, że  $\alpha$  wtw istnieje taki świat możliwy  $w$ , że  $w$  jest osiągalne ze świata aktualnego @ i  $\alpha$  w  $w$ ;
- (C) jest przygodne, że  $\alpha$  wtw  $\alpha$  w @ i istnieje taki świat możliwy  $w$ , że  $w$  jest osiągalne z @ i nieprawda, że  $\alpha$  w  $w$ .

Niewątpliwą zaletą powyższych schematów jest to, że pozwalają one na analizowanie funktorów modalnych za pomocą tak prostych środków logicznych, jak kwantyfikacja pierwszego rzędu wiążąca zmienne przebiegające po światach możliwych.

Co więcej, powyższe schematy pozwalają również na analizowanie tych twierdzeń modalnych języka potocznego, które są niewyraźne w

---

<sup>1</sup>W rozprawie wyrażenie „twierdzenie” będzie używane na oznaczenie zdań, które zostały poddane semantycznej interpretacji. Same zdania utożsamiane będą z poprawnie zbudowanymi ciągami symboli.

ramach aparatury pojęciowej logiki modalnej S5. Braku środków pojęciowych, o którym tutaj mowa, nie można w prosty sposób uzupełnić poprzez wprowadzenie aksjomatów dla funktorów pierwotnie niewystępujących w systemie logiki modalnej S5. Notorycznie powracającym przykładem pojęcia modalnego nieposiadającego swojego odpowiednika wśród funktorów modalnych przyjmowanych w ramach logiki modalnej S5, jest pojęcie *aktualności* w sensie, w jakim występuje np. w twierdzeniu „Aktualnie, deficyt budżetowy przekroczył próg 50 % PKB, choć mógł nie przekraczać tego progu”. Chcąc poddać analizie twierdzenia tego typu, można uzupełnić aksjomaty logiki S5 o aksjomat aktualności:

(A) Aktualnie,  $\alpha$  wtw  $\alpha$  w @.

Przywołane wyżej twierdzenie byłoby podstawieniem schematu „Aktualnie,  $\alpha$  i możliwe, że nieprawda, że  $\alpha$ ”.

Problem jednak w tym, że wciąż można wskazać twierdzenia modalne języka potocznego, które nie poddają się analizie nawet z punktu widzenia logiki S5 uzupełnionej o aksjomat aktualności. Przykładowo, w twierdzeniu „Mój samochód mógłby mieć ten sam kolor, który ma aktualnie twój samochód ” funktor aktualności przyjmuje argument nazwowy, a nie zdaniowy, jak wymaga tego aksjomat (A)<sup>2</sup>.

Wraz z pojawieniem się aparatury światów możliwych dostarczenie analizy dla tego twierdzenia nie nastroczało już żadnych trudności. Zostało ono sprowadzone do następującej postaci: „Istnieje taki świat możliwy, że mój samochód ma w tym świecie możliwym taki sam kolor, jaki twój ma w świecie aktualnym”. Nie dość, że terminologia światów możliwych pozwala na analizę funktorów modalnych przy użyciu prostych środków logicznych, to jeszcze systematyzuje ona zachowanie znacznie większej liczby pojęć modalnych języka potocznego niż pozwalała na to którakolwiek z dostępnych logik modalnych,

---

<sup>2</sup>W sprawie szerszego omówienia przykładu zob. §5.1.

włącznie z najpopularniejszą wśród filozofów logiką modalną S5. Na marginesie za każdym razem gdy będę mówił o logice modalnej w tej pracy, będę miał na myśli logikę modalną S5.

Na tym nie kończą się korzyści, jakie mogą płynąć z możliwościowatych<sup>3</sup> analiz. Wraz z rozwojem filozoficznych koncepcji światów możliwych podejmowano również próby wykorzystania ich do redukcji szeregu innych bytów, zwłaszcza tych, dla których sformułowanie zadowalających kryteriów identyczności nastroczało trudności. Za pomocą światów możliwych starano się redukować takie byty jak własności, relacje czy sądy w sensie logicznym. Wszystkie sprowadzano do teoriomnogościowych konstrukcji, których elementami były same światy możliwe lub zamieszkujące je byty. I tak własności redukowano do zbiorów indywiduów, relacje  $n$ -argumentowe — do zbiorów  $n$ -tek uporządkowanych, a sądy w sensie logicznym — do zbiorów światów możliwych. Tym samym — jak utrzymywano — byty o intensjonalnych kryteriach identyczności były zastępowane bytami o ekstensjonalnych kryteriach identyczności. Rosnące wśród filozofów zainteresowanie światami możliwymi nie powinno więc dziwić.

Sukces możliwościowatych analiz jest jednak okupiony pewnymi kosztami. Aby zachować wszelkie dobrodziejstwa płynące z tych analiz, *prima facie* nieustannie należy odnosić się do wielości bytów konkretnych, wśród których znajdują się i takie, których istnienie budzi intuicyjny sprzeciw. Załóżmy, że uznajemy twierdzenie modalne „Mogłyby istnieć jednorożce” za prawdziwe. Akceptując jego możliwościowatą parafrazę, *prima facie* jesteśmy zmuszeni do uznania prawdziwości twierdzenia, zgodnie z którym istnieje świat możliwy, który zamieszkują jednorożce. Co z kolei wydaje się pociągać za sobą konieczność zaakceptowania istnienia takich bytów jak jednorożce.

---

<sup>3</sup>Przez możliwościowate analizy rozumiem analizy, w których występują odniesienia do światów możliwych i bytów istniejących w tych światach oraz które sformułowane są w języku, w którym spełnione są postulaty semantyki relacyjnej Kripkego.



Podobne przykłady można mnożyć, a granice naszej fantazji wyznacza jedynie spójność logiczna. W konsekwencji, jeśli akceptujemy możliwościowe parafrazy twierdzeń modalnych, *prima facie* jesteśmy zmuszeni do uznania bytów, o których nawet nie przeszło nam przez myśl, iż mogłyby istnieć.

Uświadomiwszy sobie takie konsekwencje akceptacji twierdzeń możliwościowych, wielu filozofów znalazło się w dość niekomfortowej sytuacji. Z jednej strony, ciężko im zarzucić możliwościowe analizy, tym bardziej że wydają się one wyjaśniać funkcjonowanie pojęć modalnych, tak obficie wykorzystywanych w filozoficznych rozważaniach. Z drugiej strony, możliwość korzystania z tych analiz jest okupiona kosztami, na które niewielu z nich chce się godzić.

Lekiem często zalecanym na takie rozterki mają być tzw. *umiarkowane stanowiska realistyczne*, redukujące istnienie światów możliwych do aktualnie istniejących bytów. W ramach jednej z nich światy możliwe redukuje się do maksymalnych stanów rzeczy (tj. takich stanów rzeczy  $s$ , że dla dowolnego stanu rzeczy  $s_2$  albo koniecznie, jeśli  $s$  zachodzi, to  $s_2$  zachodzi, albo koniecznie, jeśli  $s$  zachodzi, to  $s_2$  nie zachodzi i dla każdego  $s_3$  i  $s_4$ , jeśli jest tak, że koniecznie, jeśli  $s$  zachodzi, to  $s_3$  i  $s_4$  zachodzą, to możliwe, że  $s_3$  i  $s_4$  zachodzą), których elementami konstytutywnymi są stany rzeczy, na które z kolei składają się własności i zbiory własności. Ostatnie wykorzystuje się do redukcji możliwych indywiduów. Próby takie — jak będę argumentował niżej (§4.4.) — są jednak niezadowalające, gdyż prowadzą do konieczności uznania intuicyjnie prawdziwych twierdzeń modalnych, jak np. „Mogłyby istnieć wyłącznie dwa krążące wokół siebie elektrony o identycznych własnościach nierelacyjnych”, za fałszywe.

W ramach innej redukcyjnej propozycji światy możliwe są utożsamiane z maksymalnymi zbiorami zdań (tj. takimi zbiorami zdań, że dla dowolnego zdania atomowego albo to zdanie, albo jego negacja jest elementem tego zbioru), a możliwe indywidua redukowane są

do nazw własnych. Redukcja taka pozwala na oddanie sprawiedliwości intuicji, że przywołane wyżej twierdzenie modalne o istnieniu dwóch jakościowo nieodróżnialnych elektronów jest prawdziwe. Wystarczy bowiem, by istniał maksymalny zbiór zdań taki, że występują w nim zdania, w których te same predykaty przyjmują za argumenty odmienne nazwy, z którymi identyfikuje się nasze dwa jakościowo nieodróżnialne elektrony.

Trudnością tej koncepcji jest jednak to, że wymaga ona zaciągnięcia zobowiązań ontologicznych co do istnienia idealnego języka, który z pewnością nie może być tworem ludzkiej działalności, choćby dlatego, że w jego słowniku musi znajdować się nieskończenie wiele nazw. W tym nieskończenie wiele nazw, do których redukuje się możliwe indywidua. W każdym razie powinno tak być, jeśli tylko jesteśmy skłonni do zaakceptowania prawdziwości twierdzenia modalnego „Mogłoby istnieć wyłącznie nieskończenie wiele indywiduów, z których żadne nie istnieje aktualnie”. Twierdzenie takie będzie bowiem analizowane jako głoszące, że istnieje maksymalny zbiór zdań, którego elementami jest między innymi nieskończenie wiele zdań o postaci „istnieje takie  $x = a$ ”, w których występują wyłącznie nazwy nieoznaczające żadnego z aktualnie istniejących indywiduów.

W niełatwej sytuacji jest zatem ktoś, komu nie służy żadne z powyższych remediów, a kto wciąż nie może pogodzić się z istnieniem światów możliwych i możliwych indywiduów. Na szczęście z pomocą przychodzą mu filozofowie, którzy składają odmienne propozycje uniknięcia konieczności przyjmowania istnienia kontrowersyjnych bytów.

W słynnej książce *Nazywanie a konieczność* Saul Kripke napisał:

Możliwy świat to nie odległy kraj, który przemierzamy lub oglądamy przez teleskop. [...] Możliwy świat jest *dany przez warunki opisowe, które z nim wiążemy*. Co mamy na myśli, gdy mówimy: „W jakimś innym możliwym świecie nie miałbym dziś tego wykładu”? Wyobrażamy sobie po

prostu sytuację, w której nie postanowiłbym mieć tego wykładu lub postanowiłbym mieć go jakiegoś innego dnia. Nie wyobrażamy sobie oczywiście całej prawdy czy fałszu, lecz jedynie rzeczy istotne do wygłaszania przeze mnie tego wykładu [...]. Nie możemy naprawdę wyobrazić sobie tego inaczej niż częściowo; to jest zatem „świat możliwy”. [...] „Światy możliwe” *ustanawia* się, nie zaś *odkrywa* za pomocą potężnych teleskopów (Kripke 1980/2001, s. 63–64).

Zgodnie z proponowanym przez Kripkego ujęciem światy możliwe redukuje się do bytów istniejących na mocy tego, że konstruujemy opisy sytuacji, które mogłyby mieć miejsce. Podkreśla się przy tym, że opisy te nie są wyczerpujące, tj. nie jest prawdą, że dla dowolnego twierdzenia  $\alpha$  albo to twierdzenie, albo jego negacja są prawdziwe w danym świecie możliwym. Redukcja proponowana przez Kripkego odbiega od przywołanych wyżej propozycji przede wszystkim pod tym względem, że nie stara się zachować obiektywnego istnienia światów możliwych i przez uzależnienie ich istnienia od naszych sposobów myślenia jest bliższa fikcjonalizmowi modalnemu, któremu poświęcona jest ta rozprawa. Niemniej jednak zachodzi ważka różnica pomiędzy propozycją Kripkego a propozycją fikcjonalistów modalnych. Ostatni starają się bowiem zachować standardowe pojęcie świata możliwego, jakie przyjmuje się w ramach semantyki relacyjnej Kripkego, tj. pojęcie świata możliwego, zgodnie z którym jest to obiekt zupełny w tym sensie, że dla dowolnego twierdzenia  $\alpha$  albo to twierdzenie, albo jego negacja jest prawdziwa w tym świecie. Z kolei w cytowanej propozycji Kripkego rezygnuje się z takiego ujęcia światów możliwych na rzecz tego, co niekiedy nazywa się *małymi światami* bądź po prostu możliwościami. Nie oznacza to, iż nie można zbudować semantyki dla języka, w którym użytkownicy posługują się pojęciem małych

światów<sup>4</sup>. Choć samo zagadnienie tego, czy powinniśmy pojmować światy możliwe zgodnie z tradycyjnym ujęciem, czyli jako obiekty dookreślone pod każdym względem, czy też zgodnie z intuicjami, że myśląc o możliwościach mamy jedynie do czynienia z częściowo dookreślonymi obiektami, jest bardzo interesujące, to nie będę się nim zajmował w tej rozprawie. Zamiast tego na potrzeby prowadzonych niżej rozważań założę, że nasz targany rozterkami filozof ma powody, by chcieć zachować standardową semantykę relacyjną Kripkego, która zakłada, że światy możliwe są dookreślone pod każdym względem.

Co w takim razie może zrobić ktoś, kto nie godząc się z istnieniem światów możliwych, jednocześnie ceni analizy możliwościów? Niektórzy filozofowie udzielają następującej odpowiedzi: „Może poszukiwać takiej interpretacji języka możliwościowego, która pozwoliłaby na zachowanie jego użyteczności, nie wymagając jednocześnie przyjmowania ani istnienia światów możliwych, ani żadnych innych bytów, które miałyby pełnić rolę ich substytutów”.

Zasadniczym celem takiego podejścia jest poszukiwanie takiej interpretacji warunków prawdziwości twierdzeń możliwościowych, która spełniałaby trzy dezyderaty:

- (a) twierdzenia tego dyskursu można uznać za twierdzenia w  *pewnym sensie*  prawdziwe, dzięki czemu mogą pełnić one rolę przesłanek w rozumowaniach;
- (b) wbrew realizmowi modalnemu<sup>5</sup> twierdzenia te nie pociągają za sobą istnienia światów możliwych i indywiduów istniejących w tych światach;
- (c) wbrew umiarkowanemu realizmowi modalnemu twierdzenia te nie pociągają istnienia substytutów światów możliwych, a ich

---

<sup>4</sup>W celu zaznajomienia się z różnicami zachodzącymi pomiędzy standardową semantyką relacyjną Kripkego a semantyką dla dyskursu małych światów możliwych zob. (Humberstone 1981).

<sup>5</sup>Więcej o modalnym realizmie będę pisał w §3.1.1.

treść powinna być brana dosłownie, tj. wyrażenia referencjalne występujące w tych twierdzeniach mają na celu odnoszenie się do światów możliwych i możliwych indywiduów, aczkolwiek nie osiągają one tego celu.

Jednym ze stanowisk, które stawiają sobie za cel spełnienie tych warunków, jest fikcjonalizm modalny. Stanowisko to występuje w dwóch wersjach. W ramach wersji *metafikcjonalistycznej* twierdzenie możliwościświatowe  $\alpha^*$  jest eliptycznie wyrażonym twierdzeniem o postaci „Zgodnie z opowieścią o istnieniu wielości światów  $p^*$ ”, które dotyczy treści pewnej opowieści i jej relacji do treści pewnego twierdzenia, a nie światów możliwych i możliwych indywiduów. Z kolei w ramach wersji *przedmiotowej* twierdzenie możliwościświatowe  $\alpha^*$  jest twierdzeniem pierwszego rzędu, tyle że akceptowanym w kontekście, w którym użytkownicy tych twierdzeń wyobrażają sobie, że istnieją światy spełniające te twierdzenia i akceptują je jako prawdziwe na gruncie tych wyobrażeń.

Oba stanowiska opierają się na przekonującej intuicji, zgodnie z którą treść fikcji czy treść naszych wyobrażeń nie muszą opisywać rzeczywistości. Mimo to pewne twierdzenia będą prawdziwe zgodnie z treścią fikcji lub na gruncie wyobrażeń. W pierwszym wypadku będą to twierdzenia, które wynikają z treści fikcji, w drugim – twierdzenia, które spełniają kryteria akceptacji przyjęte w ramach tzw. gry w udawanie. W skrócie gra w udawanie to rządzona pewnymi regułami aktywność wyobrażania sobie, że istnieją rzeczy, które musiałyby istnieć, aby pewna opowieść lub teoria była prawdziwa. Wśród reguł można wyróżnić reguły postulatywne, które określają ogólne ramy tego, co można i co należy sobie wyobrazić na gruncie pewnej teorii. Przykładowo, gdy zaangażujemy się w grę w udawanie, na gruncie której mechanika Newtona jest prawdziwa, wśród reguł postulatywnych pojawi się m.in. nakaz wyobrażania sobie, że istnieje przestrzeń o absolutnym układzie współrzędnych. Obok reguł postulatywnych

wyróżnia się również tzw. reguły generowania. Są to reguły określające dopuszczalne sposoby generowania nowych wyobrażeń w oparciu o realnie zachodzące stany rzeczy. Przykładowo, gdy obserwujemy ruch ciała z punktu A do punktu B, w ramach gry, w której wyobrażamy sobie, że świat jest taki, jakim opisuje go mechanika Newtona, możemy ruch tego ciała opisać np. za pomocą serii czwórek uporządkowanych  $\langle x, y, z, t \rangle$ , które reprezentują położenie obiektu w chwili  $t$  w absolutnym układzie współrzędnych trójwymiarowej przestrzeni.

W kwestii istnienia światów możliwych fikcjonalista również będzie utrzymywał, że nie muszą istnieć światy możliwe, aby pewne twierdzenia były prawdziwe zgodnie z opowieścią o istnieniu światach możliwych lub w ramach gry w udawanie, w której wyobrażamy sobie, że istnieją światy możliwe. Sama opowieść lub wyobrażenia mogą być fałszywe, a mimo to pewne twierdzenia o światach możliwych mogą być prawdziwe na ich gruncie. Szczegółom tego stanowiska będzie poświęcony rozdział 3. Będę tam argumentował, że zwolennicy — uznawanej za najbardziej obiecującą — umiarkowanej wersji fikcjonalizmu modalnego muszą uciec się do wprowadzenia takich własności drugiego rzędu jak własność *bycia możliwe egzemplifikowanym przez wiele indywiduów*. Akceptacja istnienia takiej własności najprawdopodobniej wiąże się z koniecznością zaakceptowania istnienia kontrowersyjnych możliwych indywiduów. W związku z tym projekt fikcjonalistów modalnych w swej umiarkowanej formie, mimo że atrakcyjny na pierwszy rzut oka, najprawdopodobniej ponosi fiasko. Jako że jest to najbardziej obiecująca forma tego stanowiska, pesymistycznie można się zapatrywać na powodzenie całego projektu.

Rozprawa powstawała w ramach realizacji grantu promotorskiego Ministerstwa Nauki i Szkolnictwa Wyższego nr N N101 198838 pt. *Fikcjonalizm modalny w filozofii analitycznej*, którego kierownikiem był dr hab. Tadeusz Szubka. W rozprawie fragmenty rozdziału 4. oraz § 5. bazują na fragmentach moich uprzednio opublikowanych artykułów

— odpowiednio — *Dwa typy modalnego fikcjonalizmu* (Warzoszczak 2009) oraz *Antyrealizm modalny: instrumentalizm versus fikcjonalizm* (Warzoszczak 2011) — lub stanowią ich kopię.

Struktura samej rozprawy przedstawia się następująco. Rozdział 1. ma charakter wprowadzający. Jego głównym zadaniem jest próba objaśnienia pojęcia dyskursu niezobowiązującego ontologicznie.

W rozdziale 2. przybliżę zasadnicze idee wspólne stanowiskom fikcjonalistycznym w ogólności oraz wyróżnię dwie wersje stanowisk fikcjonalistycznych: metafikcjonalizm i fikcjonalizm przedmiotowy (§2.2.), które zestawię ze sobą (§2.3.), przychylając się do argumentacji, zgodnie z którą druga z nich lepiej odpowiada praktyce językowej osób, które akceptują pewne twierdzenia, nie chcąc jednocześnie przyjmować związanych z nimi zobowiązań ontologicznych.

W rozdziale 3. przejdę do prezentacji modalnego metafikcjonalizmu oraz przedmiotowego fikcjonalizmu modalnego, by następnie poddać te stanowiska krytycznej analizie. W tym rozdziale bliżej przyjrzemy się również pojęciu świata możliwego w rozumieniu fikcjonalisty modalnego.

Ostatni rozdział 4. zostanie poświęcony nieredukcjonistycznym stanowiskom alternatywnym względem fikcjonalizmu modalnego, mianowicie instrumentalizmowi oraz modalizmowi. Będę tam argumentował na rzecz tezy, że w zestawieniu z każdym z tych stanowisk można wskazać przekonujące racje przemawiające za fikcjonalizmem modalnym.

Nim przejdę dalej, chcę poczynić jeszcze jedną uwagę dotyczącą ogólnego założenia, które przyświecało mi podczas pisania tej rozprawy, a które brzmi: Podstawowe pojęcia z zakresu ontologii i filozofii języka posiadają swoje intuicyjne interpretacje, nawet jeśli poszczególni filozofowie proponują ich odmienne definicje. W związku z tym założeniem pozwolę sobie na wprowadzenie takich pojęć jak dyskurs, warunki prawdziwości czy zobowiązania ontologiczne w ich intuicyj-

nym rozumieniu, zanim podam ich definicje. Jestem przekonany, że nie ma niczego karygodnego w takiej praktyce. Przeciwnie, uważam, że sam fakt, iż potrafimy klasyfikować definicje tych pojęć jako mniej lub bardziej adekwatne, opiera się właśnie na tym, że posiadamy intuicyjne rozumienie tych pojęć, w oparciu o które formułujemy pewne oczekiwania wobec adekwatnych definicji.



# Rozdział I

## Dyskurs niezobowiązujący ontologicznie

Większość akceptowanych przez nas twierdzeń ma na celu opisywanie sposobu, na jaki mają się rzeczy, o których staramy się mówić. W większości przypadków przyjmuje się założenie, że przedmiot, którego to twierdzenie dotyczy, *realnie* istnieje oraz posiada własności przypisywane mu przez to twierdzenie. Gdy ktoś jest przekonany, że twierdzenie, że Szczecin leży na północny-zachód od Warszawy, jest prawdziwe, zwykle zakłada, że zarówno Szczecin, jak i Warszawa *realnie* istnieją oraz pozostają do siebie we wspomnianej relacji. Podobnie, szczerzej akceptacji twierdzenia „Istnieją kruki” zwykle towarzyszy założenie, że dziedzina kwantyfikacji jest niepusta, a niektóre spośród znajdujących się w niej przedmiotów są krukami.

Choć przekonanie o realnym istnieniu przedmiotów, o których mówimy, zazwyczaj idzie w parze z akceptacją twierdzeń pierwszego rzędu, istnieją przypadki, które wymykają się tej ogólnej tendencji. Niepodobna przypuszczać, by wygłaszaniu takich twierdzeń jak „Sherlock Holmes był wyśmienitym detektywem”, „Jednorożec przypomina koniowate” itp. towarzyszyło przekonanie, że istnieją przedmioty, o których mowa w tych twierdzeniach. Co ciekawe, zachowanie rezerwy co

do istnienia tych przedmiotów nie przeszkadza użytkownikom języka w uznawaniu za prawdziwe niektórych z takich twierdzeń, a innych — za fałszywe. Panuje dość powszechne przekonanie, że twierdzenie, że Sherlock Holmes był wyśmienitym detektywem, jest prawdziwe, podczas gdy twierdzenie, że Sherlock Holmes był kamerdynerem Watsona, jest fałszywe. Podobnie twierdzenie, że jednorożec jest podobny do koniowatych, jest powszechnie uznawane za prawdziwe, podczas gdy twierdzenie, iż jednorożec ma cechy słoniowatych, jest traktowane jako fałszywe. Mamy zatem interesujące przypadki dyskursów, co do których panują intuicje, że dziedziny przedmiotowe tych dyskursów są puste, a mimo to użytkownicy tych dyskursów wyróżniają prawdziwe i fałszywe twierdzenia wyrażalne w ramach tych dyskursów.

Ostatni fenomen jest szeroko rozpoznawany w literaturze filozoficznej, w której pojawiło wiele prób jego wyjaśnienia. Począwszy od zanegowania intuicji, że dziedziny przedmiotowe tych dyskursów są puste, po zachowanie tej intuicji i próbę przyporządkowania twierdzeniom takich dyskursów niestandardowych warunków prawdziwości. Ich niestandardowość ma się przejawiać w tym, że mają one umożliwić mówienie o prawdziwości bądź fałszywości twierdzeń bez przyjmowania zobowiązań ontologicznych co do istnienia bytów, których twierdzenia te zdają się dotyczyć. Pierwsze z wymienionych podejść traktuje dyskurs o przedmiotach fikcyjnych jako zobowiązujący ontologicznie, drugie — jako niezobowiązujący ontologicznie.

Fikcjonalizm modalny wpisuje się w ostatni z tych nurtów. Dlatego dalsza dyskusja o nim wymaga eksplikacji wprowadzonego wyżej rozróżnienia na dwa typy dyskursów. Właśnie jemu zostanie poświęcony ten rozdział.

Niech *dyskursem* będzie język lub fragment języka o lepiej lub gorzej zdefiniowanych regułach formowania poprawnie zbudowanych ciągów wyrażeń, regułach ustalających znaczenia stałych logicznych i stałych pozalogicznych oraz regułach wyznaczających warunki praw-

dziwości twierdzeń sformułowanych w tym fragmencie języka.

Gdy mówi się o tym, że dyskurs  $D$  jest *zobowiązujący ontologicznie*, zwykle wyraża się ideę, że prawdziwość twierdzeń tego dyskursu pociąga za sobą zobowiązania ontologiczne co do istnienia pewnej klasy bytów  $K$ .

Dwa kluczowe pytania narzucają się niemal automatycznie: na czym polega implikowanie zobowiązań ontologicznych przez prawdziwość twierdzeń dyskursu  $D$  co do istnienia bytów klasy  $K$ ? Na podstawie jakich kryteriów powinniśmy ustalać, które byty należą do klasy  $K$ ?

Zacznijmy od pierwszego z postawionych pytań. Zwykle gdy mówi się, że prawdziwość twierdzenia  $p$  pociąga za sobą zobowiązania ontologiczne co do istnienia bytów klasy  $K$ , ma się na myśli to, że  $p$  nie mogłoby być prawdziwe, gdyby nie istniały byty należące do klasy  $K$ . Zależność prawdziwości twierdzenia od tego, co istnieje i na jaki sposób istnieje, próbuje się charakteryzować na rozmaite sposoby. Cztery z nich zasługują na szczególną uwagę: *teoria uprawdziwiaczy*, koncepcja *globalnej superweniencji* prawdziwości na bycie, koncepcja *lokalnej superweniencji* prawdziwości na bycie oraz koncepcja *ufundowania* prawdziwości na bycie.

## 1 Teoria uprawdziwiaczy

Zgodnie z teorią uprawdziwiaczy

dla każdego prawdziwego twierdzenia<sup>6</sup>  $p$ , istnieje pewien byt, na mocy istnienia którego twierdzenie  $p$  jest prawdziwe (por. Armstrong 2004, s. 5, 7).

---

<sup>6</sup>Armstrong mówi o sądach pojmowanych jako byty abstrakcyjne, ja preferuję mówienie o twierdzeniach rozumianych jako zinterpretowane zdania z powodów, o których szerzej będę mówił w §4.4. Przyjęcie, iż nośnikami prawdziwości są twierdzenia, sprawia również, że (TM) wymaga przeformułowania do postaci (TM\*).

Byt ten jest *uprawdziwiaczem* dla twierdzenia  $p$ . Co jednak rozumie się przez *prawdziwość twierdzenia na mocy istnienia pewnego bytu*? Zdaniem Davida Armstronga jest to następujący konieczny związek zachodzący pomiędzy prawdziwością twierdzenia a jego uprawdziwiaczem: niech  $U$  będzie uprawdziwiaczem dla twierdzenia  $p$ , wówczas, z konieczności, jeśli  $U$  istnieje, to  $p$  jest prawdziwe (Armstrong 2004, s. 6). Daje to nam następujące sformułowanie *zasady uprawdziwiania*:

(TM) dla każdego prawdziwego twierdzenia  $p$ , istnieje takie  $U$ , że z konieczności, jeśli  $U$  istnieje, to  $p$  jest prawdziwe.

Samo uznanie, że relacja uprawdziwiania stanowi konieczny związek pomiędzy prawdziwością twierdzenia  $p$  a istnieniem uprawdziwacza  $U$ , nie stanowi jeszcze jej filozoficznie interesującej charakterystyki. Nie wyklucza bowiem utożsamienia uprawdziwacza  $U$  z całością bytu. Przy takim utożsamieniu prawdą będzie, że dla każdego prawdziwego twierdzenia  $p$ , z konieczności, jeśli całość bytu (taka, jaką ona aktualnie jest) istnieje, to  $p$  jest prawdziwe. Problem jednak w tym, że taka teoria kompletnie nie przystaje do naszych intuicji co do tego, w jaki sposób prawdziwość określonego twierdzenia zależy od bytu; intuicji, do których wyjaśnienia została powołana. Wydaje się bowiem, że prawdziwość twierdzenia, że psy szczekają, przede wszystkim zależy od pewnego *fragmentu* rzeczywistości, mianowicie od psów oraz od tego, czy szczekają. W związku z tym, jeśli teoria uprawdziwacza ma wyjaśniać nasze intuicje odnośnie do tego, na czym polega prawdziwość danego twierdzenia, kandydatami na uprawdziwacze powinny być byty, które nie obejmują wszystkiego, co istnieje.

Ponadto, jeśli teoria uprawdziwaczy ma być pomocna przy ustalaniu zobowiązań ontologicznych implikowanych przez twierdzenia pewnego dyskursu, uprawdziwacze nie powinny być identyfikowane z całością bytu. W przeciwnym razie nie sposób nadać sensu niektórym powszechnym praktykom w filozofii. Przykładowo, odrzucanie koncepcji filozoficznej z powodu związanych z nią zobowiązań ontologicznych

byłoby równoznaczne z odrzuceniem jej z powodu pociągania za sobą zobowiązań co do istnienia... całości bytu. Rzadko jednak słyszy się o tym, by ktoś twierdził, iż powodem odrzucenia pewnego stanowiska jest to, że pociąga ono za sobą zobowiązania co do istnienia całości bytu. Raczej podkreśla się, że pociąga ono za sobą zobowiązania co do istnienia takich bytów, których nie chcielibyśmy widzieć wśród składników rzeczywistości (por. Armstrong 2004, s. 18).

Jak w takim razie powinna się przedstawiać filozoficznie interesująca charakterystyka uprawdziwiaczy? Uprawdziwiaczami powinny być byty, które nie tylko pozostają w koniecznym związku z uprawdziwianymi twierdzeniami, ale również są — dodaje Armstrong — *minimalne*, gdzie

- (MU)  $U$  jest minimalnym uprawdziwiaczem dla twierdzenia  $p$  wtw, gdy  $U$  jest uprawdziwiaczem dla  $p$  i nie istnieje takie  $U^*$ , że  $U^*$  jest w sposób właściwy zawarte w  $U$  i  $U^*$  jest uprawdziwiaczem dla twierdzenia  $p$  (Armstrong 2004, s. 20).

Warto tutaj odnotować trzy rzeczy. Po pierwsze, naturalnie nasuwającą się interpretacją relacji zawierania się jednego uprawdziwiacza w drugim jest teoriomogociowa inkluzja. Wówczas natychmiast natrafiamy na kwestię tego, czym są elementy zbiorów, pomiędzy którymi zachodzi relacja zawierania się. Zwolennicy teorii uprawdziwiaczy odpowiadają, że chodzi tutaj o fakty. Nie chcąc się angażować w spór o naturę faktów, zauważę jedynie, że z dwóch najpopularniejszych koncepcji faktów teoretyk uprawdziwiaczy może wybrać tylko jedną. Fakty najczęściej traktuje się jako pary uporządkowane, których elementami jest  $n$ -argumentowa własność oraz  $n$ -tka uporządkowana przedmiotów *egzemplifikujących* tę własność. Alternatywnie bywają traktowane jako prawdziwe sądy w sensie logicznym (zob. np. Frege 1918/1977, s. 125). Wybór drugiej opcji postawiłby zwolennika teorii uprawdziwiaczy przed karkołomnym zadaniem wyjaśniania związku pomiędzy prawdziwością a bytem w kategoriach posiadania przez

pewien sąd w sensie logicznym własności bycia prawdziwym. Przy założeniu, że twierdzenia wyrażają sądy w sensie logicznym, takie „wyjaśnienie” stanowiłoby raczej trywialną konstatację niż wartościową eksplikację relacji pomiędzy prawdziwością a bytem. Nie mówiłoby bowiem niczego ponad to, że z konieczności, jeśli sąd  $p$  jest prawdziwy, to twierdzenie wyrażające ten sąd jest prawdziwe.

Po drugie, interpretacja relacji zawierania się pomiędzy uprawdziwiaczami w kategoriach teoriomnogościowej inkluzji nie może obowiązywać w całej rozciągłości. Wymagałaby ona bowiem przyjęcia założenia, że dowolny uprawdziwiacz jest zbiorem faktów. Tymczasem nietrudno o przykład dyskursu, w którym występują twierdzenia, dla których uprawdziwiaczami nie mogą być zbiory faktów. Załóżmy, że akceptując teorię uprawdziwiaczy, chcemy zidentyfikować zobowiązania ontologiczne teorii mnogości. Twierdzenia uniwersalne dotyczące wszystkich zbiorów powinny być uprawdziwiane m.in. przez wszystkie fakty, których elementami konstytutywnymi są poszczególne zbiory i własności przypisywane im przez dane twierdzenie uniwersalne. Gdyby jednak uprawdziwiacze były zbiorami, zbiór faktów uprawdziwiających twierdzenia uniwersalne teorii mnogości miałby cechy wspólne z antynomiorodnymi zbiorami. Podobnie jak zbiór wszystkich zbiorów czy klasa wszystkich klas niebędących swoimi własnymi elementami zbiór faktów uprawdziwiających uniwersalne twierdzenia teorii mnogości musiałby zawierać sam siebie. Niech  $r$  będzie twierdzeniem uniwersalnym prawdziwym o wszystkich zbiorach i niech  $U'$  będzie zbiorem faktów uprawdziwiających to twierdzenie. Z założenia elementami  $U'$  mają być wyłącznie fakty, od których zachodzenia prawdziwość  $r$  jest zależna. W takim razie, w zbiorze tym powinien być również zawarty fakt, dla którego elementem konstytutywnym jest sam zbiór  $U'$ . W końcu,  $U'$  również jest zbiorem i twierdzenie  $r$  również dotyczy tego zbioru. Zbiory tworzą jednak hierarchię, w której żaden z nich nie może być swoim własnym elementem. Zatem, jeśli teoria uprawdzi-

wiaczy ma pozwalać na identyfikowanie zobowiązań ontologicznych teorii mnogości, uprawdziwiacze nie powinny być zbiorami.

Podobnie nie powinno się traktować faktów w całej rozciągłości jako  $n$ -tek uporządkowanych. Załóżmy bowiem, że jesteśmy skłonni zaakceptować zobowiązania ontologiczne implikowane przez twierdzenie, że klasa właściwa nie jest elementem żadnej klasy. Zgodnie z teorią uprawdziwiaczy powinien zachodzić fakt uprawdziwiający to twierdzenie. Fakt ten nie może jednak być  $n$ -tką uporządkowaną, której elementami są: dopełnienie własności bycia elementem jakiegokolwiek klasy i klasa właściwa, ponieważ klasa właściwa nie może być elementem jakiegokolwiek zbioru. Wydaje się zatem, że chcąc sformułować teorię faktów, którą zwolennik teorii uprawdziwiaczy mógłby wykorzystać w identyfikacji zobowiązań ontologicznych dyskursu poświęconego klasom właściwym, trzeba je utożsamiać ze złożonymi bytami, których konstytucja nie jest analizowana za pomocą narzędzi teorii mnogości.

Ktoś mógłby zapytać, z jakiego typu złożonymi bytami powinny w ogólności być utożsamiane fakty? Nie jest to jednak pytanie, które chciałbym podejmować w tej rozprawie. Zamiast tego, odnotowawszy ograniczenie standardowej koncepcji faktów, przyjmę, że w przypadku dyskursu światów możliwych dopuszczalna jest zarówno teoriomnościowa interpretacja relacji zawierania się, zachodząca pomiędzy uprawdziwiaczami, jak i traktowanie faktów jako  $n$ -tek uporządkowanych<sup>7</sup>.

---

<sup>7</sup>Na marginesie warto odnotować, że założenie takie posiada filozoficznie sporne konsekwencje, stoi ono bowiem w sprzeczności z tzw. *nieograniczoną zasadą rekombinacji*. Zasada ta głosi, że (NZR) dla dowolnych obiektów z dowolnych światów istnieje świat zawierający dowolną liczbę duplikatów tych obiektów. W następujący sposób można pokazać, iż (NZR) wyklucza istnienie zbioru wszystkich możliwych indywiduów: założmy bowiem, że istnieje zbiór  $Z$  wszystkich możliwych indywiduów. Niech liczbą kardynalną zbioru  $Z$  będzie  $C$ . Na mocy aksjomatu zbioru potęgowego istnieje zbiór potęgowy zbioru  $Z$  o mocy kardynalnej równej  $2^C$ . Wówczas na mocy (NZR) powinien istnieć świat możliwy, w którego dziedzinie znajduje się  $2^C$  indywiduów. W związku z tym moc kardynalna zbioru wszystkich możliwych indywiduów powinna wynosić co najmniej  $2^C$ . Z założenia jednak zbiór wszystkich możliwych indywiduów miał mieć moc kardynalną równą  $C$ .

Po trzecie, nie każde prawdziwe twierdzenie będzie miało minimalny uprawdziwiacz. Greg Restall (zob. Armstrong 2004, s. 21–22) podaje przykład twierdzenia  $q$ , że istnieje nieskończenie wiele elektronów, które nie posiada minimalnego uprawdziwacza. Zgodnie ze standardowym ujęciem nieskończoności o zbiorze przedmiotów powiemy, że jest nieskończony, gdy jest on równoliczny ze swoim podzbiorem. Skoro istnieją podzbiory nieskończonego zbioru elektronów, które są z nim równoliczne, to każdy z tych podzbiorów posiada nieskończenie wiele elementów będących elektronami. W związku z tym, każdy z tych podzbiorów nieskończonego zbioru elektronów również jest uprawdziwaczem dla twierdzenia  $q$ . Z uwagi jednak na nieskończoną liczbę elementów w każdym z tych podzbiorów każdy z nich jest równoliczny ze swoim podzbiorem właściwym itd. W konsekwencji, o żadnym proponowanym uprawdziwaczu dla twierdzenia  $q$  nie można powiedzieć, że nie istnieje zbiór faktów, który jest w nim w sposób właściwy zawarty i jest uprawdziwaczem dla tego twierdzenia. Zatem, twierdzenie  $q$  nie ma minimalnego uprawdziwacza.

---

Żaden zbiór nie może jednak być większy od samego siebie. Zatem, jeśli nieograniczona zasada rekombinacji obowiązuje, nie istnieje zbiór wszystkich możliwych indywiduów (zob. Nolan 2002, rozdz. VII).

Jeśli interpretować relację zawierania się zachodzącą pomiędzy uprawdziwaczami w kategoriach teoriomnogościowej inkluzji, konieczne jest przyjęcie założenia, że dowolny uprawdziwiacz dla dowolnego prawdziwego twierdzenia możliwościowego stanowi zbiór faktów. Weźmy dowolne twierdzenie możliwościowe  $s$  prawdziwe o wszystkich możliwych indywiduach. Uprawdziwaczem dla  $s$  będzie zbiór faktów  $Z$  co najmniej zawierający wszystkie fakty, których elementami konstytutywnymi są poszczególne możliwe indywidua oraz własności przypisywane im przez twierdzenie  $s$ . Na mocy aksjomatu sumy powinien istnieć zbiór  $Z^*$ , którego elementami są byty będące elementami zbiorów (tj. faktów utożsamionych z  $n$ -tkami uporządkowanymi) należących do zbioru  $Z$ . Podzbiorem zbioru  $Z^*$  będzie zbiór wszystkich singletonów, które jako swoje elementy zawierają możliwe indywidua. Na mocy aksjomatu sumy istnieje zbiór  $Z'$ , którego elementami są elementy zbiorów należących do  $Z^*$ . Zbiór  $Z'$  będzie zatem zbiorem wszystkich możliwych indywiduów. W związku z tym akceptacja teoriomnogościowej interpretacji relacji zawierania się pomiędzy uprawdziwaczami dla twierdzeń dyskursu możliwościowego wymaga odrzucenia nieograniczonej zasady rekombinacji.



Gdy już przyjmiemy, że niektóre twierdzenia mają minimalne uprawdziwiacze, kolejnym narzucającym się pytaniem jest: „W jaki sposób ustala się, który zbiór faktów jest minimalnym uprawdziwiaczem dla danego twierdzenia?”. Wydaje się, iż naturalna procedura typowania bytów pełniących rolę uprawdziwiaczy powinna bazować na znaczeniach uprawdziwianych twierdzeń. Odwołanie się do znaczeń uprawdziwianych twierdzeń daje również naturalne wyjaśnienie koniecznego związku pomiędzy istnieniem pewnego bytu a prawdziwością pewnego twierdzenia — związek pomiędzy istnieniem zbioru faktów  $U$  a prawdziwością twierdzenia  $p$  jest konieczny, ponieważ znaczenia wyrażeń składających się na twierdzenie  $p$  są takie, że *odnoszą się* do bytów stanowiących elementy konstytutywne faktów należących do  $U$ , a twierdzenie głosi, iż pozostają one do siebie w takiej relacji, w jakiej pozostają do siebie w faktach należących do  $U$ .

Jeśli konieczny związek zachodzący pomiędzy znaczeniami twierdzeń a ich uprawdziwiaczami będzie interpretowany w absolutystycznym sensie, tj. jako związek, który zachodzi z konieczności niezależnie od jakichkolwiek założeń dodatkowych, to uznanie (TM) za prawdziwe wymusza uznanie znaczeń za byty abstrakcyjne istniejące z konieczności. Jedynie pod tym warunkiem możemy uznać, że z konieczności, jeśli istnieje  $U$ , to  $p$  jest prawdziwe. W przeciwnym bowiem razie, tj. gdyby twierdzenia nie istniały z konieczności, możliwa jest sytuacja, w której istniałoby  $U$ , ale nie istniałoby  $p$  i w związku z tym nie byłoby ani prawdziwe, ani fałszywe. Zatem, na proponowaną przez Armstronga charakterystykę relacji uprawdziwiania możemy przystać jedynie wówczas, gdy nośniki wartości logicznych są bytami istniejącymi z konieczności (por. Merricks 2007, s. 6–11).

Poniżej będę twierdził, że fikcjonalista modalny nie powinien przyjmować tego założenia<sup>8</sup>. W związku z tym optymalne byłoby takie sformułowanie rozróżnienia na dyskurs zobowiązujący ontologicznie i

---

<sup>8</sup>Zob. §4.4.

dyskurs niezobowiązujący ontologicznie, które nie wymagałoby przyjmowania koniecznie istniejących nośników wartości logicznych. Szczęśliwie, nietrudno doprowadzić (TM) do postaci, w której będzie wolne od tego założenia. Wystarczy zrelatywizować konieczność do istnienia twierdzenia  $p$ :

(TM\*) dla każdego prawdziwego twierdzenia  $p$ , istnieje takie  $U$ , że z konieczności, jeśli  $U$  istnieje i  $p$  istnieje, to  $p$  jest prawdziwe (por. Merricks 2007, s. 7).

Zgodnie z powyższymi sugestiami, konieczny związek pomiędzy  $p$  a  $U$  zachodzi ze względu na to, że znaczenia wyrażeń występujących w twierdzeniu  $p$  odnoszą się do bytów składających się na fakty należące do  $U$  i byty te pozostają do siebie w takiej relacji, jak głosi  $p$ .

Warto tutaj podkreślić, że odnoszenie się wyrażenia do bytu jest tutaj rozumiane niestandardowo w tym sensie, że nie jest zarezerwowane wyłącznie dla relacji zachodzącej pomiędzy nazwami a indywiduami, ale również obejmuje relacje zachodzącą pomiędzy predykatami a własnościami oraz zmiennymi związanymi a ciągami przedmiotów.

O znaczeniu wyrażenia  $v$  powiemy, że odnosi się do bytu  $x$  wtw, gdy odwołanie się do bytu  $x$  jest niezbędne do wyjaśnienia roli, jaką wartość semantyczna wyrażenia  $v$  pełni przy wyznaczaniu warunków prawdziwości twierdzenia, w którym wyrażenie  $v$  występuje. I tak, jeśli uważamy, że twierdzenie „ $a$  jest  $F$ ” jest prawdziwe wtw, gdy byt  $a$  przyporządkowany stałej „ $a$ ” jest  $F$ , to możemy powiedzieć „ $a$ ” odnosi się do  $a$ . Jeśli dodatkowo — jak czyni to np. Armstrong (2004, s. 23–24) — uważamy, że wartościami semantycznymi predykatów są własności bądź relacje, powiemy wówczas, że jednoargumentowy predykat „ $F$ ” odnosi się do własności  $F$ , dwuargumentowy predykat „ $R$ ” odnosi się do binarnej relacji  $R$  itd.

W przypadku twierdzeń ogólnych, tj. twierdzeń egzystencjalnych (wyrażonych przy użyciu zmiennych związanych kwantyfikatorem

szczegółowym) bądź twierdzeń uniwersalnych (wyrażonych przy użyciu zmiennych związanych kwantyfikatorem ogólnym), w ogólności powiemy, że zmienne związane odnoszą się do  $m$ -tych elementów  $n$ -elementowych ciągów przedmiotów, do których należą byty znajdujące się w dziedzinie kwantyfikacji. Rzecz jasna, liczba elementów należących do tych ciągów będzie wyznaczana przez  $n$ -arność predykatów, które przyjmują jako swoje argumenty zmienne związane danym kwantyfikatorem. Podobnie,  $m$ -ty element, do którego będzie się odnosiła zmienna związana, będzie wyznaczany przez kolejność tej zmiennej wśród zmiennych związanych kwantyfikatorem.

Podsumowując dotychczasowe ustalenia, powiemy, że uprawdziwizacjami są zbiory faktów, których elementami konstytutywnymi są przedmioty lub ciągi przedmiotów, do których odnoszą się — odpowiednio — nazwy własne lub zmienne związane kwantyfikatorami oraz własności lub relacje, do których odnoszą się predykaty (por. Armstrong 2004, s. 24–25). Przykładowo, uprawdziwizaczem dla twierdzenia jednostkowego, że  $a$  jest  $F$ , będzie singleton, którego elementem jest fakt, że  $a$  jest  $F$ . Elementami konstytutywnymi ostatniego będzie przedmiot  $a$  i własność  $F$ . Z kolei uprawdziwizaczem dla twierdzenia egzystencjalnego (np. istnieje takie  $x$ , że  $Fx$ ) będzie zbiór tylko tych faktów, których elementami konstytutywnymi są ciągi przedmiotów (do których odnoszą się zmienne związane kwantyfikatorem szczegółowym) spełniające własność (do której odnosi się predykat). Wreszcie, uprawdziwizaczem dla twierdzenia uniwersalnego (np. dla każdego  $x$ ,  $Fx$ ) będzie zbiór wszystkich faktów takich, że dla dowolnego ciągu przedmiotów (znajdującego się pośród ciągów przedmiotów, do których odnoszą zmienne związane kwantyfikatorem ogólnym) istnieje taki fakt, należący do tego zbioru, że ciąg ten spełnia własność (do której odnosi się predykat). W przypadku przygodnie prawdziwych twierdzeń uniwersalnych w zbiorze faktów uprawdziwiających takie twierdzenia powinien również znaleźć się *negatywny fakt uniwersalny*,

że *nie* ma więcej  $n$ -argumentowych ciągów przedmiotów.

Bez negatywnego faktu uniwersalnego powinno być tak, że sam zbiór faktów takich, że określona własność jest spełniona przez wszystkie ciągi przedmiotów, powinien być uprawdziwiaczem dla twierdzenia uniwersalnego. Zatem powinno być tak, że z konieczności, jeśli zbiór tych faktów istnieje i twierdzenie to istnieje, to jest ono prawdziwe. Mógłby jednak istnieć świat, w którym zachodzą jakieś dodatkowe fakty, które sprawiałyby, że twierdzenie to byłoby fałszywe.

Rozważmy przykład. Twierdzenie, że dla każdego  $x$ , nieprawda, że  $x$  jest jednorożcem, jest przygodnie prawdziwym twierdzeniem uniwersalnym. Gdyby zbiór wszystkich aktualnie zachodzących faktów, na które składa się jednoelementowy ciąg przedmiotów i dopełnienie własności bycia jednorożcem, stanowił uprawdziwiacz dla tego twierdzenia, to niezależnie od tego, czy zachodziłyby jakieś dodatkowe fakty, wartość tego twierdzenia nie powinna ulec zmianie. Mógłby jednak istnieć świat, który różni się od świata aktualnego tylko tym, że istnieje w nim jakaś odległa, dodatkowa planeta zamieszkała przez jednorożce. W świecie tym twierdzenie, że dla dowolnego  $x$ , nieprawda, że  $x$  jest jednorożcem, byłoby twierdzeniem fałszywym, gdyż istniałyby takie jednoelementowe ciągi przedmiotów, które nie spełniałyby dopełnienia własności bycia jednorożcem. Zatem zbiór wszystkich faktów, których elementami konstytutywnymi są ciągi, do których elementów odnoszą się zmienne związane, i własności, do których odnoszą się predykaty, nie stanowi uprawdziwacza dla tego twierdzenia.

O jaki fakt powinniśmy uzupełnić zbiór faktów uprawdziwiających nasze twierdzenie uniwersalne? Cechą odróżniającą świat aktualny od świata jednorożców jest to, że w pierwszym w ogóle nie istnieją przedmioty posiadające własność bycia jednorożcem, a wszystkim istniejącym przedmiotom można przypisać dopełnienie tej własności. W związku z tym kandydatem na uzupełnienie zbioru faktów, pełniącego rolę uprawdziwacza dla tego twierdzenia, jest negatywny fakt

uniwersalny, że w świecie nie istnieje więcej przedmiotów.

Niektórzy filozofowie wzdragają się przed przyjmowaniem uniwersalnego faktu negatywnego, który należałby do wszystkich zbiorów faktów uprawdziających przygodnie prawdziwe twierdzenia uniwersalne. Preferują oni poszukiwanie negatywnych faktów, które pozostałyby w bliższym związku z typem przedmiotów, których istnienie neguje się. Analizując podobne zagadnienie, Armstrong pisze:

Dość często zakłada się – nawet wśród zwolenników uprawdziaczy – że cały świat jest wymagany jako uprawdziacz dla [negatywnych – P.W.] prawd [egzystencjalnych – P.W.]. Wydaje się to zbyt pesymistyczne. Minimalnym uprawdziaczem dla [twierdzenia, że – P.W.] nie istnieją arktyczne pingwiny, wydają się być arktyczne zwierzęta [...] wraz ze stanem rzeczy, że stanowią one całość takich zwierząt. Każde z arktycznych zwierząt jest – jak powiedziałby Platon – na mocy swojej natury *odmienne* od pingwina. Zatem, ten ogólny stan rzeczy wydaje się wystarczającym uprawdziaczem dla tego negatywnego [twierdzenia – P.W.] egzystencjalnego (Armstrong 2004, s. 75 – 76).

Podobnie, być może minimalnym uprawdziaczem dla prawdziwości twierdzenia negującego istnienie jednorożców powinny być fakty dotyczące zwierząt wraz z faktem, że stanowią one całość zwierząt? Problem jednak w tym, że egzemplifikowana przez wszystkie arktyczne zwierzęta własność *bycia całością arktycznych zwierząt*, podobnie jak egzemplifikowana przez wszystkie zwierzęta własność *bycia całością zwierząt*, podlega dalszej analizie. Załóżmy, że elementy pewnej klasy bytów wspólnie egzemplifikują pierwszą z tych własności. Wówczas na własność *bycia całością arktycznych zwierząt* będzie składała się m.in. własność *bycia zwierzęciem*, własność *zasiedlania Arktyki*, które przysługują każdemu bytowi należącemu do tej klasy. Obie

własności jednak nie wystarczą do tego, by elementy tej klasy były wszystkimi arktycznymi zwierzętami, musimy bowiem dodać, że w świecie nie ma innych arktycznych zwierząt poza należącymi do tej klasy. Wspólnie elementy tej klasy powinny zatem posiadać własność *bycia takimi, że nie ma innych arktycznych zwierząt w tym świecie*. Ostatnie wydaje się równoważne z tym, że świat posiada własność *bycia takim, że nie ma więcej arktycznych zwierząt*. Ostatnia własność znowu wydaje się podlegać dalszej analizie. Kiedy bowiem światu przysługuje własność *bycia takim, że nie ma więcej arktycznych zwierząt*? Naturalnie nasuwająca się odpowiedź brzmi: „Wtedy, gdy elementy pewnej klasy bytów posiadają własność *bycia arktycznym zwierzęciem*, wszystkie byty nienależące do tej klasy nie są arktycznymi zwierzętami i *nie ma więcej przedmiotów*”. W związku z tym, analiza minimalnych uprawdziwiaczy dla negatywnych twierdzeń egzystencjalnych wymaga odwołania się do *negatywnego faktu uniwersalnego* (por. Merricks 2007, s. 58 – 59). Analogiczne rozważania stosują się do własności *bycia całością zwierząt*.

Co ciekawe, w przypadku koniecznie prawdziwych twierdzeń uniwersalnych nie powstaje potrzeba włączania negatywnego faktu uniwersalnego do zbioru faktów stanowiących uprawdziwiacz dla takiego twierdzenia. Żadne nowe fakty nie zmieniają bowiem wartości logicznej koniecznie prawdziwego twierdzenia. Weźmy twierdzenie, że każdy kawaler jest nieżonatym mężczyzną. Niezależnie od tego, czy będą istniały jakieś byty rodzajowo odmienne od bytów aktualnie istniejących, tak długo jak będą one bytami logicznie możliwymi, tak długo fakty, których są elementami konstytutywnymi, będą należały do zbioru faktów stanowiącego uprawdziwiacz dla tego twierdzenia.

Dysponując tak zarysowaną teorią uprawdziwiaczy, możemy pokusić się o następujące sformułowanie kryterium zobowiązań ontologicznych:

twierdzenia teorii  $T$  wyrażonej w ramach dyskursu  $D$  pociągają

zobowiązania co do istnienia bytów klasy  $K$  wtw, gdy byty należące do klasy  $K$  stanowią elementy konstytutywne faktów należących do zbioru faktów uprawdziającego twierdzenia egzystencjalne teorii  $T$  wyrażonej w ramach dyskursu  $D$ .

Powody ograniczenia do twierdzeń egzystencjalnych klasy twierdzeń istotnych dla identyfikacji zobowiązań ontologicznych dyskursu  $D$  są dość oczywiste.

Po pierwsze, w odróżnieniu od twierdzeń uniwersalnych twierdzenia egzystencjalne pozwalają na rozpoznanie większej liczby klas bytów, których istnienie przyjmuje się w ramach dyskursu  $D$ . Opierając się na twierdzeniach uniwersalnych o postaci „dla każdego  $x$ ,  $Gx$ ”, można ustalić co najwyżej niektóre z klas bytów, których istnienie przyjmuje się w  $D$  — mianowicie tylko te klasy, których elementami są wszystkie byty znajdujące się w dziedzinie kwantyfikacji. Najczęściej jednak spory ontologiczne dotyczą tych klas bytów, do których nie należą wszystkie byty znajdujące się w dziedzinie kwantyfikacji. I tak światy możliwe czy przedmioty wewnętrznie sprzeczne stanowią przedmioty sporów, choć nigdy nie utrzymuje się, że do klasy światów możliwych czy przedmiotów wewnętrznie sprzecznych należą wszystkie byty.

Po drugie, twierdzenia jednostkowe z niepustymi nazwami własnymi o tyle tylko odgrywają istotną rolę w identyfikowaniu zobowiązań ontologicznych, o ile pociągają za sobą twierdzenia egzystencjalne. Jest tak dlatego, że debaty ontologiczne z zasady mają dotyczyć ogółu bytów należących do danego rodzaju, a nie jednostkowych przypadków należących do danego rodzaju. Gdy stawia się pytanie o istnienie liczb, przedmiotów materialnych itd., stawia się pytanie o to, czy istnieją jakiekolwiek przedmioty będące liczbami, przedmiotami materialnymi itd., a nie o to, czy ta oto liczba istnieje, czy ten oto przedmiot materialny istnieje.

Położenie nacisku na rolę twierdzeń egzystencjalnych przybliża sformułowane wyżej kryterium do popularnego Quine’owskiego kryte-

rium zobowiązań ontologicznych. Zgodnie z nim twierdzenie  $p$  pociąga za sobą zobowiązania ontologiczne co do istnienia bytów klasy  $K$  wtw, gdy, jeśli twierdzenie to jest wyrażone w rachunku predykatów pierwszego rzędu, a wszelkie wystąpienia nazw własnych są zastąpione zmiennymi pierwszego rzędu związanymi kwantyfikatorami szczegółowymi, to, jeśli twierdzenie  $p$  jest prawdziwe, to byty klasy  $K$  znajdują się pośród wartości zmiennych związanych (por. Quine 1969, Quine 1948/2000).

Jak zwraca uwagę Armstrong (2004, s. 23), zachodzi istotna różnica pomiędzy oboma kryteriami. Jeśli elementami konstytutywnymi faktów będą również własności i relacje pojmowane jako byty intensjonalne, to sformułowane wyżej kryterium pozwoli na zidentyfikowanie zobowiązań ontologicznych co do istnienia tego typu bytów. Tymczasem dobrze znanym ograniczeniem Quine'owskiego kryterium zobowiązań ontologicznych jest to, że pozwala ono na zidentyfikowanie zobowiązań wyłącznie co do bytów, które są wartościami zmiennych związanych pierwszego rzędu. Wśród takich bytów z pewnością nie znajdziemy własności pojmowanych jako byty intensjonalne.

Rzecz jasna, teoria uprawdziwiaczy i sformułowane na jej podstawie kryterium zobowiązań ontologicznych same z siebie nie pociągają za sobą zobowiązań co do istnienia bytów intensjonalnych. To, czy pojawiają się one wśród zobowiązań dyskursu  $D$ , w znacznej mierze zależy od tego, w jaki sposób będziemy interpretowali elementy konstytutywne faktów. Można poszukiwać takiej koncepcji własności i relacji, w ramach której będą one bytami doskonale nadającymi się na wartości zmiennych związanych pierwszego rzędu. W takiej sytuacji zasadnicza różnica pomiędzy oboma kryteriami będzie opierała się na tym, że pierwsze współgra z koncepcją faktów, zgodnie z którą ich elementami konstytutywnymi są byty intensjonalne, podczas gdy wersja Quine'owska pozostaje neutralna w tym względzie.



Zdaniem samego Armstronga teoria uprawdziwiaczy stanowi doskonale narzędzie identyfikowania zobowiązań ontologicznych również dla tych, którzy nie akceptują faktów:

niektórzy mogą argumentować, że [...] świat jest światem rzeczy, a nie faktów. [...] Jest to sporne stanowisko, ale przyjmując je na potrzeby argumentu, wciąż może ono być pogodzone z doktryną uprawdziwiaczy. Niech świat będzie światem rzeczy. Fundamentalne prawdy (te, które mają minimalne uprawdziwacze) będą wówczas posiadały formę ' $x$  istnieje', a  $x$ -y, czymkolwiek mogą one być, będą uprawdziwaczami dla tych prawd (Armstrong 2004, s. 24).

Być może w ontologii, zgodnie z którą świat jest światem rzeczy, znajdują się nie będące zbiorami faktów uprawdziwacze dla twierdzeń, w których głosi się, że pewien byt istnieje. Problem jednak w tym, że trudno sobie wyobrazić ontologię, której tezy formułowane są wyłącznie przy pomocy takich twierdzeń. Jeśli teoria ontologiczna ma być choć odrobinę wiarygodna, powinna również zawierać twierdzenia, w których orzeka się pewne własności o rzeczach istniejących w danym świecie. Ale gdy w zbiorze naszych tez ontologicznych pojawiają się twierdzenia tego typu, następnym krokiem powinno być wyjaśnienie, co jest ich uprawdziwaczami.

Gdyby przyjąć sugerowane przez Armstronga ograniczenie, prawdziwość takich twierdzeń powinna być uznana, jedynie gdy twierdzenia, w których orzeka się pewne własności o rzeczach, da się zredukować do twierdzeń, w których stwierdza się tylko istnienie pewnych rzeczy. Załóżmy, że chcemy znaleźć uprawdziwacze dla niezwykle ubogiej teorii metafizycznej, która czerpie swoje inspiracje z nihilizmu mereologicznego. Niech teoria ta ogranicza się do dwóch twierdzeń: (1) każda rzecz jest identyczna z samą sobą i (2) każda rzecz jest bytem prostym (nie posiadającym części). Załóżmy dodatkowo, że —

o dziwo — nasza uboga teoria trafia w sedno i nie myli się ani co do identyczności rzeczy, ani co do ich mereologicznej natury. Idąc za sugestią Armstronga, powinniśmy uznać, że wszystkie rzeczy są uprawdziwiaczem zarówno w przypadku tezy (1), jak i w przypadku tezy (2). Z pewnością wartości logiczne obu twierdzeń są od siebie niezależne. Czy powinny być również od siebie metafizycznie niezależne? Jeśli teoria uprawdziwiaczy ma powiedzieć nam coś filozoficznie interesującego o relacji prawdziwości twierdzeń do bytu, powinna również być w stanie wyjaśnić, co po stronie rzeczy odpowiada niezależnej prawdziwości tez (1) i (2). Przyjmując jednak, że uprawdziwiaczami są rzeczy, a nie fakty (czy zbiory faktów), niezależność prawdziwości tych twierdzeń możemy wyjaśniać jedynie w kategoriach odmiennej liczby rzeczy, które uprawdziwiają twierdzenia (1) i (2). Z założenia jednak, nasze twierdzenia — jak przystało na twierdzenia metafizyczne — są koniecznie prawdziwe. Dlatego też nie sposób wyjaśnić niezależność prawdziwości tych twierdzeń w kategoriach odmiennej liczby rzeczy, które je uprawdziwiają. Zatem tak ograniczona teoria uprawdziwiaczy nie jest w stanie dostarczyć metafizycznego wyjaśnienia dla niezależności prawdziwości obu twierdzeń.

Nie twierdzę, że przekreśla to hipotetyczną propozycję Armstronga. Chcę jedynie wskazać na niedosyt, który pozostaje po obietnicach dostarczenia filozoficznie interesującego wyjaśnienia związku pomiędzy prawdą a bytem. Jak to w chwilach niedosytu, zwykle łakomym okiem patrzymy na teorie, które są w stanie powiedzieć nam więcej o naturze tej relacji. Dlatego jeśli istnieje jakaś teoria uprawdziwiaczy, która z jednej strony odwołuje się wyłącznie do bytów, które mogą być wartościami zmiennych związanych pierwszego rzędu, a z drugiej strony mogą powiedzieć, co po stronie rzeczy odpowiada niezależnej prawdziwości tez (1) i (2), to powinna ona być preferowana względem hipotetycznej propozycji Armstronga.

Dającym nadzieję sposobem na uniknięcie ograniczeń hipotetycznej propozycji Armstronga jest uznanie, że:

- uprawdziwiacze są zbiorami faktów,
- fakty są parami uporządkowanymi takimi, że ich elementami jest  $n$ -argumentowa własność i  $n$ -elementowy ciąg przedmiotów oraz  $n$ -argumentowa własność jest zrealizowana przez  $n$ -elementowy ciąg przedmiotów,
- $n$ -argumentowe własności są identyfikowane ze zbiorami  $n$ -tek uporządkowanych,
- realizacja  $n$ -argumentowej własności przez  $n$ -tkę uporządkowaną przedmiotów jest interpretowana w kategoriach należenia tej  $n$ -tki uporządkowanej do zbioru  $n$ -tek uporządkowanych, z którym zidentyfikowano daną własność.

Tak zinterpretowane własności legitymują się ekstensjonalnymi warunkami identyczności i jako takie nadają się do tego, by być wartościami zmiennych pierwszego rzędu. W odróżnieniu od hipotetycznej propozycji Armstronga, korzystając z tak zarysowanej koncepcji, można wskazać, co po stronie bytu odpowiada za niezależność prawdziwości tez (1) i (2). Otóż uprawdziwiają je zupełnie inne zbiory faktów. W pierwszym wypadku mamy do czynienia ze zbiorem, którego elementami są fakty będące parami uporządkowanymi, których elementami są zbiór wszystkich par uporządkowanych o postaci  $\langle x, x_i \rangle$  oraz para uporządkowana o postaci  $\langle x, x_i \rangle$ . W drugim wypadku elementami tego zbioru będą fakty, których elementami będą zbiór wszystkich rzeczy prostych i jednoelementowy ciąg o postaci  $\langle x_i \rangle$ .

Choć w wielu wypadkach da się wskazać odmienne uprawdziwiacze, w wielu innych koncepcja ta cierpi na podobne ograniczenia co hipotetyczna propozycja Armstronga. W kontekście poszukiwania uprawdziwaczy dla tez metafizycznych pierwszych trudności nastrocza już

wskazanie, co po stronie bytu odpowiada za niezależną prawdziwość twierdzeń, w których tym samym obiektom przypisuje się odmienne, ale z konieczności koekstensywne własności monadyczne. Załóżmy, że z konieczności wszystkie przedmioty proste są przedmiotami materialnymi. Rozważmy teraz twierdzenie (2) i twierdzenie (3), głoszące, że każda rzecz jest materialna. W obu wypadkach elementami faktów będą te same zbiory własności oraz te same jednoelementowe ciągi przedmiotów. Podobnie jak w hipotetycznej propozycji Armstronga możemy mieć trudność ze wskazaniem, co po stronie rzeczy odpowiada niezależnej prawdziwości twierdzeń (2) i (3). Rzecz jasna, teoria uprawdziwaczy, opierająca się na teoriomnogościowej koncepcji własności, ma tę zaletę, że zawsze pozostawia pole do dostarczenia przekonującej analizy własności bycia prostym lub własności bycia materialnym, która pozwoli na wskazanie odmiennych uprawdziwaczy dla każdego z tych twierdzeń.

Jednak nawet gdyby udało się przedstawić takie analizy tych własności, z punktu widzenia fikcjonalizmu modalnego akceptacja takiej koncepcji uprawdziwaczy ma pewną istotną wadę. Chodzi mianowicie o to, że poszukiwania uprawdziwaczy dla twierdzeń, w których rzeczom przypisuje się odmienne, przygodnie koekstensywne własności, szybko doprowadzą nas do ontologii możliwych indywiduów, którą fikcjonalista odrzuca. W ramach powyższej propozycji o dwóch własnościach monadycznych powiemy, że są różne, gdy różne są zbiory rzeczy, z którymi je zidentyfikowano. Problem jednak w tym, że przygodnie koekstensywne własności są one identyfikowane z dokładnie tymi samymi zbiorami rzeczy aktualnie istniejących. Jednym z najbardziej znanych sposobów odróżnienia zbiorów rzeczy, z którymi identyfikujemy odmienne, ale przygodnie koekstensywne własności, jest wprowadzenie w Lewisowskim stylu do naszej ontologii konkretnych możliwych indywiduów. Jeśli bowiem własności są przygodnie koekstensywne, z pewnością znajdą się możliwe rzeczy, które posiadają

jedną z nich, nie posiadając drugiej. Mając tak bogatą ontologię, z łatwością znajdziemy odmienne uprawdziwiacze dla twierdzeń przypisyjących przygodnie koekstensywne własności. Takie rozwiązanie jest jednak nieakceptowalne dla fikcjonlisty modalnego, który za wszelką cenę stara się odrzucić istnienie możliwych indywiduów.

Z punktu widzenia fikcjonalizmu modalnego trudno znaleźć satysfakcjonującą postać teorii uprawdziwiaczy, która spełniałaby Quine'owskie standardy mówienia wyłącznie o bytach, które mogą być wartościami zmiennych związanych pierwszego rzędu. W ramach fikcjonalizmu modalnego koncepcja taka musiałaby wymagać przyjęcia nieakceptowalnej ontologii możliwych indywiduów albo nie dostarczałaby filozoficznie interesującego wyjaśnienia relacji pomiędzy prawdą a bytem.

Można jednak postawić nieco bardziej fundamentalne pytanie: co poza Quine'owską ortodoksją może przemawiać na niekorzyść teorii uprawdziwiaczy? W podręcznym arsenale jej przeciwników znajdują się przede wszystkim obiekcje podnoszone przeciw zastosowaniom tej teorii do wyjaśnienia prawdziwości negatywnych twierdzeń egzystencjalnych, takich jak twierdzenie, że nie istnieją jednorożce czy że nie istnieje Sherlock Holmes itd. Z każdego nietrywialnie spełnionego przypadku zarówno (TM), jak i (TM\*) wynika, że dla dowolnego prawdziwego twierdzenia powinien istnieć jakiś byt, uprawdziwiający to twierdzenie. W przypadku negatywnych twierdzeń egzystencjalnych wymóg istnienia uprawdziwacza wydaje się jednak niewiarygodny. Czy negatywne twierdzenia egzystencjalne nie dlatego są prawdziwe, że nie istnieje nic, co posiadałoby cechy, o których mowa w tych twierdzeniach? W związku tym intuicyjnie prawdziwa wydaje się teza, że nie istnieje nic, co mogłoby uprawdziwiać negatywne twierdzenia egzystencjalne (Merricks 2007, s. 40).

Jak widzieliśmy wyżej, wielu zwolenników teorii uprawdziwiaczy postawiło sobie za zadanie znalezienie uprawdziwaczy dla prawdzi-

wych negatywnych twierdzeń egzystencjalnych. Proponowali więc, by były nimi zbiory faktów, w których znajdziemy dwa typy faktów. Po pierwsze, będą to fakty, których elementem jest dowolny jednoelementowy ciąg przedmiotów spełniający dopełnienie własności, do której odnosi się predykat. Po drugie, będzie to negatywny fakt uniwersalny, że w świecie nie ma więcej przedmiotów. Jeśli uznać, że ostatni jest analizowalny w kategoriach posiadania przez całość bytu własności *bycia takim, że nie ma więcej przedmiotów*, to natrafimy na kilka trudności.

W pierwszym rzędzie czeka nas problem tego, czy własność ta jest pierwotna, czy podlega dalszej analizie. Zdaniem Trentona Merricka (por. 2007, s. 61–62) własność ta przypomina własność *bycia takim, że istnieją elektrony*. Ostatnią własność można przypisać światom, których elementami bądź częściami są elektrony<sup>9</sup>, a wydaje się przysługiwać im na mocy tego, że elektrony są ich elementami bądź częściami. Gdyby miała zachodzić taka analogia, świat powinien posiadać własność *bycia takim, że nie ma niczego więcej*, na mocy tego, że pozostaje w jakiejś relacji do czegoś, co nie jest światem. Problem jednak w tym, że jeśli własność ta faktycznie przysługuje światu, to nie może on pozostawać w relacji do czegoś innego, gdyż poza tym, co jest w świecie, nie istnieje nic innego. Jeśli przyjmiemy, że własność *bycia takim, że nie ma więcej przedmiotów*, jest własnością relacyjną, to — z uwagi na nieistnienie bytu, w relacji do którego miałyby pozostawać całość faktów — nie powinna ona istnieć. W konsekwencji, nie powinno być również negatywnego faktu uniwersalnego. Z uwagi na to, że miał on należeć do zbioru faktów uprawdziwiających

---

<sup>9</sup>W przypadku światów rozumianych jako maksymalne sumy mereologiczne istnienie elektronu w takim świecie jest reprezentowane za pomocą relacji *bycia częścią*, w jakiej elektron ten pozostaje do sumy utożsamionej ze światem. Gdy świat jest natomiast rozumiany jako całość faktów, a ostatnie są traktowane jako pary uporządkowane, wówczas relacja zachodząca pomiędzy elektronem a światem będzie relacją *bycia elementem co najmniej jednego faktu należącego do całości faktów identyfikowanej ze światem*.

negatywne twierdzenia egzystencjalne, twierdzenia te — wbrew teorii uprawdziwiaczy — nie miałyby swoich uprawdziwiaczy.

Być może zatem własność ta jest własnością posiadaną przez świat na mocy jego wewnętrznej konstytucji? Byłaby ona wówczas własnością wewnętrzną podobnej do własności bycia przedmiotem złożonym, przysługującej krzesłu na mocy tego, że posiada ono takie części jak siedzisko, oparcie itd. Posiłkując się przywołaną wyżej argumentacją na rzecz konieczności wprowadzenia negatywnego faktu uniwersalnego, łatwo można wykazać, iż własność bycia takim, że nie ma więcej przedmiotów, nie jest własnością wewnętrzną. Załóżmy bowiem, że świat  $w$  jest sumą mereologiczną, której częściami są wszystkie istniejące przedmioty oraz że każdy z tych przedmiotów stanowi elementy zbioru  $Z$ . Wówczas wystarczającym warunkiem przysługiwania tej własności światu  $w$  powinno być to, że składają się na niego przedmioty należące do zbioru  $Z$ . Mógłby jednak istnieć świat  $w^*$ , w którym poza przedmiotami należącymi do  $Z$  istnieją jeszcze jakieś dodatkowe przedmioty. Wówczas fragmentowi świata  $w^*$ , na który składają się wyłącznie byty należące do zbioru  $Z$ , nie powinna przysługiwać własność bycia takim, że nie ma więcej przedmiotów.  $Z$  założenia bowiem w świecie  $w^*$  jest więcej przedmiotów. Gdyby własność ta była własnością wewnętrzną, powinna ona przysługiwać temu fragmentowi świata  $w^*$ . Zatem, własność ta nie jest własnością wewnętrzną.

Wydaje się, że nie pozostaje nic innego, jak uznać własność bycia takim, że nie ma więcej przedmiotów, za własność pierwotną przysługującą światu. To jednak sprawia, że teoria uprawdziwaczy staje się jeszcze bardziej kontrowersyjna. Nie tylko wymaga przyjęcia, że musi istnieć byt uprawdziwiający negatywne twierdzenia egzystencjalne, ale również wymusza wprowadzenie do naszej ontologii negatywnego faktu uniwersalnego, na który składa się własność wymykająca się standardowym ontologicznym klasyfikacjom własności na relacyjne

i wewnętrzne. Tym samym sposób, na jaki przysługuje ona światu, jest na tyle tajemniczy, że każda teoria, która może poradzić sobie z wyjaśnieniem fenomenu prawdziwości negatywnych twierdzeń egzystencjalnych bez odwoływania się do tego typu własności, powinna być preferowana względem teorii uprawdziwiaczy.

## 2 Superweniencja prawdziwości na bycie

David Lewis w *Truthmaking and Difference-Making* (2001) proponował, aby zależność prawdziwości twierdzeń od tego, co istnieje, definiować w kategoriach globalnej superweniencji prawdziwości na bycie:

(G-TSB) Dla każdego twierdzenia  $p$ , dla każdego świata  $w$  i  $w^*$ , jeśli  $p$  jest prawdziwe w  $w$  i  $p$  nie jest prawdziwe w  $w^*$ , to albo coś istnieje w świecie  $w$ , co nie istnieje w świecie  $w^*$ , albo pewne  $n$ -tki przedmiotów pozostają do siebie w pewnych fundamentalnych relacjach w świecie  $w$  i nie pozostają do siebie w tych relacjach w świecie  $w^*$  (Lewis 2001, s. 612).

Zgodnie z zaproponowaną przez Lewisa koncepcją zależność prawdziwości od bytu ma w całości wyrażać się w tym, że zmiany pod względem tego, co istnieje, lub zmiany pod względem tego, na jaki sposób istnieje, będą skutkowały zmianami w tym, co jest prawdziwe.

Zdaniem Lewisa akceptacja (G-TSB) pozwala na wyjaśnienie prawdziwości negatywnych twierdzeń egzystencjalnych bez odwoływania się do jakiegokolwiek bytu, który miałby istnieć w świecie, w którym twierdzenie to jest prawdziwe. Teza o globalnej superweniencji prawdziwości na bycie wymaga jedynie, by w celu wskazania zależności prawdziwości od bytu wskazać na różnicę pomiędzy światami możliwymi, w których odmienne twierdzenia są prawdziwe. Weźmy twierdzenie  $t$ , że nie istnieją jednorożce oraz negację tego twierdzenia,



czyli  $t$ . W naszym świecie prawdziwe jest twierdzenie  $t$ , ponieważ różni się on od wszystkich światów, w których prawdziwe jest twierdzenie  $t$ . Różnica, o której tutaj mowa, jest różnicą pod względem tego, co istnieje w naszym świecie, a co istnieje w światach, w których prawdziwe jest twierdzenie  $t$ . W innych światach musi istnieć coś, co nie istnieje w naszym świecie, mianowicie jednorożce. Pod względem istnienia jednorożców świat ten różni się od naszego świata, w którym nie istnieje nic, co sprawiałoby, że  $t$  byłoby prawdziwe (Lewis 2001, s. 610).

Zgodnie z Lewisowską koncepcją prawdziwości negatywnych twierdzeń egzystencjalnych twierdzenie takie jest prawdziwe, gdy nie istnieje nic, na czego istnieniu mogłaby superweniować prawdziwość negacji tego twierdzenia. Charles B. Martin (1996, s. 61) stawia takiej koncepcji prawdziwości negatywnych twierdzeń egzystencjalnych zarzut, że kompletnie nie wyjaśnia ona, dlaczego twierdzenia te są prawdziwe. Próbuje się bowiem wyjaśniać prawdziwość negatywnych twierdzeń egzystencjalnych przez odwołanie się do pewnych negatywnych twierdzeń, o których zakłada się, że są prawdziwe. Przykładowo, twierdzenie  $t$  jest prawdziwe, gdy nie istnieje nic, co sprawiałoby, że negacja tego twierdzenia jest prawdziwa. Innymi słowy, twierdzenie to jest prawdziwe, gdy nie istnieją jednorożce. Ostatnie samo jest negatywnym twierdzeniem egzystencjalnym, którego prawdziwość zakłada się przy okazji wyjaśniania prawdziwości negatywnych twierdzeń egzystencjalnych w ogólności. Zatem proponowane przez Lewisa wyjaśnienie prawdziwości negatywnych twierdzeń egzystencjalnych nie jest żadnym wyjaśnieniem, gdyż zakłada się w nim to, co miało być wyjaśnione. Lewis zgadza się z Martinem, że proponowana przez niego koncepcja nie jest wyjaśnieniem prawdziwości negatywnych twierdzeń egzystencjalnych, ale dodaje, że wcale nie lepiej mają się zwoleńnicy teorii uprawdziwiaczy, którzy z kolei w wyjaśnianiu prawdziwości pozytywnych twierdzeń egzystencjalnych odwołują się do pewnych

pozytywnych twierdzeń egzystencjalnych, o których zakładają, że są prawdziwe. Jest tak dlatego, że gdy wyjaśniają prawdziwość pozytywnych twierdzeń egzystencjalnych, stwierdzając, że istnieje pewien byt będący uprawdziwiaczem tego twierdzenia, to ostatnie twierdzenie samo jest pozytywnym twierdzeniem egzystencjalnym, którego prawdziwość się zakłada (Lewis 2001, s. 611–612). Wciąż jednak — w odróżnieniu od (TM) bądź (TM\*) — (G-TSB) nie pociąga konieczności poszukiwania jakiegoś specjalnego bytu, który istnieje w świecie, w którym prawdziwe jest negatywne twierdzenie egzystencjalne, i który pełni rolę uprawdziwacza tego twierdzenia.

Czy na podstawie (G-TSB) można sformułować kryterium zobowiązań ontologicznych? W znanych mi pracach Lewisa nie spotkałem się z próbą sformułowania kryterium zobowiązań ontologicznych, jednak wydaje się w oparciu o jego koncepcję globalnej superweniencji prawdziwości na bycie można pokusić się o sformułowanie następującej propozycji:

twierdzenia teorii  $T$  wyrażonej w ramach dyskursu  $D$  implikują zobowiązania ontologiczne co do istnienia bytów klasy  $K$  wtw, gdy byty klasy  $K$  istnieją w każdym świecie, w którym twierdzenia egzystencjalne kluczowe dla teorii  $T$  wyrażonej w ramach dyskursu  $D$  są prawdziwe, i światy, w których twierdzenia te są fałszywe, różnią się od światów, w których są one prawdziwe, ze względu na istnienie bytów klasy  $K$ .

O twierdzeniu egzystencjalnym  $p$  powiemy, że jest kluczowe dla teorii  $T$ , gdy stwierdza się w jego ramach istnienie bytów postulowanych w ramach teorii  $T$ .

Rzucającą się w oczy wadą tego kryterium zobowiązań ontologicznych jest to, że nie pozwala ono na odróżnienie zobowiązań ontologicznych implikowanych przez prawdy konieczne. Dla dwóch dowolnych koniecznych prawdziwych twierdzeń  $p$  i  $q$ , klasa bytów, które istnieją w każdym świecie, w którym  $p$  jest prawdziwe, jest identyczna z klasą

bytów, które istnieją w każdym świecie, w którym  $q$  jest prawdziwe. Na gruncie sformułowanego wyżej kryterium powinny one pociągać za sobą te same zobowiązania ontologiczne. Wydaje się jednak, że niektóre twierdzenia pretendujące do miana prawd koniecznych implikują odmienne zobowiązania ontologiczne. Przykładowo, dla współczesnego arystotelika będzie prawdą konieczną zarówno to, że istnieją własności, jak i to, że istnieją indywidua egzemplifikujące te własności. Oba twierdzenia pociągają za sobą jednak odmienne zobowiązania ontologiczne, pierwsze dotyczą bowiem własności, drugie — indywiduów. Nasuwającym się sposobem obejścia tej trudności jest uznanie, iż zależność prawdziwości twierdzeń od bytu nie powinna być charakteryzowana w kategoriach różnic pomiędzy *całymi* światami, a raczej w kategoriach zależności ich prawdziwości od *niektórych* spośród bytów istniejących w tych światach. Wówczas można byłoby wskazać byty, od których zależna jest prawdziwość twierdzeń dotyczących własności, oraz odmienne od ostatnich byty, od których zależna jest prawdziwość twierdzeń dotyczących indywiduów. Wprowadzenie takiej modyfikacji byłoby jednak równoznaczne z uznaniem, że zależność prawdziwości od bytu nie może w ogólności być charakteryzowana za pomocą globalnych różnic pomiędzy światami możliwymi. Muszą być uwzględnione również zależności od bytów składających się na te światy. Samo (G-TSB) nie uwzględnia tego typu zależności pomiędzy prawdziwością a bytem, nie jest więc jasne, jak systematyczna charakterystyka tej zależności miałaby się przedstawiać.

Poza koniecznością rozwikłania ostatniej kwestii koncepcja globalnej superweniencji prawdziwości na bycie ma — z punktu widzenia rozważań, które będą prowadzone poniżej — jeszcze jedną zasadniczą wadę. Otóż przy formułowaniu podstawowego rozróżnienia na dyskursy ontologicznie zobowiązujące i ontologicznie niezobowiązujące, na którym będzie się opierała koncepcja fikcjonalizmu modalnego, dobrze byłoby nie odwoływać się do aparatury światów możliwych.

Fikcjonałści modalni są sceptyczni wobec istnienia światów możliwych. Twierdzą również, że dyskurs światów możliwych można interpretować w taki sposób, by można było akceptować twierdzenia tego dyskursu bez przyjmowania implikowanych przez nie zobowiązań ontologicznych. Tylko w takim sensie fikcjonalista może zaakceptować (G-TSB). Nawet jeśli aparatura światów możliwych za sprawą (G-TSB) pozwala w przystępny sposób opisać rozróżnienie na dyskurs zobowiązujący i niezobowiązujący ontologicznie, to dla fikcjonalisty modalnego jest ona jedynie fikcją, za pomocą której uławiamy sobie zrozumienie natury tej zależności. Wypadałoby jednak powiedzieć coś o naturze zależności prawdziwości od bytu bez angażowania się w dyskurs światów możliwych. Powinno się to zrobić choćby po to, by wyjaśnić, dlaczego uważa się, że dyskurs światów możliwych pozwala w przystępny sposób opisać tę zależność. W związku z tym fikcjonałści modalni powinni poszukiwać odmiennej od (G-TSB) charakterystyki zależności prawdziwości od bytu.

Co więcej, jeśli (G-TSB) faktycznie ma wyrażać ideę zależności prawdziwości od bytu, to konieczne jest przyjęcie jednej z realistycznych koncepcji światów możliwych: albo koncepcji Lewisa, albo koncepcji Plantingi<sup>10</sup> (por. Merricks 2007, s. 85–87). Zgodnie z pierwszą z nich światy możliwe są bytami tego samego rodzaju co świat aktualny, w każdym z nich istnieją — podobnie jak w świecie aktualnym — konkretne przedmioty i każdy z tych światów jest równie realny jak świat aktualny. Zgodnie z drugą z nich światy możliwe są bytami abstrakcyjnymi — maksymalnymi niesprzecznymi stanami rzeczy, a istniejące w

---

<sup>10</sup>Tutaj (G-TSB) przyjęłoby następującą postać: dla każdego świata  $w$  i  $w^*$ , jeśli  $p$  jest prawdziwe w  $w$  i fałszywe w  $w^*$ , to istnieje taki stan rzeczy  $s$ , że  $s$  jest implikowane przez  $w$  i  $s$  jest wykluczane przez  $w^*$ . Samo kryterium zobowiązań ontologicznych może przyjąć następującą postać: twierdzenia teorii  $T$  wyrażonej w ramach dyskursu D pociągają za sobą zobowiązania ontologiczne co do istnienia bytów klasy  $K$  wtw, gdy dowolny świat, w którym twierdzenia tej teorii są fałszywe, różni się od świata, w którym twierdzenia tej teorii są prawdziwe, co najmniej tym, że wyklucza wszystkie pozytywne stany rzeczy, które zawierają byty klasy  $K$  jako swoje elementy konstytutywne.

nich indywidua są reprezentowane za pomocą tzw. indywidualnych istot. Ostatnie mogą być charakteryzowane jako zbiory zrelatywizowanych do poszczególnych światów własności, które z konieczności mogą przysługiwać tylko jednemu przedmiotowi.

Założmy, że nie chcąc akceptować żadnej z tych koncepcji, wybieramy jedną z alternatyw. Wśród najlepiej sprzedających się ofert na rynku filozoficznych koncepcji światów możliwych znajdziemy koncepcję utożsamiającą je z maksymalnymi niesprzecznymi sędami w sensie logicznym (gdzie sąd w sensie logicznym  $p$  jest maksymalny, gdy dla dowolnego sądu w sensie logicznym,  $p$  implikuje ten sąd bądź jego negację i sąd  $p$  jest niesprzeczny, gdy wszystkie sądy implikowane przez ten sąd mogą być zarazem prawdziwe) oraz z maksymalnymi niesprzecznymi zbiorami zdań (gdzie zbiór zdań jest maksymalny wtw, gdy dla dowolnego zdania, albo to zdanie, albo jego negacja jest elementem tego zbioru i zbiór zdań jest niesprzeczny wtw, gdy nie istnieje takie zdanie, że ono i jego negacja są elementami tego zbioru). W pierwszej koncepcji (G-TSB) przyjmuje postać:

Dla dowolnego sądu  $r$ , dla dowolnych maksymalnych niesprzecznych sądów  $p$  i  $q$ , jeśli  $r$  jest implikowane przez  $p$  i  $r$  nie jest implikowane przez  $q$ , to istnieje taki sąd  $s$ , że  $s$  jest implikowane przez  $q$  i  $s$  wyklucza się z  $r$ , i negacja  $s$  jest implikowana przez  $p$ .

Wówczas jednak (G-TSB) wyraża relację wynikania zachodzącą pomiędzy sędami; relację, która ma zachodzić m.in. pomiędzy wartościami logicznymi twierdzeń wyrażających te sądy. W związku z tym o prawdziwości sądu mówi jedynie w relacji do prawdziwości innego sądu, a co za tym idzie, relacja ta w ogóle nie wyraża idei zależności prawdziwości od bytu. Ostatnia jest bowiem relacją zachodzącą pomiędzy prawdziwością sądu a bytem, który w ogóle nie jest nośnikiem wartości logicznych.

Podobnie rzecz przedstawia się, gdy zaakceptujemy erzacym lingwistyczny, wówczas bowiem (G-TSB) przyjmuje następującą postać:

Dla dowolnego zdania  $r$ , dla dowolnych maksymalnych niesprzecznych zbiorów zdań  $Z$  i  $Q$ , jeśli  $r$  jest elementem  $Z$  i  $r$  nie jest elementem  $Q$ , to istnieje takie zdanie  $s$ , że  $s$  jest elementem  $Q$  i  $s$  jest negacją  $r$ .

Na gruncie tego stanowiska (G-TSB) dotyczy jedynie relacji pomiędzy zbiorami zdań, nie mówi zaś niczego ani o tym, czy relacje te w jakikolwiek sposób odzwierciedlają relacje pomiędzy bytami, ani o tym, jaka jest natura zależności prawdziwości od bytu.

Argumentacja w powyższym stylu daje się uogólnić na wszystkie koncepcje światów możliwych, w ramach których identyfikuje się je z maksymalnymi niesprzecznymi reprezentacjami, w których zawarte są inne reprezentacje. Jeśli bowiem cokolwiek będzie mówiło się o werydyczności jakiegokolwiek reprezentacji (tj. takiej własności tej reprezentacji, że reprezentowany obiekt posiada wszystkie własności, jakie wyszczególnione są w treści tej reprezentacji), to będzie mowa o jej werydyczności jedynie w relacji do werydyczności jakiejś innej reprezentacji. Zatem (G-TSB) nie może wyrażać idei zależności prawdziwości od bytu, gdy światy utożsamia się z maksymalnymi reprezentacjami.

Wydaje się jednak, że gdy formułuje się kryterium zobowiązań ontologicznych, pożądaną praktyką jest poszukiwanie takiej jego formy, by zachowana została jak najdalej idąca neutralność względem szczegółowych sporów ontologicznych. Pod tym względem kryterium zobowiązań ontologicznych sformułowane na podstawie (G-TSB) nie wypada najlepiej, gdyż wymaga uprzedniego rozstrzygnięcia sporu o naturę światów możliwych na rzecz realistycznych stanowisk. Biorąc pod uwagę, iż światy możliwe w ramach dyskutowanych ujęć nie są bytami pierwotnymi, wybranie którejkolwiek z realistycznych opcji wiązałoby się z koniecznością rozstrzygnięcia szeregu sporów

powiązanych ze sporami o istnienie światów możliwych. W przypadku sformułowania (G-TSB) na podstawie stanowiska Lewisa nasze kryterium zobowiązań ontologicznych wymagałoby odrzucenia nihilizmu mereologicznego (tj. stanowiska, zgodnie z którym nie istnieją przedmioty złożone), zajęcia stanowiska w sporze o naturę indywiduów, stanowiących części światów, zajęcia stanowiska w sporze o naturę aktualności, istnienia indywiduów zamieszkujących wiele światów itd. Nie lepiej rzecz ma się ze sformułowaniem (G-TSB) na podstawie stanowiska Plantingi, wówczas bowiem nasze kryterium zobowiązań ontologicznych wymagałoby zaakceptowania istnienia stanów rzeczy, uniwersaliów itd. Dobrze jednak byłoby operować takim sformułowaniem idei zależności prawdziwości od bytu, która pozwoliłoby na wyrażenie kryterium zobowiązań ontologicznych zachowującego neutralność przynajmniej w przypadku niektórych z wyżej wymienionych sporów ontologicznych.

Jednym ze stanowisk, które wydaje się zachowywać neutralność wobec sporu o naturę światów możliwych, jest stanowisko lokalnej superweniencji prawdziwości na bycie, zgodnie z którym:

(L-TSB) Z konieczności, każde prawdziwe twierdzenie  $p$  jest takie, że, z konieczności, jeśli istnieją wszystkie aktualne byty i posiadają takie własności, jakie aktualnie posiadają, to  $p$  jest prawdziwe (Merricks 2007, s. 72–73).

W odróżnieniu od (G-TSB) w (L-TSB) w definiowaniu zależności prawdziwości od bytu nie porównuje się ze sobą dwóch odmiennych światów możliwych, lecz mówi się o związkach pomiędzy prawdziwością a tym, co aktualnie istnieje. Celem uniknięcia zarzutu postawionego przed chwilą Lewisowi należy przyjąć założenie, że pojęcie konieczności występujące w (L-TSB) jest pojęciem nieanalizowalnym w kategoriach kwantyfikacji ogólnej po światach możliwych. W przeciwnym razie, jeśli w standardowy sposób dokonamy analizy pojęć

modalnych możliwości i konieczności jako — odpowiednio — kwantyfikacji szczegółowej oraz kwantyfikacji ogólnej po światach możliwych, sprowadzimy (L-TSB) do postaci:

Dla każdego świata możliwego  $w$ , dla każdego twierdzenia  $p$ , jeśli dla każdego  $x$  takiego, że  $x$  istnieje w świecie aktualnym,  $x$  istnieje w świecie  $w$  i posiada te same własności, które posiadało w świecie aktualnym, i  $p$  jest prawdziwe w świecie aktualnym, to  $p$  jest prawdziwe w świecie  $w$ .

W tej formie (L-TSB) wyraża relację pomiędzy bytem a prawdziwością jedynie wówczas, gdy przyjmie się realistyczną koncepcję światów możliwych. W przeciwnym razie (L-TSB) wyrażałoby — jak wyżej — relację pomiędzy sądami w sensie logicznym bądź zbiorami zdań i ich elementami.

Jak pokazuje Merricks (2007, s. 73), łatwo znaleźć znaczącą różnicę pomiędzy (L-TSB) i (G-TSB). Załóżmy, że nie istnieje własność *bycia takim, że nie ma więcej przedmiotów*, oraz że możliwe jest, że każdy aktualnie istniejący przedmiot posiada dokładnie te same własności, które posiada aktualnie, i dodatkowo istnieje jednorożec. Jeśli prawdziwość lokalnie superweniuje na bycie, to istnienie aktualnie istniejących przedmiotów posiadających te same własności, jakie aktualnie posiadają, powinno pociągać za sobą prawdziwość tych samych twierdzeń, które prawdziwe są w świecie aktualnym. Jednym z tych twierdzeń jest twierdzenie, że nie istnieją jednorożce. Z uwagi jednak na to, że twierdzenie to byłoby fałszywe, gdyby zachodziła wprowadzona wyżej możliwość, okazuje się, że jego prawdziwość nie superweniowałaby lokalnie na tym, co istnieje. Fałszywość tego twierdzenia nie rodzi jednak najmniejszych problemów, gdy przyjmie się (G-TSB). Opisana wyżej możliwość zachodzi bowiem w świecie, który różni się od świata aktualnego pod względem tego, co w nim istnieje, i jako taki powinien różnić się od niego również pod względem tego, co jest w nim prawdziwe.



Powyższe świadectwo na rzecz odmienności (G-TSB) oraz (L-TSB) ujawnia zarazem piętę achillesową ostatniej koncepcji. Okazuje się bowiem, że prawdziwość negatywnych twierdzeń egzystencjalnych może lokalnie superweniować na tym, co aktualnie istnieje, jedynie wówczas, gdy przyjmie się, że całość tego, co aktualnie istnieje, posiada własność bycia takim, że nie ma więcej przedmiotów. W przeciwnym bowiem razie, zgodnie z (L-TSB), twierdzenie, że nie istnieje jednorożec, będzie prawdziwe w każdym świecie, w którym istnieją wszystkie aktualnie istniejące przedmioty i posiadają te same własności, które aktualnie posiadają, mimo że w niektórych z tych światów będą istniały jednorożce.

Jak starałem się wykazać powyżej, istnienie takiej własności jest jednak wysoce podejrzane. Nie twierdzę, że powinniśmy odrzucić (L-TSB) z powodu tego, że wymaga przyjęcia istnienia tej własności — być może jest tak, że żadna z osiągalnych koncepcji tego, na czym polega zależność prawdziwości od bytu, nie może uniknąć odwoływania się do jej istnienia.

Wielu filozofów żywi przekonanie, że ani (TM) bądź (TM\*), ani teza o superweniencji prawdziwości na bycie, czy to globalnej, czy lokalnej, w ogóle nie wyrażają idei zależności prawdziwości od tego, co istnieje. W pierwszym wypadku, jeśli implikacja występująca zarówno w (TM), jak i (TM\*) jest po prostu ścisłą implikacją, to dowolny byt  $U$  będzie pełnił rolę uprawdziwacza dla dowolnego koniecznie prawdziwego twierdzenia  $p$ , o ile tylko istnieje twierdzenie  $p$ , ponieważ dla prawdziwości tej implikacji wystarczającą jest prawdziwość jej następnika. Ścisła implikacja będzie prawdziwa niezależnie od tego, jakie twierdzenie egzystencjalne wystąpi w jej poprzedniku. Warunki prawdziwości tej implikacji dopuszczają zatem sytuację, w której w jej poprzedniku stwierdza się istnienie bytu  $U$ , który jest nieistotny dla prawdziwości twierdzenia  $p$ . Filozofowie, zdając sobie sprawę z tej trudności, proponowali, aby implikacji występującej w (TM) i

(TM\*) nie traktować jako ścisłej implikacji, ale jako np. implikację relewantną bądź implikację parakonsystentną (zob. Lewis 2001, s. 604). W tej rozprawie kwestię tę pozostawiam już na boku.

W drugim wypadku sama superweniencja nie mówi niczego o tym, który z członów tej relacji ma być bardziej fundamentalny w porządku wyjaśniania. Idea zależności prawdziwości twierdzeń od tego, co istnieje, jest ideą wyjaśniania prawdziwości twierdzeń poprzez odwołanie się do istnienia pewnych bytów, które pozostają w odpowiednich relacjach do znaczeń tych twierdzeń. Z pewnością idea zależności prawdziwości twierdzeń od tego, co istnieje, nie jest ideą wyjaśniania istnienia pewnych bytów poprzez odwołanie się do prawdziwości pewnych twierdzeń. W związku z tym zależność prawdziwości od bytu powinna być relacją niezwrotną. Relacja superweniencji jest natomiast relacją symetryczną. Weźmy (G-TSB). Jeśli dwa światy różnią się pod względem tego, co jest w nich prawdziwe, to różnią się pod względem tego, co istnieje, bądź tego, jakie własności posiadają przedmioty istniejące w tych światach, ale i odwrotnie: jeśli dwa światy będą różniły się pod względem tego, co w nich istnieje, bądź tego, jakie własności posiadają przedmioty istniejące w tych światach, to będą różniły się pod względem tego, co jest w nich prawdziwe. Podobnie rzecz przedstawia się z (L-TSB): z konieczności jest tak, że dla każdej prawdy  $p$ , jeśli istnieją wszystkie aktualne byty i posiadają takie własności, jakie aktualnie posiadają, to  $p$  jest prawdziwe, ale jeśli relacja ta faktycznie jest relacją superweniencji, powinno również być odwrotnie, z konieczności, dla każdej prawdy  $p$ , jeśli  $p$  jest prawdziwe, to istnieją wszystkie aktualne byty i posiadają takie własności, jakie aktualnie posiadają (por. Armstrong 2004, s. 8).

### 3 Ufundowanie prawdziwości na bycie

Zgodnie z propozycją, na której chcę zakończyć powyższy przegląd stanowisk, zależność prawdziwości od bytu może być oddana za pomocą relacji *fundowania* zachodzącej pomiędzy prawdziwością pewnego twierdzenia a istnieniem pewnego bytu (zob. Rosen 2010, s. 109–117 i 114, przypis 3). Gdy mówimy o relacji fundowania zachodzącej pomiędzy czymś a czymś, zwykle wyrażamy to za pomocą takich fraz, jak „coś zachodzi na mocy czegoś”, „coś jest całkowicie zależne od czegoś” itd. Metafizycy zwykle utrzymują, że sama relacja fundowania jest relacją pierwotną. Gdy chce się rzucić nieco więcej światła na dowolną pierwotną relację, dobrze wyjść od przykładów jej zachodzenia, by na ich podstawie zrekonstruować jej własności. Jednym z przykładów zachodzenia relacji fundowania jest właśnie relacja zachodząca pomiędzy prawdziwością pewnego twierdzenia a istnieniem bytu, na mocy istnienia którego twierdzenie to jest prawdziwe. Można również wyliczyć kilka kolejnych przykładów:

- (1) Mówi się, że własności dyspozycyjne są ufundowane na własnościach kategoryalnych, np. własność bycia tłukliwą posiadana przez szklankę jest ufundowana na tym, jaką ma ona strukturę molekularną.
- (2) Niektórzy filozofowie twierdzą, że fakty dotyczące znaczeń wyrażań ufundowane są na faktach dotyczących praktyki posługiwania się tymi wyrażeniami, wyznaczającej ich rolę inferencyjne itd.
- (3) Niekiedy mówi się, że to, że pewien czyn jest zły, jest ufundowane na tym, że czyn ten posiada pewną własność, która sprawia, że jest zły, np. łamanie obietnic jest moralnie niedozwolone, ponieważ uniemożliwia społeczną współpracę.

- (4) Fizykaliści twierdzą, że takie fakty w naszym świecie jak np. to, że ludzie myślą, mają pierwszoosobowe doznania, nie stanowią fundamentalnych właściwości rzeczywistości, lecz są ufundowane na takich faktach, jak pobudzenia takiej a takiej grupy neuronów itd.
- (5) Niekiedy mówi się, że posiadanie przez pewien przedmiot pewnej własności jest ufundowane na tym, że pozostaje on w pewnych relacjach do innych przedmiotów, np. przedmiot posiada własność bycia najwyższą budowlą na mocy tego, że jest budowlą i jest wyższy niż każda inna budowla; przedmiot posiada własność bycia sławnym w pewnej grupie ludzi m.in. na mocy tego, że pozostaje w relacji bycia rozpoznawanym przez ludzi należących do tej grupy itd.
- (6) Metafizycy zwykle stawiają sobie za zadanie nie tyle powiedzenie, co istnieje, ale co istnieje w najbardziej fundamentalnym sensie, tj. istnienie których bytów nie jest ufundowane na istnieniu żadnych innych bytów.

Co powyższe przykłady mówią o relacji fundowania? W pierwszym rzędzie sugerują, że relacja fundowania jest *relacją zachodzącą pomiędzy faktami*, tj. fakt, że istnieją byty klasy  $K$ , jest ufundowany na fakcie istnienia bytów klasy  $K^*$ , albo fakt, że  $A$  jest  $B$ , jest ufundowany na fakcie, że  $G$  jest  $F$ . Przykładowo, fakt, że predykat „jest oponą” stosuje się do opon, jest ufundowany na społecznej praktyce posługiwania się tym wyrażeniem.

Mając w pamięci, że nasze zainteresowanie fundowaniem prawdziwości na bycie wzięło się z poszukiwań uniwersalnego kryterium zobowiązań ontologicznych, które by w miarę możliwości nie przesądzało sporów ontologicznych, pojawiające się wyżej zobowiązanie co do faktów może budzić pewien niepokój, gdyż wydaje się przesądzać spór ontologiczny o istnienie faktów. Nie chcąc jednak wikłać

się w ten spór, zwróćę jedynie uwagę na to, że pojęcie faktów jest na tyle silnie zakorzenione w naszych rozważaniach teoretycznych, że dowolny przeciwnik włączania ich do ontologii powinien być w stanie przeprowadzić ich redukcyjną analizę. Analiza ta powinna być na tyle systematyczna, by możliwe było nadanie takiej treści teoretycznie istotnym twierdzeniom z pozoru dotyczącym faktów, która pozwoliłaby na uniknięcie zobowiązań ontologicznych co do nich. Jeśli zgodzić się, że twierdzenia dotyczące zachodzenia relacji fundowania są teoretycznie istotne, redukcyjna analiza powinna również objąć twierdzenia tego typu. Być może wiara w możliwość dostarczenia takiej analizy jest naiwna, wówczas jednak mielibyśmy silne racje na rzecz akceptacji jakiejś ontologii faktów.

Wracając do rekonstrukcji własności relacji fundowania można dostrzec, iż niektóre z powyższych przykładów pokazują, że jeden fakt może być ufundowany wspólnie przez wiele faktów. Przykładowo, fakt, że osoba  $a$  jest rozpoznawana w społeczności  $S$ , jest ufundowany na takiej kolekcji faktów, że każdy z faktów należących do niej jest faktem, że dany członek tej społeczności rozpoznaje osobę  $a$  i nie istnieje taki fakt, że jakiś członek społeczności  $S$  rozpoznaje osobę  $a$  i fakt ten nie należy do tej kolekcji.

Mając relację fundowania pomiędzy faktem, że  $p$ , oraz kolekcją faktów  $X$  możemy również zdefiniować pojęcie bycia *częściowo ufundowanym* przez pewne fakty. O fakcie, że  $p$ , powiemy, że jest częściowo ufundowany przez fakty  $Y$  wtw, gdy istnieje taka kolekcja faktów  $X$  i  $Y$ , że fakt, że  $p$ , jest ufundowany na kolekcji faktów  $X$  i kolekcja faktów  $Y$  jest zawarta w sposób właściwy w kolekcji faktów  $X$ . Przykładowo, fakt, że osoba  $a$  jest rozpoznawalna w społeczności  $S$ , jest częściowo ufundowany na fakcie, że osoba  $x$  ze społeczności  $S$  rozpoznaje osobę  $a$ .

Drugą rzeczą, którą warto odnotować, jest to, że zwykle mówi się o fundowaniu, gdy próbuje się *wyjaśnić*, na czym polega zachodzenie

pewnego faktu poprzez odwołanie się do zachodzenia pewnego innego faktu (zob. Fine 2001, s. 16). Przykładowo, zachodzenie faktu, że osoba  $a$  jest znana w społeczności  $S$ , wyjaśnia się poprzez zachodzenie szeregu faktów takich, że poszczególni członkowie tej społeczności rozpoznają osobę  $a$ ; zachodzenie faktu, że predykat „jest oponą” stosuje się do opon, wyjaśnia się poprzez zachodzenie takich a takich faktów składających się na praktykę językową danej społeczności itd. Dlatego też, jeśli zgodzić się z metafizykami, że relacja fundowania jest pierwotna, to w następujący sposób można zdefiniować pojęcie wyjaśniania:

zachodzenie faktu, że  $p$ , jest wyjaśniane przez zachodzenie faktu, że  $q$  wtw, gdy fakt, że  $p$ , jest ufundowany na fakcie, że  $q$ .

Wyjaśnianie zdefiniowane w taki sposób wydaje się oddawać intuicję, że zachodzenie pewnych faktów powinno być wyjaśniane przez zachodzenie faktów, które bierze się za fakty bardziej fundamentalne względem wyjaśnianych faktów. Przykładowo, fakt, że fakt, że ruch molekuł wody jest wzmożony, stanowi wyjaśnienie dla faktu, że woda jest ciepła, jedynie wówczas, gdy pierwszy z tych faktów traktuje się jako fakt bardziej fundamentalny od drugiego z nich. Podobnie, fakt, że pewien przedmiot pozostaje w pewnej relacji do innych przedmiotów, stanowi wyjaśnienie faktu, iż posiada on pewną własność relacyjną, jedynie wówczas, gdy pierwszy z tych faktów jest bardziej fundamentalny od drugiego. Zatem relacja fundowania jest relacją zachodzącą pomiędzy faktami, z których jeden jest faktem bardziej fundamentalnym. W związku z powyższym, relacja fundowania powinna być relacją *niezwrotną* — żaden fakt nie może być bardziej fundamentalny od samego siebie.

Jeśli fakt, że  $p$ , jest ufundowany na innym fakcie, że  $q$ , to pierwszy nie może fundować ostatniego. Jest tak dlatego, że relacja fundowania zachodzi pomiędzy dwoma faktami tylko wtedy, gdy fakt, na

którym ufundowane jest zachodzenie innego faktu, jest bardziej fundamentalny. Gdyby więc fakt, że  $p$ , fundował zachodzenie faktu, że  $q$ , pierwszy musiałby być faktem bardziej fundamentalnym od drugiego. Z założenia jednak to fakt, że  $q$ , miał fundować fakt, że  $p$ , zatem ten pierwszy fakt jest bardziej fundamentalny. Przykładowo, jeśli fakt, że pewien przedmiot posiada własność bycia najwyższym budynkiem, jest ufundowany na fakcie, że przedmiot ten pozostaje w relacji bycia wyższym w stosunku do dowolnego budynku, to ostatni fakt nie może być ufundowany na pierwszym fakcie. Mamy zatem kolejną formalną własność relacji fundowania, mianowicie *antysymetryczność*, tj. jeśli fakt, że  $p$ , jest ufundowany na fakcie, że  $q$ , to nieprawda, że fakt, że  $q$ , jest ufundowany na fakcie, że  $p$ .

Relacja fundowania wydaje się również być relacją *przechodnią*, tj. jeśli fakt, że  $p$ , jest ufundowany na fakcie, że  $q$  i fakt, że  $q$ , jest ufundowany na fakcie, że  $r$ , to fakt, że  $p$ , jest ufundowany na fakcie, że  $r$ . Przykładowo, jeśli fakt, że prawnie zakazana jest jazda samochodem pod wpływem alkoholu jest ufundowany na intencjonalnych działaniach ustawodawców, a działania ustawodawców są ufundowane na pewnych ruchach ich ciał i pobudzeniach pewnych grup neuronów w ich mózgach, to pierwszy fakt jest ufundowany na ostatnim.

Jeśli relacja fundowania ma pełnić rolę w wyjaśnianiu, to jeśli kolekcja faktów  $X$  funduje fakt, że  $p$ , to kolekcja faktów  $X$  nie może zawierać faktów, od których zachodzenia zachodzenie faktu, że  $p$ , nie jest zależne. Fakty takie są bowiem irrelewantne przy udzielaniu odpowiedzi na pytanie o to, dlaczego fakt, że  $p$ , zachodzi. Na tej podstawie możemy zidentyfikować kolejną własność relacji fundowania, a mianowicie jej *niemonotoniczność*, gdzie o relacji  $R$  powiemy, że jest monotoniczna wtw, gdy jeśli  $R$  zachodzi pomiędzy  $x$  i  $y$ , to  $R$  zachodzi pomiędzy  $x$  i  $(y$  i  $z)$ . Gdybyśmy uznali relację fundowania za monotoniczną w tym sensie, musielibyśmy uznać, że jeśli fakt, że  $p$ , jest ufundowany przez kolekcję faktów  $X$ , to fakt, że  $p$ , jest ufundowany

przez kolekcję faktów  $X$  i  $Y$ , gdzie do  $Y$  należy dowolny fakt, który nie należy do kolekcji faktów  $X$  i jest różny od faktu, że  $p$ . Jeśli na fakty należące do kolekcji  $Y$  nakłada się jedynie te dwa warunki, to do kolekcji faktów  $Y$  będą należały wszystkie fakty, które nie należą do  $X$  i są różne od faktu, że  $p$ . W skrajnym przypadku do sumy kolekcji faktów  $X$  i  $Y$  będą należałyby wszystkie fakty różne od faktu, że  $p$ , który byłby wówczas fundowany przez wszystkie pozostałe fakty. Wśród faktów będą znajdowały się również fakty, które są ufundowane na fakcie, że  $p$ , co prowadzi do pogwałcenia antysymetryczności relacji fundowania.

Z uwagi na to, że zachodzenie relacji fundowania stanowi wyjaśnienie zachodzenia pewnego faktu w kategoriach zachodzenia pewnego innego faktu, związek pomiędzy tymi faktami powinien być metafizycznie konieczny. Zatem, jeśli fakt, że  $p$ , jest ufundowany na fakcie, że  $q$ , to z konieczności, jeśli fakt, że  $q$ , istnieje, to istnieje fakt, że  $p$  (przy założeniu, że metafizyczna natura bytów stanowiących elementy konstytutywne obu faktów nie ulega zmianie).

Jeśli przyjąć, że relacja fundowana jest wystarczająco zrozumiałą relacją, możemy uznać, iż relacja zależności prawdziwości twierdzeń od bytu jest relacją fundowania zachodzącą pomiędzy prawdziwością twierdzenia a bytem:

(FT) Dla każdego twierdzenia  $p$ , jeśli twierdzenie  $p$  jest prawdziwe, to istnieje taki fakt  $x$ , że  $x$  funduje prawdziwość twierdzenia  $p$ .

Wówczas możemy powiedzieć, że

twierdzenia teorii  $T$  wyrażonej w ramach egzystencjalnego dyskursu  $D$  pociągają za sobą zobowiązania ontologiczne co do istnienia bytów klasy  $K$  wtw, gdy prawdziwość twierdzeń egzystencjalnych teorii  $T$  wyrażonej w ramach dyskursu  $D$  jest



ufundowana na zachodzeniu faktów, dla których byty klasy  $K$  są elementami konstytutywnymi,<sup>11</sup>

gdzie do klasy  $K$  należą byty, do których odnoszą się wyrażenia pierwotne występujące w twierdzeniach tego dyskursu.

Z analogicznych powodów jak w przypadku (TM), (TM\*) i (L-TSB), (FT) nie pozwala na uniknięcie kłopotliwych zobowiązań ontologicznych implikowanych przez negatywne twierdzenia egzystencjalne oraz przygodnie prawdziwe twierdzenia uniwersalne.

Podsumowując powyższe rozważania, możemy w następujący sposób scharakteryzować pojęcie dyskursu zobowiązującego ontologicznie:

(DZ) Dyskurs  $D^*$  jest ontologicznie zobowiązujący wtw, gdy istnieje taka teoria  $T$  wyrażona w ramach tego dyskursu, że prawdziwość jej pozytywnych twierdzeń egzystencjalnych jest ufundowana na faktach, dla których byty klasy  $K^*$  są elementami konstytutywnymi oraz byty klasy  $K^*$  są fundamentalnymi elementami rzeczywistości,

gdzie byty klasy  $K^*$  są bytami, do których odnoszą się wyrażenia dyskursu  $D^*$ .

Zaletą (DZ) jest to, że wprowadzenie niezwrotnej relacji fundowania dobrze oddaje — w odróżnieniu od relacji superwencji — intuicję, że to prawdziwość powinna być wyjaśniana przez odwołanie się do istnienia pewnych bytów, a nie odwrotnie. Z kolei niemonotoniczność relacji fundowania sprawia, że — w odróżnieniu od relacji uprawdziwiania — prawdziwość prawd koniecznych nie jest ufundowana na istnieniu dowolnego przedmiotu. Wreszcie, dzięki temu, że relacja fundowania jest taka, że fundujący fakt wyjaśnia zachodzenie faktu fundowanego, relacja ta zachodzi wyłącznie pomiędzy faktami, dla których istnienie jednego z nich jest istotne dla istnienia drugiego z nich.

---

<sup>11</sup>W takim kierunku podążają rozważania K. Fine'a w jego artykule (2001).

Mając zdefiniowane pojęcie dyskursu zobowiązującego ontologicznie, musimy się zastanowić, w jaki sposób powinniśmy teraz zdefiniować pojęcie dyskursu niezobowiązującego ontologicznie, by stanowiło ono formę zaprzeczenia temu, że prawdziwość twierdzeń tego dyskursu zależna jest od tego, co istnieje jako fundamentalny element rzeczywistości. Poszukując takiego stanowiska, najlepiej przyjrzeć się formom antyrealistycznych stanowisk w metafizyce. Najpopularniejszą formą stanowiska antyrealistycznego jest redukcjonizm, zgodnie z którym byty klasy  $K$  nie istnieją, ponieważ redukowalne są bytów klasy  $K^*$ . Należy tutaj zauważyć, że jeśli zdefiniujemy redukcję w kategoriach relacji fundowania, tj. uznamy, że byty jednej klasy redukują się do bytów drugiej klasy wtw, gdy istnienie pierwszych jest ufundowane na istnieniu drugich, to w celu nadania jakiegokolwiek formie redukcjonistycznego antyrealizmu antyrealistycznego charakteru należy dodać, iż jedynymi realnie istniejącymi bytami są byty stanowiące elementy fundamentalnej rzeczywistości. W literaturze występuje też klasa stanowisk antyrealistycznych przyjmujących nieredukcjonistyczne formy, z których na wyróżnienie zasługują dwie z nich. Zgodnie z pierwszą z nich prawdziwość (w pewnym sensie) twierdzeń danego dyskursu nie jest fundowana na istnieniu bytów, do których mają odnosić się wyrażenia tego dyskursu, ponieważ byty te nie istnieją i nie istnieje taka klasa bytów, do której dałoby się je zredukować. W tę klasę stanowisk antyrealistycznych wpisują się stanowiska fikcjonalistyczne (zob. rozdział 2). Zgodnie z drugą formą nieredukcjonistycznych stanowisk antyrealistycznych prawdziwość (w pewnym sensie) twierdzeń danego dyskursu nie jest ufundowana na istnieniu żadnych bytów, ponieważ funkcją wyrażen tego dyskursu w ogóle nie jest odnoszenie się do jakichkolwiek bytów. W tę klasę stanowisk wpisują się stanowiska instrumentalistyczne (zob. §5.1.). Niech dyskurs  $D$  będzie dyskursem, który interpretuje się na antyrealistyczny sposób, wówczas wszystkie wyróżnione wyżej formy antyrealizmu wydają się spełniać następujące

kryterium:

dla dowolnego pozytywnego twierdzenia egzystencjalnego teorii  $T$  wyrażonej w ramach tego dyskursu, nie jest prawdą, że istnieje klasa bytów  $K$  taka, że prawdziwość tych twierdzeń jest ufundowana na istnieniu tych bytów i byty klasy  $K$  stanowią elementy fundamentalnej rzeczywistości,

gdzie byty klasy  $K$  są bytami bądź stanowią elementy konstytutywne bytów, do których odnoszą się wyrażenia występujące w twierdzeniach egzystencjalnych tego dyskursu, o ile w ogóle wyrażenia dyskursu  $D$  pełnią taką funkcję.

Na kanwie powyższych rozważań można sformułować następującą definicję dyskursu niezobowiązującego ontologicznie:

DN) Dyskurs  $D$  jest ontologicznie niezobowiązujący wtw, gdy dla dowolnej teorii  $T$  wyrażonej w ramach tego dyskursu, nie istnieje taka klasa bytów  $K$ , że prawdziwość pozytywnych twierdzeń egzystencjalnych tej teorii jest ufundowana na istnieniu bytów klasy  $K$  i byty klasy  $K$  są fundamentalnymi elementami rzeczywistości.

# Rozdział II

## Dwa typy fikcjonalizmu

Stanowiskiem, w ramach którego twierdzenia dyskursu D są interpretowane jako twierdzenia nie pociągające za sobą zobowiązań ontologicznych co do istnienia bytów klasy  $K$ , jest *fikcjonalizm*. Współczesna forma tego stanowiska pojawiła się wraz z opublikowaniem *Science without Numbers* Hartryego Fielda (1980) oraz *Scientific Image* Basa van Fraassen'a (1980), w których podjęto się obrony fikcjonalizmu w sprawie — odpowiednio — bytów matematycznych i przedmiotów nieobserwowalnych. Choć stanowiska sformułowane w tych rozprawach dotyczyły określonych dyskursów, to każde z nich posiada cechy charakterystyczne wspólne wszystkim wersjom fikcjonalizmu.

Gdy przybliża się główne idee fikcjonalizmu, bardzo często wychodzi się od kreślenia analogii pomiędzy twierdzeniami dyskursu D, poddanymi fikcjonalistycznej interpretacji, a twierdzeniami fikcyjnymi, w których mówi się o fikcyjnych postaciach, fikcyjnych zdarzeniach itp. W przypadku ostatnich powszechny ma być pogląd, że można akceptować te twierdzenia jako *prawdziwe w pewnym sensie*, mimo iż nie przyjmuje się istnienia przedmiotów, których twierdzenia te miałyby dotyczyć, gdyby były odczytywane dosłownie. O twierdzeniu fikcyjnym powiemy, że jest *odczytywane dosłownie*, gdy bierze się je

za twierdzenie, z którego wynikają wszystkie konsekwencje standardowo wynikające z treści tych twierdzeń. Przykładowo, gdy fikcyjne twierdzenie głosi, że istnieją elfy zamieszkujące lasy, to wynika z niego, że istnieją takie przedmioty. Choć fikcjoniści chętnie powołują się na tę analogię (zob. Rosen 1990, s. 331–332; Rosen 2005, s. 15; Field 1989, s. 3) i choć powszechna jest intuicja, że akceptując twierdzenia fikcyjne, nie przyjmuje się istnienia fikcyjnych przedmiotów, to w filozofii sprawa jest nieco bardziej skomplikowana. Już krótki przegląd filozoficznej literatury poświęconej dyskursom o przedmiotach fikcyjnych pokazuje, że fikcjonalizm jest tylko jednym ze stanowisk, które starają się oddać sprawiedliwość tej intuicji, a stanowiska negujące tę intuicję należą do zdecydowanej mniejszości.

Większość realistów w sprawie istnienia przedmiotów fikcyjnych (zob. van Inwagen 1977, Parsons 1991, Zalta 2000) nie utożsamia ich z przedmiotami, które nie odbiegałyby znacząco od nas pod względem ich rodzaju. Jedynym wyjątkiem jest tu D. Lewis, który w swoim artykule *Prawda w fikcji* (1978/1991) przyjmuje istnienie przedmiotów fikcyjnych, nie różniących się pod względem rodzaju od przedmiotów znajdujących się w naszym otoczeniu. W ogólności jednak realiści w sprawie istnienia przedmiotów fikcyjnych utożsamiają je z bytami abstrakcyjnymi. Parsons przyjmuje, że przedmioty fikcyjne są zbiorami tych własności, które przypisuje się im w powieściach, w których się pojawiają. Zalta identyfikuje je z bytami abstrakcyjnymi, które w pewien szczególny sposób, zwany *kodowaniem*, posiadają własności przypisywane im w powieściach, w których się pojawiają.

W ramach większości tych stanowisk wychodzi się zatem od odrzucenia dosłownego odczytania fikcyjnych twierdzeń jako mówiących o przedmiotach niewiele różniących się od nas samych i uznania, że de facto mówią one o pewnych przedmiotach abstrakcyjnych, które mają z nami niewiele wspólnego. Najprawdopodobniej już samo utożsamienie przedmiotów fikcyjnych z odpowiednimi przedmiotami abstrakcyj-

nymi wystarczy do tego, by oddać sprawiedliwość potocznej intuicji, że przedmioty fikcyjne (traktowane jako podobne do nas i naszego otoczenia) nie istnieją, a mimo to można uważać twierdzenia fikcyjne za prawdziwe — za prawdziwe, gdy odczyta się je jako twierdzenia dotyczące odpowiednich przedmiotów abstrakcyjnych.

Na tle tych stanowisk realistycznych w sprawie istnienia przedmiotów fikcyjnych fikcjonalizm wyróżnia to, że zgodnie z nim najlepszym sposobem odczytania twierdzeń dotyczących przedmiotów fikcyjnych jest ich dosłowne odczytanie jako twierdzeń mówiących właśnie o przedmiotach niewiele różniących się pod względem rodzaju od nas samych i naszego otoczenia. Z racji tego, że fikcjonałiści zarazem odrzucają istnienie takich przedmiotów lub pozostają agnostykami w sprawie ich istnienia, dosłownie odczytane twierdzenia fikcyjne są uznawane przez nich za literalnie fałszywe bądź zawieszają oni sąd w sprawie ich wartości logicznej. Są jednak przekonani, że twierdzenia fikcyjne mogą być akceptowane z uwagi na pewne racje niezależne od ich literalnej prawdziwości. Dodają nawet, że można uznać te twierdzenia za *prawdziwe w pewnym sensie*, mianowicie za *prawdziwe zgodnie z pewną fikcją* bądź za *prawdziwe w ramach pewnej gry w udawanie*.

W tym rozdziale wyjdę od wskazania cech wspólnych wszystkim wersjom fikcjonalizmu (§2.1), następnie przejdę do nakreślenia różnicy pomiędzy dwiema wersjami fikcjonalizmu: metafikcjonalizmem i fikcjonalizmem przedmiotowym (§2.2), by na koniec przejść do krótkiego zestawienia obu stanowisk (§2.3).

## 1 Wspólne cechy różnych odmian fikcjonalizmu

Czym zatem jest fikcjonalizm? Niezależnie od tego, z jaką formą fikcjonalizmu mamy do czynienia, Gideon Rosen wyróżnia następujące

wspólne cechy tych stanowisk:

1. „W opozycji do instrumentalizmu i nonkognitywizmu fikcjonalista utrzymuje, że twierdzenia [...] danego dyskursu są autentycznymi reprezentacjami tego, jak mają się rzeczy, i — z uwagi na ostatnie — są zdolne do bycia prawdziwymi lub fałszywymi” (Rosen 2005, s. 14).
2. „W opozycji do jednego rodzaju redukcjonizmu, fikcjonalista utrzymuje, że język tego dyskursu powinien być interpretowany *řdosłownie*. Twierdzenia [...] tego dyskursu autentycznie implikują to, co uważa się za naturalnie wynikające z nich. Zatem, jeśli teoria wydaje się głosić, że np. każdy ma anioła stróża w niebie, to jest ona prawdziwa jedynie, gdy istnieją anioły w niebie” (Rosen 2005, s. 14).
3. „W opozycji do jednego rodzaju realizmu, fikcjonalista utrzymuje, że ostatecznym celem dyskursu w danej dziedzinie nie jest (lub nie musi być) dostarczenie prawdziwej koncepcji tej dziedziny, ale raczej dostarczenie koncepcji o pewnych *řzaletach* — zaletach, które teoria może posiadać, nie będąc prawdziwą” (Rosen 2005, s. 14)<sup>12</sup>.

Frederick Kroon podsumowuje trzy powyższe punkty, proponując następującą charakterystykę spełnianą przez dowolną wersję fikcjonalizmu:

Stanowisko *S* w sprawie twierdzeń dyskursu *D* jest *fikcjonalizmem* wtw, gdy zgodnie z *S* twierdzenia dyskursu *D* posiadają reprezentacjonistyczną treść, wyznaczaną przez ich warunki prawdziwości, nie istnieją przedmioty, których

---

<sup>12</sup>Zob. również (Field 1989, s. 2–3; Kalderon 2005, s. 1–5; Kalderon 2005b, s. 112; Kroon 2011, s. 787, 791–792; Sainsbury 2010, s. 173–174; van Fraassen 1980, s. 12).

mają dotyczyć twierdzenia tego dyskursu oraz właściwą postawą wobec twierdzeń dyskursu D nie jest przekonanie o ich prawdziwości, ale ich akceptacja ze względu na zalety niezależne od prawdziwości (Kroon, 2011, s. 791).

W obu przypadkach pojawiają się tezy wymagające dalszego wyjaśnienia. W tym paragrafie podejmę się próby udzielenia odpowiedzi na następujące pytania: „W jakim sensie twierdzenia dyskursu D są autentycznymi reprezentacjami tego, jak rzeczy się mają?” oraz „Czym jest dosłowna interpretacja twierdzeń dyskursu D?”. Kwestię tego, w jaki sposób powinna być pojmowana ontologicznie niezobowiązująca akceptacja tych twierdzeń, odkładam do następnego paragrafu, gdyż odmienne formy fikcjonalizmu różnią się pod tym właśnie względem.

Zacznijmy zatem od pierwszego pytania: *w jakim sensie twierdzenia dyskursu D są autentycznymi reprezentacjami tego, jak rzeczy się mają?* Zgodnie z (1) o twierdzeniu *p* dyskursu D powiemy, że jest autentyczną reprezentacją tego, jak się rzeczy mają, gdy *p* jest *zdolne do posiadania wartości logicznej (truth-apt)*. Gdy filozofowie mówią o zdolności do posiadania wartości logicznej przez twierdzenia dyskursu D, mogą mieć na myśli co najmniej trzy rzeczy:

1. (a) to, że twierdzenia dyskursu D należą do gramatycznej kategorii zdań oznajmujących (por. Jackson, Oppy i Smith 1994, §2.);
2. (b) to, że twierdzenia dyskursu D należą do gramatycznej kategorii zdań oznajmujących oraz istnieje społeczna praktyka wyznaczająca standardy poprawnego użycia tych twierdzeń (por. Jackson, Oppy i Smith 1994, §3.);
3. (c) to, że twierdzenia dyskursu D posiadają gramatyczną formę zdań oznajmujących, są rządzone przez usankcjonowane społeczną praktyką reguły wyznaczające poprawne ich użycia oraz



*mogą stanowić treści przekonań* (por. Jackson, Oppy i Smith 1994, §4.).

Z uwagi na to, że przyjęcie tezy, iż twierdzenia dyskursu D są reprezentacjami tego, jak się rzeczy mają, ma odróżniać fikcjonalizm od nonkognitywizmu, zarówno stanowisko wyrażone w (a), zwane *syntaktycyzmem*, jak i stanowisko wyrażone w (b), zwane *zdyscyplinowanym syntaktycyzmem*, nie powinny stanowić zamierzonych interpretacji pojęcia zdolności do posiadania wartości logicznej. Oba stanowiska są bowiem spójne z nonkognitywizmem.

Zgodnie z nonkognitywizmem twierdzenia dyskursu D nie posiadają treści, która byłaby wartościowalna logicznie, ponieważ twierdzenia te nie są reprezentacjami tego, jak się rzeczy mają, a jedynie wyrazami niepropozycjonalnych postaw, jakie osoba posiadająca dyspozycję do ich wygłoszenia może zajmować wobec pewnych przedmiotów, zdarzeń, czynów itd. Przykładowo, twierdzenie „Tortury są złe” jest w ramach nonkognitywizmu interpretowane jako wyraz dezaprobaty dla tortur, przez co jest treściowo bliskie nieoznajmującemu twierdzeniu „Nie dla tortur!”. Syntaktycyzm nakłada na twierdzenia dyskursu D zdolne do posiadania wartości logicznej jedynie warunek posiadania formy gramatycznej zdania oznajmującego. Warunek ten jest spełniony przez wszystkie twierdzenia posiadające nonkognitywistyczną interpretację, a mimo to przy ich nonkognitywistycznej interpretacji nie posiadają wartości logicznej.

W gruncie rzeczy wiarygodne przykłady twierdzeń przywoływanych przez nonkognitywistów można traktować jako kontrprzykłady do syntaktycyzmu. Odsłaniają one największą wadę tego stanowiska. Polega ona na tym, że przy ustalaniu, czy dane twierdzenie w ogóle kwalifikuje się do posiadania wartości logicznej, zupełnie ignoruje się jego znaczenie. Jak pokazują przykłady nonkognitywistów, zdania w formie oznajmującej mogą mieć wiele odmiennych znaczeń, z których niektóre nie są wartościowalne logicznie.

Zdyscyplinowany syntaktycyzm wydaje się stanowić krok w pożądanym kierunku, gdyż do powyższego warunku dodaje również warunek podlegania społecznie usankcjonowanym regułom rządzącym poprawnym użyciem danego twierdzenia, które powinny (przynajmniej do pewnego stopnia) wyznaczać znaczenia twierdzeń dyskursu D. Warunek ten nadal nie jest wystarczający do wykluczenia nonkognitywistycznej interpretacji zdań tego dyskursu. Nonkognitywiści mogą wskazać na istniejącą praktykę społeczną wyznaczającą okoliczności, w których użycie zdania „Tortury są złe” jest zasadne — jest ono zasadne zawsze, gdy osoba wypowiadająca to zdanie odczuwa dezaprobatę względem tortur (zob. Jackson, Oppy i Smith 1994, §3.).

Co więcej, ani zdyscyplinowany syntaktycyzm, ani tym bardziej syntaktycyzm nie oddają sprawiedliwości intuicji, że tym, co wyróżnia wartościowalne logicznie twierdzenia, jest to, że mogą one stanowić treści naszych przekonań. Gdy ktoś jest przekonany, że psy szczekają, jest zarazem przekonany o prawdziwości twierdzenia wyrażającego treść tego przekonania. Nałożenie warunku możliwego wyrażania treści przekonań na znaczenie każdego twierdzenia zdolnego do posiadania wartości logicznej pozwala wykluczyć nonkognitywistyczne interpretacje takich twierdzeń z grona interpretacji, zgodnie z którymi twierdzenia te zdolne są do posiadania wartości logicznych. Twierdzenia poddane nonkognitywistycznej interpretacji z zasady mają bowiem wyrażać niepropozycjonalne postawy i jako takie nie mogą być treściami przekonań. Zaletą stanowiska (c) jest to, że przypisuje się w nim twierdzeniom dyskursu D wszystkie własności, które w naturalny sposób łączą się ze zdolnością do posiadania wartości logicznej. Jednak uwzględnienie możliwości bycia treścią przekonań sprawia, że pojęcie zdolności do posiadania wartości logicznej zaczyna być pojęciem uwikłanym w takie przedmiotowe kwestie jak to, czy stan, w którym znajduje się podmiot wypowiadający dane twierdzenie, jest stanem przekonania. Z ostatnim z kolei wiążą się takie kwestie jak to,

czy treść tego przekonania ma na celu „pasowanie” do tego, jak rzeczy się mają, czy treść tego przekonania pozostaje w odpowiedniej relacji do zdolności podejmowania racjonalnego działania na jej podstawie itd. (zob. Jackson, Oppy i Smith 1994, §4.).

Ostatnie wydaje się pozostawać w napięciu z minimalistycznymi koncepcjami prawdziwości takimi jak deflacionizm. Zgodnie z deflacyjną koncepcją prawdy twierdzenie, że zdanie „ $p$ ” jest prawdziwe, jest poznawczo równoważne z twierdzeniem, że  $p$  (tj. reguły inferencyjne upoważniają użytkownika języka do przejścia od jednego z tych twierdzeń do drugiego i odwrotnie). Predykat „jest prawdziwy” ma mieć zastosowanie tylko do zdań, które użytkownik języka rozumie, a jego znaczenie w pełni oddaje schemat (T):

(T) „ $p$ ” jest prawdziwe wtw, gdy  $p$  (por. Field 1994/2001, s. 222).

W kontekście sporu o to, na ile bogate powinno być pojęcie prawdy, deflacionista może utrzymywać, że przyjęcie jego koncepcji prawdziwości dostarcza podstaw do łatwego obalenia wszelkiej maści stanowisk nonkognitywistycznych. Ostatnie wymagają bowiem, by treść predykatu prawdziwości była bogatsza, niż wyraża to schemat (T). Jest tak dlatego, że zgodnie z nonkognitywizmem istnieją twierdzenia, które nie są prawdziwe, mimo iż *prima facie* mogą być podstawione w schemacie (T).

Z drugiej jednak strony, istnienie nonkognitywistycznych interpretacji zdań dyskursu D może być traktowane jako wyzwanie rzucone właśnie zwolennikowi deflacyjnej koncepcji prawdziwości. Nawet jeśli przyjąć, że treść samego predykatu prawdziwości jest oddawana przez schemat (T), to jeszcze pozostaje kwestią otwartą to, do których twierdzeń predykat prawdziwości w ogóle ma zastosowanie. Z tego punktu widzenia wysoce niesatysfakcjonująca byłaby ewentualna odpowiedź deflacionisty, że twierdzenia takie jak „Tortury są złe” są zdolne do posiadania wartości logicznej, ponieważ są zrozumiałe i *prima facie* można je podstawić za zmienne zdaniowe w schemacie (T). Można

bowiem kontrargumentować, że jeśli twierdzenia te są wyrazami niepropozycjonalnych postaw, to nie powinny one stanowić podstawień za zmienne zdaniowe w schemacie (T). Przy próbie ustalenia, które z twierdzeń mogą stanowić podstawienia schematu (T), okaże się, że konieczne jest odwołanie się do takich właściwości tych twierdzeń, o których nie ma mowy w schemacie (T). Przykładem dodatkowej cechy, którą powinny posiadać twierdzenia podstawialne w schemacie (T), może być to, że istnieje społeczna praktyka używania twierdzeń dyskursu D, która nadaje im takie znaczenia, że mogą one wyrażać treści przekonań.

Wszystko to sprawia, że pojęcie prawdziwości przestaje legitymować się minimalistyczną treścią w pełni oddawaną przez schemat (T). Frank Jackson, Graham Oppy i Michael Smith podsumowują sformułowane wyżej uwagi w następujący sposób:

Minimalizm jest czasami charakteryzowany jako uogólnienie idei, że bycie prawdziwym przez *śnieg jest biały* nie polega na niczym innym, jak na tym, że *śnieg jest biały*. Jeśli rozumieć to jako doktrynę dotyczącą zdań, to jest ona fałszywa. Koniecznym warunkiem bycia prawdziwym zdaniem jest to, że jest ono zdolne do posiadania wartości logicznej, co — jak widzieliśmy — obejmuje znacznie więcej niż to, że *śnieg jest biały*. Bogaty wzorzec użyć i konwencji wymagany przez bycie zdolnym do posiadania wartości logicznej dalece wykracza poza to, że *śnieg jest biały* (zob. Jackson, Oppy i Smith 1994, §6).

Powyższe rozważania wydają się pokazywać, że do konsekwencji pojęcia zdolności do posiadania wartości logicznej należy m.in. to, że pojęcie prawdy musi być rozumiane w jakimś nieminimalistycznym sensie. W związku z tym, gdy będziemy mówić o tym, że twierdzenia dyskursu D mogą być prawdziwe, pojęciu prawdziwości będziemy nadawali nieminimalistyczną interpretację.

Zgodnie z tezą (2) twierdzenia dyskursu D powinny być interpretowane *dosłownie*. Z uwagi na to, że teza ta ma stanowić wyraz przeciwstawienia fikcjonalizmu *redukcjonizmowi* w ramach ontologii, jednym z warunków, jakie musi spełnić dosłowna interpretacja twierdzeń dyskursu D, jest to, że powinna z niej wynikać negacja redukcjonizmu. Załóżmy, że twierdzenia dyskursu D mają na celu mówienie o bytach klasy  $K$  oraz że twierdzenia dyskursu  $D^*$  mają na celu mówienie o bytach klasy  $K^*$ . O twierdzeniach dyskursu D powiemy, że są *redukowalne* wtw, gdy istnieje dyskurs  $D^*$  taki, że istnieje taka klasa bytów  $K^*$ , że twierdzenia tego dyskursu mają na celu mówienie o bytach klasy  $K^*$ , i twierdzenia dyskursu D mają na celu mówienie o bytach klasy  $K$ , klasy  $K$  i  $K^*$  są rozłączne, i prawdziwość twierdzeń dyskursu D jest ufundowana na prawdziwości twierdzeń dyskursu  $D^*$  (por. Rosen 2010, §10). Negacją ostatniej tezy jest teza o *nieredukowalności* twierdzeń dyskursu D. W tym wypadku będzie ona równoznaczna z tezą, że nie istnieje dyskurs  $D^*$  taki, że istnieje klasa bytów  $K^*$  takich, że twierdzenia dyskursu  $D^*$  mają na celu mówienie o bytach tej klasy, lub nie jest prawdą, że klasy  $K$  i  $K^*$  są rozłączne, lub nie jest prawdą, że prawdziwość twierdzeń dyskursu D jest ufundowana na prawdziwości twierdzeń dyskursu  $D^*$ .

Jak przyjąłem w poprzednim rozdziale, to, do jakich bytów mają na celu odnosić się twierdzenia jakiegokolwiek dyskursu, powinno być wyznaczane na podstawie tego, jakie są warunki prawdziwości tych twierdzeń. O twierdzeniach pewnego dyskursu powiemy, że mają na celu odnoszenie się do bytów klasy  $K$  wtw, gdy twierdzenia, w których występują wyrażenia referencjalne tego dyskursu, nie byłyby prawdziwe, gdyby nie istniały byty tej klasy. Wprowadzone wyżej pojęcia pozwalają na dokonanie następującej analizy tezy, że twierdzenia dyskursu D są *autentycznymi reprezentacjami tego, jak rzeczy się mają*:

Twierdzenia dyskursu D mają na celu *reprezentowanie* bytów

klasy  $K$  wtw, gdy twierdzenia dyskursu  $D$  mają na celu odnoszenie się do bytów klasy  $K$ , predykaty dyskursu  $D$  mają na celu odnoszenie się do własności, które byty klasy  $K$  posiadałyby, gdyby istniały i celem formułowania twierdzeń dyskursu  $D$  jest przypisanie bytom klasy  $K$  własności, które posiadałyby, gdyby istniały.

Dotychczas posługiwałem się w intuicyjny sposób pojęciem warunków prawdziwości zdań dyskursu  $D$ , nie mówiąc wiele o jego ewentualnym związku z pojęciem znaczenia twierdzeń dyskursu  $D$ . Zwykle jednak fikcjonałści, przyjmując tezę, że treść twierdzeń dyskursu  $D$  reprezentuje to, jak rzeczy się mają, przyjmują warunkowoprawdziwościową koncepcję znaczenia twierdzeń tego dyskursu. Taką też koncepcję znaczenia dla zdań dyskursu  $D$  przyjmę w tej rozprawie<sup>13</sup>. U podstaw tej koncepcji znaczenia leży następująca idea:

W celu powiedzenia, czym *jest* znaczenie, możemy zapytać, jaką funkcję pełni znaczenie i znaleźć coś, co pełni tę funkcję. Znaczeniem zdania jest coś, co wyznacza warunki, pod którymi zdanie to jest prawdziwe lub fałszywe (Lewis 1970/1983, s. 193)

Zatem zgodnie z warunkowoprawdziwościową koncepcją znaczenia znaczeniem zdania są jego warunki prawdziwości. Czym są warunki prawdziwości zdania?

---

<sup>13</sup>Jedynym znanym mi wyjątkiem jest H. Field (zob. 1994/2001), który choć jest fikcjonałistą, akceptuje również deflacyjną koncepcję prawdy. Nie chcę jednak podejmować się tutaj kwestii tego, w jaki sposób w szczegółach mógłby przedstawiać się fikcjonalizm Fielda. Zamiast tego przyjmę, że w przypadku fikcjonalizmu modalnego warunkowoprawdziwościowa koncepcja znaczeń twierdzeń mających na celu mówienie o światach możliwych jest koncepcją właściwą. Za przyjęciem tego założenia może przemawiać fakt, że fikcjonałści starają się poddać fikcjonałistycznej interpretacji dyskurs modalnego realisty. Czołowy przedstawiciel tego nurtu — D. Lewis był zwolennikiem warunkowoprawdziwościowej koncepcji znaczenia (zob. Lewis 1970/1983). Dlatego jestem przekonany, że pozostawienie na boku szczegółów stanowiska Fielda nie przedstawi w fałszywym świetle głównego stanowiska, które będzie dyskutowane w tej rozprawie.

W rozważaniach na temat warunków prawdziwości ograniczę się do rekonstrukcji poglądów Carnapa, wyrażonych w książce *Meaning and Necessity* (1956), które uzupełnię o inspiracje z pracy Davida Lewisa *Languages and Language* (1970/1983). Jestem przekonany, że koncepcje warunków prawdziwości zaproponowane przez tych autorów są w wystarczające do tego, aby w oparciu o nie przybliżyć zręby koncepcji znaczenia akceptowanej przez fikcjonalistów.

Zgodnie z propozycjami tych filozofów warunki prawdziwości danego zdania mogą być reprezentowane jako zbiory możliwych stanów rzeczy, w których zdanie to jest prawdziwe bądź fałszywe<sup>14</sup>. Znaczenie zdania powinno zatem wyznaczać jego wartości logiczne relatywnie do odmiennych możliwych stanów rzeczy. Innymi słowy, powinno oddawać zależność pomiędzy wartościami logicznymi zdania a możliwymi stanami rzeczy. Naturalnym kandydatem na byty, które wyznaczają sposób, na jaki coś jest zależne od czegoś innego, są *funkcje*. Funkcje są szczególnym przypadkiem relacji dwuargumentowych takim, że każdy pierwszy człon tej relacji (zwany jej argumentem) pozostaje w tej relacji tylko i wyłącznie do jego bytu (zwanego jej wartością).

Znaczenia zdań mogą być reprezentowane za pomocą funkcji z możliwych stanów rzeczy w wartości logiczne. O dwóch funkcjach  $f_1$  i  $f_2$  mówi się, że są identyczne wtw, gdy dla każdego  $x$  i  $y$ , jeśli  $f_1(x) = y$ , to  $f_2(x) = y$  i odwrotnie. W ramach takiej koncepcji znaczenia o dwóch zdaniach powiemy, że są równoznaczne, gdy posiadają tę samą wartość logiczną relatywnie do tych samych możliwych stanów rzeczy.

W rezultacie otrzymujemy jednak zbyt gruboziarnistą koncepcję znaczeń zdań. Prowadzi ona bowiem do uznania wszystkich prawd koniecznych za równoznaczne. Jest tak dlatego, że prawda konieczna będzie prawdziwa w dowolnej możliwej sytuacji, a jeśli znaczenia zdań są reprezentowane jedynie za pomocą funkcji z możliwych sytuacji w wartości logiczne, to wszystkie prawdy konieczne miałyby to samo

---

<sup>14</sup>Zob. (Carnap 1956; Kripke 1963; Lewis 1970/1983).

znacznie. Rezultat ten jest o tyle niepożądany, że taka koncepcja znaczenia ma niewiele wspólnego z tym, w jaki sposób używamy języka. Zwykle nie mamy najmniejszych problemów z rozpoznaniem różnicy pomiędzy znaczeniem zdania „Każdy kawaler jest nieżonatym mężczyzną” a znaczeniem zdania „ $2 + 2 = 4$ ”. Intuicyjnie rzecz biorąc, różnica pomiędzy znaczeniami tych zdań bierze się z tego, że dotyczą one zupełnie odmiennych przedmiotów. Teoria znaczenia dla dowolnego języka powinna respektować intuicje użytkowników tego języka co do odmienności znaczeń poszczególnych zdań, w przeciwnym bowiem razie nie miałyby nawet szans na wyjaśnienie praktyki językowej użytkowników tego języka.

Naturalnym sposobem oddania sprawiedliwości intuicji, że znaczenia zdań „Każdy kawaler jest nieżonatym mężczyzną” i „ $2 + 2 = 4$ ” są odmienne, jest uznanie, iż znaczenia tych zdań zależne są od znaczeń terminów w nich występujących. Odmienność ostatnich powinna pociągać za sobą odmiennność znaczeń całych zdań. Jednym ze sposobów wyrażenia zależności zachodzącej pomiędzy znaczeniami zdań a znaczeniami wyrażeń składających się na te zdania jest przyjęcie tzw. *zasady kompozycjonalności*. Zgodnie z nią znaczenie złożonego wyrażenia jest wspólnie wyznaczone przez znaczenia wyrażeń składających się na nie oraz przez strukturę tego złożonego wyrażenia<sup>15</sup>.

Przyjęcie tej zasady wymaga jednak od teoretyków znaczenia udzielenia odpowiedzi na pytanie, czym są znaczenia wyrażeń występujących w zdaniach. Bazując na idei, od której wyszliśmy wyżej, gdy pytamy o to, czym jest znaczenie danego wyrażenia, powinniśmy pytać o to, jaką rolę pełni znaczenie tego wyrażenia, a następnie wskazać byt, który może pełnić taką rolę. Okazuje się, że odpowiedzi na pytania, czym jest znaczenie nazw własnych oraz czym jest znaczenie predykatów, przedstawiają się podobnie do odpowiedzi na pytanie, czym jest znaczenie zdania. Znaczeniem nazwy własnej jest to, co

---

<sup>15</sup>W sprawie tego sformułowania zasady kompozycjonalności oraz jego krytyki, zob. (Gendler Szabó 2000).



wyznacza jej ekstensję, tj. indywiduum, do którego się ona odnosi. Z kolei znaczeniem predykatu jest to, co wyznacza jego ekstensję, tj. zbiór przedmiotów, do których predykat ten się stosuje. Podobnie jak wyżej, byt, który mógłby pełnić rolę znaczenia nazw własnych i predykatów, jest funkcją. Jaką funkcją? W przypadku nazw własnych jest to funkcja z możliwych sytuacji na zbiór, którego elementami są: indywiduum, do którego odnosi się dana nazwa, lub „puste indywiduum”, które jest ekstensją nazwy we wszystkich tych możliwych sytuacjach, w których nie istnieje indywiduum, do którego nazwa ta się odnosi (por. Carnap 1956, s. 37). W przypadku predykatów jest to funkcja z możliwych sytuacji na zbiór, którego elementami są zbiory indywiduów (dla wszystkich tych możliwych sytuacji, w których predykat ma zastosowanie) oraz zbiór pusty (dla wszystkich tych możliwych sytuacji, w których predykat ten nie stosuje się do żadnego indywiduum).

Weźmy teraz zdanie „ $Fa$ ”, gdzie „ $a$ ” jest nazwą własną, a „ $F$ ” — predykatem „jest szary”. Co jest znaczeniem zdania „ $Fa$ ”? Zgodnie z wprowadzoną wyżej zasadą kompozycjonalności oraz założeniem, że znaczenie zdania wyznacza jego warunki prawdziwości, znaczenie zdania powinno być funkcją ze znaczeń wyrażeń składających się na to zdanie w wartości logiczne: niech znaczeniem nazwy „ $a$ ” będzie funkcja  $f_a$  i znaczeniem predykatu będzie funkcja  $f_F$ , wówczas znaczeniem zdania będzie funkcja  $f : \langle f_a, f_F \rangle \mapsto \{0, 1\}$ , gdzie  $\langle f_a, f_F \rangle$  jest parą uporządkowaną taką, że oddaje ona strukturę zdania „ $Fa$ ” i jej elementami są znaczenia wyrażeń składających się na to zdanie (odpowiednio znaczenie nazwy własnej „ $a$ ” i znaczenie predykatu „ $F$ ”). W rezultacie uzyskujemy następujące warunki prawdziwości dla zdania „ $Fa$ ”:

Zdanie „ $Fa$ ” jest prawdziwe wtw, gdy do ekstensji predykatu „ $F$ ” należy indywiduum a będące odniesieniem nazwy „ $a$ ”,

czyli zdanie „ $Fa$ ” jest prawdziwe wtw, gdy indywiduum  $a$  należy do zbioru szarych rzeczy.

Z racji tego, że  $f_a$  i  $f_F$  wyznaczają ekstensje nazwy własnej „ $a$ ” i predykatu „ $F$ ” relatywnie do możliwych sytuacji, wartość logiczna zdania również jest wyznaczana relatywnie do możliwych sytuacji. Powiedzieć teraz, że zdanie „ $Gb$ ” ma to samo znaczenie, co zdanie „ $Fa$ ”, to powiedzieć, że funkcje, za pomocą których reprezentuje się znaczenia tych zdań, są *intensjonalnie izomorficzne*. Za Carnapem (1956, s. 59) o znaczeniach dwóch  $n$ -członowych zdań  $p$  i  $q$  powiemy, że są intensjonalnie izomorficzne wtedy i tylko wtedy gdy spełniają warunki dwa warunki:

1. Zdania  $p$  i  $q$  są logicznie równoważne, tj. dla dowolnej możliwej sytuacji przyjmują te same wartości logiczne.
2. Każdy  $m$ -ty człon tego zdania  $p$  (dla  $m : 1 \leq m \leq n$ ) jest logicznie równoważny  $m$ -temu członowi zdania  $q$ .

O dwóch członach odmiennych zdań powiemy, że są logicznie równoważne, gdy dla dowolnej możliwej sytuacji przyjmują tę samą ekstensję.

Niech znaczeniem zdania „ $Fa$ ” będzie funkcja  $f_1 : \langle f_a, f_F \rangle \mapsto \{0, 1\}$ , a znaczeniem zdania „ $Gb$ ” będzie funkcja  $f_2 : \langle f_b, f_G \rangle \mapsto \{0, 1\}$ , wówczas o znaczeniach tych zdań powiemy, że są identyczne, gdy funkcje  $f_1$  i  $f_2$  są intensjonalnie izomorficzne, czyli dla dowolnej możliwej sytuacji funkcje  $f_1$  i  $f_2$  przyjmują te same wartości oraz  $f_a = f_b$  i  $f_F = f_G$ .

Warunki prawdziwości dla zdań, w których występują zmienne związane kwantyfikatorami, wyznacza się relatywnie do wartości przyporządkowanych tym zmiennym. Weźmy zdanie „Istnieje takie  $x_1, \dots, x_n$ , że  $Px_1, \dots, x_n$ ” i zdanie „Dla każdego  $x_1, \dots, x_n$ ,  $Px_1, \dots, x_n$ ” wyrażone w klasycznym rachunku predykatów. Wspólnym założeniem definicji prawdziwości dla tych zdań jest przyjęcie istnienia listy

zmiennych indywiduowych należących do słownika języka, w którym wyrażono te zdania. Listę tę będziemy reprezentowali jako ciąg  $SV = \langle x_1, x_2, x_3, \dots \rangle$ . Każdy ciąg indywiduów  $s = \langle a_1, a_2, a_3, \dots \rangle$  może być skorelowany z ciągiem  $SV$  w taki sposób, że  $m$ -ty element ciągu  $s$  jest przyporządkowany  $m$ -temu wyrazowi ciągu  $SV$  jako jego wartość. W zgodzie z powyższą praktyką przyporządkowanie może być charakteryzowane jako funkcja  $f_v$  z możliwych sytuacji w zbiór ciągów indywiduów. Funkcja ta będzie znaczeniem zmiennych związanych. Dalej, znaczeniem zdania „Istnieje takie  $x_1, \dots, x_n$ , że  $Px_1, \dots, x_n$ ” powinna być funkcja ze znaczeń składowych tego zdania w wartości logiczne. Standardowo warunki prawdziwości tego zdania definiuje się w następujący sposób:

Zdanie „Istnieje takie  $x_1, \dots, x_n$ , że  $Px_1, \dots, x_n$ ” jest prawdziwe wtw, gdy istnieje taki  $n$ -elementowy ciąg przedmiotów  $s$ , że  $m$ -ty wyraz ciągu  $SV$  jest przyporządkowany  $m$ -temu elementowi ciągu  $s$  (dla  $1 \leq m \leq n$ ) i  $s$  należy do ekstensji predykatu  $P$ .

Podobnie znaczeniem zdania „Dla każdego  $x_1, \dots, x_n$ , że  $Px_1, \dots, x_n$ ” powinna być funkcja ze znaczeń składowych tego zdania w wartości logiczne. Standardowo warunki prawdziwości dla tego zdania definiuje się w następujący sposób:

Zdanie „Dla każdego  $x_1, \dots, x_n$ , że  $Px_1, \dots, x_n$ ” jest prawdziwe wtw, gdy dla każdego  $n$ -elementowego ciągu przedmiotów  $s$ , jeśli  $m$ -ty wyraz element tego ciągu  $SV$  jest przyporządkowany  $m$ -temu elementowi ciągu  $s$  (dla  $1 \leq m \leq n$ ), to  $s$  należy do ekstensji predykatu  $P$ .

Na podstawie przywołanych wyżej koncepcji znaczeń zdań można z łatwością pokazać, że funkcje stanowiące znaczenia takich zdań jak „Dla każdego  $x$ , jeśli  $x$  jest kawalerem, to nieprawda, że  $x$  jest żonaty” oraz „ $2 + 2 = 4$ ” nie są intensjonalnie izomorficzne. By

takimi były, funkcje konstytuujące znaczenia tych zdań powinny być logicznie równoważne. Tymczasem wśród funkcji składających się na znaczenie pierwszego zdania nie występuje funkcja, które byłaby logicznie równoważna z jakąkolwiek z funkcji składających się na znaczenie drugiego zdania. Zatem zdania te mają odmienne znaczenia, choć są konieczne prawdziwe.

Bazując na takiej koncepcji znaczenia zdań, możemy nadać następującą treść tezie (2):

twierdzenia dyskursu D są *odczytywane dosłownie* wtw, gdy przyjmuje się, że znaczenia twierdzeń dyskursu D są wyznaczone przez ich warunki prawdziwości (w powyższym sensie) i twierdzenia dyskursu D nie są redukowalne.

Wprowadzona wyżej koncepcja znaczenia zdań dyskursu D rodzi wiele pytań, wśród których jedno jest szczególnie interesujące w kontekście obecnych rozważań, mianowicie: *jaka jest metafizyczna natura funkcji przekształcającej możliwe sytuacje w ekstensje wyrażeń?* Gdyby w standardowy sposób utożsamić funkcje ze zbiorami par uporządkowanych, to znaczenia wyrażeń powinny być zbiorami par uporządkowanych, których elementami konstytutywnymi są możliwe stany rzeczy i ekstensje wyrażeń. Zakładając, że para uporządkowana istnieje tylko o tyle, o ile istnieją jej elementy konstytutywne, natrafiamy na dylemat dla fikcjonalisty modalnego: albo możliwe sytuacje istnieją i wówczas można dokonać rozróżnienia pomiędzy znaczeniami koekstensywnych predykatów, nazw pustych itd., albo możliwe sytuacje nie istnieją, a wówczas nie można dokonać rozróżnienia pomiędzy znaczeniami aktualnie koekstensywnych predykatów, nazw pustych itd. Jeśli istnieją wyłącznie aktualne stany rzeczy, to funkcje, z którymi utożsamia się znaczenia wyrażeń, stanowiłyby zbiory par uporządkowanych, których elementami konstytutywnymi byłyby aktualne stany rzeczy i ekstensja. Sprawiałoby to, że znaczenia koekstensywnych wyrażeń byłyby utożsamione z tymi samymi zbiorami par

uporządkowanych. Prowadziłoby do niepożądanego skutku uznawania koekstensywnych wyrażeń za równoznaczne. Jeśli natomiast istnieją możliwe stany rzeczy różne od aktualnych stanów rzeczy, to projekt fikcjonalisty modalnego byłby skazany na niepowodzenie. Przyjęcie istnienia możliwych sytuacji jest bowiem równoznaczne z zaciągnięciem zobowiązań ontologicznych co do bytów, których istnienie fikcjonałści kwestionują. Przykładowo, znaczeniem zdania „Istnieje świat możliwy, w którym istnieją krasnoludki” byłaby funkcja z możliwej sytuacji, która w tym wypadku byłaby tożsama ze światem możliwym, w którym istnieją krasnoludki. Tymczasem cały projekt fikcjonalistów modalnych miał pozwalać na odrzucenie światów możliwych.

W § 4.4. będę się starał uniknąć tego dylematu, proponując redukcję funkcji wyznaczających znaczenia wyrażeń do odpowiednich reprezentacji mentalnych, które — w odróżnieniu od funkcji — nie będą identyfikowane ze względu na posiadanie tej samej dziedziny i przeciwdziedziny oraz posiadanie tej samej struktury. Będą natomiast identyfikowane ze względu na funkcje, które pełnią w życiu kognitywnym podmiotu posługującego się dyskursem D. Propozycja ta będzie się opierała na — jak sądzę — wiarygodnym założeniu, że można posiadać reprezentacje mentalne niezależnie od tego, czy są one spełnione przez jakikolwiek przedmiot. Założenie to powinno pozwolić na uniknięcie konieczności odwoływania się do dziedzin i przeciwdziedzin funkcji, za pomocą których obecnie reprezentujemy znaczenia wyrażeń. Dzięki temu będzie można uniknąć zarówno konieczności przyjmowania istnienia możliwych sytuacji występujących w dziedzinach tych funkcji, jak i kontrowersyjnych bytów, takich jak „puste indywiduum”. Wciąż jednak charakteryzowanie znaczeń wyrażeń poprzez odwołanie się do funkcji z możliwych sytuacji w ekstensje tych wyrażeń jest o tyle pożyteczne, że pozwoli na modelowanie funkcji pełnionych przez odpowiednie reprezentacje mentalne w życiu kognitywnym podmiotu.

## 2 Metafikcjonalizm versus fikcjonalizm przedmiotowy

Zgodnie z tezą (c) „fikcjonalista utrzymuje, że ostatecznym celem dyskursu w danej dziedzinie nie jest (lub nie musi być) dostarczenie prawdziwej koncepcji tej dziedziny, ale raczej dostarczenie koncepcji o pewnych *řaletach* — zaletach, które teoria może posiadać, nie będąc prawdziwą” (Rosen 2005, s. 14).

Zapotrzebowanie na fikcjonalizm pojawia się zwykle tam, gdzie pewien dyskurs spełnia dwa warunki. Z jednej strony chce się zachować możliwość posługiwania się twierdzeniami tego dyskursu, gdyż uważa się, że dyskurs ten legitymuje się pewnymi zaletami niezależnymi od prawdziwości jego twierdzeń. Z drugiej strony zobowiązania ontologiczne implikowane przez twierdzenia tego dyskursu są na tyle kontrowersyjne, że chce się uniknąć akceptowania jego twierdzeń jako literalnie prawdziwych.

Rzecz jasna, posługiwanie się twierdzeniami dyskursu D wymaga od użytkownika tych twierdzeń, by w jakimś sensie je *akceptował*. Akceptację twierdzeń zwykle charakteryzuje się jako zajęcie rządzonej pewnymi normami postawy wobec treści tych twierdzeń. Jakkolwiek by się te normy przedstawiały, w przypadku fikcjonalizmu nie mogą nakładać na akceptowane twierdzenia warunku, który tylnymi drzwiami wprowadzałby konieczność uznania ich za prawdziwe. Na tym jednak nie koniec. Fikcjonalistę od realisty odróżnia to, że jego zdaniem ostatecznym celem posługiwania się dyskursem D jest dostarczenie pewnej koncepcji, która posiada zalety niezależne od jej prawdziwości. W związku z tym fikcjonalista musi twierdzić, że akceptacja twierdzenia *p* dyskursu D może spełniać każdą normę, jaką nakłada się na akceptację tych twierdzeń w ramach dyskursu D, nawet jeśli *p* jest fałszywe. Dodatkowo, aby twierdzenia dyskursu D mogły być teoretycznie użyteczne, ich akceptacja musi wiązać się z możliwością ich

wykorzystywania jako przesłanek w uzasadnianiu kolejnych twierdzeń (niekoniecznie będących twierdzeniami dyskursu  $D^{16}$ ) jako racji do podejmowania działań itd.

O akceptacji twierdzenia  $p$  dyskursu  $D$ , która spełnia wszystkie normy nakładane na akceptację twierdzeń w ramach dyskursu  $D$ , powiemy, że jest *idealną akceptacją* tego twierdzenia w ramach tego dyskursu (Rosen 2005, s. 16). Normy akceptacji, o których tutaj mowa, są normami *wewnętrznymi* względem dyskursu  $D$  w tym sensie, że to z punktu widzenia tego dyskursu wyznaczają one okoliczności, które stanowią *ostateczne* uzasadnienie dla akceptacji twierdzeń tego dyskursu.

Celem wprowadzenia wewnętrznych norm akceptacji jest skonstruowanie fikcjonalizmu z realizmem w kwestii tego, jakie warunki muszą spełnić twierdzenia, które są idealnie akceptowalne. Nawet realista może zgodzić się, że pewne fałszywe twierdzenia mogą być akceptowane, ponieważ ułatwiają myślenie o pewnej klasie fenomenów. Nie będą to jednak twierdzenia idealnie akceptowalne, jeśli ostatecznym celem realistycznie interpretowanego dyskursu jest dostarczenie prawdziwego obrazu świata. W ramach takiego dyskursu idealnie akceptowalnymi twierdzeniami będą jedynie prawdy, a wewnętrzne normy akceptacji tego dyskursu będą dopuszczały pełną akceptację tylko takich twierdzeń. Z kolei fałszywe twierdzenia mogą być akceptowane przez realistę na gruncie norm akceptacji *zewnętrznych* względem realistycznie interpretowanego dyskursu (por. Rosen 2005, s. 16).

Przyjęcie przez realistę zewnętrznych standardów akceptacji wcale nie jest rzadkim przypadkiem. Realiści najczęściej akceptują nie tyle w pełni wyrażoną teorię, co program badawczy, którego celem ma być dostarczenie prawdziwego obrazu świata (zob. van Fraassen 1980,

---

<sup>16</sup>Relacje pomiędzy twierdzeniami dyskursu  $D$  a twierdzeniami dyskursu  $D^*$  mogą być wyznaczane za pośrednictwem tzw. reguł generowania, o których więcej można przeczytać w § 2.2.2.

s. 12). W początkowej fazie programu badawczego wiele spośród akceptowanych przez realistę twierdzeń ma charakter metaforyczny i przez to akceptowane są jedynie tymczasowo. Oczywiście ostatecznym celem programu jest wyeliminowanie metafor naukowych na rzecz literalnie prawdziwych twierdzeń. W artykule *Postscript on Metaphor* (1981) Quine podaje już klasyczny przykład takiego procesu. We wcześniejszej teorii fal świetlnych naukowcy utrzymywali, że światło jest falą rozchodzącą się w jakiejś substancji. Miał nią być eter. Gdy jednak okazało się, że eter nie istnieje, nie porzucono teorii fal świetlnych, tylko nadal się nią posługiwano, uznając koncepcję fal świetlnych za użyteczną metaforę, którą warto akceptować, nawet jeśli opiera się na fałszywym założeniu, iż istnieje substancja, w której rozchodzą się fale świetlne. Z ostatniego powodu realista mógł akceptować twierdzenia, w których występowało tak rozumiane pojęcie fali świetlnej, jedynie posługując się normami akceptacji zewnętrznymi względem realistycznie interpretowanego dyskursu naukowego. Ostatecznie wprowadzono techniczne pojęcie fali, które wolne było od presupozycji istnienia substancji, w której fala świetlna się rozchodzi. Dzięki temu twierdzenia formułowane przy użyciu nowego pojęcia fali mogły pretendować do bycia akceptowanymi ze względu na normy wewnętrzne względem realistycznie interpretowanego dyskursu naukowego, czyli ze względu na to, że są twierdzeniami prawdziwymi. Realista naukowy jest zatem przekonany, że nie istnieją idealnie akceptowalne twierdzenia naukowe, które byłyby fałszywe. Uważa również, że wszystkie metafory naukowe zostaną ostatecznie wyeliminowane. Jak pisał Quine:

Błędem jest myślenie [...] o uzusie językowym jako literalnym w jego głównej części i metaforycznym na peryferiach (*trimmings*). Metafora lub coś do niej zbliżonego rządzi rozwojem języka i naszym jego nabywaniem. Kognitywny dyskurs w jego czystej literalności jest raczej w większości wypadków oczyszczeniem [z metaforycznej treści — P.W.],



charakterystycznym dla starannie wypracowanych centralnych fragmentów nauki. Stanowi otwartą przestrzeń w tropikalnej dżungli stworzoną przez usuwanie przenośni (Quine 1978/1981, s. 188–189).

W związku z istnieniem zewnętrznych norm akceptacji należy podkreślić, że — w odróżnieniu od realistów — dla fikcjonalisty idealna akceptacja twierdzeń dyskursu D jest kompatybilna z ich fałszywością (Rosen 2005, s. 16).

W związku z powyższym fikcjonalista musi przyjąć rozróżnienie pomiędzy *pełną akceptacją* twierdzeń dyskursu D a przekonaniem o ich prawdziwości. O podmiocie *S* powiemy, że w pełni akceptuje twierdzenie *p* dyskursu D, gdy akceptacja twierdzenia *p* jest idealną akceptacją relatywnie do wewnętrznych norm akceptacji w ramach dyskursu D, z których żadna nie nakłada na twierdzenie *p* warunku bycia prawdziwym bądź warunku bycia uznawanym za prawdziwy. Z kolei o podmiocie *S* powiemy, że jest przekonany, że *q* (gdzie „*q*” jest twierdzeniem dyskursu D\*) wtw, gdy akceptacja twierdzenia *q* jest idealną akceptacją relatywnie do wewnętrznych norm akceptacji w ramach dyskursu D\*, wśród których znajduje się norma nakładająca na *q* warunek bycia prawdziwym bądź warunek bycia uznawanym za prawdziwy (por. Rosen 2005, s. 16–17). Przykładowo, zgodnie z fikcjonalizmem van Fraassena twierdzenia o nieobserwowalnych obiektach są idealnie akceptowalne, gdy są konsekwencjami empirycznie adekwatnej teorii. Teoria jest empirycznie adekwatna, gdy wszystkie dotyczące obserwowalnych bytów twierdzenia, które wynikają z tej teorii, są prawdziwe (zob. van Fraassen 1980, s. 12). Z kolei zgodnie z realizmem naukowym twierdzenia o nieobserwowalnych obiektach są idealnie akceptowalne, gdy istnieją obiekty fundujące prawdziwość tych twierdzeń.

Niektórzy (np. Kalderon 2005b, s. 119–129; Rosen 2005, s. 17; Yablo 2001/2010, s. 179–183, 186–187) wskazują na kolejne rozróżnie-

nie, do którego przyjęcia fikcjonalista jest zobowiązany. Twierdzą, że po stronie aktów mowy zwykle wyrazem akceptacji twierdzenia  $p$  polegającej na przekonaniu o jego prawdziwości jest asercja twierdzenia  $p$ . Z uwagi na to, że fikcjonalista akceptuje twierdzenie  $p$ , nie będąc przekonany o jego prawdziwości, wyrazem jego akceptacji powinna być *quasi*-asercja twierdzenia  $p$ .

Większość koncepcji asercji faktycznie wiąże ją z wyrazem przekonania o prawdziwości stwierdzanego twierdzenia. Zgodnie z koncepcją *asercji jako wyrazu postawy*, wypowiadając  $p$ , podmiot  $S$  stwierdza, że  $p$  wtw, gdy  $S$  jest przekonany, że  $p$ , i chce sprawić, by jego słuchacze również byli przekonani, że  $p$  (zob. MacFarlane 2011, s. 81). Zatem warunkiem dokonania asercji twierdzenia  $p$  jest przekonanie o jego prawdziwości.

Podobnie jest w ramach koncepcji *asercji jako aktywności rządzonej pewnymi regułami*. Asercja jest tutaj aktem mowy, którym rządzi reguła wiedzy, zgodnie z którą  $S$  może stwierdzić  $p$  tylko wtedy, gdy  $S$  wie, że  $p$ . Alternatywnie może to być reguła racjonalnego przekonania, zgodnie z którą  $S$  może stwierdzić  $p$  tylko wtedy, gdy przekonanie, że  $p$ , jest racjonalne (zob. MacFarlane 2011, s. 85). Oba te podejścia przekreślają możliwość stwierdzenia takiego twierdzenia, o którym podmiot  $S$  jest przekonany, że jest ono fałszywe, lub w stosunku do którego zachowuje postawę agnostyczną. Gdy  $S$  uważa, że  $p$  jest fałszywe, bądź jest agnostykiem w sprawie jego wartości logicznej, nie może zarazem wiedzieć, że  $p$ , o ile tylko wiedza, że  $p$ , zakłada przekonanie o prawdziwości twierdzenia  $p$ . Przyjmując, że  $S$  jest racjonalnym podmiotem i posiada racje na rzecz przekonania, że  $p$  jest fałszywe, bądź posiada racje na rzecz zawieszenia sądu w sprawie wartości logicznej twierdzenia  $p$ ,  $S$  nie może być zarazem racjonalnie przekonany o tym, że  $p$ .

Nie inaczej sprawy mają się w przypadku koncepcji *asercji jako zobowiązania*. W jej ramach stwierdzenie przez podmiot  $S$ , że  $p$ , jest

zaciągnięciem przez niego zobowiązań co do prawdziwości twierdzenia *p*. Zaciągnięcie zobowiązania co do prawdziwości twierdzenia *p* ma polegać na gotowości *S*-a do szczerzej obrony twierdzenia *p*, gdy staje się ono przedmiotem kontrargumentacji oraz na gotowości *S*-a do zezwolenia innym na powoływanie się na jego autorytet, gdy stwierdzają oni, że *p* (zob. MacFarlane 2011, s. 90–91, 94). Podejście to wyklucza możliwość dokonania asercji twierdzenia, o którym *S* jest przekonany, iż jest ono fałszywe, bądź w sprawie którego *S* zawiesza sąd. Jest tak dlatego, że co do żadnego z takich twierdzeń *S* nie zaciągałby zobowiązań w powyższym sensie.

Wyjątkiem w tym względzie jest Robert Stalnaker „koncepcja”<sup>17</sup> asercji jako propozycji dodania twierdzenia do zbioru sądów, które uczestnicy konwersacji przyjmują na potrzeby jej prowadzenia. Każdy z uczestników konwersacji presuponuje pewien zbiór sądów, gdzie „sąd jest presuponowany, gdy mówiący ma dyspozycję do działania, jak gdyby zakładał lub był przekonany, że sąd ten jest prawdziwy i jak gdyby zakładał lub był przekonany, że jego słuchacze również zakładają lub są przekonani, że jest on prawdziwy. Presupozycje są tym, co jest brane przez mówiącego za *wspólny grunt* uczestników konwersacji; tym, co jest traktowane jako ich *wspólna* lub *wzajemna* wiedza” (Stalnaker 1978/1999, s. 84). Każda konwersacja, w której w ogóle może nastąpić wymiana informacji pomiędzy użytkownikami języka, dokonuje się w kontekście, w którym wiedza milcząco zakładana przez uczestników konwersacji jest taka sama bądź różna ze względu na informacje, które nie mają wpływu na kwestie podnoszone w ramach konwersacji. W przeciwnym razie łatwo o brak zrozumienia pomiędzy uczestnikami konwersacji (zob. Stalnaker 1978/1999, s. 85).

---

<sup>17</sup>Wyrażenie „koncepcja” pojawia się w cudzysłowie, gdyż sam Stalnaker uważa, że w swojej analizie ogranicza się jedynie do wskazania wpływu, jaki asercja ma na kontekst, w którym prowadzi się konwersację. Nie podejmuje się jednak próby jej zdefiniowania. Co więcej, argumentuje, że inne akty, np. przyjmowanie założeń, również mają podobne do asercji wpływy na kontekst, w jakim odbywa się konwersacja (Stalnaker 1978/1999, s. 87).

W sytuacjach, gdy wiedza milcząco zakładana przez uczestników danej konwersacji różni się w stopniu znaczącym, łatwo o nieporozumienia zarówno co do tego, w jakich znaczeniach używa się poszczególnych wyrażeń, jak i co do tego, co w ogóle jest przedmiotem dyskusji. Przykładowo, gdy uczestnicy dyskusji o idealizmie przyjmują na tyle odmienne presupozycje, że jeden z nich przez „idealizm” rozumie podejmowanie próby wcielania w praktykę wzniosłych idei reformowania społeczeństwa, których nie da się urzeczywistnić, a drugi rozumie przez „idealizm” stanowisko filozoficzne kwestionujące istnienie świata materialnego, to bardzo łatwo o błędną interpretację wypowiedzi uczestników takiej konwersacji. Wystarczy wziąć tezę, że idealizm jest niewiarygodny, ponieważ kłóci się ze zdroworozsądkowym podejściem do rzeczywistości. W ustach drugiego z nich twierdzenie to wyraża tezę, że negacja istnienia przedmiotów materialnych nie współgra ze zdroworozsądkowym przekonaniem o istnieniu tych przedmiotów w rzeczywistości. Natomiast pierwszy z nich odczyta to twierdzenie jako trywialnie głoszące, że próba zrealizowana nierealizowalnych idei reformowania społeczeństwa nie jest zdroworozsądkowym podejściem do rzeczywistości. Pomędzy oboma rozmówcami może dochodzić do werbalnej zgody, ale faktycznie nie doszło nawet do ustalenia poglądów poszczególnych stron w kwestii filozoficznie i potocznie rozumianego idealizmu.

Zbiór presuponowanych sądów wyznacza zbiór możliwości, które uczestnicy konwersacji uznają — przynajmniej z punktu widzenia celów komunikacji — za kandydatów na bycie aktualnymi. Zbiór ten nazywany jest zbiorem kontekstowym (*context set*). Na gruncie takich założeń na temat komunikacji „stwierdzić coś to ograniczyć zbiór kontekstowy w pewien szczególny sposób, pod warunkiem że nie ma obiekcji ze strony innych uczestników konwersacji. Szczególny sposób, na jaki zbiór kontekstowy jest ograniczony, to taki, że eliminowane są wszystkie sytuacje niekompatybilne z tym, co zostało już

powiedziane. Innymi słowy, istotnym skutkiem asercji jest zmiana presupozycji uczestników konwersacji poprzez dodanie treści tego, co stwierdzone, do tego, co presuponowane. Skutku tego można uniknąć jedynie, gdy stwierdzenie zostanie odrzucone” (Stalnaker 1978/1999, s. 86). Asercja twierdzenia  $p$  jest zatem propozycją ograniczenia zbioru kontekstowego do tych możliwości, które są spójne z twierdzeniem  $p$ . Jest to propozycja, która pojawia się w trakcie konwersacji dokonującej się na gruncie pewnego zbioru presupozycji podzielanych przez jej uczestników.

Czy warunkiem stwierdzenia twierdzenia  $p$  przez podmiot  $S$  jest to, że podmiot  $S$  jest przekonany o prawdziwości tego twierdzenia? Wydaje się, iż odpowiedź na to pytanie jest przecząca, gdyż Stalnaker dodaje, że „w zamierzonym rozumieniu presuponowane sądy nie muszą stanowić wspólnej lub wzajemnej wiedzy, [a — P.W.] mówiący nie musi w nie wierzyć. Może presuponować dowolny sąd, który uważa za wygodny do założenia z uwagi na cele konwersacji, pod warunkiem że gotowy jest założyć, że jego słuchacze przyjmą go razem z nim” (Stalnaker 1978/1999, s. 84). Jeśli zatem podmiot  $S$  stwierdzi  $p$  w kontekście dyskusji opierającej się na presupozycjach przyjętych *jedynie* z uwagi na ich użyteczność, to ograniczenie zbioru kontekstowego wyznaczonego przez takie presupozycje wcale nie musi być równoznaczne z byciem przekonanym o prawdziwości  $p$ . Można bowiem dojść do takiego ograniczenia zbioru możliwości, z których żadna nie jest traktowana przez podmiot  $S$  jako kandydatka do bycia aktualną. Jeśli taka interpretacja stanowiska Stalnakera jest poprawna, to również fikcjonałści mogą dokonywać asercji przy Stalnakerowskim rozumieniu tego pojęcia.

Koncepcja ta stanowi jednak wyjątek na tle tych, których zwolennicy przyjmują za truizm to, że asercja twierdzenia jest wyrazem przekonania o jego prawdziwości. Z tymi koncepcjami asercji fikcjonałista chce skonstrastować swoje stanowisko w sprawie wyrażania

akceptacji twierdzeń dyskursu D. Akt ten nazywany jest przez fikcjonalistów *quasi*-asercją. Już sama sugeruje, że pomiędzy nim a asercją zachodzi pewien związek.

Jaki to związek? Choć fikcjonalista nie przyjmuje tezy, że twierdzenia dyskursu D są prawdziwe, to jest on zobowiązany do akceptacji stanowiska, że pewne twierdzenia tego dyskursu legitymują się pewnymi zaletami niezależnymi od prawdziwości. Jeśli — jak uczyniłem to w pierwszym rozdziale — przez dyskurs rozumiemy po prostu język, którego zdaniom przyporządkowane są znaczenia, to ściśle rzecz biorąc dyskursowi nie możemy przypisać zalet takich, jak bycie empirycznie adekwatnym, bycie logicznie spójnym itd. Własności takie możemy przypisać jakiejś teorii wyrażonej w tym dyskursie.

Niech  $V$  będzie zbiorem zalet, którymi legitymuje się teoria  $T$  wyrażona w dyskursie D<sup>18</sup>. Wówczas nasuwają się analogie pomiędzy *quasi*-asercją a asercją w rozumieniu trzech pierwszych koncepcji asercji, o których wspomniałem wyżej. W pierwszym wypadku, wypowiadając  $p$ , podmiot  $S$  *quasi*-stwierdza, że  $p$  wtw, gdy  $S$  jest przekonany, że  $p$  jest *poprawnym* twierdzeniem teorii  $T$  legitymującej się zaletami należącymi do zbioru  $V$  i chce sprawić, by jego słuchacze również byli przekonani, że  $p$  jest *poprawnym* twierdzeniem tej teorii. W drugim wypadku  $S$  może *quasi*-stwierdzić  $p$  tylko wtedy, gdy  $S$  wie, że  $p$  jest *poprawnym* twierdzeniem teorii  $T$ , bądź  $S$  może *quasi*-stwierdzić  $p$  tylko wtedy, gdy jest racjonalnie przekonany o tym, że  $p$  jest *poprawnym* twierdzeniem teorii  $T$ . W trzecim wypadku *quasi*-stwierdzenie przez podmiot  $S$ , że  $p$ , jest zaciągnięciem przez niego zobowiązań wobec poprawności twierdzenia  $p$  będącego twierdzeniem teorii  $T$ .  $S$  zaciąga zobowiązania wobec poprawności twierdzenia  $p$ , gdy  $S$  jest gotowy do obrony tezy, że twierdzenie  $p$  jest *poprawnym* twierdzeniem

---

<sup>18</sup>W dalszej części będę skrótowo pisał w taki sposób, jak gdyby dyskurs, a nie teoria wyrażona w ramach pewnego dyskursu, legitymował się pewnymi zaletami. Zawsze będę miał wówczas na myśli, że teoria wyrażona w tym dyskursie legitymuje się pewnymi zaletami.

teorii  $T$ , gdy jego poprawność staje się przedmiotem kontrargumentacji i  $S$  zezwala innym na powoływanie się na jego autorytet, gdy przypisują oni twierdzeniu  $p$  własność bycia poprawnym twierdzeniem teorii  $T$ <sup>19</sup>.

We wszystkich powyższych sformułowaniach pojawiło się pojęcie *poprawności* twierdzeń dyskursu  $D$ . Wyjaśnienie tego, kiedy można uznać twierdzenie dyskursu  $D$  za poprawne, jest źródłem rozbieżności pomiędzy *metafikcjonalizmem* a *fikcjonalizmem przedmiotowym*.

## 2.1 Metafikcjonalizm

Motto metafikcjonalisty głosi, że w przypadku ontologicznie kontrowersyjnego dyskursu  $D$ , gdy jego użytkownik wyraża akceptację pewnego twierdzenia pierwszego rzędu  $p$ , robi to dlatego, że jest przekonany, że w rzeczywistości akceptuje twierdzenie „Zgodnie z fikcją  $f$   $p$ ”. Innymi słowy, twierdzenie  $p$ , które z pozoru wygląda na twierdzenie pierwszego rzędu, jest jedynie *eliptycznie* wyrażonym twierdzeniem, w którym mowa o relacji pomiędzy treścią fikcji  $f$  a treścią twierdzenia  $p$ . Wartością zmiennej  $f$  ma tutaj być teoria<sup>20</sup>  $T$  wyrażona w dyskursie  $D$ , której metafikcjonalista przypisuje zalety należące do zbioru  $V$ . Zatem,

twierdzenie  $p$  dyskursu  $D$  jest poprawne wtw, gdy  $p$  jest prawdziwe zgodnie z fikcją  $f$ .

---

<sup>19</sup>Z pewnością *quasi*-asercja nie jest nieszczerą asercją. Za B. Williamsem (2002, s. 73–74) ostatnią można scharakteryzować w następujący sposób: podmiot  $S$  nieszczerze stwierdza  $p$ , gdy  $S$  udaje, że żywi przekonanie, że  $p$  jest prawdziwe i jego celem jest wywołanie w słuchaczu przekonania, że  $S$  żywi przekonanie, że  $p$  jest prawdziwe. Fikcjonalista nie ma jednak na celu zwodzenia słuchaczy w sprawie jego stosunku do prawdziwości twierdzeń dyskursu  $D$ . Wręcz przeciwnie, wprost deklaruje, że twierdzenia dyskursu  $D$  są fałszywe, lub że zachowuje postawę agnostycyzmu wobec nich, a jednocześnie uważa, że legitymują się one pewnymi zaletami niezależnymi od ich prawdziwości.

<sup>20</sup>W rozprawie, mówiąc o teoriach poddanych fikcjonalistycznej interpretacji, będę nazywał je również hipotezami, opowieściami bądź fikcjami.

Z uwagi na to, że — w ramach metafikcjonalizmu — twierdzenie  $p$  jest po prostu eliptycznie wyrażonym twierdzeniem „Zgodnie z  $f$   $p$ ” oraz że metafikcjonalista przyjmuje literalną prawdziwość ostatniego, *quasi*-asercję można zdefiniować w kategoriach asercji:

Podmiot  $S$  *quasi*-stwierdza  $p$  wtw, gdy  $S$  stwierdza, że zgodnie z fikcją  $f$   $p$  (zob. Yablo 2001/2010, §4).

Poprzedzenie twierdzeń dyskursu  $D$  *funktozem opowieści* „Zgodnie z fikcją  $f$ ” ma umożliwiać akceptację kontrowersyjnych twierdzeń tego dyskursu bez zaciągania niepożądanych zobowiązań ontologicznych. Antyrealistyczny charakter metafikcjonalizmu ma bazować na relacji zachodzącej pomiędzy wartością logiczną twierdzenia o formie „Zgodnie z fikcją  $f$   $p$ ” a wartością logiczną twierdzenia „ $p$ ”. W powszechnym mniemaniu twierdzenie „Zgodnie z fikcją  $f$   $p$ ” można uznać za prawdziwe, nawet gdy — wzorem metafikcjonalistów — samo  $p$  uznaje się za fałszywe bądź pozostaje się agnostykiem w sprawie jego wartości logicznej. Twierdzenia wyrażone przy użyciu funktora opowieści mówią bowiem o relacji treści fikcji  $f$  do treści twierdzeń poprzedzonych tym funktorem, a samo mówienie o takiej relacji nie wymaga uznawania twierdzeń fikcji  $f$  za prawdziwe. Dzięki temu, jeśli wyrażenia występujące w twierdzeniu  $p$  kwantyfikują po bytach klasy  $K$ , to akceptacja prawdziwości twierdzenia „Zgodnie z fikcją  $f$   $p$ ” nie wymaga uznawania istnienia bytów klasy  $K$ . Przykładowo, z faktu, iż twierdzenia „Zgodnie z *Monadologią* Leibniza istnieją wzajemnie odizolowane monady” lub „Zgodnie z mechaniką Newtona środki ciężkości układu słonecznego porusza się z absolutną stałą prędkością lub pozostaje w absolutnym spoczynku w absolutnej przestrzeni” są prawdziwe, nie wynika, że — odpowiednio — istnieją monady, lub że istnieje absolutna przestrzeń i absolutne parametry układów fizycznych. W ramach metafikcjonalistycznego stanowiska unika się zatem zobowiązań ontologicznych co do istnienia bytów klasy  $K$  dzięki



uznaniu, iż w kontekstach, w których korzysta się z teorii, której prawdziwości nie uznaje się, nie tyle akceptuje się twierdzenia tej teorii w jakimś osłabionym sensie akceptacji, co raczej akceptuje się twierdzenia o takiej treści, że nie mają one ontologicznie niepokojących konsekwencji.

Warto podkreślić, że funktor opowieści, o którym mówią metafikcjoniści, jest rozumiany szerzej niż wówczas, gdy używa się tego zwrotu w zastosowaniu do takich fikcji jak opowieści C. Doyle’a o Sherlocku Holmesie. W niniejszym kontekście jako fikcyjne opowieści traktuje się również hipotezy metafizyczne, koncepcje filozoficzne, a nawet teorie naukowe.

Nie ma również wśród metafikcjonalistów jednolitego poglądu na to, w jaki sposób interpretować funktor opowieści „Zgodnie z fikcją  $f$ ”. Kwestię możliwych interpretacji tego funktora pragnę jednak odłożyć do §4.3.

## 2.2 Fikcjonalizm przedmiotowy

Inaczej na sprawę unikania kontrowersyjnych zobowiązań ontologicznych zapatruje się zwolennik fikcjonalizmu przedmiotowego. Według niego podczas ontologicznie niezobowiązującej akceptacji treści twierdzeń dyskursu  $D$  nie ulegają zmianie i wciąż są wyznaczone przez ich warunki prawdziwości. Twierdzenia te albo są przez niego uznawane za literalnie fałszywe, albo zawiesza on sąd w sprawie ich wartości logicznej. Mimo przekonania o nieistnieniu bytów klasy  $K$  bądź zawieszenia sądu w sprawie ich istnienia można akceptować twierdzenia dyskursu  $D$ , gdyż ich akceptacja polega na zajęciu odpowiedniej postawy wobec tych twierdzeń, mianowicie polega na uznaniu ich za prawdziwe w ramach tzw. gry w udawanie. Zatem,

podmiot  $S$  *quasi*-stwierdza  $p$  wtw, gdy podmiot  $S$  akceptuje  $p$  jako prawdziwe w ramach odpowiedniej gry w udawanie  $G$  (innymi słowy, jako *fikcyjnie* prawdziwe).

Na czym polega akceptacja pewnego twierdzenia jako prawdziwego w ramach gry w udawanie? Gry w udawanie, podobnie jak wszystkie inne gry, są aktywnościami rządzonymi pewnymi regułami. Wyróżnia je to, że „są jednym z rodzajów aktywności wyobrażania sobie — są działaniami (*exercises of*) wyobraźni obejmującymi *rekwizyty*” (Walton 1990, s. 12). Zatem, gdy mówimy o akceptacji twierdzenia *p* jako prawdziwego w ramach odpowiedniej gry w udawanie *G*, akceptacja ta musi wiązać się z wyobrażaniem sobie czegoś w związku z pewnymi rekwizytami. Ponadto zakłada się, że gry w udawanie są aktywnościami, w których może uczestniczyć wiele osób. W tym kontekście nasuwają się dwa zasadnicze pytania: „Co wyobraża sobie podmiot *S*, gdy akceptuje pewne twierdzenie jako prawdziwe w ramach pewnej gry w udawanie?” oraz „Czym są rekwizyty i w jaki sposób wpływają one na to, co wyobraża się sobie w ramach danej gry w udawanie?”.

Założmy, że pojęcie wyobrażania sobie jest wystarczająco zrozumiałe. Wówczas w odpowiedzi na pierwsze z postawionych pytań można wysunąć następującą propozycję: podmiot *S*, akceptując twierdzenie *p* dyskursu *D* jako prawdziwe w ramach gry w udawanie *G*, wyobraża sobie, że zachodzą okoliczności, w których *p* jest prawdziwe. To, w jakich okolicznościach twierdzenie *p* jest prawdziwe, jest zależne od warunków prawdziwości tego twierdzenia. W przypadku twierdzenia z wyrażeniami referencjalnymi, mającymi na celu odnoszenie się do bytów klasy *K*, podmiot *S*, akceptując takie twierdzenie jako fikcyjnie prawdziwe, będzie wyobrażał sobie, że istnieją odniesienia tych wyrażen referencjalnych i posiadają takie własności lub pozostają do siebie w takich relacjach, jak głosi akceptowane twierdzenie. W przypadku twierdzenia, w którym występują kwantyfikatory mające na celu przebieganie po bytach klasy *K*, podmiot *S*, akceptując takie twierdzenie jako fikcjonalnie prawdziwe, będzie wyobrażał sobie, że istnieją przedmioty stanowiące wartości zmiennych związanych występujących w tym twierdzeniu i posiadają one takie własności lub

pozostają do siebie w takich relacjach, jak głosi to twierdzenie.

Na tym nie koniec. W ramach gry w udawanie *G* przynajmniej część wyobrażeń żywionych przez jej uczestników zależna jest od tego, jakie są *rekwizyty* tej gry. Tu „rekwizyty są generatorami fikcyjnej prawdziwości — rzeczami, które na mocy ich natury lub samego ich istnienia czynią sądy fikcyjnie prawdziwymi” (Walton 1990, s. 37). Zwykle rekwizyty traktuje się jako realnie istniejące rzeczy bądź obiektywnie zachodzące okoliczności, których istnienie bądź zachodzenie stanowi konieczny i wystarczający warunek fikcyjnej prawdziwości tych twierdzeń. To, czy istnienie danej rzeczy bądź zachodzenie pewnych okoliczności stanowi warunek fikcyjnej prawdziwości, zależne jest od tzw. *reguł generowania* (zob. Walton 1990, s. 38). Mają one wiązać fikcyjną prawdziwość twierdzeń z istnieniem pewnych rzeczy lub zachodzeniem pewnych okoliczności. W praktyce ich celem jest uzależnianie aktywności akceptowania pewnych twierdzeń jako prawdziwych w ramach gry *G* od okoliczności zachodzących niezależnie od założeń przyjmowanych w ramach tej gry. W związku z tym mają one postać warunkowych preskrypcji: „*Jeśli* pewne okoliczności zachodzą, należy wyobrazić sobie pewne rzeczy” (Walton 1990, s. 40).

Aby rzucić nieco więcej światła na dotychczasową dość abstrakcyjną charakterystykę gier w udawanie, pozwolę sobie przytoczyć kilka przykładów tych gier. Z grami w udawanie mamy styczność już od dzieciństwa i angażujemy się w nie również jako dorośli, aczkolwiek w ostatnim wypadku przyjmują one znacznie bardziej wysublimowane formy. Przykładem gry w udawanie, w którą angażują się dzieci, może być gra w Indian i kowbojów. Uczestnicy takiej gry dzielą się na Indian i kowbojów. Podział dokonuje się na mocy pewnej umowy pomiędzy uczestnikami tej gry, dzięki której część z nich uznawana jest za Indian, a część — za kowbojów. Angażując się w taką grę, dzieci mogą posługiwać się takimi przedmiotami jak plastikowe łuki i strzały zakończone gumową przyssawką, plastikowe pistolety strze-

lające plastikowymi kulkami itd. W trakcie gry udają, że mierzą do siebie autentycznymi łukami bądź pistoletami, a wystrzeliwane z nich strzały bądź kulki mogą śmiertelnie ranić. Przyjąwszy takie założenia co do przedmiotów, których używają uczestnicy tej gry w udawanie, szereg dalszych twierdzeń dotyczących łuków i pistoletów może stać się fikcyjnie prawdziwymi na mocy tego, w jakich okolicznościach przedmioty te są używane oraz jakie własności posiadają poszczególne z nich. Gdy plastikowa strzała wystrzelona z plastikowego łuku trafi jakiegokolwiek uczestnika tej gry w nogę, powinien on symulować zranienie. Wówczas twierdzenie, że ktoś został zraniony w nogę strzałą wystrzeloną z tego a tego łuku, będzie fikcyjnie prawdziwe w ramach tej gry w udawanie. Gdy natomiast strzała trafi któregokolwiek z uczestników w klatkę piersiową, powinien on co najmniej symulować ciężkie zranienie, jeśli nie agonię. Wówczas twierdzenie, że ktoś został ciężko zraniony lub jest w stanie agonalnym z powodu trafienia strzałą wystrzeloną z tego a tego łuku, będzie fikcyjnie prawdziwe w ramach tej gry w udawanie. Zarówno związek pomiędzy własnościami tych przedmiotów, jak i obiektywnie zachodzącymi okolicznościami, w których przedmioty takie się znajdują, a fikcyjną prawdziwością twierdzeń jest ustalany za pomocą wprowadzenia odpowiednich reguł generowania. Niektóre z reguł generowania wiążące zachodzenie tego typu sytuacji z tym, jakie twierdzenia uznaje się za fikcyjnie prawdziwe, mogą być dla uczestników tej gry na tyle oczywiste, że nie muszą uzgadniać między sobą ich obowiązywania. Po prostu milcząco zakładają, że są one na tyle jasne, że każdy z uczestników gry w udawanie będzie się do nich stosował. Inne z reguł generowania mogą wymagać uzgodnienia *explicite*. Zarówno przedmioty wykorzystywane w grze, jak i realnie zachodzące okoliczności, w które uwikłane są te przedmioty, stanowią rekwizyty tej gry w udawanie.

W podobne gry angażują się czytelnicy powieści kryminalnych, a przynajmniej ci spośród nich, którzy pełni emocji oddają się ich

lekturze i z zapartym tchem śledzą losy bohaterów rozgrywające się w jednym z realnie istniejących miast. Ludzie ci uczestniczą w grach w udawanie, w których wyobrażają sobie, że czytana przez nich powieść jest opisem zdarzeń, które faktycznie miały miejsce w danym mieście. Jeśli czytelnikowi znane jest miasto, w którym rozgrywa się akcja powieści, może sobie na podstawie znajomości jego topografii w danym okresie historycznym wyobrażać wiele rzeczy, o których autor nawet nie wspomina, ale które podsuwa czytelnikowi, odwołując się w powieści do miasta i lokalizując jej akcję w pewnym kontekście historycznym. Czytelnik zaangażowany w grę w udawanie wywołaną przez lekturę opowieści, wyobrażając sobie pewne rzeczy dzięki znajomości topografii miasta wskazanego przez autora, musi zaakceptować (świadomie bądź nie) pewne reguły generowania, które pozwalają mu na przejście od tego, jakim miasto było w danym okresie historycznym, do tego, co należy sobie wyobrazić o miejscach, w których rozgrywa się akcja czytanej przez niego powieści. Niektóre spośród faktów związanych z topografią miasta w danym okresie historycznym można zatem uznać za rekwizyty tej gry w udawanie.

Jak pokazują powyższe przykłady, fikcyjna prawdziwość niektórych twierdzeń jest częściowo zależna od istnienia i własności rekwizytów. W związku z tym dzięki znajomości reguł generowania i wiedzy na temat tego, które z twierdzeń uznano za fikcyjnie prawdziwe, można odtworzyć, na podstawie jakich obiektywnie zachodzących okoliczności akceptowano te twierdzenia. Dla dowolnego twierdzenia  $p$ , jeśli  $p$  jest prawdziwe w ramach pewnej gry w udawanie i istnieją reguły generowania, łączące fikcyjną prawdziwość  $p$  z zachodzeniem okoliczności  $k$ , to treść literalnie prawdziwego twierdzenia  $q$ , w którym stwierdza się zachodzenie okoliczności  $k$ , stanowi *realną treść* twierdzenia  $p$  (por. Yablo 2001/2010, s. 182–183).

Nie wszystkie fikcyjnie prawdziwe twierdzenia muszą mieć realną treść. Niektóre z nich mogą dotyczyć wyłącznie przedmiotów, których

istnienie przyjmuje się jedynie w ramach gry w udawanie, i przypisywać im takie własności, że nie istnieją reguły generowania, które pozwoliłyby na przyporządkowanie tym twierdzeniom realnej treści. Załóżmy, że jesteśmy agnostykami w sprawie istnienia bytów matematycznych oraz że akceptujemy istnienie obiektów konkretnych i zarazem — jak Steven Yablo w (2005) — uważamy, że język liczb jest nieeliminowalny z rozważań teoretycznych<sup>21</sup>. Będąc fikcjonalistami przedmiotowymi, akceptowalibyśmy wówczas twierdzenia czystej matematyki jako prawdziwe jedynie w pewnej grze w udawanie. Same twierdzenia czystej matematyki pozbawione byłyby realnej treści, gdyż dopiero w rezultacie zastosowań matematyki do opisu pewnych aspektów pozamatematycznej rzeczywistości w grze w udawanie mogą pojawić się twierdzenia, których prawdziwość zależna jest od tego, jakimi są obiektywnie zachodzące pozamatematyczne okoliczności. O okolicznościach powiemy, że są pozamatematyczne, gdy żaden z bytów, składających się na te okoliczności, nie jest bytem należącym do uniwersum czystej matematyki. By zależności pomiędzy fikcyjną prawdziwością zmatematyzowanych twierdzeń a pozamatematycznymi okolicznościami były systematyczne, fikcjonalista przedmiotowy powinien wprowadzić reguły generowania uzależniające fikcyjną prawdziwość zmatematyzowanych twierdzeń od pewnych właściwości pozamatematycznych okoliczności. Twierdzenia, które będą literalnie prawdziwe o odpowiednich aspektach tych pozamatematycznych okoliczności, można uznać wówczas za realną treść zmatematyzowanych twierdzeń, dla których zachodzenie tych okoliczności jest koniecznym i wystarczającym warunkiem ich fikcyjnej prawdziwości. Jeśli na pozamatematyczne fakty składają się wyłącznie fakty dotyczące konkretów, to realną treścią zmatematyzowanych twierdzeń będą prawdziwe twierdzenia, w których występują wyłącznie odwołania do konkretów. Przykładowo,

---

<sup>21</sup>W sprawie przeglądu i krytycznej analizy stanowisk fikcjonalistycznych co do istnienia bytów matematycznych, zob. (Wójtowicz 2003), w szczególności rozdziały 2. i 3.

gdy zmatematyzowane twierdzenia dotyczące pozamatematycznych okoliczności mają postać „liczba  $F$ -ów  $= n$ ”, reguły generowania mogą przyjąć następującą postać: „jeśli istnieje <sub>$n$</sub>   $x$ ,  $Fx$ , to należy wyobrazić sobie, że liczba  $F$ -ów  $= n$ ” (por. Yablo 2006, s. 182–183, Yablo 2005, s. 104), gdzie kwantyfikacja numeryczna „istnieje <sub>$n$</sub>   $x$ , ...” jest definiowana w następujący sposób:

Istnieje <sub>$n$</sub>   $x$ , że  $Fx =_{df}$ . Istnieje  $x_1, \dots, x_n$ ,  $Fx_1$  i ... i  $Fx_n$  i  $x_1 \neq x_2$  i ... i  $x_1 \neq x_n$  i  $x_m \neq x_n$ ,

gdzie ( $1 < m < n$ ). Jeśli tak zdefiniowane twierdzenia z kwantyfikacją numeryczną mają wyrażać realną treść twierdzeń matematyki stosowanej o postaci „liczba  $F$ -ów  $= n$ ”, to należy przyjąć założenie, że zmienne związane występujące w definiensie tej definicji przebiegają po obiektach konkretnych.

Dopiero wraz z przyjęciem fikcyjnej prawdziwości zmatematyzowanych twierdzeń o konkretach twierdzenia czystej matematyki mogą występować jako kroki w rozumowaniach, których wnioskami są zmatematyzowane twierdzenia o realnej treści wyrażanej przez literalnie prawdziwe twierdzenia o pozamatematycznych faktach. Bardzo prosty przykład takiego rozumowania przedstawia się następująco:

P1. \*liczba jabłek  $= 2^{*22}$ ,

P2. \*liczba gruszek  $= 2^*$ ,

P3.  $*2 + 2 = 4^*$ ,

W. \*liczba jabłek + liczba gruszek  $= 4^*$ .

Na mocy reguł generowania można wskazać okoliczności, które stanowiły konieczny i wystarczający warunek fikcyjnej prawdziwości

---

<sup>22</sup>Twierdzenie występujące w wolnym miejscu w  $* \dots *$  jest twierdzeniem akceptowanym jako prawdziwe w ramach odpowiedniej gry w udawanie.

twierdzenia (W). Okoliczności te w sposób prawdziwy opisuje twierdzenie, że istnieje<sub>4</sub>  $x$ ,  $x$  jest jabłkiem lub  $x$  jest gruszką, które wyraża realną treść twierdzenia (W).

W ramach samej matematycznej gry w udawanie (tj. gry, w której wyobraża się sobie istnienie bytów, na mocy których twierdzenia matematyczne są fikcyjnie prawdziwe) będą również występowały takie twierdzenia o postaci „liczba  $F$ -ów  $= n$ ”, którym nie będzie można przyporządkować realnej treści. Weźmy twierdzenie (PN), że liczba parzystych liczb pierwszych jest równa jeden. Twierdzenie to można również wyrazić następująco: istnieje takie  $x$ , że jest parzystą liczbą pierwszą i dla każdego  $y$ , jeśli  $y \neq x$ , to nieprawda, że  $y$  jest parzystą liczbą pierwszą. Z pewnością fikcjonalista w sprawie istnienia bytów matematycznych nie chciałby, by to ostatnie twierdzenie wyrażało realną treść twierdzenia, że liczba parzystych liczb pierwszych jest równa jeden. Gdyby tak było, obiektywnie musiałyby zachodzić okoliczność, że tylko jedna liczba jest liczbą pierwszą, a wówczas realnie istniałyby liczby i fikcjonalizm w sprawie istnienia bytów matematycznych byłby stanowiskiem samoobalającym się.

Fikcjonalista musi zatem uznać, że twierdzenie (PN) jest twierdzeniem tak jedynie fikcyjnie prawdziwym. Dlaczego miałoby być pozbawione realnej treści? Naturalnie nasuwająca się odpowiedź jest taka, że twierdzenie (PN) nie spełnia nakładanego wyżej warunku, aby zmienne twierdzeń wyrażających realną treść przebiegały po konkretnych, czy — ogólniej — po bytach istniejących niezależnie od tego, co zakłada się w matematycznej grze w udawanie.

Powstaje jednak pytanie, w jaki sposób mogłoby przedstawiać się ogólne kryterium odróżniania twierdzeń, którym można przyporządkować realną treść, od twierdzeń, którym treści takiej przyporządkować nie sposób? Pozostając jeszcze przy interpretacji matematyki stosowanej w ramach fikcjonalizmu przedmiotowego, w pierwszym rzędzie należy zauważyć, że byty matematyczne mogą funkcjonować w ramach



matematycznej gry w udawanie na dwa zasadniczo odmienne sposoby. Z jednej strony mogą one pełnić rolę *narzędzi pomocnych* w *reprezentowaniu* obiektywnie zachodzących okoliczności, które nie zawierają bytów matematycznych jako swoich elementów konstytutywnych. Z drugiej strony, byty matematyczne same mogą być przedmiotami, które są *reprezentowane* w ramach matematycznej gry w udawanie, a wówczas twierdzenia dotyczące tych bytów są prawdziwe jedynie na mocy własności, które należało sobie wyobrazić o tych bytach, gdy używało się ich jako narzędzi pomocnych w reprezentowaniu obiektywnie zachodzących okoliczności. Zatem twierdzenia, w których występują odniesienia do bytów matematycznych występujących jako narzędzia pomocne w reprezentowaniu bytów pozamatematycznych, będą twierdzeniami posiadającymi realną treść. Z kolei twierdzenia, w których występują odniesienia do bytów matematycznych występujących jako reprezentowane byty, będą twierdzeniami pozbawionymi realnej treści.

S. Yablo rozróżnia — odpowiednio do tego, czy mamy do czynienia z pierwszą funkcją bytów matematycznych, czy z drugą — dwa typy gier w udawanie: „[Gra w udawanie — P.W.]  $G$  jest *podstawowa*, jeśli akceptacja [twierdzeń — P.W.] w  $G$  jest funkcją tego, jak rzeczy naprawdę się mają [...].  $G^*$  jest *wtórna* (*parasitic*), jeśli akceptacja [twierdzeń — P.W.] w  $G^*$  jest zależna od tego, jakie rzeczy są wyobrażone w pewnej innej grze (jak mogłoby się okazać, grze  $G$ )” (Yablo 2001/2010, s. 189). W ramach terminologii przyjętej w tej rozprawie każda gra w udawanie zawiera w sobie zarówno grę bazową, jak i wtórną, o ile tylko będą akceptowane w niej twierdzenia, którym nie można przyporządkować (*via* reguły generowania) realnej treści.

Uogólniając uwagi z dwóch ostatnich paragrafów, możemy wprowadzić następujące warunki:

Byty klasy  $K$ , których istnienie zakłada się w ramach gry w udawanie  $G$ , pełnią rolę *narzędzi pomocnych* w *reprezentowaniu*

sposobów, na jakie są realnie istniejące byty klasy  $K^*$  wtw, gdy zachodzą takie okoliczności  $k$ , że byty klasy  $K^*$  są elementami konstytutywnymi tych okoliczności, byty klasy  $K^*$  istnieją niezależnie od założeń przyjmowanych w ramach gry  $G$  i istnieją twierdzenia dyskursu  $D$ , w których występują terminy odnoszące się bądź kwantyfikujące po bytach klasy  $K$ , a zachodzenie okoliczności  $k$  stanowi konieczny i wystarczający warunek fikcyjnej prawdziwości tych twierdzeń.

Byty klasy  $K$ , których istnienie zakłada się w ramach gry  $G$ , pełnią rolę *reprezentowanych* bytów wtw, gdy twierdzenia odnoszące się do bytów klasy  $K$  są prawdziwe na mocy założeń czynionych w ramach gry  $G$  i nie istnieją okoliczności  $k$  zachodzące niezależnie do założeń przyjmowanych w ramach gry  $G$ , które są koniecznymi i wystarczającymi warunkami fikcyjnej prawdziwości tych twierdzeń.

Rola bytów klasy  $K$  może zmieniać się w zależności od tego, jakie są warunki fikcyjnej prawdziwości twierdzeń odnoszących się do bytów tej klasy bądź kwantyfikujących po nich. W prezentowanym wyżej przykładzie, gdy mowa o liczbie jabłek, liczba dwa pełni rolę narzędzia pomocnego w reprezentowaniu pozamatematycznych faktów, gdy natomiast mowa o parzystej liczbie pierwszej, liczba 2 pełni rolę reprezentowanej rzeczy.

Interesującą własnością fikcyjnie prawdziwych twierdzeń, które mają na celu reprezentowanie bytów klasy  $K$ , jest to, że są one fikcyjnie prawdziwe niezależnie od okoliczności, w jakich rozgrywa się dana gra w udawanie. Może to być podstawą do wprowadzenia nowego typu reguł konstytuujących gry w udawanie, które w sposób bezwarunkowy nakazują wyobrażenie sobie pewnych rzeczy każdemu, kto tylko chce być uczestnikiem danej gry w udawanie. Nazwijmy takie reguły *regułami postulatywnymi*.

Podsumowując dotychczasowe rozważania, możemy powiedzieć, że grą w udawanie jest wszelka — rządzona regułami postulatywnymi oraz regułami generowania — aktywność wyobrażania sobie, że pewne rzeczy mają miejsce.

Mając reguły postulatywne i reguły generowania, możemy powiedzieć, że:

twierdzenie  $p$  dyskursu  $D$  jest poprawne wtw, gdy akceptacja  $p$  jako prawdziwego w ramach gry w udawanie jest nakazana przez reguły postulatywne lub reguły generowania.

Odwołanie się do reguł generowania jest konieczne wszędzie tam, gdzie przyjmuje się, że fikcyjnie prawdziwe twierdzenia teorii  $T$  są użyteczne do opisu jakiejś klasy fenomenów, których istnienie przyjmuje się na gruncie założeń logicznie niezależnych od założeń akceptowanych w ramach gry w udawanie, w której postulaty teorii  $T$  są fikcyjnie prawdziwe. Właśnie dzięki wskazaniu, w jakim sensie fikcyjna prawdziwość twierdzeń tej teorii jest zależna od prawdziwości twierdzeń wyrażających ich realną treść, można starać się wyjaśnić, w jaki sposób teoria  $T$  pozwala rzucić więcej światła na naturę bytów, których dotyczą twierdzenia wyrażające realną treść jej twierdzeń.

Gry w udawanie, które powoływane są do istnienia jedynie w celu dostarczenia użytecznych narzędzi reprezentowania pewnej dziedziny przedmiotów, różnią się pod ważnym względem od gier w udawanie, w które angażują się dzieci, czytelnicy powieści, widzowie spektaklów teatralnych itp. W grach ostatniego typu zainteresowanie rekwizytami bierze się głównie z tego, że pozwalają one na wzbogacenie treści samej gry w udawanie. Treścią gry w udawanie jest klasa wszystkich twierdzeń, które są fikcyjnie prawdziwe w ramach tej gry. Dzięki temu, że na deskach sceny teatralnej prezentowane są rekwizyty, możemy wzbogacić treść gry w udawanie poprzez uznanie za fikcyjnie prawdziwe wszystkich tych twierdzeń, które można zasadnie uznać za fikcyjnie prawdziwe, bazując na własnościach przedstawionych nam

rekwizytów. Walton w swoim artykule (1993/2005) nazywał gry w udawanie tego typu grami w udawanie *zorientowanymi na treść* i kontrastował je z grami w udawanie *zorientowanymi na rekwizyty*. Ostatnie tworzy się, aby ułatwić poznanie własności rekwizytów tych gier. Wszystkie gry w udawanie, które konstruowane są po to, by można było w ich ramach zaakceptować twierdzenia pewnych teorii, których zobowiązań ontologicznych nie chce się akceptować, a które uważa się za użyteczne, są właśnie grami *zorientowanymi na rekwizyty*. Ich podstawowym celem jest dostarczenie opisu pewnych okoliczności, których zachodzenie akceptuje się na podstawie niezależnych przesłanek; opisu, który — jak ufają fikcjonałiści — ułatwia myślenie o tych obiektywnie zachodzących okolicznościach<sup>23</sup>.

---

<sup>23</sup>W artykule (1993/2005) Walton, posługując się pojęciem gry w udawanie *zorientowanej na rekwizyty*, charakteryzuje metaforyczne użycia wyrażeń. Zgodnie z proponowanym przez niego ujęciem twierdzenie *p* jest użyte w sposób metaforyczny, gdy jest ono literalnie fałszywe, jest prawdziwe w ramach odpowiedniej gry w udawanie, dla której rekwizytami są rzeczy, o których mówi się w sposób metaforyczny. Przykładowo, w żargonie hydraulików rury nagwintowane na zewnątrz nazywa się rurami męskimi, a rury nagwintowane od wewnątrz — rurami żeńskimi. Rzecz jasna, rurom nie można w sposób literalny przypisywać cech płciowych, dlatego też — zdaniem Waltona — można im przypisać te własności, jedynie angażując się w grę w udawanie, w której zakłada się, że rury mogą posiadać cechy płciowe i w ramach której prawdziwe są twierdzenia przypisujące im takie cechy. Fikcyjna prawdziwość tych twierdzeń jest uzależniona od sposobu, na jaki nagwintowana jest dana rura. Dlatego też gra w udawanie, w którą angażują się hydraulicy, zawiera reguły generowania nakazujące przyjęcie założenia, iż dana rura jest żeńska lub męska, jeśli jest nagwintowana — odpowiednio — od wewnątrz lub na zewnątrz. Yablo w (2001/2010) akceptuje stanowisko Waltona w sprawie metaforycznego użycia wyrażeń i fikcjonalizm przedmiotowy (w scharakteryzowanym wyżej sensie) nazywa *figuralizmem*, traktując twierdzenia przyjmowane w ramach *zorientowanych na rekwizyty* gier w udawanie jako metafory. Dla Yablo powiedzenie, że liczba jabłek jest równa dwa, jest metaforycznym sposobem powiedzenia, że istnieje takie *x* i *y*, że *x* jest jabłkiem i *y* jest jabłkiem, i  $x \neq y$ , i dla każdego *z*, jeśli *z* jest jabłkiem, to  $z = x$  lub  $z = y$ . Choć faktycznie niektórzy filozofowie odrzucający zobowiązania ontologiczne pewnych koncepcji, np. istnienie światów możliwych, mają tendencję do mówienia, iż wypowiedzi dotyczące kontrowersyjnych bytów są wypowiedziami metaforycznymi, w tej rozprawie pozostawiam na boku kwestię słuszności takiego poglądu. Bez wątpienia rozważania poświęcone temu, na ile koncepcje metafor mogłyby okazać się owocne w charakteryzowaniu ontologicznie niezobowiązującej akceptacji twierdzeń egzystencjalnych, są bardzo interesujące. Wykraczają one jednak poza obszar badań w tej rozprawie.

### 3 Spór w fikcjonalistycznej rodzinie

Pomiędzy zwolennikami fikcjonalizmu przedmiotowego i metafikcjonalizmu toczą się spory o to, dla którego z tych stanowisk bilans zysków i strat wypada korzystniej. W artykule (2001/2010, s. 181) Yablo podnosi kilka zarzutów przeciwko metafikcjonalizmowi. Zgodnie z pierwszym z nich, zwanym *problemem zachowania konieczności*,

gdy twierdzenia akceptowane przez fikcjonalistów jako prawdziwe zgodnie z pewną fikcją wyrażają konieczne relacje pomiędzy obiektami, metafikcjonalista natrafia na trudność w wyjaśnieniu, w jaki sposób proponowana przez niego parafraza może zachować konieczną prawdziwość tych twierdzeń. Fikcja, zgodnie z którą twierdzenia te są prawdziwe, mogłaby nigdy nie powstać, albo mogłaby powstać w odmiennej formie. Wtedy jednak niektóre z koniecznych związków zachodzących pomiędzy obiektami mogłyby nie być wyrażane przez żadne z twierdzeń, które aktualnie są prawdziwe zgodnie z tą fikcją.

Przykładowo, niech metafikcjonalista w sprawie istnienia bytów matematycznych utrzymuje, że twierdzenia matematyczne są prawdziwe zgodnie ze standardową matematyką. Wówczas — twierdzi Yablo — metafikcjonalista ma kłopot z wyjaśnieniem, dlaczego  $2 + 3$  wydaje się z konieczności równe 5. Nie jest *a priori* oczywiste, że standardowa matematyka nie mogłaby być inna, choć wydaje się *a priori* oczywiste, że  $2 + 3$  nie mogłoby nie być równe 5. Standardowa matematyka jest stworzoną przez człowieka teorią, która mogłaby w sposób błędny ujmować relacje ilościowe pomiędzy obiektami. Dlatego nie jest oczywiste, że zgodnie ze standardową matematyką  $2 + 3 = 5$ .

Wydaje się jednak, że zwolennik metafikcjonalizmu mógłby w pełni zasadnie odpowiedzieć na ten zarzut: „*Tu quoque!*”. Tak samo jak standardowa matematyka mogłaby być taka, że twierdzenie, że  $2 + 3$

= 5, nie byłoby prawdziwe zgodnie z nią, tak też nie jest oczywiste, że reguły postulatywne matematycznej gry w udawanie nie mogłyby różnić się od obecnie przyjmowanych w ten sposób, że żadna z nich nie nakazywałaby wyobrażenia sobie istnienia bytów matematycznych, spełniających twierdzenie, że  $2 + 3 = 5$ . Podobnie jak standardowa matematyka mogłaby być odmienna od tej, jaką jest obecnie, tak też matematyczna gra w udawanie mogłaby różnić się od tej, jaką jest obecnie.

Być może intencją powyższego zarzutu było wskazanie na to, że podczas gdy metafikcjonalista uzależnia prawdziwość danego twierdzenia matematycznego od treści standardowej matematyki, to przedmiotowy fikcjonalista konstruuje byty matematyczne przede wszystkim jako narzędzia pomocne w reprezentowaniu pozamatematycznych faktów (por. Yablo 2005, s. 104–108). Pomiedzy tymi faktami mogą zachodzić pewne konieczne relacje, a zmatematyzowane reprezentacje tych relacji odzwierciedlają tylko (za pośrednictwem reguł generowania) konieczny charakter relacji zachodzących pomiędzy pozamatematycznymi bytami.

Nie jest jednak jasne, dlaczego metafikcjonalista nie mógłby powiedzieć czegoś podobnego o treści standardowej matematyki. Dlaczego nie mógłby twierdzić, że jej treść jest formułowana głównie z intencją stosowania jej jako narzędzia do reprezentowania pozamatematycznych faktów i że jeśli pomiędzy tymi faktami zachodzą konieczne relacje, to konieczności te zostaną oddane w treści standardowej matematyki? Field twierdzi, że standardowa matematyka jest rozumiana w taki sposób, że z jej twierdzeń i pozamatematycznych faktów dotyczących Układu Słonecznego wynikają np. takie zmatematyzowane twierdzenia jak twierdzenie, że liczba planet bliższych Słońcu niż Ziemia jest równa dwa (zob. Field 1989, s. 3, przyp. 3). Jeśli tak jest, to w treści standardowej matematyki muszą występować twierdzenia bądź schematy twierdzeń, nazywane przez Fielda *prawami pomo-*

stowymi, które pozwalają na skolerowanie prawdziwości twierdzeń matematycznych zgodnych ze standardową matematyką z zachodzeniem odpowiednich pozamatematycznych faktów. Stąd już tylko krok do tego, by — na wzór Yablo — twierdzić, że własności bytów, o których mówią twierdzenia prawdziwe zgodnie z treścią standardowej matematyki, są im przypisywane głównie ze względu na rolę, jaką pełnią w reprezentowaniu bytów pozamatematycznych i reprezentacje te dziedziczą konieczne związki zachodzące pomiędzy tymi pozamatematycznymi bytami. Jeśli w ogóle taki projekt ma cień szans na powodzenie w przypadku fikcjonalizmu przedmiotowego, to nie jest jasne, dlaczego miałby zakończyć się fiaskiem w przypadku jego metafikcjonalistycznego odpowiednika. Przy takiej interpretacji pierwszy z zarzutów Yablo wydaje się chybiony.

Drugi z zarzutów Yablo, nazywany *problemem fenomenologicznym*, przedstawia się następująco:

gdy stosujemy matematykę, np. mówiąc, że liczba głodujących wzrasta, to wydaje się, że twierdzenia takie przede wszystkim dotyczą ludzi i problemu głodu. Tymczasem proponowana przez metafikcjonalistę parafraza tego twierdzenia jako eliptycznie wyrażonego twierdzenia o postaci (SP) „Zgodnie ze standardową matematyką liczba głodujących wzrasta”, sprawia, że twierdzenie to wydaje się przede wszystkim dotyczyć treści fikcji.

Yablo niestety nie mówi niczego więcej o problemie fenomenologicznym, warto więc zastanowić się, gdzie znajduje się ostrze jego zarzutu. Nim jednak przejdziemy do jego oceny, najpierw odróżnijmy matematykę standardową od stosowanej. W pierwszej występują wyłącznie twierdzenia, które mają na celu charakteryzowanie własności obiektów matematycznych, w drugiej dodatkowo występują zmatematyzowane twierdzenia reprezentujące właściwości obiektów pozamatematycznych. Niezależnie od tego, czy metafikcjonalista wygłosi twierdzenie matematyki czystej, czy stosowanej, w obu wypadkach będzie chciał

uniknąć zobowiązań ontologicznych co do istnienia bytów matematycznych. Zgodnie ze swoją ulubioną praktyką każde twierdzenia standardowej matematyki poprzedzi funktorem opowieści „Zgodnie ze standardową matematyką”, a twierdzenia stosowanej matematyki — funktorem „Zgodnie ze stosowaną matematyką”. (SP) z pewnością nie jest twierdzeniem standardowej matematyki, bo w jej słowniku nawet nie występuje pojęcie ludzi. Dlatego jest ono twierdzeniem matematyki stosowanej i powinno mieć postać (SP\*) „Zgodnie ze stosowaną matematyką liczba głodujących wrasta”. Wspomniane wyżej prawa pomostowe wyznaczają prawidłowy sposób matematycznego parametryzowania zjawiska głodu w populacji ludzi, a czysto matematyczne relacje pomiędzy liczbami pozwalają ustalić, czy mamy do czynienia z nasileniem, czy osłabieniem zjawiska. W każdym razie, za sprawą praw pomostowych twierdzenie (SP\*) pośrednio dotyczy zjawiska głodu, a nie wyłącznie treści jakiejś oderwanej od tego zjawiska fikcji. Powstaje więc pytanie, co powoduje, że twierdzenie (SP\*) sprawia wrażenie twierdzenia dotyczącego treści stosowanej matematyki, a nie głodujących ludzi?

Jedno z przychodzących na myśl wyjaśnień przedstawia się następująco: wygłoszenie tego twierdzenia w kontekście dyskusji o głodujących ludziach prowadzi do pogwałcenia jednej z Grice’owskich maksym konwersacyjnych, mianowicie *maksymy ilości*<sup>24</sup>, zgodnie z którą wkład w dyskusję powinien być na tyle informatywny, na ile jest to konieczne z punktu widzenia celu wymiany informacji podczas komunikacji (zob. Grice 1989, s. 26–27). Gdy mówi się o głodujących ludziach, twierdzenie (SP) wydaje się zanadto informatywne — mówi nie tylko o liczbie

---

<sup>24</sup>Z uwagi na to, że metafikcjoniści w sprawie bytów matematycznych wprowadzają prawa pomostowe, przypadek wygłoszenia twierdzenia (SP\*) w kontekście dyskusji o głodujących ludziach nie wydaje się przykładem pogwałcenia Grice’owskiej maksymy relacji. Zgodnie z nią wkład w dyskusję powinien być relewantny, tj. dotyczyć tego, o czym w niej mowa. Na mocy istnienia odpowiednich praw pomostowych niektóre ze zmatematyzowanych twierdzeń prawdziwych (zgodnie ze stosowaną matematyką) można traktować jako twierdzenia dotyczące problemu głodujących ludzi.



głodujących ludzi, ale również *explicite* wspomina o relacji pomiędzy treścią stosowanej matematyki a treścią pewnego twierdzenia. Sprawia to, że słuchacz zaczyna doszukiwać się powodu, dla którego mówiący (o którym zakłada, że jest racjonalnym uczestnikiem dyskusji) wygłosił akurat twierdzenie (SP\*). W rezultacie prowadzi to do odwrócenia uwagi słuchacza od przedmiotu dyskusji, którym są głodujący ludzie, i zwrócenia jej ku treści stosowanej matematyki, wywołując wrażenie, że (SP\*) nie dotyczy przedmiotu dyskusji.

Choć faktycznie wygłoszenie (SP\*) w kontekście dyskusji o głodujących ludziach może prowadzić do takich skutków, to metafikcjonalista posiada środki, dzięki którym może wygłosić twierdzenie (SP\*) w taki sposób, by nie powodowało ono odwracania uwagi od tematu tej dyskusji. Jak widzieliśmy w § 2.2.1., w kontekstach, w których akceptuje się twierdzenie  $p$  dyskursu D, metafikcjonalista traktuje je jako eliptycznie wyrażone twierdzenie o postaci „Zgodnie z  $f$   $p$ ”. Dlatego też w kontekście dyskusji o głodujących ludziach celem uniknięcia wspomnianego efektu może on wygłosić twierdzenie (SP\*) w jego eliptycznej formie, która *pozornie* dotyczy wyłącznie głodujących ludzi i ich wzrastającej liczby, nie wspominając niczego o treści stosowanej matematyki. Jeśli powodem powstania wrażenia, że wygłoszenie (SP\*) nie dotyczy głodujących ludzi, było to, że (SP\*) jest zanedo informatywne, można obejść problem fenomenologiczny przez wygłoszenie tego twierdzenia w eliptycznej formie. Jeśli natomiast zarzut ten opierał się na założeniu, że teoria metafikcjonalisty nie wprowadza zależności pomiędzy prawdziwością zgodnie z fikcją a tym, jakie są obiektywnie zachodzące okoliczności, to — ponownie — zarzut ten jest chybiony.

Czy można wykazać, że nawet eliptycznie wyrażone twierdzenie (SP\*) wciąż nie jest tym, co chcielibyśmy wygłosić w kontekście dyskusji o głodujących ludziach? Trzeci zarzut, jaki można podnieść przeciwko metafikcjonalizmowi, przedstawia się następująco:

uznanie, że każde fikcyjne twierdzenie poddaje się metafikcjonalistycznej interpretacji prowadzi do zatarcia różnicy pomiędzy *mówieniem o fikcji* a *opieraniem się na fikcji* w tym, co się mówi (por. Yablo 2001/2010, s. 181, Joyce 2005, s. 291).

Richard Joyce (2005) podkreśla, że zatarcie różnicy pomiędzy mówieniem o fikcji i opieraniem się na fikcji w tym, co się mówi, sprawia, że nie możemy wyjaśnić tego, co robi narrator, gdy opowiada powieść. Choć gdy mówimy o treści *Procesu* F. Kafki, wygłaszamy twierdzenia typu „Zgodnie z *Procesem* Kafki Józefowi K. nie przedstawiono zarzutów”, to już teza, że dowolne fikcyjne twierdzenie poddaje się takiej interpretacji, uniemożliwia oddanie sprawiedliwości naszym intuicjom co do tego, czym jest narracja. Ta zaś wydaje się formą udawania czegoś. Gdy narrator opowiada *Proces* Kafki, udaje, że stwierdza, że Józefowi K. nie postawiono zarzutów. Uduje, że jest osobą, która ma wiedzę o zdarzeniach, o których opowiada. Gdyby twierdzenie, że Józefowi K. nie postawiono zarzutów, było tylko eliptycznie wyrażonym twierdzeniem o postaci „Zgodnie z *Procesem* Kafki Józefowi K. nie przedstawiono zarzutów”, to nie można byłoby powiedzieć o narratorze, że udaje, że coś stwierdza. W jakim bowiem celu ktokolwiek miałby udawać, że stwierdza literalnie prawdziwe twierdzenie dotyczące treści *Procesu* Kafki?

Jeśli zgodzić się na rozróżnienie na mówienie o treści fikcji i mówienie, w którym opieramy się na treści fikcji, to można je rozciągnąć na dyskursy, w których formułowane są ontologicznie kontrowersyjne twierdzenia. Niech dyskursem takim będzie dyskurs D i wyrażona w nim teoria *T*, której zobowiązań ontologicznych fikcjonałści nie chcą akceptować. Z jednej strony, można twierdzić, że gdy mówimy o treści teorii *T*, możemy wygłaszać w sposób eliptyczny twierdzenie *p* z intencją rozumienia go jako twierdzenia o postaci „Zgodnie z *T* *p*”. Z drugiej jednak strony, gdy używamy teorii *T* do mówienia o obiektywnie zachodzących okolicznościach, posługujemy się jej twier-

dzeniami, nie mówiąc niczego o jej treści. Mówimy tak, jak gdyby teoria  $T$  stanowiła adekwatną reprezentację faktów. Druga z tych form akceptacji twierdzeń może właśnie być wyjaśniana w kategoriach akceptacji ich jako prawdziwych w ramach gry w udawanie, tj. o podmiocie  $S$  powiemy, że *udaje*, że stwierdza, że  $p$  wtw, gdy  $S$  akceptuje  $p$  jako prawdziwe w ramach odpowiedniej gry w udawanie (por. Joyce 2005, s. 291–292).

Kolejnym powodem, dla którego — zdaniem Joyce’a (2005, s. 292) — powinniśmy preferować fikcjonalizm przedmiotowy, jest to, że w łatwy sposób możemy wyjaśnić jeszcze jeden fenomen. Otóż osoby akceptujące twierdzenia, w których prawdziwość nie wierzą, mają tendencję do wycofywania się z tego, co powiedziały, gdy tylko kontekst stanie się nieco bardziej krytyczny. Gdy ktoś kwestionuje prawdziwość wygłaszanych przez nie twierdzeń, odpowiadają, że ściśle rzecz biorąc wygłoszone przez nich twierdzenia były fałszywe. Gdy mamy do czynienia z akceptacją twierdzenia w ramach gry w udawanie, zmiana nie ulega jego treści, a jedynie nasza postawa wobec niego — akceptujemy twierdzenie o pewnej treści, wyobrażając sobie istnienie bytów, które sprawiają, że jest ono prawdziwe. Tymczasem, zgodnie z metafikcjonalizmem, ontologicznie niezobowiązująca akceptacja twierdzenia  $p$  jest możliwa, ponieważ w rzeczywistości akceptuje się twierdzenie „Zgodnie z fikcją  $f$   $p$ ”. Wtedy jednak nie jest jasne, co robią osoby, które w bardziej krytycznych kontekstach wycofują swoje stwierdzenia. Z pewnością nie wycofują się uznania prawdziwości twierdzenia o postaci „Zgodnie z fikcją  $f$   $p$ ”, gdyż jest to twierdzenie literalnie prawdziwe. Być może — zdaniem metafikcjonalisty — w krytycznych kontekstach powiedzenie, że nie akceptuje się prawdziwości twierdzenia  $p$ , jest formą powiedzenia, że nie tyle wygłosiło się  $p$ , co raczej eliptycznie sformułowane twierdzenie o postaci „Zgodnie z fikcją  $f$   $p$ ”. Nie wyjaśnia to jednak, dlaczego użytkownicy twierdzeń, które pociągają za sobą — w ich mniemaniu — nieakceptowalne zobowiązania

ontologiczne, wycofują się z nich, ogłaszając ich fałszywość. Mogliby przecież stwierdzić, że *de facto* ogłosili twierdzenie o odmiennej treści. Wycofanie się z tego, co uprzednio się stwierdziło, jest zrozumiałe, gdy treść twierdzenia nie zmienia się, lecz w krytycznym kontekście zmieniają się standardy jego akceptacji. Gdy w debacie metafizycznej zaczynamy mówić wyłącznie o tym, co istnieje, i nie chcemy akceptować zobowiązań ontologicznych implikowanych przez pewne twierdzenia, możemy wycofać się z nich, ponieważ w kontekście metafizycznej debaty zmieniają się standardy akceptacji. Zmianami standardów akceptacji nie sposób natomiast wyjaśnić wycofania się z twierdzeń o postaci „Zgodnie z  $f$   $p$ ”, gdyż są one akceptowane jako prawdziwe, czy to w mniej, czy to bardziej krytycznym kontekście.

Wydaje się zatem, że fikcjonalizm przedmiotowy lepiej współgra z naszą praktyką językową i intuicyjnymi przekonaniem dotyczącymi tego, na czym polega wykonywanie takich czynności, jak opowiadanie opowieści bądź posługiwanie się pewną teorią (w odróżnieniu od mówienia o niej).

# Rozdział III

## Fikcjonalizm modalny

Fikcjonalista modalny stawia przed sobą dwa zasadnicze cele. Po pierwsze, chce uniknąć przyjmowania zobowiązań ontologicznych co do istnienia światów możliwych i możliwych indywiduów. Po drugie, chce zachować prawo do posługiwania się dyskursem światów możliwych. Uważa bowiem, że dyskurs ten doskonale sprawdza się w analizie pojęć modalnych, pozwalając na zredukowanie większości z nich do kwantyfikacji po światach możliwych i możliwych indywiduach.

W literaturze filozoficznej wyróżnia się modalności *de dicto* i modalności *de re*. Zwykle pojęcia modalne *de dicto* identyfikowane są z funktorami konieczności lub możliwości, których argumentami są wartościowalne logiczne zdania, jak np. w twierdzeniu „Z konieczności, wszyscy w tym pokoju są ludźmi”. Z kolei pojęcia modalne *de re* interpretuje się jako te funktory konieczności lub możliwości, których argumentami są wyłącznie predykaty, jak np. w twierdzeniu „Wszyscy w tym pokoju są z konieczności ludźmi”. Należy zauważyć, że oba zdania diametralnie różnią się warunkami prawdziwości. By pierwsze z nich mogło być prawdziwe, w każdym świecie możliwym, w którym istnieje dany pokój, musi być prawdą, że w pokoju tym znajdują się wyłącznie ludzie. Rzecz jasna, nie przekreśla możliwości tego, by w pokoju znajdowali się nie tylko ludzie. Zatem pierwsze z tych zdań jest

falszywe. By drugie z nich było prawdziwe, wystarczy, że w świecie aktualnym w danym pokoju znajdują się wyłącznie ludzie oraz że z konieczności człowiek jest człowiekiem. Jeśli w świecie aktualnym, w danym pokoju znajdują się wyłącznie ludzie, drugie z tych twierdzeń będzie prawdziwe, nawet jeśli w innych światach w pokoju będą znajdowali się nie tylko ludzie.

W ramach dyskursu światów możliwych każda z modalności jest analizowalna w kategoriach kwantyfikacji po światach możliwych lub po możliwych indywiduach. I tak:

„Z konieczności,  $p$ ” jest prawdziwe wtw, gdy dla każdego świata możliwego  $w$ ,  $p$  w  $w$

„Możliwie,  $p$ ” jest prawdziwe wtw, gdy istnieje świat możliwy  $w$  taki, że  $p$  w  $w$

„Dla każdego (istnieje takie)  $x$ , (że) z konieczności  $Fx$ ” jest prawdziwe wtw, gdy dla każdego świata  $w$  i dla każdego  $x$ , jeśli  $x$  istnieje w  $w^*$  i  $Fx$  w  $w^*$ , to jeśli  $x$  istnieje w  $w$ , to  $Fx$  w  $w$

„Dla każdego (istnieje takie)  $x$ , (że) możliwie  $Fx$ ” jest prawdziwe wtw, gdy istnieje świat możliwy taki, że jeśli  $x$  istnieje w  $w^*$ , to  $x$  istnieje w  $w$  i  $Fx$  w  $w$ ,

gdzie  $w^*$  jest światem aktualnym.

Każda z tych modalności jest analizowana za pomocą dobrze znanej i niezbyt skomplikowanej aparatury logicznej, jaką jest kwantyfikacja pierwszego rzędu. Biorąc pod uwagę powszechność występowania pojęć modalnych w filozofii, sympatia do możliwości zredukowania ich przy użyciu prostych środków logicznych nie powinna dziwić. Dlatego filozofowie niechętnie porzucają możliwość korzystania z możliwości światowych analiz. Jednak równie niechętnie akceptują zobowiązania ontologiczne, które wydają się wynikać z takich analiz.

Żywot fikcjonalizmu modalnego rozpoczął się, gdy podejmowano jedno z pierwszych prób pogodzenia możliwości korzystania z możliwościówświatowych analiz z jednoczesnym odrzuceniem istnienia światów możliwych lub zawieszeniem sądu w tej sprawie. Wszystko na dobre rozpoczęło się, gdy G. Rosen w *Modal Fictionalism* (1990) zaproponował metafikcjonalistyczną interpretację dyskursu światów możliwych. Wraz z rozwojem fikcjonalizmu przedmiotowego, za sprawą artykułu S. Yablo *How in the World?* (1996), pojawiła się również przedmiotowa odmiana fikcjonalizmu modalnego. Następujące cechy są wspólne obu wersjom fikcjonalizmu modalnego:

- (1<sub>MF</sub>) wbrew instrumentalizmowi w sprawie istnienia światów możliwych (zob. §5.1.) fikcjonalista modalny traktuje twierdzenia możliwościświatowe, tj. twierdzenia mające na celu odnoszenie się do światów możliwych lub możliwych indywiduów, jako autentyczne reprezentacje tego, jak rzeczy się mają;
- (2<sub>MF</sub>) wbrew redukcjonizmowi twierdzenia możliwościświatowe powinny być odczytywane dosłownie jako implikujące istnienie światów możliwych i możliwych indywiduów;
- (3<sub>MF</sub>) wbrew realizmowi modalnemu celem używania dyskursu światów możliwych nie jest opis rzeczywistości, a sformułowanie teorii o zaletach niezależnych od jej prawdziwości.

Tym, co odróżnia oba stanowiska, jest koncepcja tego, na czym polega ontologicznie niezobowiązująca akceptacja twierdzeń możliwościświatowych.

## 1 Metafikcjonalizm modalny

Metafikcjonalista modalny utrzymuje, że twierdzeniem

- (1) „Istnieje świat możliwy, w którym istnieją niebieskie łabędzie”

jest eliptycznie wyrażonym twierdzeniem o postaci:

- (1\*) „Zgodnie z opowieścią o wielości światów istnieje świat możliwy, w którym istnieją niebieskie łabędzie”.

Prawdziwość ostatniego twierdzenia wymaga jedynie — twierdzi metafikcjonalista — przyjęcia istnienia opowieści o wielości światów możliwych o takiej treści, że twierdzenie „Istnieje świat możliwy, w którym istnieją niebieskie łabędzie” jest prawdziwe zgodnie z nią. Zarówno istnienie opowieści o wielości światów, jak i to, że z jej treści wynikają pewne twierdzenia, jest niezależne od jej prawdziwości — jest niezależne od istnienia światów możliwych i możliwych indywiduów. Zatem metafikcjonałiści modalni, nie przyjmując istnienia światów możliwych bądź nie uznając przyjęcia ich istnienia za zasadne, wierzą jedynie w istnienie opowieści, zgodnie z którą prawdziwe są niektóre spośród twierdzeń dotyczących światów możliwych i możliwych indywiduów (Rosen 1990, s. 332).

Gdy twierdzenie (1) nie jest wygłoszone z intencją wyrażenia twierdzenia (1\*), jest ono odrzucane z uwagi na przekonanie o jego fałszywości bądź sąd w jego sprawie jest zawieszany z uwagi na agnostyczną postawę wobec kwestii istnienia światów możliwych<sup>25</sup>.

Uznanie, iż twierdzenia możliwościświatowe są prawdziwe jedynie zgodnie z opowieścią o wielości światów możliwych, stanowi pierwszy krok metafikcjonalistycznej strategii unikania zobowiązań ontologicznych co do światów możliwych i możliwych indywiduów. Kolejnym powinno być powiedzenie czegoś więcej o treści hipotezy o istnieniu światów możliwych. Wskazanie, co jest treścią tej opowieści, pozwoli na ocenę tego, czy dyskurs możliwościświatowy poddany metafikcjonalistycznej interpretacji może pełnić rolę, jakie fikcjonalista modalny

---

<sup>25</sup>W § 4.2. przedstawiona zostanie trzecia propozycja, pochodząca od (Rosen 1995), zgodnie z którą przynajmniej niektóre spośród twierdzeń możliwościświatowych niepoprzedzone funktorem opowieści są pozbawione wartości logicznych.



chciałby, by dyskurs ten pełnił — w szczególności, czy nadaje się do analizowania wspomnianych wyżej pojęć modalnych.

W związku z powyższym, kryteria, którymi metafikcjonałści modalni będą kierowali się w poszukiwaniu odpowiedniej opowieści o wielości światów, będą przypominały kryteria, które powinna spełnić dobra teoria modalności. Zdaniem Rosena będą one przedstawiały się następująco:

- (K<sub>1</sub>) równoważność materialna twierdzeń modalnych z ich możliwoświatowymi przekładami;
- (K<sub>2</sub>) wystarczające moce ekspresywne opowieści o wielości światów, które pozwalają na przełożenie dowolnego, poprawnie zbudowanego twierdzenia modalnego na twierdzenie możliwoświatowe poprzedzone funktorem opowieści „Zgodnie z opowieścią o wielości światów”;
- (K<sub>3</sub>) niekolistość opowieści o wielości światów — jeśli opowieść o wielości światów ma dostarczyć redukcji zdań modalnych do wypowiedzi o światach możliwych, to nie powinna sama zakładać pojęć modalnych.

## 1.1 Fikcja realisty modalnego

### 1.1.1 Lewisowska koncepcja światów możliwych

W przekonaniu Rosena, powyższe warunki najlepiej spełnione są przez D. Lewisa koncepcję światów możliwych. Na treść tej opowieści będą składały się następujące twierdzenia:

- (2a) Rzeczywistość składa się z wielości wszechświatów czy „światów”.

- (2b) Jeden z nich jest tym, co zwykle nazywamy *tym światem* — największym czasoprzestrzennie powiązany systemem, którego jesteśmy częściami.
- (2c) Inne [światy — P.W] są w przybliżeniu rzeczami tego samego rodzaju [co nasz świat — P.W.]: są systemami obiektów, w większości konkretnych, połączonych przez sieć zewnętrznych relacji, podobnych czasoprzestrzennym odległościom, które łączą obiekty w naszym świecie.
- (2d) Każdy wszechświat jest odizolowany od pozostałych, tj. partykularia należące do odmiennych światów nie pozostają do siebie w relacjach czasoprzestrzennych [...].
- (2e) Ogół wszechświatów jest zamknięty na zasadę rekombinacji. Luźno mówiąc, dla dowolnego zestawu obiektów z dowolnej liczby wszechświatów istnieje wszechświat zawierający dowolną liczbę duplikatów każdego z nich, pod warunkiem że istnieje czasoprzestrzeń wystarczająco duża, by pomieścić je.
- (2f) Nie ma arbitralnych (*arbitrary*) ograniczeń dotyczących wielości wszechświatów.
- (2g) Nasz wszechświat nie jest szczególny. Nie ma w nim niczego niezwykłego z perspektywy systemu wszechświatów (Rosen 1990, s. 333).

W treści tej fikcji musi znaleźć się jeszcze coś, co zagwarantuje nam, że pewne światy możliwe będą zawierały różowe słonie, niebieskie łabędzie itd. Same postulaty (2a) — (2g) nie mogą wygenerować takiej mnogości światów, dlatego muszą być uzupełnione o tzw. „*encyklopedię* — listę niemodalnych prawd o wewnętrznym charakterze tego wszechświata” (Rosen 1990, s. 335), gdzie tym światem jest świat aktualny. Wzbogacona o encyklopedię opowieść o wielości

światów, dzięki postulatowi o zamknięciu ogółu wszechświatów na zasadę rekombinacji, gwarantuje, że zgodnie z opowieścią o wielości światów, istnieją światy możliwe, w których istnieją niebieskie łabędzie itp. Niech opowieść, składająca się z listy postulatów (2a) — (2g) oraz katalogu niemodalnych prawd, nazywa się „PW”.

Treść opowieści PW różni się od treści realizmu modalnego tym, że nie odgrywa w niej żadnej roli akceptowana przez realistę modalnego teza, że światy różne od świata aktualnego są tym, co możliwe. Teza ta stanowi wyraz redukcyjnej koncepcji modalności proponowanej przez realistę modalnego, zgodnie z którą sposób, na jaki pewien świat mógłby być, jest utożsamiony ze sposobem, na jaki jest pewien świat osiągalny z tego świata. Jednak metafikcjonalisci modalni o redukcyjnych aspiracjach analizują warunki prawdziwości twierdzeń modalnych nie w kategoriach istnienia pewnego świata możliwego, ale w kategoriach prawdziwości pewnego twierdzenia zgodnie z treścią fikcji PW, która jest w wystarczającym stopniu wyznaczona przez listę postulatów (2a) — (2g) oraz encyklopedię (por. Rosen 1990, s. 334).

Tym, co wyróżnia realizm modalny na tle konkurencyjnych koncepcji światów możliwych, jest to, że w ramach tego stanowiska przy definiowaniu światów możliwych nie występują odwołania explicite do pojęć modalnych. Lewis utrzymuje bowiem, że światy możliwe są maksymalnymi sumami mereologicznymi możliwych indywiduów, z których każdy pozostaje względem każdego innego w relacjach czasoprzestrzennych (por. Lewis 1986, s. 71). Suma mereologiczna  $m$  jest maksymalna ze względu na relację  $R$  wtw, gdy dla każdego  $x$  i  $y$ , jeśli  $xRy$  lub  $yRx$ , to  $x$  i  $y$  są częściami  $m$  oraz dla dowolnego  $x$ , jeśli  $x$  jest częścią  $m$ , to nie istnieje takie  $z$ , że  $z$  jest częścią  $m$  i nieprawda, że  $xRz$  lub  $zRx$ <sup>26</sup>. Same możliwe indywidua mają być

<sup>26</sup>Lewis był zwolennikiem uniwersalizmu mereologicznego głoszącego, że dowolne obiekty – niezależnie od ich rodzaju – mogą stanowić części jednej sumy mereologicznej. Dlatego w powyższej definicji maksymalnej sumy mereologicznej nie

bytami konkretnymi, które zasadniczo przypominają nas samych i przedmioty znajdujące się w naszym otoczeniu.

Taka koncepcja światów możliwych wzbudziła liczne wątpliwości wśród filozofów. W jednym z często podnoszonych zarzutów zwracano uwagę na to, że do konsekwencji tej koncepcji należy to, że nie mogłyby istnieć przedmioty pozbawione czasoprzestrzennej lokalizacji. Skoro częściami światów mogą być wyłącznie przedmioty, które pozostają do siebie w relacjach czasoprzestrzennych, to w każdym z nich mogą istnieć jedynie przedmioty znajdujące się w czasoprzestrzeni. Jeśli światy możliwe mają w sposób wyczerpujący reprezentować ogół możliwości, to istnienie przedmiotów pozbawionych czasoprzestrzennej lokalizacji byłoby niemożliwe. Tymczasem nie bardzo wiadomo, dlaczego niemożliwe miałyby być istnienie świata, w którym znajdowałyby się — powiedzmy — pozbawione czasoprzestrzennej lokalizacji dusze. Hipoteza taka nie nosi znamion sprzeczności, powinna więc mieć swoją reprezentację w ramach koncepcji światów możliwych, o ile jej moce ekspresywne mają pozwalać na dostarczenie analiz dla modalności logicznych.

Aby obejść tę trudność, Lewis proponuje uogólnić wprowadzoną wyżej koncepcję światów możliwych do postaci:

*Światy możliwe są maksymalnymi sumami mereologicznymi, których części powiązane są relacjami czasoprzestrzennymi lub relacjami analogicznymi do relacji czasoprzestrzennych.*

Relacja  $R$  jest analogiczna do relacji czasoprzestrzennej, gdy łącznie spełnia cztery warunki:

(A1) relacja  $R$  jest *naturalna* (Lewis 1986, s. 75–76),

---

nakładałem żadnych warunków wartości zmiennych  $x$  i  $y$ . Mogą nimi być indywidua, jak to ma miejsce w przypadku maksymalnych sum mereologicznych, których części powiązane są relacjami czasoprzestrzennymi. Mogą nimi być np. treści – u Lewisa rozumiane zbiory światów możliwych – jak to ma miejsce w przypadku maksymalnych sum mereologicznych treści powiązanych relacją równoważności logicznej.

(A2) relacja  $R$  jest wszechogarniająca (Lewis 1986, s. 76),

(A3) relacja  $R$  jest wyróżniająca (Lewis 1986, s. 76),

(A4) relacja  $R$  jest zewnętrzną (Lewis 1986, s. 76).

Zacznijmy od warunku (A1). Zasadniczą ideą stojącą za podziałem relacji na naturalne i nienaturalne jest próba oddania różnicy pomiędzy takimi relacjami, których wspólne posiadanie przez sumy mereologiczne indywiduów sprawia, że sumy te są do siebie jakościowo podobne — oraz takimi relacjami, których wspólne posiadanie przez sumy mereologiczne indywiduów nie sprawia, że sumy te są jakościowo podobne (Lewis 1986, s. 59–60).

Rozróżnienie to można zilustrować za pomocą przykładów takich relacji jak *bycie oddalonym o 10 metrów* oraz *bycie oddalonym o 10 metrów lub o 10 lat*. Najmniejsze sumy mereologiczne indywiduów, pomiędzy którymi zachodzi pierwsza relacja, są jakościowo podobne do siebie w tym sensie, że ich częściami są indywidua, które oddalone są od siebie o 10 metrów. Natomiast w przypadku posiadania drugiej relacji przez dwie sumy mereologiczne indywiduów nie mamy gwarancji, że będą one do siebie jakościowo podobne w tym sensie, że ich części będą pozostawały do siebie w tych samych niedysjunktywnych relacjach. Może bowiem być tak, że części jednej sumy mereologicznej będą pozostawały do siebie w tej relacji na mocy tego, że są oddalone od siebie o 10 metrów, podczas części innej sumy mereologicznej będą pozostawały do siebie w tej relacji na mocy tego, że są oddalone od siebie o 10 lat.

Nie należy przeceniać roli jakościowego podobieństwa w eksplicacji rozróżnienia na relacje naturalne i nienaturalne. Nie stanowi ono bowiem faktu pierwotnego. Przeciwnie, jakościowe podobieństwo najczęściej wyjaśniane jest w kategoriach posiadania wspólnych cech, które są stosunkowo jednorodne, czyli zaliczają się do relacji naturalnych. Można nawet pokusić się o uogólnienie i powiedzieć, że sumy

$m$  i  $m^*$  są jakościowo podobne do siebie ze względu na relację  $R$ , gdy jeśli relacja  $R$  jest relacją naturalną, to obie sumy mereologiczne posiadają części, które pozostają do siebie w relacji  $R$  lub jeśli relacja  $R$  jest relacją dysjunktywną, na którą składają się relacje naturalne  $R_1$  i  $R_2$ , to obie sumy posiadają takie części, które pozostają do siebie co najmniej w jednej z tych relacji.

Samo rozróżnienie na relacje naturalne i nienaturalne najczęściej traktowane jest jako rozróżnienie stopniowalne. Najbardziej naturalnymi relacjami są relacje niedefiniowalne w kategoriach bardziej pierwotnych relacji. Nazwijmy je *doskonale* naturalnymi relacjami. Inne relacje będą naturalne jedynie w pochodnym sensie, tj. będą takimi relacjami, które da się zdefiniować za pomocą niezbyt skomplikowanego ciągu definicji, w których występują wyłącznie odwołania do doskonale naturalnych relacji. Jeszcze inne relacje będą zachodziły pomiędzy indywiduami w na tyle niejednorodny sposób, że ciężko będzie wskazać jakości, które dzielą między sobą sumy mereologiczne, których częściami są te indywidua. Relacje te mogą być na tyle niejednorodne, że w języku nie będą występowały predykaty na ich nazwanie.

Czy pojęcie naturalności relacji (i oczywiście własności) jest pojęciem pierwotnym? Lewis udziela pozytywnej odpowiedzi na to pytanie:

Wielu filozofów jest sceptycznych w sprawie rozróżnienia na naturalne i niejednorodne (*gruesome*) własności. Uważają, że jest ono niezasadne, o ile nie można jakoś przeprowadzić go przy użyciu kategorii, które nie zakładają tego rozróżnienia. Sądzę, że nie można tego dokonać, ponieważ stale zakładamy to rozróżnienie. Czy powiemy, że naturalne własności są tymi, które figuruje w prawach natury? Nie, jeśli będziemy używać naturalności własności, gdy będziemy przeprowadzać rozróżnienie pomiędzy prawami natury a akcydentalnymi regularnościami. Czy

powiemy, że są one tymi, które figurują w treści naszych myśli? Nie, jeśli będziemy mówili, że uniknięcie bezzasadnej niejednorodności jest częścią tego, co konstytuuje poprawność przypisania treści. Czy powiemy, że są one tymi, których egzemplifikacje są zjednoczone podobieństwem? Nie, jeśli będziemy mówili, że podobieństwo jest dzieleniem naturalnych własności. Jeśli nie jesteśmy gotowi do obejścia się bez pewnych użyć rozróżnienia pomiędzy naturalnymi i nienaturalnymi własnościami, nie będziemy mieć łatwej drogi do jego niekolistego zdefiniowania. Nie jest to powodem do odrzucenia tego rozróżnienia, a raczej, jeśli jest nieuniknione, jest to powodem do zaakceptowania go jako pierwotnego (Lewis 1986, s. 63).

Zatem jednym z pojęć pierwotnych w naszej opowieści o wielości światów jest pojęcie naturalności relacji.

Zgodnie z warunkiem (A2) relacja analogiczna do relacji czasoprzestrzennej jest relacją wszechobejmującą. Oznacza to, że jeśli w systemie rzeczy istnieje ciąg relacji typu  $R$  biegnący od indywiduum  $a$  do indywiduum  $b$ , to relacja typu  $R$  zachodzi bezpośrednio pomiędzy  $a$  i  $b$ . By uniknąć nieporozumień, warto w tym miejscu zaznaczyć, że w tym wypadku zmienna  $R$  musi przyjmować jako swoje wartości tak nieokreślone relacje jak np. relacja czasoprzestrzennego oddalenia. Gdyby przyjmowała jako swoje wartości tak określone relacje jak np. relacja bycia oddalonym o 10 m, to warunek (A2) miałby absurdalne konsekwencje. Wymagałby, by w przypadku gdy indywiduum  $a$  jest oddalone o 10 m od indywiduum  $b$  i  $b$  jest oddalone o 10 m od  $c$ , to  $a$  jest oddalone o 10 m od  $c$ . A taka sytuacja wcale nie musi mieć miejsca.

Z kolei zgodnie z warunkiem (A3) relacja analogiczna do relacji czasoprzestrzennej jest relacją wyróżniającą. O relacji  $R$  powiemy, że jest relacją wyróżniającą, gdy nie istnieją takie dwie rzeczy  $x$  i  $y$  w

systemie rzeczy powiązanych relacją  $R$ , że  $x$  i  $y$  mają to samo miejsce w tym systemie. Innymi słowy, nie jest prawdą, że  $x$  i  $y$  pozostają w dokładnie takich samych relacjach  $R$  do tych samych rzeczy.

Ostatni warunek — (A4) wymaga od relacji  $R$ , by była relacją zewnętrzną. W filozoficznej literaturze relacje zewnętrzne kontrastuje się z relacjami wewnętrznymi. O relacji  $R$  powiemy, że jest relacją wewnętrzną względem jej członów (wziętych z osobna) wtw, gdy „kiedykolwiek  $a$  i  $a^*$  są duplikatami (lub są identyczne) oraz  $b$  i  $b^*$  są duplikatami (lub są identyczne), to obie pary  $\langle a, b \rangle$  i  $\langle a^*, b^* \rangle$  pozostają w relacji [ $R$  — P.W.] bądź żadna z nich” (Lewis 1983b, s. 356—przyp. 16). Zgodnie z Lewisa teoriomnogościową koncepcją własności elementy pary  $\langle a, b \rangle$  pozostają w relacji  $R$  do siebie, gdy para ta należy do zbioru par uporządkowanych, z którym identyfikuje się relację  $R$ . Z kolei  $x$  jest duplikatem  $y$  wtw, gdy  $x$  i  $y$  posiadają dokładnie te same własności wewnętrzne (Lewis 1983a, s. 355). Własność  $F$  jest własnością wewnętrzną indywiduum  $x$  wtw, gdy  $x$  posiada  $F$  niezależnie od relacji, w jakich pozostaje do otaczającego go świata (Lewis 1986, s. 199). Podsumowując, relacjami wewnętrznymi są te relacje, które zachodzą pomiędzy indywiduami na mocy posiadanych przez nie własności wewnętrznych.

W przeciwieństwie do nich relacje zewnętrzne są relacjami, które nie zachodzą wyłącznie ze względu na wewnętrzne własności członów tych relacji. W (1983a) Lewis definiuje pojęcie relacji zewnętrznej w następujący sposób:

relacja  $R$  jest relacją zewnętrzną wtw, gdy  $R$  jest relacją wewnętrzną względem par uporządkowanych, których elementami są członowie tej relacji.

Z kolei relacja ma być relacją wewnętrzną względem par uporządkowanych, gdy dla dowolnych par uporządkowanych  $\langle x, y \rangle$  i  $\langle x^*, y^* \rangle$ , jeśli pary te są duplikatami, to ich członowie albo pozostają, albo nie pozostają do siebie w relacji  $R$  (Lewis 1983a, s. 356, przyp.



16). Para  $\langle x, y \rangle$  jest duplikatem pary  $\langle x^*, y^* \rangle$  wtw, gdy (a)  $x$  i  $x^*$  posiadają te same własności wewnętrzne,  $y$  i  $y^*$  posiadają te same własności wewnętrzne oraz (b) wewnętrzne relacje pomiędzy  $x$  i  $y$  są takie same jak wewnętrzne relacje pomiędzy  $x^*$  i  $y^*$  (por. Lewis 1983a, s. 356–przyp. 17, 357).

Zaproponowana przez Lewisa definicja budzi wątpliwości. W pierwszym rzędzie warunek (b) jest redundantny. Na mocy definicji relacji wewnętrznej zachodzi ona pomiędzy dwoma indywiduami, gdy posiadają one te same własności wewnętrzne. Jeśli tak, to warunek (b) jest implikowany przez warunek (a). Jeśli bowiem pomiędzy dwoma przedmiotami zachodzą pewne relacje wewnętrzne, dokładnie te same relacje będą zachodziły pomiędzy duplikatami tych przedmiotów.

Proponowana przez Lewisa definicja relacji zewnętrznej jest również nieadekwatna. Weźmy paradygmatyczny przykład takiej relacji, jakim jest relacja czasoprzestrzenna. Niech  $a$  i  $b$  będą niebieskimi kulami o tej samej wielkości i masie, a relacja przestrzenna  $R$  będzie relacją bycia oddalonym o 10 m. Jeśli zgodzić się z Lewisem, że kształt jest wewnętrzną własnością indywiduów  $a$  i  $b$ , ale już nie jest nią ani ich kolor, ani masa, wówczas jedynymi relacjami wewnętrznymi zachodzącymi pomiędzy tymi kulami będą relacja posiadania tego samego kształtu i być może relacja posiadania tej samej wielkości. Z łatwością można znaleźć duplikat tej pary uporządkowanej. Wystarczy wziąć dowolną parę, której elementami będą czerwone kule  $c$  i  $d$  o tej samej wielkości co kule  $a$  i  $b$ . Problem jednak w tym, że nie sposób dostrzec, jakie własności tych par uporządkowanych miałyby sprawić, że kule pozostawałyby do siebie w tych samych relacjach czasoprzestrzennych. Co więcej, nic nie stoi na przeszkodzie, by kule  $a$  i  $b$  były oddalone od siebie o 5 m, mimo że posiadają te same własności wewnętrzne.

Niewystarczające jest zatem nałożenie warunku, by odpowiadające sobie człony par uporządkowanych, pomiędzy którymi zachodzi relacja zewnętrzna  $R$ , były duplikatami. Jeśli jednak relacja zewnętrzna  $R$

zachodząca pomiędzy członami pary uporządkowanej ma być relacją wewnętrzną względem par, których elementami są członowie tej relacji, to wymaga to przyjęcia istnienia wewnętrznych własności par uporządkowanych, ze względu na które relacja  $R$  zachodzi pomiędzy ich elementami. Musiałyby to być takie własności wewnętrzne tych par, które nie są wyznaczane wyłącznie przez własności wewnętrzne elementów tych par. Idąc za tą sugestią, powinniśmy powiedzieć, że para  $\langle a, b \rangle$  jest duplikatem  $\langle a^*, b^* \rangle$  wtw, gdy (a)  $a$  jest duplikatem  $a^*$ ,  $b$  jest duplikatem  $b^*$  oraz (b) pary  $\langle a, b \rangle$  i  $\langle a^*, b^* \rangle$  mają te same własności wewnętrzne.

Jakie własności wewnętrzne można przypisać parom uporządkowanym? Wydaje się, że jedynymi własnościami par uporządkowanych, które nie są zależne od relacji, w jakich pozostają one do innych obiektów, są ich własności formalne. Wśród nich znajdują się takie własności jak własność posiadania dwóch elementów, własność posiadania elementu wyróżnionego itd. Trudno jednak dostrzec jakąkolwiek własność wewnętrzną par uporządkowanych, która w ogóle mogłaby być traktowana jako podstawa zachodzenia relacji czasoprzestrzennych pomiędzy elementami tych par. Jeśli zaakceptować Lewisowską teoriomnogościową koncepcję relacji, zgodnie z którą  $n$ -członowa relacja jest identyfikowana ze zbiorami  $n$ -elementowych par uporządkowanych, to tym trudniej dostrzec, jak własności wewnętrzne pary uporządkowanej mogłyby stanowić jakąkolwiek podstawę dla zachodzenia np. relacji bycia oddalonym o 5 m. Wiele relacji binarnych może zachodzić pomiędzy dwoma indywiduami. Mogą one być oddalone o 5 m od siebie, mogą one nieustannie myśleć o sobie itd. Zgodnie z teoriomnogościową koncepcją relacji tym, co sprawia, że relacje te różnią się od siebie, jest to, że pary uporządkowane są elementami odmiennych zbiorów par uporządkowanych. Wówczas jednak to, czy pomiędzy członami danej pary uporządkowanej zachodzi określona relacja zewnętrzna, jest warunkowane nie tyle własnościami wewnętrznymi tej

pary, co jej relacją do określonego zbioru par uporządkowanych.

Jeśli chcemy zachować możliwie najwięcej z Lewisowskiej definicji relacji zewnętrznych, naturalnym posunięciem wydaje się zastąpienie par uporządkowanych jakimiś innymi obiektami, których wewnętrzne własności mogłyby stanowić podstawę dla zachodzenia relacji czasoprzestrzennych. W *On the Plurality of Worlds* (1986, s. 62) Lewis podsuwa kierunek takich zmian i definiuje relacje zewnętrzne jako relacje wewnętrzne względem sum mereologicznych członów relacji  $R$ . Problem jednak w tym, że wciąż nie jest to warunek, który zagwarantowałyby nam zachodzenie relacji czasoprzestrzennych pomiędzy obiektami. W ramach klasycznej mereologii, której Lewis był zwolennikiem, dla dwóch dowolnych obiektów istnieje suma mereologiczna, której są one częściami. W przypadku naszych dwóch kul będzie istniała ich suma mereologiczna, której jedynymi częściami będą nasze kule. Trudno oczekiwać, by na mocy własności wewnętrznych takiej sumy mereologicznej miały zachodzić określone relacje czasoprzestrzenne. Suma taka bowiem istniałaby i posiadałaby dokładnie takie same własności wewnętrzne niezależnie od tego, czy jedna z naszych kul znajdowałaby się w Paryżu, a druga w Nowym Jorku, czy obie znajdowałyby się w tej samej szufladzie.

Znacznie lepiej zdefiniować relację zewnętrzną w następujący sposób:

relacja  $R$  zachodząca pomiędzy obiektami  $x$  i  $y$  jest relacją zewnętrzną wtw, gdy (a)  $R$  nie zachodzi na mocy wewnętrznych własności  $x$  i  $y$ , (b)  $R$  jest relacją wewnętrzną względem takich sum mereologicznych, których częściami są  $x$  i  $y$  i które posiadają tylko te części, których posiadanie jest konieczne dla zachodzenia relacji  $R$ .

O relacji  $R$  powiemy, że jest wewnętrzną względem sumy mereologicznej  $m$ , gdy dla dowolnej sumy mereologicznej  $m^*$ , jeśli  $m^*$  jest duplikatem  $m$  i pomiędzy częściami sumy  $m$  zachodzi relacja  $R$ ,

to  $R$  zachodzi pomiędzy częściami  $m^*$ . Suma mereologiczna  $m^*$  jest duplikatem sumy mereologicznej  $m$ , gdy każda część sumy  $m^*$  jest duplikatem pewnej części sumy  $m$  i odwrotnie oraz  $m^*$  posiada takie same własności wewnętrzne jak  $m$ .

Przy tak zdefiniowanej relacji zewnętrznej kluczowym pytaniem jest to, czy sumy mereologiczne posiadają własności wewnętrzne, które nie są determinowane przez ich własności formalne, tj. te spośród ich własności, które wyznaczone są przez aksjomaty klasycznej mereologii. Stanowisko Lewisa wyraźnie wspiera twierdzącą odpowiedź na to pytanie. W swej ontologii Lewis przyjmuje istnienie sum mereologicznych, które posiadają pewne własności wewnętrzne, nie będące zarazem ich formalnymi własnościami. Przykładami takich sum są niemal<sup>27</sup> wszystkie teoretycznie interesujące indywidua, które posiadają części właściwe — indywidua w skali makro, światy możliwe itd. I tak, własność posiadania kształtu przez obiekty w skali makro jest przykładem własności wewnętrznej sumy mereologicznej, która nie jest zarazem jej własnością formalną.

Idąc tym tropem, relacje przestrzenne zachodzące pomiędzy indywiduami  $a$  i  $b$  można interpretować jako relacje wewnętrzne względem regionu przestrzennego  $r$ , będącego najmniejszą sumą punktów przestrzeni, które obejmują przestrzeń pomiędzy indywiduami  $a$  i  $b$  oraz wszystkie punkty przestrzeni zajmowane przez części tych indywiduów. Relacja przestrzenna jest relacją wewnętrzną względem regionu przestrzennego  $r$ , gdy dla dowolnego regionu przestrzennego  $r^*$  będącego duplikatem regionu przestrzennego  $r$ , jeśli pomiędzy indywiduami w  $r$  zachodzi relacja przestrzenna bycia oddalonym o  $x$  metrów, to zachodzi ona również pomiędzy indywiduami w regionie  $r^*$  i odwrotnie. Z kolei, region przestrzenny  $r^*$  jest duplikatem regionu przestrzennego  $r$  wtw, gdy pomiędzy punktami regionu  $r$  i regionu  $r^*$  zachodzi jedno-

---

<sup>27</sup>Klauzula „niemal” została wprowadzona w celu wykluczenia badań teoretycznych (takich jak klasyczna mereologia), dla których dowolna suma mereologiczna jest teoretycznie interesująca.

jednoznaczne przyporządkowanie, relacje pomiędzy poszczególnymi punktami przestrzeni w obu regionach są takie same, zachodzi jedno-jednoznaczne przyporządkowanie punktów przestrzeni zajmowanych przez indywidua bądź ich części i relacje pomiędzy zajmowanymi punktami przestrzeni są takie same w obu regionach. W związku z tym relacje przestrzenne są relacjami wewnętrznymi względem regionów przestrzennych, które są sumami mereologicznymi punktów przestrzeni zajmowanych przez człony relacji przestrzennych oraz możliwie najmniejszej ilości punktów niezajmowanych przez człony tej relacji.

Podsumowując dotychczasowe rozważania, można by przeformułować Lewisowską koncepcję światów możliwych do postaci:

*Światy możliwe są maksymalnymi sumami mereologicznymi powiązanymi taką relacją  $R$ , że  $R$  jest naturalna, wszechobejmująca, wyróżniająca oraz zewnętrzna.*

Warto w tym miejscu również zaznaczyć, że Lewisowskie światy możliwe są bytami o ekstensjonalnych warunkach identyczności: świat  $w$  jest identyczny ze światem  $w^*$  wtw, gdy dla dowolnego indywiduum  $x$ , jeśli  $x$  jest częścią świata  $w$ , to  $x$  jest częścią świata  $w^*$  i odwrotnie. Z uwagi na to, że relacje, wiążące wszystkie indywidua istniejące w świecie  $w$ , są relacjami wewnętrznymi względem sumy tych indywiduów, nie musimy odwoływać się do tej relacji przy okazji formułowania warunków identyczności dla świata  $w$ .

W związku z powyższym bardzo często utrzymuje się, że Lewisowska koncepcja światów możliwych pozwala na pełną redukcję pojęć modalnych do pojęć ekstensjonalnych. Choć przy formułowaniu realistycznej koncepcji światów możliwych *explicite* nie używa się pojęć modalnych, to jednak twierdzenie, że z pomocą Lewisowskiej koncepcji światów możliwych można dokonać całkowitej redukcji pojęć modalnych, przypomina nieco ogłaszanie zwycięstwa w bitwie jeszcze przed jej zakończeniem. Okazuje się bowiem, że milczącym

założeniem koncepcji Lewisa jest to, że częściami maksymalnych sum mereologicznych, z którymi utożsamia się światy, muszą być przedmioty możliwe. Nie bardzo jednak wiadomo, jak wyrazić to założenie bez odwoływania się do pojęć modalnych. Niezbyt satysfakcjonująca byłaby np. odpowiedź, że wszystkie przedmioty istniejące w światach są możliwe, ponieważ są częściami światów możliwych. To, czy dana suma mereologiczna jest możliwa, jest wtórne względem tego, czy posiada ona możliwe części. Suma mereologiczna kwadratowego koła z moim samochodem jest przecież tak samo niemożliwa jak samo kwadratowe koło. Lewis musi zatem przyjmować, że części światów są możliwe i nie wydaje się, by mógł dostarczyć analizy tego założenia, w której nie musiałby odwoływać się do pojęcia możliwości. Fikcja realisty modalnego wydaje się zatem nie spełniać warunku (K<sub>3</sub>), zgodnie z którym powinna dostarczać niekolistej redukcji pojęć modalnych.

Nie jest to jakiś szczególny problem. Najbardziej obiecująca — przynajmniej w mojej opinii — z konkurencyjnych koncepcji światów możliwych również wymaga przyjęcia pierwotnych pojęć modalnych. Dodatkowo wymusza przyjęcie istnienia bytów intensjonalnych, które nie pojawiają się ontologii realisty modalnego. Fakt ten można uznać za rację do preferowania treści realizmu modalnego jako treści opowieści o wielości światów możliwych.

Koncepcja, którą mam na myśli, to umiarkowany realizm Alвина Plantingi (1974). Zgodnie z nim światy możliwe utożsamia się z maksymalnymi możliwymi stanami rzeczy. Stan rzeczy  $x$  jest maksymalnym stanem rzeczy wtw, gdy dla dowolnego stanu rzeczy  $y$ ,  $y$  jest zawarte w stanie rzeczy  $x$  lub  $y$  jest wykluczone ze stanu rzeczy  $x$ . O stanie  $y$  powiemy, że jest wykluczony z maksymalnego stanu rzeczy  $x$  wtw, gdy *nie jest możliwe*, by  $x$  zachodziło i  $y$  zachodziło. Maksymalny stan rzeczy  $x$  jest możliwy wtw, gdy wszystkie stany rzeczy zawarte w tym stanie rzeczy *mogą* zarazem zachodzić. W koncepcji światów możliwych Plantingi przyjmuje się zatem pierwotne pojęcie

możliwości.

Elementami konstytutywnymi stanów rzeczy wchodzących w skład maksymalnych stanów rzeczy są takie byty intensjonalne, jak własności, n-argumentowe relacje oraz *indywidualne istoty*. Ostatnie „są kompletnymi i spójnymi zbiorami własności zrelatywizowanych do świata” (Plantinga 1974, s. 77). Indywidualne istoty pełnią rolę substitutów dla transświatowych indywiduów, tj. indywiduów istniejących w więcej niż jednym świecie możliwym. Każde z takich indywiduów posiada kompletną jakościową charakterystykę w każdym ze światów, w których istnieje. Innymi słowy, dla dowolnej własności, albo własność ta przysługuje temu indywiduum, albo mu nie przysługuje w danym świecie. Z uwagi jednak na to, że indywidua istniejące w wielu światach możliwych posiadają odmienne własności w odmiennych światach (różnią się ze względu na przygodnie posiadane własności), zbiór własności posiadanych przez to indywiduum w jednym świecie różni się od zbioru własności posiadanych przez to indywiduum w innym świecie. Celem uniknięcia sprzeczności posiadania przez jedno i to samo indywiduum wykluczających się własności (np. posiadania prawej ręki i nieposiadania prawej ręki) posiadanie tych własności relatywizuje się do świata możliwego, w którym dane indywiduum istnieje. Wymusza to przyjęcie stanowiska, zgodnie z którym posiadanie własności nie jest relacją dwuargumentową, której członami są indywiduum i własność, ale jest relacją trójargumentową, której członami są indywiduum, własność i świat możliwy.

Zaletą proponowanej przez Plantingę koncepcji indywidualnych istot jest to, że nie stoi ona w rażącej sprzeczności z naszymi intuicjami. Jeśli np. uważamy, że krasnoludki mogłyby istnieć, to w ramach koncepcji Plantingi utożsamia się je z takimi abstrakcyjnymi bytami jak kompletne zbiory własności zrelatywizowanych do światów możliwych. Istnienie tak abstrakcyjnych bytów zwykle nie stoi w rażącej sprzeczności z potocznymi intuicjami, że krasnoludki nie istnieją.

Jest tak dlatego, że intuicje te zazwyczaj są wyrazem sprzeciwu wobec przyjmowania istnienia krasnoludków jako bytów konkretnych, które nie różnią się zasadniczo od nas samych i przedmiotów, które znamy z naszego otoczenia. Intuicje pozostają natomiast niesprecyzowane co do istnienia takich bytów jak kompletne zbiory własności zrelatywizowanych do świata. Dla odmiany Lewisowska koncepcja możliwych indywiduów stoi w jaskrawej sprzeczności z tymi intuicjami. Lewis proponuje, by krasnoludki utożsamiać z konkretnymi indywiduami, które różnią się od przedmiotów znanych nam z otoczenia tym, że istnieją w innym świecie możliwym. *Nota bene* sam świat możliwy również jest bytem tego samego rodzaju co świat, w którym przyszło nam żyć i który uznajemy za aktualny tylko dlatego, że jesteśmy jego częściami.

Gdy poddać Lewisowskie stanowisko fikcjonalistycznej interpretacji, okazuje się, że znaczenie lepiej nadaje się ono na treść fikcji o wielości światów możliwych niż propozycja Plantingi. Ostatnia nie tylko nie spełnia warunku  $(K_3)$ , ale również  $(K_1)$ , zgodnie z którym powinna zachodzić równoważność materialna pomiędzy twierdzeniami modalnymi a twierdzeniami możliwościświatowymi. Utożsamienie możliwych indywiduów z kompletnymi i spójnymi zbiorami własności zrelatywizowanych do świata sprawia, że nie posiada się środków do reprezentowania światów, w których istnieją dwa jakościowo nieodróżnialne indywidua. Wyobraźmy sobie, że istnieje świat możliwy, w którym istnieją wyłącznie dwa elektrony krążące wokół siebie po orbitach o identycznym kształcie, usytuowanych względem siebie równolegle (por. Black 1954). Dowolna własność wewnętrzna bądź zewnętrzna posiadana przez jeden z tych elektronów w tym świecie jest zarazem własnością posiadaną przez drugi z nich i odwrotnie. Niech światem tym będzie świat *w*, wówczas w świecie *w* oba elektrony będą reprezentowane za pomocą tego samego zbioru własności zrelatywizowanych do świata *w*. By w ramach ontologii Plantingi



móc odróżnić od siebie te elektrony, należy wskazać świat możliwy, w którym jeden z tych elektronów posiada własności, których nie posiada drugi z nich. Problem jednak w tym, że jakiegokolwiek własności posiadałby jeden z tych elektronów w jakimś świecie, powinien istnieć świat możliwy, w którym drugi z nich posiada te same własności. Ciężko bowiem wskazać powód, dla którego to, co jest możliwe dla jednego z jakościowo nieodróżnialnych (w świecie  $w$ ) elektronów, miałyby być niemożliwe dla drugiego z nich. Zawsze powinien istnieć świat paralelny, w którym jeden z tych elektronów posiada dokładnie te same własności wewnętrzne i zewnętrzne. Takie paralelne światy byłyby również od siebie jakościowo nieodróżnialne. Cokolwiek bowiem działałoby się w świecie jednego z tych elektronów, będzie działało się w świecie drugiego z nich. W związku z tym dla każdej własności zrelatywizowanej do świata, którą posiada jeden z tych elektronów, istnieje ta sama własność zrelatywizowana do świata, którą posiada drugi z nich.

Skoro nie możemy odróżnić elektronów na podstawie ich własności, być może dałoby się przeprowadzić takie rozróżnienie na podstawie odmienności światów, w których posiadają one te własności. Wówczas posiadałyby odmienne własności zrelatywizowane do światów, w których istnieją, i można byłoby na tej podstawie dokonać ich odróżnienia. Widzieliśmy jednak, że paralelne światy są jakościowo nieodróżnialne. A przyjmując koncepcję Plantingi, nie sposób dokonać numerycznego odróżnienia jakościowo identycznych światów. Ich warunki identyczności są bowiem zależne od tego, jakie stany rzeczy składają się na nie. O światach  $w$  i  $w^*$  powiemy, że są identyczne wtw, gdy dla dowolnych stanów rzeczy  $s$  i  $s^*$ , jeśli  $s$  zawarte jest w  $w$ , to  $s$  zawarte jest w  $w^*$  i odwrotnie oraz jeśli  $w$  wyklucza się z  $s^*$ , to  $w^*$  wyklucza się z  $s^*$  i odwrotnie. Warunki identyczności stanów rzeczy są z kolei zależne od warunków identyczności ich elementów konstytutywnych i relacji, w jakich pozostają one do siebie. Innymi

słowy, jeśli stan rzeczy  $s$  i stan rzeczy  $s^*$  posiadają te same elementy i pozostają one do siebie w tej samej relacji, to  $s$  jest identyczne z  $s^*$ . Weźmy nasze paralelne światy  $w$  i  $w^*$ . Miałyby one różnić się od siebie jedynie tym, że elementem konstytutywnym niektórych spośród stanów rzeczy zawartych w świecie  $w$  jest indywidualna istota jednego z naszych elektronów, a elementem konstytutywnym niektórych spośród stanów rzeczy zawartych w świecie  $w^*$  jest indywidualna istota drugiego z nich. Widzieliśmy jednak, że w ramach ontologii Plantingi nie ma środków do tego, by odróżnić obie indywidualne istoty. Zatem światy  $w$  i  $w^*$  są identyczne, a w ramach ontologii Plantingi w ogóle nie mamy środków do reprezentowania możliwości istnienia dwóch jakościowo nieodróżnialnych elektronów. Nic nie wskazuje na to, by możliwość taka była iluzją, dlatego koncepcja Plantingi nie spełnia warunku ( $K_1$ ).

W ramach koncepcji Lewisa nie ma najmniejszego problemu z odróżnieniem jakościowo identycznych elektronów. Takie cechy posiadane są bowiem przez numerycznie odmienne konkretne indywidua, których warunki identyczności nie są wtórne względem tego, czy posiadają one te same własności. Co prawda, identyczność dwóch indywiduów pociąga jakościową nieodróżnialność, ale nie odwrotnie.

### 1.1.2 Teoria odpowiedników

Kolejnym elementem treści fikcji o wielości światów możliwych jest teoria odpowiedników. Jak już wspomniałem, jednym z podstawowych zastosowań dyskursu możliwoświatowego jest analiza pojęć modalnych *de dicto* i *de re*. Pierwsze są funktorami przyjmującymi za argumenty całe zdania, np. „Możliwe, że istnieją niebieskie łabędzie”, „Konieczne, że każdy kawaler jest nieżonatym mężczyzną”. Drugie są funktorami przyjmującymi za swoje argumenty predykaty, jak np. w zdaniach „Jan mógłby być kierowcą rajdowym” czy „Liczba dwa jest z konieczności liczbą pierwszą”. Twierdzenia modalne *de re* mówią

o tym, jaka jest kwalifikacja modalna sposobu przysługiwania własności danemu indywiduum bądź grupie indywiduów. Wyróżnia się własności przysługujące z konieczności oraz własności przysługujące przygodnie. Twierdzenia modalne *de dicto* mówią o tym, jaka jest kwalifikacja modalna twierdzeń poprzedzonych funktorem modalnym. Wyróżnia się tutaj twierdzenia koniecznie prawdziwe oraz twierdzenia przygodnie prawdziwe. Co uzasadnia poważne traktowanie tego *prima facie* obowiązującego rozróżnienia?

Jednym z powodów jego wprowadzenia jest to, że wzajemne implikowanie się twierdzeń modalnych zawierających pojęcia modalne *de dicto* oraz twierdzeń modalnych zawierających pojęcia modalne *de re* zachodzi jedynie na gruncie pewnego założenia, o którym intuicje podpowiadają nam, że jest ono jest fałszywe. Warunkiem koniecznym zachodzenia wzajemnego implikowania się tych twierdzeń jest prawdziwość następującej równoważności:

(BF) Dla każdego  $x$ , koniecznie  $Px$  wtw, gdy koniecznie dla każdego  $x$ ,  $Px$ ,

która stanowi koniunkcję tzw. formuły Barcan (jeśli dla każdego  $x$ , koniecznie  $Px$ , to koniecznie dla każdego  $x$ ,  $Px$ ) (zob. Barcan 1946) oraz tzw. odwróconej formuły Barcan (jeśli koniecznie dla każdego  $x$ ,  $Px$ , to dla każdego  $x$ , koniecznie  $Px$ ). Zwykle utrzymuje się, że obie formuły są prawdziwe jedynie wówczas, gdy założymy, że dziedzina przedmiotowa  $d$  jest niezmienna dla wszystkich światów, tj. dla dwóch dowolnych światów możliwych  $w$  i  $w^*$ ,  $d(w) = d(w^*)$ , a więc, że we wszystkich światach istnieją te same przedmioty (zob. Girle 2003, s. 66–67). Załóżmy, że istnieją tylko trzy światy  $w$ ,  $w^*$  oraz  $w^{**}$  takie, że  $d(w) = \{a, b, c\}$ ,  $d(w^*) = a, b, c$  i  $d(w^{**}) = \{a, b, c\}$ . Załóżmy również, że relacja osiągalności pomiędzy tymi światami jest zwrotna, symetryczna, przechodnia i euklideska. Niech twierdzenie „Dla każdego  $x$ , koniecznie  $Px$ ” będzie prawdziwe w świecie  $w$  wtw, gdy dla każdego  $x$ , jeśli  $x$  istnieje w  $w$ , to dla każdego świata  $v$  osiągalnego ze świata

$w$ , to  $Px$  w  $v$ . Z kolei twierdzenie „Koniecznie dla każdego  $x$ ,  $Px$ ” jest prawdziwe w świecie  $w$  wtw, gdy dla dowolnego świata  $v$  osiągalnego ze świata  $w$ , dla każdego  $x$ ,  $Px$  w  $v$ . Jako że przyjęliśmy założenie, iż dziedzina jest niezmienna dla wszystkich światów, to wszystkie indywidua, które zmienna  $x$  występująca po lewej stronie równoważności (BF) przyjmuje jako wartości, istnieją w każdym świecie osiągalnym ze świata  $w$ . Zatem jeśli lewa strona równoważności (BF) jest prawdziwa w świecie  $w$ , to jest prawdą, że wszystkie te indywidua są  $P$  w każdym świecie osiągalnym ze świata  $w$ . Z uwagi na niezmienność dziedziny przedmiotowej w każdym świecie jest prawdą, że wszystkie indywidua, istniejące w tym świecie, są  $P$ . Stąd prawdziwa jest prawa strona równoważności BF. Zatem lewa strona równoważności (BF) materialnie implikuje prawą. Z uwagi na to, że implikacja ta jest prawdziwa we wszystkich z naszych światów, możemy w sposób nie budzący wątpliwości zastosować regułę konieczności do tej materialnej implikacji i twierdzić, że lewa strona równoważności (BF) pociąga za sobą prawą, ponieważ pierwsza ściśle implikuje drugą: Koniecznie, jeśli dla każdego  $x$ , koniecznie  $Px$ , to koniecznie dla każdego  $x$ ,  $Px$ ). Poprzednik odwróconej formuły Barcan jest prawdziwy, ponieważ dla dowolnego świata  $v$  osiągalnego ze świata  $w$ , dla każdego  $x$ ,  $Px$  w  $v$ . Następnik zaś jest prawdziwy, ponieważ — z uwagi na niezmienność dziedziny — dla każdego  $x$ , jeśli  $x$  istnieje w świecie  $w$ , dla dowolnego świata  $v$  osiągalnego z  $w$ ,  $Px$  w  $v$ . Ponownie, stosując regułę konieczności do odwróconej formuły Barcan, możemy powiedzieć, że poprzednik pociąga za sobą następnik, ponieważ pierwszy ściśle implikuje drugi. W rezultacie na gruncie założenia, że dziedzina przedmiotowa jest niezmienna dla wszystkich światów, obie strony równoważności (BF) wzajemnie się implikują.

Gdy jednak porzucimy ostatnie założenie, fałszywa będzie odwrócona formuła Barcan. Załóżmy, że mamy trzy światy  $w$ ,  $w^*$ ,  $w^{**}$  takie, że  $d(w) = \{a, b, c\}$ ,  $d(w^*) = \{a, b, d\}$   $d(w^{**}) = \{d\}$ . Załóżmy,

że w każdym z naszych światów prawdą jest, że każde indywiduum istniejące w tym świecie jest  $P$ . Prawdziwy jest zatem poprzednik odwróconej formuły Barcan. Jednak nie wszystkie indywidua istniejące w świecie  $w$  istnieją we wszystkich światach osiągalnych ze świata  $w$ . Nie jest więc prawdą, iż wszystkie indywidua istniejące w świecie  $w$  są  $P$  w każdym ze światów osiągalnych ze świata  $w$ . Zatem fałszywy jest następnik odwróconej formuły Barcan, co — przy prawdziwości jej poprzednika — pociąga za sobą fałszywość odwróconej formuły Barcan. Fałszywość ostatniej pociąga za sobą z kolei równoważność (BF), która jest koniecznym warunkiem wzajemnego implikowania się zdań modalnych *de dicto* i zdań modalnych *de re*.

Gdyby ktoś protestował przeciwko zaproponowanym wyżej warunkom prawdziwości dla twierdzenia „Dla każdego  $x$ , koniecznie  $Px$ ” i proponował, by uznać je za prawdziwe w świecie  $w$  wtw, gdy dla każdego  $x$ , jeśli  $x$  istnieje w  $w$ , to dla każdego świata  $v$  osiągalnego ze świata  $w$ , jeśli  $x$  istnieje w  $v$ , to  $Px$  w  $v$ , wówczas formuła Barcan może być prawdziwa jedynie na gruncie założenia, że dziedziny przedmiotowe światów są niezmiennie. Załóżmy bowiem, że mamy trzy światy  $w$ ,  $w^*$  oraz  $w^{**}$  takie, że  $d(w) = \{a, b\}$ ,  $d(w^*) = \{a, b, c\}$ ,  $d(w^{**}) = \{d\}$  i załóżmy, że „Dla każdego  $x$ , koniecznie  $Px$ ” jest prawdziwe w świecie  $w$ , wówczas będzie ono również prawdziwe w świecie  $w^*$  (na mocy istnienia indywiduów  $a$  i  $b$  w tym świecie) i  $w^{**}$  (na mocy trywialnej prawdziwości okresu warunkowego „jeśli  $x$  istnieje w  $v$ , to  $Px$  w  $v$ ), jeśli jednak indywidua  $c$  i  $d$  są takie, że żadne z nich nie jest  $P$ , wówczas fałszywy będzie następnik formuły Barcan indexBarcan R., co — przy prawdziwości jej poprzednika — sprawia, iż jest ona fałszywa. Celem nadania wiarygodności ostatniej tezie, załóżmy, że indywidua  $a$  i  $b$  są ludźmi, a predykat  $P$  jest równoznaczny z predykatem „jest rozumny”. Jeśli zgodzić się, że ludzie z konieczności są rozumni, to twierdzenie, że dla każdego  $x$ , z konieczności,  $x$  jest rozumne, będzie prawdziwe w świecie  $w$ . Jeśli jednak indywidua  $c$  i  $d$  nie są ludźmi, a

— powiedzmy — psami, to w światach  $w^*$  i  $w^{**}$  nie będzie prawdą, że wszystkie indywidua są rozumne, co jest wystarczającą podstawą do odrzucenia prawdziwości twierdzenia, że koniecznie, dla każdego  $x$ ,  $x$  jest rozumne. Fałszywość formuły Barcan pociąga za sobą oczywiście fałszywość równoważności (BF).

Fałszywość tej równoważności daje z kolei powód do uznania różniczenia na te dwa typy modalności za obowiązujące, pod warunkiem że światy mogą mieć odmienne dziedziny przedmiotowe. Na rzecz ostatniego założenia przemawiają silne intuicje modalne co do tego, że mogłoby istnieć więcej obiektów, niż aktualnie istnieje. Przyjmując jakąkolwiek formę realizmu w sprawie światów możliwych, powiemy, że intuicja ta co najmniej oznacza, iż istnieją takie światy  $w$  i  $v$ , że dziedzina pierwszego z nich jest równoliczna z podzbiorem dziedziny drugiego z nich, co równoważne jest porzuceniu warunku niezmienności dziedzin przedmiotowych dla wszystkich światów.

W dotychczasowej dyskusji o relacji modalności *de dicto* do modalności *de re* przyjmowaliśmy założenie, że istnieją indywidua, które istnieją w wielu światach możliwych, tymczasem Lewis jest zwolennikiem tezy o wzajemnym odizolowaniu światów możliwych. Skoro światy są maksymalnymi sumami mereologicznymi powiązanymi relacjami czasoprzestrzennymi bądź analogicznymi do nich, żadne indywiduum nie powinno być częścią dwóch światów. W przeciwnym razie przechodniość relacji czasoprzestrzennych (bądź analogicznych do nich) doprowadziłaby do zespolenia wszystkich tych światów, które (a) mają to samo indywiduum za swoją część i (b) których części powiązane są relacjami dokładnie tego samego typu. Załóżmy, że indywidua w światach  $w$  i  $w^*$  powiązane są standardowymi relacjami czasoprzestrzennymi. Niech  $a$  będzie wspólną częścią światów  $w$  i  $w^*$ ; a pozostaje w relacjach czasoprzestrzennych zarówno z częściami świata  $w$ , jak i świata  $w^*$ . Relacje te są przechodnie, więc dla dowolnego  $x$  i  $y$ , jeśli  $x$  pozostaje w relacji czasoprzestrzennej do  $a$ ,  $y$  pozostaje w

relacji czasoprzestrzennej do  $a$  i  $x$  istnieje w świecie  $w$ , a  $y$  w świecie  $w^*$ , to na mocy przechodniości relacji czasoprzestrzennych  $x$  pozostaje w tej relacji do  $y$ . Wówczas jednak ani  $w$ , ani  $w^*$  nie jest maksymalną sumą mereologiczną części powiązanych relacjami czasoprzestrzennymi. Istnieją bowiem byty, które pozostają do części tej sumy w relacji czasoprzestrzennej, ale z założenia nie są ich częściami. Jeśli chcemy uniknąć spajania odmiennych światów możliwych o tej samej czasoprzestrzeni, powinniśmy odrzucić istnienie indywiduów, które są ich wspólnymi częściami. Uniknięcie tej konsekwencji jest o tyle ważne, że większość z reprezentowanych przez nas możliwości dotyczy właśnie światów, które nie różnią się pod względem czasoprzestrzeni od naszego własnego świata.

Jak w takim razie powinniśmy interpretować warunki prawdziwości twierdzeń modalnych *de re*? Standardowo o twierdzeniu modalnym *de re* „Dla każdego  $x$ , koniecznie  $Px$ ” mówiliśmy, że jest prawdziwe w świecie  $w$  wtw, gdy dla każdego  $x$ , jeśli  $x$  istnieje w świecie  $w$ , to dla każdego świata  $v$  osiągalnego ze świata  $w$ ,  $x$  jest  $P$  w  $v$ , bądź jest prawdziwe w świecie  $w$  wtw, gdy dla każdego  $x$  istniejącego w świecie  $w$ , dla każdego świata  $v$  osiągalnego ze świata  $w$ , jeśli  $x$  istnieje w  $v$ , to  $x$  jest  $P$  w  $v$ . Nie mając jednak indywiduów, które istniałyby w wielu światach możliwych, musimy inaczej zdefiniować warunki prawdziwości twierdzeń modalnych *de re*. Identyczność indywiduów w poprzek światów możliwych nie może już odgrywać roli przy wartościowaniu zdań modalnych *de re*, tj. wartościując zdania modalne *de re*, nie sprawdzamy już, czy indywidua przyporządkowane w świecie  $w$  jako wartości zmiennych związanych występujących w zdaniach modalnych *de re* posiadają pewną własność w światach osiągalnych ze świata  $w$ , jeśli w nich istnieją. Lewis proponuje zastąpić relacje identyczności w poprzek światów relacją bycia odpowiednikiem:

Relacja bycia odpowiednikiem jest naszym substytutem dla relacji identyczności pomiędzy rzeczami z różnych świa-

tów. Gdy niektórzy powiedzieliby, że istniejesz w różnych światach, w których masz nieco odmienne własności i przydarzają ci się nieco odmienne rzeczy, wolę powiedzieć, że istniejesz w świecie aktualnym i w żadnym innym, ale masz odpowiedniki w wielu światach możliwych. Twoje odpowiedniki silnie przypominają cię pod względem istotnych aspektów treści, jak i kontekstu. Przypominają cię bardziej niż inne rzeczy w ich światach, ale naprawdę nie są tobą, ponieważ każdy z nich jest w swoim świecie, a ty jesteś tutaj — w świecie aktualnym (Lewis 1968/1983, s. 27–28).

Relacja bycia odpowiednikiem jest więc relacją podobieństwa zarówno ze względu na własności zewnętrzne (aspekty kontekstu), jak i wewnętrzne (aspekty treści).

Jak utrzymuje Lewis, taka relacja podobieństwa — w odróżnieniu od relacji osiągalności będącej relacją podobieństwa ze względu na pewien określony aspekt — w ogólności nie jest relacją przechodnią (tj. z faktu, iż pewne indywiduum  $x$  istniejące w świecie  $w^*$  jest podobne do mnie bardziej niż jakiegokolwiek inne indywiduum w świecie  $w^*$  oraz że pewne indywiduum  $y$  istniejące w świecie  $w^{**}$  jest bardziej podobne do indywiduum  $x$  niż jakiegokolwiek inne indywiduum ze świata  $w^{**}$ , nie wynika jeszcze, że w świecie  $w^{**}$  nie istnieje takie indywiduum  $z$ , które byłoby bardziej podobne do mnie niż indywiduum  $y$ ). Owa relacja nie jest też relacją symetryczną (tj. choć indywiduum  $x$  istniejące w świecie  $w^*$  przypomina mnie najbardziej ze wszystkich indywiduów istniejących w świecie  $w^*$ , to wciąż może istnieć indywiduum  $y$  istniejące w świecie aktualnym, które bardziej przypomina indywiduum  $x$  niż ja). Co więcej, niewiarygodny jest również postulat, że dla dowolnego świata  $w$ , istnieje tylko jeden mój odpowiednik, jeśli w ogóle istnieją w tym świecie indywidua przypominające mnie. Przeciwnie, nic nie wyklucza istnienia świata, w którym istnieją indywidua  $x$  i  $y$ ,



z których żadne z nich nie przypomina mnie bardziej od drugiego z nich, a każde z nich przypomina mnie bardziej niż wszystkie pozostałe indywidua istniejące w tym świecie. W takich okolicznościach oba te indywidua będą traktowane jako moje odpowiedniki. Z podobnych powodów nic nie przekreśla istnienia świata możliwego, w którym pewne indywiduum  $x$  jest wspólnym odpowiednikiem dwóch indywiduów istniejących w świecie aktualnym. Wreszcie, nic nie przekreśla istnienia świata, w którym pewne indywidua istniejące w świecie aktualnym — powiedzmy ludzie — nie mają odpowiedników w tym świecie, ponieważ w świecie tym w ogóle nie istnieją ludzie (Lewis 1968/1983, s. 28–29).

Obok hipotezy, że relacja bycia odpowiednikiem jest relacją podobieństwa, wczesna teoria odpowiedników obejmowała następujące postulaty dotyczące tej relacji:

- (O<sub>1</sub>) Dla każdego  $x, y$ , jeśli  $x$  jest odpowiednikiem  $y$ , to istnieje taki świat  $w$ , że  $y$  istnieje w świecie  $w$ .
- (O<sub>2</sub>) Dla każdego  $x, y$ , jeśli  $x$  jest odpowiednikiem  $y$ , to istnieje taki świat  $w$ , że  $x$  istnieje w świecie  $w$ .
- (O<sub>3</sub>) Dla każdego  $x, y$  i dla każdego świata  $w$ , jeśli  $x$  istnieje w  $w$  i  $y$  istnieje w świecie  $w$  i  $x$  jest odpowiednikiem  $y$ , to  $x = y$ .
- (O<sub>4</sub>) Dla każdego  $x$ , dla każdego  $w$ , jeśli  $x$  istnieje w  $w$ , to  $x$  jest swoim własnym odpowiednikiem (Lewis 1968/1983, s. 27).

Postulaty (O<sub>1</sub>), (O<sub>2</sub>) oraz (O<sub>4</sub>) raczej nie budzą wątpliwości. Dwa pierwsze ustalają jedynie dziedzinę i przeciwdziedzinę relacji bycia odpowiednikiem w taki sposób, że, cokolwiek jest elementem pierwszej, jest indywiduum istniejącym w jakimś świecie, a cokolwiek jest elementem drugiej, również jest indywiduum istniejącym w pewnym świecie. Z uwagi na to, że relacja bycia odpowiednikiem jest wprowadzona jako substytut dla relacji identyczności, która miała zachodzić

między indywiduami istniejącymi w różnych światach, ograniczenie pola relacji bycia odpowiednikiem do indywiduów istniejących w światach możliwych jest zrozumiałe samo przez się. Z kolei ostatni postulat wydaje się po prostu konsekwencją znaczenia pojęcia bliskiego podobieństwa ze względu na własności wewnętrzne i zewnętrzne — każde indywiduum jest podobne do siebie samego ze względu na własności wewnętrzne i zewnętrzne.

Nim przejdziemy do postulatu ( $O_3$ ), przedstawmy krótko Lewisa analizę warunków prawdziwości twierdzeń modalnych *de re*:

Dla każdego  $x$ , koniecznie  $Px$  jest prawdziwe w świecie  $w$  wtw, gdy dla każdego świata  $v$  (osiągalnego ze świata  $w$ ), dla każdego  $x$ , istnieje takie  $y$ , że jeśli  $x$  istnieje w świecie  $w$ ,  $y$  istnieje w świecie  $v$  i  $y$  jest odpowiednikiem  $x$ , to  $Px$  w  $v$ .

Istnieje takie  $x$ , że koniecznie  $Px$  jest prawdziwe w świecie  $w$  wtw, gdy dla każdego świata  $v$ , dla pewnego  $x$ , istnieje takie  $y$ , że  $x$  istnieje w świecie  $w$  i  $y$  istnieje w świecie  $v$ , i  $y$  jest odpowiednikiem  $x$ , i  $Px$  w  $v$ .

Dla każdego  $x$ , możliwie  $Px$  jest prawdziwe w świecie  $w$  wtw, gdy istnieje taki świat  $v$ , że dla każdego  $x$ , istnieje takie  $y$ , że  $x$  istnieje w świecie  $w$  i  $y$  istnieje w świecie  $v$ , i  $y$  jest odpowiednikiem  $x$ , to  $Px$  w  $v$ .

Istnieje takie  $x$ , że możliwie  $Px$  jest prawdziwe w świecie  $w$  wtw, gdy istnieje taki świat  $v$ , dla pewnego  $x$ , istnieje takie  $y$ , że  $x$  istnieje w świecie  $w$  i  $y$  istnieje w świecie  $v$ , i  $y$  jest odpowiednikiem  $x$ , i  $Px$  w  $v$  (por. Lewis 1968/1983, s. 31).

Bazując na powyższych przykładach, można w następujący sposób pokazać, iż — na gruncie tezy o wzajemnym odizolowaniu światów — akceptacja postulatu ( $O_3$ ) albo prowadzi do przyjęcia istnienia jakościowo nieodróżnialnych światów możliwych, jeśli uważa się, iż

indywidua mogłyby wymieniać się rolami, albo prowadzi do odrzucenia możliwości tego, by pewne indywidua mogłyby się wymienić rolami, jeśli odrzuci się istnienie jakościowo nieodróżnialnych światów możliwych.

Założmy, że indywidua  $x_1$  i  $x_2$  mogłyby wymienić się rolami, tj. dla pewnego indywiduum  $x_1$ , możliwe jest, iż indywiduum to posiada wszystkie własności wewnętrzne i zewnętrzne posiadane aktualnie przez indywiduum  $x_2$ . Możliwość wymienienia się rolami przez indywidua wydaje się oddawać przekonanie, iż jest logicznie możliwe, bym żył życiem innej osoby, tj. posiadał wszystkie własności, które posiada ta osoba, pozostawał w takich relacjach do innych osób, w jakich pozostaje ta osoba itd. Zgodnie z wprowadzonymi wyżej schematami przekładu zdań modalnych *de re*, powinniśmy powiedzieć, że twierdzenie modalne *de re* „ $x_1$  mogłoby posiadać wszystkie własności wewnętrzne i zewnętrzne  $x_2$ ” jest prawdziwe w świecie aktualnym wtw, gdy istnieje takie  $x_1$ ,  $x_2$ ,  $x_3$  i  $x_4$ , istnieje taki świat  $v$ , że  $x_1$  i  $x_2$  istnieją w świecie aktualnym,  $x_3$  i  $x_4$  istnieją w świecie  $v$ ,  $x_3$  jest odpowiednikiem  $x_1$  i  $x_4$  jest odpowiednikiem  $x_2$  i dla każdej własności  $F$ , jeśli  $Fx_1$ , to  $Fx_3$  i dla każdej własności  $G$ , jeśli  $Gx_2$ , to  $Gx_4$ . Zgodnie z postulatem ( $O_3$ ) żadne indywiduum nie posiada żadnego odpowiednika w świecie, w którym istnieje, który byłby różny od tego indywiduum. Zatem jeśli indywidua  $x_1$  i  $x_2$  mogą wymieniać się rolami i ( $O_3$ ) obowiązuje, to  $w \neq v$ . Przyjęcie istnienia jakościowo nieodróżnialnych światów wraz z tezą o rozłączności ich dziedzin przedmiotowych sprawia, że relacja bycia odpowiednikiem przestaje być relacją podobieństwa ze względu na własności wewnętrzne i zewnętrzne. Gdyby bowiem była to relacja podobieństwa, powinna raczej zachodzić pomiędzy  $x_1$  a  $x_4$  oraz pomiędzy  $x_2$  a  $x_3$ . Powstaje więc pytanie: czym jest relacja bycia odpowiednikiem, jeśli nie jest relacją podobieństwa ze względu na wewnętrzne i zewnętrzne własności? Musiałaby być jakąś relacją, która nie zachodzi pomiędzy indywidu-

ami z odmiennych światów na podstawie tego, jakie posiadają one własności. Zdaniem Lewis *index* Lewis D.a, jeśli jednak nie zachodzi na podstawie własności posiadanych przez członki tej relacji, to staje się czymś tajemniczym, ze względu na co ta relacja zachodzi, przez co tajemnicza staje się również jej natura (Lewis 1986, s. 229–230). Czy można jakoś uniknąć konieczności wprowadzania relacji bycia odpowiednikiem nie zachodzącej ze względu na własności wewnętrzne i zewnętrzne jej członków?

Możemy tutaj wykonać jeden z dwóch kroków: albo możemy odrzucić tezę o wzajemnym odizolowaniu światów i przyjąć, że mogłyby istnieć indywidua stanowiące części wielu światów możliwych, albo odrzucić postulat ( $O_3$ ). Akceptując tezę o wzajemnym odizolowaniu światów możliwych, Lewis wybiera odrzucenie postulatu ( $O_3$ ) i dopuszcza sytuacje, w których indywiduum posiada odpowiednik (różny od tego indywiduum) istniejący w tym samym świecie (zob. Lewis 1968/1986, s. 43; Lewis 1986, s. 230–232, 232 — przyp. 22), a w związku z tym to, co mogłoby przydarzyć się pierwszemu, może być reprezentowane przez to, co przydarza się drugiemu z tych indywiduów, przy czym oba mogą istnieć w tym samym świecie. Jeśli  $x_1$  i  $x_2$  mogą pozostawać do siebie w relacji bycia odpowiednikiem, pomimo iż istnieją w tym samym świecie, nie musimy przyjmować istnienia świata  $v$  jakościowo nieodróżnialnego od świata  $w$ , ponieważ możliwość wymienienia się rolami przez indywidua  $x_1$  i  $x_2$  jest reprezentowana przez to, co aktualnie przydarza się obu z nich, a nie przez to, co przydarza się jakimś dwóm jakościowo od nich nieodróżnialnym indywiduom, istniejącym w jakimś innym świecie możliwym. Oczywiście narzuca się pytanie, ze względu na jakie własności wewnętrzne i zewnętrzne zachodzi relacja bycia odpowiednikiem. Lewis odpowiada, że jest to relacja bycia odpowiednikiem ze względu na własność bycia rzeczami tego samego rodzaju (Lewis 1986, s. 232). Na gruncie takiej relacji bycia odpowiednikiem poza indywiduum  $x_2$  również inne indywidua

są odpowiednikami indywiduum  $x_1$ , ponieważ jedyny warunek, jaki nakłada się na indywiduum, będące odpowiednikiem indywiduum  $x_1$ , jest taki, by było ono rzeczą tego samego rodzaju. Wciąż jednak relacja bycia odpowiednikiem zachodzi pomiędzy danym indywiduum a obiektem najbardziej podobnym do niego ze względu na pewną klasę własności. W tym wypadku odpowiednikami są wszystkie indywidua, które są najbardziej podobne do naszego indywiduum ze względu na klasę własności przyporządkowanych rodzajowi obiektów, do którego należy nasze indywiduum. W przypadku bogatszej klasy własności odpowiednio mniej indywiduów będzie pozostawało w relacji bycia odpowiednikiem do naszego indywiduum. I nie ma nic dziwnego w tym, że przeciwdziedzina relacji podobieństwa może zmieniać się w zależności od tego, ze względu na jakie własności relacja ta zachodzi.

Ostatnie jest szczególnie istotnie, gdyby ktoś chciał upatrywać w teorii odpowiedników Lewisa odrodzenia esencjalistycznej metafizyki Arystotelesa, tj. upatrywać w teorii odpowiedników koncepcji, która pozwoli w kategoriach światów możliwych oraz relacji bycia odpowiednikiem wyjaśnić różnicę pomiędzy własnościami przysługującymi w sposób istotny i przygodny. Co prawda, można pokusić się o następujące definicje własności istotnie oraz przygodnie przysługujących:

Własność  $F$  przysługuje indywiduum  $x$  w sposób istotny wtw, gdy dla każdego  $y$ , jeśli  $y$  jest odpowiednikiem indywiduum  $x$ , to  $Fy$ .

Własność  $F$  przysługuje indywiduum  $x$  w sposób przygodny wtw, gdy istnieje takie  $y$ , że  $y$  jest odpowiednikiem  $x$  i  $Fy$ , i istnieje takie  $z$ , że  $z$  jest odpowiednikiem  $x$  i  $Fz$  (por. Lewis 1968/1983, s. 35).

Dopóki jednak nie wyróżnimy jednej z wielu relacji podobieństwa, które wyznaczają odmienne klasy odpowiedników danego indywiduum,

dopóty nie powiemy niczego o esencjalizmie w klasycznym sensie. Zgodnie z relacją podobieństwa ze względu na rodzaj naturalny istotną cechą naszego indywiduum będzie to, że posiada ono własności kojarzone z jego rodzajem naturalnym. Zgodnie z relacją podobieństwa ze względu na biologiczne pochodzenie istotną cechą naszego indywiduum będzie posiadanie rodziców o takim a takim materiale genetycznym. Wreszcie, w ekstremalnych przypadkach relacje podobieństwa mogą nie nakładać praktycznie żadnych wymogów na to, jakie własności powinien posiadać odpowiednik, a wówczas każde indywiduum byłoby odpowiednikiem każdego innego indywiduum (Lewis 1968/1983, s. 43). Wtedy jedynymi istotnymi własnościami naszego indywiduum są własności, które posiada każde indywiduum. Na przeciwległym biegunie będą znajdowały się takie „hiperesencjalistyczne” relacje podobieństwa, że jedynym odpowiednikiem danego indywiduum będzie ono samo, a wówczas każda własność tego indywiduum będzie jego własnością istotną (Lewis 1968/1983, s. 43). Z pewnością na podstawie dwóch ostatnich relacji podobieństwa wyznaczających klasę odpowiedników  $x$ -a — odpowiednio — wszystkie indywidua bądź wyłącznie indywiduum  $x$  nie sformułowano by definicji własności istotnie lub przygodnie przysługujących, które oddawałyby intuicje zwolenników klasycznego esencjalizmu. Być może relacja podobieństwa ze względu na rodzaj naturalny byłaby bliższa intuicjom kryjącym się za esencjalizmem, a być może istotne jest również biologiczne pochodzenie danego indywiduum. Sama teoria odpowiedników może jedynie dostarczyć środków do zdefiniowania własności istotnych i przygodnych ze względu na pewną relację podobieństwa, nie mówi już jednak niczego o tym, która z relacji podobieństwa jest relacją odpowiednią do rekonstruowania rozróżnień właściwych klasycznemu esencjalizmowi.

Relacja bycia odpowiednikiem dziedziczy po relacji podobieństwa niezdeterminowanie właściwe ostatniej — niezdeterminowanie co do własności, ze względu na które podobieństwo ma zachodzić; co do

tę, jaka jest waga jednej z tych własności względem drugiej; co do tego, jakie są wymagane minimalne standardy podobieństwa, wreszcie — co do zakresu, w jakim można eliminować z klasy odpowiedników indywidua w razie istnienia indywiduów bardziej podobnych (Lewis 1968/1983, s. 42). Zwykle kontekst, w którym stwierdza się twierdzenie modalne *de re*, pozwala rozwikłać powyższe aspekty niezdecydowania. Gdy powiem, że Marysia nie mogła zjeść całego kotleta, prawdopodobnie będę miał na myśli relację podobieństwa do Marysi ze względu na jej konstytucję fizjologiczną (w szczególności wielkość żołądka), jej nawyki żywieniowe itd. Gdy jednak powiem, że Marysia mogłaby zjeść nawet trzy kotlety, będę odwoływał się do innej relacji podobieństwa zachodzącej pomiędzy Marysią a jej odpowiednikami. Prawdopodobnie będzie to relacja podobieństwa ze względu na bycie kobietą w takim i takim przedziale wiekowym. Taka kontekstowa wrażliwość relacji bycia odpowiednikiem ponownie pokazuje, że sama teoria odpowiedników nie stanowi odrodzenia klasycznego esencjalizmu.

Wracając do kwestii odróżnienia modalności *de re* i modalności *de dicto*, przytoczony wyżej argument na rzecz prawomocności tego odróżnienia można odtworzyć na gruncie teorii odpowiedników, gdy w miejsce założenia o identyczności dziedzin przedmiotowych światów możliwych przyjmie się założenie, że dla dwóch dowolnych światów możliwych  $w$  i  $v$ , dla każdego  $x$ , jeśli  $x$  należy do  $d(w)$ , istnieje takie  $y$ , że  $y$  należy do  $d(v)$  i  $y$  jest odpowiednikiem  $x$ , i odwrotnie. Wówczas twierdzenia modalne *de re* będą pociągały za sobą twierdzenia modalne *de dicto* i odwrotnie. Gdy jednak odrzuci się powyższe założenie, to — z uwagi na możliwość nieistnienia odpowiedników danego indywiduum w innych światach — formuła Barcan nie będzie obowiązywała dla wszystkich relacji podobieństwa. Choć może być tak, że w każdym świecie możliwym  $v$  wszystkie indywidua istniejące w tym świecie są takie, że wszystkie odpowiedniki tych indywiduów

istniejące w innych światach są  $P$  — to jeszcze nie wynika z tego, że dla dowolnego świata  $v$ , dla każdego indywiduum  $x$  istniejącego w  $v$ ,  $x$  jest  $P$ ;  $x$  może bowiem nie być odpowiednikiem żadnego z tych indywiduów.

Podsumowując, treścią fikcji o wielości światów możliwych jest realistyczna koncepcja światów możliwych w wydaniu D. Lewisa oraz jego teoria odpowiedników.

## 1.2 Metafikcjonalistyczne zastosowania fikcji realisty modalnego

Podobnie jak realista modalny metafikcjonalista modalny również jest zainteresowany zachowaniem możliwościwiatowej analizy pojęć modalnych bądź zachowaniem możliwości przechodzenia od twierdzeń modalnych do twierdzeń możliwościwiatowych i odwrotnie, z tym jednakże wyjątkiem, że wszelkie odniesienia do światów możliwych lub możliwych indywiduów będą dokonywały się w zasięgu funktora opowieści. Stąd też, proponuje się, by twierdzenia modalne analizować w ramach schematu:

(MMF)  $p \equiv \text{Zgodnie z } PW, p^*$ ,

gdzie „ $p$ ” jest dowolnym twierdzeniem modalnym, natomiast „ $p^*$ ” jest realistycznym przekładem twierdzenia modalnego „ $p$ ” na twierdzenie o światach możliwych lub możliwych indywiduach.

W zależności od sposobu, na jaki metafikcjonalista modalny interpretuje schemat (MMF), można wyróżnić dwie wersje metafikcjonalizmu modalnego. Zgodnie z jego *silną* odmianą przed metafikcjonalizmem modalnym stawia się dwa cele: z jednej strony, ma umożliwić mówienie o światach możliwych bez przyjmowania ich istnienia, a z drugiej strony, ma dostarczyć redukcijnej analizy warunków prawdziwości twierdzeń modalnych (zob. Rosen 1990, s. 353–354; Nolan 1997, s. 263). W związku z tym, prawa strona schematu (MMF) jest



interpretowana jako redukcyjna analiza warunków prawdziwości twierdzeń występujących po jego lewej stronie. Zgodnie z *umiarkowaną* wersją metafikcjonalizmu modalnego zasadniczym celem tego stanowiska jest zachowanie możliwości mówienia o światach możliwych bez przyjmowania ich istnienia, ale już nie stawia się mu celu dostarczenia redukcyjnej analizy warunków prawdziwości twierdzeń modalnych. Co prawda, umiarkowany fikcjonalista modalny może dopuszczać przechodzenie od prawd modalnych do twierdzeń możliwościwiatowych i odwrotnie, ale jego teoria nie będzie analizą twierdzeń modalnych, a jedynie teorią „łączącą modalne fakty z faktami o opowieści *PW*” (Rosen 1990, s. 354). W ramach tego stanowiska to „twierdzenia dotyczące światów możliwych (albo, by być precyzyjniejszym, tego, jak światy możliwe są opisywane w fikcji) opierają się raczej na faktach dotyczących modalności niż odwrotnie” (Nolan 1997, s. 262). Zatem schemat (MMF) pełni jedynie rolę ustalania wzajemnych zależności pomiędzy wartościami logicznymi twierdzeń modalnych i twierdzeń możliwościwiatowych poprzedzonych funktorem opowieści „Zgodnie z *PW*” bez dostarczania redukcyjnej analizy warunków prawdziwości którejkolwiek z nich. Dystynktywną cechą umiarkowanego metafikcjonalizmu modalnego jest również akceptacja obiektywnego istnienia faktów modalnych, które fundują prawdziwość twierdzeń modalnych i od których zależna jest również prawdziwość (zgodnie z fikcją *PW*) twierdzeń możliwościwiatowych. Oddanie ostatniej zależności będzie wymagało wprowadzenia odpowiednich praw pomostowych. Zauważmy, że gdyby umiarkowany metafikcjonalista modalny nie wprowadził praw pomostowych, czymś zupełnie niejasnym byłoby to, dlaczego w ogóle mielibyśmy być zainteresowani zachowaniem możliwości mówienia o światach możliwych. Powodem, dla którego umiarkowani metafikcjonałiści modalni chcą zachować możliwość posługiwania się dyskursem możliwościwiatowym, zwykle jest to, że w ich przekonaniu dostarcza on narzędzi do owocnego opisu faktów modalnych. Do

kwestii tej wróć w §4.6<sup>28</sup>.

## 2 Przedmiotowy fikcjonalizm modalny

Przedmiotowy fikcjonalizm modalny występuje wyłącznie w umiarkowanej wersji, zgodnie z którą przyjmuje się istnienie faktów modalnych. Odwołując się do gier w udawanie, przedmiotowy fikcjonalista modalny dokonuje następującej charakterystyki dyskursu możliwościowego:

[Twierdzenie — P.W.] „Istnieją światy, w których istnieją niebieskie łabędzie” sugeruje grę, w której symulujemy, że wszystkie i tylko te rzeczy, które *mogłyby* mieć miejsce w tym świecie, *dzieją się* w jakimś świecie. Celem tej wypowiedzi jest powiedzenie, że modalne fakty są takie, że sprawiają, iż [twierdzenie — P.W.] „istnieją światy niebieskich łabędzi” jest warte symulacji w grze (Yablo 1996, s. 279).

---

<sup>28</sup>Niekiedy wyróżnia się również szeroki fikcjonalizm modalny (zob. Nolan 1997, s. 261–262), na który składają się trzy tezy: (a) twierdzenia możliwościowe są akceptowane jako prawdziwe zgodnie z fikcją *PW*, a istnienie samych światów możliwych jest odrzucane, bądź sąd w tej sprawie jest zawieszany; (b) twierdzenia modalne również są twierdzeniami akceptowanymi jako prawdziwe zgodnie z pewną modalną fikcją, a istnienie faktów modalnych jest odrzucane bądź sąd w tej sprawie jest zawieszany oraz (c) przechodzenie od twierdzeń prawdziwych zgodnie z modalną fikcją do twierdzeń prawdziwych zgodnie z hipotezą *PW* jest umożliwiane dzięki schematowi: zgodnie z modalną fikcją, *p* wtw, gdy zgodnie z hipotezą *PW*, *p*<sup>\*</sup>. Przy czym zakłada się, że żadna ze stron tego schematu nie ma stanowić redukcijnej analizy warunków prawdziwości drugiej z nich. Szerokim fikcjonalizmem modalnym mogą być zainteresowani wszyscy filozofowie, którzy są modalnymi sceptykami, twierzącymi, iż nasze zdolności kognitywne nie pozwalają nam na poznawanie faktów modalnych. Pogląd taki jest bez wątpienia interesującym stanowiskiem filozoficznym, jednakże z dwóch powodów nie będzie dyskutowany w tej rozprawie. Po pierwsze, nie ma wśród fikcjonalistów modalnych zdeklarowanych zwolenników tego stanowiska. Po drugie, dyskusja nad takim stanowiskiem wymaga podjęcia wielu zagadnień z zakresu epistemologii modalności, które na chwilę obecną znajdują się poza zakresem moich kompetencji.

Użytkownicy dyskursu możliwościświatowego, akceptując twierdzenia odnoszące się do światów możliwych bądź kwantyfikujące po nich, angażują się w taką grę w udawanie, w której wyobrażają sobie istnienie bytów spełniających te twierdzenia. W szczególności zatem wyobrażają sobie istnienie światów możliwych oraz możliwych indywiduów. O twierdzeniu możliwościświatowym powiemy, że jest fikcyjnie prawdziwe wtw, gdy istnieje taka gra, w której istnieje reguła nakazująca wyobrażenie sobie istnienia bytów spełniających to twierdzenie, gdy kwestia fikcyjnej prawdziwości tego twierdzenia powstanie. W ramach możliwościświatowej gry w udawanie światom możliwym oraz możliwym indywiduom przypisuje się jedynie fikcyjne istnienie, gdyż mają one jedynie status bytów wyobrażonych, dzięki czemu można w jej kontekście akceptować pewne twierdzenia możliwościświatowe bez przyjmowania zobowiązań ontologicznych co do istnienia światów możliwych. Możliwościświatową grę w udawanie oznaczmy skrótem „*GPW*”.

W rozdziale 2. wspomniałem, że granie w udawanie jest aktywnością wyobrażania sobie czegoś, rządzoną regułami postulatywnymi oraz regułami generowania. Gdy mamy do czynienia z konkretnym przypadkiem gry w udawanie takim jak gra *GPW*, można nieco więcej powiedzieć o zasadach rządzących tą grą. Wydaje się, że wśród reguł postulatywnych powinny znaleźć się takie reguły, które nakazują wyobrażenie sobie, iż istnieją byty, które spełniają postulaty (2a) — (2g), jeśli powstaje kwestia fikcyjnej prawdziwości tych postulatów. Z kolei na reguły generowania gry *GPW*, mające formę warunkowych nakazów wyobrażania sobie pewnych rzeczy, jeśli zachodzą odpowiednie okoliczności, mogą składać się następujące preskrypcje:

(RG-@) Dla każdego faktu  $x$ , jeśli  $x$  funduje prawdziwość twierdzenia niemodalnego „ $p$ ”, to  $x$  jest koniecznym i wystarczającym warunkiem prawdziwości twierdzenia „ $W @, p$ ” (gdzie „@” desygnuje świat aktualny).

(RG-□) Dla każdego faktu modalnego  $y$ , jeśli  $y$  funduje prawdziwość

twierdzenia, „Konieczne, że  $q$ ”, to  $y$  jest koniecznym i wystarczającym warunkiem fikcyjnej prawdziwości twierdzenia „We wszystkich światach możliwych,  $q$ ”.

- (RG- $\diamond$ ) Dla każdego faktu modalnego  $z$ , jeśli  $z$  funduje prawdziwość twierdzenia „Możliwe, że  $r$ ”, to  $z$  jest koniecznym i wystarczającym warunkiem fikcyjnej prawdziwości twierdzenia „W pewnym świecie  $w$ ,  $r$ ”.

Wprowadzenie reguł generowania w grze *GPW* z pewnością jest warunkiem koniecznym możliwości dostarczenia wyjaśnienia, w jaki sposób fikcyjna prawdziwość twierdzeń możliwościświatowych może zależeć od tego, jakie są fakty modalne fundujące literalną prawdziwość twierdzeń modalnych. Wprowadzenie powyższych reguł generowania nie wyjaśnia jeszcze, w jaki sposób możliwe jest wnioskowanie o istnieniu lub własnościach faktów modalnych na podstawie tego, co uznaje się za prawdziwe w ramach *GPW*. Przedmiotowy fikcjonalista modalny twierdzi jednak, że użyteczność możliwościświatowych analiz zasadza się na tym, iż dyskurs ten dostarcza środków pozwalających na ujawnianie informacji o faktach modalnych. W związku z tym musi również przyjąć założenie, że można przechodzić od fikcyjnie prawdziwych twierdzeń o światach możliwych do literalnie prawdziwych twierdzeń o faktach modalnych. Dlatego też w ramach jego koncepcji powinny pojawić się jakieś reguły ustalające poprawne przejścia od fikcyjnej prawdziwości twierdzeń możliwościświatowych do prawdziwości twierdzeń modalnych. Nazwijmy te reguły *odwróconymi regułami generowania* i wstępnie nadajmy im następującą postać:

- (conv-RG- $\square$ ) Dla każdego twierdzenia  $q$ , jeśli „We wszystkich światach możliwych,  $q$ ” jest fikcyjnie prawdziwe, to istnieje fakt modalny  $y$  fundujący prawdziwość twierdzenia „Konieczne, że  $q$ ”;

oraz

(conv-RG- $\Diamond$ ) Dla każdego twierdzenia  $r$ , jeśli „W pewnym świecie  $w$ ,  $r$ ” jest fikcyjnie prawdziwe, to istnieje fakt modalny  $z$  fundujący prawdziwość twierdzenia „Możliwe, że  $r$ ”.

Dopóki jednak przedmiotowy fikcjonalista modalny nie powie niczego więcej o naturze faktów modalnych, nie jest jasne, jakich dokładnie informacji twierdzenia możliwościwiatowe dostarczają o faktach modalnych oraz czy własności faktów modalnych, które można przypisać im na podstawie możliwościwiatowych analiz, są faktycznie własnościami posiadanymi przez te fakty modalne. Do zagadnienia tego wrócę w §4.6.

### 3 Fikcjonalizm modalny — rewolucyjny projekt czy zastana praktyka?

Istnieje jeszcze jeden podział, który obowiązuje wśród stanowisk fikcjonalistycznych w ogólności. Wyróżnia się mianowicie *hermeneutyczny* fikcjonalizm i *rewolucyjny* fikcjonalizm (zob. Burgess i Rosen 1997, s. 6–7; Kalderon 2005b, s. 136–142). Ten drugi wpisuje się w klasę teorii błędu, zgodnie z którymi użytkownicy dyskursu D błędnie przekonani są o prawdziwości twierdzeń tego dyskursu, gdyż twierdzenia te są albo fałszywe, albo akceptacja ich jako prawdziwych jest nieuzasadniona i w związku z tym powinni oni poddać rewizji swoje przekonania co do wartości logicznej tych twierdzeń (zob. Kalderon 2005b, s. 106). Choć użytkownicy dyskursu D są błędnie przekonani o prawdziwości twierdzeń tego dyskursu, to w obliczu złej nowiny, jaką jest nieistnienie bądź niezasadność przyjęcia istnienia bytów klasy  $K$ , których dotyczą twierdzenia tego dyskursu, mogą oni postąpić na dwa sposoby. Z jednej strony mogą zarzucić posługiwanie się dyskursem D, z drugiej — mogą chcieć zachować prawo posługiwania się tym dyskursem — który wedle ich przekonania posiada zalety niezależne od jego

prawdziwości — osłabiając standardy idealnej akceptacji twierdzeń w ramach tego dyskursu. Rewolucyjny fikcjonalista zaleca właśnie przeprowadzenie rewizji dotychczasowej praktyki językowej przez wyrugowanie przekonania, iż twierdzenia dyskursu D są prawdziwe, oraz przez zmianę standardów akceptacji twierdzeń tego dyskursu na taką, która będzie spójna z przekonaniem o fałszywości tych twierdzeń bądź z zawieszeniem sądu w tej sprawie.

W odróżnieniu od rewolucyjnego fikcjonalisty hermeneutyczny fikcjonalista nie przypisuje użytkownikom dyskursu D błędnego lub nieuzasadnionego przekonania o prawdziwości twierdzeń tego dyskursu, nie zaleca również zmiany standardów akceptacji twierdzeń dyskursu D, twierdzi natomiast, że użytkownicy tego dyskursu nigdy nie traktowali tych twierdzeń jako literalnie prawdziwych, standardy idealnej akceptacji w ramach tego dyskursu zawsze były spójne z przekonaniem o fałszywości tych twierdzeń bądź zawieszeniem sądu w sprawie ich wartości logicznej. Hermeneutyczny fikcjonalista twierdzi zatem, że proponowana przez niego koncepcja stanowi tylko opis sposobu, na jaki użytkownicy dyskursu D używają twierdzeń tego dyskursu.

Podsumowując, rewolucyjny fikcjonalista utrzymuje, że użytkownicy dyskursu D powinni przestać uważać twierdzenia tego dyskursu za prawdziwe, hermeneutyczny fikcjonalista — że nigdy nie zaczęli traktować ich w ten sposób.

Czy modalnemu fikcjonalizmowi jest bliżej do stanowiska o hermeneutycznym, czy rewolucyjnym charakterze? Nieintuicyjny charakter koncepcji Lewisa, na którą — jak sam pisał — najczęstszą reakcją było wytrzeszczanie oczu, może sugerować, że użytkownicy dyskursu możliwościowego raczej nie akceptują jego twierdzeń jako literalnie prawdziwych. Nie chcę twierdzić, że fikcjonalizm modalny wpisuje się w klasę hermeneutycznych stanowisk fikcjonalistycznych, gdyż wymagałoby to przeprowadzenia badań empirycznych, niemniej jednak

prima facie wydaje się dość bliski tym ostatnim.

## 4 Fikcjonalizm modalny D. Armstronga

Fikcjonalizm modalny D. Armstronga pojawił się w jego pracy (1989). Upraszczając, Armstrong utożsamia światy możliwe z koniunkcjami możliwych atomowych stanów rzeczy (1989, s. 47–48). Same stany rzeczy są ustrukturyzowanymi bytami, których elementami konstytutywnymi są indywidua, własności lub relacje. Stan rzeczy zaś jest atomowy wtw, gdy jego elementami konstytutywnymi są wyłącznie proste indywidua, tj. takie indywidua, które nie mają innych indywiduów jako swoich części właściwych oraz proste własności lub relacje, tj. takie własności lub relacje, które nie posiadają elementów konstytutywnych, które same byłyby własnościami lub relacjami (1989, s. 38). W związku z tym z grona elementów konstytutywnych atomowych stanów rzeczy wykluczone są wszystkie negatywne własności i relacje, gdyż mają one jako swoje elementy konstytutywne inne własności lub relacje. Atomowe stany rzeczy mają być tutaj bytami, których struktura odpowiada formie takich twierdzeń, jak to, że  $a$  jest  $F$ ,  $a$  jest  $G$ ,  $a$  pozostaje w relacji  $R$  do  $b$  itd., jeśli tylko odniesieniami wyrażen składających się na te twierdzenia są proste indywidua, własności lub relacje (1989, s. 45). Wśród atomowych stanów rzeczy nie znajdują się również negatywne stany rzeczy takie, jak to, że nie jest tak, że  $a$  jest  $H$  itd. Niektóre spośród atomowych stanów rzeczy istnieją, wówczas stanowią one uprawdziwiacze dla twierdzeń, których formy odpowiadają strukturom tych stanów rzeczy, inne nie istnieją i są jedynie możliwymi stanami rzeczy. Zgodnie z propozycją Armstronga, „możliwe stany rzeczy są wszystkimi kombinacjami prostych indywiduów, własności i relacji, które respektują formę atomowych stanów rzeczy” (1989, s. 48), przy czym zakłada się, że w przypadku prostych indywiduów, własności i relacji dowolne z nich są wzajemnie współmożliwe, tj.

każde z nich z każdym innym może stanowić element konstytutywny pewnego możliwego stanu rzeczy. Armstrong przyjmuje założenie, że wszystkie proste indywidua, własności i relacje są bytami aktualnie istniejącymi. W związku z tym możliwe stany rzeczy są nieistniejącymi kombinacjami tych aktualnie istniejących bytów. Utożsamiając światy możliwe z koniunkcjami możliwych stanów rzeczy, których istnienie się odrzuca, staje się również przed koniecznością odrzucenia istnienia tych światów. W jaki zatem sposób możliwe jest mówienie o nich? Odpowiedź Armstronga przedstawia się podobnie do przedstawionego wyżej stanowiska metafikcjonalisty modalnego. Traktuje on reprezentację sposobu, na jaki świat aktualny mógłby być, jako fikcję, zgodnie z którymi świat aktualny jest na pewien sposób. Dodatkowo wprowadza również dużą fikcję, która stanowi reprezentację uniwersum światów możliwych. W ramach takiego podejścia twierdzenie możliwości światowej  $p^*$  jest elipitycznie wyrażonym twierdzeniem o postaci: „Zgodnie z dużą fikcją, istnieje fikcja, zgodnie z którą  $p^*$ ”. Treść dużej fikcji i fikcji reprezentujących poszczególne światy możliwe jest determinowana przez możliwe kombinacje aktualnie istniejących bytów (1989, s. 50).

Zasadniczą trudnością, przed którą staje taka koncepcja, jest to, że ograniczając się do kombinacji na aktualnie istniejących bytach, nie sposób utworzyć koniunkcji stanów rzeczy, w której suma indywiduów stanowiących elementy konstytutywne tych stanów rzeczy byłaby większa od sumy aktualnie istniejących indywiduów. W rezultacie fałszywe będzie każde twierdzenie o postaci: „Zgodnie z dużą fikcją, istnieje fikcja, zgodnie z którą istnieje  $n$  indywiduów”, gdzie  $n$  jest liczbą indywiduów większą od liczby aktualnie istniejących indywiduów. Wydaje się jednak, że twierdzenie modalne „Mogłoby istnieć więcej indywiduów niż aktualnie istnieje” jest prawdziwe, a w związku z tym treść dużej fikcji nie spełnia warunku ( $K_1$ ). Z uwagi na tę konsekwencję w dalszej części rozprawy skupię się wyłącznie na fikcjonalizmie modalnym bazującym na realizmie modalnym Lewisa.



# Rozdział IV

## Fikcjonalizm modalny pod lupą

Czas przyjrzeć się najczęściej wymienianym wadom i zaletom fikcjonalizmu modalnego. Zaczniemy od zalet.

### 1 Wiedza o światach możliwych

Flagową zaletą fikcjonalizmu modalnego ma być, już wspomniana, możliwość uniknięcia przyjmowania niewiarygodnej ontologii realizmu modalnego. Kolejnym argumentem za fikcjonalizmem jest spójne z naturalizmem wyjaśnienie możliwości poznawania światów możliwych. Tego typu wyjaśnienie jest nieosiągalne dla realisty modalnego. Podejmując kwestię poznania sposobów, na jakie są światy możliwe, Lewis pisze:

Sądzę, że nasze codzienne przekonania modalne są w znacznej mierze konsekwencjami zasady rekombinacji [...]. Można by sobie wyobrazić rygorystyczne rozumowanie wychodzące od jej precyzyjnego sformułowania, ale w rzeczywistości nasze rozumowanie jest bliższe temu, by przyjąć formę eksperymentów myślowych. Próbuje się myśleć, w

jaki sposób duplikaty rzeczy zaakceptowanych już jako możliwe — przykładowo dlatego, że są aktualne — mogłyby być zaaranżowane, by pasowały do opisu rzekomej możliwości. Wyobrazivszy sobie wiele aranżacji — oczywiście, nie w zupełnych szczegółach — rozważamy, w jaki sposób mogłyby być odpowiednio opisane (Lewis 1986, s. 113–114).

Zasada rekombinacji głosi, że dla dowolnych indywiduów nie posiadających części wspólnych, istnieje świat możliwy, w którym istnieje dowolna liczna duplikatów tych indywiduów, pod warunkiem że dziedzina przedmiotowa tego świata jest w stanie pomieścić te indywidua (por. Lewis 1986, s. 87–89).

Poza eksperymentami myślowymi Lewis podkreśla również rolę wyobraźni w poznaniu sposobów, na jakie są światy możliwe. Lewis musi tutaj zakładać, że zasady rządzące naszą wyobraźnią spełnione są w uniwersum światów możliwych. Rodzi się jednak pytanie, w jaki sposób możliwe jest, by „zasady rządzące naszymi wyobrażeniami prawdziwie opisywały dziedzinę obiektów, z którą ludzkie byty nie mają absolutnie żadnego kontaktu, podczas gdy zasady te zostały prawdopodobnie ukształtowane przez zupełnie naturalny proces ewolucyjny?” (Rosen 1990, s. 340).

W przypadku fikcjonalizmu modalnego powyższy problem w ogóle nie powstaje, ponieważ zasady rządzące naszą wyobraźnią wyznaczają jedynie treść fikcji *PW* lub gry *GPW*, nie roszcząc pretensji do bycia zasadami prawdziwymi o jakiegokolwiek dziedzinnie przedmiotowej (Rosen 1990, s. 340).

Dla realisty modalnego problemem jest również wyjaśnienie, w jaki sposób możliwe jest ustalenie, że światy możliwe w ogóle istnieją. Główna strategia argumentacyjna przyjęta przez Lewisa w (1986) ma postać wniosku do najlepszego wyjaśnienia. Lewis twierdzi, że proponowana przez niego koncepcja światów możliwych najlepiej

wyjaśnia, jakie są warunki prawdziwości twierdzeń modalnych, jaka jest natura własności i relacji, a także jaka jest natura sądów w sensie logicznym i wielu innych. Faktycznie, system metafizyczny zaproponowany przez Lewisa pozwala, jak żaden inny, dzięki przyjęciu relatywnie niewielkiej ilości pojęć pierwotnych, w tym w większości pojęć ekstensjonalnych, na przeprowadzenie redukcyjnych analiz szeregu bytów. Jednakże w obliczu reagowania z niedowierzaniem na jego koncepcję powstaje pytanie, czy prostota i teoretyczna użyteczność tego systemu stanowi wystarczającą podstawę do wnioskowania o istnieniu wielości światów?

Wnioskowanie do najlepszego wyjaśnienia jest niededukcyjnym schematem wnioskowania. Z tego powodu należy jeszcze wykazać, że zastosowanie wnioskowania do najlepszego wyjaśnienia w argumentacji na rzecz istnienia światów możliwych należy do niezawodnych zastosowań tej argumentacji. Potrzebny jest zatem *meta-argument przemawiający za niezawodnością tego zastosowania wnioskowania do najlepszego wyjaśnienia*. Jaką postać mógłby przyjąć taki meta-argument? Niestety, Lewis już nie udziela odpowiedzi na to pytanie.

Najpopularniejszym zastosowaniem wnioskowania do najlepszego wyjaśnienia jest argumentacja na rzecz istnienia nieobserwowalnych obiektów. Utrzymuje się, że obserwacje, które najlepiej wyjaśniać przy użyciu pewnej teorii postulującej istnienie przedmiotów nieobserwowalnych, stanowią świadectwo na rzecz niezawodności zastosowania wnioskowania do najlepszego wyjaśnienia w argumentacji na rzecz istnienia bytów nieobserwowalnych. Nawet jeśli zgodzić się z takim podejściem, to w przypadku koncepcji Lewisa meta-argument tego typu jest nieosiągalny. Nie sposób wskazać na taką klasę obserwacji, do których wyjaśnienia najlepiej byłoby zastosować realizm modalny. Wydaje się również, że w przypadku realizmu nie sposób wyjść od empirycznych przesłanek przy konstrukcji odpowiedniego meta-argumentu. Z zasady fakty empiryczne nie mogą być świadectwami istnienia światów

możliwych, gdyż z założenia są one kauzalnie odizolowane od naszego świata. Nie wydaje się również, by meta-argument mógł stanowić rozumowanie wychodzące od apriorycznych przesłanek. Ciężko wskazać, jakie aprioryczne racje mogłyby przemawiać za poprawnością wnioskowania z zalet koncepcji Lewisa o jej prawdziwości (por. Shalkowski 2010, s. 179). W obliczu niedostarczenia meta-argumentu na rzecz poprawności zastosowania wnioskowania do najlepszego wyjaśnienia w argumentacji na rzecz istnienia światów możliwych można mieć wątpliwości, czy istnieje jakiekolwiek uzasadnienie dla zaakceptowania ich istnienia.

Fikcjonalista modalny nie staje przed analogicznym problemem, gdyż nie przeprowadza on wnioskowania z pragmatycznych zalet fikcji *PW* o istnieniu bytów postulowanych przez tę fikcję. W ramach jego stanowiska, zalety fikcji *PW* uważa się za racje do korzystania z niej. Stanowią one jedynie rację do akceptacji jej twierdzeń jako prawdziwych zgodnie z fikcją *PW* bądź jako prawdziwych w ramach gry *GPW*. Motywując swoje stanowisko, fikcjonalista modalny mógłby podzielać wszystkie przywołane wyżej obawy, wyrażone wobec możliwości sformułowania odpowiedniego meta-argumentu przez realistę modalnego. W trudności dostarczenia takiego meta-argumentu fikcjonalista może nawet upatrywać wsparcia dla przyjmowanej przez niego tezy o nieistnieniu światów możliwych bądź dla pozostania na stanowisku agnostycyzmu w sprawie ich istnienia. W szczególności agnostyczna wersja fikcjonalizmu modalnego wydaje się tutaj atrakcyjna.

## 2 Obiekcja Brocka-Rosena

Nieoczekiwanie fikcjonalistyczna analiza funktora konieczności może stanowić poważne zagrożenie dla projektu uniknięcia przyjmowania istnienia światów możliwych. Weźmy funktor konieczności *de dicto*. Jest to funktor zdaniotwórczy od argumentu zdaniowego o

takiej własności, że odniesienia do bytów w zdaniu będącym jego argumentem są ontologicznie zobowiązujące. Poprawna analiza tego funktora powinna zachowywać tę własność. Jako że fikcjonalista chce przeprowadzić redukcję tego funktora, będzie musiał zaproponować taki schemat przekładu zdań z tym funktorem, w którym zdania występujące jako jego argumenty będą pociągały za sobą swoje standardowe zobowiązania ontologiczne. Problem pojawia się, gdy argumentami takich funktorów będą zdania, w których czyni się odniesienia do bytów klasy światów możliwych (por. Nolan i O’Leary-Hawthorne 1996, s. 26–27).

Za realistą metafikcjonalista analizuje funktor modalny konieczności w kategoriach kwantyfikacji po wszystkich światach możliwych, z tym że twierdzenia kwantyfikujące po światach możliwych są akceptowane przez niego jako prawdziwe jedynie wówczas, gdy wszystkie wystąpienia kwantyfikatorów, wiążących zmienne przebiegające po światach możliwych i możliwych indywiduach, znajdują się w zasięgu funktora opowieści. Zatem metafikcjonalista modalny powinien akceptować jako prawdziwe wszystkie podstawienia schematu (MMF) o następującej postaci:

Konieczne, że  $p$  wtw, gdy zgodnie z fikcją  $PW$ , we wszystkich światach możliwych,  $p$ .

Zwykle uważa się, że prawdziwość lewej strony tej równoważności pociąga za sobą zobowiązania ontologiczne co do istnienia bytów, o których mowa w twierdzeniu  $p$ , gdyż z jej prawdziwości wynika prawdziwość twierdzenia  $p$  w świecie aktualnym. Wynikanie ostatniego z pierwszego jest sankcjonowane w każdej logice modalnej, w której przyjmuje się aksjomat T:  $\Box p \rightarrow p$ . Założenie o literalnej prawdziwości argumentu zdaniowego będzie rodzić problemy, gdy (a) w twierdzeniach modalnych z funktorem konieczności wystąpią odniesienia do światów możliwych i możliwych indywiduów oraz (b) podstawienie prawej strony schematu (MMF) będzie prawdziwe. Wtedy zakładając

prawdziwość równoważności wyrażanej przez podstawienie schematu (MMF), metafikcjonalista modalny będzie musiał uznać konieczną prawdziwość twierdzeń kwantyfikujących po światach możliwych. Czy istnieje takie podstawienie schematu (MMF), które spełnia warunki (a) i (b)?

Zgodnie z obiekcją Brocka-Rosena, sformułowaną niezależnie od siebie przez Stuarta Brocka (1993) i Gideona Rosena (1993), poniższe rozumowanie pokazuje, że istnieją takie niepożądane podstawienia schematu (MMF). Weźmy twierdzenie modalne:

(BR<sub>1</sub>) Jest konieczne, że jest przygodne, że istnieją kangury,

któremu na gruncie realistycznego schematu przekładu ma odpowiadać:

(BR<sub>2</sub>) We wszystkich światach możliwych, istnieją takie światy, w których kangury istnieją i takie, w których kangury nie istnieją,

które — na gruncie metafikcjonalizmu modalnego — jest eliptycznie wyrażonym twierdzeniem o postaci:

(BR<sub>3</sub>) Zgodnie z *PW*, we wszystkich światach możliwych, istnieją takie światy, w których kangury istnieją i takie, w których kangury nie istnieją.

Z twierdzenia (BR<sub>2</sub>) wynika twierdzenie:

(BR<sub>4</sub>) We wszystkich światach, istnieją przynajmniej dwa światy,

które ma swój metafikcjonalistyczny odpowiednik o postaci:

(BR<sub>5</sub>) Zgodnie z *PW*, we wszystkich światach, istnieją przynajmniej dwa światy.

Na mocy schematu (MMF), twierdzenie (BR<sub>5</sub>) jest równoważne twierdzeniu modalnemu:

(BR<sub>6</sub>) Koniecznie, istnieją przynajmniej dwa światy możliwe.

Z ostatniego można z kolei wyprowadzić twierdzenie:

(BR<sub>7</sub>) Istnieją przynajmniej dwa światy możliwe.

Jeśli w ramach metafikcjonalistycznej analizy twierdzeń modalnych zakłada się literalną prawdziwość lewej strony równoważności (MMF), to zastosowanie tego schematu do (BR<sub>6</sub>) prowadzi do uznania literalnej prawdziwości twierdzenia (BR<sub>7</sub>), która z kolei przekreśla szanse na powodzenie metafikcjonalistycznego projektu uniknięcia zobowiązań ontologicznych względem światów możliwych (por. Rosen 1993, s. 75–77).

W odpowiedzi na powyższą obiekcję H. Noonan (1994) argumentuje, że opiera się ona na błędnym założeniu (Z), zgodnie z którym twierdzenie (BR<sub>2</sub>) pociąga za sobą twierdzenie (BR<sub>4</sub>). W języku logiki modalnej z kwantyfikatorami twierdzeniu (BR<sub>1</sub>) odpowiada twierdzenie:

$$(BR_1\text{-QM}) \quad \Box(\Diamond \exists x Kx \& \Diamond \exists x Kx),$$

gdzie predykat „*K*” jest równoznaczny z predykatem języka potocznego „jest kangurem”. Na gruncie teorii odpowiedników, przedstawionej w (Lewis 1968/1983, s. 31–32), realistycznemu nieformalnemu przekładowi twierdzenia (BR<sub>1</sub>-QM), oznaczonemu wyżej jako (BR<sub>2</sub>), odpowiada:

$$(BR_2\text{-CT}) \quad (w)[Ww \rightarrow (\exists w_1(Ww_1 \& \exists x(I(x, w_1) \& Kx) \& \exists w_2(Ww_2 \& \exists y(I(y, w_2) \& Ky)))],$$

gdzie „*W*” jest równoznaczny z predykatem „jest światem możliwym”, a relacja „*I*” — z predykatem „jest częścią”.

Z kolei sformalizowanym przekładem twierdzenia (BR<sub>6</sub>) jest zdanie:

$$(BR_6\text{-QM}) \quad \Box[\exists w_1 \exists w_2(Ww_1 \& Ww_2 \& w_1 \neq w_2)],$$

któremu odpowiada nieformalny realistyczny przekład, oznaczony wyżej jako  $(BR_4)$ , którego formalnym odpowiednikiem w ramach teorii odpowiedników jest:

$$(BR_4-CT) \quad (w)[Ww \rightarrow (\exists w_1(Ww_1 \& I(w_1, w)) \& \exists w_2(Ww_2 \& I(w_2, w))) \& w_1 \neq w_2].$$

Zgodnie z opowieścią  $PW$ , twierdzenie  $(BR_2-CT)$  jest prawdziwe, natomiast twierdzenie  $(BR_4-CT)$  jest fałszywe, gdyż wymaga przyjęcia — wbrew tezie o wzajemnym odizolowaniu światów — że jeden świat może być częścią drugiego świata (Noonan 1994, s. 135). Powyższe zastosowania schematu przekładu z (Lewis 1968/1983) dla twierdzeń modalnych pokazują, że założenie  $(Z)$  obiekcji Brocka-Rosena nie jest poprawne, ponieważ twierdzenie  $(BR_2-CT)$  nie pociąga za sobą twierdzenia  $(BR_4-CT)$ .

Kiedy porówna się twierdzenia  $(BR_2)$  i  $(BR_2-CT)$ , widać, że w drugim nie występuje wyrażenie relacyjne „ $I$ ”, które odpowiadałoby frazie „w świecie”, składającej się na frazę „we wszystkich światach” na początku pierwszego twierdzenia. W związku z tym  $(BR_2)$  nie powinno być traktowane jako poprawny nieformalny przekład twierdzenia  $(BR_2-CT)$ . Przekład taki powinien przyjąć następującą postać:

$$(BR_2^*) \quad \text{Dla wszystkich światów } w, \text{ istnieje świat możliwy } w_1, \text{ w którym istnieją kangury i istnieje świat możliwy } w_2, \text{ w którym nie istnieją kangury.}$$

Twierdzenie  $(BR_4-CT)$  jest natomiast poprawnym przekładem  $(BR_4)$ , ponieważ każdemu wystąpieniu wyrażenia „w świecie” w  $(BR_4)$  odpowiada w wyrażenie relacyjne „ $I$ ” w  $(BR_4-CT)$ . Twierdzenie  $(BR_4)$  nie wynika jednak z twierdzenia  $(BR_2^*)$ , ponieważ „z sądu, że istnieje kilka światów o odmiennych własnościach nie wynika, że wszystkie [...] światy zawierają w sobie te odmienne światy” (Noonan 1994, s. 136). Zatem założenie obiekcji Brocka-Rosena jest niepoprawne, a sam zarzut można odrzucić.



Nie powinno się jednak przedwcześnie świętować sukcesu, gdyż zaproponowana przez Noonana strategia radzenia sobie z obiekcją Brocka-Rosena opiera się na takim schemacie przekładu twierdzeń modalnych na twierdzenia możliwościświatowe, który prowadzi do niepożądanych konsekwencji. W przypadku twierdzeń modalnych, w których występuje funktor modalny możliwości poprzedzający twierdzenie egzystencjalne, schemat przekładu, na którym opiera się jego kontrargumentacja, przedstawia się następująco:

- (MR-M) Możliwie, istnieje takie  $x$ , że  $Gx$  wtw, gdy istnieje takie  $w$ , że  $w$  jest światem możliwym i istnieje takie  $x$ , że  $x$  jest częścią świata  $w$  i  $Gx$ .

Gdyby jednak w taki sposób przekładać *wszystkie* twierdzenia modalne z funktorem możliwości, to realista modalny popadłby w tarapaty. Z jednej bowiem strony, realista modalny musi zaakceptować prawdziwość twierdzenia:

- (CR) Istnieje wiele światów możliwych,

z którego — w każdej logice modalnej, w której występuje aksjomat T — wyprowadzić można twierdzenie:

- (CR-M) Możliwie, istnieje wiele światów możliwych,

które również powinno być prawdziwe. Okazuje się jednak, że na gruncie powyższego schematu przekładu twierdzeniu (CR-M) odpowiada twierdzenie:

- (CR-CT)  $\exists w[Ww \& \exists w_1, w_2, w_3, \dots (Ww_1 \& Ww_2 \& Ww_3 \& \dots \& I(w_1, w) \& I(w_2, w) \& I(w_3, w) \& \dots)]$ ,  
 $w_1 \& w \neq w_2 \& w \neq w_3 \& \dots \& w_1 \neq w_2 \& w_1 \neq w_3 \& w_2 \neq w_3 \& \dots]$ ,

które jest niespójne z przyjmowaną w ramach realizmu modalnego tezą o wzajemnym odizolowaniu światów możliwych, wykluczającą to,

by jeden świat możliwy mógł być częścią drugiego świata możliwego. Na gruncie realizmu modalnego (CR-CT) musi być nie tylko fałszywe, ale i koniecznie fałszywe. Schemat przekładu wykorzystany przez Noonana nie wydaje się zatem zamknięty na regułę wnioskowania: „ $p \rightarrow \Diamond p$ ”, w związku z czym konieczna jest jego modyfikacja.

W swoim artykule (1968/1983) Lewis nie podejmował kwestii analizy twierdzeń modalnych o *transświatowym* zakresie, tj. takich twierdzeń modalnych, w których kwantyfikacja nie jest ograniczona do bytów znajdujących się w dziedzinie pojedynczego świata, a przebiega po bytach znajdujących się w dziedzinach przedmiotowych wielu światów możliwych. Jak pokazuje schemat przekładu (MR-M), funktor możliwości jest nie tylko analizowany jako kwantyfikacja egzystencjalna po światach możliwych, ale również wprowadza on (za pośrednictwem wyrażenia relacyjnego „ $I$ ”) ograniczenie kwantyfikacji, znajdującej się w jego zasięgu, do bytów, które są częściami jednego świata. W związku z tym standardową rolą funktorów modalnych jest również ograniczanie dziedziny kwantyfikacji występującej w twierdzeniach modalnych do części jakiegoś świata możliwego.

Funktor możliwości nie zawsze jednak pełni swoją standardową funkcję. W (1986, s. 13–14) Lewis rozważa przykład takiego porównawczego twierdzenia modalnego, jak „Możliwe, że pewna czerwona rzecz bardziej przypomina pewną pomarańczową rzecz niż przypomina jakąkolwiek niebieską rzecz”, które wydaje się prawdziwe, nawet jeśli uznać, iż porównywane indywidua są częściami odmiennych światów możliwych. Jeśli tak jest, to funktor modalny możliwości nie ogranicza tutaj dziedziny kwantyfikacji występującej w zdaniu, będącym jego argumentem, do dziedziny przedmiotowej jakiegoś świata. Z tego powodu realista modalny proponuje, by twierdzenie to wyrazić bez użycia jakichkolwiek funktorów modalnych za pomocą twierdzenia „Istnieje takie  $x$  i  $y$ , że  $x$  jest czerwone i  $y$  jest pomarańczowe i dla każdego  $u$  i  $v$ , jeśli  $u$  jest czerwone i  $v$  jest niebieskie, to  $u$  bardziej

przypomina  $y$  niż  $u$  przypomina  $v$ ". Przyjmuje się tutaj, że kwantyfikacja ma nieograniczony zasięg i przebiega po wszystkich możliwych indywiduach.

John Divers w (1999, s. 322–325) zwraca uwagę, że twierdzenie (CR-M) również jest twierdzeniem modalnym o transświatowym zakresie. Z tezy o wzajemnym odizolowaniu światów możliwych wynika, że funktory modalne poprzedzające twierdzenia, w których kwantyfikuje się po światach możliwych, nie pełnią swojej standardowej roli ograniczania zakresu kwantyfikacji do dziedziny jakiegoś świata. Z uwagi na to — kontynuuje Divers — że funktor modalny możliwości w (CR-M) nie pełni swojej standardowej funkcji, można uznać go za funktor *semantycznie redundantny* na gruncie realizmu modalnego i traktować wszystkie twierdzenia modalne, w których występują takie funktory, jako twierdzenia równoważne twierdzeniom powstałym przez opuszczenie tych funktorów. Wówczas (CR-M) odpowiadałoby twierdzenie:

$$(CR-CT^*) \exists w_1, w_2, w_3, \dots (Ww_1 \& Ww_2 \& Ww_3 \& \dots \& w_1 \neq w_2 \& w_1 \neq w_3 \& w_2 \neq w_3 \& \dots).$$

Aby uwzględnić semantyczną redundantność funktorów modalnych w twierdzeniach o transświatowym zakresie, fikcjoniści modalni — twierdzi Divers — powinni zmodyfikować schemat (MMF), który jest niewrażliwy na przypadki wystąpień takich funktorów modalnych, rozbijając go na dwa następujące:

(MMF<sub>1</sub>) Jeśli twierdzenie „Możliwe, że  $p$ ” jest twierdzeniem modalnym o wewnątrzświatowym zakresie<sup>29</sup>, to możliwe, że  $p$  wtw, gdy zgodnie z  $PW$ , istnieje świat  $w$ , w którym  $p$ .

Jeśli twierdzenie „Konieczne, że  $p$ ” jest twierdzeniem modalnym o wewnątrzświatowym zakresie, to konieczne, że  $p$  wtw, gdy zgodnie z  $PW$ , dla każdego świata  $w$ ,  $p$  w  $w$ .

<sup>29</sup>Tj. zakres zmiennych związanych występujących w twierdzeniu  $p$  jest ograniczony do dziedziny jakiegoś świata.

(MMF<sub>2</sub>) Jeśli twierdzenie „Możliwe, że  $p$ ” jest twierdzeniem modalnym o transświatowym zakresie, to możliwe, że  $p$  wtw, gdy zgodnie z  $PW$ ,  $p$ .

Jeśli twierdzenie „Konieczne, że  $p$ ” jest twierdzeniem modalnym o transświatowym zakresie, to konieczne, że  $p$  wtw, gdy zgodnie z  $PW$ ,  $p$  (por. Divers 1999, s. 325).

W kontekście powyższego rozróżnienia obiekcja Brocka-Rosena opiera się na nieuprawnionym przejściu od twierdzenia (BR<sub>1</sub>) „Jest konieczne, że przygodnie istnieją kangury” do twierdzenia (BR<sub>2</sub>) „We wszystkich światach możliwych, istnieją takie światy, w których kangury istnieją i takie, w których kangury nie istnieją”, gdyż wymaga ono przyjęcia założenia, że funktor konieczności, występujący w (BR<sub>1</sub>), nie jest semantycznie redundantny. Tymczasem jest przeciwnie, gdyż funktor ten nie pełni tutaj swojej standardowej roli. Jeśli funktor modalny konieczności jest semantycznie redundantny w (BR<sub>1</sub>), to poprawnym realistycznym przekładem (BR<sub>1</sub>) nie będzie (BR<sub>2</sub>), ale twierdzenie „Istnieje taki świat, w którym istnieją kangury i istnieje taki świat, w którym nie istnieją kangury”, któremu odpowiada fikcjonalistyczne „Zgodnie z  $PW$ , istnieje świat, w którym istnieją kangury i istnieje świat, w którym nie istnieją kangury”. Na podstawie schematu (MMF<sub>2</sub>) ostatnie może być interpretowane jako wyrażające tezę o koniecznym istnieniu takich światów bądź tezę o ich możliwym istnieniu. W obliczu tych dwóch opcji każdy, kto odrzuca istnienie światów, powinien wybrać drugą z nich i uznać prawdziwość twierdzenia „Możliwie, istnieją światy możliwe”. Zatem — konkluduje Divers — rozbiecie schematu (MMF) na schematy (MMF<sub>1</sub>) i (MMF<sub>2</sub>) pozwala uniknąć obiekcji Brocka-Rosena, gdyż usuwa konieczność akceptowania problematycznego twierdzenia (BR<sub>6</sub>).

Propozycja Diversa natrafia na dwie zasadnicze trudności. Pierwsza z nich dotyczy tego, że akceptacja takiej strategii wymaga od fikcjonalisty modalnego przyjęcia tezy, iż kwantyfikacja w zasięgu

funktora możliwości jest ontologicznie niezobowiązująca, akceptuje on bowiem jako prawdziwe twierdzenie modalne „Możliwie, istnieją światy możliwe” — nie przyjmując ich istnienia bądź pozostając w tej sprawie agnostykiem. Powstaje wówczas pytanie, dlaczego nie uogólnić ostatniego sposobu unikania zobowiązań ontologicznych co do istnienia światów możliwych i nie porzucić fikcjonalizmu modalnego? Pytanie to jest tym bardziej palące, że w literaturze występuje stanowisko modalizmu, zgodnie z którym można uniknąć zobowiązań ontologicznych co do istnienia światów możliwych, jeśli się uzna, iż kwantyfikacja po światach dokonuje się właśnie w zasięgu funktorów modalnych. Szerzej o tej kwestii będę mówił w §5.2., gdzie zestawię obie wersje fikcjonalizmu modalnego z tą koncepcją.

Druga i istotniejsza trudność bierze się z tego, że rozumowanie *Diversa* opiera się na założeniu, iż za pośrednictwem schematu (MMF2) można z twierdzenia „Zgodnie z *PW*, istnieje świat, w którym istnieją kangury i istnieje świat, w którym nie istnieją kangury” wyprowadzić twierdzenie „Możliwie, istnieje świat, w którym istnieją kangury i istnieje świat, w którym nie istnieją kangury”, które może zaakceptować fikcjonalista modalny. Problem jednak w tym, że zastosowania schematu (MMF<sub>2</sub>) opierają się na założeniu, iż funktory modalne występujące w twierdzeniach po lewej stronie tego schematu są semantycznie redundantne. Co prawda, założenie to obowiązuje na gruncie realizmu modalnego, ale nie powinien go przyjmować metafikcjonalista modalny, gdyż musiałby uznać twierdzenie „Możliwie, istnieje świat, w którym istnieją kangury i istnieje świat, w którym nie istnieją kangury” za twierdzenie, które jest równoważne na mocy semantycznej redundantności funktorów modalnych w nim występujących z twierdzeniem „Istnieje świat, w którym istnieją kangury i istnieje świat, w którym nie istnieją kangury”, czego oczywiście nie powinien robić. Celem odrzucenia drugiego twierdzenia jako fałszywego i uznania pierwszego za prawdziwe fikcjonalista modalny musi zatem założyć, że

funktor modalny występujący w pierwszym z tych twierdzeń nie jest semantycznie redundantny. Wówczas jednak bez wyjaśnienia pozostaje kwestia tego, w jaki sposób na podstawie schematu ( $\text{MMF}_2$ ) można wyprowadzić z twierdzenia „Zgodnie z  $PW$ , istnieje świat, w którym istnieją kangury i istnieje świat, w którym nie istnieją kangury” twierdzenie modalne, w którym występuje semantycznie nieredundantny funktor modalny — ale Divers nie podejmuje już tego zagadnienia.

W dotychczasowej ekspozycji odpowiedzi na obiekcję Brocka-Rosena przedstawiałem wyłącznie takie strategie, które zmierzały do jej oddalenia przez wykazanie, iż jest ona sformułowana na podstawie niepoprawnego schematu przekładu twierdzeń modalnych na twierdzenia możliwości światowe. Oprócz strategii tego rodzaju podjęto również próby dokonania takich przekształceń schematu ( $\text{MMF}$ ), by wprowadzona modyfikacja pozwoliła uniknąć pozornego samo-obalenia fikcjonalizmu modalnego, nawet gdy przyjmie się, że rozumowanie prowadzące do wniosku omawianej obiekcji jest poprawne.

Założmy zatem, że wnioskowanie przeprowadzone w ramach obiekcji Brocka-Rosena jest poprawne. Jak widzieliśmy wyżej, trudność pojawiała się, gdy metafikcjonalista akceptował równoważność twierdzeń ( $\text{BR}_6$ ) i ( $\text{BR}_5$ ), tj.:

- (A) Koniecznie, istnieją co najmniej dwa światy możliwe wtw, gdy zgodnie z hipotezą  $PW$ , dla wszystkich światów, istnieją co najmniej dwa światy możliwe.

Wówczas, zakładając poprawność przekładu, jeśli prawa strona tej równoważności jest prawdziwa, to metafikcjonalista jest zmuszony do akceptacji prawdziwości jej lewej strony, która wydaje się implikować istnienie wielości światów możliwych.

Zdaniem R. Woodwarda, sposobem na uniknięcie tej konsekwencji jest zastąpienie schematu ( $\text{MMF}$ ) przez następujące schematy:

( $\text{MMF}_3$ ) „ $p$ ” jest prawdziwe wtw, gdy zgodnie z  $PW$ ,  $p^*$ ,

(MMF<sub>4</sub>) „ $p^*$ ” jest prawdziwe wtw, gdy zgodnie z  $PW$ ,  $p^*$  (Woodward 2008, s. 277),

gdzie „ $p$ ” jest dowolnym zdaniem modalnym, a „ $p^*$ ” jego możliwościowym przekładem. Najistotniejszą cechą obu tych schematów jest to, że ich prawe strony wyznaczają warunki prawdziwości dla — odpowiednio — twierdzeń modalnych i twierdzeń możliwościowych — warunki prawdziwości, które nie wymagają postulowania istnienia światów możliwych.

Dysponując takimi schematami, można zaakceptować przejście od (BR<sub>6</sub>) do (BR<sub>7</sub>) „Istnieją co najmniej dwa światy możliwe” bez przyjmowania istnienia światów możliwych, gdyż warunki prawdziwości tych twierdzeń są takie, że pierwsze jest prawdziwe wtw, gdy zgodnie z  $PW$ , dla wszystkich światów, istnieją co najmniej dwa światy, podczas gdy drugie jest prawdziwe wtw, gdy zgodnie z  $PW$ , istnieją co najmniej dwa światy. Założywszy, iż treść hipotezy  $PW$  jest zamknięta na implikację, możemy zaakceptować przejście od (BR<sub>6</sub>) do (BR<sub>7</sub>), odrzucając istnienie światów możliwych bądź zachowując postawę agnostycyzmu wobec ich istnienia. Celem uznania prawdziwości obu twierdzeń musimy — na gruncie tych warunków prawdziwości — jedynie przyjąć istnienie hipotezy  $PW$  (Woodward 2008, s. 279–282).

Wadą propozycji Woodwarda jest to, że nie bardzo wiadomo, jak wygląda przekład twierdzeń modalnych, w których odniesienia do światów występują w zasięgu funktora modalnego możliwości. Jeśli redukcyjną analizą twierdzenia modalnego „Możliwie, istnieją co najmniej dwa światy możliwe” jest twierdzenie „Zgodnie z  $PW$ , istnieje świat możliwy, w którym istnieją co najmniej dwa światy możliwe”, to ostatnie twierdzenie jest fałszywe, gdy jest niespójne z przyjmowaną w ramach opowieści  $PW$  tezą o wzajemnym odizolowaniu światów. W związku z tym fałszywe powinno również być twierdzenie „Możliwie, istnieją co najmniej dwa światy możliwe”. W obliczu tego, że Woodward akceptuje jako prawdziwy przekład twierdzenia „Koniecznie,

istnieją co najmniej dwa światy możliwe”, fałszywość twierdzenia „Możliwie, istnieją co najmniej dwa światy możliwe” sprawiałaby, że proponowana przez niego redukcyjna analiza warunków prawdziwości twierdzeń modalnych dotyczących światów możliwych nie spełniałaby twierdzenia „ $\Box p \rightarrow \Diamond p$ ”, które jest prawdziwe w każdej logice modalnej, w której obowiązuje aksjomat T.

Daniel Nolan i John O’Leary-Hawthorne w (1996, s. 29–30) sugerowali, iż akceptując umiarkowany metafikcjonalizm modalny, można przyjąć wyłącznie schemat (MMF<sub>4</sub>) i uniknąć obiekcji Brocka-Rosena. W ich przekonaniu wystarczy odrzucić schemat (MMF<sub>3</sub>), by uniknąć konieczności akceptowania kluczowych dla tej obiekcji przejść od (BR<sub>1</sub>) do (BR<sub>3</sub>) oraz od (BR<sub>5</sub>) do (BR<sub>6</sub>). Propozycja ta wydaje się jednak chybiona, gdyż — jak podkreślałem wyżej — czym innym jest porzucenie ambicji definiowania warunków prawdziwości twierdzeń modalnych w kategoriach prawdziwości twierdzeń możliwoświatowych zgodnie z opowieścią *PW*, a czym innym jest odrzucenie możliwości przechodzenia od prawdziwości twierdzeń modalnych do twierdzeń możliwoświatowych prawdziwych zgodnie z opowieścią *PW*. W rezultacie nie jest jasne, dlaczego przejście od prawdziwości twierdzenia (BR<sub>1</sub>) do (BR<sub>3</sub>), czy od (BR<sub>5</sub>) do (BR<sub>6</sub>) miałyby być nieuprawnione, skoro w treści opowieści *PW* mogą istnieć prawa pomostowe uzależniające wartości logiczne (BR<sub>3</sub>) i (BR<sub>5</sub>) od wartości logicznych — odpowiednio — (BR<sub>1</sub>) i (BR<sub>6</sub>). Wciąż zatem należy wyjaśnić, jakie są warunki prawdziwości twierdzenia (BR<sub>6</sub>) „Koniecznie, istnieją przynajmniej dwa światy możliwe”, skoro nie implikują one istnienia światów możliwych. Niestety Nolan i O’Leary-Hawthorne nie rozwijają tej kwestii.

Niemniej jednak ich propozycja wydaje się krokiem w kierunku najbardziej obiecującej odpowiedzi na obiekcję Brocka-Rosena. Strategia Nolana i O’Leary-Hawthorne’a wymaga uzupełnienia o rozróżnienie na takie twierdzenia modalne, dla których metafikcjonalista modalny



proponuje redukcyjną analizę i takie, dla których jej nie proponuje. Mówiąc w rozdziale 2. o fikcjonalizmie przedmiotowym, wspomniałem o rozróżnieniu na bazowe i wtórne gry w udawanie. W pierwszych postuluje się istnienie bytów o takich własnościach, by mogły one pełnić rolę narzędzi pomocnych w reprezentowaniu okoliczności zachodzących niezależnie do tego, co przyjmuje się w ramach gry w udawanie. W drugich, byty, których istnienie postulowano w ramach gier pierwszego typu, same zaczynają pełnić rolę reprezentowanych przedmiotów, a wszystkie twierdzenia reprezentujące te byty są akceptowane jako prawdziwe w ramach gier w udawanie drugiego typu. Umiarkowany metafikcjonalista modalny może analogicznie wyróżnić dwa odmienne typy twierdzeń modalnych i — odpowiednio — dwie odmienne funkcje, jakie mogą pełnić twierdzenia możliwościowe prawdziwe zgodnie z fikcją *PW*. Z jednej strony, może wyróżnić *faktualne* twierdzenia modalne, które reprezentują sposoby, na jakie świat aktualny mógłby być lub musi być, i których prawdziwość jest ufundowana na istnieniu odpowiednich faktów modalnych. Pozwoli to na wyróżnienie twierdzeń możliwościowych prawdziwych zgodnie z *PW*, które pełniła (za pośrednictwem praw pomostowych) rolę narzędzi użytecznych w opisie tych faktów modalnych. Z drugiej zaś strony, może wyróżnić fikcyjne twierdzenia modalne, w których reprezentuje się modalne własności światów możliwych, a wówczas twierdzeniom możliwościowym prawdziwym zgodnie z *PW* można przypisać rolę redukcyjnych analiz warunków prawdziwości takich twierdzeń modalnych.

Za Diversem twierdzenia modalne ostatniego typu można uznać za twierdzenia, w których funktory modalne są semantycznie redundantne i zaproponować dla nich następującą redukcyjną analizę warunków ich prawdziwości:

- (MMF<sub>2</sub><sup>\*</sup>) Jeśli twierdzenie „Możliwe, że *p*” jest twierdzeniem modalnym o transświatowym zakresie i nie reprezentuje wyłącznie sposobu, na jaki świat aktualny mógłby być, to „Możliwe, że *p*” jest

prawdziwe wtw, gdy zgodnie z  $PW$ ,  $p$ .

Jeśli twierdzenie „Konieczne, że  $p$ ” jest twierdzeniem modalnym o transświatowym zakresie i nie reprezentuje wyłącznie sposobu, na jaki świat aktualny musi być, to „Konieczne, że  $p$ ” jest prawdziwe wtw, gdy zgodnie z  $PW$ ,  $p$ .

Jeśli uzupełni się schemat  $(MMF_2^*)$  o schematy  $(MMF_4)$ ,  $(MMF_1)$  i  $(MMF_2)$  oraz ograniczy klasę podstawień lewej strony dwóch ostatnich schematów do faktycznych twierdzeń modalnych, można nawet zaakceptować jako poprawne następujące rozumowanie zbliżone do obiekcji Brocka-Rosena  $(BR^*)$ :

- $(BR_1)$  Jest konieczne, że jest przygodne, iż istnieją kangury.
- $(BR_3^*)$  Zgodnie z  $PW$ , istnieją takie światy, w których kangury istnieją, i takie, w których kangury nie istnieją (na mocy  $MMF_2$ ).
- $(BR_5^*)$  Zgodnie z  $PW$ , istnieją przynajmniej dwa światy (z  $BR_3^*$ ).
- $(BR_6)$  Konieczne, istnieją przynajmniej dwa światy możliwe (na mocy  $MMF_2^*$ ).
- $(BR_7)$  Istnieją przynajmniej dwa światy możliwe.

Rozumowanie to nie prowadzi jednak do konieczności zaakceptowania istnienia światów możliwych, gdyż — w ramach propozycji metafikcjonalisty — warunki prawdziwości ostatniego twierdzenia są analizowane zgodnie ze schematem  $(MMF_4)$  i w związku z tym są identyczne z warunkami prawdziwości twierdzenia  $(BR_6)$  analizowanego zgodnie ze schematem  $(MMF_2^*)$ .

Zaletą takiego podejścia do obiekcji Brocka-Rosena jest to, że (a) w odróżnieniu od propozycji Woodwarda i Noonana jest ona spójna z tezą o wzajemnym odizolowaniu światów możliwych oraz (b) w

odróżnieniu od propozycji Diversa pozwala na zablokowanie przeprowadzenia nieuprawnionego wnioskowania o prawdziwości twierdzenia „Możliwie, istnieje wiele światów możliwych”, w którym funktor możliwości nie jest semantycznie redundantny, na podstawie przesłanki, że w ramach realizmu modalnego funktory modalne w twierdzeniach modalnych o transświatowym zakresie są semantycznie redundantne. Z uwagi na to, że semantyczna redundantność funktorów modalnych w twierdzeniach modalnych o transświatowym zakresie jest ważna tylko w ramach realizmu modalnego, obowiązuje ona jedynie dla tych twierdzeń modalnych, których warunki prawdziwości poddają się redukcijnej analizie za pomocą schematu ( $MMF_2^*$ ), czyli obowiązuje dla fikcyjnych twierdzeń modalnych. Gdy z kolei antyrealista bądź agnostyk w sprawie istnienia światów możliwych mówi o tym, że mogłyby one istnieć, nie wygłasza twierdzenia, którego warunki prawdziwości poddawałyby się analizie zgodnie z powyższym schematem. Mówi on bowiem nie o relacji treści fikcji *PW* do pewnego twierdzenia możliwościowego, ale o tym, że sama fikcja *PW* mogłaby być prawdziwa — o tym, jaką mogłaby być rzeczywistość.

Rzecz jasna powraca pytanie, czy uznanie, iż prawdziwość twierdzenia modalnego „Możliwie, istnieją co najmniej dwa światy możliwe” w kontekstach, w których mówi się o tym, jaka mogłaby być rzeczywistość, nie pociąga za sobą zobowiązań ontologicznych co do istnienia światów możliwych, nie powinno stanowić wystarczającej podstawy do uogólnienia strategii unikania kwantyfikowania po światach możliwych i możliwych indywiduach w zasięgu funktorów modalnych na wszystkie twierdzenia możliwościowe? Obok ostatniej trudności do niewątpliwych wad powyższej propozycji należy również to, iż wymaga ona komplikowania i tak już skomplikowanego stanowiska poprzez wprowadzenie trzech schematów przekładów twierdzeń modalnych. Jeśli jednak metafikcjonalista modalny chce uniknąć samoobalenia tej koncepcji, nie ma innego wyjścia, jak pogodzić się z koniecznością

wprowadzenia tak wielu schematów przekładu.

Odpowiednio przeformułując obiekcję Brocka-Rosena, można sformułować zarzut przeciw przedmiotowej wersji fikcjonalizmu modalnego w postaci następującego rozumowania:

(BR<sub>1</sub>) Jest konieczne, że jest przygodne, że istnieją kangury,

któremu — na mocy reguł generowania (RG- $\Box$ ) i (RG- $\Diamond$ ) — odpowiada fikcyjnie prawdziwe:

(BR<sub>2</sub>) Dla wszystkich światów możliwych, istnieją takie światy, w których kangury istnieją i takie, w których kangury nie istnieją.

Z twierdzenia (BR<sub>2</sub>) powinno wynikać fikcyjnie prawdziwe twierdzenie:

(BR<sub>4</sub>) We wszystkich światach, istnieją przynajmniej dwa światy,

z którego na mocy odwróconej reguły generowania (conv-RG $\Box$ ) wyprowadzić można twierdzenie:

(BR<sub>6</sub>) Koniecznie, istnieją przynajmniej dwa światy możliwe,

które z kolei pociąga za sobą twierdzenie:

(BR<sub>7</sub>) Istnieją przynajmniej dwa światy możliwe.

Korzystając z rozróżnienia na bazowe gry w udawanie i wtórne gry w udawanie, przedmiotowy fikcjonalista modalny powinien ograniczyć stosowalność reguł (conv-RG- $\Box$ ) i (conv-RG- $\Diamond$ ) do faktualnych twierdzeń modalnych. W przeciwnym razie będzie musiał przyjąć istnienie faktu modalnego fundującego prawdziwość twierdzenia (BR<sub>6</sub>), co najprawdopodobniej wiązałoby się z koniecznością przyjęcia istnienia światów możliwych. Odrzucając ich istnienie, musi traktować twierdzenie modalne (BR<sub>6</sub>) jako prawdziwe w ramach pewnej gry w udawanie

(nazwijmy ją „*GM*”), która jest wtórna względem możliwościwiatowej gry *GPW*, następnie zaś wprowadzić takie reguły generowania w grze *GM*, by koniecznym i wystarczającym warunkiem fikcyjnej prawdziwości tego twierdzenia modalnego była prawdziwość twierdzenia ( $BR_7$ ) w grze *GPW*. Co więcej, aby uniknąć przyjmowania schematu przekładu Noonana dla twierdzenia ( $BR_2$ ), przedmiotowy fikcjonalista modalny powinien przyjąć wiarygodne założenie, iż w dziedzinie faktów modalnych spełnione są aksjomaty logiki modalnej *S5*. Dzięki temu prawdziwość każdego twierdzenia modalnego, w którym występują iteracje funktorów modalnych, jest ufundowana na tym samym fakcie modalnym, co prawdziwość twierdzenia modalnego powstałego z tamtego twierdzenia przez opuszczenie wszystkich funktorów modalnych, które mają za swoje argumenty twierdzenia modalne. Wówczas istnienie faktu modalnego fundującego prawdziwość twierdzenia ( $BR_1$ ) będzie — na mocy reguły generowania ( $RG-\Diamond$ ) — konieczną i wystarczającą podstawą fikcyjnej prawdziwości twierdzenia „Istnieją takie światy, w których kangury istnieją i istnieją takie światy, w których kangury nie istnieją”, a nie twierdzenia ( $BR_2$ ). W rezultacie przedmiotowy fikcjonalista modalny będzie w stanie zaakceptować rozumowanie analogiczne do rozumowania ( $BR^*$ ), dostarczając takiej interpretacji każdego z kroków tego rozumowania, że na żadnym z jego etapów nie pojawi się konieczność przyjmowania istnienia światów możliwych.

Jeśli tylko fikcjonalista przedmiotowy sympatyzuje z agnostycyzmem w sprawie istnienia światów możliwych i akceptuje twierdzenie „Możliwie, istnieją światy możliwe”, musi również dopuszczać tezę, że to, czy dane twierdzenie modalne, w którym występują odwołania do światów, jest jedynie fikcyjnie prawdziwe, czy literalnie prawdziwe — zależy od tego, czy osoba wygłaszająca to twierdzenie ma na celu mówienie o tym, jaką rzeczywistość mogłaby być. Doprowadzi to także do konieczności udzielenia odpowiedzi na pytanie, dlaczego nie

uogólnić strategii unikania zobowiązań ontologicznych przez kwantyfikowanie po niepożądanych bytach w zasięgu funktorów modalnych. Koncepcja przedmiotowego fikcjonalisty modalnego nie grzeszy również prostotą, ale podobnie jak przypadku metafikcjonalizmu — jej brak jest ceną za obietnicę uniknięcia przyjmowania zobowiązań co do istnienia światów i za możliwość zachowania obiektywnie istniejących faktów modalnych.

### 3 Problemy z funktorem opowieści

Strategia metafikcjonalisty modalnego opiera się na poprzedzeniu twierdzeń możliwoswiatowych funktorem opowieści „Zgodnie z *PW*”. W jaki sposób powinniśmy rozumieć ten funktor? Intuicyjnie, twierdzenia o postaci „Zgodnie z fikcją *f*, *p*” mówią coś o relacji pomiędzy treścią fikcji *f* a treścią twierdzenia *p*. Jaka to relacja?

W większości wypadków, a przynajmniej w przypadku opowieści *PW*, relacja ta nie jest relacją konsekwencji logicznej. Niezależnie od tego, jaką koncepcję relacji konsekwencji logicznej przyjmujemy, funktor opowieści „Zgodnie z *PW*” wydaje się nie wyrażać tej relacji. W literaturze standardowo wyróżnia się dwa formalne ujęcia relacji konsekwencji logicznej: teoriowodowe i teoriomodelowe (por. Shapiro 2005b, s. 652). Zgodnie z pierwszym z nich, formuła  $\varphi$  wynika ze zbioru formuł  $\lambda$  relatywnie do systemu dedukcyjnego *SD* wtw, gdy istnieje taki ciąg formuł zakończony formułą  $\varphi$ , że każdy element tego ciągu jest albo elementem zbioru  $\lambda$ , albo aksjomatem systemu *SD* lub wynika z poprzednich elementów tego ciągu na mocy reguł inferencyjnych systemu *SD*. Zgodnie z drugim z nich, formuła  $\varphi$  wynika ze zbioru zdań  $\lambda$  wtw, gdy każda interpretacja, która spełnia wszystkie elementy zbioru  $\lambda$ , spełnia również  $\varphi$ , gdzie interpretacje utożsamia się z modelami będącymi parami uporządkowanymi  $\langle D, I \rangle$ , gdzie *D* jest dziedziną, *I* jest funkcją przyporządkowującą wyrażeniom eks-

tensje w  $D$ , a spełnianie zachodzące pomiędzy modelem a formułą definiuje się w standardowy sposób.

W obu ujęciach dla zachodzenia relacji konsekwencji logicznej istotne są znaczenia stałych pozalogicznych. W pierwszym wypadku aksjomaty i reguły inferencji mają wyznaczać wyłącznie znaczenia stałych logicznych (Shapiro 2005b, s. 667–668). W związku z tym tylko ich znaczenia są istotne dla zachodzenia relacji konsekwencji logicznej. W drugim wypadku ignorowane są wszelkie różnice pomiędzy izomorficznymi modelami. Model  $M = \langle D, I \rangle$  jest izomorficzny z modelem  $M' = \langle D', I' \rangle$  wtw, gdy istnieje jedno-jednoznaczna funkcja  $g$  z modelu  $M$  na model  $M'$ , która zachowuje relacje. Przykładowo, niech  $g$  będzie jedno-jednoznaczna funkcja z  $D$  na  $D'$  i niech  $R$  będzie dwuargumentowym predykatem pierwszego rzędu, wówczas  $g$  zachowuje relację  $R$ , gdy para uporządkowana  $\langle a, b \rangle$ , której elementy należą do  $D$ , należy do  $I(R)$  wtw, gdy para uporządkowana  $\langle g(a), g(b) \rangle$ , której elementy należą do  $D'$ , należy do  $I'(R)$ . Z uwagi na to, że przedmioty należące do dziedziny  $D$  mogą różnić się od przedmiotów należących do dziedziny  $D'$ , ekstensja przyporządkowana predykatowi  $R$  w modelu  $M$  może różnić się od ekstensji przyporządkowanej mu w modelu  $M'$ . Mimo to oba modele są nieodróżnialne z logicznego punktu widzenia, co sprawia, że w świetle teoriomodelowej definicji relacji konsekwencji znaczenia stałych pozalogicznych również są nieistotne. Zachowane są jedynie znaczenia stałych logicznych, gdyż są one wyznaczane przez rekursywne definicje spełniania formuł. Przykładowo, znaczenie stałej logicznej koniunkcji „&” jest wyznaczane przez rekursywną standardową definicję spełniania formuły  $\varphi \& \psi$ , która jest niezmienna dla wszystkich interpretacji. W związku z tym stałe logiczne są jedynymi wyrażeniami, których znaczenie jest istotne dla zachodzenia relacji konsekwencji logicznej w ujęciu teoriomodelowym (por. Shapiro 2005b, s. 662–663).

Gdyby funktor opowieści wyrażał relację konsekwencji logicznej,

metafikcjonalista modalny powinien raczej zaakceptować jej teoriomodelową interpretację, gdyż wydaje się ona opierać wyłącznie na syntaktycznych własnościach twierdzeń logicznie wynikających z siebie. Teoriomodelowa interpretacja relacji konsekwencji logicznej ma *prima facie* tę wadę, iż w ramach tego ujęcia jest ona relacją zachodzącą pomiędzy twierdzeniami na mocy tego, że twierdzenie wynikające z danego zbioru twierdzeń jest spełnione we wszystkich modelach, w których spełnione są twierdzenia należące do tego zbioru. Z kolei relacja spełniania twierdzenia w modelu wydaje się wymagać przyjęcia, że w dziedzinie tego modelu istnieją odpowiednie byty spełniające te twierdzenia. Gdyby metafikcjonalista modalny przyjął istnienie modeli, w których spełnione byłyby wynikające z siebie twierdzenia możliwościowe, to stanąłby on przed koniecznością zaakceptowania istnienia bytów znajdujących się w dziedzinie tych modeli. Jest tak dlatego, że takie byty teoriomnogościowe jak para uporządkowana  $\langle D, I \rangle$  istnieją jedynie, gdy istnieją ich elementy, czyli  $D$  i  $I$ , które z kolei istnieją jedynie wówczas, gdy istnieją ich elementy, którymi w przypadku  $D$  są byty spełniające twierdzenia możliwościowe, a w przypadku  $I$  wyrażenia oraz byty stanowiące ich ekstensje i należące do  $D$ . Z uwagi na to, że twierdzenia możliwościowe spełniane są przez światy możliwe i możliwe indywidua, gdyby funktor opowieści „Zgodnie z  $PW$ ” interpretować jako teoriomodelową relację konsekwencji, wówczas najprawdopodobniej wiązałoby się to z koniecznością uznania istnienia światów możliwych, czego metafikcjonalista modalny oczywiście nie chce czynić.

Wydaje się również, że twierdzeniami prawdziwymi zgodnie z fikcją  $PW$  powinny być twierdzenia, które ani *explicite* nie pojawiają się w jej treści, ani nie da się ich z niej wyprowadzić, stosując jedynie reguły inferencyjne wybranego systemu dedukcyjnego. W związku z tym zdania z funktorem opowieści nie wydają się wyrażać relacji konsekwencji logicznej. Mogłyby ją wyrażać, pod warunkiem że treść



fikcji *PW* zawierałaby kompletne opisy światów możliwych, a wszystkie twierdzenia, których nie da się wyprowadzić z innych za pomocą samych dedukcyjnych reguł inferencji, byłyby *explicite* wyrażone. Z uwagi jednak na to, że najwiarygodniejszym stanowiskiem, jakie może zająć fikcjonalista w sprawie treści fikcji *PW*, jest — jak twierdzą w §4.5. — stanowisko, zgodnie z którym jest ona wytworem ludzkiej aktywności o skończonej mocy twórczej, w treści tej fikcji nie będą znajdowały się pełne opisy światów możliwych, które musiałyby się tam znajdować, gdyby funktor opowieści miał wyrażać relację konsekwencji logicznej. Ponadto prawdziwość twierdzeń zgodnie z fikcją *PW* często zależy od znaczenia stałych pozalogicznych występujących w tych twierdzeniach. Przykładowo, jeśli zgodnie z fikcją *PW*, istnieje taki świat, że istnieje w tym świecie osioł i istnieją w tym świecie mówiące byty, to zgodnie z fikcją *PW*, istnieje taki świat, w którym istnieją mówiące osły. Ostatnie twierdzenie „wynika” z pierwszego na mocy tego, że w ramach fikcji *PW* akceptuje się zasadę rekombinacji. Obowiązywanie tej zasady w fikcji *PW* jest ograniczone przez znaczenia predykatów występujących w twierdzeniach prawdziwych zgodnie z *PW*, w szczególności, jej obowiązywanie jest ograniczone do tych predykatów, których znaczenia się nie wykluczają.

W (1978/1991) Lewis proponuje, by twierdzenia o postaci „Zgodnie z fikcją *f* *p*” traktować jako kontrfaktyczne okresy warunkowe o postaci „Gdyby *f* było prawdziwe, *p* byłoby prawdziwe”. Kontrfaktyczna zależność tego typu tym różni się od logicznej konsekwencji, że pozwala na uwzględnienie szeregu pozalogicznych faktów przy próbie oceny wartości logicznej wyrażającego ją twierdzenia. Zgodnie z możliwoświatową semantyką dla kontrfaktycznych okresów warunków zaproponowaną w (Lewis 1973, cz. 1)

(CF) kontrfaktyczny okres warunkowy „Gdyby było tak, że *p*, to byłoby tak, że *q*” jest prawdziwy wtw, gdy pewien świat, w którym „*p*” i „*q*” są prawdziwe, jest bliższy światu aktualnemu

niż jakikolwiek ze światów, w którym „ $p$ ” jest prawdziwe i „ $q$ ” jest fałszywe<sup>30</sup>.

Relacja bliskości zachodząca pomiędzy światami jest stopniowalną relacją definiowaną w kategoriach całościowego podobieństwa tych światów, tj. podobieństwa pod względem sposobu, na jaki całość rzeczy ma się w tych światach. Im świat  $w$  jest bliższy światu  $w^*$ , tym bardziej światy te są do siebie podobne ze względu na sposób, na jaki ma się w nich całość rzeczy. W wielu wypadkach kwestia uszeregowania światów ze względu na całościowe podobieństwo do pewnego świata może nastęrczać problemów, niemniej jednak zasadnicza intuicja stojąca za wprowadzeniem relacji bliskości pomiędzy światami w kontekście analizy kontrfaktycznych okresów warunkowych wydaje się wystarczająco jasna z punktu widzenia tego, o czym dalej będzie mowa. Chodzi mianowicie o to, by starając się oszacować wartość logiczną kontrfaktycznego okresu warunkowego „Gdyby  $p$ , to  $q$ ”, próbować ustalać wartość logiczną następnika tego kontrfaktycznego okresu warunkowego w tych światach, które różnią się od

---

<sup>30</sup>Podobna analiza warunków prawdziwości kontrfaktycznych okresów warunkowych była proponowana przez Stalnaker (zob. np. Stalnaker 1975/1999, s. 68–69) z tym wyjątkiem, że w ramach jego propozycji kontrfaktyczny okres warunkowy „Gdyby  $p$ , to byłoby  $q$ ” jest prawdziwy wtw, gdy „ $q$ ” jest prawdziwe w najbliższym świecie osiągalnym ze świata aktualnego, w którym „ $p$ ” jest prawdziwe. Jak zauważa Lewis, analiza ta „opiera się na w ogólności niewiarygodnym założeniu, że nigdy nie będzie więcej niż jednego najbliższego  $p$ -świata” (Lewis 1973/1986, s. 6). Założenie to jest niewiarygodne, ponieważ z uwagi na to, że — jak zobaczymy poniżej — relacja bliskości jest relacją całościowego podobieństwa pomiędzy światami, nie wydaje się, by drobne różnice pomiędzy światami mogły mieć zasadniczy wpływ na relację bliskości zachodzącą pomiędzy światami, w których prawdziwy jest zarówno poprzednik, jak i następnik kontrfaktycznego okresu warunkowego — a światem aktualny. Weźmy kontrfaktyczny okres warunkowy „Gdyby Szczecin leżał nad morzem Bałtyckim, byłby pod względem powierzchni największym nadmorskim miastem w Polsce”. Niech w światach  $w$  i  $w^*$  będzie prawdziwy zarówno poprzednik, jak i następnik tego kontrfaktycznego okresu warunkowego i niech jedyną różnicą pomiędzy tymi światami będzie to, że w świecie  $w$  szerokość szczecińskiej plaży wynosiłaby 100 metrów w chwili  $t$ , a w świecie  $w^*$  — 95 metrów i 2 centymetry w chwili  $t$ . Nie wydaje się, by którykolwiek ze światów, w którym Szczecin leży nad morzem, różniących się jedynie pod względem szerokości plaży, mógłby być uznany za bliższy światu aktualnemu.

świata aktualnego jedynie w takim stopniu, jaki był wymagany przez uznanie prawdziwości twierdzenia „ $p$ ”. Nałożenie na świat — w którym prawdziwy jest zarówno poprzednik, jak i następnik kontrfaktycznego okresu warunkowego — warunku, by był bliski światu aktualnemu, pozwala przy ustalaniu kontrfaktycznej zależności pomiędzy poprzednikiem i następnikiem uwzględnić szereg pozalogicznych faktów. Gdy próbujemy oszacować wartość logiczną kontrfaktycznego okresu warunkowego o postaci „Gdyby Anna pojechała z nami w góry, Karol również pojechałby z nami”, musimy próbować oszacować wartość logiczną twierdzenia „Karol pojechał z nami w góry”, w światach, w których nie tylko prawdziwe jest twierdzenie „Anna pojechała z nami góry”, ale również relacje osobiste pomiędzy odpowiednikami Karola i Anny są podobne do tych, które pomiędzy nimi aktualnie zachodzą. Poza obszarem naszego zainteresowania są światy, w których odpowiednik Karola nigdy nie poznał odpowiednika Anny albo obecność Anny w pewnym miejscu stanowi dla Karola wystarczającą rację do niepojawiania się nim.

Gdy przyjdzie nam szacować wartość logiczną kontrfaktycznych okresów warunkowych o postaci „Gdyby fikcja  $f$  była prawdziwa,  $p$  byłoby prawdziwe”, będziemy brali pod uwagę wszystkie światy bliskie światu aktualnemu, w których fikcja  $f$  jest prawdziwa, i próbowali oszacować, jaka jest wartość logiczna twierdzenia „ $p$ ”. W przypadku większości fikcji będziemy zatem rozważać znacznie mniejszą klasę światów możliwych, niż rozważalibyśmy, gdybyśmy sprawdzali, czy z treści fikcji  $f$  logicznie wynika twierdzenie  $p$ , wówczas bowiem — zakładając, że możliwe interpretacje traktujemy jako światy możliwe — sprawdzalibyśmy wartość logiczną okresu warunkowego „jeśli  $f$ , to  $p$ ” w każdym logicznie możliwym świecie.

Metafikcjonalista modalny nie może zaakceptować proponowanej przez Lewisa analizy twierdzeń z funktorem opowieści w niezmienionej postaci, ponieważ nieustannie odwołuje się ona do światów możliwych.

W przeciwnym razie albo doszłoby do samoobalenia jego stanowiska, albo stałoby się ono koliste. Pierwsza sytuacja zachodziłaby, gdyby za Lewisem uznać, że warunki prawdziwości twierdzenia „Zgodnie z *PW*, *p*” wymagają przyjęcia istnienia światów możliwych. Druga sytuacja miałaby miejsce, gdyby — ponownie stosując swoją strategię unikania zobowiązań ontologicznych co do istnienia światów — metafikcjonalista poprzedził twierdzenie „We wszystkich bliskich światach, w których jest tak, że *PW*, jest tak, że *p*” funktorem opowieści „Zgodnie z *PW*”, sprawiłoby to bowiem, że analiza tego funktora opowieści byłaby akceptowalna jedynie wówczas, gdyby poprzedziło się ją tym funktorem opowieści. Wówczas jednak analiza ta zakładałaby to, co miało być analizowane.

Naturalnie nasuwającym sposobem obejścia tej trudności jest przyjęcie założenia, że warunki prawdziwości takich kontrfaktycznych okresów warunkowych nie powinny być analizowane w kategoriach relacji bliskości zachodzącej pomiędzy światami i prawdziwości twierdzeń w światach możliwych.

Założmy na chwilę, że można zdefiniować warunki prawdziwości dla kontrfaktycznych okresów warunkowych bez odwoływania się do światów możliwych. Czy, przyjąwszy to założenie, metafikcjonalista modalny mógłby nieco optymistyczniej zapatrywać się na szanse powodzenia swojego projektu? Zdaniem Boba Hale’a (1995, s. 65) powodów do świętowania sukcesu nie byłoby, gdyż metafikcjonalizm modalny stałby się ofiarą dylematu „tak prostego, jak śmiertelnego”, znanego jako *dylemat Hale’a*.

Dylemat ten opiera się na założeniu, że metafikcjonalista modalny akceptuje tezę, że fikcja *PW* jest fałszywa, i stanowi rezultat rozważań nad statusem modalnym jej fałszywości:

(H<sub>1</sub>) Jeśli *PW* jest koniecznie fałszywe, to kontrfaktyczny okres warunkowy (CMF) „Gdyby *PW* było prawdziwe, to *p* byłoby prawdziwe” będzie — zgodnie ze standardową semantyką —

trywialnie prawdziwy na mocy niemożliwości jego poprzednika. Wówczas (CMF) będzie prawdziwe niezależnie od tego, jakie twierdzenie znajdzie się w następniku tego okresu. Metafikcjonalista modalny chciałby jednak utrzymywać, że nie wszystkie twierdzenia są prawdziwe zgodnie z fikcją *PW*. Przykładowo, uważa on twierdzenie „Zgodnie z *PW*, nie istnieją światy możliwe” za fałszywe, jeśli jednak poprzednik tego okresu jest koniecznie fałszywy, to — zgodnie ze standardową semantyką — okres ten będzie prawdziwy.

- (H<sub>2</sub>) Jeśli fikcja *PW* jest przygodnie fałszywa, to powstaje pytanie, w jaki sposób metafikcjonalista modalny zamierza wyrazić tezę o jej możliwej prawdziwości. Jeśli standardowo postanowi zastosować schemat (MMF), wówczas otrzyma twierdzenie „Zgodnie z *PW*, istnieje świat możliwy, w którym *PW* jest prawdziwe”, które następnie jest analizowane jako „Gdyby *PW* było prawdziwe, to istniałby świat, w którym *PW* jest prawdziwe”. Odkąd jednak zakłada się prawdziwość poprzednika ostatniego kontrfaktycznego okresu warunkowego w jednym ze światów, to konsekwencją tego okresu jest okres o następującej postaci: „Gdyby *PW* było prawdziwe w świecie *w*, *PW* byłoby prawdziwe w świecie *w*”. Ostatni okres jest podstawieniem schematu „Gdyby *p* było prawdziwe w świecie *w*, to *p* byłoby prawdziwe w świecie *w*”, który jest prawdziwy dla dowolnego podstawienia twierdzenia za zmienną *p*, z twierdzeniem koniecznie fałszywym włącznie. Metafikcjonalista modalny — konkluduje Hale — nie jest zatem w stanie nadać sensu tezie, że *PW* jest przygodnie fałszywe, gdyż proponowana przez niego analiza jest spełniona nawet wówczas, gdy *PW* jest koniecznie fałszywe. Skoro nie może nadać sensu tezie o przygodnej fałszywości *PW*, nie jest jasne, czy metafikcjonalista modalny jest w ogóle w stanie unikać pierwszego członu tego dylematu.

Choć w oryginalnym sformułowaniu dylemat Hale'a wydaje się dotyczyć wyłącznie tej wersji metafikcjonalizmu modalnego, której zwolennik odrzuca istnienie światów możliwych, to stosuje się on również do tej jego wersji, w której metafikcjonalista zachowuje agnostyczną postawę wobec ich istnienia. Zgodnie z agnostyczną wersją metafikcjonalizmu modalnego użytkownicy dyskursu możliwościowego nie mają wystarczających świadectw na rzecz fałszywości bądź prawdziwości hipotezy *PW* i w związku z tym zawieszają swój sąd w sprawie jej wartości logicznej. Jednocześnie przyjmują, że hipoteza ta jest albo prawdziwa, albo fałszywa. Wówczas sformułowany przez Hale'a dylemat można odczytać jako zarzut, zgodnie z którym zwolennik agnostycznej wersji metafikcjonalizmu modalnego nie jest w stanie nadać takiego sensu hipotezie o fałszywości *PW*, by nie prowadziła ona do sformułowanego wyżej dylematu. Jeśli — można by kontynuować — dylemat ten ujawnia, iż hipoteza o fałszywości *PW* prowadzi do niepożądanych konsekwencji, to można potraktować go jako świadectwo na rzecz odrzucenia tezy o fałszywości *PW*, co z kolei może stanowić — na gruncie zasady dwuwartościowości — świadectwo na rzecz prawdziwości *PW*. Wydaje się zatem, że jeśli dylemat Hale'a jest „tak prosty, jak śmiertelny”, to jest on równie śmiertelny dla obu wersji metafikcjonalizmu modalnego.

Metafikcjonalista modalny — jak wylicza Rosen w (1995) — może próbować wybrnąć z kleszczy tego dylematu na kilka sposobów:

- (S<sub>1</sub>) Może odrzucić standardową semantykę dla kontrfaktycznych okresów warunkowych i przyjąć istnienie nietrywialnie prawdziwych oraz fałszywych kontrfaktycznych okresów warunkowych z niemożliwymi poprzednikami (w skrócie: kontrmożliwych okresów warunkowych).
- (S<sub>2</sub>) Może uznać, że hipoteza *PW* nie jest wartościowalna logicznie, gdyż niektóre z jej twierdzeń pozbawione są wartości logicznych.

- (S<sub>3</sub>) Może uznać, że funktor opowieści nie jest analizowalny jako kontrfaktyczny okres warunkowy.
- (S<sub>4</sub>) Może ograniczyć zakres zastosowań schematu (MMF) w taki sposób, by nie stosował się on do modalnego statusu samej hipotezy *PW*.

Zacznijmy od propozycji (S<sub>1</sub>). Rosen w (1995, s. 69–70) wskazuje na to, że każdy, kto zaakceptuje Noonana schemat przekładu twierdzeń modalnych na twierdzenia dotyczące treści hipotezy *PW*, zmuszony jest do zaakceptowania jej koniecznej fałszywości, gdyż — jak widzieliśmy wyżej — na gruncie tego schematu, twierdzenie „Możliwie, istnieją światy możliwe” jest fałszywe. Jest to jednak wada schematu przekładu proponowanego przez Noonana, gdyż sprawia, iż realista modalny byłby zmuszony do odrzucenia niektórych podstawień schematu: „ $p \rightarrow \Diamond p$ ”. Tym samym treść hipotezy *PW* nie byłaby zamknięta na tę implikację, co jest wysoce niepożądanym rezultatem. Ta własność przekładu Noonana sprawia, że powinien on być zarzucony, a już z pewnością nie powinien być wykorzystywany do motywowania akceptacji opcji (S<sub>1</sub>). Na tym nie koniec. Jak widzieliśmy, propozycja Noonana jest eskstrapolacją Lewis owskiego schematu przekładu z (1968/1983) na twierdzenia modalne o transświatowym zakresie, tymczasem Lewis koncentrował się w (1968/1983) na analizie twierdzeń modalnych o wewnątrzświatowym zakresie. Kwestię twierdzeń modalnych o transświatowym zakresie podjął w (1986), gdzie proponował, by twierdzenia te analizować zgodnie ze schematem (MMF<sub>2</sub>) i uznać funktory modalne w nich występujące za semantycznie redundantne. Ostatnie pozwala na wyprowadzanie (w ramach opowieści *PW*) z twierdzenia, że *p*, twierdzenia, że możliwe, że *p*, gdyż są one traktowane jako równoważne w tej opowieści. Rozróżnienia na schematy (MMF<sub>1</sub>) i (MMF<sub>2</sub>) lub dyskutowane wyżej modyfikacje tych schematów pozwalają więc na uniknięcie niepożądanych konsekwencji

schematu Noonana i powinny być przyjęte w miejsce jego propozycji. Wówczas jednak same zastosowania metafikcjonalistycznego schematu przekładu twierdzeń modalnych na twierdzenia dotyczące treści hipotezy *PW* nie uzasadniałyby jej koniecznej fałszywości.

Co prawda, w literaturze wskazuje się przypadki nietrywialnie prawdziwych i fałszywych kontrmożliwych okresów warunkowych. Przykładami okresów pierwszego typu są twierdzenia: „Gdyby każda liczba miała co najmniej dwa dzielniki, to każda liczba miałaby więcej niż jeden dzielnik”, „Gdyby Jasiowi udało się narysować kwadratowe koło, matematycy byłiby zdziwieni” itd. Przykładami okresów drugiego typu są twierdzenia: „Gdyby każda liczba miała co najmniej dwa dzielniki, to każda liczba miałaby co najwyżej jeden dzielnik”, „Gdyby istniało kwadratowe koło, to nie byłoby kołem” itd. Choć wydaje się, że istnieją przypadki nietrywialnie prawdziwych bądź nietrywialnie fałszywych kontrmożliwych okresów warunkowych, to nie bardzo wiadomo, dlaczego mielibyśmy traktować twierdzenie „Gdyby *PW* było prawdziwe, to *p* byłoby prawdziwe” jako kontrmożliwy okres warunkowy. Co miałyby nas motywować do uznania, że poprzednik tego twierdzenia jest koniecznie prawdziwy? Jeśli motywacją miało być — jak sugeruje Rosen — zastosowanie schematu przekładu Noonana do twierdzenia „Możliwie, istnieją światy możliwe”, to z czystym sumieniem możemy odrzucić hipotezę o koniecznej fałszywości *PW*.

Co więcej, jak zauważa sam Rosen, uznanie, iż hipoteza *PW* jest niemożliwa, nie jest uznaniem jej logicznej niemożliwości, wszystko bowiem wskazuje na to, że jest ona wewnętrznie spójna. Jeśli tak, to hipoteza *PW* najprawdopodobniej powinna być traktowana jako metafizycznie niemożliwa. Gdy dopuści się, by niektóre z metafizycznie niemożliwych hipotez były zarazem logicznie możliwe, można paść ofiarą drugiego członu dylematu Hale’a, dla którego postawienia — twierdzi Hale (1995b, s. 78) — wystarczy przyjęcie założenia o logicznej



możliwości hipotezy *PW*. Zatem, jeśli uznamy, że realizm modalny jest logicznie możliwy, wówczas rozwijanie strategii ( $S_1$ ) stanie się bezcelowe, a metafikcjonalista modalny wciąż będzie musiał uporać się z drugim członem dylematu Hale'a.

Rdzeniem strategii ( $S_2$ ) jest odrzucenie założenia, że hipoteza *PW* posiada wartość logiczną. Dopiero na gruncie tego założenia powstaje pytanie o status modalny tej hipotezy, który następnie można wykorzystać do sformułowania dylematu Hale'a. Co mogłoby uzasadniać wybór tej opcji? W rozdziale 3. mówiłem o tym, że Lewis definiuje światy możliwe jako maksymalne sumy mereologiczne indywiduów powiązanych pewnymi zewnętrznymi relacjami. W przypadku naszego świata są to relacje czasoprzestrzenne, ale w przypadku innych światów możliwych mogą to być relacje, które nie są egzemplifikowane w naszym świecie. Nazwijmy relacje ostatniego typu relacjami „bycia współświatowcem”<sup>31</sup>:

Relacja bycia współświatowcem — pisze Rosen — [...] jest fikcyjną relacją lub, precyzyjniej, predykat „jest współświatowcem” jest słowem, którego pierwsze sensowne użycie pojawiło się w kontekście opowieści. [...] wiarygodnym poglądem dotyczącym predykatów tego rodzaju jest to, że choć mogą one być sensownie używane w prawdziwych zdaniach o treści ich rodzimej opowieści, to nieopowiedzone [funktozem opowieści — P.W.] zdania, w których

<sup>31</sup>Wyrażenie „jest współświatowcem” jest żywą kalką angielskiego *world-mate*. Zdecydowałem się na wprowadzenie tak sztucznego terminu, ponieważ chodzi o to, by wskazać na klasę relacji, które nie zachodzą pomiędzy indywiduami istniejącymi w naszym świecie. Co prawda, korzystając z Lewis index Lewis D.owskiego pojęcia naturalności relacji, można byłoby zdefiniować relację bycia współświatowcem jako maksymalną zewnętrzną, naturalną relację zachodzącą pomiędzy wszystkimi częściami pewnego świata, wówczas jednak relacje czasoprzestrzenne również wpisywałyby się w klasę tych maksymalnych zewnętrznych i naturalnych relacji. Tymczasem chodzi o to, by mówić o wszystkich maksymalnych, naturalnych relacjach, które nie są relacjami czasoprzestrzennymi, stąd obecność tak nieporęcznego predykatu, jak „jest współświatowcem”.

one występują, są pozbawione wartości logicznej. [...] Po-  
gląd ten znany jest z Fregego jako teoria znaczenia nazw  
fikcyjnych i jeśli jest tam wiarygodny, wydaje się, że nie  
ma powodu, dla którego nie można byłoby rozciągnąć go  
na predykaty (Rosen 1995, s. 72).

Jeśli uznać hipotezę *PW* za koniunkcję twierdzeń, wówczas jednym  
z kluczowych założeń strategii ( $S_2$ ) jest kontrowersyjne:

(Z- $S_2$ ) koniunkcja twierdzeń jest pozbawiona wartości logicznej wtw,  
gdy wśród jej członów znajdują twierdzenia pozbawione wartości  
logicznej.

Na gruncie tego założenia oraz tezy, że twierdzenia (pierwszego  
rzędu) z predykatem „jest współświatowcem” są pozbawione war-  
tości logicznych, możemy uznać, iż hipoteza *PW* jest pozbawiona  
wartości logicznej. Jest tak dlatego, że hipoteza *PW* nie stanowi wy-  
łącznie koniunkcji twierdzeń pozbawionych wartości logicznych. Wśród  
jej twierdzeń znajdują się również prawdziwe twierdzenia dotyczące  
świata aktualnego. Tworzą one wspomnianą wyżej encyklopedię. Za-  
tem opowieść *PW* powinna być pozbawiona wartości logicznej, mimo  
że na jej twierdzenia składają się również twierdzenia prawdziwe. Teza  
taka nie budzi zastrzeżeń. Jednak część z twierdzeń opowieści *PW*  
można uznać za fałszywe. Będą to wszystkie twierdzenia, w których  
stwierdza się istnienie światów możliwych, które różnią się od naszego  
świata np. przebiegiem pewnych wydarzeń historycznych, nie różniąc  
się co do fundamentalnych relacji wiążących indywidua istniejące w  
tych światach. Przytoczony przez Rosena powód do uznawania takich  
twierdzeń za pozbawione wartości logicznej nie będzie miał wówczas  
zastosowania, a w obliczu wyboru pomiędzy ich prawdziwością a ich  
fałszywością metafikcjonalista z pewnością zdecyduje się na to drugie.  
W związku z tym hipoteza *PW* powinna być pozbawiona wartości lo-  
gicznej, mimo że stanowi koniunkcję tez, z których część jest fałszywa.

Ostatnie jest już dość kontrowersyjnym założeniem, które nijak ma się do zasad, którymi posługujemy się przy wartościowaniu koniunkcji. Ilekolwiek wartości logicznych przyjąlibyśmy, zazwyczaj całej koniunkcji przypisujemy „najniższą” spośród wartości, jakie przyjmują jej człony.

W ocenie strategii ( $S_2$ ) w pierwszym rzędzie należy się zastanowić, co w ogóle oznacza teza, że hipoteza *PW* jest pozbawiona wartości logicznej z uwagi na to, że kluczowy dla niej predykat „jest współświatowcem” posiada znaczenie jedynie w kontekście opowieści *PW*. Jak twierdzi Rosen, stanowisko to ma być uogólnieniem Fregego koncepcji nazw fikcyjnych na predykaty. Ostatni pisał:

Zdanie „Odyseusz został wysadzony w głębokim śnie na ląd Itaki” ma oczywiście sens. Jest jednak wątpliwe, czy występująca w nim nazwa „Odyseusz” ma znaczenie, a tym samym jest też wątpliwe, czy ma je całe zdanie. Pewne jest tyle, że kto zdanie to uznaje serio za prawdę lub za fałsz, ten przyznaje nazwie „Odyseusz” nie tylko sens, lecz również znaczenie. Orzeczenie bowiem przypisuje się — albo się go odmawia — właśnie znaczeniu. Jeżeli nie uznaje się znaczenia, to nie można mu ani przypisać orzeczenia, ani odmówić (Frege 1892/1977, s. 68–69).

Dowiadujemy się zatem, że twierdzenia, w których występują nazwy własne, są prawdziwe lub fałszywe, jedynie gdy nazwy mają odniesienia (w terminologii Fregego — znaczenia). Uogólniając tę koncepcję na predykaty mamy do wyboru dwie możliwości:

- (1) Twierdzenie (pierwszego rzędu), w których występuje  $n$ -argumentowy predykat „ $F$ ”, jest pozbawione wartości logicznej, ponieważ nie istnieją przedmioty, do których predykat ten ma zastosowanie.

- (2) Twierdzenie (pierwszego rzędu), w których występuje  $n$ -argumentowy predykat „ $F$ ”, jest pozbawione wartości logicznej, ponieważ nie istnieje byt ogólny, do którego odnosi się ten predykat.

Obie opcje wydają się opierać na semantycznej koncepcji presupozycji<sup>32</sup>, zgodnie z którą twierdzenie  $q$  jest presupozycją twierdzenia  $p$  wtw, gdy twierdzenie  $q$  nie jest konsekwencją logiczną twierdzenia  $p$ , twierdzenie  $q$  nie stanowi części treści wyrażanej przez twierdzenie  $p$  i prawdziwość twierdzenia  $q$  jest koniecznym warunkiem prawdziwości lub fałszywości twierdzenia  $p$ . Stąd o niektórych twierdzeniach możliwościowych powiemy, że są pozbawione wartości logicznej wtw, gdy nie są spełnione ich presupozycje. W przypadku opcji (1) nie jest spełnione twierdzenie „Istnieje takie  $x_1, \dots, x_n$ , że  $Fx_1, \dots, x_n$ ”, a w przypadku (2) nie jest spełnione twierdzenie „Istnieje taki a taki  $n$ -argumentowy byt ogólny”. Nie sędzę jednak, by strategia ( $S_2$ ) była poprawnym sposobem obejścia dylematu Hale’a.

Zacznijmy od rozważenia opcji (1). Podstawową trudnością, przed którą stanie jej zwolennik, jest to, że z samego faktu, iż aktualnie nie istnieją przedmioty, do których predykat „jest współświatowcem” miałby zastosowanie, nie wynika jeszcze, że nie mogłyby one istnieć. W związku z tym nie wynika z tego jeszcze, że presupozycje twierdzeń, w których występuje ten predykat, nie mogłyby być spełnione, a one same nie mogłyby być prawdziwe lub fałszywe. Nawet jeśli niektóre z twierdzeń hipotezy  $PW$  aktualnie nie posiadają wartości logicznych, nie wynika z tego jeszcze, że nie mogą one posiadać wartości logicznych, a sama hipoteza  $PW$  nie mogłaby być prawdziwa lub fałszywa. W zupełności wystarczy to do postawienia dylematu Hale’a, z tym że obecnie nie opierałby on się na założeniu, że hipoteza  $PW$  jest fałszywa, a jedynie na założeniu, że mogłaby być fałszywa. Wówczas w członie ( $H_1$ ) tego dylematu wyprowadzano by konsekwencje tezy,

<sup>32</sup>W sprawie semantycznej koncepcji presupozycji, zob. (Strawson 1950/1967, s. 392–393) i (Frege 1892/1977, s. 77–78).

że gdyby spełnione były presupozycje wszystkich twierdzeń składających się na hipotezę *PW*, to byłyaby ona z konieczności fałszywa<sup>33</sup>. Natomiast w członie ( $H_2$ ) wyprowadzano by konsekwencje z tezy, że gdyby spełnione były presupozycje tych twierdzeń, to hipoteza *PW* byłaby przygodnie fałszywa.

Metafikcjonalista, który zdecyduje się na opcję (1), musi również wyjaśnić, co oznacza teza, że nie istnieją przedmioty, do których stosuje się predykat „jest współświatowcem”. Z założenia ma on nie stosować się do indywiduów istniejących w świecie aktualnym. Dlaczego tak miałoby być? Czy któreś z cech indywiduów istniejących w świecie aktualnym wykluczają możliwość wchodzenia w relacje bycia współświatowcem? Jeśli znaczenie tego predykatu jest zgodne z zasadą rekombinacji przyjmowaną w *PW*, to predykat ten powinien mieć zastosowanie do aktualnie istniejących indywiduów, gdyż — zgodnie z tą zasadą — powinien istnieć świat możliwy, w którym istnieją duplikaty wszystkich aktualnie istniejących indywiduów (i być może kilka innych indywiduów) i są one powiązane relacją bycia współświatowcem. Nie wydaje się zatem, by wewnętrzne cechy aktualnie istniejących indywiduów wykluczały możliwości stosowania do nich predykatu „jest współświatowcem”. Być może mówienie o tym, że dane indywiduum jest powiązane relacją bycia współświatowcem z pewnym innym indywiduum, jest jedynie luźnym sposobem mówienia, gdyż z uwagi na to, że relacja ta ma maksymalny charakter (tj. wszystkie części pewnej sumy mereologicznej pozostają do siebie w tej relacji i jako relacja zewnętrzna stanowi wewnętrzną własność pewnej sumy, której częściami są członowie tej relacji), powinniśmy przyjąć, iż stosuje się ona wyłącznie do maksymalnych sum mereologicznych. Niewiele to jednak zmienia. Nawet jeśli założyć, że predykat „pozo-

---

<sup>33</sup>Tj. gdyby aktualnie istniejące indywidua były powiązane relacją bycia współświatowcem, to hipoteza *PW* byłaby koniecznie fałszywa. Zauważmy, że z założenia o istnieniu indywiduów powiązanych relacją bycia współświatowcem nie wynika teza o istnieniu wielości światów.

stają w relacji bycia współświatowcem” będzie prawdziwy o sumie mereologicznej aktualnie istniejących indywiduów, to czy wynika z tego, że nie stosuje się on do tej sumy? Czy nie można go fałszywie o niej orzec? Nie bardzo wiem, jaki miałby być powód uznania, że nie można go fałszywie orzec o sumie aktualnie istniejących indywiduów.

Założmy nawet, że zwolennik opcji (1) wskazał przekonujące racje na rzecz tezy, iż predykat „jest współświatowcem” stosuje się tylko do tych sum mereologicznych, o których można go prawdziwie orzec. Jeśli założymy, że presupozycje twierdzeń, w których pojawia się ten predykat, mogłyby być spełnione, musimy również założyć, że twierdzenia te mogłyby posiadać wartości logiczne. W rezultacie dylemat Hale’a ma na nowo zastosowanie do stanowiska metafikcjonalisty.

Jeśli strategia ( $S_2$ ) w ogóle ma prowadzić do uniknięcia dylematu Hale’a, metafikcjonalista najprawdopodobniej będzie musiał przyjąć, że twierdzenia, w których występuje ten predykat, nie tylko aktualnie, ale z zasady są pozbawione wartości logicznych. W ramach proponowanego wyżej ujęcia twierdzeniami z zasady pozbawionymi wartości logicznych będą twierdzenia, których presupozycje są z konieczności fałszywe. Problem jednak w tym, że w obecnej sytuacji wiązałoby się to z uznaniem, że twierdzenie, że istnieją sumy mereologiczne, o których predykat „jest współświatowcem” jest prawdziwy, jest z konieczności fałszywe. Z uwagi na to, że każda suma mereologiczna, o której predykat ten byłby prawdziwy, byłaby Lewisowskim światem możliwym, uznanie, że presupozycja tego twierdzenia jest z konieczności fałszywa, jest uznaniem, że niektóre z tez składający się na hipotezę *PW* (mianowicie wszystkie te tezy, w których mówi się, że istnieją sumy mereologiczne indywiduów powiązanych relacją bycia współświatowcem) są z konieczności fałszywe. Jeśli hipoteza *PW* ma być koniunkcją twierdzeń, to wówczas będzie ona koniecznie fałszywa z uwagi na konieczną fałszywość jednego z jej członów. Nawet jeśli dopuszcza się luki prawdziwościowe, dowolna koniunkcja, której członem

jest fałszywe twierdzenie, powinna być fałszywa. Jeśli hipoteza *PW* jest koniecznie fałszywa, to w zależności od tego, czy mamy tutaj do czynienia z niemożliwością logiczną, czy niemożliwością metafizyczną, zastosowanie ma — odpowiednio — albo pierwszy, albo drugi człon dylematu Hale’a.

Powyższe rozumowanie stoi w sprzeczności z założeniem (Z-S<sub>2</sub>), że koniunkcja twierdzeń jest pozbawiona wartości logicznej wtw, gdy wśród jej członów znajdują twierdzenia pozbawione wartości logicznej. Jest ono jednak niewiarygodne i powyższy przykład rozumowania dobitnie to pokazuje. Z tego powodu należy odrzucić zarówno to założenie, jak i sformułowaną na jego podstawie strategię radzenia sobie z dylematem Hale’a.

Przejdźmy teraz do opcji (2), zgodnie z którą twierdzenia pierwszego rzędu, w których występuje predykat „jest współświatowcem”, są pozbawione wartości logicznej, ponieważ nie istnieje relacja, do której odnosi się ten predykat. W zależności od tego, jakie stanowisko w sprawie własności przyjmiemy, teza ta będzie miała odmienne znaczenia. Jeśli utożsamić relacje ze zbiorami *n*-tek uporządkowanych, których elementami są indywidua pozostające do siebie w tej relacji, to odrzucając istnienie takich *n*-tek uporządkowanych, automatycznie popadnie się we wszystkie problemy opcji (1). Jeśli relacje traktować po arystotelesowsku i uznać, iż istnieją tylko te z nich, które są egzemplifikowane przez indywidua, wówczas odrzucenie istnienia relacji bycia współświatowcem będzie wiązało się z przyjęciem tezy, iż nie istnieją indywidua egzemplifikujące tę relację, co — ponownie — doprowadzi do wszystkich problemów opcji (1). Z podobnych względów odrzucenia wymaga również propozycja utożsamienia tej relacji z tropami, które, jeśli założyć istnienie indywiduów, indywidualizowane są ze względu na to, które indywidua je posiadają. Ostatnią z możliwości jest uznanie relacji bycia współświatowcem za byt platoński. Z uwagi na to, że platonicy zwykle przyjmują istnienie nieegzempli-

fikowanych własności, odrzucenie istnienia tej nieegzemplifikowanej własności wydaje się nie prowadzić do trudności właściwych opcji (1). Ponadto, zwykle utrzymuje się, że byty platońskie należą do tych bytów, które istnieją bądź nie istnieją z konieczności, toteż nieistnienie naszej platońskiej relacji sprawi, że twierdzenia, w których występuje ten predykat, z zasady będą pozbawione wartości logicznych. Opcja (2) nie jest jednak pozbawiona problemów.

Nawet jeśli uznać, że predykat „jest współświatowcem” odnosi się do platońskiej relacji, to niektóre z Lewisowskich światów możliwych istniałyby, gdyby istniały sumy mereologiczne indywiduów powiązanych tą platońską relacją. Jeśli twierdzi się, że relacja ta nie istnieje z konieczności, to również z konieczności nie istnieją sumy mereologiczne indywiduów powiązanych tą relacją. Zatem z konieczności nie istnieją niektóre z Lewisowskich światów możliwych. Wśród twierdzeń składających się na hipotezę *PW* powinno się pojawić twierdzenie egzystencjalne, w którym stwierdza się istnienie takich Lewisowskich światów. Zakładając jednak, że hipoteza *PW* stanowi koniunkcję twierdzeń, będzie ona koniecznie fałszywa na mocy koniecznej fałszywości niektórych z jej członów i w zależności od tego, czy mamy tutaj do czynienia z niemożliwością logiczną, czy niemożliwością metafizyczną, zastosowanie ma — odpowiednio — albo pierwszy, albo drugi człon dylematu Hale’ a.

Jeśli metafikcjonalista modalny chciałby dalej obstawać przy strategii ( $S_2$ ), musiałby zrezygnować z proponowanego przez Rosena wyjaśnienia, dlaczego niektóre z twierdzeń egzystencjalnych pozbawione są wartości logicznych, i zaproponować taką koncepcję, która spełniałaby również zmodyfikowane założenie ( $Z-S_2$ ) o postaci:

- ( $Z-S_2^*$ ) koniunkcja twierdzeń jest pozbawiona wartości logicznej wtw, gdy wśród jej członów znajdują wyłącznie twierdzenia pozbawione wartości logicznej lub twierdzenia prawdziwe.

Wydaje się bowiem, iż jedynie koniunkcje spełniające warunek



( $Z-S_2^*$ ) można wiarygodnie uznać za koniunkcję pozbawione wartości logicznej. Niestety, celem spełnienia warunku ( $Z-S_2^*$ ) strategii ( $S_2$ ) musiałby utrzymywać, że wszystkie twierdzenia składające się na fikcję *PW* — poza twierdzeniami składającymi się na encyklopedię — są twierdzeniami pozbawionymi wartości logicznych, gdyż z pewnością nie mógłby akceptować ich prawdziwości. Tezy takiej nie da się jednak uzasadnić, nawet jeśli zgodzić się na to, że twierdzenia, w których występują wyrażenia o fikcyjnym rodowodzie, są twierdzeniami pozbawionymi wartości logicznych. Na opowieść *PW* składają się bowiem tezy, które są formułowane wyłącznie przy użyciu wyrażeń, z których każde ma zastosowanie do realnie istniejących indywiduów — są to twierdzenia, które opisują światy bliskie światu aktualnemu ze względu na zewnętrzne relacje zachodzące w tych światach i własności posiadane przez indywidua zamieszkujące te światy. Dlaczego mielibyśmy traktować takie twierdzenia jako pozbawione wartości logicznych, a nie po prostu fałszywe? Wyjaśnieniem nie może być fikcyjny rodowód wyrażeń składających się na te twierdzenia. Trudno również utrzymywać, że takie twierdzenie egzystencjalne jak „Istnieje świat możliwy, w którym polscy politycy obniżają podatki w 2013 roku” jest pozbawione wartości logicznych, ponieważ nie istnieją przedmioty, o których mają na celu mówić te twierdzenia. Wydaje się raczej, że dla kogoś, kto wątpi w istnienie światów możliwych, twierdzenie takie jest albo fałszywe, albo mogłoby okazać się, że jest ono fałszywe. Na ostatnim stanowisku powinien stać agnostyk. Wówczas jednak nie jest spełniony warunek ( $Z-S_2^*$ ), gdyż wymagane jest przyjęcie fałszywości lub możliwej fałszywości twierdzenia, że istnieje taki a taki świat. Wydaje się zatem, że strategia ( $S_2$ ) nie ma większych szans na powodzenie.

Co więcej, przyjęcie tezy, że hipoteza *PW* nie jest wartościowalna logicznie, sprawia, iż bycie metafikcjonalistą modalnym jest nie do pogodzenia z byciem agnostykiem w sprawie istnienia światów możli-

wych. Ostatni bowiem nie wyklucza ani prawdziwości, ani fałszywości hipotezy *PW*, podczas gdy zwolennik strategii ( $S_2$ ) jest zmuszony do przyjęcia tezy, iż hipoteza ta z konieczności jest pozbawiona wartości logicznej. Wykluczenie agnostycznej wersji metafikcjonalizmu modalnego wydaje się wysoce niepożądane, ponieważ — jak widzieliśmy w § 4.1. — jedną z motywacji do przyjęcia metafikcjonalizmu modalnego były rozważania epistemologiczne, w których wskazywano na to, że — z uwagi na czasoprzestrzenne odizolowanie od światów możliwych oraz brak niezależnych świadectw — nie mamy epistemicznych podstaw do uznania tej hipotezy za prawdziwą. Z podobnych względów nie mamy również podstaw do uznawania jej za fałszywą, nawet jeśli nasze intuicje podpowiadają nam, że jest ona fałszywa. W takich okolicznościach najrozsądniejszą postawą jest zawieszenie sądu w sprawie prawdziwości bądź fałszywości tej hipotezy, przy jednoczesnym dopuszczeniu możliwości, że w razie pojawienia się odpowiednich świadectw będziemy w stanie oszacować, czy jest ona prawdziwa, czy fałszywa.

Kolejną strategią, którą mógłby podążyć zwolennik metafikcjonalizmu modalnego, jest strategia ( $S_3$ ), w której odrzuca się interpretowanie twierdzeń z funktorem opowieści w kategoriach kontrfaktycznych okresów warunkowych. Jak w takim razie powinno się interpretować twierdzenia z funktorem opowieści? Jedna z propozycji przedstawia się następująco:

... wyjaśnienie fikcjonalistycznego funktora opowieści w kategoriach kontrfaktycznych okresów warunkowych nigdy nie było częścią oficjalnego stanowiska fikcjonalistycznego. Raczej było odrzucane dokładnie z tych powodów [jakie Hale wyeksponował w swoim dylemacie — P.W.]. Oficjalnie funktor opowieści jest pierwotny (Rosen 1995, s. 70, zob. również Rosen 1990, s. 346).

Proponuje się zatem, by uznać funktor opowieści za pierwotny. Z

dotychczasowych rozważań możemy jedynie powiedzieć, że funktor ten ma być (a) funktorem *niefaktaulnym*, tj. twierdzenia, które przyjmuje za argumenty, nie muszą być twierdzeniami prawdziwymi, by twierdzenia, które wyrażane są przy jego użyciu, można było uznać za prawdziwe, oraz (b) twierdzenia sformułowane przy jego użyciu nie stanowią przypadku twierdzeń, do których miałby zastosowanie dylemat Hale’a, tj. nie poddają się one parafrazie ani jako kontrfaktyczne okresy warunkowe, ani jako ścisłe implikacje.

Sam Rosen przyznaje, że „jest coś głęboko *niesatysfakcjonującego* w takim wyborze tego, co pierwotne” (1990, s. 347). Uznanie funktora opowieści za pierwotny jest nie tylko niesatysfakcjonujące, ale również niewiarygodne. Funktor ten wydaje się wyrażać jakąś relację pomiędzy twierdzeniami analizowalną poprzez odwołanie się do pojęć modalnych i pojęcia prawdziwości. Intuicje podpowiadają nam, że funktor ten wyraża konieczny związek pomiędzy prawdziwością opowieści *PW* a pewnymi twierdzeniami możliwościowymi. Mało tego, jest to związek, który najlepiej zdają się oddawać kontrfaktyczne okresy warunkowe, ale wtedy stanowisko metafikcjonalisty modalnego na nowo podatne jest na dylemat Hale’a. W kontekście dyskusji z tym dylematem, uznanie funktora opowieści za pierwotny jest tym bardziej niepokojące, że cała strategia odpierania zarzutu Hale’a wydaje się obciążona błędnym kołem. Przyjmuje się bowiem, że funktor opowieści spełnia założenia (a) i (b), tym samym zakłada się bowiem to, co miało być wykazane — mianowicie, że do funktora opowieści nie stosuje się dylemat Hale’a. Z uwagi na niewiarygodność tej propozycji przechodzę do strategii ( $S_4$ ).

Jeśli odrzucić tezę o logicznej niemożliwości hipotezy *PW* jako niewiarygodną, źródłem postawionego przez Hale’a problemu będzie analizowanie tezy o możliwej prawdziwości hipotezy *PW* za pomocą schematu (MMF). Możliwym sposobem uniknięcia tej trudności jest ograniczenie zakresu zastosowań tego schematu w taki sposób, by

możliwa prawdziwość samej hipotezy  $PW$  znajdowała się poza tym zakresem.

Naturalnie, nasuwa się pytanie, jakie racje, poza dylematem Hale'a, motywują ograniczenie zastosowań schematu (MMF). W odpowiedzi można zauważyć, że stosując schemat (MMF) w ogóle nie da się sformułować tezy równoważnej twierdzeniu modalnemu „Możliwe, że  $PW$  jest prawdziwe”, o ile nie uzna się funktora modalnego możliwości za semantycznie redundantny. Jeśli bowiem będzie ono analizowane jako twierdzenie, że zgodnie z  $PW$  istnieje świat  $w$ , w którym  $PW$  jest prawdziwe, to jest to twierdzenie prawdziwe o każdym świecie, którego istnienie przyjmuje się w ramach hipotezy  $PW$ . Jest tak dlatego, że poprawną i zgodną z tezą o wzajemnym odizolowaniu światów interpretacją twierdzenia „Zgodnie z  $PW$  istnieje świat  $w$ , w którym  $PW$  jest prawdziwe” jest twierdzenie, że zgodnie z  $PW$  ze świata  $w$  osiągalne są światy, które spełniają koniunkcję twierdzeń składających się na  $PW$ . Jeśli relację osiągalności zdefiniujemy jako relację podobieństwa ze względu na obowiązywanie praw klasycznej logiki, to ze świata  $w$  osiągalny będzie każdy ze światów, w którym, zgodnie z  $PW$ , spełnione są twierdzenia składające się na  $PW$ , czyli *de facto* każdy ze światów, których istnienie przyjmuje się zgodnie z fikcją  $PW$ . Jeśli  $PW$  jest prawdziwe w świecie  $w$ , to — w ramach fikcji  $PW$  — dla każdego świata  $v$ , osiągalnego ze świata  $w$ ,  $PW$  również powinno być prawdziwe w  $v$ . Gdyby tak nie było, tj. gdyby w świecie  $v$  było prawdą, że  $PW$  jest fałszywe, to jedynym wyjaśnieniem tego stanu rzeczy byłoby to, że ze świata  $v$  nie jest osiągalny żaden inny świat, a w związku z tym, w świecie  $v$  nie jest prawdziwe żadne z twierdzeń, które składają się na hipotezę  $PW$  i dotyczą innych światów możliwych. Ostatnie jest jednak fałszywe na gruncie realizmu modalnego. Jeśli relacją osiągalności jest relacja podobieństwa ze względu na obowiązywanie praw klasycznej logiki, to gdyby ze świata  $v$  nie był osiągalny żaden świat, oznaczałoby to, że zgodnie

z *PW* istniałby tylko jeden świat możliwy, ponieważ warunkiem koniecznym bycia światem możliwym jest bycie takim przedmiotem, w którym spełnione są prawa klasycznej logiki. Z uwagi jednak na to, że zgodnie z *PW* istnieje wiele światów możliwych i w każdym z nich spełnione są prawa klasycznej logiki, ze świata *v* osiągalny jest każdy świat możliwy, którego istnienie przyjmuje się w fikcji *PW*. Jeśli więc relacją osiągalności jest podobieństwo ze względu na obowiązywanie praw klasycznej logiki, to *PW* jest prawdziwe w każdym świecie możliwym, co na gruncie możliwościwiatowej analizy twierdzeń modalnych powinno być równoważne twierdzeniu „Konieczne, że *PW* jest prawdziwe”. Mamy więc do czynienia z sytuacją, w której zarówno ostatnie twierdzenie modalne, jak i twierdzenie modalne „Możliwe, że *PW* jest prawdziwe” są równoważne temu samemu twierdzeniu, mianowicie „Zgodnie z *PW* istnieje świat *w*, w którym *PW* jest prawdziwe”. Stanowi to podstawę do uznania semantycznej redundantności funktorów modalnych występujących w tego typu twierdzeniach modalnych. Dlatego wydaje się, iż takie podstawienie schematu (MMF), że jego lewą stroną mogłoby być prawdziwe twierdzenie modalne „Możliwe, że *PW* jest prawdziwe”, wyraża prawdziwą równoważność jedynie wówczas, gdy uzna się funktor modalny możliwości za semantycznie redundantny<sup>34</sup>.

Rzecz jasna, antyrealista bądź agnostyk w sprawie istnienia światów możliwych nie przyjąłby jako literalnie prawdziwego twierdzenia „Możliwe, że *PW* jest prawdziwe”, w którym funktor możliwości byłby semantycznie redundantny, gdyż twierdzenie takie byłoby równoważne twierdzeniu „*PW* jest prawdziwe”. Niemniej jednak mogą

<sup>34</sup>Gdyby nawet przyjąć, że funktor modalny możliwości nie jest tutaj semantycznie redundantny, to na mocy prawdziwości *PW* w każdym świecie możliwym postulowanym w *PW* twierdzenie „Możliwie, *PW* jest prawdziwe” byłoby prawdziwe na mocy tego, że byłoby konsekwencją twierdzenia „Koniecznie, *PW* jest prawdziwe”. Wówczas, stosując schemat (MMF), nie sposób dać wyraz tezie, że *PW* jest możliwie prawdziwe, której warunki prawdziwości byłyby — jak chciałby fikcjonalista modalny — spójne z tezą o nieistnieniu światów możliwych.

oni zaakceptować pierwsze z tych twierdzeń, odrzucając ostatnie bądź zawieszając jego akceptację. W związku z powyższym, metafikcjonalista modalny, który akceptuje pierwsze z tych twierdzeń jako literalnie prawdziwe, powinien również utrzymywać, że nie poddaje się ono analizie w kategoriach prawdziwości zgodnie z hipotezą *PW*.

Prowadzi to do tego, że metafikcjonalista modalny jest zobowiązany do przyjęcia pojęć modalnych nieanalizowalnych za pomocą schematu (MMF). Gdyby metafikcjonalista dodatkowo uznał, że funktor możliwości jest tutaj funktorem pierwotnym, wówczas mógłby przyjąć tezę o możliwej prawdziwości hipotezy *PW*, unikając jednocześnie drugiego członu dylematu Hale'a, który opierał się na analizie pojęcia możliwości za pomocą schematu (MMF). Ograniczając zakres zastosowania schematu (MMF), metafikcjonalista może sformułować tezę, że możliwie, hipoteza *PW* jest prawdziwa, nie traktując jej jako równoznacznej z kontrfaktycznym okresem warunkowym, który prowadził do trudności podnoszonej przez Hale'a. Metafikcjonalista modalny staje oczywiście przed koniecznością przyjęcia założenia, że pierwotny funktor możliwości jest niefaktualny, w związku z czym uznanie, że twierdzenie, że możliwe, że *PW* jest prawdziwe, jest prawdziwe — jest spójne z fałszywością *PW*. Zajęcie takiego stanowiska na powrót każe zadać pytanie o to, dlaczego nie uogólnić takiej strategii i nie zacząć po prostu kwantyfikować egzystencjalnie po światach w zasięgu funktora modalnego możliwości. Kwestię tę odkładam do §5.2.

Na początku tego paragrafu przyjąłem założenie, iż metafikcjonalista modalny jest w stanie sformułować jakąś analizę kontrfaktycznego okresu warunkowego „Gdyby *PW* było prawdziwe, to *p* byłoby prawdziwe”, która nie wymagałaby odwoływania się do światów możliwych i relacji bliskości pomiędzy nimi. Czas przyjrzeć się temu założeniu.

Nim przejdę do nakreślenia potrzebnej metafikcjonalistycznej analizy kontrfaktycznych okresów warunkowych, zacznę od wskazania racji przemawiających za tym, że okres „Gdyby *PW* było prawdziwe, to

$p$  byłoby prawdziwe” nie poddaje się możliwościwiatowej analizie. W (1990, s. 345) Rosen wskazuje racje za tym, że twierdzenia z funktorem opowieści „Zgodnie z  $PW$ ” nie poddają się możliwościwiatowej analizie. Choć w (1978/1991) Lewis zaproponował, by twierdzenia z funktorem opowieści, np. „Zgodnie z opowieścią o Sherlocku Holmesie Sherlock Holmes jest detektywem”, interpretować jako kontrfaktyczne okresy warunkowe o postaci (SH) „Gdyby opowieść o Sherlocku Holmesie była prawdziwa, to twierdzenie, że Sherlock Holmes jest detektywem, byłoby prawdziwe”, którym następnie przyporządkował następujące warunki prawdziwości:

(SH) jest prawdziwe wtw, gdy pewien świat, w którym zarówno opowieść o Sherlocku Holmesie jest prawdziwa, jak i twierdzenie, że jest on detektywem, jest prawdziwe, jest bliższy światu aktualnemu niż jakikolwiek świat, w którym opowieść o Sherlocku Holmesie jest prawdziwa i twierdzenie, że jest on detektywem, jest fałszywe.

Podobnie jak w większości kontrfaktycznych okresów warunkowych relacja podobieństwa pomiędzy światami nie będzie wyznaczała jednego świata jako tego najbliższego, ale zbiór wszystkich tych światów, które różnią się od świata aktualnego w stopniu na tyle niewielkim, na ile wymaga tego uznanie prawdziwości poprzednika tego okresu w tych światach (a być może również uznanie, skądinąd znanych, intencji autora nie wyrażonych w tekście). Niech więc fikcja o Sherlocku Holmesie będzie prawdziwa w pewnym zbiorze światów możliwych  $W_{SH}$ . Wówczas o twierdzeniu  $p$  powiemy, że jest prawdziwe zgodnie z tą fikcją wtw, gdy  $p$  jest prawdziwe we wszystkich światach należących do zbioru  $W_{SH}$ . Uogólniając,

(PZF) o twierdzeniu  $q$  powiemy, że jest prawdziwe zgodnie z fikcją  $f$  wtw, gdy twierdzenie  $p$  jest prawdziwe we wszystkich światach należących do zbioru  $W_f$ .

Rosen w (1990, s. 345–346) zauważa, że choć proponowana przez Lewisa analiza funktora opowieści dobrze radzi sobie z fikcjami, które reprezentują sposoby, na jakie rzeczy aktualnie mogłyby być, to już znacznie gorzej wypadają jej zastosowania do opowieści o transświatowej treści, które reprezentują również sposoby, na jakie są pewne światy możliwe. W takich wypadkach fikcja *PW* jest prawdziwa albo we wszystkich światach możliwych, albo w żadnym z nich. Przykładowo, wyobraźmy sobie, że zwolennik realizmu modalnego pisze opowieść zatytułowaną *Ten samotny świat*, w której proponuje wyobrażenie sobie, że istnieje tylko jeden świat, a następnie wyprowadza z tego szereg konsekwencji, zgodnie z którymi wszystko, co ma miejsce w tym świecie, ma w nim miejsce z konieczności. Załóżmy teraz, że chcemy przyjąć twierdzenie, że zgodnie z *Tym samotnym światem* żał jest nieuzasadniony. Gdyby zastosować do tego twierdzenia Lewisowską analizę, musielibyśmy poszukiwać takiego zbioru światów, w którym opowieść *Ten samotny świat* jest prawdziwa. Problem jednak w tym, że zgodnie z *Tym samotnym światem* istnieje tylko jeden świat, powinien to więc być zbiór światów, w którym nie ma innych światów. Widzieliśmy jednak wyżej, że realista modalny nie może przyjąć istnienia takich światów, gdyż musiałby założyć, że istnieją światy, z których nie są osiągalne żadne inne światy, co pozostaje w napięciu z tezą o istnieniu wielości światów, tezą, że relacja osiągalności jest relacją podobieństwa, oraz zasadą rekombinacji. Jakikolwiek ograniczenia nałożylibyśmy na relację podobieństwa, zawsze znajdzie się świat, który będzie podobny do danego świata pod względem własności wyróżnionych w warunkach, które nałożylibyśmy na relację podobieństwa pomiędzy światami. Jest tak dlatego, że zgodnie z zasadą rekombinacji, dla dowolnego świata *w*, istnieje taki świat *v*, w którym istnieją duplikaty indywiduów, istniejących w świecie *w*, duplikaty sum mereologicznych tych indywiduów itd. Zatem, zbiór światów możliwych, w których opowieść *Ten samotny świat* byłaby



prawdziwa, jest zbiorem pustym, a w związku z tym opowieść *Ten samotny świat* jest koniecznie fałszywa i zgodnie z Lewisowską semantyką dla kontrfaktycznych okresów warunkowych z niemożliwym poprzednikiem (zob. Lewis 1973, § 1.6.) kontrfaktyczny okres warunkowy „Gdyby opowieść *Ten samotny świat* była prawdziwa, to  $p$  byłoby prawdziwe” będzie trywialnie prawdziwy dla dowolnego podstawienia za  $p$ , pomimo że opowiadanie *Ten samotny świat* wydaje się mieć ustaloną treść, zgodnie z którą pewne twierdzenia ewidentnie nie są prawdziwe. Powyższa analiza funktora opowieści „Zgodnie z *Tym samotnym światem*” jest więc nieadekwatna.

Gdy z kolei mamy opowieści o światach możliwych, które są spójne z założeniami realizmu modalnego, wówczas zbiór światów, w których opowieść ta jest prawdziwa, jest zbiorem wszystkich światów możliwych, mimo że opowieści te mogą różnić się co do swoich treści. Wyobraźmy sobie, że ktoś pisze opowieść *Punkt zwrotny*, w której mówi się o dwóch światach możliwych takich, że mają one identyczne historie do pewnego momentu, potem następują wydarzenia różniące bieg historii, jaka dalej toczy się w każdym z tych światów. Powstaje pytanie, w jakich światach możliwych prawdziwa jest ta opowieść. Zdaniem Lewisa twierdzenia dotyczące dowolnego świata są prawdziwe we wszystkich pozostałych światach, a jeśli tak, to treść opowieści *Punkt zwrotny* jest utożsamiana, podobnie jak treść *PW*, ze zbiorem wszystkich światów możliwych. Konsekwencją (PZF) jest to, że dowolne twierdzenie, które jest prawdziwe zgodnie z *PW*, jest prawdziwe zgodnie z *Punktem zwrotnym* i odwrotnie. Treść *PW* różni się jednak od treści *Punktu zwrotnego*. Dlatego Lewisowska analiza funktora opowieści i w tym wypadku jest nieadekwatna.

Czym można byłoby zastąpić Lewisowską analizę funktora opowieści? Naturalnym sposobem rozwiązania przywołanych wyżej trudności jest wprowadzenie w przypadku opowieści o transświatowej treści szeroko pojętej relacji wynikania zachodzącej pomiędzy zbiorami

twierdzeń, które *explicite* zostały wyrażone w tych opowieściach, a twierdzeniami, które uznaje się za prawdziwe zgodnie z tymi opowieściami, ale które ani nie muszą należeć do zbioru twierdzeń *explicite* wyrażonych w tej opowieści, ani nie muszą dedukcyjnie wynikać z *explicite* wyrażonych twierdzeń. Jak wspomniałem na początku tego paragrafu, relacja prawdziwości twierdzenia zgodnie z treścią pewnej fikcji zwykle jest relacją zachodzącą na mocy znaczeń stałych pozalogicznych występujących w twierdzeniach składających się na *explicite* wyrażoną treść fikcji i znaczeń stałych pozalogicznych występujących w twierdzeniach prawdziwych zgodnie z tą fikcją. Z uwagi na to, że analiza twierdzeń z funktorem opowieści jako kontrfaktycznych okresów warunkowych zasadniczo pozwala na uwzględnienie znaczeń stałych pozalogicznych, metafikcjonalista modalny potrzebuje takiej analizy kontrfaktycznych okresów warunkowych, by wyrażały one pewną relację zachodzącą pomiędzy treściami twierdzeń, która nie będzie wymagała odwoływania się do semantyki światów możliwych.

Analiza taka — jak sądzę — wpisywałaby się w klasę Goodmanowskich koncepcji kontrfaktycznych okresów warunkowych. N. Goodman utrzymywał, że kontrfaktyczny okres warunkowy „Gdyby było tak, że  $p$ , to byłoby tak, że  $q$ ” jest prawdziwy wtw, gdy istnieje pewne prawdziwe twierdzenie  $s$ , spełniające warunki (a) i (b), takie, że  $q$  jest implikowane przez  $(p \& s)$  w połączeniu z koniunkcją twierdzeń wyrażających prawa natury rządzące światem aktualnym (por. Bennett 2003, s. 302). Wspomniane wyżej warunki przedstawiają się następująco:

- (a) Nie istnieje taki człon koniunkcji  $s$ , że jego prawdziwość jest przyczynowo zależna od fałszywości twierdzenia  $p$ , tj. z twierdzenia  $p$  i koniunkcji twierdzeń wyrażających prawa nie wynika żaden z członów koniunkcji  $s$  (por. Bennett 2003, s. 314).
- (b) Każdy człon koniunkcji  $s$  jest taki, że jest on logicznie niezależny od  $p$  w tym sensie, że dla dowolnego członu koniunkcji  $s$  (1) jest możliwe, że jest on prawdziwy i  $p$  jest fałszywe, (2) jest możliwe

, że jest on fałszywy i  $p$  jest prawdziwe, (3) jest możliwe, że jest on fałszywy i  $p$  jest fałszywe i (4) jest możliwe, że jest on prawdziwy i  $p$  jest prawdziwe (por. Bennett 2003, s. 317–318).

Na gruncie tej analizy kontrfaktyczny okres warunkowy „Gdybym przekręcił kluczyk w stacyjce, silnik auta rozpocząłby pracę” jest prawdziwy wtw, gdy z twierdzenia „Przekręciłem kluczyk w stacyjce” wraz z koniunkcją twierdzeń prawdziwych o okolicznościach, w których miałyby się dokonać przekręcenie kluczyka, a które mogą obejmować takie twierdzenia, jak: „Instalacja elektryczna w samochodzie jest sprawna”, „Akumulator jest naładowany”, „Pompa paliwa jest sprawna”, „Rozrusznik jest sprawny”, „Komory spalania w silniku utrzymują odpowiednie ciśnienie”, „Układ wydechowo-ssący jest sprawny”, „Układ zasilania paliwem jest sprawny” itd., w połączeniu z twierdzeniami wyrażającymi odpowiednie prawa fizyki, wynika twierdzenie „Silnik auta zapalił”.

W ramach Goodmanowskiej analizy nie przyjmuje się założenia, że twierdzenia stanowiące człony koniunkcji  $s$  muszą być znane osobie formułującej dany kontrfaktyczny okres warunkowy. Osoba ta, formułując dany kontrfaktyczny okres warunkowy, może po prostu zakładać, że istnieje koniunkcja twierdzeń (być może znanych ekspertom w danej dziedzinie), która pozwala na wyprowadzenie następnika tego okresu z jego poprzednika. Dlatego w Goodmanowskiej analizie warunków prawdziwości kontrfaktycznych okresów warunkowych „ $s$ ” występuje jako zmienna związana przebiegającą po takich koniunkcjach twierdzeń, że spełniają one warunki (a) i (b), oraz z  $(p \& s)$  w połączeniu z prawami rządzącymi światem aktualnym wynika twierdzenie  $q$  (zob. Bennett 2003, s. 302–303). Zostając przy przywołanym wyżej przykładzie, osoba formułująca ten kontrfaktyczny okres warunkowy może mieć mgliste pojęcie o mechanice samochodowej, prawach rządzących przepływem prądu przez przekładniki czy wreszcie o procesach termodynamicznych zachodzących w silniku podczas jego pracy. Jednak

wygaszając takie twierdzenie, zakłada, że istnieje koniunkcja twierdzeń  $s$ , która pozwala na wyprowadzenie następnika z poprzednika.

Konieczność odwoływania koniunkcji twierdzeń  $s$  powstaje jedynie w przypadku tzw. *zależnych* kontrfaktycznych okresów warunkowych, tj. tych kontrfaktycznych okresów warunkowych, których wartość szacuje się z uwzględnieniem partykularnych, pozajęzykowych faktów zachodzących w świecie aktualnym. Tymczasem kontrfaktyczny okres warunkowy „Gdyby  $PW$  było prawdziwe, to  $p$  byłoby prawdziwe” stanowi przypadek tzw. *niezależnego* kontrfaktycznego okresu warunkowego, przy szacowaniu wartości logicznej którego nie jest wymagane uwzględnianie partykularnych, pozajęzykowych faktów zachodzących w świecie aktualnym. W związku z tym, szacując wartość logiczną „Gdyby  $PW$  było prawdziwe, to  $p$  byłoby prawdziwe”, musimy odwołać się jedynie do treści tej opowieści (z uwzględnieniem znaczeń stałych pozalogicznych) i sprawdzić, czy *możliwe* jest, by  $PW$  było prawdziwe i  $p$  było fałszywe. Zatem „Gdyby  $PW$  było prawdziwe, to  $p$  byłoby prawdziwe” jest prawdziwe wtw, gdy nie jest *możliwe*, że  $PW$  jest prawdziwe i  $p$  jest fałszywe, gdzie pojęcie możliwości występuje jako pojęcie pierwotne. W takim kierunku powinna podążać próba rozwinięcia fikcjonalistycznej analizy twierdzeń z funktorem opowieści.

Na chwilę obecną trudno jest oszacować koszty takiej analizy funktora opowieści, gdyż wciąż nie jest jasne, jaka jest koncepcja treści fikcji  $PW$  i treści twierdzeń prawdziwych zgodnie z nią, pomiędzy którymi ma zachodzić relacja bycia takimi, że nie mogłoby być tak, że treść pierwszej jest spełniona i treść ostatniego nie jest. Do problemu treści fikcji  $PW$  przejdziemy w następnym paragrafie.

Do tej pory mówiłem jedynie o metafikcjonalizmie modalnym, należy jednak zadać pytanie, czy dylemat Hale’a ma swoje zastosowanie również do przedmiotowego fikcjonalizmu modalnego? Załóżmy, że nasz fikcjonalista przedmiotowy modalny stoi na stanowisku, że hipoteza  $PW$  jest możliwie prawdziwa. Nawet jeśli nieopatrznie zde-

cyduje się on rzucić więcej światła na znaczenie tego twierdzenia w kategoriach fikcyjnej prawdziwości twierdzeń możliwościowych, wówczas stosując reguły (conv-RG- $\diamond$ ) i (RG- $\diamond$ ), uzyskaliby równoważność (MGF):

Możliwie,  $PW$  jest prawdziwe wtw, gdy  $\ast$  istnieje taki świat  $w$ ,  
że  $PW$  jest prawdziwe w tym świecie $\ast$

(gdzie  $\ast p^\ast$  wskazuje, że twierdzenie  $p$  jest akceptowane jako prawdziwe w ramach gry w udawanie). W odróżnieniu od prawej strony równoważności (MMF)<sup>35</sup> prawa strona równoważności (MGF) nie jest podstawieniem takiego schematu zdaniowego, że jest on prawdziwy dla dowolnego podstawienia twierdzeń za zmienne zdaniowe, niezależnie od tego, czy podstawiane są twierdzenia przygodnie czy koniecznie fałszywe. Powyższa równoważność nie posiada zatem własności, na której opierał się drugi człon dylematu Hale’a.

Co prawda, przedmiotowy fikcjonalista modalny powinien być zdolny do wyrażenia pewnych tez o treści  $GPW$ , tj. o klasie twierdzeń, co do których uczestnicy gry w udawanie są zobowiązani wyobrażać sobie, iż istnieją byty spełniające te twierdzenia. Mógłby np. utrzymywać, że jeśli z koniunkcji twierdzeń składających się na  $PW$  wynika twierdzenie  $p^\ast$ , to uczestnik gry  $GPW$ , akceptując twierdzenia składające się na  $PW$ , powinien również zaakceptować twierdzenie  $p^\ast$ . Następnie można argumentować, że pojawienie się pojęcia wynikania wymaga uznania prawdziwości kontrfaktycznych okresów warunkowych o postaci „Gdyby  $PW$  było prawdziwe, to  $p^\ast$  byłoby prawdziwe”, ale — zakładając, iż wykluczamy opcję koniecznej fałszywości hipotezy  $PW$  — akceptacja takiego kontrfaktycznego okresu warunkowego nie otwiera drogi do dylematu Hale’a, ponieważ

---

<sup>35</sup>Tj.

(MMF)  $p \equiv$  Zgodnie z  $PW$ ,  $p^\ast$ ,

żaden z takich kontrfaktycznych okresów warunkowych nie jest proponowany przez zwolennika fikcjonalizmu przedmiotowego jako korelat jakiegokolwiek twierdzenia modalnego. Proponowane przez niego korelacje (via reguły generowania) przypominają przekłady proponowane przez realistów modalnych:

$$p \text{ wtw, gdy } p^*,$$

gdzie „ $p$ ” jest twierdzeniem modalnym, a „ $p^*$ ” jest możliwoświatowym przekładem, z tym wyjątkiem, że fikcjonalista przedmiotowy nakłada na prawą stronę tej równoważności warunek bycia fikcyjnie prawdziwą.

Ponadto — jak argumentowano wyżej — twierdzenie po prawej stronie równoważności „Możliwe, że  $PW$  jest prawdziwe wtw, gdy \*istnieje taki świat  $w$ , że  $PW$  jest prawdziwe w tym świecie\*” jest spełnione dla dowolnego świata przyporządkowanego zmiennej  $w$  jako jej wartość, którego istnienie wyobraża się sobie w  $GPW$ , a w związku z tym powinno być równoważne twierdzeniu „Konieczne, że  $PW$  jest prawdziwe”. W momencie, gdy twierdzenia „Możliwe, że  $PW$  jest prawdziwe” i „Konieczne, że  $PW$  jest prawdziwe” są równoważne z takim samym możliwoświatowym przekładem, wówczas funktory modalne występujące w tych twierdzeniach są semantycznie redundantne. Fikcjonalista modalny może zaakceptować oba twierdzenia modalne jedynie jako prawdziwe w ramach jakiejś modalnej gry w udawanie GM, w której istnieją takie reguły, że koniecznym i wystarczającym warunkiem prawdziwości tego twierdzenia modalnego jest prawdziwość (w  $GPW$ ) koniunkcji twierdzeń składających się na  $PW$ . Jeśli więc funktor modalny możliwości występujący po lewej stronie powyższej równoważności nie jest semantycznie redundantny, to nie istnieją reguły generowania, które mogłyby powiązać prawdziwość tego twierdzenia z fikcyjną prawdziwością pewnego twierdzenia możliwoświatowego, gdyż w ramach  $GPW$  dowolne niemodalne twierdzenie o światach możliwych jest z konieczności prawdziwe lub z konieczności

falszywe. W związku z tym zwolennik przedmiotowego fikcjonalizmu modalnego musi utrzymywać, że istnieją twierdzenia modalne, na których treść możliwościwiatowa analiza nie rzuca światła. Podobnie jak w metafikcjonalizmie modalnym, również w przedmiotowej wersji fikcjonalizmu modalnego trzeba zaakceptować twierdzenie, że możliwe, że istnieje wiele światów możliwych, jako prawdziwe, utrzymując zarazem, że nie pociąga ono za sobą zobowiązań ontologicznych co do istnienia światów możliwych. Natrafi tym samym na powtarzające się pytanie, dlaczego nie uogólnić tego podejścia i nie zacząć po prostu kwantyfikować po światach w zasięgu funktorów modalnych. Do kwestii tej wrócimy w §5.2.

## 4 Problem drobnoziarnistości treści fikcji

W swojej koncepcji ontologicznie niezobowiązujących warunków prawdziwości dla twierdzeń możliwościwiatowych fikcjonałści modalni muszą przyporządkować jakąś treść tym twierdzeniom. W metafikcjonalizmie modalnym konieczność przyporządkowania treści tym twierdzeniom pojawia się wraz z mówieniem o tym, które z nich są, a które nie są prawdziwe zgodnie z fikcją *PW*. W przedmiotowym fikcjonalizmie modalnym natomiast konieczność przyporządkowania treści tym twierdzeniom zjawia się za sprawą przypisania treści myślom uczestników gry *GPW*. W tym kontekście pojawia się palące pytanie o to, jaka jest ontologiczna natura treści (nazwijmy je — *treściami fikcji PW*), do których odwołują się zwolennicy tych stanowisk.

Udzielenie odpowiedzi na to pytanie jest tym bardziej pożądane, że fikcjonałści modalni muszą wykazać, że można ukuć taką koncepcję treści fikcji *PW*, która będzie spełniała dwa dezyderaty: ( $D_1$ ) nie będzie wymagała przyjmowania istnienia światów możliwych i możliwych indywiduów oraz ( $D_2$ ) będzie wystarczająco drobnoziarnista, by

nie doprowadzała do utożsamiania ze sobą światów możliwych bądź możliwych indywiduów, które można rozróżnić na gruncie realizmu modalnego.

Zanim przejdziemy dalej, chcę wyraźnie podkreślić, że w tej rozprawie ograniczę się jedynie do zarysowania kierunku, który mógłby obrać fikcjonalista modalny, poszukując koncepcji treści fikcji *PW*. Dlatego osoby oczekującego kompleksowej koncepcji treści fikcji z pewnością poczują się rozczarowane. Moim celem jest jednak wskazanie kilku fenomenów, które mogą stanowić dobry punkt wyjścia w konstruowaniu rozwiniętej koncepcji treści fikcji *PW*.

Zgodnie z jedną z najbardziej naturalnych koncepcji, gdy mówimy o treści fikcji, mówimy o zbiorze sądów w sensie logicznym, którego elementami są sądy wyrażane przez zdania *explicite* występujące w tej fikcji i — jeśli treść fikcji jest domknięta na implikację — sądy, które są przez nie pociągane. Z uwagi na to, że treść fikcji *PW* ma być domknięta na implikację pierwszego rzędu (zob. §5.1.), elementami zbioru sądów, z którym identyfikowano by jej treść, powinny być zarówno sądy wyrażane przez zdania możliwoświatowe *explicite* pojawiające się w tej fikcji, jak i sądy pociągane przez nie.

Gdy przyjrzeć się trzem najpopularniejszym koncepcjom sądów w sensie logicznym, żadna z nich nie spełnia obu dezyderatów ( $D_1$ ) i ( $D_2$ ).

Zgodnie z jedną z koncepcji sądów w sensie logicznym są one zbiorami światów możliwych<sup>36</sup>. Koncepcja taka nie byłaby najlepszym wyborem dla fikcjonalisty modalnego, ponieważ prowadzi do pogwałcenia dezyderatu ( $D_1$ ).

Dwie kolejne koncepcje sądów w sensie logicznym traktują je jako ustrukturyzowane byty, posiadające pewne elementy konstytutywne. Pierwszą z nich jest neorussellowska koncepcja sądów w sensie

---

<sup>36</sup>Do najbardziej wpływowych zwolenników takiej koncepcji można zaliczyć Lewisa (zob. 1986, §1.4) oraz Stalnaker (zob. 1976).



logicznym<sup>37</sup>, zgodnie z którą sądy w sensie logicznym są  $n$ -tkami uporządkowanymi, których elementami mogą być indywidua, własności i relacje. Najczęściej elementami danej  $n$ -tki uporządkowanej są obiekty identyfikowane jako wartości semantyczne wyrażeń składających się na twierdzenie wyrażające ten sąd. Dla przykładu, twierdzenie jednostkowe „Jan jest ojcem” wyraża sąd w sensie logicznym o postaci  $\langle j, F \rangle$ , gdzie  $j$  jest indywiduum przyporządkowanym nazwie „Jan” jako wartość semantyczna, a  $F$  jest własnością bycia ojcem przyporządkowaną predykatowi „jest ojcem” jako wartość semantyczna. Gdyby treść fikcji  $PW$  utożsamiać ze zbiorem tak rozumianych sądów w sensie logicznym, dowolne twierdzenie jednostkowe dotyczące pewnego świata możliwego bądź możliwego indywiduum wyrażałoby sąd, którego elementami konstytutywnymi powinien być dany świat możliwy lub dane możliwe indywiduum. Jako że sąd jest w ramach tego ujęcia obiektem teoriomnogościowym, który istnieje, gdy istnieją jego elementy konstytutywne, przyjęcie zobowiązań ontologicznych co do takich sądów będzie zarazem wymagało przyjęcia istnienia światów możliwych i możliwych indywiduów. W związku z tym utożsamienie treści fikcji  $PW$  ze zbiorem neorussellovskich sądów prowadzi do pogwałcenia dezyderatu ( $D_1$ ).

Odmianą powyższej propozycji może być koncepcja sądów w sensie logicznym, która również identyfikowałaby je z  $n$ -tkami uporządkowanymi, przy czym ich elementami konstytutywnymi miałyby być byty, które nie są kontrowersyjne z punktu widzenia fikcjonalizmu modalnego. Jednym ze sposobów uniknięcia zobowiązań ontologicznych co do istnienia możliwych indywiduów i światów możliwych jest zredukowanie ich do — odpowiednio — indywidualnych istot oraz konstrukcji powstałych poprzez wprowadzenie odpowiednich relacji pomiędzy indywidualnymi istotami. Indywidualne istoty są zbiorami takich własności, które stanowiłyby pełną jakościową charakterystykę

<sup>37</sup>Do zwolenników tej koncepcji można zaliczyć m.in. Davida Kaplana (zob. 1977/1989) oraz Marka Crimmins i Johna Perry’ego (zob. 1989).

możliwych indywiduów, gdyby te istniały. Zakłada się przy tym, że własności te są czysto ogólne, tj. można je scharakteryzować bez odwołania się do indywiduów.

Niezależnie od tego, czy utożsamimy czysto ogólne własności i relacje z platońskimi uniwersaliami, czy z jakimiś innymi bytami, propozycja konstruowania sądów składających się na fikcję *PW* przy wykorzystaniu indywidualnych istot i konstrukcji na nich prowadzi do następującego dylematu:

- a) Jeśli można z powodzeniem wykorzystać indywidualne istoty i konstrukcje na nich do utworzenia takiego zbioru sądów, z którym można utożsamić treść fikcji *PW*, to fikcjonalizm modalny powinien być zarzucony na rzecz stanowiska redukcjonistycznego, w którego ramach możliwe indywidua są redukowane do indywidualnych istot, a światy możliwe — do odpowiednich konstrukcji na indywidualnych istotach.
- b) Jeśli nie można z powodzeniem wykorzystać indywidualnych istot i konstrukcji na nich do utworzenia takiego zbioru sądów, z którym można utożsamić treść fikcji *PW*, to dowolna próba utożsamienia jej treści z takim zbiorem sądów doprowadzi do tego, że treść fikcji *PW* będzie zbyt gruboziarnista.

Przez „możliwość wykorzystania indywidualnych istot i konstrukcji na nich do utworzenia odpowiedniego zbioru sądów” rozumiem to, że po zastąpieniu światów możliwych bądź możliwych indywiduów innymi bytami będzie można przeprowadzić wszystkie rozróżnienia pomiędzy możliwymi sytuacjami, które da się przeprowadzić na gruncie realizmu modalnego.

Pierwszy człon dylematu: założmy, że indywidualne istoty mogą pełnić rolę możliwych indywiduów oraz że konstrukcje na relacjach i indywidualnych istotach mogą pełnić rolę światów możliwych. Wówczas treść fikcji *PW*, reprezentowana jako zbiór sądów, których ele-

mentami konstytutywnymi są indywidualne istoty bądź odpowiednie konstrukcje na nich, powinna być wystarczająco bogata, by można było przeprowadzić dowolne rozróżnienie pomiędzy możliwościami, które można odróżnić na gruncie realizmu modalnego. W obliczu tego, że fikcjonalista modalny przyjmowałby istnienie tak pojętej treści fikcji *PW*, powstaje pytanie, co stanowi motywację do tego, by nie zarzucić jego stanowiska na rzecz stanowiska redukcjonistycznego, zgodnie z którym światy możliwe i możliwe indywidua są redukowane do odpowiednich konstrukcji na bytach intensjonalnych?

W tym kontekście trudno byłoby o przekonujące racje na rzecz odstąpienia od akceptacji takiego redukcjonizmu. Motywując swoje stanowisko w sprawie dyskursu możliwoświatowego, fikcjonalista modalny nie mógłby np. twierdzić, że akceptacja jego pozwala na zdefiniowanie pojęcia świata możliwego bez odwoływania się do pojęć modalnych<sup>38</sup>. Co prawda, inspirowany realizmem modalnym mógłby charakteryzować światy możliwe jako takie pary uporządkowane, że ich pierwszym elementem jest relacja czasoprzestrzenna *R* bądź analogiczna do niej, a drugim elementem *n*-tka uporządkowana indywidualnych istot powiązanych relacją *R*, oraz twierdzić, że wszystkie indywidualne istoty, należące do tej *n*-tki uporządkowanej, wspólnie posiadają własność bycia maksymalną sumą mereologiczną ze względu na relację *R*<sup>39</sup>. Choć często utrzymuje się, że koncepcja Lewisa pozwala na całkowitą redukcję pojęć modalnych, to — jak wcześniej widzieliśmy — w proponowanej przez niego analizie zakłada się, że wszystkie maksymalne

<sup>38</sup>Argumentację taką można znaleźć w (Kim 2005).

<sup>39</sup>Dla uproszczenia zakładam, że wśród własności składających się na indywidualną istotę znajdują się również własności zewnętrzne, które indywidualne istoty nabywają na mocy relacji, w jakich pozostają z innymi indywidualnymi istotami. W przeciwnym razie światy możliwe należałoby traktować jako zbiory par uporządkowanych, w których obok pary uporządkowanej, której pierwszym elementem jest relacja *R* i *n*-tka uporządkowana wszystkich indywidualnych istot istniejących w tym świecie, znajdują się jeszcze pary uporządkowane, których pierwszym elementem jest dowolna z relacji zachodzących w danym świecie, a drugim *n*-tka indywidualnych istot pozostających do siebie w tej relacji.

sumy mereologiczne, z którymi utożsamia się światy możliwe, są możliwe. Założenie to nie poddaje się już możliwości światowej analizie. Nie jest więc przesądzone, że zastąpienie indywiduów indywidualnymi istotami wymusi taką charakterystykę światów możliwych, że będzie ona odwoływała się do pojęć modalnych tam, gdzie realista modalny nie musi uciekać się do nich. Świat możliwy mógłby być bowiem charakteryzowany za pomocą takich własności, które będą odpowiadały własnościom przypisywanym światom przez realistę modalnego. Weźmy pewien świat *w*, który jest maksymalną sumą mereologiczną jednego elektronu i jednego fotonu, pomiędzy którymi zachodzą relacje czasoprzestrzenne. Wówczas świat ten będzie stanowił parę uporządkowaną, której pierwszy elementem będzie relacja czasoprzestrzenna, a drugim para uporządkowana, której elementami będą indywidualna istota elektronu i indywidualna istota fotonu. Akceptując uniwersalizm mereologiczny (zgodnie z którym dla dwóch dowolnych obiektów istnieje suma mereologiczna tych obiektów), możemy uznać, że suma mereologiczna tych indywidualnych istot jest maksymalna ze względu na tę relację czasoprzestrzenną. W tym kontekście, gdyby Lewisowska koncepcja światów możliwych nie wymagała wprowadzenia pierwotnego pojęcia modalnego możliwości, to zaproponowana wyżej koncepcja abstrakcyjnego świata możliwego również nie wymagałaby odwoływania się do tego pojęcia. Jeśli zatem jedyną motywacją do niezarzucania fikcjonalizmu modalnego na rzecz stanowiska redukcjonistycznego jest właśnie próba uniknięcia przyjmowania pierwotnych pojęć modalnych w definicji pojęcia świata możliwego, to nie jest to motywacja wystarczająca, ponieważ stanowisko redukcjonistyczne radzi sobie na tym polu równie dobrze jak fikcjonalizm modalny.

Co więcej, można wskazać racje przemawiające na rzecz stanowiska redukcjonistycznego: jeśli indywidualne istoty mogą pełnić rolę możliwych indywiduów oraz da się skonstruować abstrakcyjne byty pełniące rolę światów możliwych, to jedyna różnica pomiędzy redukcjonizmem

a fikcjonalizmem modalnym byłaby taka, że pierwszy uznaje byty abstrakcyjne za byty, do których można zredukować światy możliwe i możliwe indywidua, podczas gdy w drugim uznaje się je za byty, które mają pełnić funkcję reprezentowania nieistniejących światów możliwych. Wprowadzenie ostatniego rozróżnienia ma tę wadę, że fikcjonalista modalny — w odróżnieniu od zwolennika redukcjonizmu — musi przyjąć, iż istnieją dwa sposoby, na jakie można rozumieć wypowiedzi o indywidualnych istotach i abstrakcyjnych światach. Z jednej strony, można je traktować jako wypowiedzi, których przedmiotem są same indywidualne istoty i ich własności (czyli — na gruncie poczynionych wyżej założeń — jako wypowiedzi, których przedmiotem jest treść fikcji *PW*). Wówczas pewne wypowiedzi o nich będą prawdziwe lub fałszywe na mocy własności posiadanych przez wspomniane indywidualne istoty i abstrakcyjne światy. Z drugiej strony, można je traktować jako wypowiedzi, w których ma się na celu reprezentowanie sposobów, na jakie mogłyby być nieistniejące konkretne światy możliwe bądź możliwe indywidua. Wówczas wypowiedzi te są traktowane jako fałszywe bądź zawiesza się sąd w ich sprawie, ponieważ nie ma wystarczających racji za tezą, że istnieją byty, które egzemplifikowałyby własności indywidualnych istot i skonstruowanych na ich podstawie światów możliwych. W konsekwencji fikcjonalizm modalny wymaga wprowadzenia dwóch sposobów, na jakie można wartościować twierdzenia możliwościowiatowe. Samo wprowadzenie odmiennych sposobów wartościowania równokształtnych twierdzeń może być zasadne, ale w momencie, gdy — jak na wstępie założyliśmy — ma się wszystkie środki po temu, by uniknąć wprowadzania tego rozróżnienia, redukując twierdzenia o światach możliwych do twierdzeń o bytach abstrakcyjnych, preferowane powinno być stanowisko teoretycznie oszczędniejsze, którym jest tutaj redukcjonistyczna koncepcja światów możliwych.

Drugi człon dylematu: powiedzenie, że treść fikcji *PW* jest zbyt

gruboziaриста, oznacza, że pojęcia, przy których użyciu konstruuje się jej treść, nie pozwalają na oddanie wszystkich rozróżnień, które z zamierzenia miały być oddane w tej fikcji. Jakie rozróżnienia fikcja *PW* ma z zamierzenia oddawać? Jej treść powinna być na tyle bogata, by dało się w niej przeprowadzić wszystkie rozróżnienia, które da się przeprowadzić na gruncie realizmu modalnego. W przeciwnym razie treść fikcji *PW* nie byłaby — wbrew tezom przyjmowanym przez fikcjonalistów modalnych — treścią realizmu modalnego. Powstaje więc pytanie: czy konsekwencją utożsamienia treści fikcji *PW* ze zbiorem sądów w sensie logicznym, będących *n*-tkami uporządkowanymi, których elementami są indywidualne istoty, abstrakcyjne światy możliwe, własności i relacje, nie będzie to, że jej treść stanie się zbyt gruboziaриста? Istnieją racje przemawiające za pozytywną odpowiedzią na to pytanie. Indywidualne istoty mają być zbiorami własności, które z założenia charakteryzowałoby w sposób jednoznaczny jakieś indywiduum, o ile istniałoby coś, co posiada wszystkie te własności. Problem w tym, że zastąpienie możliwych indywiduów indywidualnymi istotami zawsze prowadzi do utożsamienia jakościowo nieodróżnialnych indywiduów. Czy w ramach realizmu modalnego można odróżnić jakościowo nieodróżnialne indywidua? W paragrafie 3.1.1.1 mówiłem o świecie dwóch elektronów, krążących wokół siebie po paralelnych orbitach. Świat ten zawiera dwa jakościowo nieodróżnialne indywidua. Gdy weźmiemy wewnętrzne własności elektronów, takie np. jak ich ładunek, okaże się, iż oba indywidua posiadają takie same własności wewnętrzne. Oba elektrony pozostają do siebie również w takich samych relacjach, a w związku z tym posiadają te same własności zewnętrzne. Zgodnie z realizmem modalnym, świat dwóch elektronów jest jednym z wielu światów możliwych, a w związku z tym teza o istnieniu takiego świata powinna móc znaleźć się w treści fikcji *PW*. Problem jednak w tym, że gdy zastąpi się indywidua indywidualnymi istotami, doprowadzi to do utożsamiania elektronów istniejących w

tym świecie, ponieważ indywidualne istoty tych elektronów będą utożsamiane ze zbiorami, których elementami są takie same własności wewnętrzne i zewnętrzne, co sprawia, że — na mocy teorii mnogościowego aksjomatu ekstensjonalności — zbiory te będą identyczne. W rezultacie, dwóm odmiennym indywiduom przyporządkowuje się tę samą indywidualną istotę, przez co niemożliwe będzie — pod groźbą pogwałcenia aksjomatu ekstensjonalności — uznanie twierdzenia, że elektrony te są różne, za zdanie prawdziwe zgodnie z *PW* bądź w ramach *GPW*. Jednak twierdzenie takie powinno być prawdziwe zgodnie z fikcją *PW* bądź w ramach *GPW*. Zatem utożsamienie treści tej fikcji ze zbiorem tak rozumianych sądów sprawia, że jest ona zbyt gruboziarnista, co stanowi wystarczającą rację na rzecz odrzucenia tego utożsamienia.

Niekiedy sugeruje się, by indywidualną istotę utożsamiać nie tyle ze zbiorem własności, które mogłyby być pełną jakościową charakterystyką pewnego indywiduum relatywnie do pewnego świata *w*, gdyby świat *w* był światem aktualnym, ale z takim zbiorem własności, który stanowi pełną jakościową charakterystykę danego indywiduum relatywnie do wszystkich światów, w których ono istnieje (zob. Plantinga 1974, s. 72–77). Zmiana ta mogłaby pozwolić na odróżnienie naszych elektronów *a* i *b*, gdyby tylko istniał pewien świat możliwy, w którym elektrony te posiadają odmienne własności, bądź gdyby istniał taki świat, w którym tylko jeden z nich by istniał. Istnienie takich światów gwarantowałoby, iż zbiory własności, z którymi utożsamia się indywidualne istoty elektronów *a* i *b*, zawierają odmienne własności i w związku z tym nie są identyczne. W paragrafie 3.1.1.1 argumentowałem, że relatywizacja własności do światów również nie dostarczy podstaw do odróżnienia wspomnianych elektronów. Gdyby nawet pozostawić tę trudność na boku, utożsamienie tych elektronów z indywidualnymi istotami, które zawierają własności zrelatywizowane do światów, jest nieosiągalne dla fikcjonalisty modalnego. Z

założenia treść fikcji *PW* powinna zawierać tezę o nieistnieniu trans-światowych indywiduów (tj. indywiduów istniejących w więcej niż jednym świecie), czego konsekwencją jest to, że wśród niemodalnych własności, należących do indywidualnych istot, z którymi utożsamia się elektrony *a* i *b*, należą wyłącznie własności posiadane przez nie w świecie dwóch elektronów. Jeśli chcemy rozważać więcej niż jeden świat, gdy ustalamy własności, które mogłyby stanowić elementy danej indywidualnej istoty, to jedyne, co możemy zrobić, to uzupełnić indywidualne istoty elektronów *a* i *b* o ich własności modalne. Zgodnie z teorią odpowiedników, elektrony te posiadają własności modalne na mocy pozostawania w pewnej relacji podobieństwa do indywiduów (tutaj znowu reprezentowanych za pośrednictwem indywidualnych istot) istniejących w innych światach. Z uwagi jednak na to, że elektrony te są jakościowo nieodróżnialne, jakkolwiek weźmiemy relację podobieństwa, te same własności modalne będą elementami indywidualnych istot tych elektronów, co sprawia, że wciąż będzie identyfikowało się indywidua, które odróżnia się na gruncie realizmu modalnego. Z tego powodu nie jest spełniony dezyderat ( $D_2$ ). Okazuje się zatem, że zbiory sądów w sensie logicznym, których elementami konstytutywnymi są indywidualne istoty, nie mogą pełnić roli treści fikcji *PW*.

Jeśli projekt fikcjonalistów modalnych ma mieć choć cień szans na powodzenie, muszą oni zarysować taką koncepcję treści fikcji, która pozwoli na spełnienie dezyderatów ( $D_1$ ) i ( $D_2$ ). Powyższe rozważania pokazują, że treść twierdzeń, w których mowa o możliwych indywiduach, nie powinna być indywidualizowana przez odwołanie się do jakościowej charakterystyki tych indywiduów, ponieważ prowadzi to do utożsamienia odmiennych, aczkolwiek jakościowo nieodróżnialnych indywiduów. Treść owa również nie powinna być indywidualizowana poprzez odwołanie się do samych tych indywiduów, ponieważ — jak zakładają fikcjonałiści modalni — one nie istnieją.



W obliczu dotychczasowych rozważań jedyny znany mi kierunek, który mógłby obrać fikcjonalista modalny w analizie treści twierdzeń możliwościowych, przedstawia się następująco: (1) treść twierdzeń jednostkowych dotyczących światów możliwych i możliwych indywiduów powinna być reprezentowana za pomocą  $n$ -tek uporządkowanych, których elementami są jednostkowe reprezentacje mentalne, zwane *pojęciami jednostkowymi*, oraz mentalne reprezentacje własności bądź relacji wyrażanych przez predykaty deskryptywne występujące w tych twierdzeniach; (2) treść twierdzeń ogólnych, w których kwantyfikuje się po światach możliwych i możliwych indywiduach powinna być reprezentowana za pomocą  $n$ -tek uporządkowanych, których elementami są pojęcia nieokreślonych przedmiotów lub nieokreślonych wielości przedmiotów oraz mentalne reprezentacje własności lub relacji wyrażanych przez predykaty występujące w tych zdaniach.

Poniżej zamierzam wyjść od wprowadzenia za R. Mark Sainsburym (2005) pojęć jednostkowych, które zostaną wykorzystane do analizy treści jednostkowych twierdzeń możliwościowych (§ 4.4.1). Następnie, również za Sainsburym, wprowadzone zostaną pojęcia pluralne, które mają na celu odnoszenie się do określonych wielości przedmiotów (§ 4.4.1.). Na podstawie pojęć jednostkowych i pojęć pluralnych zostaną wprowadzone pojęcia nieokreślonych przedmiotów oraz pojęcia nieokreślonych wielości przedmiotów, które będą wykorzystywane w analizie treści możliwościowych twierdzeń ogólnych (§ 4.4.2). Wszystkie dyskutowane poniżej koncepcje będą miały jedynie szkicowy charakter, gdyż zasadniczym celem ich wprowadzenia jest nakreślenie kierunku, w jakim mogliby podążać fikcjonałiści modalni, a nie rozwijanie koncepcji treści fikcji *PW*.

### 4.1 Treść możliwościowych twierdzeń jednostkowych

Możliwoświatowe twierdzenia jednostkowe są twierdzeniami, w których występują wyrażenia referencjalne mające na celu odnoszenie się do określonego świata możliwego bądź określonego możliwego indywiduum. Do przykładów takich twierdzeń należą głównie twierdzenia, w których występują deskrypcje określone w użyciu referencjalnym, a nie atrybutywnym. Kontrast pomiędzy wspomnianymi zastosowaniami deskrypcji określonych można oddać, wskazując na ich odmienne zachowania w kontekstach modalnych. Z jednej strony, twierdzenie „Jedyny wynalazca szkieł dwuogniskowych mógłby w ogóle nie zajmować się optyką” można traktować jako twierdzenie fałszywe, ponieważ dowolne indywiduum, które spełnia deskrypcję określoną występującą na początku tego twierdzenia, spełnia również predykat „zajmuje się optyką”. W takich wypadkach deskrypcja określona „jedyny wynalazca szkieł dwuogniskowych” jest użyta w sposób atrybutywny. Z drugiej jednak strony, twierdzenie to może być interpretowane jako prawdziwe, gdyż indywiduum, które aktualnie spełnia tę deskrypcję, mogłoby nie zajmować się optyką. W takich wypadkach, deskrypcja określona „jedyny wynalazca szkieł dwuogniskowych” jest użyta w sposób referencjalny.

Bazując na tym kontraście, jako przykład jednostkowego twierdzenia możliwościowego można podać twierdzenie:

- (A) Świat dwóch elektronów mógłby być światem bardziej zaludnionym.

Z uwagi na to, że Lewis odrzucał istnienie jakościowo nieodróżnialnych światów możliwych, wyrażenie „świat dwóch elektronów” funkcjonuje w ramach fikcji *PW* jako deskrypcja określona. Z kolei z uwagi na to, że ze świata dwóch elektronów osiągalne są światy możliwe, w których istnieje więcej indywiduów niż dwa elektrony,

twierdzenie to wydaje się prawdziwe. Jednakże jest ono prawdziwe jedynie wówczas, gdy występująca w nim deskrypcja określona „świat dwóch elektronów” jest użyta w sposób referencjalny.

Kolejnym przykładem twierdzenia, w którym występują wyrażenia mające na celu odnoszenie się do poszczególnych przedmiotów, jest twierdzenie:

- (B) Oba jakościowo nieodróżnialne elektrony byłyby jakościowo odróżnialne, gdyby istniało takie trzecie indywiduum, do którego pozostawałyby w odmiennych relacjach.

Choć nie jest to twierdzenie jednostkowe *sensu stricto*, gdyż występuje w nim odwołanie się do wielości dwóch przedmiotów, to wydaje się, iż tworząc mentalną reprezentację tej wielości dwóch elektronów, podmiot będzie musiał utworzyć mentalne reprezentacje każdego z tych dwóch elektronów. Powodem tego jest to, że jeśli zgodzić się, iż predykat „pozostawałyby w odmiennych relacjach do trzeciego indywiduum” jest analizowalny jako „pozostawałyby w takich relacjach do trzeciego indywiduum, że relacje te byłyby odmienne”, to predykat „pozostawałyby w relacjach do trzeciego indywiduum” jest predykatem o dystrybutywnym znaczeniu, tj. przypisuje się w nim każdemu z tych dwóch elektronów własność pozostawania w jakiejś relacji do trzeciego indywiduum. Użycie takiego dystrybutywnego predykatu wymaga od podmiotu utworzenia reprezentacji obu elektronów, które wchodzi w skład tej wielości.

Należy tutaj podkreślić, że w kontekście jakościowej nieodróżnialności tych elektronów podmiot nie reprezentuje ich raczej za pośrednictwem treści deskryptywnych. Ponadto, z uwagi na to, że w ramach fikcji *PW* istnieje tylko jedna taka para elektronów i pluralna deskrypcja „dwa jakościowo nieodróżnialne elektrony” została tutaj użyta w sposób referencjalny, podmiot nie tworzy reprezentacji dowolnej dwójki elektronów spełniającej treść tej pluralnej deskrypcji, a raczej tworzy jednostkowe reprezentacje każdego elektronu z osobna.

Zgodnie ze sformułowaną wyżej sugestią treść jednostkowych twierdzeń możliwościowych (oraz twierdzeń, których zrozumienie wymaga od podmiotu utworzenia jednostkowych reprezentacji) jest częściowo analizowana przez odwołanie się do tzw. pojęć jednostkowych. Czym są pojęcia jednostkowe? Zgodnie z koncepcją Sainsbury'ego (2005) mają one stanowić mentalne reprezentacje określonych obiektów, które, jeśli w ogóle istnieją, stanowią odniesienia tych pojęć. Powodem dla wprowadzenia tego typu pojęć do analizy zdolności kognitywnych człowieka jest próba zdania sprawy z jego umiejętności gromadzenia i rewidowania informacji o określonych przedmiotach oraz — powiązanej z ostatnimi — umiejętności rozpoznawania przedmiotów jako tych samych w odmiennych chwilach czasu czy zdolności rozpoznania, że informacje pochodzące z rozmaitych źródeł dotyczą jednego i tego samego przedmiotu. Na gruncie założenia, że istnieją pojęcia jednostkowe, zbieranie informacji o przedmiotach jest następującym procesem: podmiot *S* ma jednostkową reprezentację przedmiotu, o którym myśli. Z reprezentacją tą kojarzone są pewne informacje o tym przedmiocie, które podmiot *S* zdobywa na takiej czy innej drodze. Załóżmy, że gromadzone przez podmiot informacje mają postać otwartych formuł '*x* jest *F*' itp. Wówczas o pewnej informacji powiemy, że jest kojarzona z reprezentacją jednostkową na sposób *s*, gdy podmiot *S* podciągający tę informację pod dane pojęcie jednostkowe posiada reprezentację typu *s* odniesienia tego pojęcia (jeśli takie istnieje) jako posiadającego własności wyszczególnione przez tę informację (Sainsbury 2005, s. 222). Jeśli sposobem podciągania informacji '*x* jest *F*' pod dane pojęcie jednostkowe jest przekonanie, wówczas podmiot powinien żywić przekonanie, że odniesienie tego pojęcia jednostkowego spełnia tę informację. Jeśli sposobem podciągania informacji '*x* jest *F*' pod dane pojęcie jednostkowe jest wyobrażanie, wówczas podmiot powinien wyobrazić sobie, że odniesienie tego pojęcia jednostkowego spełnia tę informację itd. (por. Sainsbury 2005, s. 229). Z punktu

widzenia zdolności do gromadzenia informacji „kluczową właściwością pojęć jednostkowych jest to, że są one rozciągnięte w czasie i dzięki temu służą jako stabilne punkty, do których informacje [o danym przedmiocie — P.W.] są dodawane lub [od których są — P.W.] odejmowane” (Sainsubry 2005, s. 223).

Choć wydaje się, że istnienie jednostkowych reprezentacji określonych obiektów nie jest zbyt kontrowersyjne, to już nie jest jasne, czy pojęcia jednostkowe mogą przydać się fikcjonalizmowi modalnemu. W pierwszym bowiem rzędzie pojęcia jednostkowe nie powinny stanowić zakamuflowanych deskrypcji określonych, reprezentujących dany obiekt tylko dzięki temu, iż jako jedyny posiada wszystkie własności wyróżnione w tej deskrypcji. Gdyby pojęcia jednostkowe stanowiły zamaskowane deskrypcje określone, treść fikcji *PW* — zgodnie z argumentacją na rzecz drugiego członu powyższego dylematu — byłaby zbyt gruboziarnista i nie pozwoliłaby na reprezentowanie naszych dwóch elektronów. W drugim rzędzie warunkiem posiadania pojęcia jednostkowego nie powinno być istnienie przedmiotu, który stanowi odniesienie tego pojęcia. W przeciwnym razie odwołanie się do pojęć jednostkowych byłoby bezużyteczne dla modalnego fikcjonisty, który poszukuje odpowiedzi na pytanie, co stanowi treść fikcji *PW*, zawierającej reprezentacje nieistniejących przedmiotów.

Zacznijmy od pierwszej kwestii. Otóż wydaje się, że pojęcia jednostkowe nie tylko nie stanowią zakamuflowanych deskrypcji określonych, ale nawet nie muszą zawierać w sobie informacji o istotnych cechach obiektów, do których się odnoszą, jeśli w ogóle do czegokolwiek się odnoszą. Kluczową racją przemawiającą za tym, iż pojęcia jednostkowe nie zawierają w sobie żadnej treści deskryptywnej, jest to, że podmiot *S* może żywić fałszywe przekonania o określonym przedmiocie; fałszywe do tego stopnia, że nawet informacje dotyczące rodzaju tego przedmiotu mogą okazać się informacjami fałszywymi. Wciąż jednak przedmiot ten jest przedmiotem jego myśli, a w związku z tym

podmiot posiada pojęcie jednostkowe, które miało na celu odnoszenie się do tego przedmiotu, mimo że informacje podciągane pod to pojęcie nie były trafne.

Dla przykładu, spacerując po parku, podmiot *S* mógł zobaczyć pewien przedmiot, o którym — na podstawie przeprowadzonej przezeń obserwacji — był przekonany, iż przedmiot ten jest królikiem. Na gruncie wprowadzonego wyżej modelu gromadzenia informacji o określonych przedmiotach podmiot *S* utworzył pojęcie jednostkowe, którego celem jest odnoszenie się do przedmiotu dostrzeżonego w parku, następnie pod nowoutworzone pojęcie jednostkowe podciągnął on informacje, zgodnie z którymi widziany przez niego przedmiot był królikiem. Gdyby nawet okazało się, iż widziany przez niego obiekt w rzeczywistości nie był królikiem, a — powiedzmy — zającem, wciąż niepodobna zaprzeczyć temu, iż na skutek percepcyjnego kontaktu z przedmiotem jego myśli utworzył on pojęcie jednostkowe, którego odniesieniem był ten przedmiot. Gdyby był on w towarzystwie kogoś zdolnego do odróżnienia zająca od królika, wówczas, opierając się na autorytecie tej osoby, mógłby zrewidować przekonanie, które żywił zaraz po zobaczeniu tego przedmiotu. Rewizja jego przekonania, że miał do czynienia z królikiem, byłaby niemożliwa, gdyby pojęcia jednostkowe były zakamuflowanymi deskrypcjami określonymi, ponieważ pojęcie jednostkowe, które utworzył na skutek percepcyjnego kontaktu z przedmiotem jego myśli, odnosiłoby się wówczas do tego obiektu jedynie, gdyby obiekt ten spełniał treść zamaskowanej deskrypcji określonej, którą wyrażałoby to jednostkowe pojęcie. Jeśli jej treścią byłoby m.in. to, iż widziany przedmiot jest królikiem, to w momencie, gdy wyszło na jaw, że widziany przez niego przedmiot był jednak zającem, pojęcie jednostkowe, które utworzył na skutek percepcyjnego kontaktu z tym przedmiotem, nie odnosiłoby się do tego przedmiotu. Nie mógłby wówczas zrewidować swojego poglądu na naturę widzianego przez niego przedmiotu, ponieważ nie posiadałby niezależnego

od fałszywych informacji sposobu odnoszenia się do tego przedmiotu (por. Sainsbury 2005, s. 223–224).

Gdyby ktoś próbował argumentować przeciw powyższym rozważaniom, twierdząc, iż deskrypcja określona, którą wyraża dane pojęcie jednostkowe, powinna raczej dotyczyć okoliczności, w których pojęcie jednostkowe zostało nabyte, niż jakościowej charakterystyki przedmiotu, z którym miało się styczność, to w odpowiedzi należy zwrócić uwagę na to, że podmiot *S* może się mylić co do okoliczności, w których nabył dane pojęcie jednostkowe. Wyobraźmy sobie, że ktoś nabył pojęcie jednostkowe odnoszące się do pewnej kobiety w trakcie sylwestrowej fety. Osoba ta może nie być w stanie dostarczyć takiego opisu okoliczności, w których nabyła to pojęcie jednostkowe, który spełniałaby tylko i wyłącznie ta okoliczność, w której doszło do poznania tej określonej kobiety. Osoba ta mogła poznać również kilka innych kobiet podczas sylwestrowej zabawy, w okolicznościach, które w równym stopniu spełniają opis, którego osoba ta jest w stanie dostarczyć. Ciężko jednak twierdzić, że osoba ta nie nabyła pojęcia jednostkowego tej kobiety na skutek kontaktu z nią. Osoba ta mogłaby również mylić się co do szeregu informacji, które — w jej mniemaniu — spełnione są przez tę określoną kobietę. Jakikolwiek typ treści deskryptywnej próbowalibyśmy zawrzeć w pojęciu jednostkowym, do pomyslenia są prawdopodobne scenariusze, w których osoba posiadająca dane pojęcie jednostkowe może mieć mylne informacje na temat odniesienia tego pojęcia (jeśli takie w ogóle istnieją) albo nie być w stanie dostarczyć takich informacji, które opisywałyby tylko jeden przedmiot bądź tylko jedną okoliczność nabycia pojęcia itd.

Co więcej, można posiadać jednostkowe pojęcie, mające na celu odnoszenie się do określonego obiektu, nawet wówczas, gdy nie posiada się wiedzy, która pozwoliłaby odróżnić ten obiekt od wszystkich innych obiektów. Mogłaby mieć miejsce sytuacja, w której podmiot *S* poznaje dwóch bliźniaków. Choć nie jest w stanie odróżnić ich od

siebie, w kontakcie z jednym z nich z pewnością utworzy jednostkową reprezentację tej osoby i dzięki temu może żywić pewne przekonania na jej temat. W przeciwnym razie musielibyśmy zaakceptować intuicyjnie fałszywe twierdzenie, iż w kontakcie z jednym z tych bliźniaków podmiot ów nie żywił żadnej jednostkowej myśli na jego temat, ponieważ nie był w stanie odróżnić go od drugiego z bliźniaków (por. Sainsbury 2005, s. 226).

Wszystko wskazuje więc na to, że pojęcia jednostkowe stanowią jednostkowe reprezentacje przedmiotów, które nie są zamaskowanymi deskrypcjami określonymi. Nie oznacza to jednak, iż przy rozpoznawaniu pewnego przedmiotu jako tego samego nie liczą się żadne informacje o tym przedmiocie. Chodzi raczej o to, że choć takie informacje biorą udział w identyfikacji danego przedmiotu, to w ogólności koniecznym warunkiem posiadania pojęcia jednostkowego nie jest jego powiązanie z określoną klasą informacji, które miałyby być spełnione tylko i wyłącznie przez jeden przedmiot. Jak widzieliśmy wyżej, jednym z powodów niemożliwości trwałego skojarzenia pewnej klasy informacji z danym pojęciem jednostkowym jest to, że z biegiem czasu można zdobyć świadectwa uprawniające do rewizji informacji podciąganych pod to pojęcie, włącznie z informacjami dotyczącymi rodzaju obiektu będącego przedmiotem jego odniesienia.

W związku z powyższym zasady rządzące odnoszeniem się pojęć jednostkowych do przedmiotów przypominają zasady rządzące odnoszeniem się nazw własnych w sensie Kripkego. Podobnie jak pojęcia jednostkowe nie są one skróconymi deskrypcjami określonych i z zamierzenia odnoszą się tylko i wyłącznie do jednego przedmiotu, jeśli w ogóle odnoszą się do czegokolwiek.

Cele teoretyczne, do których wykorzystuje się pojęcia jednostkowe (tj. wyjaśnienie zdolności gromadzenia informacji o danym przedmiocie, zdolności do rozpoznawania obiektu znanego z odmiennych okoliczności jako tego samego itd.), sugerują, że następujące zasady po-



winny rządzić odnoszeniem się pojęcia jednostkowego do określonego przedmiotu (jeśli pojęcie to w ogóle się odnosi do jakiegoś przedmiotu):

- a) *Odmienne pojęcia jednostkowe mogą mieć to samo odniesienie.*  
 Źródłem tej zasady jest to, że podmiot może nabyć w odmiennych okolicznościach odmienne pojęcia jednostkowe, których odniesieniem jest ten sam przedmiot. Gdy dopuścimy taką możliwość, wówczas zdolność do rozpoznania, że gromadziło się informacje na temat tego samego przedmiotu, mimo że początkowo nie zdawano sobie z tego sprawy, może być reprezentowana jako uświadomienie sobie, że dwa pojęcia jednostkowe odnoszą się do tego samego przedmiotu (zob. Sainsbury 2005, s. 232). Konsekwencją przyjęcia tej zasady jest to, że pojęcia jednostkowe nie są indywidualizowane wyłącznie ze względu na to, do jakiego przedmiotu się odnoszą, tj. wystarczającym warunkiem identyczności pojęć jednostkowych nie jest posiadanie tego samego odniesienia. Indywidualizowanie pojęć jednostkowych ze względu na to, do czego się odnoszą, wyklucza możliwość tego, by podmiot nie wiedział, że jest w posiadaniu kilku pojęć jednostkowych, których odniesieniem jest ten sam przedmiot. Jednak sytuacja taka jak najbardziej może mieć miejsce.
- b) *Odniesienia pojęć jednostkowych nie ulegają zmianom.* Niekiedy argumentuje się, że świadectwem na rzecz tezy, iż odniesienia pojęć jednostkowych mogą ulegać zmianom, jest następująca sytuacja: załóżmy, że pojęcie jednostkowe J pierwotnie zostało utworzone, gdy podmiot *S* odnosił się do pewnej osoby o imieniu Jim. Następnie podmiot *S* spotykał wyłącznie bliźniaka Jima — Tima, którego błędnie identyfikował z Jimem. Posługując się jakimś pojęciem jednostkowym, podmiot *S* był w stanie ponownie identyfikować Tima. Ostatnią zdolność niekiedy próbuje się wyjaśniać przez uznanie, iż odniesienie pojęcia jednostkowego J uległo zmianie. W momencie, gdy podmiot *S* dowiedział się, że

błędnie identyfikował Jima z Timem, dowiedział się również, że szereg informacji o Timie traktował jako informacje o Jimie, podobnie jak szereg informacji, które zdobył o Jimie, przypisywał Timowi. Posiadłszy wiedzę, że Jim nie jest Timem, podmiot *S* dążyłby również do zrewidowania swoich poglądów na temat każdego z nich. W ramach koncepcji pojęć jednostkowych rewidowanie informacji o pewnym przedmiocie polega na wyłączaniu informacji spod pojęcia odnoszącego się do danego przedmiotu. Mając jednak w pamięci, że założyliśmy, że odniesienie pojęcia *J* ulega zmianie, zachowanie podmiotu *S* po odkryciu, że Jim nie jest Timem, można próbować opisać w następujący sposób: stwierdzenie, że podmiot mylnie traktował informacje o Timie jako informacje o Jimie najprawdopodobniej byłoby traktowane jako stwierdzenie, że podmiot od pewnej chwili błędnie traktował pojęcie jednostkowe *J* jako odnoszące się do Jima, podczas gdy odnosiło się ono wtedy już do Tima. Na skutek błędnego przekonania co do tego, kto jest odniesieniem pojęcia *J*, podmiot *S* podciągał pod pojęcie odnoszące się do Tima informacje, które brał za informacje dotyczące Jima. Podobnie, stwierdzenie, że część informacji o Jimie traktował jako informacje o Timie, byłoby traktowane jako stwierdzenie, że podmiot *S* nie dostrzegł zmiany odniesienia pojęcia *J* i kontynuował podciąganie informacji o Timie pod pojęcie, które w tamtej chwili nie odnosiło się już do Jima. Na czym jednak polega rewidowanie informacji o Jimie i Timie, gdy odniesienia pojęć jednostkowych mogą ulegać zmianie? Gdy podmiot *S* rewiduje informacje o Jimie, być może miejsce ma sytuacja, w której ponownie zmianie ulega odniesienie pojęcia jednostkowego *J* i teraz ponownie odnosi się ono do Jima, a podmiot *S* wyłącza spod tego pojęcia wszystkie informacje, które dotyczyły Tima. W czasie, gdy odniesieniem pojęcia *J* był już Tim, część informacji podciągniętych pod to

pojęcie (mianowicie te, które dotyczyły Jima) również wymaga rewizji. W związku z tym po tym, jak podmiot *S* zrewidował już informacje o Jimie i przyszła kolej na rewidowanie informacji o Timie, pojęcie *J* ponownie powinno zmienić odniesienie i teraz powinny być wyłączone spod niego wszystkie (przed momentem podciągnięte) informacje o Jimie i podciągnięte wszystkie (przed momentem wyłączone) informacje o Timie. Problem w tym, że ciężko dostrzec, w jaki sposób tak funkcjonujące pojęcia jednostkowe w ogóle mogłyby pełnić rolę stabilnych punktów, wokół których podmiot *S* gromadzi informacje. Przypadek rewidowania informacji o przedmiotach pokazuje, że dopuszczenie, by odniesienia pojęć jednostkowych ulegały zmianie, sprawia, że odniesienia pojęć jednostkowych są na tyle niestabilne, że informacje zdobyte o pewnym przedmiocie mogą okazać się po kilku chwilach informacjami błędnie podciągniętymi pod dane pojęcie jednostkowe, ponieważ właśnie zmieniło ono swoje odniesienie. Co więcej, rodzi się pytanie: jakie jednostkowe reprezentacje towarzyszyłyby podmiotowi *S*, gdyby wypowiedział zdanie „Jim i Tim są bliźniakami”? Czy podmiot odnosiłby się do Jima i Tima za pomocą tego samego pojęcia *J*, które zmienia swoje odniesienie przy okazji wystąpienia każdej z tych nazw? Oczywiście, kuszące jest powiedzenie, że już co najmniej podczas rewidowania informacji podmiot utworzył dwa pojęcia jednostkowe, z których jedno odnosi się do Jima, a drugie — do Tima. Powiedzenie czegoś takiego jest jednak o tyle niezasadne, że gdy przyjrzymy się, według jakiej zasady zmienia się odniesienie pojęcia jednostkowego *J*, to okaże się, że nic nie uzasadnia założenia, że podmiot utworzył dwa pojęcia jednostkowe w momencie, gdy dowiedział się, że Jim nie jest Timem. Gdy zadamy sobie bowiem pytanie, w jaki sposób dokonała się zmiana odniesienia pojęcia *J* z Jima na Tima, okaże się, iż najbardziej adekwatna

odpowieź na to pytanie przedstawia się następująco: odkąd większość informacji podciąganych pod to pojęcie było spełnianych przez Tima, a nie Jima, odtąd pojęcie to zmieniło swoje odniesienie z Jima na Tima. Uogólniając, możemy powiedzieć, że w ramach tej koncepcji to klasa informacji podciągniętych pod dane pojęcie jednostkowe wyznacza jego odniesienie, a nie odwrotnie. Jeśli jednak klasa informacji podciągniętych pod dane pojęcie jednostkowe wyznacza jego odniesienie, to nic nie stoi przeszkodzie temu, by rewizję informacji o Jimie i Timie opisywać w kategoriach zmian odniesienia pojęcia J. Ostatecznie wracamy więc do pytania o to, w jaki sposób tak rozumiane pojęcia mogą pełnić rolę stabilnych punktów, wokół których podmiot gromadzi informacje? Wydaje się, że jedyną drogą wyjścia z tej niestabilności pod względem odniesienia jest przyjęcie tezy, że odniesienia pojęć jednostkowych nie ulegają zmianom (por. Sainsbury 2005, s. 232–234).

- c) *Pojęcie jednostkowe ma co najwyżej jedno odniesienie.* Zasadniczą rolą pojęć jednostkowych jest umożliwienie podmiotowi *S* myślenia tylko o jednym przedmiocie i zbieranie o nim informacji. Gdyby pojęcie jednostkowe mogło odnosić się do więcej niż jednego obiektu, pojęcia jednostkowe nie mogłyby pełnić tej funkcji. W związku z tym przyjmuje się założenie, iż pojęcia jednostkowe odnoszą się co najwyżej do jednego obiektu (zob. Sainsbury 2005, s. 234). Klauzula „co najwyżej” pojawia się tutaj, dlatego że próba sformułowania koncepcji treści fikcji *PW* wymaga, by dopuszczalne były pojęcia jednostkowe, które nie posiadają odniesień.

Przejdźmy teraz do kwestii pustości pojęć jednostkowych. O pojęciu jednostkowym powiemy, że jest puste, wtw, gdy nie istnieje odniesienie tego pojęcia jednostkowego. Przyjęcie istnienia takich pojęć jednostkowych każe postawić następujące pytania: ze względu

na co puste pojęcia jednostkowe są indywidualizowane? Kiedy, jeśli kiedykolwiek, powiemy o dwóch pustych pojęciach jednostkowych, że mają na celu odnoszenie się do tego samego przedmiotu?

Zacznijmy od pierwszego z postawionych wyżej pytań. Istnieje tendencja do tego, by pojęcia jednostkowe indywidualizować częściowo przez odwołanie się do ich odniesień. O pojęciu jednostkowym powiemy, że jest częściowo indywidualizowane ze względu na jego odniesienie wtw, gdy koniecznym warunkiem identyczności pojęć jednostkowych jest posiadanie tego samego odniesienia. Jeśli fikcjonalista modalny ma mieć pożytek z pojęć jednostkowych przy próbie charakteryzowania treści fikcji *PW*, musi sprzeciwić się tej tendencji. Jeśli treścią fikcji *PW* jest treść realizmu modalnego, to powinny w niej znajdować się np. jednostkowe reprezentacje elektronów *a* i *b*, nawet jeśli elektrony te nie istnieją. Fikcjonalista modalny może co najwyżej twierdzić, iż praktyka indywidualizowania pojęć jednostkowych, w której uwzględnia się ich odniesienie, jest uzasadniona w jedynie w przypadku niepustych pojęć jednostkowych.

Formułując kryteria identyczności dla pojęć jednostkowych, zwolennik istnienia pustych pojęć powinien dokonać takiej modyfikacji warunku posiadania tego samego odniesienia, by ostatni mógł być trywialnie spełniony w przypadku, gdy pojęcia jednostkowe nie posiadają odniesienia. Niech *I* i *I\** będą pojęciami jednostkowymi, wówczas zmodyfikowany warunek posiadania tego samego odniesienia może przedstawiać się następująco:

- (O) *I* i *I\** posiadają to samo odniesienie wtw, gdy jeśli *I* odnosi się do *x* i *I\** odnosi się do *y*, to  $x = y$ .

Na mocy prawdziwości implikacji materialnej z fałszywym poprzednikiem warunek ten może być trywialnie spełniony przez pojęcia jednostkowe nie posiadające odniesień. Z uwagi na to, że prawa strona równoważności (O) nie jest warunkiem wystarczającym identyczności pojęć jednostkowych, zarówno w przypadku pustych, jak i niepustych

pojęć jednostkowych należy wskazać na kolejne warunki, których spełnienie pozwoli uznać pojęcia jednostkowe za identyczne. Jakie to warunki?

Pojęć jednostkowych  $I$  i  $I^*$  nie powinno się traktować jako identycznych, gdy spełniają one warunek (O) oraz warunek identyczności informacji podciąganych pod te pojęcia, tj. dla każdej informacji  $i$ , jeśli  $i$  jest podciągnięta pod  $I$ , to  $i$  jest podciągnięta pod  $I^*$ , i odwrotnie. Przeciw takim kryteriom identyczności pojęć jednostkowych przemawia nasza zdolność do utworzenia reprezentacji mentalnych jakościowo nieodróżnialnych elektronów  $a$  i  $b$ . Pod oba pojęcia podciąga się te same informacje, oba spełniają również warunek (O), ale mają one na celu odnoszenie się do odmiennych obiektów.

Niemniej jednak informacje podciągane pod pojęcia jednostkowe mają wpływ na funkcje, jakie pełnią one w życiu kognitywnym podmiotu. Zdaniem Sainsbury'ego pojęcia jednostkowe powinny być indywidualizowane właśnie ze względu na to, jaką funkcję pełnią w życiu kognitywnym podmiotu (zob. Sainsbury 2005, s. 236–237), oraz ze względu na mentalną historię podmiotu, który jest w posiadaniu danego pojęcia jednostkowego (Sainsbury 2005, s. 239).

Poniżej skupię się na dwóch nasuwających się sposobach rozumienia obu tych warunków, które przedstawię — odpowiednio — jako kryteria identyczności pojęć jednostkowych (IPJ) oraz (IPJ-3). Będę argumentował, że przeciw każdemu z tych kryteriów można sformułować kontrprzykłady, które stanowią powód do ich odrzucenia.

Zacznijmy od pierwszego sposobu odczytania warunków sformułowanych przez Sainsbury'ego. Zgodnie z nim pojęcia jednostkowe nie mogą być identyczne, gdy pełnią odmienne role w życiu kognitywnym podmiotu. Z uwagi na to, że podstawową rolą pojęć jednostkowych jest gromadzenie informacji o określonych obiektach, tożsamość ról w życiu kognitywnym podmiotu powinna być sformułowana w kategoriach podciągania i wyłączania informacji spod pojęć jednostkowych.

Z kolei uwzględnienie mentalnej historii podmiotu może być przestawione poprzez relatywizację podciągnięcia bądź wyłączenia pewnej informacji spod pojęcia jednostkowego do określonej chwili. Wówczas oba warunki można łącznie sformułować w następujący sposób:

- (MK) Pojęcia jednostkowe  $I$  i  $I^*$  pełnią te same role w życiu kognitywnym podmiotu  $S$  wtw, gdy dla dowolnej chwili  $t$  i dowolnej informacji  $i$ ,  $i$  jest podciągane pod pojęcie jednostkowe  $I$  w chwili  $t$  przez podmiot  $S$  wtw, gdy  $i$  jest podciągane pod pojęcie jednostkowe  $I^*$  w chwili  $t^*$  przez podmiot  $S$  oraz dla dowolnej chwili  $t^*$  i dowolnej informacji  $i^*$ ,  $i^*$  jest wyłączane spod pojęcia  $I$  w chwili  $t$  przez podmiot  $S$  wtw, gdy  $i^*$  jest wyłączane spod pojęcia  $I^*$  w chwili  $t^*$  przez podmiot  $S$ .

Jeśli za Sainsburym przyjąć założenie, że w jednej chwili nie można nabyć dwóch pojęć jednostkowych (zob. Sainsbury 2005, s. 239), to można pokusić się o sformułowanie następującego kryterium identyczności pojęć jednostkowych:

- (IPJ) Pojęcia jednostkowe  $I$  i  $I^*$  są identyczne wtw, gdy  $I$  i  $I^*$  spełniają warunek (MK) i warunek (O).

(IPJ) nie stanowi wystarczającego warunku identyczności pojęć jednostkowych, gdy mamy do czynienia z pustymi pojęciami jednostkowymi. Wprowadzenie warunku (O) jest niezbędne, ponieważ rola, jaką dane pojęcie jednostkowe pełni w życiu kognitywnym podmiotu, może być wyznaczona przez błędne przekonania podmiotu na temat przedmiotu, do którego pojęcie to ma na celu się odnosić. Wróćmy do przypadku bliźniaków Jima i Tima. Podmiot nieustannie mylił ich ze sobą. Pod pojęcie jednostkowe  $J$  (odnoszące się do Jima) podciągnął pewien zbiór informacji  $Z$ . Spośród nich jedynie niewielka część dotyczyła Jima. Choć na skutek kontaktu z Timem podmiot faktycznie wytworzył nowe pojęcie jednostkowego  $J^*$ , którego odniesieniem jest właśnie Tim, to nie rozpoznając różnicy pomiędzy Jimem

i Timem, w błędny sposób podciągał informacje o Timie pod pojęcie jednostkowe *J*. Zarazem, posiadając pojęcie jednostkowe *J\** (którego nie odróżniał od pojęcia *J*), również podciągał łącznie informacje o Jimie i Timie pod pojęcie *J\**. Zanim podmiot dowiedział się, że Jim i Tim są braćmi bliźniakami, pojęcia *J* i *J\** pełniły te same role w jego życiu kognitywnym. Jeśli to, jaką rolę pełni pojęcie jednostkowe w życiu kognitywnym podmiotu, jest wyznaczane przez jego przekonania na temat tego, jakie informacje i w jakim czasie podciąga on pod to pojęcie, wówczas w takich wypadkach, jak przypadek pojęć *J* i *J\**, pojęcia te (w czasie sprzed dokonania odkrycia, że Tim nie jest Jimem) będzie można odróżnić jedynie ze względu na to, że posiadają odmienne odniesienia. Gdyby podmiot nigdy nie dowiedział się, że Tim nie jest Jimem, role pojęć *J* i *J\** zawsze byłyby tożsame z punktu widzenia jego życia kognitywnego, jednak są to odmienne pojęcia, ponieważ nie spełniają warunku (O).

W tym miejscu nasuwa się następująca kontrargumentacja: pojęcia jednostkowe *J* i *J\** nie mogą pełnić tej samej roli w życiu kognitywnym podmiotu *S*, ponieważ podmiot utworzył pojęcie jednostkowe *J* wcześniej niż pojęcie jednostkowe *J\**. Wcześniej spotkał przecież Jima. Zarzut taki byłby słuszny, gdyby role pojęć jednostkowych nie były transparentne dla podmiotu *S*. Można byłoby wówczas twierdzić, iż pojęcia jednostkowe różnią się względem siebie rolami w życiu kognitywnym, choć z różnicy tej nie zdaje sobie sprawy podmiot *S*. Może nie wiedzieć, że utworzył pojęcia *J* i *J\** w odmiennych chwilach czasu i że w związku z tym nie jest spełnione kryterium (IPJ). Jeśli jednak to, kiedy i jakie informacje są podciągane bądź wyłączane spod pojęcia jednostkowego, ma mieć wpływ na racjonalne działania podmiotu *S*, to role pojęć jednostkowych muszą być dla podmiotu *S* transparentne. Racje racjonalnego działania powinny być znane podmiotowi podejmującemu działanie. Z uwagi na to, że podmiot może racjonalnie działać pod wpływem fałszywego przekonania co do tego, kiedy i jakie



informacje podciągał pod które pojęcie jednostkowe, rola pełniona przez to pojęcie w jego życiu kognitywnym powinna być zależna od przekonań, które żywi on na temat tego, kiedy i jakie informacje były podciągane pod dane pojęcie jednostkowe. W powyższym przykładzie podmiot *S* jest przekonany, że Tim jest Jimem, i w konsekwencji jest przekonany, że pod pojęcie jednostkowe *J\** podciąga informacje poczynawszy już od chwili, gdy spotkał Jima, mimo że *de facto* pojęcie jednostkowe *J\** zostało przez niego utworzone w późniejszej chwili, gdy spotkał Tima. Za tym, że to właśnie przekonania podmiotu co do tego, kiedy i jaką informację podciąga pod jakie pojęcie, wyznaczają role, jakie pełnią pojęcia *J* i *J\** w jego życiu kognitywnym, wydaje się również przemawiać to, że po dowiedzeniu się o tym, iż Jim nie jest Timem, podmiot najprawdopodobniej zreflektowałby się, iż żywił szereg fałszywych przekonań o Jimie. Co więcej, gdy dowiedział się, że poznał Tima w późniejszym czasie niż Jima, odkrył, że błędne było jego przekonanie co do tego, od jakiego momentu podciągał informacje pod pojęcie *J\**. Samo mówienie o żywieniu błędnych przekonań zarówno o Jimie, jak i Timie sugeruje, że nieprawidłowy byłby następujący opis tej sytuacji: podmiot nie tyle podciąga informacje zarówno pod pojęcie *J* i *J\**, ale w rzeczywistości podciąga wszystkie informacje o Timie pod pojęcie *J\**, a wszystkie informacje o Jimie pod pojęcie *J*, przy czym myli się co do tego, co jest odniesieniem pojęcia *J\**. Na czym miałyby tutaj polegać błąd podmiotu? Zapewne na tym, że żywi przekonanie, że gromadzi informacje o Jimie, podczas gdy w rzeczywistości gromadził informacje o Timie. Jednak w ramach powyższego modelu gromadzenie jakichkolwiek informacji o Jimie, czy to prawdziwych, czy fałszywych, polega na ich podciąganiu pod pojęcie jednostkowe, którego odniesieniem jest właśnie Jim, czyli pod pojęcie *J*. Podobnie, z informacjami gromadzonymi o Timie.

Jeśli we właściwy sposób ujmuję transparentność ról pojęć jednostkowych w życiu kognitywnym podmiotu, to warunek (MK) powinien

zostać przeformułowany do postaci:

- (MK<sub>2</sub>) Pojęcia jednostkowe  $I$  i  $I^*$  pełnią te same role w życiu kognitywnym podmiotu  $S$  wtw, gdy podmiot  $S$  jest przekonany, że dla dowolnej chwili  $t$  i dowolnej informacji  $i$ ,  $i$  jest podciągane przez niego pod pojęcie jednostkowe  $I$  w chwili  $t$  wtw, gdy  $i$  jest podciągane przez niego pod pojęcie jednostkowe  $I^*$  w chwili  $t$  oraz dla dowolnej chwili  $t^*$  i dowolnej informacji  $i^*$ ,  $i^*$  jest wyłączane przez niego spod pojęcia  $I$  w chwili  $t^*$  wtw, gdy  $i^*$  jest wyłączane przez niego spod pojęcia  $I^*$  w chwili  $t^*$ .

Warunek identyczności pojęć (IPJ) uzyska teraz postać:

- (IPJ<sub>2</sub>) Pojęcia jednostkowe  $I$  i  $I^*$  są identyczne wtw, gdy pojęcia  $I$  i  $I^*$  spełniają warunki (MK<sub>2</sub>) oraz (O).

Dysponując takim kryterium, możemy odróżnić niepuste pojęcia o tych samych rolach w życiu kognitywnym podmiotu na podstawie tego, że nie jest spełniony warunek (O). W przypadku pustych pojęć jednostkowych droga ta jest zamknięta. W związku z tym — na mocy kryterium (IPJ<sub>2</sub>) — wystarczającym warunkiem identyczności pustych pojęć jednostkowych powinno być spełnianie warunku (MK<sub>2</sub>).

Jednak z łatwością można wyobrazić sobie sytuacje, w których mamy do czynienia z odmiennymi pustymi pojęciami jednostkowymi, mimo że spełniają one warunek (MK<sub>2</sub>). Wyobraźmy sobie, że podmiot  $S$  dość pobieżnie czyta powieść, w której występuje dwóch bohaterów  $c$  i  $d$ , oraz że na skutek swojej nieuwagi doszedł do błędnego przekonania, iż  $c$  i  $d$  są tą samą postacią. Celem uwiarygodnienia tego scenariusza możemy założyć bohaterowie ci mają podobne imiona, powiedzmy, Jim i James. Na jego nieszczęście w fabule książki już do jej końca nie znajdują się żadne świadectwa, które przeczyłyby jego przekonaniu. Wręcz przeciwnie, losy bohaterów  $c$  i  $d$  są na tyle blisko związane, że im dalej brnie w treść tej książki, tym bardziej utwierdza się w

przekonaniu, że *c* i *d* są tym samym bohaterem pod odmiennymi nazwami. Jednak natrafiając w tekście po raz pierwszy na wzmiankę o bohaterze *c*, podmiot *S* tworzy pojęcie jednostkowe, pod które podciąga wszystkie informacje dotyczące bohatera *c*. Gdy następnie trafia na pierwszą wzmiankę o bohaterze *d*, powinien utworzyć pojęcie jednostkowe, pod które podciąga informacje o bohaterze *d*. Z uwagi na to, że nie odróżnia bohaterów *c* i *d*, pod oba pojęcia podciąga te same informacje. Gdyby w rozmowie z kimś, kto czytał tę samą książkę dowiedział się, że jednak bohaterowie *c* i *d* są odmiennymi postaciami, dowiedziałby się, iż w błędny sposób utożsamiał ze sobą dwa pojęcia jednostkowe. Po uzyskaniu tej informacji *S* z pewnością zacząłby odróżniać oba pojęcia jednostkowe.

Jednak powstaje pytanie, czy gdyby podmiot *S* nigdy nie dowiedział się, że bohaterowie *c* i *d* nie są jednym i tym samym bohaterem, to czy w ogóle utworzyłby odmienne pojęcia jednostkowe na skutek pierwszego kontaktu z nazwami „Jim” i „James” w tekście czytanej przez niego powieści? Z uwagi na to, że podmiot zreflektowałby się, że identyfikując ze sobą postacie *c* i *d*, żywił o nich błędne przekonania, gdyby dowiedział się o ich odmienności, mamy tutaj do czynienia z odmiennymi pojęciami jednostkowymi. W takim razie pytanie brzmi: co odróżnia puste pojęcia jednostkowe o tej samej roli w życiu kognitywnym podmiotu i w jaki sposób są one tworzone, że podmiot mógł w ogóle nie zdawać sobie sprawy z tego, że utworzył dwa pojęcia jednostkowe?

Poszukując odpowiedzi na pierwszy człon powyższego pytania, przejdźmy do drugiego sposobu rozumienia warunków sformułowanych przez Sainsbury’ego. Dotychczas konieczność odwoływania się do mentalnej historii podmiotu *S* prezentowana była przez relatywizowanie gromadzonych informacji do chwili, o których podmiot *S* był przekonany, że podciąga w nich dane informacje pod dane pojęcie jednostkowe. Można jednak zająć inne stanowisko w tej sprawie i

odwołania do mentalnej historii podmiotu  $S$  traktować po prostu jako chwile, w których podmiot  $S$  nabył dane pojęcie. Przy czym nieistotne byłyby tutaj przekonania podmiotu co do, kiedy nabył to pojęcie. Strategia powinna działać, jeśli tylko podzieli się założenie Sainsbury'ego, że podmiot nie może nabyć dwóch pojęć jednostkowych w tej samej chwili. Wówczas kryterium identyczności pojęć jednostkowych przedstawiałoby się następująco:

- (IPJ<sub>3</sub>) Pojęcia jednostkowe  $I$  i  $I^*$  są identyczne wtw, gdy pojęcia  $I$  i  $I^*$  zostały utworzone w tej samej chwili, spełniają warunek (MK<sub>3</sub>) oraz warunek (O),

gdzie warunek (MK<sub>3</sub>) jest następującą modyfikacją warunku (MK<sub>2</sub>):

- (MK<sub>3</sub>) Pojęcia jednostkowe  $I$  i  $I^*$  pełnią te same role w życiu kognitywnym podmiotu  $S$  wtw, gdy podmiot  $S$  jest przekonany, że dla dowolnej informacji  $i$ ,  $i$  jest podciągane przez niego pod pojęcie jednostkowe  $I$  wtw, gdy  $i$  jest podciągane przez niego pod pojęcie jednostkowe  $I^*$  oraz dla dowolnej informacji  $i^*$ ,  $i^*$  jest wyłączane przez niego spod pojęcia  $I$  wtw, gdy  $i^*$  jest wyłączane przez niego spod pojęcia  $I^*$ .

Czy kryterium takie jest adekwatne? Z pewnością pozwala ono na odróżnienie pojęć jednostkowych, które podmiot nabył w trakcie lektury powieści o Jimie i Jamesie. Choć na skutek błędnego przekonania o ich identyczności podmiot  $S$  może podciągać te same informacje pod oba pojęcia, to jeśli w przypadku pustych pojęć jednostkowych tworzonych w trakcie czytania lektury chwilą ich nabycia jest pierwsze zetknięcie się z daną postacią w tekście. W takim wypadku wiarygodna jest teza, że podmiot  $S$  nabył te pojęcia jednostkowe w odmiennych chwilach, gdyż w odmiennych chwilach zetknął się z fragmentami tekstu, w których wprowadzano te postacie. Niemniej jednak adekwatność tego kryterium zależy od poprawności założenia, zgodnie z

którym niemożliwe jest, by jednej chwili nabyć wiele pustych pojęć jednostkowych, pod które podciąga się te same informacje.

Można mieć wątpliwości co do prawdziwości tego założenia. O tym, że można utworzyć więcej niż jedno pojęcie jednostkowe w jednej chwili wydają się świadczyć następujące badania. W 1886 roku James McKeen Cattell zajął się fenomenem zdolności do oszacowania na pierwszy rzut oka liczności grup przedmiotów, w których nie występują więcej niż trzy przedmioty. Wraz z rozwojem technologii naukowcy ustalili, iż dorosła osoba jest zdolna do ustalenia liczności tak małych grup przedmiotów w czasie wynoszącym około pół sekundy. Po przekroczeniu limitu trzech przedmiotów, dorosłe osoby potrzebują dodatkowych 200–300 milisekund na doliczenie każdego kolejnego przedmiotu, który występuje w grupie przedmiotów liczniejszej niż trzy przedmioty.

Pomiędzy psychologami toczy się spór o to, w jaki sposób odbywa się ustalanie liczności grup przedmiotów do trzech przedmiotów łącznie. Zgodnie z jednym poglądem, gdy ustalamy licznosc tak małych grup przedmiotów, dodajemy do siebie kolejno każdy obiekt z osobna, choć nie uświadamiamy sobie tego procesu. W ramach tego modelu dorośli ludzie posiadają zdolność bardzo szybkiego dodawania kolejno poszczególnych przedmiotów, o ile ich liczba nie jest większa od trzech. Zgodnie z konkurencyjnym stanowiskiem, ustalanie liczności tak małych grup przedmiotów dokonuje się w jednym akcie. W ramach tego modelu jest to możliwe dzięki temu, że człowiek posiada zdolność (zwaną przeduwagowym mechanizmem integracji cech przedmiotów) do jednoczesnego wydzielenia kilku przedmiotów z pola percepcyjnego. W odróżnieniu od poprzedniego modelu szacowania liczności tak małych grup model ten nie wymaga przyjmowania założenia, że licząca osoba skupia uwagę na każdym przedmiocie z osobna, by następnie zastosować do nich algorytm kolejnego ich dodawania. Eksperymentalnym poparciem dla tezy, że uwaga nie jest skierowana

na każdy przedmiot z osobna przy ustalaniu liczności tak małych grup przedmiotów, są badania przeprowadzone na osobach, które na skutek uszkodzenia mózgu cierpiały na niezdolność do koncentrowania uwagi na przedmiotach w polu percepcyjnym. Osoby te miały duże trudności z ustaleniem liczności grup przedmiotów, począwszy od tych, które zawierały zaledwie cztery przedmioty<sup>40</sup>. Jednak zaskakująco dobrze radziły sobie z ustalaniem liczności grup zawierających nie więcej niż trzy przedmioty<sup>41</sup> (zob. Dehaene 1997, s. 66–70). Jeśli zdolność do ustalenia liczności grup przedmiotów polega m.in. na utworzeniu pojęć jednostkowych tych przedmiotów w liczbie odpowiadającej liczbie przedmiotów w danej grupie, to powyższe badania sugerują, iż podmiot powinien być w stanie utworzyć do trzech pojęć jednostkowych jednocześnie.

Zatem jedynym kontrprzykładem do (IPJ<sub>3</sub>) byłby wiarygodny scenariusz, w którym podmiot S tworzy dwa puste pojęcia jednostkowe, pod które będzie podciągał identyczne informacje. Może mieć to miejsce, gdy podmiot S natrafi w powieści na pluralne reprezentacje niewielkiej ilości jakościowo nieodróżnialnych przedmiotów. Załóżmy, że nasz podmiot czyta powieść o następującej treści: „Dwa elektrony krążą wokół siebie po paralelnych orbitach o identycznym kształcie. Elektrony te niczym nie różnią się od siebie, a poza nimi nie istnieje nic innego. Koniec”. Co dzieje się, gdy podmiot tworzy mentalną reprezentację opisanej sytuacji? Wiarygodna wydaje się następująca odpowiedź: podmiot S tworzy dwa pojęcia jednostkowe reprezentujące te elektrony i podciąga pod każde z nich te same informacje. Pojęcia te są puste, więc w trywialny sposób spełniają warunek (O). Podciągane są pod nie te same informacje, więc spełniają również warunek (MK<sub>3</sub>). Pojęcia te zostały także utworzone w tej samej chwili

<sup>40</sup>Badane osoby w 75% udzielały niepoprawnej odpowiedzi na pytanie „Ile widzisz przedmiotów?”, gdy w polu percepcyjnym znajdowały się cztery przedmioty.

<sup>41</sup>Te same osoby tylko w 8% udzielały niepoprawnej odpowiedzi na pytanie „Ile widzisz przedmiotów?”, gdy w polu percepcyjnym znajdowały się trzy przedmioty.

na skutek zetknięcia się z pluralną frazą „dwa elektrony”, a zatem spełniają one kryterium identyczności ( $IPJ_3$ ). Mimo to wydaje się, że podmiot utworzył dwa pojęcia jednostkowe, z których każde ma na celu odnoszenie się do jednego z tych dwóch przedmiotów. Jeśli tak, kryterium ( $IPJ_3$ ) nie oddaje adekwatnie warunków identyczności pojęć jednostkowych. Jak w takim może przedstawiać się odmienne kryterium?

W udzieleniu odpowiedzi na to pytanie pomocne jest przyjrzenie się temu, co w literaturze mówi się o warunkach identyczności przedmiotów fikcyjnych. Badając tę kwestię, będziemy mogli dowiedzieć się, ze względu na co ustala się tożsamość danego przedmiotu fikcyjnego, i spróbować utworzyć bazujące na tym kryterium identyczności pojęć jednostkowych.

Już w przedbiegach możemy odrzucić dwie koncepcje warunków identyczności obiektów fikcyjnych. Lewis w (1978/1991) przedmioty fikcyjne reprezentował za pomocą takich zbiorów możliwych indywiduów, które posiadają wszystkie własności przypisane im przez autora opowieści (o ile ten nie przypisywał im własności wzajemnie wykluczających się) oraz istnieją w takich światach możliwych, które różnią się od świata aktualnego w nie większym stopniu niż wymaga tego prawdziwość opowieści, w której pojawiły się te przedmioty. W ramach tej koncepcji o dwóch przedmiotach fikcyjnych powiemy, że są identyczne wtw, gdy zbiory indywiduów, za których pomocą reprezentuje się te przedmioty fikcyjne, są identyczne. Budowanie jakiegokolwiek koncepcji pojęć jednostkowych na podstawie takich kryteriów prowadziłoby do zaciągnięcia zobowiązań ontologicznych co do istnienia możliwych indywiduów. Oczywiście, fikcjonalista modalny nie może skorzystać z takiego rozwiązania. Równie bezużyteczna jest abstrakcjonistyczna koncepcja przedmiotów fikcyjnych<sup>42</sup>, zgodnie z którą — w uproszczeniu — przedmiot fikcyjny utożsamiany jest ze

---

<sup>42</sup>Zob. (Parsons 1991).

zbiorem własności, które autor przypisał temu przedmiotowi. Budowanie kryterium identyczności pojęć jednostkowych na podstawie takiej koncepcji przedmiotów prowadziłoby do utożsamienia pojęć jednostkowych odnoszących się do jakościowo nieodróżnialnych przedmiotów. Każde z nich byłoby utożsamione z reprezentacją tego samego zbioru własności. Wówczas jednak jedno pojęcie jednostkowe stanowiłoby reprezentację jakościowo nieodróżnialnych przedmiotów. Tymczasem wydaje się, że podmiot tworzy reprezentacje dwóch przedmiotów, gdy czyta przywołaną wyżej opowieść o dwóch elektronach.

Jody Azzouni zwraca uwagę, że zwykle w ramach fikcji przedmioty fikcyjne są wprowadzane przez ich twórców albo przez zastosowanie nowej nazwy własnej albo przez wprowadzenie nowej deskrypcji określonej pewnego przedmiotu. Występowanie tych samych przedmiotów w późniejszych epizodach fikcji rozpoznaje się dzięki temu, że ponownie pojawia się wprowadzona uprzednio nazwa własna, bądź dzięki temu, że pojawiło się odwołanie do uprzednio wprowadzonej deskrypcji określonej. Identyczność pomiędzy fikcyjnymi przedmiotami rządzona jest w fikcji następującą konwencją:

- (IK) Jeśli przedmioty fikcyjne są wprowadzane za pomocą odmiennych nazw bądź odmiennych deskrypcji określonych, to przedmioty te są różne, chyba że w fikcji *explicite* stwierdza się ich identyczność.

Konwencja (IK) jest praktycznie niezbędna zarówno w tworzeniu fikcji, jak i jej odbiorze. Bez niej każde wprowadzenie fikcyjnego przedmiotu powinno być uzupełnione o listę twierdzeń, w których albo uznaje się nowowprowadzony przedmiot za identyczny z jakimś uprzednio wprowadzonym przedmiotem fikcyjnym, albo uznaje się, że nie jest on identyczny z żadnym z uprzednio wprowadzonych przedmiotów fikcyjnych (zob. Azzouni 2010, s. 130–132). Wprowadzanie takich list nie tylko nadszarpnęłoby wartość literacką dzieła, ale przede wszystkim odebrało autorom możliwość wykorzystywania konwencji



(IK) do zwodzenia czytelnika co do tego, które z pozornie odmiennych obiektów ostatecznie okażą się identyczne. Z uwagi na to, że konwencja (IK) reguluje praktykę społeczną, można sformułować następujący warunek identyczności przedmiotów fikcyjnych:

(IPF) Przedmiot fikcyjny  $x$  jest identyczny z przedmiotem fikcyjnym  $y$  wtw, gdy istnieje społeczna praktyka uznawania za koekstensywne nazw bądź deskrypcji określonych, za pomocą których wprowadzono przedmioty fikcyjne  $x$  i  $y$ .

Choć konwencja (IK) jest niezbędna w recepcji fikcji, to nie sposób traktować sformułowanego na jej podstawie (IPF) jako kryterium identyczności przedmiotów fikcyjnych. Kryterium to zakłada bowiem, że wszystkie przedmioty fikcyjne są wprowadzane za pośrednictwem nazw własnych bądź deskrypcji określonych, tymczasem niekiedy przedmioty fikcyjne są również wprowadzane za pomocą takich wyrażań pluralnych jak „dwa elektrony”. (IPF) nie ma zastosowania do takich przypadków. W związku z tym nie może stanowić podstawy do sformułowania zadowalającego kryterium identyczności pojęć jednostkowych. Niemniej (IPF) wskazuje kierunek poszukiwań adekwatniejszego kryterium identyczności przedmiotów fikcyjnych, ponieważ chcąc uwzględnić przedmioty fikcyjne wprowadzane za pomocą takich środków językowych jak frazy pluralne bądź deskrypcje nieokreślone, również powinniśmy bazować na językowej praktyce społecznej posługiwania się takimi formami językowymi przy wprowadzeniu fikcyjnych przedmiotów do treści fikcji. Sugeruje to naturalne przeformułowanie kryterium (IPF) do postaci:

(IPF\*) Przedmiot fikcyjny  $x$  jest identyczny z przedmiotem fikcyjnym  $y$  wtw, gdy istnieje społeczna praktyka uznawania za koekstensywne wyrażań (nazw, deskrypcji określonych, deskrypcji nieokreślonych), za których pomocą wprowadzono przedmioty fikcyjne  $x$  i  $y$ .

U podstaw znaczenia pluralnych sposobów mówienia o przedmiotach fikcyjnych leży praktyka nieutożsamiania ze sobą przedmiotów, które należą do wielości, do której forma pluralna ma na celu się odnosić. Gdy w powieści pojawia się np. twierdzenie „Dwa psy przeraźliwie ujadają”, przedmioty, do których ma na celu odnosić się wyrażenie „dwa psy”, nie będą spełniały kryterium (IPF\*) i będą od siebie różne.

Na podstawie powyższych rozważań można pokusić się o następujące sformułowanie kryterium identityczności pojęć jednostkowych:

- (IPJ<sub>4</sub>) Pojęcia jednostkowe I i I\* są identityczne wtw, gdy I i I\* spełniają warunek (MK<sub>3</sub>), zostały nabyte w tej samej chwili oraz I i I\* zostały nabyte w okoliczności, która w ramach społecznej praktyki językowej jest podstawą do utożsamienia przedmiotów, do których pojęcia te mają na celu się odnosić.

W konsekwencji, gdy podmiot *S* zetknie się w powieści z twierdzeniem „Dwa psy przeraźliwie ujadają”, zgodnie ze społeczną praktyką powinien utworzyć pojęcia jednostkowe dwóch ujadających psów. Racją przemawiającą na rzecz tego, że podmiot utworzy dwa pojęcia jednostkowe na skutek zetknięcia się z taką pluralną formą językową, jest jego gotowość do gromadzenia dalszych informacji o dwóch ujadających psach, w tym gromadzenia informacji o każdym z nich osobna. Podobnie, gdy podmiot zetknie się z w fikcji *PW* z twierdzeniem „Dwa elektrony krążą wokół siebie po orbitach o takim samym kształcie”, również tworzy dwa pojęcia jednostkowe tych elektronów, a dodatkową racją, dla której powinno się zakładać, iż utworzy on tutaj dwa pojęcia jednostkowe, jest jego gotowość do dalszego gromadzenia informacji o tych elektronach, w tym informacji o każdym z nich z osobna. Wystarczy, że podmiot skupi swoją uwagę na arbitralnie wybranym elektronie.

Dzięki odwołaniu się do społecznej, a nie indywidualnej praktyki językowej kryterium (IPJ<sub>4</sub>) pozwala również na odróżnienie pojęć jednostkowych, które podmiot mylnie identyfikuje ze sobą. Choć

sam podmiot może być w błędzie, to w ramach praktyki społecznej okoliczności, w których nabył te pojęcia, stanowią podstawę do tego, by je odróżnić.

Zgodnie z powyższymi sugestiami, gdy w ramach fikcji *PW* podmiot *S* styka się ze wzmianką o świecie dwóch krążących wokół siebie elektronów, wystąpienie formy pluralnej prowadzi go do utworzenia dwóch pojęć jednostkowych, które mają na celu odnoszenie się do tych dwóch odmiennych przedmiotów nieaktualnych. Naturalne pytanie, idące za taką sugestią, brzmi: „Które z utworzonych pojęć jednostkowych odnosi się do którego elektronu?”. Jedyna nasuwająca się odpowiedź jest taka, że jest to kwestia arbitralnego wyboru. Z teoretycznego punktu widzenia istotne jest jedynie to, by pewną sytuację reprezentować jako zawierającą dwa określone przedmioty, a nie jeden. W przypadku, gdy nie pojawiają się żadne dalsze informacje deskryptywne, które pozwoliłyby odróżnić wspomniane przedmioty, nie można zrobić niczego więcej ponad arbitralny wybór, które pojęcie ma na celu odnoszenie się do którego przedmiotu.

Odwołanie się do pojęć jednostkowych rozwiązuje jedynie część problemów związanych z drobnoziarnistością treści możliwościowej fikcji. Strategia analizowania treści twierdzeń formułowanych w ramach fikcji *PW* poprzez odwołanie się do pojęć jednostkowych cierpi bowiem na poważne ograniczenie — wymaga przyjęcia założenia, że podmiot tworzy pojęcia jednostkowe, które mają na celu odnoszenie się do każdego z przedmiotów stanowiących elementy danej możliwej sytuacji. Można jednak wskazać światy możliwe, które mają jako swoje części tak wiele nieodróżnialnych indywiduów, że hipoteza o tym, że podmiot *S* byłby w stanie utworzyć tyle pojęć jednostkowych, ile indywiduów jest w tym świecie, jest pozbawiona jakiegokolwiek wiarygodności.

Niech *w* będzie światem możliwym, w którym istnieje nieskończenie wiele elektronów umiejscowionych w przestrzeni w taki sposób, że

każdy elektron znajduje się w takiej samej odległości od najbliższego mu elektronu, a wszystkie one są rozmieszczone w taki sposób, jak gdyby były punktami na obwodzie okręgu. Załóżmy dodatkowo, że świat *w* jest światem fizycznie niemożliwym i pomiędzy elektronami zachodzą tylko takie oddziaływania, że każdy elektron oddziałuje na każdy inny elektron w tym świecie w taki sam sposób. Zakładając dodatkowo, że każdy z tych elektronów posiada takie same własności wewnętrzne, jedyną postawą do odróżnienia tych elektronów pomiędzy sobą byłoby wówczas wskazanie jakichś własności zewnętrznych, których nie można przypisać każdemu z nich. Z uwagi jednak na to, że elektrony te pozostają do siebie w tych samych relacjach, każdemu z nich można przyporządkować dokładnie takie same własności zewnętrzne. Zmianie ulegają tutaj nie relacje, w których pozostają do siebie elektrony w świecie *w*, a jedynie to, które elektrony w tym świecie pozostają do których elektronów w tych relacjach. Właśnie z uwagi na to, że elektrony te pozostają do siebie w tych samych relacjach i posiadają te same własności wewnętrzne, są one jakościowo nieodróżnialne i nie mogą być identyfikowane ze względu na to, że każdy z nich spełnia jakąś deskrypcję, której nie spełnia żaden inny. Przedmioty te można odróżnić jedynie wówczas, gdy przyjmie się, że można odróżnić je na jakiejś innej podstawie niż jakościowa charakterystyka.

Jeśli chcemy móc mówić o takich przedmiotach, musimy przyjąć, że można próbować odnosić się do nich bez pośrednictwa deskrypcji określonych lub nazw własnych. Korzystanie z pierwszych jest niemożliwe, ponieważ nie jest prawdą, że każdy z tych przedmiotów spełnia jakąś unikalną dla niego deskrypcję. Korzystanie z nazw własnych jest niemożliwe, ponieważ żaden skończony podmiot nie może przyporządkować nazw własnych nieskończonej liczbie przedmiotów. Podobnie podmiot *S*, słysząc o nieskończonej ilości elektronów w świecie *w*, nie jest w stanie utworzyć nieskończonej liczby pojęć jednostkowych, z których każde miałyby na celu odnoszenie się do jednego z elektronów

w świecie  $w$ . Musimy zatem rozejrzeć się za innymi sposobami reprezentowania przedmiotów istniejących w takich światach możliwych jak świat  $w$ .

Jednym ze środków językowych, który pozwala na mówienie o określonych wielościach obiektów, przypisując wszystkim przedmiotom należącym do tej wielości pewne własności, są pluralne deskrypcje określone. Pluralne deskrypcje określone są formami językowymi o następującej postaci gramatycznej: „Te oto  $F$ -y...”, których analiza logiczna przedstawia się następująco: „Pewne rzeczy  $\theta$  są takie, że  $F\theta$  i dla każdego  $Y$ , jeśli  $FY$ , to  $Y$  znajdują się pośród  $\theta$ , i...” (por. Sainsbury 2005, s. 193). Pluralne deskrypcje określone mają tę niewątpliwą zaletę, że pozwalają na wyrażenie pewnych informacji o pewnych wielościach przedmiotów bez odnoszenia się do każdego poszczególnego przedmiotu znajdującego się w tej wielości. Jeśli przyjąć wprowadzone wyżej rozróżnienie na atrybutywne i referencjalne użycia deskrypcji określonych, można przenieść to rozróżnienie na pluralne deskrypcje określone. Następnie referencjalne użycia pluralnych deskrypcji określonych można traktować jako wyrazy posiadania reprezentacji mentalnych, które mają na celu odnoszenie się do określonych wielości przedmiotów. Odwołując się do reprezentacji mentalnych tego typu, można próbować nadać treść tym twierdzeniom fikcji  $PW$ , w których mówi się o wszystkich indywiduach istniejących w takich światach jak świat  $w$ . Założenie, że przy opisie takich światów posługujemy się pluralnymi deskrypcjami określonymi, jest wspierane przez przyjmowanie w ramach fikcji  $PW$  tezy o nieistnieniu transświatowych indywiduów oraz nieistnieniu jakościowo nieodróżnialnych światów. W związku z tym, zgodnie z  $PW$ , istnieje tylko jedna taka wielkość elektronów.

Zgodnie z hipotezą, którą chcę tutaj zaproponować, obok pojęć jednostkowych użytkownicy języka są również w stanie tworzyć reprezentacje określonych wielości przedmiotów; reprezentacje te będę nazywał *pojęciami pluralnymi*. Pod wieloma względami pojęcia plu-

ralne przypominają pojęcia jednostkowe. Po pierwsze, ich rolą w życiu kognitywnym podmiotu *S* jest umożliwienie mu gromadzenia oraz rewidowania informacji o określonych wielościach przedmiotów. Po drugie, nie zawierają one zakamuflowanych treści deskryptywnych. Podobnie jak w przypadku pojęć jednostkowych fakt, że pewna wielość nie spełnia informacji, które podmiot mylnie przypisuje tej wielości, nie wyklucza posiadania pojęcia pluralnego, które ma na celu odniesienie się do tej wielości przedmiotów. Po trzecie, zasady odnoszenia się pojęć pluralnych do wielości przedmiotów są analogiczne do zasad odnoszenia się pojęć jednostkowych do konkretnych przedmiotów. W analogii do zasad (a)–(c) rządzących odniesieniem pojęć jednostkowych można sformułować następujące zasady rządzące odniesieniem pojęć pluralnych:

- a) *Odmienne pojęcia pluralne mogą mieć to samo odniesienie.* Źródłem tej zasady jest to, że podmiot może nabyć w odmiennych okolicznościach odmienne pojęcia pluralne, których odniesieniem jest ta sama wielość przedmiotów. O dwóch wielościach przedmiotów powiemy, że są identyczne wtw, gdy dla każdego przedmiotu *x*, jeśli jest jednym z przedmiotów składających się na pierwszą wielość przedmiotów, to jest jednym z przedmiotów składających się na drugą wielość przedmiotów i odwrotnie. Konsekwencją przyjęcia tej zasady jest to, że pojęcia pluralne nie są indywidualizowane wyłącznie ze względu na to, do jakiej wielości obiektów się odnoszą, tj. wystarczającym warunkiem identyczności pojęć pluralnych nie jest posiadanie tego samego odniesienia.
- b) *Odniesienia pojęć pluralnych nie ulegają zmianom.* Uzasadnienie tego założenia jest analogiczne do uzasadnienia pojęć jednostkowych: przyjęcie założenia odwrotnego sprawiałoby, że trudno byłoby wyjaśnić, w jaki sposób pojęcia pluralne mogą pełnić

rolę stabilnych punktów, wokół których podmiot gromadzi informacje o określonych wielościach przedmiotów.

- c) *Pojęcie pluralne ma co najwyżej jedno odniesienie.* Zasadniczą rolą pojęć pluralnych jest umożliwienie podmiotowi  $S$  myślenia o określonej wielości przedmiotów i zbieranie informacji o nich. Pełnienie tej roli byłoby niemożliwe, gdyby pojęcie pluralne mogło odnosić się do więcej niż jednej wielości przedmiotów. W związku z tym przyjmuje się założenie, iż pojęcia pluralne odnoszą się co najwyżej do jednej wielości przedmiotów. Klauzula „co najwyżej” pojawia się tutaj dlatego, że próba sformułowania koncepcji treści fikcji  $PW$  wymaga, by dopuszczalne były pojęcia pluralne, które nie posiadają odniesień.

Po czwarte, kryteria identyczności pojęć pluralnych są analogiczne do kryteriów identyczności pojęć jednostkowych:

- (IPP) Pojęcia pluralne  $P$  i  $P^*$  są identyczne wtw, gdy  $P$  i  $P^*$  pełnią te same role w życiu kognitywnym podmiotu  $S$ ,  $P$  i  $P^*$  zostały nabyte w tej samej chwili oraz  $P$  i  $P^*$  zostały nabyte w okoliczności, która w ramach społecznej praktyki językowej jest podstawą do utożsamienia wielości przedmiotów, do których pojęcia te mają na celu się odnosić.

Pojęcia pluralne  $P$  i  $P^*$  pełnią te same role w życiu kognitywnym podmiotu  $S$  wtw, gdy podmiot  $S$  jest przekonany, że dla dowolnej informacji  $i$ ,  $i$  jest podciągane przez niego pod pojęcie pluralne  $P$  wtw, gdy  $i$  jest podciągane przez niego pod pojęcie pluralne  $P^*$  oraz  $i^*$  jest wyłączane przez niego spod pojęcia  $P$  wtw, gdy  $i^*$  jest wyłączane przez niego spod pojęcia  $P^*$ .

Jeśli zgodzić się na hipotezę istnienia pojęć jednostkowych oraz pojęć pluralnych i jakąś koncepcję mentalnych reprezentacji własności oraz relacji, to będzie się posiadało środki do tego, by analizować

za ich pomocą treści jednostkowych wyrażeń referencjalnych, treści pluralnych wyrażeń referencjalnych oraz zdań utworzonych za pomocą  $n$ -argumentowego predykatu poprzedzającego ciąg  $n$  jednostkowych wyrażeń referencjalnych i zdań utworzonych za pomocą  $n$ -argumentowego pluralnego predykatu poprzedzającego ciąg  $n$  pluralnych wyrażeń referencjalnych. Przykładowo, treść takich zdań jak zdanie „ $Ft_1, \dots, t_n$ ”, gdzie  $t_1, \dots, t_n$  są jednostkowymi wyrażeniami referencjalnymi, a predykat „ $F$ ” jest  $n$ -argumentowym predykatem, można reprezentować za pomocą  $n$ -tki uporządkowanej o postaci:  $\langle \mathbf{F}, \langle I_1, \dots, I_n \rangle \rangle$ , gdzie  $I_1, \dots, I_n$  są pojęciami jednostkowymi, a  $\mathbf{F}$  mentalną reprezentacją własności bądź relacji, którą predykat „ $F$ ” ma na celu wyrażać. Treść zdania „ $Gu_1, \dots, u_n$ ”, gdzie  $u_1, \dots, u_n$  są pluralnymi wyrażeniami referencjalnymi, a predykat „ $G$ ” jest  $n$ -argumentowym predykatem pluralnym, można reprezentować za pomocą  $n$ -tki uporządkowanej o postaci:  $\langle \mathbf{G}, \langle P_1, \dots, P_n \rangle \rangle$ , gdzie  $P_1, \dots, P_n$  są pojęciami pluralnymi, a  $\mathbf{G}$  mentalną reprezentacją własności bądź relacji, którą predykat „ $G$ ” ma na celu wyrażać.

## 4.2 Treść ogólnych twierdzeń możliwościowych

Większość twierdzeń składających się na fikcję  $PW$  to twierdzenia ogólne, które mówią o dowolnym przedmiocie z pewnej dziedziny, który spełnia treść deskryptywną tych twierdzeń. W odróżnieniu od twierdzeń, w których występują jednostkowe wyrażenia referencjalne, w twierdzeniach ogólnych takich jak twierdzenia uniwersalne o postaci „Dla każdego  $x$ ,  $Fx$ ” bądź twierdzenia egzystencjalne o postaci „Istnieje takie  $x$ , że  $Fx$ ” występują zmienne indywiduowe, które nie mają na celu odnoszenia się do określonego indywiduum znajdującego się w dziedzinie kwantyfikacji. Standardowo warunki prawdziwości twierdzeń uniwersalnych definiuje się kategoriach spełniania formuł, powstałych przez opuszczenie kwantyfikatora ogólnego, przez dowolny ciąg przedmiotów (znajdujących się w dziedzinie kwantyfikacji) taki,



że  $n$ -ty element tego ciągu jest przyporządkowany jako wartość  $n$ -tej zmiennej występującej w tej formule. Z kolei warunki prawdziwości twierdzeń egzystencjalnych definiuje się w kategoriach spełnienia formuł, powstałych przez opuszczenie kwantyfikatora egzystencjalnego, przez pewien ciąg przedmiotów (znajdujących się w dziedzinie kwantyfikacji) taki, że  $n$ -ty element tego ciągu jest przyporządkowany jako wartość  $n$ -tej zmiennej występującej w tej formule. Używając twierdzeń uniwersalnych bądź egzystencjalnych, podmiot  $S$  zazwyczaj nie dysponuje jednostkowymi reprezentacjami obiektów, po których kwantyfikuje w tych twierdzeniach. Zazwyczaj nie dysponuje również pojęciami pluralnymi, które miałyby za swoje odniesienia wielości, na które składają się przedmioty znajdujące się w dziedzinie kwantyfikacji. W jaki sposób powinniśmy reprezentować treści twierdzeń ogólnych, występujących w fikcji  $PW$ ?

W ramach jednej z koncepcji (zob. Glanzberg 2006, § 31.1) kwantyfikatorom można przyporządkować własności drugiego rzędu jako ich wartości semantyczne. Niech  $D$  będzie dziedziną kwantyfikacji, do której należą wyłącznie indywidua, wówczas — upraszczając — wartością semantyczną kwantyfikatora ogólnego jest własność bycia taką  $n$ -argumentową własnością, że podpada pod nią każda  $n$ -tka uporządkowana indywiduów należących do  $D$ . Z kolei wartością semantyczną kwantyfikatora szczegółowego jest własność bycia taką  $n$ -argumentową własnością, że podpada pod nią co najmniej jedna  $n$ -tka uporządkowana indywiduów należących do  $D$ . O twierdzeniu „Dla każdego  $x$ ,  $Fx$ ” powiemy, że jest prawdziwe wtw, gdy własności  $F$  przysługuje własność bycia taką własnością, że podpada pod nią każde indywiduum należące do  $D$ . Natomiast o twierdzeniu „Istnieje takie  $x$ ,  $Fx$ ” powiemy, że jest prawdziwe wtw, gdy własności  $F$  przysługuje własność bycia taką własnością, że podpada pod nią co najmniej jedno indywiduum należące do  $D$ . Pozostaje tylko pytanie, czy podmiot  $S$  posiada reprezentacje mentalne takich własności drugiego rzędu.

Gdyby przyjąć, iż podmiot posiada mentalne reprezentacje takich własności, można by próbować w następujący sposób scharakteryzować treść twierdzeń ogólnych, które występują w fikcji *PW*: niech  $U_D$  będzie własnością bycia taką  $n$ -argumentową własnością, że podpada pod nią wszystkie  $n$ -tki indywiduów należące do dziedziny  $D$ ,  $E_D$  będzie własnością bycia taką  $n$ -argumentową własnością, że podpada pod nią co najmniej jedna  $n$ -tka indywiduów należący do dziedziny  $D$ ,  $N_D$  będzie własnością bycia  $n$ -argumentową własnością taką, że nie podpada pod nią żadna  $n$ -tka indywiduów należących do dziedziny  $D$ . Niech  $\mathbf{U}_D$  będzie reprezentacją mentalną własności  $U_D$ ,  $\mathbf{E}_D$  będzie reprezentacją mentalną własności  $E_D$ ,  $\mathbf{N}_D$  będzie reprezentacją mentalną własności  $N_D$ . Wówczas treść twierdzenia uniwersalnego „Dla każdego  $x$ ,  $Fx$ ” może być reprezentowana jako  $\langle \mathbf{U}_D, \mathbf{F} \rangle$ , treść zdania egzystencjalnego „Istnieje takie  $x$ , że  $Fx$ ” może być reprezentowana za pomocą  $\langle \mathbf{E}_D, \mathbf{F} \rangle$  i treść negatywnego zdania egzystencjalnego może być reprezentowana za pomocą  $\langle \mathbf{N}_D, \mathbf{F} \rangle$ .

Odwołania do indywiduów należących do dziedziny  $D$  sprawiają, że własności drugiego rzędu  $U_D$ ,  $E_D$ ,  $N_D$  są własnościami zewnętrznymi, które przysługują  $n$ -argumentowym własnościom niższego rzędu na mocy tego, iż pozostają one w relacji  $R$  (*bycia posiadanym przez*) do  $n$ -tek przedmiotów należących do dziedziny  $D$ . W związku z tym reprezentacja mentalna własności  $U_D$ ,  $E_D$ ,  $N_D$  powinna obejmować reprezentację mentalną relacji  $R$  oraz  $n$ -tek indywiduów należących do dziedziny  $D$ . Widzieliśmy jednak, że w ogólności nie można reprezentować wielości indywiduów należących do dziedziny kwantyfikacji ani za pomocą szeregu pojęć jednostkowych mających na celu odnoszenie się do każdego z nich z osobna, ani za pomocą pojęcia pluralnego, które miałoby na celu odnoszenie się do określonej wielości przedmiotów. W obu przypadkach musielibyśmy bowiem założyć, że podmiot  $S$ , dając wyraz swoim myślom przy użyciu twierdzeń ogólnych, ma na celu

odnoszenie się do określonych przedmiotów bądź określonej wielości przedmiotów. Zwykle jednak twierdzeń ogólnych używa się właśnie w celu mówienia o jakichkolwiek przedmiotach, które spełniają pewną treść deskrypcyjną wyrażoną przez te twierdzenia.

Gdyby zatem istniały mentalne reprezentacje własności  $U_D$ ,  $E_D$ ,  $N_D$ , byłyby one ustrukturyzowanymi reprezentacjami mentalnymi, których elementami konstytutywnymi powinna być mentalna reprezentacja relacji  $R$  oraz — chciałoby się rzec — reprezentacje nieokreślonych wielości przedmiotów i reprezentacje nieokreślonych pojedynczych przedmiotów. Czy coś może przemawiać za tym, że użytkownicy języka są w stanie utworzyć reprezentacje nieokreślonych przedmiotów bądź nieokreślonych wielości przedmiotów?

Jednym z fenomenów, które mogłyby wspierać zasadność wprowadzenia reprezentacji takiego typu, jest umiejętność konstruowania przedmiotów myśli na podstawie pewnego zbioru informacji. W szczególności chodzi mi o myślenie o niesprecyzowanych przedmiotach, które wywoływane jest przez takie sposoby mówienia jak „Niech pewien przedmiot będzie  $F$ ”, „Niech  $x$  będzie  $G$ ” itd. W analogii do koncepcji pojęć jednostkowych można próbować opisać zdolność do konstruowania takiego przedmiotu myśli jako podciąganie informacji pod *pojęcie nieokreślonego przedmiotu*. Kognitywna funkcja pojęć nieokreślonych przedmiotów odpowiada funkcji pełnionej przez pojęcia jednostkowe, mianowicie zarówno jedne, jak i drugie pełnią rolę punktów, pod które podciąga się pewne informacje. Pojęcia nieokreślonych przedmiotów odróżniają się od pojęć jednostkowych tym, że nie mają na celu odnoszenia się do określonego przedmiotu, ale do dowolnego przedmiotu, który spełnia treść deskryptywną. W związku z tym pojęcia nieokreślonych przedmiotów z zasady nie stanowią stabilnych i rozciągłych w czasie punktów, wokół których gromadzi się informacje o określonym przedmiocie. Nawet gdy wraca się do rozważań o własnościach wyrażanych przez jednoargumentowy predykat „ $F$ ”, nie

mówi się „Wróćmy do przedmiotu, który rozważaliśmy wczoraj” — raczej mówi się o tym jedynie, by myśleć o przedmiocie, który spełnia tę samą deskryptywną charakterystykę, o której wczoraj była mowa.

Jeśli rozważania te uznać za motywację do wprowadzenia pojęć nieokreślonych przedmiotów, to kolejnym zagadnieniem, przed którym staje zwolennik ich wprowadzenia, jest sformułowanie kryterium identyczności dla tych pojęć. Nie mają one na celu odnoszenia się do określonych przedmiotów, a tylko do przedmiotów spełniających informacje deskryptywne podciągane pod te pojęcia. Dlatego nie ma obaw, że utożsamiając pojęcia nieokreślonych przedmiotów pełniące tę samą rolę w życiu kognitywnym podmiotu, utożsamimy dwa intuicyjnie odmienne pojęcia. W związku z tym kryterium ich identyczności może przedstawiać się następująco:

- (IPO) Pojęcia nieokreślonych przedmiotów  $O$  i  $O^*$  są identyczne wtw, gdy dla dowolnej chwili  $t$  i dowolnej informacji  $i$ ,  $i$  jest podciągane przez podmiot  $S$  pod pojęcie nieokreślonego przedmiotu  $O$  w chwili  $t$  wtw, gdy  $i$  jest podciągane przez podmiot  $S$  pod pojęcie nieokreślonego przedmiotu  $O^*$  w chwili  $t$  oraz dla dowolnej chwili  $t^*$  i dowolnej informacji  $i^*$ ,  $i^*$  jest wyłączane przez niego spod pojęcia  $O$  w chwili  $t^*$  wtw, gdy  $i^*$  jest wyłączane przez niego spod pojęcia  $O^*$  w chwili  $t^*$ .

Opierając się na podobnych rozważaniach, można *mutatis mutandis* motywować wprowadzenie pojęć nieokreślonych wielości przedmiotów, dla których można sformułować analogiczne kryteria identyczności, z zastrzeżeniem, że informacje podciągane pod te pojęcia mają pluralną treść w tym sensie, że mają one zastosowanie wyłącznie do wielości przedmiotów. W zależności od tego, czy mamy do czynienia z predykatami o pluralnej treści dystrybutywnej, czy z predykatami o pluralnej treści kolektywnej, podciągane informacje albo będą dotyczyły każdego z przedmiotów należących do tej wielości z osobna, albo

będą dotyczyły wszystkich przedmiotów należących do tej wielości razem wziętych. Analogicznie do kryterium (IPO) będzie przedstawiało się również kryterium identyczności pojęć nieokreślonych wielości przedmiotów.

Dysponując zarówno pojęciami nieokreślonych przedmiotów, jak i nieokreślonych wielości przedmiotów, można próbować w następujący sposób charakteryzować treść twierdzeń ogólnych. Treść twierdzenia egzystencjalnego „Istnieje takie  $x$ , że  $Fx$ ” może być reprezentowana za pomocą istnienia takiego pojęcia nieokreślonego przedmiotu  $O$  i takiego pojęcia nieokreślonej wielości przedmiotów  $V_D$  (należących do dziedziny kwantyfikacji  $D$ ), że  $\mathbf{R} < O, \mathbf{F} >$  i  $\mathbf{S} < O, V_D >$ , gdzie  $\mathbf{R}$  jest mentalną reprezentacją relacji drugiego rzędu zachodzącą pomiędzy przedmiotem a własnością wyrażoną przez predykat „ $F$ ”, a  $\mathbf{S}$  jest mentalną reprezentacją relacji *bycia jednym z*. Z kolei treść twierdzenia uniwersalnego „Dla każdego  $x$ ,  $Fx$ ” byłaby reprezentowana następująco: istnieje takie pojęcie nieokreślonej wielości przedmiotów  $V_D$  (należących do dziedziny kwantyfikacji  $D$ ), że dla dowolnego pojęcia nieokreślonego przedmiotu  $O$ , jeśli  $\mathbf{S} < O, V_D >$ , to  $\mathbf{R} < O, \mathbf{F} >$ .

Dla przykładu, treść twierdzenia „Istnieje świat  $w$ , w którym istnieją jakościowo nieodróżnialne przedmioty” może być reprezentowana w następujący sposób: istnieje takie  $O_1$  i istnieje takie  $V_{d(w)}$  i  $V_D$ , że  $\mathbf{S} < O_1, V_D >$  i  $\mathbf{R} < O_1, \mathbf{W} >$ , i dla każdego pojęcia nieokreślonego przedmiotu  $O_2$  i  $O_3$ , jeśli  $\mathbf{S} < O_2, V_{d(w)} >$ , i  $\mathbf{S} < O_3, V_{d(w)} >$ , i  $\mathbf{R} < < O_3, O_1 >, \mathbf{I} >$ , to  $\mathbf{R} < O_2, \mathbf{F} >$  wtw, gdy  $\mathbf{R} < O_3, \mathbf{F} >$ .  $\mathbf{W}$  jest tutaj mentalną reprezentacją własności bycia światem możliwym,  $V_{d(w)}$  jest mentalną reprezentacją wielości przedmiotów istniejących w świecie  $w$ ,  $O_1$  jest reprezentacją nieokreślonego świata możliwego,  $\mathbf{F}$  jest mentalną reprezentacją własności przypisywanych indywiduom, indywidua reprezentowane są za pomocą pojęć nieokreślonych przedmiotów  $O_2$  i  $O_3$ ,  $\mathbf{I}$  jest mentalną reprezentacją relacji *bycia różnym od*.

Zaproponowane wyżej rozróżnienia na pojęcia jednostkowe, pluralne, nieokreślonych przedmiotów oraz nieokreślonych wielości przedmiotów wymagają rzecz jasna dalszej analizy, na którą niestety nie ma tutaj miejsca. Celem niniejszego rozdziału było jedynie wskazanie kierunku, w którym mógłby podążać fikcjonalista modalny w próbie rekonstrukcji treści fikcji *PW*.

### 4.3 Co z ontologią reprezentacji mentalnych?

Przedstawione wyżej zręby koncepcji treści fikcji *PW* wciąż mogą budzić liczne obawy ontologów. Ograniczę się do rozważania kwestii tego, czy posiadanie reprezentacji mentalnej musi być efektem interakcji podmiotu z pewnymi niezależnie istniejącymi przedmiotami. Gdyby tak było, powstawałby problem, czy aby tylnymi drzwiami nie wprowadza się bytów niepożądanych dla fikcjonalisty modalnego. Dla przykładu, posiadanie pojęcia psa może polegać na uchwyceniu Fregowskiego sensu psa, na mentalnym skierowaniu myśli ku Meinongowskiemu przedmiotowi posiadającemu własność psa czy wreszcie na skierowaniu myśli na zbiór Lewisowskich możliwych psów. Możliwości interpretacyjnych jest wiele, ale żadnej z nich nie jest po drodze z projektem fikcjonalisty modalnego.

W próbie uporania się z tą trudnością warto wyjść od odróżnienia dwóch problemów. Jednym jest filozoficzna interpretacja pojęć ogólnych takich jak pojęcia psa. Innym problemem jest np. filozoficzna interpretacja pojęć jednostkowych, mających stanowić reprezentację mentalną określonego przedmiotu. Traktowanie tych kwestii jako oddzielnych jest o tyle zasadne, że zwolennik najbardziej obiecującej, umiarkowanej wersji fikcjonalizmu modalnego może proponować odmienne koncepcje nabywania dla każdego z tych dwóch typów pojęć. Umiarkowany fikcjonalista modalny przyjmuje bowiem istnienie takich bytów jak platońskie własności czy relacje. Gdy ma się je w ontologii, można również założyć, że podmiot może pozostawać do nich w

pewnej relacji intencjonalnej (zwykle nazywanej *uchwytywaniem*). Na gruncie takich założeń wiarygodne wydaje się przyjęcie, że podmiot *S* może posiadać pojęcie psa wtedy i tylko wtedy, gdy posiada mentalną reprezentację własności bycia psem wywołaną uchwyceniem tej własności. Analogicznie może przedstawiać się koncepcja pozostałych pojęć ogólnych.

Sytuacja komplikuje się, gdy zaczniemy rozważać pojęcia jednostkowe. Są one reprezentacjami mentalnymi określonych przedmiotów, z którymi podmiot koreluje reprezentacje mentalne własności przysługujących tym przedmiotom. Z uwagi na to, że pojęcia jednostkowe pełnią istotną rolę w konstruowaniu reprezentacji mentalnych dla jakościowo nieodróżnialnych przedmiotów, nie powinny być utożsamiane z reprezentacjami mentalnymi powstałymi na skutek uchwycenia pewnej wiązki własności. Co najwyżej można byłoby je traktować jako reprezentacje powstałe na skutek skierowania myśli na Lewisowskie indywiduum. Taka strategia jest nie tylko nieakceptowalna dla fikcjonalisty modalnego, ale również nie może stanowić ogólnej koncepcji pojęć jednostkowych. Mają one być stabilnymi punktami, z którymi podmiot może korelować informacje zdobywane w różnych chwilach o danym przedmiocie. Zatem zawsze gdy podmiot będzie wyobrażał sobie, że pewnemu określonemu przedmiotowi przysługują pewne własności lub będzie wracał do wyobrażeń o tym przedmiocie i wzbogacał swoje wyobrażenia, przypisując temu przedmiotowi kolejne własności, będzie w posiadaniu pojęcia jednostkowego mającego na celu odnośnienie się do tego przedmiotu. Zdarza się jednak, że podmiot celowo bądź przez własną ignorancję wyobraża sobie, że przedmiot posiada wykluczające się własności. Możliwość popadania w modalne iluzje bądź tworzenia wewnętrznie sprzecznych fikcji wydaje się świadczyć o tym, że posiadanie pojęcia jednostkowego jest niezależne od tego, czy istnieje Lewisowskie możliwe indywiduum, do którego podmiot miałby pozostawać w pewnej relacji intencjonalnej. Jest tak dlatego,

że pojęciom jednostkowym, pod które podciąga się wykluczające się reprezentacje ogólne, nie odpowiada żadne Lewisowskie indywiduum. To z kolei wystarczy fikcjonalistom modalnemu do tego, by twierdzić, iż posiadanie jednostkowej reprezentacji mentalnej w ogólności nie polega na uchwytowaniu jakiegoś bytu.

Jedyną znaną mi koncepcją, która posiadanie dowolnego pojęcia jednostkowego uzależniałaby od pozostawania w pewnej intencjonalnej relacji do określonego przedmiotu, jest noneizm Grahama Priesta (zob. 2005). Jednak ceną za wprowadzenie takiej zależności pomiędzy posiadaniem pojęć jednostkowych a istnieniem odpowiedniego przedmiotu jest konieczność akceptowania istnienia przedmiotów posiadających wykluczające się własności. W ontologii Priesta przedmioty takie zaludniają światy niemożliwe. W ramach jego stanowiska przyjmuje się również istnienie tzw. świata trywialnego, w którym dowolna sprzeczność jest prawdziwa, włącznie z tezą, że samo to stanowisko jest zarazem fałszywe i prawdziwe. Wydaje się zatem, że mamy do czynienia ze stanowiskiem niekoherentnym w tym sensie, że w jego ramach jest ono zarówno prawdziwe, jak i fałszywe. Jeśli zatem założenie, że każdemu pojęciu jednostkowemu odpowiada niezależnie istniejący byt, prowadzi do sprzeczności, to mamy dobre racje na rzecz tezy, że zależność taka w ogólności nie zachodzi.

## 5 Inne problemy z treścią fikcji

Na początku tego rozdziału wspomniałem o dwóch wersjach fikcjonalizmu modalnego: silnej oraz umiarkowanej. W tym paragrafie podejmę się ich zestawienia, i oceny przez pryzmat tego, jak radzą sobie z *problemem kompletności* treści fikcji *PW* oraz z *problemem sztuczności* treści tej fikcji. Nim przejdę do rozważań poświęconych tym problemom, przypomnę, czym różnią się od siebie obie wersje fikcjonalizmu modalnego.



Zgodnie z jego *silną* odmianą fikcjonalizm modalny ma pełnić dwie funkcje. Z jednej strony ma umożliwić mówienie o światach możliwych bez przyjmowania ich istnienia. Z drugiej zaś strony ma dostarczyć redukcyjnej analizy pojęć modalnych (zob. Rosen 1990, s. 353–354; Nolan 1997, s. 263). W literaturze tak ambitna wersja fikcjonalizmu modalnego pojawia się jedynie w jego metafikcjonalistycznej odmianie. W związku z tym pierwsze z zadań będzie wypełniane poprzez analizę warunków prawdziwości dla twierdzeń możliwościwiatowych „ $p^*$ ” w zgodzie z wprowadzonym wyżej schematem:

(MMF<sub>4</sub>) „ $p^*$ ” jest prawdziwe wtw, gdy zgodnie z  $PW$ ,  $p^*$ .

Z kolei redukcyjna analiza twierdzeń modalnych „ $p$ ” będzie się dokonywała przez wyznaczenie warunków prawdziwości tych twierdzeń w zgodzie z wprowadzonym wyżej schematem:

(MMF<sub>3</sub>) „ $p$ ” jest prawdziwe wtw, gdy zgodnie z  $PW$ ,  $p^*$ .

Przedmiotowy fikcjonalizm modalny nie występuje w silnej wersji, ponieważ jego zwolennicy przyjmują, że reguły generowania uzależniają fikcyjną prawdziwość od obiektywnie zachodzących faktów modalnych, na których zachodzeniu ufundowana jest prawdziwość twierdzeń modalnych. Tymczasem silny fikcjonalista modalny postępuje dokładnie odwrotnie, uzależniając prawdziwość twierdzeń modalnych od tego, co jest prawdziwe zgodnie z fikcją  $PW$ .

Zgodnie z *umiarkowaną* wersją fikcjonalizmu modalnego zasadniczym celem tego stanowiska jest zachowanie możliwości mówienia o światach możliwych bez przyjmowania ich istnienia, ale nie jest już nim dostarczenie analizy warunków prawdziwości twierdzeń modalnych. Mimo to umiarkowany fikcjonalista modalny dopuszcza przechodzenie od prawd modalnych do twierdzeń możliwościwiatowych. Jest to możliwe dzięki temu, że w ramach tego stanowiska fikcyjna prawdziwość niektórych twierdzeń możliwościwiatowych jest zależna od tego, jakie

zachodzą fakty modalne. Fakty modalne fundują również prawdziwość twierdzeń modalnych. Wszędzie tam, gdzie zachodzenie faktu modalnego funduje wyłącznie prawdziwość twierdzenia modalnego  $p$ , a fikcyjna prawdziwość twierdzenia możliwościowego  $p^*$  jest zależna tylko od zachodzenia tego faktu modalnego, tam uzasadnione jest przejście od  $p$  do  $p^*$  i odwrotnie. Słowem, „twierdzenia dotyczące światów możliwych (albo, by być bardziej precyzyjnym, tego, jak światy możliwe są opisywane w fikcji) opierają się raczej na faktach dotyczących modalności niż odwrotnie” (Nolan 1997, s. 262).

Umiarkowany fikcjonalizm modalny występuje w dwóch formach:

- (UMFM) Umiarkowany metafikcjonalizm modalny, zgodnie z którym: (1) warunki prawdziwości twierdzeń możliwościowych wyznaczone są przez schemat (MMF<sub>4</sub>), (2) istnieją fakty modalne, na których ufundowana jest prawdziwość twierdzeń modalnych i (3) prawdziwość niektórych twierdzeń z funktorem opowieści „Zgodnie z  $PW$ ” jest zależna (*via* prawa pomostowe) od tego, jakie są fakty modalne.
- (UPFM) Umiarkowany przedmiotowy fikcjonalizm modalny, zgodnie z którym: (1) twierdzenia możliwościowe są prawdziwe w ramach gry  $GPW$ , (2) istnieją fakty modalne, na których ufundowana jest prawdziwość twierdzeń modalnych i (3) fikcyjna prawdziwość niektórych twierdzeń możliwościowych jest zależna (*via* reguły generowania) od tego, jakie są fakty modalne.

## 5.1 Problem niekompletności treści fikcji

Jeśli fikcja  $PW$  ma być dziełem człowieka, siłą rzeczy jej wyrażona *explicite* treść musi być ograniczona w tym sensie, że fikcjonalista modalny życzyłby sobie, aby więcej sądów było prawdziwymi zgodnie z fikcją  $PW$  niż jest w niej *explicite* wyrażone. Daniel Nolan (1997, s.

266–268) argumentuje, że ograniczenie pod względem *explicite* wyrażonej treści fikcji *PW* jest źródłem problemów dla zwolenników silnej wersji fikcjonalizmu modalnego. Jeśli bowiem przyjmiemy założenie, że

(a) istnieje nieskończenie wiele możliwości,

to — twierdzi Nolan — dla zwolennika silnego metafikcjonalizmu modalnego jest to równoznaczne z tym, że

(b) istnieje nieskończenie wiele twierdzeń możliwościowych prawdziwych zgodnie z fikcją *PW*.

*Explicite* wyrażona treść fikcji nie jest jednak nieskończona, dlatego też zwolennik silnego fikcjonalizmu modalnego musi uznać, że *explicite* wyrażona treść pociąga za sobą nieskończenie wiele twierdzeń możliwościowych. Z punktu widzenia silnego fikcjonalizmu modalnego wadą takiego rozwiązania jest to, że wymaga ono uznania relacji wynikania za niepodlegającą analizie w kategoriach prawdziwości zgodnie z fikcją *PW*. W charakterystykę relacji wynikania uwikłane są jednak pojęcia modalne. Dlatego silny fikcjonalista modalny będzie musiał zaakceptować istnienie pojęć modalnych, do których nie stosuje się analiza (MMF<sub>3</sub>). Jak starałem się wskazywać wyżej, samo w sobie nie musi to jeszcze przekreślać aspiracji silnego fikcjonalizmu modalnego, gdyż nie jest jasne, czy jakakolwiek koncepcja światów możliwych jest w stanie obyć się bez przyjmowania pierwotnych pojęć modalnych. Jednak Nolan stawia pytanie, dlaczego w ogóle mielibyśmy przyjmować silny fikcjonalizm modalny, skoro „wygląda na to, że pozostałe pojęcia modalne są definiowalne w kategoriach implikacji”? (Nolan 1997, s. 268). Jako przykład, podaje następującą definicję pojęcia modalnego możliwości: możliwe, że *p* wtw, gdy nieprawda, że *p* implikuje *p*.

Niejasne jest dla mnie, w jaki sposób powinno się odczytywać argumentację Nolana. Jeśli silny fikcjonalizm modalny „stawia sobie za

cel bycie *teorią modalności*, tj. wyjaśnieniem warunków prawdziwości twierdzeń modalnych” (Rosen 1990, s. 354), czy też „aspiruje do bycia deflacyjną<sup>43</sup> teorią twierdzeń modalnych, która redukuje twierdzenia modalne do twierdzeń niemodalnych (w tym wypadku prawdziwych w fikcji)” (Nolan 1997, s. 263), to jeśli z założenia (a) ma wynikać (b), to powinno ono być interpretowane jako teza, że istnieje nieskończenie wiele prawdziwych twierdzeń modalnych. Dopiero wówczas na mocy schematu (MMF<sub>3</sub>) wynikałoby z niego (b). Problem jednak w tym, że jeśli twierdzenia modalne są formułowane przez ludzi, to są one jako twory ludzkie tak samo ograniczone w swej liczbie jak twierdzenia tworzące treść fikcji *PW*. Przy takim odczytaniu teza (a) staje się niewiarygodna, a w związku z tym niewiarygodna staje się również argumentacja Nolana. Co więcej, silny fikcjonalista modalny mógłby utrzymywać, że twierdzenie „Istnieje nieskończenie wiele możliwości” jest prawdziwe wtw, gdy zgodnie z opowieścią *PW* istnieje nieskończenie wiele światów możliwych. Ostatnie twierdzenie wydaje się prawdziwe w ramach fikcji *PW*, a Nolan nie wyjaśnia, dlaczego twierdzenie to miałoby stanowić niezadowalającą analizę twierdzenia, że istnieje nieskończenie wiele możliwości? Jedyne nasuwająca się odpowiedź jest taka, że przyjęcie, iż istnieje nieskończenie wiele możliwości, jest wyrażeniem przekonania, iż rzeczywistość modalna jest taka, że istnieje nieskończenie wiele możliwości, z których każda jest dookreślona pod każdym względem, a silny fikcjonalista modalny nie jest w stanie przeprowadzić *ontologicznej* redukcji tak pojętych możliwości do prawdziwości twierdzeń możliwoświatowych zgodnie z fikcją *PW*. Jeśli taki jest powód, dla którego Nolan uznawał, iż z (a) można wyprowadzić (b), to jego argumentacja przeciw silnemu fikcjonalizmowi modalnemu staje się kolistą. Wymaga bowiem przy-

---

<sup>43</sup>W tym kontekście o interpretacji twierdzeń dyskursu D powiemy, że jest deflacyjna wtw, gdy interpretacja twierdzeń dyskursu D pozwala na uznanie ich za prawdziwe bez przyjmowania zobowiązań ontologicznych co do istnienia bytów, o których wydają się mówić te twierdzenia (zob. Rosen 1990, s. 329–330).

jęcia — wbrew silnemu fikcjonalizmowi modalnemu — założenia, że prawdziwość twierdzeń modalnych polega na czymś innym niż na prawdziwości (zgodnie z fikcją *PW*) twierdzeń możliwościowych. Gdyby bowiem — jak utrzymuje zwolennik silnego fikcjonalizmu modalnego — prawdziwość twierdzeń modalnych polegała na prawdziwości (zgodnie z *PW*) twierdzeń możliwościowych, to nie bardzo wiadomo, co uzasadniałoby interpretację tezy, że istnieje nieskończenie wiele możliwości, jako tezy, że możliwości te są dookreślone pod każdym względem.

Nie jest również jasne, w jaki sposób można byłoby uogólnić propozycję zaprzęgnięcia pojęcia pociągania za sobą do analizy funktorów modalnych. Co prawda, wydaje się, że można zachować wzajemną definiowalność funktorów modalnych *de dicto*: konieczne, że  $p$  wtw, gdy możliwe, że  $p$ . Można bowiem wprowadzić następującą definicję: konieczne, że  $p$  wtw, gdy  $p$  pociąga  $p$ . Nie bardzo wiadomo, w jaki sposób skonstruować podobne definicje dla modalności *de re*. Weźmy twierdzenie modalne *de re*:  $(x)\Box Px$ . Ani twierdzenie „ $(x)Px$  pociąga za sobą  $(x)Px$ ”, ani twierdzenie „ $(x) Px$  pociąga za sobą  $(x) Px$ ” nie posiadają tych samych warunków prawdziwości, co twierdzenie „ $(x)\Box Px$ ”, a przynajmniej nie posiadają ich, gdy przyjmiemy, iż dziedziny przedmiotowe światów możliwych mogą różnić się od siebie. Załóżmy, że istnieje świat możliwy, w którym istnieją wyłącznie elektrony, oraz że elektrony z konieczności posiadają własność  $F$ , wówczas w świecie tym będzie prawdą, że  $(x)\Box Fx$ . Nie oznacza to jednak, że negacja tego twierdzenia bądź negacja predykatu występującego w tym twierdzeniu zawsze będzie pociągała za sobą zaprzeczenie tego twierdzenia. W świecie, w którym nie ma elektronów, można bowiem zaprzeczyć twierdzeniu  $(x)\Box Px$ , nie popadając w sprzeczność. Niewykluczone, że twierdzenia modalne *de re* mają swoje odpowiedniki pod postacią twierdzeń modalnych *de dicto*, nie chcę jednak angażować się tutaj w rozważanie tych kwestii. Pragnę jedynie zwrócić uwagę

na to, że jeśli dla dowolnego twierdzenia modalnego *de re*, istnieje jego parafraza w postaci twierdzenia *de dicto*, to wydaje się, iż w wielu wypadkach będzie trzeba uciec się do dyskursu możliwościowego celem ustalenia, jakie twierdzenie modalne *de dicto* mogłoby pełnić rolę takiej parafrazy. Z ostatniego silny modalny fikcjonalista mógłby wnosić, iż dyskurs możliwościowy pozwala na analizowanie warunków prawdziwości twierdzeń modalnych.

W literaturze pojawia się jeszcze inny problem związany z niekompletnością treści fikcji, który został sformułowany przez Rosena przy okazji pierwszej publikacji poświęconej fikcjonalizmowi modalnemu (1990, § 7). Zarzut Rosena bierze się z tego, że realista modalny, chcąc odeprzeć zarzut Peter Forresta i David Armstronga (1984), musi przyjąć, iż istnieje górna granica liczności dziedzin przedmiotowych światów możliwych. W przeciwnym razie do realizmu modalnego stosuje się następująca argumentacja Forresta i Armstronga: zgodnie z Lewisowską teorią światów możliwych, każdy świat możliwy jest maksymalną sumą mereologiczną indywiduów i jako taka suma sam jest indywiduum. Weźmy teraz *nieograniczoną zasadę rekombinacji* głoszącą, że:

(NZR) dla dowolnej klasy indywiduów, istnieje świat zawierający jako swoje części dowolną liczbę duplikatów tych indywiduów.

Stosując (NZR) do klasy wszystkich światów możliwych, otrzymujemy *duży świat*, który zawiera wszystkie wyjściowe światy jako swoje nieprzecinające się części. Z uwagi na to, że wychodziliśmy od zbioru wszystkich światów możliwych, duży świat musi być elementem tego zbioru. Problem z dużym światem jest jednak taki, że gdy przyjmiemy założenie, iż liczba elektronów istniejących w tym świecie jest równa  $C$ , to istnieje  $2^C - 1$  niepustych podzbiorów elektronów istniejących w dużym świecie. Dla każdego takiego podzbioru — na mocy (NZR) — istnieje świat możliwy, którego częściami są duplikaty elektronów należących do tego podzbioru. Mamy zatem  $2^C - 1$  takich światów,

z których każdy zawiera co najmniej jeden elektron. Każdy z tych światów stanowi również część dużego świata, z czego wynika, że w dużym świecie istnieje co najmniej  $2^C - 1$  elektronów. Z założenia jednak duży świat miał mieć  $C$  elektronów, musi więc posiadać więcej części, niż posiada. Niemożliwe jednak jest, by duży świat był większy od samego siebie. Zatem duży świat nie istnieje. Jeśli Lewisowska koncepcja światów możliwych dopuszcza istnienie dużego świata, to stanowi ono *reductio ad absurdum* założenia, że istnieją Lewisowskie światy.

W odpowiedzi na ten zarzut Lewis (1986, s. 89–90, 102–103) twierdzi, że należy odrzucić (NZR) na rzecz *ograniczonej zasady rekombinacji*:

- (OZR) dla dowolnej klasy indywiduów, istnieje świat zawierający jako swoje części dowolną liczbę duplikatów tych indywiduów, pod warunkiem że pozwala na to wielkość dziedziny przedmiotowej danego świata.

W przypadku światów możliwych, których części powiązane są relacjami czasoprzestrzennymi, ograniczenie liczby duplikatów opiera się na założeniu, iż istnieje górna granica ilości obiektów, które mogą pomieścić się w jednej czasoprzestrzeni. Niech granicą tą będzie liczba  $C^*$ .

Niech  $\kappa$  będzie liczbą większą od liczby indywiduów istniejących w świecie aktualnym. Powstaje pytanie, jaka jest wartość logiczna twierdzenia (GG) „Mogłoby istnieć  $\kappa$  indywiduów nieposiadających części wspólnych”. Twierdzenie to posiada realistyczny przekład o postaci (GG-MR) „Istnieje świat możliwy, w którym istnieje  $\kappa$  indywiduów, nieposiadających wspólnych części”. Dla realisty modalnego twierdzenie to ma określoną wartość logiczną. Zależy ona od tego, czy istnieje świat, w którym istnieje  $\kappa$  indywiduów. Z kolei jego istnienie zależy od tego, czy jego dziedzina jest liczniejsza niż dziedzina przedmiotowa świata, w którym istnieje  $C^*$  indywiduów. Choć możemy

nie wiedzieć, czy istnieje świat  $\kappa$  indywiduów, i związku z tym, jaka jest wartość logiczna twierdzenia (GG), to wciąż jest ona określona ze względu na to, jaka jest wielość światów możliwych.

Rosen zauważa, że dla silnego fikcjonalisty modalnego fakt, iż wartość logiczna twierdzenia (GG) jest nam nieznana, jest znacznie bardziej problematyczny. Fikcjonalistycznym przekładem twierdzenia (GG) jest twierdzenie (GG-MF) „Zgodnie z fikcją *PW* istnieje świat możliwy, w którym istnieje  $\kappa$  indywiduów nieposiadających wspólnych części”. Z uwagi na to, że ani postulaty (2a) — (2g), ani encyklopedia nie mówią niczego ani o świecie  $C^*$  indywiduów, ani o świecie  $\kappa$  indywiduów, nie jesteśmy w stanie przyjąć twierdzenia (GG-MF) jako prawdziwego, gdyż jego treść sugeruje, że *PW* mówi coś na temat którejkolwiek z tych dwóch kwestii. Wydaje się zatem, że (GG-MF) nie jest prawdziwe.

Następnie Rosen rozważa dwie opcje, do których mógłby uciec się silny fikcjonalista modalny. Z jednej strony mógłby on utrzymywać, iż twierdzenie (GG-MF) jest fałszywe, co wydaje się naturalnym wyborem, gdyż twierdzenie to orzeka o opowieści *PW* własność, której ona nie posiada. Wówczas jednak fałszywe powinno być również twierdzenie (GG), którego analizą jest twierdzenie (GG-MF). Z fałszywości (GG) możemy wnosić o prawdziwości jego negacji (GG) „Nie mogłoby istnieć  $\kappa$  indywiduów nieposiadających wspólnych części”. (GG) jest z kolei analizowane za pomocą fikcjonalistycznego twierdzenia (GG-MF) „Zgodnie z *PW* nie istnieje świat, w którym istnieje  $\kappa$  indywiduów nieposiadających wspólnych części”. Z uwagi jednak na to, że hipoteza *PW* nie mówi niczego o światach  $\kappa$  indywiduów, (GG-MF) wydaje się fałszywe. Z jego fałszywości możemy wnosić o fałszywości (GG). Sprawia to, że silny fikcjonalista modalny jest zobowiązany do uznania fałszywości (GG) i fałszywości jego negacji. Twierdzenia są jednak ze sobą sprzeczne, a zatem prawdziwość jednego pociąga za sobą fałszywość drugiego i odwrotnie. W związku z tym uznanie



falszywości (GG-MF) nie jest opcją, z której mógłby skorzystać silny metafikcjonalista modalny.

Alternatywnie silny fikcjonalista modalny mógłby uznać (GG-MF) za pozbawione wartości logicznej. Konsekwencją przyjęcia tej opcji jest konieczność uznania, że stałe logiczne nie są funkcjami prawdziwościowymi, gdy ich członami są twierdzenia modalne. Jeśli (GG-MF) stanowi analizę (GG) i jest ono pozbawione wartości logicznej, to (GG) również jest pozbawione wartości logicznej. Podobnie jest z  $(\neg \text{GG})$  i  $(\neg \text{GG-MF})$  — oba pozbawione są wartości logicznych. Jeśli jednak założyć, że zgodnie z *PW* prawdziwe są podstawienia praw logicznych, wówczas prawdziwe powinno być twierdzenie „Zgodnie z *PW* istnieje świat możliwy, w którym istnieje  $\kappa$  indywiduów lub nieprawda, że istnieje świat możliwy, w którym istnieje  $\kappa$  indywiduów”. Ostatnie twierdzenie stanowi analizę twierdzenia „GG lub  $\neg \text{GG}$ ” i skoro analiza tego twierdzenia jest prawdziwa, to i ono powinno być prawdziwe. Silny fikcjonalizm modalny prowadzi zatem do konieczności uznania prawdziwości alternatywy, w której żaden z członów nie jest prawdziwy. Rosen podkreśla, że taka konsekwencja jest o tyle niewiarygodna, że w przypadku twierdzeń modalnych „nasze zwykłe sposoby myślenia nie dają nam powodu, by spodziewać się skomplikowanej logiki propozycjonalnej [dla tych twierdzeń — P.W.]” (Rosen 1990, s. 343).

Źródłem wszystkich podnoszonych przez Rosena trudności jest zaakceptowanie tezy, iż do celów metafikcjonalizmu modalnego należy dostarczenie redukcijnej teorii modalności. Jak widzieliśmy, wszystkie niepożądane konsekwencje wynikały z tego, że próbowano analizować warunki prawdziwości twierdzeń modalnych w kategoriach prawdziwości twierdzeń możliwoswiatowych zgodnie z fikcją *PW*. Gdy porzuci się ambicje silnego fikcjonalizmu modalnego, akceptując umiarkowaną odmianę fikcjonalizmu modalnego, można uniknąć powyższych trudności (zob. Nolan 2011b). Umiarkowani fikcjonałiści modalni akceptują bo-

wiem istnienie faktów modalnych fundujących prawdziwość twierdzeń modalnych. W związku z tym mogą uznać, że twierdzenia modalne są prawdziwe lub fałszywe niezależnie od tego, czy mamy wiedzę na temat ich wartości logicznej. Co więcej umiarkowani fikcjonałści modalni uzależniają to, czy dane twierdzenie możliwościwiatowe jest prawdziwe zgodnie z fikcją *PW* bądź jest prawdziwe w grze *GPW*, od tego, jakie są fakty modalne. W związku z tym to, czy twierdzenie (GG-MF) jest prawdziwe, czy fałszywe, zależy nie od tego, czy w treści opowieści *PW* mówi się cokolwiek o światach  $\kappa$  indywiduów, ale od tego, czy zachodzą fakty modalne fundujące prawdziwość twierdzenia (GG). Podobnie to, czy twierdzenie „Istnieje świat  $\kappa$  indywiduów” jest prawdziwe w grze *GPW*, jest zależne nie od tego, czy w grze *GPW* istnieją postulaty bazowe, które nakazują wyobrażenie sobie świata  $\kappa$  indywiduów, ale od tego, czy zachodzą fakty modalne fundujące prawdziwość twierdzenia (GG), które stanowią zarazem konieczny i wystarczający warunek fikcyjnej prawdziwości twierdzenia „Istnieje świat  $\kappa$  indywiduów”. W obu wypadkach konieczne jest uznanie za wiarygodne założenia, że zgodnie z treścią fikcji *PW* bądź w ramach gry *GPW* może być prawdziwe twierdzenie, które — odpowiednio — nie wynika (w szerokim sensie) z twierdzeń składających się na treść fikcji *PW* bądź nie wynika z żadnych twierdzeń, których prawdziwość wyobrażają sobie uczestnicy gry *GPW*. Gdy jednak uzależnia się prawdziwość (zgodnie z fikcją bądź w ramach gry) od obiektywnie zachodzących okoliczności, konsekwencja taka nie powinna szczególnie dziwić.

Zresztą zwolennicy umiarkowanych stanowisk fikcjonalistycznych z pełną świadomością akceptują tę konsekwencję. Dyskutując z Rosena problemem niekompletności, Nolan pisał, że „umiarkowany fikcjonałsta powinien być w stanie uczynić fikcję określoną w tym zakresie, w jakim korelaty twierdzeń modalnych są reprezentowane [w tej fikcji — P.W.]” (2011b). Wszystkie twierdzenia możliwościwiatowe poprzedzone

funktozem opowieści „Zgodnie z *PW*”, które pełnią rolę narzędzi użytecznych w reprezentowaniu faktów modalnych i — w związku z czym — których wartość logiczna jest zależna od tego, jakie są fakty modalne, można uznać za określone pod względem ich wartości logicznej, o ile tylko uznaje się, że fakty modalne są zdeterminowane.

W przypadku przedmiotowego fikcjonalizmu modalnego konsekwencja ta równoważna jest z uznaniem, że w ramach *GPW* pewne sądy mogą mieć ustaloną fikcyjną wartość logiczną, nawet gdy użytkownicy tej gry nie wyobrażają sobie istnienia bytów spełniających te twierdzenia. W (1990) Walton podejmuje pytanie: „Czy twierdzenie *p* może być fikcyjnie prawdziwe, jedynie gdy użytkownicy danej gry w udawanie wyobrażają sobie zachodzenie okoliczności, na mocy zachodzenia których twierdzenie to byłoby prawdziwe?”. Udziela na nie negatywnej odpowiedzi, pisząc, że „sąd jest fikcyjny [...], jeśli musi być wyobrażony (w pewnym kontekście), gdy powstanie kwestia jego fikcjonalności” (Walton 1990, s. 40). W naszym wypadku kwestia fikcyjnej prawdziwości pewnego twierdzenia możliwościowego reprezentującego fakty modalne będzie powstawała, gdy dowiemy się o zachodzeniu pewnego faktu modalnego, który — na mocy reguł generowania — stanowi konieczny i wystarczający warunek fikcyjnej prawdziwości tego twierdzenia możliwościowego.

Nie wszystkie jednak twierdzenia możliwościowe prawdziwe zgodnie fikcją *PW* pełnią rolę narzędzi pomocnych w reprezentowaniu faktów modalnych, niektóre z nich dotyczą przede wszystkim samych światów możliwych i nie ze wszystkich takich twierdzeń można wyprowadzić odpowiadające im twierdzenia modalne, wówczas — twierdzi Nolan (2011b) — umiarkowany metafikcjonalista modalny może uznać, że w przypadku niektórych takich twierdzeń będzie czymś nieokreślonym to, czy są one prawdziwe, czy są one fałszywe zgodnie z fikcją *PW*. Przykładem takiego twierdzenia może być twierdzenie dotyczące liczności uniwersum światów możliwych. Wydaje się bowiem, że nie

ma jedno-jednoznacznej relacji pomiędzy ilością faktów modalnych a ilością światów możliwych. Przeciwnie, za pomocą jednego świata można reprezentować wiele faktów modalnych — świat, w którym istnieją niebieskie łabędzie i mówiące osły, może służyć jako narzędzia do reprezentowania takich faktów modalnych, jak to, że bycie osłem nie wyklucza się z byciem mówiącym, bycie łabędziem nie wyklucza się z byciem niebieskim, że możliwe jest współegzystowanie łabędzi i osłów itd.

Podobnie rzecz przedstawia się w przypadku przedmiotowego fikcjonalizmu modalnego: niektóre z twierdzeń możliwościowych będą takimi twierdzeniami bazowej części gry *GPW*, że ich wartość logiczna nie będzie determinowana przez reguły postulatywne tej gry. W związku z tym mogą posiadać nieokreśloną wartość logiczną w ramach tej gry.

Choć w obu przypadkach dopuszcza się istnienie twierdzeń możliwościowych o nieokreślonej wartości logicznej, to z twierdzeń tych nie wyprowadza się już żadnych twierdzeń modalnych. W związku z tym ich występowanie nie prowadzi do konieczności rewidowania naszych sposobów myślenia o kwestiach modalnych.

## 5.2 Problem sztuczności treści fikcji

W (1997, s. 264–266) Nolan argumentuje, że silna wersja fikcjonalizmu modalnego prowadzi do konieczności przyjęcia szeregu niewiarygodnych tez na temat prawdziwości twierdzeń modalnych. Co oczywiste, fikcja *PW* mogłaby nie powstać. Gdyby tacy filozofowie jak Leibniz, Kripke czy Lewis nie zainteresowali się światami możliwymi, dyskurs możliwościowy, a za nim i fikcja *PW* mogłyby nie ujrzeć światła dziennego. Słowem, istnienie fikcji *PW* jest czymś przygodnym. Gdy zaprzęgniemy schemat (*MMF*<sub>3</sub>) do analizy warunków prawdziwości twierdzeń modalnych, okaże się, że gdyby fikcja *PW* nie istniała, to żadne z twierdzeń modalnych nie byłoby

prawdziwe. Równoważność pomiędzy twierdzeniami modalnymi a ich fikcjonalistycznymi odpowiednikami wymaga, by oba twierdzenia posiadały taką samą wartość logiczną. Gdyby fikcja *PW* nie istniała, wszystkie twierdzenia z funktorem opowieści albo byłyby fałszywe, albo pozbawione wartości logicznych, w zależności od naszej ulubionej analizy twierdzeń o niespełnionych presupozycjach egzystencjalnych. Intuicje podpowiadają nam jednak, że niektóre twierdzenia modalne byłyby prawdziwe, nawet gdyby fikcja *PW* nie powstała. Twierdzenia modalne nie są bowiem twierdzeniami, które są przygodnie prawdziwe lub fałszywe. Zatem warunków prawdziwości twierdzeń modalnych nie powinno się analizować w kategoriach prawdziwości twierdzeń możliwościowych zgodnie z fikcją *PW*. Nazwijmy tę obiekcję *modalnym problemem sztuczności*.

Fikcje nie tylko przygodnie istnieją, ale również zaczynają istnieć w pewnym momencie. Fikcja *PW* została utworzona przez Lewisa w 1986 roku<sup>44</sup>. We wszystkich chwilach sprzed utworzenia fikcji *PW* żadne z twierdzeń dotyczących jej treści nie były prawdziwe. Ponownie, gdy zaprzęgniemy schemat ( $MMF_3$ ) do analizy warunków prawdziwości twierdzeń modalnych, okaże się, że we wszystkich chwilach sprzed powstania fikcji *PW* żadne twierdzenie modalne nie było prawdziwe. Wydaje się jednak, że twierdzenia modalne były prawdziwe, zanim powstała fikcja *PW*. Zatem warunków prawdziwości twierdzeń modalnych nie powinno się analizować w kategoriach prawdziwości twierdzeń możliwościowych zgodnie z fikcją *PW* (zob. Kim 2005, s. 119–120). Nazwijmy tę obiekcję *temporalnym problemem sztuczności*.

Wartość logiczna twierdzenia „Zgodnie z fikcją *PW*,  $p^*$ ” ulegałaby zatem zmianie w zależności od tego, czy zostałoby wygłoszone przed, czy po 1986 roku oraz w zależności od tego, czy mówimy o

---

<sup>44</sup>Na potrzeby rozważań przyjmijmy, że Lewisowska teoria światów możliwych w jej pełnej postaci powstała w 1986 roku. Rzecz jasna, historycznie jest to teza fałszywa, gdyż część treści książki *On the Plurality of Worlds* to treść Wykładów Lockowskich, wygłoszonych przez Lewisa w 1984 roku, zob. (Lewis 1986, ix).

kontrfaktycznych scenariuszach, w których *PW* nie powstało, czy o świecie aktualnym po 1986 roku. By twierdzenia dotyczące fikcji *PW* posiadały niezmienną wartość logiczną, być może należałoby przeformułować je do postaci: „*Aktualnie, obecnie, zgodnie z fikcją PW, p\**” (por. Kim 2005, s. 128–129). Następnie można zmodyfikować prawą stronę schematu ( $\text{MMF}_3$ ), w wyniku czego otrzymujemy:

( $\text{MMF}_3^*$ ) „*p*” jest prawdziwe wtw, gdy *aktualnie, obecnie, zgodnie z PW, p\**.

*Prima facie* przeformułowanie to pozwala na uniknięcie sformułowanych wyżej trudności. Dzięki wprowadzeniu modalnego funktora aktualności prawa strona schematu ( $\text{MMF}_3^*$ ) wydaje się koniecznie prawdziwa. Z kolei dzięki wprowadzeniu funktora temporalnej aktualności wydaje się ona wiecznie i odwiecznie prawdziwa, tj. niech *t* będzie chwilą obecną, wówczas prawa strona tego schematu wydaje się prawdziwa w każdej chwili wcześniejszej od *t*, w chwili *t* oraz w każdej chwili późniejszej od chwili *t*.

Trudno jednak oprzeć się wrażeniu, że mamy tutaj do czynienia z pewnym technicznym wybiegiem, który prześlizguje się po istocie całego zarzutu. W powyższej próbie obejścia modalnego problemu sztuczności wciąż wydaje się intuicyjnie prawdziwe, że:

(RMF) gdyby *aktualnie* fikcja *PW* nie istniała, to żadne twierdzenie nie byłoby prawdziwe zgodnie z nią.

W odpowiedzi Saehwa Kim twierdzi, że „powyższy kontrfaktyczny okres warunkowy ma niemożliwy poprzednik, ponieważ opowieść *PW* istnieje obecnie w naszym świecie i dlatego jest trywialnie prawdziwy” (Kim 2005, s. 130). Jak przypuszczam, Kim sugeruje, że powyższy kontrfaktyczny okres warunkowy intuicyjnie wydaje się prawdziwy, dlatego że jest trywialnie prawdziwy. Odpowiedź taka rodzi jednak więcej problemów niż dostarcza wyjaśnień.

W pierwszym rzędzie z punktu widzenia pragmatyki trywialnie prawdziwe kontrfaktyczne okresy warunkowe używane są zwykle do wyrażenia przekonania, że twierdzenie znajdujące się w poprzedniku wygłoszonego kontrfaktycznego okresu warunkowego nie może być prawdziwe. Przykładowo, powiedzenie, że gdyby istniała największa liczba pierwsza, to świnie latałyby, w większości kontekstów wydaje się twierdzeniem prawdziwym, gdyż wygłaszane jest z intencją powiedzenia czegoś w stylu: „Teza o istnieniu największej liczby pierwszej jest z konieczności fałszywa” (zob. Vander Laan 2004, s. 261). Tymczasem powiedzenie, że gdyby *aktualnie* nie istniała fikcja *PW*, to żadne twierdzenie nie byłoby prawdziwe zgodnie z nią, nie wydaje się występować w kontekście, w którym wyraża się przekonanie o niemożliwości poprzednika tego kontrfaktycznego okresu warunkowego. Raczej wyraża przekonanie, że świat aktualny mógłby być na tyle odmienny, że twierdzenie „*Aktualnie*, zgodnie z *PW*, *p*” byłoby fałszywe, gdyż w świecie tym nie istniałaby opowieść *PW*.

Pozostając jeszcze przy dyskursie możliwościświatowym, za Martinem Daviesem i Lloydem Humberstonem (1980, s. 1–3 i przyp. 3) możemy w nieco bardziej formalny sposób przedstawić intuicję, którym dano wyraz pod koniec poprzedniego akapitu. Weźmy model  $M = \langle W, w^*, V \rangle$ , gdzie  $W$  jest zbiorem światów możliwych,  $w^*$  jest światem wyróżnionym w modelu  $M$ , tj. światem, który jest światem aktualnym w modelu  $M$ ,  $V$  jest wartościowaniem. Dla dowolnego świata  $w$  należącego do zbioru  $W$ , „*Aktualnie*,  $p$ ” jest prawdziwe w świecie  $w$  wtw, gdy  $p$  jest prawdziwe w świecie  $w^*$ . Z uwagi na prawdziwość twierdzenia „*Aktualnie*,  $p$ ” w każdym świecie w modelu  $M$  możemy powiedzieć, że jest ono koniecznie prawdziwe w tym modelu. Weźmy teraz model  $M' = \langle W', w^{**}, V' \rangle$ , gdzie  $W = W'$ ,  $V = V'$ ,  $w^* \neq w^{**}$  i  $w^{**}$  jest światem wyróżnionym w modelu  $M'$ , wówczas dla dowolnego świata  $v$  należącego do  $W'$ , „*Aktualnie*,  $q$ ” jest prawdziwe w świecie  $v$  wtw, gdy „ $q$ ” jest prawdziwe w świecie  $w^{**}$ .

Podobnie o twierdzeniu „Aktualnie,  $q$ ” możemy powiedzieć, że jest koniecznie prawdziwe w modelu  $M'$ . Dalej, niech  $M'$  będzie wariantem modelu  $M$  wtw, gdy  $W = W'$  i  $V = V'$ . Dysponując relacją bycia wariantem zachodzącą pomiędzy modelami, możemy wprowadzić modalny funktor stałości (*fixedly*)  $F$ : dla dowolnego świata  $w$  należącego do zbioru  $W$ , „ $F$  (Aktualnie,  $p$ )” jest prawdziwe w  $w$  wtw, gdy dla każdego modelu  $M'$  będącego wariantem modelu  $M$ , dla dowolnego świata  $v$  należącego do zbioru  $W'$ , „Aktualnie,  $p$ ” jest prawdziwe w świecie  $v$ . Widzimy zatem, że można wyróżnić dwa sensy, w których twierdzenie „Aktualnie,  $p$ ” jest koniecznie prawdziwe. Z jednej strony może być ono koniecznie prawdziwe *relatywnie* do tego, który świat jest światem aktualnym. Z drugiej strony może ono być prawdziwe *niezależnie* od tego, który świat jest światem aktualnym. Zastanówmy się teraz, w którym z tych sensów konieczności prawdziwe jest twierdzenie „Aktualnie, zgodnie z  $PW$ ,  $p^*$ ”. Z uwagi na to, że istnienie opowieści  $PW$  wydaje się czymś przygodnym, twierdzenie jest koniecznie prawdziwe jedynie w pierwszym sensie i nie jest koniecznie prawdziwe w drugim sensie. Gdy mówimy o tym, że gdyby inny świat był światem aktualnym, to „Aktualnie, zgodnie z  $PW$ ,  $p^*$ ” mogłoby być fałszywe, mamy właśnie na myśli to, że twierdzenie to nie jest koniecznie prawdziwe w drugim z wyróżnionych wyżej sensów.

Powyższa prezentacja odmiennych sensów, w jakich twierdzenie „Aktualnie, zgodnie z  $PW$ ,  $p^*$ ” jest koniecznie prawdziwe, ma tę wadę, że może sugerować, iż pewien świat możliwy mógłby zastąpić świat aktualny. Jednak właściwy sposób odczytania różnicy pomiędzy modelami  $M$  i  $M'$  jest następujący: światy wyróżnione w tych modelach jedynie reprezentują sposoby, na jakie mógłby istnieć świat aktualny. Zatem powiedzenie, że pewien świat możliwy mógłby być światem aktualnym, jest jedynie luźnym sposobem powiedzenia, że świat możliwy mógłby być na taki sposób, na jaki reprezentowany jest pewien świat możliwy.



By strategia argumentacyjna przyjęta przez Kim mogła zakończyć się powodzeniem, musiałyby twierdzić, że twierdzenie „*Aktualnie*, zgodnie z *PW*,  $p^*$ ” jest koniecznie prawdziwe w drugim z wyróżnionych sensów konieczności. Problem jednak w tym, że utrzymywanie takiego stanowiska byłoby uzasadnione jedynie wówczas, gdy albo żaden ze światów, który mógłby być światem aktualnym, nie mógłby różnić się od świata aktualnego pod względem istnienia opowieści *PW*, albo żaden inny świat nie mógłby być światem aktualnym.

Z uwagi na intuicyjnie prawdziwe przekonanie, że opowieść *PW* istnieje przygodnie, uważam, że pierwsza opcja nie ma szans na powodzenie. Samo dopuszczenie, by inne światy mogły pełnić rolę świata aktualnego, z pewnością pozwoli na znalezienie modelu, w którym wyróżniony świat będzie taki, że nie istniałaby w nim opowieść *PW*. Zatem zwolennik strategii Kim musiałby utrzymywać, że żaden inny świat nie mógłby być światem aktualnym poza tym światem, który jest światem aktualnym. Co mogłoby to oznaczać? Biorąc pod uwagę strategię odpierania modalnego problemu sztuczności, najprawdopodobniej oznacza to, że żadne z twierdzeń, które są prawdziwe w świecie aktualnym, nie mogłoby być fałszywe w tym świecie — koniec końców, twierdzenie  $p$  jest prawdziwe w świecie aktualnym wtw, gdy twierdzenie „*Aktualnie*,  $p$ ” jest prawdziwe. Wówczas jednak wydaje się, że dochodzimy do absurdałnej konsekwencji, zgodnie z którą dla dowolnego niemodalnego twierdzenia  $q$  prawdziwego w świecie aktualnym, twierdzenie to jest koniecznie prawdziwe, gdyż jest logicznie równoważne twierdzeniu „*Aktualnie*,  $q$ ”. Konsekwencja taka stanowi wystarczającą podstawę do odrzucenia powyższej odpowiedzi na modalny problem sztuczności.

Choć wyżej wspomniałem o temporalnym problemie sztuczności, zamierzam pominąć dyskusję nad nim z dwóch powodów. Po pierwsze, samo niepowodzenie w odparciu modalnego problemu sztuczności w sposób wystarczający dyskwalifikuje silną wersję fikcjonalizmu mo-

dalnego. Po drugie, podjęcie dyskusji nad zagadnieniem odwieczności i wieczności prawd wymagałoby przedyskutowania wielu kwestii, takich jak determinizm, warunki prawdziwości twierdzeń o przyszłości, warunki prawdziwości twierdzeń o przeszłości itp., które dalece wykraczają poza zakres rozważań prowadzonych w tej rozprawie.

Gdy fikcjonalista modalny porzuci ambicje dostarczenia redukcyjnej analizy warunków prawdziwości twierdzeń modalnych i przyjmie istnienie faktów modalnych, żaden z powyższych zarzutów nie będzie miał zastosowania do jego koncepcji. Fikcja *PW* bądź gra w udawanie *GPW* powstały bowiem po to, by ułatwić nam myślenie o faktach modalnych, które fundują prawdziwość twierdzeń modalnych, nie zaś po to, by redukować fakty modalne do faktów dotyczących treści fikcji *PW* czy też symulacji czynionych w ramach gry *GPW*. Jeśli fakty modalne istnieją z konieczności oraz wiecznie i odwiecznie, wówczas twierdzenia modalne, których prawdziwość jest ufundowana na ich istnieniu, również są konieczne, wiecznie i odwiecznie prawdziwe.

## 6 Trudności umiarkowanego fikcjonalizmu modalnego

Dwa ostatnie paragrafy pokazują, że we współzawodnictwie między silnym a umiarkowanym fikcjonalizmem modalnym palmę pierwszeństwa dzierży ostatni. Do tej pory jednak niewiele dowiedzieliśmy się o szczegółach tego stanowiska. Zwolennicy umiarkowanego fikcjonalizmu modalnego albo utrzymują, że prawdziwość zgodnie z fikcją *PW* „opiera się [za pośrednictwem praw pomostowych — P.W.] na faktach dotyczących modalności” (Nolan 1997, s. 262), albo utrzymują, że fakty modalne są takie, że stanowią (za pośrednictwem reguł generowania) konieczne i wystarczające warunki fikcyjnej prawdziwości niektórych twierdzeń możliwościowych (por. Yablo 1996, s. 279). Dopóki jednak zwolennik umiarkowanego fikcjonalizmu modalnego

nie powie niczego więcej o naturze faktów modalnych, nie jest jasne, czy proponowane przez niego stanowisko nie stanie się ofiarą poniższej obiekcji.

Jason Stanley w (2001, s. 41–43) twierdzi, że zwolennik umiarkowanego fikcjonalizmu ma problem z wyjaśnieniem zdolności użytkowników dyskursu D (poddawanego fikcjonalistycznej interpretacji) do rozumienia twierdzeń, z którymi nigdy wcześniej się nie zetknęli. Zwykle zdolność tę wyjaśnia się poprzez przyjęcie, że w języku obowiązuje zasada kompozycjonalności. Zgodnie z nią znaczenie złożonego wyrażenia wyznaczane jest przez znaczenia wyrażen składających się na nie oraz przez jego strukturę, czyli sposób, w jaki poszczególne wyrażenia składają się na wyrażenie złożone. Innymi słowy, znaczenie złożonego wyrażenia jest funkcją ze znaczeń wyrażen składających się na to złożone wyrażenie. Dzięki obowiązywaniu zasady kompozycjonalności na podstawie znajomości znaczeń wyrażen składających się na złożone twierdzenie, z którym użytkownik języka nigdy wcześniej się nie zetknął, oraz na podstawie znajomości struktury tego twierdzenia użytkownik języka może ustalić znaczenie złożonej całości. Jeśli znaczeniami twierdzeń są ich warunki prawdziwości, to w językach, w których obowiązuje zasada kompozycjonalności, można ustalić warunki prawdziwości złożonego twierdzenia na podstawie znajomości znaczeń wyrażen składających się na nie. W przypadku fikcjonalizmu przedmiotowego, na którym Stanley koncentruje swoją krytykę, sama znajomość znaczeń wyrażen składających się na złożone twierdzenie nie wystarczy do tego, by wyznaczyć warunki jego fikcyjnej prawdziwości. Jest tak dlatego, że fikcjonałiści przedmiotowi za pośrednictwem reguł generowania uzależniają fikcyjną prawdziwość twierdzeń dyskursu D od faktów, których elementy konstytutywne nie stanowią wartości semantycznych przyporządkowanych wyrażeniom tego dyskursu. W konsekwencji warunki fikcyjnej prawdziwości twierdzeń dyskursu D nie są wyznaczane wyłącznie przez znacze-

nia wyrażeń składających się na te twierdzenia oraz struktury tych twierdzeń. Nie jest więc jasne, czy fikcjonalista przedmiotowy jest w stanie wyjaśnić zdolność do rozumienia twierdzeń dyskursu D, z którymi użytkownik tego dyskursu nigdy wcześniej się nie zetknął. Nie jest bowiem jasne, czy proponowane przez niego reguły generowania są wystarczająco systematyczne, by odwołując się do nich, można było wyjaśnić zdolność do rozumienia złożonych wyrażeń. Sam Stanley wyraża pesymizm co do możliwości dostarczenia takich reguł generowania.

Analogiczny zarzut można sformułować pod adresem umiarkowanego metafikcjonalizmu. To, które z twierdzeń dyskursu D jest prawdziwe zgodnie z fikcją  $f$  stojącą za tym dyskursem, jest zależne (za pośrednictwem praw pomostowych) od zachodzenia pewnych faktów. Elementy konstytutywne tych faktów również nie stanowią wartości semantycznych wyrażeń składających się na twierdzenia prawdziwe zgodnie z fikcją  $f$ . Nie wystarczy zatem znać znaczeń wyrażeń składających się na twierdzenie  $p$  dyskursu D i postulatów składających się na fikcję  $f$ , by znać warunki prawdziwości twierdzenia „Zgodnie z  $f$ ,  $p$ ”. Należy jeszcze wiedzieć, czy zachodzą fakty, od których zależy prawdziwość twierdzenia „Zgodnie z fikcją  $f$ ,  $p$ ”. Warunki prawdziwości niektórych z twierdzeń o postaci „Zgodnie z  $f$ ,  $p$ ” nie podlegają zatem zasadzie kompozycyjności. Powstaje więc pytanie, czy prawa pomostowe są wystarczająco systematyczne, by powołując się na nie, można było wyjaśnić zdolność do rozumienia twierdzeń o postaci „Zgodnie z  $f$ ,  $p$ ”, z którymi użytkownik dyskursu D wcześniej się nie zetknął. Również w tej sprawie Stanley najprawdopodobniej zająłby pesymistyczne stanowisko. Nazwijmy podniesiony przez niego problem — *problemem systematyczności*.

Na marginesie warto odnotować, że problem ten nie powstaje dla silnej wersji metafikcjonalizmu. Z uwagi na to, że silny metafikcjonalista nie wprowadza zależności pomiędzy wartościami logicznymi

twierdzeń o postaci „Zgodnie z  $f$ ,  $p$ ” niezależnych od relacji pomiędzy treścią fikcji  $f$  i twierdzenia  $p$  faktów, wartości logiczne takich twierdzeń w całości są zależne wyłącznie od tego, jaka jest treść twierdzenia  $p$ , jaka jest treść fikcji  $f$  i czy pomiędzy treścią twierdzenia  $p$  a treścią fikcji  $f$  zachodzi szeroko pojęta relacja wynikania.

Wracając do problemu systematyczności, jak zauważa Yablo (2001, s. 96), reguły generowania różnią się w zależności od tego, o jakim dyskursie mowa. Dlatego ciężko oszacować wartość zarzutów Stanleya, dopóki nie przejdzie się do rozważań nad konkretną grą w udawanie i jej regułami generowania. Choć fikcyjnie prawdziwe twierdzenia nie podlegają zasadzie kompozycjonalności, zwolennik fikcjonalizmu przedmiotowego może utrzymywać, że reguły generowania są na tyle systematyczne, że zdolność do rozumienia nowych twierdzeń dyskursu  $D$  nie jest niczym tajemniczym. Podobnie jest w przypadku umiarkowanego metafikcjonalizmu. Dopóki nie przejdzie się do rozważań nad konkretną fikcją i jej prawami pomostowymi, dopóty ciężko oszacować wartość zarzutów Stanleya. Umiarkowany metafikcjonalista również może utrzymywać, że prawa pomostowe są na tyle systematyczne, że zdolność do rozumienia twierdzeń dotyczących prawdziwości zgodnie z fikcją  $F$  nie musi być niczym tajemniczym. Znalezienie takich reguł byłoby wystarczającą podstawą do tego, by odrzucić obawy podnoszone w ramach problemu systematyczności.

W przypadku dyskursu możliwościwiatowego można oddalić problem systematyczności, wskazując na takie reguły generowania bądź prawa pomostowe, które rzuciłyby więcej światła na zdolność do rozumienia twierdzeń możliwościwiatowych  $p^*$  prawdziwych w ramach gry  $GPW$  bądź na zdolność do rozumienia twierdzeń o postaci „Zgodnie z  $PW$ ,  $p^*$ ”. Stajemy więc przed pytaniem: „Czy są one wystarczająco systematyczne?”. Trudno oczekiwać jakiegokolwiek odpowiedzi, zanim umiarkowany fikcjonalista modalny nie przybliży swojej koncepcji faktów modalnych. Od nich bowiem ma być zależna fikcyjna prawdziwość

twierdzeń możliwościowych bądź ich prawdziwość zgodnie z *PW*.

Kolejnym powodem, dla którego umiarkowany fikcjonalista modalny powinien powiedzieć coś o faktach modalnych, jest to, że musi on wskazać na taką koncepcję faktów modalnych, która będzie spełniała dwa dezyderaty: ( $D_1$ ) fakty modalne nie mogą być identyczne z żadną z proponowanych przez aktualistów redukcją światów możliwych oraz ( $D_2$ ) fakty modalne muszą być takie, że możliwe jest wyjaśnienie roli, jaką pełni dyskurs możliwościowy w myśleniu o nich. W pierwszym wypadku fakty modalne nie mogą być identyczne z maksymalnymi stanami rzeczy, tj. takimi stanami rzeczy, że dla dowolnego stanu rzeczy  $x$ ,  $x$  jest zawarty w tym stanie rzeczy lub  $x$  jest wykluczony z tego stanu rzeczy. Stan rzeczy  $x$  jest wykluczony z maksymalnego stanu rzeczy, gdy niemożliwe jest, by zachodziły wszystkie stany rzeczy zawarte w tym maksymalnym stanie rzeczy i  $x$  zachodziło, a stan rzeczy  $x$  jest zawarty w maksymalnym stanie rzeczy wtw, gdy nie jest wykluczony z tego maksymalnego stanu rzeczy. Elementami konstytutywnymi stanów rzeczy są tutaj takie byty jak indywidualne istoty (tj. takie zbiory własności, że relatywnie do danego maksymalnego stanu rzeczy, dla dowolnej własności, albo własność ta należy do tego zbioru, albo nie należy do niego), własności oraz relacje (zob. Plantinga 1974). Gdyby fakty modalne były takimi maksymalnymi stanami rzeczy, wówczas przyjęcie umiarkowanego fikcjonalizmu modalnego byłoby bezcelowe. Stanowisko to nie oferowałoby niczego ponad to, co oferuje taki umiarkowany realista modalny jak Plantinga. Dodatkowo jednak rodziłoby trudności z tytułu wprowadzania reguł generowania czy praw pomocowych. W sprawie zasadności dezyderatu ( $D_2$ ) należy zauważyć, że umiarkowany fikcjonalista modalny zwykle proponuje wprowadzenie dyskursu możliwościowego, gdyż uważa go za użyteczne narzędzie reprezentowania pewnych własności faktów modalnych. Dlatego też musi udzielić odpowiedzi na pytanie, jakie własności faktów modalnych są reprezentowane przez twierdzenia możliwościowe. Odpowiedź

wskaże zarazem, jaką funkcję pełnią reguły generowania bądź prawa pomostowe. Dzięki temu będziemy mogli oszacować, czy pozwalają one na uniknięcie problemu systematyczności.

Poniżej ograniczę się do naszkicowania pewnej koncepcji faktów modalnych (§ 4.6.1.). Następnie w relacji do tej koncepcji faktów modalnych postaram się ustalić rolę dyskursu możliwościowego i sformułować odpowiednie reguły generowania oraz prawa pomostowe (§ 4.6.2.). Na podstawie ustaleń poczynionych w tym paragrafie poddam rewizji koncepcję faktów modalnych z § 4.6.1., nakreślając jedynie niektóre z cech nowej koncepcji faktów modalnych i pokazując, że spełnia ona dezyderaty ( $D_1$ ) i ( $D_2$ ). W § 4.6.3. wskażę na zgubne dla umiarkowanych fikcjonalistów trudności, których źródłem wydaje się nowo wprowadzona koncepcja faktów modalnych.

## 6.1 Fakty modalne — pierwsze podejście

Gdy rozejrzeć się po dostępnych kategoriach bytów, najprawdopodobniej umiarkowany fikcjonalista modalny będzie musiał uciec się do jakiejś koncepcji ustrukturyzowanych faktów modalnych, których elementami konstytutywnymi będą takie byty jak platońskie uniwersalia. Ostatnie wydają się pożądane dla celów umiarkowanych fikcjonalistów modalnych z następujących powodów:

- (a) platońskie uniwersalia są bytami *ogólnymi* w tym sensie, że można je scharakteryzować bez odwoływania się do indywiduów. Dzięki temu akceptacja ich istnienia nie powinna wymagać przyjęcia istnienia możliwych indywiduów;
- (b) platońskie uniwersalia mogą istnieć niezależnie od tego, czy są egzemplifikowane przez indywidua. Dzięki temu dopuszczalne będą fakty modalne, których elementami konstytutywnymi są tzw. *obce* (*alien*) własności, tj. takie własności, które ani nie są egzemplifikowane przez żadne z aktualnie istniejących indywiduów, ani

nie są redukowalne do kombinacji aktualnie egzemplifikowanych własności. Mając takie fakty w ontologii, można nadać im rolę bytów fundujących prawdziwość takich twierdzeń modalnych jak „Pewne przedmioty mogłyby mieć zupełnie odmienną naturę od jakiegokolwiek z aktualnie istniejących przedmiotów”;

- (c) platońskie uniwersalia — jak będę argumentował — są bytami tego rodzaju, że dyskurs możliwościwiatowy reprezentuje własności zwykle kojarzone z tymi bytami. Dzięki temu można zidentyfikować sens, w jakim dyskurs możliwościwiatowy ułatwia myślenie o faktach.

Wadą wprowadzenia bytów ogólnych jako elementów konstytutywnych faktów modalnych jest to, że kolejnemu ograniczeniu ulega hasło zachowania wszystkich zalet realizmu modalnego bez przyjmowania istnienia światów możliwych i możliwych indywiduów. Teraz obok niemożliwości przeprowadzenia teoriomnogościowej redukcji sądów w sensie logicznym umiarkowany fikcjonalista modalny nie jest również w stanie — w odróżnieniu od realisty modalnego — zredukować własności do zbiorów indywiduów oraz relacji do zbiorów  $n$ -tek uporządkowanych i jest zmuszony do wprowadzenia bytów intensjonalnych tam, gdzie realista modalny odwołuje się do bytów o ekstensjonalnych warunkach identyczności. Co więcej, nie jest jasne, czy wprowadzając nieegzemplifikowane własności i relacje umiarkowany fikcjonalista modalny nadal będzie mógł twierdzić, iż jego stanowisko jest spójne z ewolucyjną koncepcją kształtowania się zdolności poznawczych człowieka. Platońskie uniwersalia raczej nie wchodzą w relacje przyczynowo-skutkowe z ludźmi, a mimo to fikcjonalista modalny musi przypisać ludziom zdolność poznawania ich.

Jakimi zatem konstrukcjami są fakty modalne? Wyjdźmy od następującej próby: atomowe fakty modalne są zbiorami współwystę-



pujących (*compresent*) uniwersaliów<sup>45</sup>. Elementami tych zbiorów są albo jednoargumentowe własności, albo zbiory, których elementami są  $n$ -argumentowa relacja i  $n$ -tka uporządkowana (której elementami są zbiory współwystępujących jednoargumentowych własności). Niech współwystępowanie (w skrócie:  $C$ ) będzie funkcją ze zbiorów własności w fakty modalne. Jeśli własności należące do zbioru  $X$  są współwystępujące, wówczas wartością  $C(X)$  będzie fakt modalny. Jeśli nie są współwystępujące, wówczas wartością funkcji  $C(X)$  będzie zbiór pusty. W pierwszym wypadku powiemy, że istnieje takie  $x$ , że  $x = C(X)$ , w drugim — że nie istnieje takie  $x$ . Celem uzyskania faktów modalnych, w których będą występowały  $n$ -argumentowe relacje, musimy zezwolić na iteracyjne stosowanie funkcji  $C$ . Wówczas fakt, że pewna  $n$ -argumentowa relacja współwystępuje z pewną  $n$ -tką uporządkowaną zbiorów współwystępujących własności, będzie wartością funkcji  $C(\{R^n, < C(X_1), \dots, C(X_n) >\})$ . Jeśli relacja  $R^n$  nie współwystępuje z  $n$ -tką uporządkowaną  $< C(X_1), \dots, C(X_n) >$ , wówczas wartością  $C(\{R^n, < C(X_1), \dots, C(X_n) >\})$  jest zbiór pusty. Podobnie jak wyżej, w pierwszym wypadku powiemy, że istnieje takie  $x$ , że  $x = C(\{R^n, < C(X_1), \dots, C(X_n) >\})$ , w drugim — że nie istnieje takie  $x$ .

Na gruncie takiej koncepcji faktów modalnych twierdzenia modalne będą — w uproszczeniu — twierdzeniami o współwystępowaniu uniwersaliów należących do pewnego zbioru. Przykładowo, twierdzenie „Pewne  $x$  mogłoby być  $F$ ” jest prawdziwe wtw, gdy własność  $F$  jest własnością współwystępującą (w skrócie  $C(\{F\})$ ) lub — precyzyjniej — istnieje takie  $x$ , że  $x = C(\{F\})$ . Z kolei twierdzenie „Relacja  $R^n$  mogłaby zachodzić pomiędzy  $x_1, \dots, x_n$ ” jest prawdziwe wtw, gdy istnieje takie  $x$ , że  $x = C(\{R^n, < C(X_1), \dots, C(X_n) >\})$ , tj. relacja  $R^n$  współwystępuje z  $n$ -tką  $< C(X_1), \dots, C(X_n) >$ , której każdy element jest zbiorem współwystępujących własności. Zbiory te — jak

<sup>45</sup>Opieram się tutaj na koncepcji faktów modalnych dyskutowanej w (Hawthorne i Sider 2006, s. 33–35).

normalnie powiedzielibyśmy — stanowią jakościową charakterystykę członów relacji  $R^n$ . W celu znalezienia bytów fundujących prawdziwość twierdzeń modalnych z funktorem konieczności najprawdopodobniej należałoby wprowadzić własności negatywne. O własności  $F$  powiemy, że jest własnością negatywną względem własności  $F$  wtw, gdy nie istnieje takie  $x$ , że  $x = C(\{ F, F \})$ . Intuicyjnie, własność  $F$  stanowi sumę wszystkich własności, które wykluczają się z własnością  $F$ . O własności  $F$  powiemy, że stanowi własność negatywną względem zbioru własności  $X$  wtw, gdy co najmniej jedna z własności należących do zbioru  $X$  jest taka, że własność  $F$  jest negatywna względem niej. Twierdzenie „Koniecznie, dla każdego  $y$ ,  $Fy$ ” jest prawdziwe wtw, gdy nie istnieje takie  $x$ , że  $x = C(\{ F \})$ , twierdzenie „Koniecznie, dla każdego  $y$ , jeśli  $Fy$ , to  $Gy$ ” jest prawdziwe wtw, gdy nie istnieje takie  $x$ , że  $x = C(\{ F, G \})$  itd.

Nakreślona wyżej koncepcja faktów modalnych jako zbiorów współwystępujących własności jest zbyt gruboziarnista, tj. doprowadza do utożsamienia *prima facie* odróżnialnych możliwości. Rozważmy dwie możliwości:

- (P<sub>1</sub>) mogłyby istnieć trzy indywidua, z których pierwszy jest  $F$  i pozostaje w relacji  $R$  do drugiego indywiduum będącego  $G$ , które z kolei pozostaje w relacji  $R$  do trzeciego indywiduum, które jest  $H$ ;
- (P<sub>2</sub>) mogłyby istnieć cztery indywidua, z których pierwsze jest  $F$  i pozostaje w relacji  $R$  do drugiego indywiduum będącego  $G$  i istnieje trzecie indywiduum będące  $G$ , które pozostaje w relacji  $R$  do czwartego indywiduum będącego  $H$ .

Dla przykładu, w pierwszym przypadku łysy mężczyzna mógłby być przełożonym wysokiego mężczyzny, który z kolei mógłby być przełożonym niskiego mężczyzny. W drugim wypadku łysy mężczyzna mógłby być przełożonym wysokiego mężczyzny i pewien inny wysoki

mężczyzna mógłby być przełożonym niskiego mężczyzny. Choć obie możliwości wydają się odmienne, w ramach powyższej koncepcji faktów modalnych są reprezentowane w ten sam sposób:  $(CP)C(\{C(\{R, < C(\{F\}), C(\{G\}) >\}), C(\{R, < C(\{G\}), C(\{H\}) >\})\})$  (por. Hawthorne i Sider 2006, s. 35). Czy akceptacja umiarkowanego fikcjonalizmu modalnego pozwoli na sformułowanie bardziej drobnoziarnistej koncepcji modalnych faktów?

## 6.2 Rola fikcji światów możliwych

Dysponując zarysowaną wyżej koncepcją faktów modalnych, możemy rzucić nieco więcej światła na relację pomiędzy istnieniem faktu modalnego a fikcyjną prawdziwością twierdzeń możliwościowatych bądź prawdziwością twierdzeń możliwościowatych zgodnie z *PW*. Jak zatem przedstawiają się reguły generowania bądź prawa pomostowe?

Zacznijmy od przedmiotowego fikcjonalizmu modalnego. W ramach gry *GPW* elementom konstytutywnym faktów modalnych przyporządkowane są następujące fikcyjne byty: własnościom — zbiory możliwych indywiduów posiadających te własności, relacjom  $n$ -argumentowym — zbiory  $n$ -tek uporządkowanych, których elementami są możliwe indywidua pozostające do siebie w tych relacjach, faktom modalnym — zbiory światów możliwych takich, że istnieją w nich indywidua posiadające własności lub pozostające do siebie w relacjach stanowiących elementy konstytutywne faktów modalnych. Na podstawie powyższych przyporządkowań możemy wprowadzić następujące reguły generowania:

- (RG-P) dla każdej własności  $F$ , jeśli własność  $F$  stanowi element konstytutywny jakiegokolwiek faktu modalnego, to jej istnienie stanowi konieczną i wystarczającą podstawę do wyobrażania sobie, że istnieje zbiór możliwych indywiduów posiadających własność  $F$ , jeśli tylko kwestia fikcyjnego istnienia (tj. kwestia istnienia w ramach *GPW*) takich zbiorów indywiduów powstanie;

- (RG-R) dla każdej relacji  $R^n$ , jeśli relacja  $R^n$  stanowi element konstytutywny jakiegokolwiek faktu modalnego, to jej istnienie stanowi konieczną i wystarczającą podstawę do wyobrażenia sobie, że istnieje zbiór  $n$ -tek uporządkowanych, których elementami są możliwe indywidua pozostające do siebie w tej relacji, jeśli tylko kwestia fikcyjnego istnienia takich zbiorów  $n$ -tek uporządkowanych powstanie;
- (RG-F) dla każdego faktu modalnego  $x$ , istnienie  $x$  stanowi konieczną i wystarczającą podstawę do wyobrażania sobie, że istnieją światy możliwe, w których istnieją indywidua posiadające własności lub pozostające do siebie w relacjach stanowiących elementy konstytutywne  $x$ -a, jeśli tylko kwestia fikcyjnego istnienia takich światów powstanie.

W ostatnim wypadku uczestnicy gry *GPW* będą sobie wyobrażali istnienie takich światów, w których będą istniały indywidua i  $n$ -tki uporządkowane możliwych indywiduów, za których pomocą reprezentuje się (*via* RG-P i RG-R) własności i relacje stanowiące elementy konstytutywne tych faktów modalnych. Wyobrażenie sobie (na podstawie faktów modalnych), że istnieją pewne światy możliwe, stanowi podstawę do zaakceptowania fikcyjnej prawdziwości pewnych twierdzeń egzystencjalnych dotyczących tych światów możliwych. Związek pomiędzy składnikami treści twierdzeń możliwościwiatowych a elementami konstytutywnymi faktów modalnych wydaje się na tyle systematyczny, że na podstawie znajomości faktów modalnych i znajomości reguł (RG-P), (RG-R) oraz (RG-F) można przewidzieć, istnienie jakich światów powinien sobie wyobrażać uczestnik gry *GPW* i jakie będą warunki fikcyjnej prawdziwości twierdzeń możliwościwiatowych, z którymi dotąd użytkownik dyskursu się nie zetknął. Dla przykładu, gdy mamy fakt modalny, że własność  $F$  współwystępuje z własnością  $G$ , wówczas za pośrednictwem reguły (RG-P) uczestnik gry *GPW* powinien wyobrazić sobie, że istnieje zbiór możliwych indywiduów, które

są  $F$ , oraz że istnieje zbiór możliwych indywiduów, które są  $G$ . Z kolei fakt modalny będzie reprezentowany przez zbiór światów możliwych, w których istnieją indywidua stanowiące wspólne elementy zbiorów, za których pomocą reprezentuje się te własności. Wyobrażenie sobie istnienia takich światów stanowi podstawę do zaakceptowania twierdzenia „Istnieje taki świat  $w$ , że istnieje takie  $x$ , że  $Fx$  i  $Gx$  w  $w$ ” jako fikcyjnie prawdziwego. Wydaje się zatem, że zależność fikcyjnej prawdziwości twierdzeń od tego, jakie fakty modalne istnieją, jest wystarczająco systematyczna, by zdolność uczestników tej gry do rozumienia twierdzeń możliwościwiatowych, z którymi się dotąd nie zetknęli, nie była owiana tajemnicą.

W przypadku umiarkowanego metafikcjonalizmu modalnego można zaproponować następujące prawo pomostowe uzależniające prawdziwość twierdzeń zgodnie z  $PW$  od tego, jakie są fakty modalne:

- (PP-F) dla każdego faktu modalnego  $x$ , istnienie  $x$  stanowi konieczną i wystarczającą podstawę prawdziwości (zgodnie z  $PW$ ) twierdzeń, w których stwierdza się istnienie światów możliwych, w których istnieją indywidua posiadające własności lub pozostające do siebie w relacjach stanowiących elementy konstytutywne  $x$ -a, jeśli kwestia opisu tego faktu w  $PW$  powstanie.

W celu uzyskania systematycznych związków pomiędzy treścią wyrażen składających się na twierdzenia możliwościwiatowe prawdziwe zgodnie z  $PW$  a elementami konstytutywnymi faktów modalnych umiarkowany metafikcjonalista modalny może zaproponować następujące reguły wiążące występowanie w fikcji  $PW$  wyrażen o pewnej treści z elementami konstytutywnymi faktów modalnych:

- (PP-P) dla każdej własności  $F$ , jeśli własność  $F$  stanowi element konstytutywny jakiegokolwiek faktu modalnego, to jej istnienie stanowi konieczną i wystarczającą podstawę do wprowadzenia do treści fikcji  $PW$  wyrażenia „zbiór możliwych indywiduów,

które są  $F''$ , którego treścią będzie reprezentacja nieokreślonej wielości indywiduów takiej, że każde z tych indywiduów stanowi element tego zbioru, jeśli tylko kwestia opisu któregośkolwiek z tych faktów w  $PW$  powstanie;

- (PP-R) dla każdej relacji  $R^n$ , jeśli relacja  $R^n$  stanowi element konstytutywny jakiegokolwiek faktu modalnego, to jej istnienie stanowi konieczną i wystarczającą podstawę do wprowadzenia do treści fikcji  $PW$  wyrażenia „zbiór  $n$ -tek uporządkowanych możliwych indywiduów, które pozostają w relacji  $R^n$ ”, którego treścią będzie reprezentacja nieokreślonej wielości indywiduów takiej, że każde z tych indywiduów stanowi element co najmniej jednej  $n$ -tki uporządkowanej należącej do tego zbioru, jeśli tylko kwestia opisu któregośkolwiek z tych faktów w  $PW$  powstanie.

Dwie ostatnie reguły nie są, ściśle rzecz biorąc, prawami pomostowymi, jeśli przez ostatnie rozumiemy powiązania pomiędzy istnieniem faktów modalnych a treścią *propozycjonalną*, która jest prawdziwa zgodnie z  $PW$ . Reguły te jednak wyznaczają relacje pomiędzy elementami konstytutywnymi faktów modalnych a znaczeniami wyrażen składających się na twierdzenia możliwościwiatowe. Tym samym sprawiają, że zachodzi systematyczny i podlegający analizie związek pomiędzy faktami modalnymi a prawdziwością (zgodnie z  $PW$ ) twierdzeń możliwościwiatowych. Weźmy fakt modalny  $C(\{F, G\})$ , wówczas na mocy (PP-F) prawdziwe będzie twierdzenie „Zgodnie z  $PW$ , istnieje taki świat możliwy  $w$ , że w którym istnieje takie indywiduum  $x$ , że  $x$  należy do zbioru indywiduów, którą są  $F$ , i  $x$  należy do zbioru indywiduów, które są  $G$ ”. W twierdzeniu tym występują takie wyrażenia, jak „zbiór indywiduów, które są  $F$ ” i „zbiór indywiduów, które są  $G$ ”, których znaczenia stanowią reprezentacje własności  $F$  i  $G$  w ramach fikcji  $PW$ . Znając zatem elementy konstytutywne faktu  $C(\{F, G\})$ , jego strukturę oraz prawa (PP-P) i (PP-F), można przewidzieć, jakie będą warunki prawdziwości (zgodnie z  $PW$ ) twierdzenia

możliwoświatowego, nawet jeśli nigdy wcześniej nie zetknęło się z tym twierdzeniem.

Umiarkowany fikcjonalista modalny nie może jednak poprzestać na nakreślonych wyżej regułach generowania lub prawach pomostowych. Z uwagi na to, że dyskurs światów możliwych ma uławiać poznawanie pewnych własności faktów modalnych, z fikcyjnej prawdziwości twierdzeń możliwościwiatowych bądź z ich prawdziwości zgodnie z *PW* powinno dać się wnioskować o własnościach faktów modalnych. W tym celu zostały wprowadzone odwrócone reguły generowania, a tutaj dodatkowo musimy prowadzić odwrócone prawa pomostowe. Celem dowiedzenia się, jaką postać powinny przyjąć odwrócone reguły generowania i odwrócone prawa pomostowe, należy postawić pytanie, czego możemy dowiedzieć się o faktach modalnych z dyskursu możliwościwiatowego?

Wróćmy do przykładów możliwości ( $P_1$ ) i ( $P_2$ ). Obie możliwości są odróżnialne w ramach dyskursu możliwościwiatowego. Jakie własności faktu modalnego (CP) są ujawniane przez twierdzenia możliwościwiatowe — odpowiednio — „Istnieje taki świat  $w$ , że istnieje takie  $x, y, z$ , że  $Fx$  i  $Gy$ , i  $Hx$ , i  $xRy$ , i  $yRz$ , i  $x \neq y$ , i  $y \neq z$  i  $x \neq z$ ” oraz „Istnieje taki świat  $w$ , że istnieje takie  $x, y, u, z$ , że  $Fx$  i  $Gy$ , i  $Gu$ , i  $Hx$ , i  $xRy$ , i  $uRz$ , i  $x \neq y$ , i  $y \neq z$ , i  $x \neq z$ , i  $x \neq u$ , i  $y \neq u$  i  $u \neq z$ ”? Tych możliwości nie odróżnia ich czysto jakościowa charakterystyka. Mogą być reprezentowane przez współwystępowanie tych samych własności i relacji. Różnią się jednak *sposobem*, na jaki te współwystępujące uniwersalia mogłyby być egzemplifikowane przez indywidua. W pierwszym wypadku mogłyby to być trzy indywidua, w drugim — cztery. Nazwijmy sposoby, na jakie współwystępujące uniwersalia mogłyby być egzemplifikowane, *wzorcami egzemplifikacji*. Wówczas można postawić następującą hipotezę: rola dyskursu możliwościwiatowego polega na ujawnianiu wzorców egzemplifikacji danego faktu modalnego.

Bez wprowadzenia wzorców egzemplifikacji dyskurs możliwościwia-

towy raczej przekłamywałby modalną rzeczywistość niż ułatwiał jej zrozumienie. Dyskurs możliwościowy pozwalałby bowiem na odróżnienie większej liczby możliwości niż pozwalają na to fakty modalne. W jego ramach odróżnialne są możliwości  $(P_1)$  i  $(P_2)$ . Gdyby jednak zbiory współwystępujących własności nie posiadały własności bycia możliwie egzemplifikowanymi na takie a takie sposoby, istniałby wyłącznie jeden fakt modalny (CP), który odpowiadałby po stronie rzeczywistości możliwościom  $(P_1)$  i  $(P_2)$ . Jeśli jednak to fakty modalne wyznaczają to, co jest możliwe, to dyskurs możliwościowy w sposób nieadekwatny reprezentowałby modalną rzeczywistość. Miałoby to — jak sądzę — zgubne konsekwencje dla tezy, że dyskurs możliwościowy pełni użyteczne narzędzie reprezentowania faktów modalnych. Pełnienie takiej funkcji powinno wiązać się ze zdolnością do wyprowadzenia z twierdzeń możliwościowych odpowiednich twierdzeń modalnych, które mają już dotyczyć faktów modalnych. Gdyby jednak dyskurs możliwościowy w sposób nieadekwatny reprezentował modalną rzeczywistość, istniałoby ryzyko, że z fikcyjnie prawdziwych bądź prawdziwych zgodnie z fikcją *PW* twierdzeń możliwościowych wyprowadza się fałszywe twierdzenia modalne. Miałoby to miejsce, gdyby np. chcieć twierdzić, że z tego, że  $(P_1)$  i  $(P_2)$  są odróżnialne w ramach dyskursu możliwościowego, wynika, że istnieją odmienne fakty modalne reprezentowane przez każdą z tych możliwości. W mojej opinii stanowi to podstawę do kwestionowania użyteczności dyskursu możliwościowego w reprezentowaniu faktów modalnych.

Jeśli zgodzić się z powyższą argumentacją i uznać konieczność odwoływania się do wzorców egzemplifikacji, palące staje się pytanie, czym one są. Czy są ufundowane na strukturze tych faktów oraz na własnościach posiadanych przez elementy konstytutywne tych faktów? Wydaje się, że następujące własności uniwersaliów częściowo wyznaczają wzorce egzemplifikacji:

- (a) własność uniwersaliów, że mogą one być egzemplifikowane przez



wiele indywiduów lub  $n$ -tek uporządkowanych indywiduów;

- (b) formalne własności relacji takie jak  $n$ -arność, zwrotność, niezwrotność, symetryczność, asymetryczność, antysymetryczność, przechodniość itd.;
- (c) metafizyczna natura relacji, która może nakładać ograniczenia na sposoby egzemplifikacji współwystępujących z nią własności.

Wzorce egzemplifikacji będą również wyznaczone przez wewnętrzną strukturę faktów modalnych. Przykładowo, gdy mamy fakt modalny  $C(\{R, < C(\{F\}), C(\{G\}) >\})$ , gdzie relacja  $R$  jest relacją niezwrotną i asymetryczną, to dopuszczalny wzorec egzemplifikacji tego faktu jest następujący: indywiduum, które jest  $F$ , pozostaje w relacji  $R$  do indywiduum, które jest  $G$ . Z klasy wzorców egzemplifikacji tego faktu modalnego będzie natomiast wykluczony następujący: indywiduum, które jest  $G$ , pozostaje w relacji  $R$  do indywiduum, które jest  $F$ . Jest on wykluczony zarówno z uwagi na niezwrotność i asymetryczność  $R$ , jak i samą strukturę faktu modalnego — to  $C(\{F\})$ , a nie  $C(\{G\})$ , stanowi pierwszy element pary uporządkowanej.

Dysponując wiedzą dotyczącą własności uniwersaliów (a)–(c) oraz wewnętrznej struktury danego faktu modalnego można przewidzieć sposoby, na jakie dany fakt modalny mógłby być egzemplifikowany. Gdy weźmiemy fakt (CP) i założymy, że  $R$  jest relacją binarną i niezwrotną zachodzącą pomiędzy indywiduami, oraz że metafizyczna natura relacji  $R$  nie przekreśla tego, że jeśli coś jest jej drugim członem, to może być jej pierwszym członem, wówczas można przewidzieć, iż fakt (CP) może być egzemplifikowany na sposoby (P<sub>1</sub>) i (P<sub>2</sub>). Na mocy (a) własność  $G$  może być egzemplifikowana przez więcej niż jedno indywiduum. Dzięki temu oba jej wystąpienia w fakcie (CP) nie muszą być egzemplifikowane przez jedno i to samo indywiduum. Z uwagi na binarność i niezwrotność relacji  $R$ , każde wystąpienie tej relacji musi być egzemplifikowane przez dwa indywidua. Na podstawie

tych własności elementów konstytutywnych faktu (CP) i jego struktury, można uznać, iż dopuszczalnym wzorcem jego egzemplifikacji jest (P<sub>2</sub>). Z drugiej strony, jeśli natura relacji  $R$  nie wyklucza tego, że indywiduum, które było jej drugim członem, może być jej pierwszym członem, to wzorzec egzemplifikacji (P<sub>1</sub>) również jest dopuszczalny. Ilustracją dopuszczalności obu wzorców egzemplifikacji było wspomniane wyżej podstawienie za  $R$  relacji bycia przełożonym, za  $F$  własności bycia łysym, za  $G$  — wysokim; za  $H$  — niskim. Niemniej jednak relacja  $R$  mogłaby mieć taką naturę, że wzorzec (P<sub>1</sub>) byłby wykluczony. Przykładowo, gdyby relacją  $R$  była relacja *bycia mężem*, (CP) nie mogłoby być egzemplifikowane na sposób (P<sub>1</sub>). Drugi człon pierwszego wystąpienia tej relacji, będący kobietą, musiałby bowiem być pierwszym członem jej drugiego wystąpienia. Jednak kobieta nie może być niczym mężem.

W związku z powyższym rolą dyskursu możliwościowego wydaje się ujawnianie sposobów, na jakie fakty modalne mogą być egzemplifikowane. Rola ta nieodzownie wiąże się z możliwością wyprowadzania z twierdzeń możliwościowych twierdzeń dotyczących sposobów, na jakie fakty modalne mogłyby być egzemplifikowane. Stanowi to podstawę do wprowadzenia następującej odwróconej reguły generowania:

(conv-RG-F) dla każdego faktu modalnego  $x$ , jeśli istnienie  $x$  stanowi konieczną i wystarczającą podstawę do wyobrażania sobie, że istnieją światy możliwe, w których istnieją indywidua posiadające własności lub pozostające do siebie w relacjach, stanowiących elementy konstytutywne faktów modalnych, i w światach tych elementy konstytutywne  $x$ -a są egzemplifikowane na odmienne sposoby  $S_1, \dots, S_n$ , to twierdzenie modalne, że  $x$  mogłoby być egzemplifikowane na sposoby  $S_1, \dots, S_n$ , jest prawdziwe

oraz odwróconego prawa pomostowego:

(conv-PP-F) dla każdego faktu modalnego  $x$ , jeśli istnienie  $x$  stanowi konieczną i wystarczającą podstawę prawdziwości (zgodnie z fikcją

*PW*) twierdzenia, że istnieją światy możliwe, w których istnieją indywidua posiadające własności lub pozostające do siebie w relacjach stanowiących elementy konstytutywne faktów modalnych, i opisy tych światów różnią się ze względu na opis sposobów  $S_1, \dots, S_n$ , na jakie egzemplifikowane są elementy konstytutywne  $x$ -a, to twierdzenie modalne, że  $x$  mogłoby być egzemplifikowane na sposoby  $S_1, \dots, S_n$ , jest prawdziwe.

Na podstawie przeprowadzonych wyżej rozważań można zrewidować wprowadzoną w poprzednim paragrafie koncepcję faktów modalnych. Chcąc bowiem mieć koncepcję faktów modalnych, która nie prowadzi do utożsamiania intuicyjnie odmiennych możliwości, fakty modalne należy utożsamić nie ze zbiorami współwystępujących własności, ale z faktem posiadania przez takie zbiory cechy bycia możliwe egzemplifikowanym na taki a taki sposób. Wówczas twierdzenie „Możliwe, że istnieje takie  $x$ , że jest  $F$ ” będzie prawdziwe wtw, gdy istnieje takie  $x$ , że  $x = \langle S, (C(\{F\})) \rangle$ .  $S$  jest własnością drugiego rzędu bycia możliwe egzemplifikowanym na taki a taki sposób. Twierdzenie „Możliwe, że istnieje takie  $x$  i  $y$ , że  $xRy$ ” jest prawdziwe wtw, gdy istnieje takie  $x$ , że  $x = \langle S, (C(\{R, \langle C(X_1), C(X_2) \rangle\})) \rangle$ , gdzie  $C(X_1)$  i  $C(X_2)$  stanowią zbiory własności, za których pomocą reprezentowane są  $x$  i  $y$ .

Zaproponowana wyżej koncepcja faktów modalnych wydaje się spełniać dezyderaty (D<sub>1</sub>) i (D<sub>2</sub>). W pierwszym wypadku wprowadzenie wzorców egzemplifikacji pozwala na znalezienie bytów fundujących prawdziwość takiego twierdzenia modalnego jak „Mogłyby istnieć jakościowo nieodróżnialne krążące wokół siebie elektrony”. Bytem fundującym prawdziwość tego twierdzenia byłby — upraszczając — fakt modalny  $\langle S, (C(X)) \rangle$ , gdzie  $X$  jest zbiorem własności wewnętrznych i zewnętrznych przysługujących tym elektronom, a  $S$  własnością bycia możliwe egzemplifikowanym dwukrotnie. Jak widzieliśmy w §4.4., w ramach ontologii Plantingi elektrony byłyby utożsamiane z in-

dywidualnymi istotami, które identyfikowane są ze zbiorami własności takimi, że dla dowolnego maksymalnego stanu rzeczy, relatywnie do tego stanu rzeczy dana własność należy do tego zbioru lub jest z niego wykluczona. Gdy utożsamia się indywidua ze zbiorami własności, nieuniknioną konsekwencją tego utożsamienia jest konieczność uznania za identyczne wszystkich bytów, które są jakościowo nieodróżnialne. Blackowskim krążącym wokół siebie elektronom przysługiwałyby te same własności, a w związku z tym byłyby one identyczne w ramach ontologii proponowanej przez Plantingę. Zachodzi więc różnica pomiędzy koncepcją światów możliwych proponowaną przez umiarkowanych realistów a wprowadzoną wyżej koncepcją faktów modalnych.

### 6.3 Problem ze wzorcami egzemplifikacji

Jedna kwestia wydaje się szczególnie problematyczna w tym miejscu. Czy da się powiedzieć coś o własności (a), tj. własności bycia możliwie egzemplifikowanym przez wiele indywiduów, która odgrywa istotną rolę przy wyznaczaniu wzorców egzemplifikacji danego zbioru współwystępujących uniwersaliów, bez kwantyfikowania po możliwych indywiduach?

Konieczność kwantyfikowania po możliwych indywiduach wydaje się nieunikniona, gdy chcemy dopuścić prawdziwość takich twierdzeń modalnych jak twierdzenie, że mogłoby istnieć więcej indywiduów, niż aktualnie istnieje. W ramach powyższego modelu prawdziwość tego twierdzenia byłaby reprezentowana za pomocą posiadania przez  $C(\{F\})$  (gdzie  $F$  jest własnością wspólną wszystkim aktualnie istniejącym indywiduom) takiego wzorca egzemplifikacji  $S$ , że własność  $F$  mogłaby być egzemplifikowana przez indywidua  $n$  razy i  $n$  jest większe od liczby egzemplifikacji tej własności w świecie aktualnym. Mówienie o takim wzorcu egzemplifikacji wymaga jednak odwoływania się do indywiduów, które aktualnie nie istnieją. Jeśli o własności (a) nie da się powiedzieć niczego bez kwantyfikowania po możliwych indywidu-

ach, to fikcjonalista modalny powinien standardowo uznać, że nie da się jej opisać bez angażowania się w grę *GPW* bądź bez mówienia o treści fikcji *PW*. Wówczas jednak powstaje następujący problem: fikcjonalista modalny chciałby twierdzić, że z twierdzeń możliwościowych, w których opisuje się sposoby egzemplifikacji własności i relacji przez indywidua istniejące w światach możliwych, można wyprowadzić literalnie prawdziwe twierdzenia dotyczące własności bycia możliwie egzemplifikowanym na sposób *S* przysługujących zbiorom współwystępujących własności. By jednak fikcjonalista modalny mógł uznać twierdzenia dotyczące tej własności za literalnie prawdziwe, własność ta musi być ogólna (tj. musi istnieć niezależnie od indywiduów). W związku z tym wnioskowanie o przysługiwaniu takich własności zbiorom współwystępujących uniwersaliów na podstawie fikcyjnej prawdziwości twierdzeń możliwościowych musi opierać się na założeniu, że własności (a) odpowiada po stronie rzeczywistości jakaś własność ogólna. Dopóki jednak fikcjonalista modalny nie powie niczego literalnie prawdziwego o naturze tej własności, założenie, że jest ona własnością ogólną, może być kwestionowane. Realista modalny mógłby twierdzić, że opis cechy (a) w kategoriach posiadania własności przez możliwe indywidua powinien być odczytywany dosłownie jako pociągający za sobą zobowiązania ontologiczne co do istnienia możliwych indywiduów i że oto właśnie fikcjonalista modalny natrafił na miejsce, gdzie będzie musiał zaakceptować istnienie tych bytów. Realista modalny mógłby dodać, że niewiarygodna jest teza fikcjonalisty modalnego, że własność bycia możliwie egzemplifikowanym jest własnością ogólną, gdyż wydaje się ona własnością relacyjną przysługującą zbiorom współwystępujących uniwersaliów na mocy ich relacji do możliwych indywiduów. W związku z tym wszystko wskazuje na to, że w jej charakterystykę muszą być uwikłane odniesienia do indywiduów.

W kontekście sporu z realistą modalnym fikcjonalista modalny

powinien zatem wykazać, że własność (a) nie należy do klasy bytów, które pociągają za sobą istnienie możliwych indywiduów. Jakiś czas temu w literaturze pojawiło się stanowisko zwane *generalizmem* (zob. Dasgupta 2009). Zgodnie z nim indywidua nie istnieją. W ich miejsce przyjmuje się istnienie funkcji z własności we własności lub z własności w stany rzeczy. Językiem koncepcji generalisty jest język logiki funktorów autorstwa Stanisława Leśniewskiego. Z formalnego punktu widzenia na logikę tę składa się sześć nieredukowalnych do siebie funkcji, które przyjmują własności za swoje argumenty i wspólnie pozwalają na parafrazę dowolnego twierdzenia klasycznego rachunku predykatów (bez stałych indywiduowych). Niech  $P^n$  będzie  $n$ -argumentową własnością, wówczas sześć wspomnianych funkcji przedstawia się następująco:

- 1) *Derelatywizacja*:  $\text{Der } Px_1, \dots, x_{n-1}$  wtw, gdy istnieje takie  $x_n$ , że  $Px_1, \dots, x_n$ .
- 2) *Duża inwersja*:  $\text{Inv } Px_1, \dots, x_n$  wtw, gdy  $Px_n, x_1, \dots, x_{n-1}$ .
- 3) *Mała inwersja*:  $\text{inv } Px_1, \dots, x_n$  wtw, gdy  $Px_1, \dots, x_{n-2}, x_n, x_{n-1}$ .
- 4) *Zwrotność*:  $\text{Ref } Px_1, \dots, x_{n-1}$  wtw, gdy  $Px_1, \dots, x_{n-1}, x_{n-1}$ .
- 5) *Negacja*:  $\text{Neg } Px_1, \dots, x_n$  wtw, gdy  $\neg Px_1, \dots, x_n$ .
- 6) *Iloczyn kartezjański*:  $P \times Qx_1, \dots, x_m, y_1, \dots, y_n$  wtw, gdy  $Px_1, \dots, x_m$  i  $Qy_1, \dots, y_n$  (por. Quine 1960, s. 344).

Skoro język logiki funktorów pozwala na przekład dowolnego twierdzenia sformułowanego w klasycznym rachunku predykatów, z pewnością pozwala na wyrażenie twierdzenia, iż zbiór uniwersaliów jest egzemplifikowany wielokrotnie. Aby uniknąć brnięcia w skomplikowane i długie formuły, założmy, że chcemy powiedzieć, iż własność  $F$  jest egzemplifikowana dwukrotnie. W języku klasycznego rachunku predykatów (bez stałych indywiduowych) wyrazilibyśmy to pomocą

twierdzenia: „Istnieje takie  $x$  i  $y$ , że  $Fx$  i  $Fy$  i  $x \neq y$ ”, któremu odpowiada następująca formuła w języku logiki funktorów:

$$\text{Der}(\text{Ref}(\text{Der}(\text{Ref}(\text{Inv}(\text{inv}(F \times F \times \text{Neg}I))))))^{46},$$

gdzie  $I$  jest relacją identyczności. Z uwagi na to, że fikcjonalista modalny mówiąc o własności (a), przyjmuje twierdzenie, wyrażalne w ramach języka logiki modalnej z kwantyfikatorami, o postaci „Możliwie, istnieje takie  $x$  i  $y$ , że  $Fx$  i  $Fy$  i  $x \neq y$ ”, można je przełożyć na twierdzenie:

$$C(\{\text{Der}(\text{Ref}(\text{Der}(\text{Ref}(\text{Inv}(\text{inv}(F \times F \times \text{Neg}I))))))\}),$$

w którym stwierdza się, że wartość funkcji  $\text{Der}$  współwystępuje.

Wiemy zatem, w jaki sposób można parafrazować twierdzenia klasycznego rachunku predykatów bez nazw i w jaki sposób od strony formalnej zachowują się nowo wprowadzone funkcje. Niewiele natomiast wiemy o tym, jakie jest wyjaśnienie tego, że funkcje te zachowują się w ten sposób. W szczególności nie wiemy, czy wyjaśnienie ich funkcjonowania ostatecznie nie będzie wymagało odwołania się do możliwych indywiduów. Wzajemną parafrazowalność twierdzeń klasycznego rachunku predykatów bez nazw i twierdzeń wyrażonych w języku logiki

---

<sup>46</sup>Przekład dokonuje się za pośrednictwem następujących kroków:

1. Istnieje takie  $x$  i  $y$ , że  $Fx$  i  $Fy$  i  $x \neq y$ .
2. Istnieje takie  $x$  i  $y$ , że  $Fx$  i  $Fy$  i  $\text{Neg } Ixy$ .
3. Istnieje takie  $x$  i  $y$ , że  $F \times F \times \text{Neg}Ixyxy$ .
4. Istnieje takie  $x$  i  $y$ , że  $\text{inv}(F \times F \times \text{Neg}I)xyyx$ .
5. Istnieje takie  $x$  i  $y$ , że  $\text{Inv}(\text{inv}(F \times F \times \text{Neg}I))xyxy$ .
6. Istnieje takie  $x$  i  $y$ , że  $\text{Ref}(\text{Inv}(\text{inv}(F \times F \times \text{Neg}I)))xyxy$ .
7. Istnieje takie  $x$ , że  $\text{Der}(\text{Ref}(\text{Inv}(\text{inv}(F \times F \times \text{Neg}I))))xx$ .
8. Istnieje takie  $x$ , że  $\text{Ref}(\text{Der}(\text{Ref}(\text{Inv}(\text{inv}(F \times F \times \text{Neg}I))))x$ .
9.  $\text{Der}(\text{Ref}(\text{Der}(\text{Ref}(\text{Inv}(\text{inv}(F \times F \times \text{Neg}I))))))$ .

funktorów można interpretować w taki sposób, że prawe strony równoważności, za których pomocą charakteryzowano te funkcje, można uznać za pierwotniejsze w porządku wyjaśniania. Ktoś, kto akceptuje istnienie indywiduów i własności, mógłby twierdzić, że funkcja *Der* jest relacją zachodzącą pomiędzy parami uporządkowanymi, których elementami są *n*-argumentowa własność oraz *n*-tka uporządkowana indywiduów, a takimi stanami rzeczy, że *n*-tka indywiduów realizuje taką a taką własność. Gdy jednak nie chcemy opisywać funkcjonowania funkcji *Der* w takich kategoriach, musimy szukać jakiegoś innego wyjaśnienia jej funkcjonowania. Co w tej sprawie ma do powiedzenia generalista? Oto co Shamik Dasgupta pisze o funkcji *Der*:

Jeśli  $L^2$  jest dwuargumentową własnością kochania, którą normalnie rozumiemy jako zachodzącą pomiędzy indywiduami *x* i *y* wtw, gdy *x* kocha *y*, to *Der*  $L^2$  jest jednoargumentową własnością bycia kochanym przez kogoś, którą normalnie rozumiemy jako egzemplifikowaną przez indywiduum *y* wtw, gdy ktoś kocha *y*. Upraszczając, *Der* częściowo „wypełnia” własność poprzez ustanowienie, jak normalnie mówimy, że coś egzemplifikuje jej pierwszą pozycję. [...] Gdy normalnie opisujemy sytuację jako będącą taką, że ktoś kocha kogoś, w jaki sposób opisze ją generalista? By odpowiedzieć na to, rozważmy jednoargumentową własność *Der*  $L^2$  opisaną przed chwilą i rozważmy rezultat kolejnego zastosowania *Der* celem otrzymania *Der* *Der*  $L^2$ . Jest to zero-argumentowa „własność”, która normalnie może być rozumiana jako stan rzeczy, mianowicie stan kochania kogoś przez kogoś. Teraz „*Der* *Der*  $L^2$ ” jest termem i jako taki nie stwierdza niczego o tym, jaki świat jest. By stwierdzić to, generalista wprowadza pierwotny predykat **x zachodzi**, który stosuje się do dowolnego zero-argumentowego termu w celu otrzymania zdania. Zatem,



gdy normalnie mówimy, że ktoś kocha kogoś, generalista powie, że  $\text{Der Der } L^2$  **zachodzi** (Dasgupta 2009, s. 53).

Cieężko dopatrzeć się w tym cytacie jakiegoś wyjaśnienia funkcjonowania  $\text{Der}$ . Nie otrzymujemy żadnych informacji poza opisem zachowania funkcji  $\text{Der}$  za pomocą metafory egzemplifikowania własności przez indywidua. Dowiadujemy się, że  $\text{Der Der } L^2$  jest termem odnoszącym się do pewnego stanu rzeczy, który normalnie opisalibyśmy jako stan taki, że jedno indywiduum kocha drugie, ale jest to opis jedynie metaforyczny. Zwykle stany rzeczy są utożsamiane z ustrukturyzowanymi bytami posiadającymi elementy konstytutywne. Gdybyśmy tutaj przyjęli, że własność  $L^2$  jest elementem konstytutywnym stanu rzeczy  $\text{Der Der } L^2$ , to automatycznie powstaje pytanie, czym są pozostałe elementy konstytutywne takiego stanu rzeczy? Nie mogą nimi być funkcje  $\text{Der}$ , gdyż z zasady są relacjami zachodzącymi pomiędzy daną własnością a pewną „blisko powiązaną” własnością o mniejszej liczbie argumentów, a nie elementami konstytutywnymi członów tych relacji. Być może stan rzeczy  $\text{Der Der } L^2$  powinien być uznany za byt pierwotny, który za pośrednictwem funkcji  $\text{Der}$  pozostaje w relacji do pierwotnej jednoargumentowej własności  $\text{Der } L^2$ ? Być może ku takiemu stanowisku skłania się generalista.

Nie sądzę jednak, by ku stanowisku temu mógł skłaniać się fikcjonalista modalny. Gdyby bowiem uznał własność bycia możliwie egzemplifikowanym przez  $n$  indywiduów za własność pierwotną, natrafiłby na co najmniej dwa problemy. Po pierwsze, im więcej mamy bytów pierwotnych w naszej ontologii, tym trudniej o informatywne wyjaśnienie, na mocy czego zachodzą relacje wiążące te byty. W rezultacie tym trudniej byłoby rzucić nieco więcej światła na naturę własności bycia możliwie egzemplifikowanym przez  $n$  indywiduów, a jest to potrzebne umiarkowanemu fikcjonalistcie modalnemu do odparcia zagrożenia ze strony realisty modalnego. Po drugie, w odróżnieniu od generalisty fikcjonalista modalny przyjmuje istnienie aktualnych

indywidualiów oraz przyjmuje, że indywidua te posiadają własności. Powinien więc zgodzić się na to, że własność bycia kochanym przez kogoś, którą posiada jakieś aktualnie istniejące indywiduum, powinna być analizowana w kategoriach zachodzenia relacji kochania pomiędzy tym indywiduum a pewnym innym indywiduum. Podobnie stan rzeczy, że  $x$  kocha  $y$ , gdzie  $x$  i  $y$  są aktualnie istniejącymi indywiduami, powinien być traktowany nie jako byt pierwotny, ale jako ustrukturyzowany byt, którego elementami konstytutywnymi są  $x$ ,  $y$  oraz relacja kochania. W przypadku możliwych indywiduów fikcjonalista modalny musiałby zupełnie zmienić swoją koncepcję stanów rzeczy. Byłyby one wartością pierwszego wystąpienia funkcji  $\text{Der}$  w  $C(\{\text{Der}(\text{Ref}(\text{Der}(\text{Ref}(\text{Inv}(\text{inv}(F \times F \times \text{Neg}I))))))\})$ , jednak elementami konstytutywnymi tego stanu rzeczy nie byłyby już możliwe indywidua. W jaki sposób wyjaśnić tę rozbieżność?

Dostarczenie przekonującej odpowiedzi na to pytanie jest o tyle trudne, że przysługiwanie własności bycia możliwie egzemplifikowanym przez dwa indywidua wydaje się z łatwością poddawać analizie w kategoriach istnienia stanu rzeczy, którego elementami są własność i dwa możliwe indywidua. Jeśli powodem, by twierdzić, że jej przysługiwanie nie podlega takiej analizie, jest to, że w przeciwnym razie jej istnienie będzie pociągało za sobą istnienie możliwych indywiduów, to w kontekście dyskusji z realistą modalnym przyjęcie takiego stanowiska jest odmową podjęcia jego zarzutu. Umiarkowany fikcjonalista modalny musiałby tutaj upierać się, że to, co wydaje się analizowalne, analizowalnym nie jest, a to, co może rodzić problemy dla fikcjonalistycznego programu unikania zobowiązań co do możliwych indywiduów, jest ontologicznie niewinne.

Własność bycia możliwie egzemplifikowanym przez wiele indywiduów sprawia, że powodzenie proponowanego przez umiarkowanego fikcjonalistę modalnego projektu uniknięcia zobowiązań ontologicznych co do istnienia światów możliwych i możliwych indywiduów

może być przedmiotem zasadnych wątpliwości. Być może umiarkowany fikcjonalista modalny jest w stanie zaproponować jakąś odmienną koncepcję faktów modalnych, która nie byłaby źródłem podobnych problemów, choć mnie nie przychodzi na myśl żadna koncepcja, która nie dziedziczyłaby trudności omawianych w tym paragrafie. Gdyby nawet był on w stanie zaproponować taką koncepcję faktów modalnych, która nie dziedziczyłaby tej trudności, musiałaby ona jednak zawierać w sobie redukcyjną analizę własności bycia możliwie egzemplifikowanym przez wiele indywiduów. Wówczas powstaje pytanie: dlaczego nie wykorzystać tej redukcyjnej analizy do utworzenia redukcyjnej koncepcji światów możliwych i zastąpienia nią umiarkowanego fikcjonalizmu modalnego? Wydaje się bowiem, że jeśli można zaproponować drobnoziarnistą redukcyjną koncepcję światów możliwych, pozwalającą na oddanie różnic pomiędzy wszystkimi intuicyjnie odmiennymi możliwościami, to raczej nic nie stoi na przeszkodzie temu, by zarzucić fikcjonalizm modalny na rzecz takiego redukcjonistycznego stanowiska. Niewątpliwą zaletą takiego stanowiska byłoby to, że nie wymagałoby ono wprowadzania praw pomostowych czy reguł generowania, które komplikują definiowanie warunków fikcyjnej prawdziwości twierdzeń możliwościowatych. W odróżnieniu od umiarkowanego fikcjonalizmu modalnego redukcyjne stanowisko unikałoby komplikacji z tytułu wprowadzenia funktora opowieści. W odróżnieniu od przedmiotowego fikcjonalizmu modalnego stanowisko to byłoby teoretycznie oszczędniejsze, gdyż nie wymagałoby wprowadzenia rozróżnienia na fikcyjną i literalną prawdziwość, pozostając wyłącznie przy ostatniej. W ramach tego stanowiska mówienie o światach możliwych i możliwych indywiduach byłoby po prostu mówieniem o bytach, do których zredukowano Lewisowskie światy możliwe i możliwe indywidua.

# Rozdział V

## Stanowiska konkurencyjne względem fikcjonalizmu modalnego

W poprzednim rozdziale starałem się przybliżyć idee konstytutywne dla fikcjonalistycznych interpretacji dyskursu możliwościowego. Fikcjonalizm modalny nie jest jednak jedynym stanowiskiem, które stawia sobie za cel zachowanie możliwości posługiwania się tym dyskursem bez akceptowania zobowiązań ontologicznych co do istnienia światów możliwych i możliwych indywiduów. W tym rozdziale zamierzam przyjrzeć się dwóm stanowiskom konkurencyjnym względem fikcjonalizmu modalnego: instrumentalizmowi (§5.1.) oraz modalizmowi (§5.2.).

### 1 Instrumentalizm *versus* fikcjonalizm

Wspólną cechą zarówno stanowisk konkurencyjnych względem fikcjonalizmu modalnego, o których będę mówił poniżej, jak i samego fikcjonalizmu modalnego jest próba poszukiwania takiej interpretacji twierdzeń możliwościowych, która spełniałaby trzy warunki:

- (a) twierdzenia możliwościowe można uznać za twierdzenia w *pewnym sensie* prawdziwe, dzięki czemu mogą pełnić one rolę przesłanek w rozumowaniach;
- (b) wbrew realizmowi modalnemu twierdzenia możliwościowe nie pociągają za sobą istnienia światów możliwych i możliwych indywiduów;
- (c) wbrew redukcjonizmowi twierdzenia możliwościowe nie pociągają za sobą istnienia substytutów światów możliwych i możliwych indywiduów (takich jak indywidualne istoty, maksymalne możliwe stany rzeczy itd.).

Jedną z prób spełnienia warunków (a)–(c) jest pewna wersja *instrumentalizmu*, proponowana przez Matti Eklunda w artykule (2005). Zgodnie z tym stanowiskiem w większości kontekstów konwersacyjnych, które nie dotyczą ontologii, użytkownicy języka pozostają obojętni na zobowiązania ontologiczne co do bytów, których zdają się dotyczyć twierdzenia wygłaszane w tych kontekstach. Użytkownicy dyskursu D, wypowiadając twierdzenia egzystencjalne, w większości wypadków nie traktują wszystkich niesionych przez nie informacji z równą powagą. Biorą na serio jedynie te, które są centralne z punktu widzenia celów komunikacyjnych, jakie stawiają sobie w danym kontekście (zob. Eklund 2005, s. 563–564). W przekonaniu Eklunda poprawnością użycia twierdzeń dyskursu D w takich kontekstach rządzą dwie zasady:

- (I<sub>1</sub>) twierdzenia dyskursu D są instrumentalnie użyteczne w danym kontekście konwersacyjnym. W ramach potocznych kontekstów konwersacyjnych, jak i w rozważaniach teoretycznych użytkownicy języka posługują się najwygodniejszymi środkami odpowiednimi do zakomunikowania danej informacji. Zobowiązania ontologiczne implikowane przez te twierdzenia mają drugorzędne znaczenie (Eklund 2005, s. 566);

(I<sub>2</sub>) istnieje taki sposób rozumienia twierdzeń dyskursu D wygłoszonych w danym kontekście, że pewna ich interpretacja jest literalnie prawdziwa (Eklund 2005, s. 573–574).

O twierdzeniu  $p$  dyskursu D powiemy, że jest w *pewnym sensie* prawdziwe w kontekście  $C$  wtw, gdy:

- (1) twierdzenie  $p$  jest instrumentalnie użyteczne ze względu na cele konwersacyjne bądź teoretyczne, jakie stawia się w kontekście  $C$ ;
- (2) istnieje takie literalnie prawdziwe twierdzenie  $q$  dyskursu  $D^*$ , że stanowi ono interpretację twierdzenia  $p$  w kontekście  $C$ .

Twierdzenie  $q$  jest literalnie prawdziwe, gdy w dziedzinie dyskursu  $D^*$  *realnie* istnieją przedmioty fundujące prawdziwość tego twierdzenia.

Należy zauważyć, iż o warunkach prawdziwości twierdzeń dyskursu D w ich instrumentalistycznej interpretacji można mówić jedynie we wtórnym sensie, tj. ich warunki prawdziwości nie są bezpośrednio wyznaczane przez ich treść, ale przez warunki prawdziwości ich literalnie prawdziwych interpretacji. Z założenia, same twierdzenia dyskursu D nie mogą tutaj być swoimi literalnie prawdziwymi interpretacjami, gdyż przekreślałoby to jakiekolwiek szanse na uniknięcie zobowiązań co do bytów klasy  $K$ , o których z pozoru mówią twierdzenia dyskursu D. Mówienie o warunkach prawdziwości twierdzeń dyskursu D w tym wtórnym sensie jest — jak ufam — spójne z tezą, że twierdzenia te same z siebie nie stanowią autentycznych reprezentacji tego, jak rzeczy się mają, za której pomocą charakteryzowałem instrumentalizm w rozdziale 2. Z kolei sama potrzeba wskazania literalnie prawdziwych interpretacji dla twierdzeń pewnego instrumentalnie użytecznego dyskursu D bierze się tutaj — jak sądzę — z faktu, że twierdzenia dyskursu D wykorzystywane są jako przesłanki w rozumowaniach

wychodzących od literalnie prawdziwych przesłanek i dochodzących do literalnie prawdziwych wniosków. Jeśli rozumowania takie w ogóle mają być uprawnione, to w jakiś sposób warunki prawdziwości (we wtórnym sensie) twierdzeń dyskursu D muszą wiązać się z warunkami prawdziwości twierdzeń stanowiących literalnie prawdziwe przesłanki oraz wnioski takich rozumowań. W ramach instrumentalistycznej propozycji związku pomiędzy wartościami logicznymi literalnie prawdziwych twierdzeń tworzących przesłanki i wnioski rozumowań a twierdzeniami dyskursu D mają zachodzić na mocy tego, jakie są warunki prawdziwości literalnie prawdziwych interpretacji twierdzeń dyskursu D.

Wracając do dyskursu możliwościowego, stanowisko Eklunda<sup>47</sup> zdaje się przybierać następującą postać: gdy w centrum zainteresowania nie ma metafizycznego sporu o istnienie światów możliwych, lecz są inne (niezależne od tego sporu) problemy, do których analizy dyskurs możliwościowy jest użytecznym narzędziem, użytkownicy tego dyskursu pozostają obojętni na zobowiązania ontologiczne implikowane przez jego twierdzenia. Interesują się natomiast ich użytecznością ze względu na cele, które sobie postawili w danym kontekście. Załóżmy, że w danym kontekście centralnym przedmiotem zainteresowania jest logiczna poprawność pewnego rozumowania zbudowanego przy użyciu potocznych pojęć modalnych. Może się okazać, że moce ekspresywne języka logiki modalnej  $S5$  z kwantyfikatorami (w skrócie:  $L_{S5}$ ) są zbyt ubogie, by w ogóle można było w nim sformułować wniosek tego rozumowania. Z kolei język możliwościowy jest na tyle bogaty, że z łatwością można sparafrazować w nim zarówno przesłanki, jak i wniosek tego rozumowania, jak również przetestować jego logiczną poprawność przy użyciu metod właściwych klasycznemu rachunkowi

---

<sup>47</sup>Sam Eklund nie rozważa zastosowania swojego stanowiska do dyskursu możliwościowego, niemniej jednak ogólny charakter, jaki zdaje się on przypisywać swojemu instrumentalizmowi, powinien pozwolić na zastosowanie go do tego dyskursu.

predykatów pierwszego rzędu. W związku z tym moce ekspresywne wraz z metodami sprawdzania logicznej poprawności rozumowań będą czynnikami konstytuującymi instrumentalną użyteczność dyskursu możliwościowego w kontekście analizy poprawności rozumowań sformułowanych przy użyciu potocznych pojęć modalnych. Przykład rozumowania, którego poprawności nie da się udowodnić w  $L_{S5}$ , a którego poprawność można wykazać, korzystając z dyskursu możliwościowego, przedstawia się następująco:

(P<sub>1</sub>) Twój samochód jest czarny.

(P<sub>2</sub>) Mój samochód mógłby być czarny.

(W) Zatem mój samochód mógłby być tego samego koloru, który ma aktualnie twój samochód.

Choć przesłanki tego rozumowania można wyrazić w języku  $L_{S5}$  jako odpowiednio:  $Bb$  (gdzie „ $B$ ” jest równoznaczne z predykatem „jest czarny”, a „ $b$ ” desygnuje twój samochód) i  $\Diamond Ba$  (gdzie „ $a$ ” desygnuje mój samochód), to już nie da się w nim wyrazić wniosku tego rozumowania. Nie pomoże tutaj nawet wprowadzenie do języka  $L_{S5}$  funktora aktualności  $\mathbf{A}$  przyjmującego za argumenty poprawnie zbudowane twierdzenia języka  $L_{S5}$ . Niech jego zachowaniem rządzi następująca reguła: „ $\mathbf{A}p$ ” jest prawdziwe w świecie  $w$  wtw, gdy  $p$  jest prawdziwe w świecie aktualnym. Treści wniosku (W) z pewnością nie oddaje twierdzenie „ $\mathbf{A}Bb$  i  $\Diamond Rab$ ” (gdzie „ $R$ ” jest równoznaczne z predykatem „jest tego samego koloru, co”), ponieważ jest ono prawdziwe nawet wówczas, gdy  $b$  jest czarne w świecie aktualnym, a oba auta posiadają ten sam kolor — powiedzmy czerwony — w świecie, do którego odsyła funktor modalny możliwości.

Gdy zaakceptuje się dyskurs możliwościowy, przedstawienie poprawności tego rozumowania nie stanowi najmniejszego problemu<sup>48</sup>:

<sup>48</sup>Pozwoliłem sobie uprościć możliwościową parafrazę i przekładać twierdzenia modalne w taki sposób, jak gdyby istniały transświatowe indywidua.



(P<sub>1</sub>) Twój samochód jest czarny.

(P<sub>2</sub>) Mój samochód mógłby być czarny.

(Z<sub>1</sub>)  $W@ \& Cb@$  (możliwoświatowa interpretacja twierdzenia P<sub>1</sub>)

(Z<sub>2</sub>)  $\exists w(Ww \& Caw)$  (interpretacja twierdzenia P<sub>2</sub>)

(Z<sub>3</sub>)  $\forall w, v \forall x, y (Ww \& Wv \rightarrow Rxyv \equiv Cxw \& Cyv)$  (założenie)

(Z<sub>4</sub>)  $Wc \& Cac$  (opuszczanie kwantyfikatora: Z<sub>2</sub>)

(Z<sub>5</sub>)  $W@$  (OK: Z<sub>1</sub>)

(Z<sub>6</sub>)  $Wc$  (OK.: Z<sub>4</sub>)

(Z<sub>7</sub>)  $Wc \& W@$  (DK: Z<sub>5</sub>, Z<sub>6</sub>)

(Z<sub>8</sub>)  $Wc \& W@ \rightarrow Racb@ \equiv Cac \& Cb@$  (opuszczanie kwantyfikatora:  
Z<sub>3</sub>)

(Z<sub>9</sub>)  $Racb@ \equiv Cac \& Cb@$  (RO: Z<sub>8</sub>, Z<sub>7</sub>)

(Z<sub>10</sub>)  $Cac \& Cb@ \rightarrow Racb@$  (OR: Z<sub>9</sub>)

(Z<sub>11</sub>)  $Cac$  (OK: Z<sub>4</sub>)

(Z<sub>12</sub>)  $Cb@$  (OK: Z<sub>1</sub>)

(Z<sub>13</sub>)  $Cac \& Cb@$  (DK: Z<sub>11</sub>, Z<sub>12</sub>)

(Z<sub>14</sub>)  $Racb@$  (RO: Z<sub>10</sub>, Z<sub>13</sub>)

(Z<sub>15</sub>)  $\exists w(Rawb@)$  (wprowadzanie kwantyfikatora: Z<sub>14</sub>)

(W) Zatem mój samochód mógłby być tego samego koloru, który  
twój samochód ma aktualnie.

„*W*” jest predykatem „jest światem”, „@” — stałą desygnującą świat aktualny, wyrażenie „*Cxw*” — „*x* jest czarny w świecie *w*”, wyrażenie „*Rxwyv*” — „*x* jest tego samego koloru w świecie *w*, co *y* jest w świecie *v*”, „*a*” — stałą indywiduową desygnującą mój samochód, „*b*” — stałą indywiduową desygnującą twój samochód.

Zgodnie z dyskutowaną wersją instrumentalizmu, w kontekstach, w których stosuje się możliwościowe analizy, użytkownik dyskursu możliwościowego będzie opierał się na przeświadczeniu, iż istnieje taka literalnie prawdziwa interpretacja twierdzeń dotyczących światów możliwych i możliwych indywiduów, która nie pociąga za sobą zobowiązań ontologicznych co do tych bytów. Eklund zastrzega, że literalnie prawdziwa interpretacja twierdzeń dyskursu możliwościowego nie musi być dokładnie tym, co wypowiadający dane twierdzenie możliwościowego miał na myśli w danym kontekście. Nie musi również być taka, że użytkownik dyskursu możliwościowego będzie w stanie podać odpowiednią interpretację tego twierdzenia. Nawet nie musi obowiązywać we wszystkich kontekstach, w których używa się dyskursu możliwościowego. Za ostatnim może przemawiać fakt, że dyskurs możliwościowy znajduje zastosowanie w wielu niesprowadzalnych do siebie dziedzinach, np. w analizach modalności i powinności.

Chcąc nadać treść zasadzie ( $I_2$ ), powinniśmy poszukiwać takiej literalnie prawdziwej interpretacji twierdzeń możliwościowych wygłaszanych w kontekstach, w których mówi się o modalnościach, która nie pociąga za sobą ani realizmu, ani redukcjonizmu w sprawie istnienia światów możliwych. Jaką literalnie prawdziwą interpretację twierdzeń dyskursu możliwościowego może zaproponować instrumentalista? Jako że instrumentalizm stanowi alternatywę względem fikcjonalizmu modalnego, wydaje się, że tylko dwa stanowiska wchodzą tutaj w grę.

Pierwszym z nich jest modalizm<sup>49</sup>. Jego zwolennicy głoszą, że twierdzenia kwantyfikujące po światach możliwych można zredukować do takich twierdzeń kwantyfikujących po światach, w których kwantyfikator egzystencjalny wiążący zmienne przebiegające po światach znajduje się w zasięgu funktora modalnego możliwości. Upraszczając, proponuje się, by twierdzenie „Istnieje świat możliwy  $w$ , w którym  $Fx$ ” zredukować do twierdzenia „Możliwie, istnieje świat aktualny  $w$ , w którym  $Fx$ ”. Pierwsze z tych twierdzeń, jeśli zostanie odczytane literalnie, zdaje się pociągać za sobą istnienie światów możliwych. Jednak drugie z nich może być prawdziwe, nawet gdy nie istnieją światy możliwe.

Drugim stanowiskiem jest uznanie, że literalnie prawdziwą interpretacją twierdzeń możliwościświatowych są odpowiednie twierdzenia modalne, których prawdziwość ufundowana jest na istnieniu odpowiednich faktów modalnych. Jak pokazuje przypadek twierdzenia modalnego (W), najprawdopodobniej będziemy mieli tutaj do czynienia z twierdzeniami modalnymi, które da się wyrazić w języku  $L_{S5}$ . Jest to o tyle niepokojące, że nie bardzo wiadomo, jakie reguły rządzą zachowaniem potocznych pojęć modalnych, jeśli nie podlegają one wyłącznie regułom logiki modalnej S5. Co więcej, uznanie, że prawdziwość twierdzeń możliwościświatowych ufundowana jest na prawdziwości odpowiednich twierdzeń modalnych, wymaga a dostarczenia takiej metafizycznej koncepcji faktów modalnych, by na ich elementy konstytutywne nie składały się światy możliwe i możliwe indywidua. Choć — jak widzieliśmy w §4.6. — nie jest jasne, czy możliwe jest dostarczenie wystarczająco drobnoziarnistej koncepcji faktów modalnych bez uciekania się do kwantyfikacji po możliwych indywiduach, założmy na potrzeby dalszej dyskusji, że instrumentalista jest w stanie zaoferować zadowalającą koncepcję faktów modalnych.

---

<sup>49</sup>Szerzej o modalizmie będę pisał w §5.2.

Jeśli powyższe opcje wyczerpują możliwe posunięcia instrumentalisty, to staje on przed poważnym dylematem, który na tym polega, że proponowane przez niego stanowisko *albo* redukuje się w całości do innego stanowiska, *albo* jest teoretycznie zbyt ubogie, by można było w jego ramach wyjaśnić fenomen teoretycznej użyteczności dyskursu możliwościowego.

Pierwszy człon dylematu: jeśli modalizm wpisuje się w klasę literalnie prawdziwych interpretacji dyskursu możliwościowego, które dorównują mu użytecznością, to zasadnym jest pytanie, dlaczego nie zarzucić instrumentalizmu na rzecz modalizmu? Podobnie jak instrumentalista zwolennik modalizmu może twierdzić, że teoretycy wykorzystujący dyskurs możliwościowy w swoich rozważaniach przykładają większą wagę do jego użyteczności niż do kontrowersyjnych zobowiązań ontologicznych. W końcu znakomitej części z nich nie przeszkadzało odwoływanie się do światów możliwych i zamieszkujących je indywiduów, mimo że większość filozofów patrzy z niedowierzaniem na ontologię Lewisa. Teoretykom tym może być również daleko do akceptacji ontologii umiarkowanego realizmu modalnego, ponieważ jest ona zbyt gruboziarnista, by można było w niej odróżnić intuicyjnie odmienne możliwości<sup>50</sup>. Dalej teza, że literalnie prawdziwa interpretacja nie musi odpowiadać temu, co użytkownik dyskursu możliwościowego ma na myśli w danym kontekście, nie jest właściwa wyłącznie tej wersji instrumentalizmu. Jak widzieliśmy w poprzednim rozdziale, w literaturze wyróżnia się bowiem rewolucyjną wersję metafikcjonalizmu modalnego, zgodnie z którą fikcjonalistyczne propozycje interpretacji twierdzeń możliwościowych mogą nie odpowiadać temu, co użytkownicy dyskursu możliwościowego mają na myśli, gdy mówią o światach możliwych. Teoretycznie modalizm również mógłby wystąpić w rewolucyjnej wersji. Jeśli tak, to trudno znaleźć rzeczowe twierdzenie, które głosiłby instrumentalista, a którego nie mogliby

---

<sup>50</sup>Ilustrowałem to w poprzednim rozdziale za pomocą przykładu możliwości istnienia dwóch krążących wokół siebie elektronów.

głosić zwolennicy modalizmu w jego rewolucyjnej wersji. Wówczas jednak wydaje się, iż instrumentalizm można w całości zredukować do modalizmu.

Drugi człon dylematu: jeśli instrumentalizm nie redukuje się do modalizmu, to zawarta w nim koncepcja cierpi na poważny brak. Można oczekiwać od interpretacji twierdzeń dyskursu możliwościowego, że wyjaśni ona, w jaki sposób twierdzenia tego dyskursu mogą pełnić te role teoretyczne, które faktycznie pełnią. Ograniczając się do omawianego wyżej rozumowania  $(P_1) \rightarrow (W)$ , można oczekiwać, że interpretacja ta da podstawy do wyjaśnienia, dlaczego do sprawdzenia poprawności rozumowań sformułowanych przy użyciu terminologii światów możliwych wystarczą jedynie metody właściwe klasycznemu rachunkowi predykatów. Tym samym powinniśmy otrzymać częściowe wyjaśnienie fenomenu użyteczności twierdzeń możliwościowych pojawiających się jako przesłanki w takich rozumowaniach jak  $(P_1) \rightarrow (W)$ . Z oczywistych względów zasada  $(I_1)$  nie stanowi wyjaśnienia instrumentalnej użyteczności twierdzeń dyskursu możliwościowego, gdyż jest on raczej jej konstatacją niż wyjaśnieniem. W związku z tym pozostaje nam jedynie zasada  $(I_2)$ , a zatem fenomen użyteczności twierdzeń dyskursu możliwościowego powinien być wyjaśniany poprzez odwołanie się do ich literalnie prawdziwych interpretacji. Jeśli będą nimi jakieś twierdzenia modalne, to ich warunki prawdziwości nie wyjaśnią fenomenu użyteczności twierdzeń dyskursu możliwościowego. Z punktu widzenia użyteczności ostatnie mają bowiem tę przewagę nad twierdzeniami modalnymi, że do sprawdzenia poprawności rozumowań sformułowanych przy użyciu twierdzeń możliwościowych można zastosować metody pierwszego rzędu właściwe klasycznemu rachunkowi predykatów. Metody te nie mają jednak zastosowania do rozumowań sformułowanych za pomocą twierdzeń modalnych. Stoimy zatem przed pytaniem: „Co odróżnia twierdzenia dyskursu możliwościowego od ich literalnie prawdziwych interpretacji, że pierwszy

legitymuje się zaletami, których nie sposób przypisać drugiemu?”, a instrumentaliscie wyczerpały się już środki, za których pomocą mógłby wyjaśnić fenomen użyteczności dyskursu możliwościowego. Owocne nie będą ani próby odwołania się do ( $I_1$ ), ani próby odwołania się do ( $I_2$ ).

W odróżnieniu od instrumentalisty fikcjonalista modalny posiada środki, które pozwalają na wyjaśnienie użyteczności dyskursu możliwościowego. Fikcjonaliscie modalni starają się bowiem przyporządkować takie warunki prawdziwości twierdzeniom tego dyskursu, by pozwoliły one na stosowanie metod właściwych klasycznemu rachunkowi predykatów podczas sprawdzania logicznej poprawności rozumowań sformułowanych przy użyciu tych twierdzeń. W przekonaniu fikcjonalisty stosowanie metod właściwych klasycznemu rachunkowi predykatów jest możliwe, ponieważ treść opowieści *PW* bądź gry *GPW* jest zamknięta na implikację pierwszego rzędu. W przypadku metafikcjonalizmu modalnego założenie to przyjmie następującą postać: jeśli twierdzenie możliwościowe  $q^*$  wynika z twierdzenia możliwościowego  $p^*$ , to jeśli „Zgodnie z *PW*,  $p^*$ ” jest prawdziwe, to „Zgodnie z *PW*,  $q^*$ ” jest prawdziwe. Dzięki temu założeniu metafikcjonalista modalny może posługiwać się metodami pierwszego rzędu, gdy sprawdza poprawność rozumowań sformułowanych przy użyciu twierdzeń możliwościowych poprzedzonych funktorem opowieści. Jest tak dlatego, że pomijając funktor opowieści, twierdzenia możliwościowe można wyrazić w języku klasycznego rachunku predykatów. Z kolei przyjęcie założenia, że treść fikcji *PW* jest zamknięta na implikację pierwszego rzędu, pozwala sprawdzać poprawność rozumowań, w których jako kroki figurują twierdzenia o postaci: „Zgodnie z *PW*,  $p^*$ ”, „Zgodnie z *PW*,  $q^*$ ”, poprzez sprawdzenie, czy z twierdzenia możliwościowego  $p^*$  wynika twierdzenie możliwościowe  $q^*$ . W przypadku przedmiotowego fikcjonalizmu modalnego założenie to przyjmie następującą postać: jeśli  $q^*$  wynika z  $p^*$ , to jeśli  $p^*$  jest fikcyjnie prawdziwe, to

$q^*$  jest fikcyjnie prawdziwe. Podobnie jak wyżej twierdzenia możliwościowe są wyrażone w języku klasycznego rachunku predykatów. Dzięki założeniu, że treść gry *GPW* jest zamknięta na implikację pierwszego rzędu, można sprawdzać poprawność rozumowań sformułowanych przy użyciu twierdzeń możliwościowych za pomocą metod właściwych klasycznemu rachunkowi predykatów. Jeśli zgodzić się, że realizm modalny jest teorią zamkniętą na implikację pierwszego rzędu, wówczas przyjęcie przez fikcjonalistę modalnego założenia, że treść fikcji *PW* bądź treść gry *GPW* jest zamknięta na implikację pierwszego rzędu, nie powinno budzić szczególnych kontrowersji.

Odpowiadając na pytanie, co decyduje o instrumentalnej użyteczności twierdzeń modalnych, można pokusić się o wysunięcie hipotezy, że jest to możliwość stosowania metod pierwszego rzędu przy sprawdzaniu poprawności rozumowań sformułowanych przy ich użyciu. Dzięki temu, że fikcjonalista modalny przyporządkowuje takie warunki prawdziwości twierdzeniom możliwościowym, że pozwalają one na wyjaśnienie możliwości stosowania metod pierwszego rzędu do sprawdzania poprawności rozumowań sformułowanych przy użyciu tych twierdzeń, może on wyjaśnić, dlaczego metody takie w ogóle mają zastosowanie do takich rozumowań. Instrumentalista nie operuje natomiast środkami, które pozwoliłyby na wyjaśnienie możliwości stosowania tych metod do sprawdzania poprawności takich rozumowań. W związku z tym wydaje się, że z tej batalii zwycięsko wychodzi fikcjonalista.

## 2 Modalizm versus fikcjonalizm modalny

Kolejnym stanowiskiem konkurencyjnym względem fikcjonalizmu modalnego jest modalizm. Jak pisze Kit Fine:

Podstawową ideą jest uznanie modalności za pierwotną i potraktowanie posybilistycznego kwantyfikatora „istnieje

możliwy przedmiot  $x$ <sup>51</sup>” jako równoważnego „możliwie, istnieje przedmiot  $x$ ”, gdzie drugi kwantyfikator (w zasięgu funktora możliwości) jest aktualistyczny — w każdym świecie przebiega po przedmiotach aktualnych w tym świecie. Zatem powiedzieć, że istnieje możliwy przedmiot, który możliwie jest mówiącym osłem, to powiedzieć, że możliwie, istnieje przedmiot, który możliwie jest mówiącym osłem (Fine 2003/2007, s. 225).

Idea ta wymaga jednak modyfikacji. Uznanie, że posybilistyczna kwantyfikacja, tj. kwantyfikacja wiążąca zmienne przebiegające po możliwych indywiduach lub światach możliwych, może być w ogólności zredukowana do kwantyfikacji po aktualnych indywiduach, znajdującej się w zasięgu funktora modalnego możliwości, prowadzi do następującej trudności: w języku realisty modalnego prawdą jest, że istnieje takie możliwe indywiduum  $x$ , które aktualnie nie istnieje. Gdyby ostatnie twierdzenie zredukować do twierdzenia, że możliwie, istnieje takie aktualne indywiduum  $x$ , że  $x$  aktualnie nie istnieje, popadlibyśmy w sprzeczność. Jak uniknąć tej niepożądanej konsekwencji?

Fine zwraca uwagę, że problem powstaje, gdyż funktor możliwości pełni taką funkcję, że formuła znajdująca się w jego zasięgu jest wartościowana w świecie, w którym istnieje indywiduum  $x$ , tymczasem w przypadku powyższego twierdzenia modalnego formuła w zasięgu funktora możliwości powinna być wartościowana w świecie aktualnym, gdyż mówi ona o tym, co znajduje się w dziedzinie świata aktualnego. W związku z tym potrzebujemy narzędzi, by powrócić ze świata, do którego odsyła nas funktor możliwości, do świata aktualnego. Rezultat ten można osiągnąć, jeśli poprzedzi się znajdującą się w zasięgu funktora możliwości kwantyfikację po aktualnych indywiduach kwan-

---

<sup>51</sup>W tym paragrafie zmienne, które są zapisane wytłuszczoną kursywą, są zmiennymi przebiegającymi po światach możliwych lub możliwych indywiduach, natomiast zakresem zmiennych, które nie są zapisane wytłuszczoną kursywą, są aktualnie istniejące indywidua lub świat aktualny.



tyfikacją egzystencjalną po świecie aktualnym: „Istnieje taki świat aktualny  $w$ ” (zob. Fine 1977/2007, s. 156; Fine 1985/2007, s. 211; Fine 2003/2007, s. 225). Dzięki temu otrzymujemy twierdzenie „Istnieje świat aktualny  $w$  taki, że możliwe, istnieje takie indywiduum  $x$ , że jego nieistnienie jest kompatybilne z tym, że świat  $w$  jest aktualny”. W tym miejscu należy podkreślić rolę, jaką pełni funktor modalny możliwości, w którego zasięgu występuje kwantyfikacja egzystencjalna po aktualnych indywiduach. W twierdzeniach o transświatowym zakresie funktor ten pełni rolę ograniczania zakresu zmiennej  $x$  występującej w jego zasięgu do bytów możliwe istniejących w możliwie istniejącym świecie aktualnym różnym od świata, który faktycznie jest światem aktualnym. Dzięki pełnieniu przez ten funktor tej funkcji możliwe istniejące indywiduum  $x$  nie musi istnieć w świecie aktualnym  $w$ . W przeciwnym razie proponowany przez Fine’a przykład twierdzenia possibility byłby paradoksalny. Indywiduum  $x$  jako aktualne musiałoby istnieć w świecie aktualnym  $w$ , co nie byłoby kompatybilne z jego nieistnieniem w tym świecie.

Powyższa forma przekładu może być uogólniona do następującej postaci:

Istnieje takie możliwe indywiduum  $x$ , że  $Fx$  wtw, gdy istnieje świat aktualny  $w$  taki, że możliwe, istnieje takie aktualne indywiduum  $x$ , że  $Fx$  jest kompatybilne z tym, że świat  $w$  jest aktualny (por. Fine 2003/2007, s. 225).

Dalszej redukcji można poddać frazę „ $Fx$  jest kompatybilne z tym, że świat  $w$  jest aktualny” (zob. Fine 1985/2007, s. 211). W pierwszym rzędzie predykat „jest aktualny” może być analizowany za pomocą predykatu „istnieje”, ponieważ modalista przyjmuje wyłącznie istnienie świata aktualnego i aktualnych indywiduów. Następnie zaś na gruncie założenia ( $M_1$ ), że jeśli pewne twierdzenie mogłoby być prawdziwe w pewnym świecie (tj. jego prawdziwość jest kompatybilna z istnieniem tego świata), to musi ono być w nim prawdziwe (Fine

1985/2007, s. 213; Fine 2003/2007, s. 226), kompatybilność  $Fx$  z istnieniem świata  $w$  można analizować w kategoriach ścisłej implikacji jako „Koniecznie, jeśli świat  $w$  istnieje, to  $Fx$ ”. Otrzymujemy zatem następującą analizę:

(EP) Istnieje takie możliwe indywiduum  $\mathbf{x}$ , że  $F\mathbf{x}$  wtw, gdy istnieje świat  $w$  taki, że jest możliwe, że istnieje takie aktualne indywiduum  $x$ , że koniecznie, jeśli  $w$  istnieje, to  $Fx$ .

Z zamierzenia ścisła implikacja „Koniecznie, jeśli  $w$  istnieje, to  $Fx$ ” ma stanowić analizę pojawiającej się w języku realisty modalnego frazy „ $Fx$  w świecie  $w$ ”.

W zasadniczo podobny sposób proponuje się eliminować kwantyfikację posybilistyczną po światach możliwych. Tam, gdzie pojawi się kwantyfikacja egzystencjalna po światach możliwych, zwolennik modalizmu będzie wprowadzał kwantyfikację egzystencjalną po światach aktualnych w zasięgu funktora modalnego możliwości. Otrzymujemy zatem następujący schemat redukcji:

(EW) Istnieje taki świat  $\mathbf{v}$ , że  $A\mathbf{v}$  wtw, gdy istnieje taki świat aktualny  $w$ , że możliwe, istnieje taki świat aktualny  $v$ , że koniecznie, jeśli  $w$  istnieje, to  $Av$  (zob. Fine 1985/2007, s. 211).

Do tej pory modalista zaproponował redukcję części twierdzeń egzystencjalnych przyjmowanych przez zwolenników istnienia światów możliwych. Jednakże ci wygłaszają również twierdzenia, w których występuje kwantyfikacja ogólna po możliwych indywiduach bądź światach możliwych. W przypadku takich twierdzeń modalista proponuje, by zastąpić kwantyfikację ogólną, wiążącą zmienne przebiegające po światach możliwych lub możliwych indywiduach, kwantyfikacją ogólną, która wiąże zmienne przebiegające po aktualnych indywiduach i aktualnych światach, z tym że — w odróżnieniu od kwantyfikacji egzystencjalnej — kwantyfikacja ogólna będzie znajdowała się w zasięgu

funktor modalnego konieczności. Proponuje się zatem następujące schematy redukcyjne:

- (UP) Dla każdego możliwego indywiduum  $\mathbf{x}$ ,  $G\mathbf{x}$  wtw, gdy istnieje taki świat aktualny  $w$ , że koniecznie, dla każdego aktualnego indywiduum  $x$ , koniecznie, jeśli  $w$  istnieje, to  $Gx$ .
- (UW) Dla każdego świata  $\mathbf{v}$ ,  $B\mathbf{v}$  wtw, gdy istnieje taki świat aktualny  $w$ , że koniecznie, dla każdego świata aktualnego  $v$ , koniecznie, jeśli  $w$  istnieje, to  $Bv$  (por. Fine 1977/2007, s. 144–145).

Funktor modalny konieczności pojawia się tutaj, ponieważ kwantyfikacja znajdująca się w jego zakresie ma przebiegać po wszystkich aktualnie istniejących indywiduach niezależnie od tego, który ze światów będzie światem aktualnym.

Wydaje się, że w odróżnieniu od redukcjonistycznej propozycji Plantingi wprowadzone wyżej schematy pozwalają na zredukowanie twierdzenia możliwościwiatowego „Istnieje taki świat  $\mathbf{v}$ , że istnieją takie możliwe indywidua  $\mathbf{x}$  i  $\mathbf{y}$ , że ( $F\mathbf{x}$  wtw, gdy  $F\mathbf{y}$ ) i  $\mathbf{x} \neq \mathbf{y}$  w świecie  $\mathbf{v}$  i dla dowolnego  $\mathbf{z}$ , jeśli  $\mathbf{z}$  istnieje w świecie  $\mathbf{v}$ , to  $\mathbf{z} = \mathbf{x}$  lub  $\mathbf{z} = \mathbf{y}$ ”, któremu w języku modalisty odpowiada twierdzenie: „Istnieje świat aktualny  $w$  taki, że możliwie, istnieje taki świat  $v$ , możliwie, istnieje aktualne indywiduum  $x$  i  $y$  oraz koniecznie, dla każdego możliwego indywiduum  $z$ , koniecznie, jeśli  $w$  istnieje, to ( $Fx$  wtw, gdy  $Fy$ ) i  $x \neq y$ ) i koniecznie, jeśli  $w$  istnieje, to  $z = x$  lub  $z = y$ ”.

Wszystkie dotychczasowe redukcje twierdzeń możliwościwiatowych opierają się na założeniu, że wartość logiczna twierdzeń egzystencjalnych bądź uniwersalnych, w których kwantyfikuje się po aktualnych indywiduach, powinna być ustalana przez sprawdzenie, czy w dziedzinie któregoś świata znajdują się byty, które spełniają pewien warunek. Nawet wartość logiczna twierdzenia, że istnieje taki świat aktualny  $w$ , że możliwie, istnieje takie indywiduum  $x$ , że koniecznie, jeśli  $w$  istnieje, to  $x$  nie istnieje, jest ustalana poprzez sprawdzenie, czy  $x$

znajduje wśród indywiduów istniejących w świecie  $w$ , tj. czy w świecie  $w$  istnieje takie  $y$ , że spełnia ono warunek „ $y = x$ ”. Tymczasem niektóre z twierdzeń możliwościowych mogą być prawdziwe, nawet jeśli nie istnieje świat, w którego dziedzinie istnieją indywidua spełniające warunek wyrażony w tym twierdzeniu. Próba ustalenia wartości logicznej takich twierdzeń powinna dokonywać się przez sprawdzanie, czy w uniwersum możliwych indywiduów istnieją takie indywidua, które spełniają warunek wyrażony w tym twierdzeniu. Przykładem takiego twierdzenia jest twierdzenie: „Istnieje taki zbiór możliwych indywiduów  $X, Y, Z$ , że dla dowolnego indywiduum  $x$  należącego do  $X$ , dla dowolnego  $y$  należącego do  $Y$  i dla dowolnego  $z$  należącego do  $Z$ ,  $x$  jest bardziej podobne do  $y$  niż do  $z$ ” itd. Twierdzenie to jest prawdziwe nawet wówczas, gdy możliwe indywidua należące do tych zbiorów nie znajdują się w jednym świecie. W jaki sposób modalista redukuje kwantyfikację po zbiorach możliwych indywiduów?

W przypadku takich twierdzeń kwantyfikację egzystencjalną po zbiorach możliwych indywiduów proponuje się zredukować w następujący sposób: niech  $X = \{x_1, x_2, \dots\}$ , wówczas twierdzenie „Istnieje takie  $X$ , że dla każdego  $x$  należącego do  $X$ ,  $Fx$ ” będzie zredukowane do twierdzenia: „Możliwie, istnieje takie aktualne indywiduum  $x_1$ , możliwe istnieje takie aktualne indywiduum  $x_2, \dots$ , że koniecznie, dla każdego  $x$ , koniecznie, jeśli  $x = x_1$  lub  $x = x_2$ , lub  $\dots$  i  $x_1 \neq x_2$ , i  $\dots$ , to  $Fx$ ”. Z uwagi na to, że każde wystąpienie kwantyfikatora egzystencjalnego jest poprzedzone oddzielnym funktorem modalnym możliwości, z których każdy może odsyłać do odmiennego świata, nie zachodzi konieczność, by  $x_1, x_2, \dots$  znajdowały się w tym samym świecie. Z kolei fakt, że formuła „dla każdego  $x$ , koniecznie, jeśli  $x = x_1$  lub  $x = x_2$ , lub  $\dots$  i  $x_1 \neq x_2$ , i  $\dots$ , to  $Fx$ ” jest poprzedzona funktorem modalnym konieczności, sprawia, że obowiązuje ona we wszystkich światach, które wybierane są przez funktory modalne możliwości poprzedzające kwantyfikatory egzystencjalne (por. Fine 1977/2007, s.

157–158).

W związku z powyższym przywołany wyżej przykład twierdzenia o transświatowym zakresie posiada następującą redukcyjną analizę: niech  $X = \{x_1, x_2, \dots\}$ ,  $Y = \{y_1, y_2, \dots\}$ ,  $Z = \{z_1, z_2, \dots\}$ . Wówczas twierdzenie to będzie redukowane do twierdzenia „Możliwie, istnieje takie aktualne indywiduum  $x_1$ , możliwe istnieje takie aktualne indywiduum  $x_2, \dots$ , możliwe, istnieje takie aktualne indywiduum  $y_1$ , możliwe istnieje takie aktualne indywiduum  $y_2, \dots$ , możliwe, istnieje takie aktualne indywiduum  $z_1$ , możliwe istnieje takie aktualne indywiduum  $z_2, \dots$ , że koniecznie, dla każdego  $x, y, z$ , koniecznie, jeśli  $x = x_1$  lub  $x = x_2$ , lub  $\dots$ ,  $y = y_1$ , lub  $y = y_2, \dots$ , lub  $z = z_1$ , lub  $z = z_2$ , lub  $\dots$ , i  $x_1 \neq x_2, i \dots, y_1 \neq y_2, i \dots, z_1 \neq z_2, i \dots$ , to  $x$  jest bardziej podobne do  $y$  niż do  $z$ ”.

Wydaje się, że proponowana przez modalistę redukcja twierdzeń dyskursu możliwościowego do odpowiednich twierdzeń modalnych dorównuje mocami ekspresywnymi językowi dyskursu możliwościowego. Redukcja ta opiera się na dwóch założeniach, mianowicie na założeniu (M1), że jeśli pewne twierdzenie mogłoby być prawdziwe w pewnym świecie, to musi ono być w nim prawdziwe, oraz założeniu (M2), że z konieczności, istnieje takie  $w$ , że  $w$  jest światem aktualnym (zob. Fine 1985/2007, s. 213 i Fine 2003/2007, s. 226). Konieczność przyjęcia pierwszego założenia ujawniła się już wyżej. Tym z kolei, co nakazuje przyjęcie drugiego z nich, jest to, że proponowane przez modalistę redukcje są w zasadzie twierdzeniami modalnymi, mówiącymi o tym, co miałoby miejsce, gdyby pewien świat, który mógłby istnieć, był światem aktualnym. Gdy przyjmie się założenie (M1), wówczas twierdzenia dotyczące tego, co miałoby miejsce, gdyby pewien świat istniał, będą twierdzeniami koniecznie prawdziwymi, ponieważ to, co mogłoby mieć miejsce w takim świecie, musi mieć w nim miejsce. Gdy jednak w ramach redukcyjnej analizy twierdzeń możliwościowych pojawiła się konieczność odwołania się do świata aktualnego,

konieczna prawdziwość tych redukcyjnych twierdzeń modalnych stała się zależna od istnienia świata aktualnego. Jeśli chcemy zachować intuicje co do koniecznej prawdziwości twierdzeń modalnych, musimy właśnie przyjąć założenie ( $M_2$ ).

Zarówno założenie ( $M_1$ ), jak i ( $M_2$ ) nie powinny budzić szczególnych zastrzeżeń. Wedle tego drugiego, jeśli pojmujemy świat aktualny jako całość istniejących przedmiotów, wówczas na gruncie stanowiska, że istnieje tylko i wyłącznie świat aktualny, przyjęcie tezy, że koniecznie istnieje jakiś świat aktualny, jest przyjęciem tezy, że z konieczności istnieją jakieś przedmioty (mianowicie te, które składają się na dany świat). Jeśli nie jesteśmy zwolennikami tezy, że mogłoby być tak, że nic nie istnieje, założenie to nie powinno budzić szczególnych wątpliwości. Z kolei pierwsze z tych założeń wyraża tezę, że sposób, na jaki jest dana możliwość, jest sposobem, na jaki jest ona z konieczności. Gdy możliwe jest, by istniały latające słonie, to niemożliwe jest, by możliwość istnienia latających słoni zachodziła, a latające słonie nie istniały.

Skoro już jesteśmy zaznajomieni z proponowaną przez modalistę redukcją twierdzeń możliwościświatowych, należy postawić pytanie, co w przeprowadzonej redukcji pozwala uniknąć konieczności przyjmowania istnienia światów możliwych i możliwych indywiduów. Odpowiedź, którą podsuwają wprowadzone wyżej schematy redukcji, przedstawia się następująco: koncepcja modalisty pozwoliła na wyeliminowanie kwantyfikacji egzystencjalnej po światach możliwych i możliwych indywiduach. Dzięki temu w twierdzeniach stanowiących redukcje twierdzeń możliwościświatowych nie występują odwołania do tego typu bytów, a w związku z tym wartość logiczna tych twierdzeń nie jest zależna od ich istnienia. Jak widzieliśmy w rozdziale 1., to, jakie zobowiązania ontologiczne są implikowane przez twierdzenia danego dyskursu, zależne jest od tego, istnienie jakich bytów funduje ich prawdziwość. Skoro w modalistycznych redukcjach nie występują

odniesienia do światów możliwych i możliwych indywiduów, to ich prawdziwość *prima facie* nie jest ufundowana na istnieniu tych bytów. Z kolei twierdzenia takie, jak „Istnieje aktualne indywiduum, które jest mówiącym osłem w świecie aktualnym *v*”, są uznawane — przynajmniej w ramach jednej z opcji (zob. Fine 1977/2007, s. 146–147) — za fałszywe, gdyż nie są spełnione przez żadne z bytów istniejących w świecie aktualnym. Mimo że twierdzenia te są fałszywe, mogą one stanowić argumenty funktora modalnego możliwości, który będzie wtedy przyjmował za wartości prawdziwe twierdzenia modalne. Dla przykładu prawdziwość twierdzenia „Możliwie, istnieje takie aktualne indywiduum, że indywiduum to jest mówiącym osłem” ma być niezależna od istnienia bytu, który spełniałby tę deskrypcję. By uzyskać ten rezultat, należy przyjąć założenie, że funktor modalny możliwości jest *niefaktualny*, tj. prawdziwość twierdzeń modalnych z tym funktorem jest niezależna od prawdziwości twierdzeń znajdujących w jego zasięgu.

Specyfika redukcji twierdzeń możliwościświatowych proponowanej przez modalistę polega na tym, że nie dokonuje się ona poprzez wskazanie bytów, które są akceptowalne dla przeciwnika istnienia światów możliwych i możliwych indywiduów, a które mogłyby pełnić ich rolę. Jak zauważa Fine, proponowana przez niego redukcja jest syntaktycznie radykalna w tym sensie, że nie zachowuje formy syntaktycznej twierdzeń możliwościświatowych, gdyż kwantyfikacja egzystencjalna po światach możliwych i możliwych indywiduach nie jest — jak w przypadku redukcji syntaktycznie konserwatywnych — zastępowana kwantyfikacją po bytach aktualnych. Jest zastępowana kwantyfikacją po bytach aktualnych w zasięgu funktorów modalnych. Z uwagi na to, że zakłada się, że nie istnieją byty spełniające twierdzenie egzystencjalne znajdujące się w zasięgu funktora modalnego możliwości, w ogólności modalistyczna redukcja twierdzeń możliwościświatowych nie pozwala na przeprowadzenie identyfikacji pomiędzy bytami, których dotyczą

redukowane twierdzenia, a bytami akceptowalnymi dla modalisty (por. Fine 1977/2007, s. 145; Fine 2003/2007, s. 224–225).

Jak widzieliśmy w poprzednim rozdziale, fikcjonalista modalny musiał uciekać się do przyjęcia tezy, że kwantyfikacja po światach w zasięgu funktora modalnego możliwości jest ontologicznie niewinna. Dlatego też nie powinien mieć zastrzeżeń co do samej idei unikania zobowiązań ontologicznych przez przyjmowanie, że kwantyfikatory w zasięgu funktora modalnego możliwości są ontologicznie niezobowiązujące. Dlaczego mimo to fikcjonalista nie chce uogólnić strategii unikania zobowiązań ontologicznych przez poprzedzenie funktorem modalnym możliwości wszystkich twierdzeń, w których występują deskrypcje światów możliwych i możliwych indywiduów? Wydaje się, że istnieją co najmniej dwie racje, by pozostać przy fikcjonalizmie i odrzucić modalizm.

Modalista przyjmuje, że kwantyfikacja posybilistyczna powinna być zredukowana do kwantyfikacji aktualistycznej w zasięgu funktorów modalnych, które z kolei nie podlegają dalszej analizie. Dlatego proponowana przez niego redukcja nie jest w stanie zachować jednej z największych zalet dyskursu możliwościświatowego — możliwości stosowania metod pierwszego rzędu do sprawdzania logicznej poprawności rozumowań sformułowanych przy użyciu twierdzeń możliwościświatowych. Przykładowo, fikcjonalista modalny wyjaśnia logiczną niepoprawność rozumowania:

(P<sub>1</sub>) Możliwie,  $p$ .

(P<sub>2</sub>) Możliwie,  $q$ .

(W) Możliwie,  $p$  i  $q$ .

przez odwołanie się do tego, że (w zgodzie z fikcją  $PW$  lub w ramach gry  $GPW$ ) z tego, że (P<sub>1PW</sub>) istnieje świat, w którym  $p$  i z tego, że (P<sub>2PW</sub>) istnieje świat, w którym  $q$ , nie wynika, że (W<sub>PW</sub>)



istnieje świat, w którym  $p$  i  $q$ . Przy założeniu, że treść fikcji  $PW$  bądź gry  $GPW$  jest zamknięta na implikację pierwszego rzędu, takie wyjaśnienie logicznej niepoprawności tego rozumowania wymaga jedynie odwołania się do intuicji dotyczących kwantyfikacji egzystencjalnej pierwszego rzędu i predykcji pierwszego rzędu. Dzięki temu unika się konieczności odwoływania się do jakichkolwiek reguł rządzących zachowaniem pojęć intensjonalnych. Modalista nie może zaproponować analogicznie oszczędnego wyjaśnienia logicznej niepoprawności tego rozumowania, ponieważ musi uznać, że jego niepoprawność nie podlega dalszemu wyjaśnieniu i jest zależna od własności funktora modalnego możliwości. W ramach proponowanej przez modalistę redukcji twierdzeń możliwoswiatowych to twierdzenia modalne, a nie twierdzenia możliwoswiatowe są pierwotne w porządku wyjaśniania, dlatego też niepoprawność tego rozumowania nie może być wyjaśniana w kategoriach kwantyfikacji po światach możliwych, lecz musi — przez odwołanie się do własności funktora modalnego możliwości, który modalista uznaje za funktor pierwotny. Uznając jednak, że niepoprawność tego rozumowania wynika z własności pierwotnego funktora modalnego możliwości, modalista pozbawia się zarazem możliwości stosowania metod pierwszego rzędu do sprawdzania poprawności rozumowań sformułowanych przy użyciu twierdzeń zaproponowanych przez niego jako redukcje twierdzeń możliwoswiatowych. Podobnie nie będzie on w stanie sprawdzić poprawności rozumowania  $(Z_1) - (Z_{15})$ , jeśli zastosuje metody pierwszego rzędu. Proponowane przez niego redukcje tych twierdzeń będą wymagały na nowo wprowadzenia pojęcia modalnego możliwości, a wnioskowanie to stanie się wnioskowaniem, w którym przesłankami będą nie twierdzenia pierwszego rzędu (czy też — precyzyjniej — twierdzenia pierwszego rzędu prawdziwe zgodnie z  $PW$  bądź w ramach  $GPW$ , których treści są zamknięte na implikację pierwszego rzędu), lecz niepodlegające dalszej redukcji twierdzenia modalne. Wydaje się zatem, że w tym miejscu propozycja fikcjonalisty

modalnego zyskuje przewagę nad propozycją modalisty.

Następnie fikcjonalista modalny mógłby zwrócić uwagę na to, czy faktycznie można nadać sens twierdzeniu „Istnieje świat aktualny *w* taki, że możliwe istnieje taki świat *v*, możliwe, istnieje aktualne indywiduum *x* i *y*, koniecznie, jeśli *w* istnieje, to ( $Fx$  wtw, gdy  $Fy$ ) i  $x \neq y$ ”, aby twierdzenie to było akceptowalne dla kogoś, kto dopuszcza wyłącznie kwantyfikowanie po aktualnie istniejących bytach. Chodzi mianowicie o wyjaśnienie, na jakiej podstawie odróżnia się indywidua *x* i *y*? W §4.4. starałem się wskazać na reprezentacje mentalne (pojęcia jednostkowe), do których mógłby odwołać się fikcjonalista modalny celem nadania sensu twierdzeniu o istnieniu świata, w którym istnieją dwa jakościowo nieodróżnialne indywidua. Nie wydaje się jednak, by modalista również mógł powołać się na takie reprezentacje mentalne celem wskazania podstawy odróżnienia *x* i *y*, gdyż prowadziłoby to do niepokojącego uzależnienia tego, co możliwe, od tego, jakie są reprezentacje mentalne użytkowników języka modalisty. Gdyby podstawą do odróżnienia *x* i *y* była zdolność utworzenia reprezentacji mentalnej dwóch numerycznie odmiennych i jakościowo nieodróżnialnych przedmiotów, to teoria modalności proponowana przez modalistę najprawdopodobniej padłaby ofiarą zarzutów podobnych do podnieszonego wyżej modalnego problemu sztuczności. Wydaje się bowiem, że to, czy użytkownicy języka modalisty posiadają pewne reprezentacje mentalne, jest przygodne, podczas gdy twierdzenia modalne proponowane przez modalistę jako redukcje twierdzeń możliwościowych przygodnie prawdziwymi nie są, mówią bowiem o tym, co z konieczności miałyby miejsce, gdyby pewne możliwości zachodziły. Gdyby stanowisko modalisty zostało podważone przez modalny problem sztuczności, stałoby się nieszczerze wiarygodną koncepcją modalności, zwłaszcza dla tych, którzy — jak sam Fine w (1994) — utrzymują, że prawdziwość twierdzeń z funktorem modalnym konieczności, a przynajmniej tych twierdzeń modalnych, w których występuje

pojęcie konieczności metafizycznej, jest ufundowana na metafizycznych istotach bądź naturach przedmiotów, których twierdzenia te dotyczą. Dla przykładu, konieczna prawdziwość twierdzenia „Sokrates jest człowiekiem” jest częściowo ufundowana na tym, że jest on ze swej istoty człowiekiem. Prawdziwość twierdzeń metafizycznie koniecznych jest więc ufundowana na faktach zachodzących niezależnie od tego, jakie reprezentacje mentalne posiadają użytkownicy języka modalisty. Wydaje się zatem, że możliwa prawdziwość pewnych twierdzeń powinna być po prostu zależna od tego, jakie są fakty. Wówczas powinien istnieć jakiś fakt, który funduje prawdziwość twierdzenia, że mogłyby istnieć jakościowo nieodróżnialne przedmioty  $x$  i  $y$ . Jeśli więc zgodzić się, że reprezentacje mentalne nie mogą stanowić podstawy do odróżnienia przedmiotów  $x$  i  $y$ , pozostajemy z pytaniem, na jakiej zasadzie można odróżnić te przedmioty? Niestety modalista nie udziela już na nie odpowiedzi.

Problem ten staje się tym bardziej palący, gdy uświadomić sobie, że gdy modalista nie wskaże podstaw odróżnienia pomiędzy jakościowo nieodróżnialnymi indywiduami, może pojawić się następujące zagrożenie ze strony realisty modalnego: mógłby on twierdzić, że jedyną podstawą przeprowadzenia takiego odróżnienia jest właśnie odwołanie się do jakościowo nieodróżnialnych możliwych indywiduów istniejących w jakimś nieaktualnym świecie. Na tej podstawie mógłby argumentować, że proponowana przez modalistę redukcja nie spełnia zadania, które ów przed nią postawił, tj. nie redukuje twierdzeń możliwościświatowych do bardziej fundamentalnych twierdzeń modalnych, gdyż próba przeprowadzenia takiej redukcji zakończyła się koniecznością odwołania się do rozróżnienia pomiędzy indywiduami, które na nowo wymaga odwołania się do nieaktualnych bytów.

Co prawda, modalista utrzymuje, że nie dokonuje redukcji twierdzeń możliwościświatowych przez wskazanie dla każdego bytu (przyjawanego przez realistę modalnego) bytu, który pełniłby jego rolę i

który byłby akceptowalny dla aktualisty i pełniłby rolę Jednak nie zwalnia go to z obowiązku wyjaśnienia, na jakiej podstawie dokonuje odróżnienia indywiduów  $x$  i  $y$ . W podobnej sytuacji znajdował się przecież fikcjonalista modalny, który również nie proponował redukcji bytów przyjmowanych przez realistę modalnego do bytów aktualnie istniejących, a mimo wszystko stanął przed trudnością wyjaśnienia, w jaki sposób możliwe jest, by treść fikcji  $PW$  była wystarczająco drobnoziarnista i pozwoliła na odróżnienie dwóch jakościowo nieodróżnialnych możliwych indywiduów, a przy tym nie pociągała za sobą ich istnienia. Jeśli zaproponowana w poprzednim rozdziale koncepcja pojęć jednostkowych ma choć cień szans na powodzenie, to umiarkowany fikcjonalista modalny nie popada w kłopoty z tytułu samego wyjaśnienia, na jakiej podstawie może on dokonać odróżnienia indywiduów  $x$  i  $y$ . Jest tak dlatego, że proponowana przez niego interpretacja dyskursu możliwościświatowego nie jest proponowana jako teoria modalności (a jedynie jako teoria ustalająca zależności pomiędzy fikcyjną prawdziwością bądź prawdziwością zgodnie z  $PW$  twierdzeń możliwościświatowych a istnieniem faktów modalnych). Tymczasem propozycja modalisty nie jest niczym innym, jak właśnie teorią modalności. Jeśli więc źródłem odróżnienia indywiduów  $x$  i  $y$  miałyby być zdolność do utworzenia odpowiednich reprezentacji mentalnych, to niewątpliwie koncepcja modalisty nie ostoja wobec modalnego problemu sztuczności. Możliwość uniknięcia tej trudności z pewnością będzie kolejną racją dla pozostania przy fikcjonalizmie modalnym.

Co ciekawe, powstaje również pytanie, czy wskazanie obiektywnej i akceptowalnej dla przeciwnika istnienia światów możliwych podstawy, na której modalista mógłby dokonać odróżnienia pomiędzy indywiduami  $x$  i  $y$ , nie sprawiałoby, że dla dowolnego bytu przyjmowanego przez zwolennika istnienia światów możliwych modalista byłby jednak w stanie wskazać byt, który pełniłby rolę tego bytu. Wówczas jego stanowisko można byłoby zarzucić na rzecz stanowiska, które

proponuje syntaktycznie konserwatywną redukcję twierdzeń możliwościowych i wprowadza byty pełniące role światów możliwych i możliwych indywiduów.

Niewykluczone więc, że modalista staje przed następującym dylematem: *albo* proponowana przez niego redukcja nie opiera się na redukcji światów możliwych i możliwych indywiduów do aktualistycznie akceptowalnych bytów, a wówczas powstaje problem z wyjaśnieniem, na jakiej podstawie odróżnia indywidua  $x$  i  $y$ , *albo* jest on w stanie wskazać na obiektywną podstawę rozróżnienia tych bytów, a wówczas nie jest jasne, dlaczego proponowana przez niego redukcja miałaby nie przyjąć konserwatywnej formy, w której istnienie światów możliwych i możliwych indywiduów jest reprezentowane przez istnienie bytów należących do tego typu bytów, do których musiał się odwołać, gdy wskazywał na podstawę odróżnienia indywiduów  $x$  i  $y$ .

Niewykluczone również, że modalista jest w stanie dostarczyć przekonującej odpowiedzi na pytanie, na jakiej podstawie odróżnia indywidua  $x$  i  $y$ , niemniej jednak mnie nie nasuwa się na myśl żadna propozycja, która pozwoliłaby obejść ten problem, a sami modalisci niewiele mówią o tej sprawie. Co więcej, gdyby nawet modaliscie udało się wskazać obiektywną postawę, na której mógłby on dokonać odróżnienia pomiędzy indywiduami  $x$  i  $y$  — która nie wymagałaby przyjmowania istnienia możliwych indywiduów — jego stanowisko wciąż mogłoby przegrywać batalię z umiarkowanym fikcjonalizmem modalnym. Ostatni mógłby bowiem, odwołując się do tych samych, przyjętych przez modalistę, obiektywnych podstaw rozróżnienia indywiduów  $x$  i  $y$ , próbować poddać analizie własność bycia możliwie egzemplifikowanym przez wiele indywiduów w taki sposób, by jej przypisanie własnościom pierwszego rzędu nie rodziło niebezpieczeństwa przyjmowania istnienia możliwych indywiduów. Choć jego koncepcja wciąż opierałaby się na konieczności wprowadzenia reguł generowania bądź praw pomostowych, to z uwagi na to, że ostatnie wydają się funk-

cjonować dość systematycznie, stanowią one koszt, który niewątpliwie warto ponieść, gdyż umożliwiają korzystanie z metod pierwszego rzędu przy sprawdzaniu poprawności rozumowań sformułowanych przy użyciu twierdzeń możliwościowych.

Podsumowując, istnieją co najmniej dwie racje, dla których fikcjonalista modalny mógłby przeciwstawić się uogólnieniu strategii unikania zobowiązań ontologicznych względem światów możliwych przez poprzedzanie funktorem możliwości twierdzeń, w których kwantyfikacja egzystencjalna wiąże zmienne przebiegające po aktualnie nieistniejących bytach. Pierwszą z nich jest to, że propozycja fikcjonalisty modalnego pozwala na stosowanie metod pierwszego rzędu w sprawdzaniu poprawności rozumowań sformułowanych przy pomocy twierdzeń możliwościowych, na co nie pozwala już propozycja modalisty. Drugim zaś jest to, że jest on w stanie wskazać na podstawy rozróżnienia indywiduów  $x$  i  $y$  bez akceptowania istnienia możliwych indywiduów, z czym modalista może mieć problemy.

# Zakończenie

Choć fikcjonalizm modalny wydaje się dobrze współgrać z intuicją wielu filozofów, według których dyskurs możliwościowy stanowi jedynie pewien sposób mówienia, który nie opisuje rzeczywistości taką, jaka ona naprawdę jest, to niestety przeprowadzone wyżej rozważania pokazują, że zwolennik tego stanowiska jest w niekorzystnym położeniu. Okazuje się bowiem, że najbardziej obiecująca wersja jego koncepcji, mianowicie umiarkowany fikcjonalizm modalny, ma tę wadę, że obok konieczności zaciągania zobowiązań ontologicznych co do istnienia uniwersaliów najprawdopodobniej pociąga za sobą również istnienie Lewisowskich możliwych indywiduów — z uwagi na konieczność wprowadzenia własności drugiego rzędu bycia możliwe egzemplifikowanym przez wiele indywiduów. Jeśli moja argumentacja w § 4.6. jest słuszna, to szanse na utrzymanie fikcjonalizmu modalnego w jego umiarkowanej wersji są nikłe. Jeśli argumentacja w § 4.5. jest słuszna, również silny fikcjonalizm modalny jest skazany na niepowodzenie. Wydaje się zatem, że projekty fikcjonalistów modalnych w dyskutowanych wyżej postaciach ponoszą fiasko.

Jeśli konkluzja ta jest słuszna, powstaje pytanie, co w takim razie ma zrobić filozof, który nie chcąc zarzucić możliwościowych analiz, wciąż nie może pogodzić się z istnieniem światów możliwych i zamieszkujących je indywiduów? Odpowiedź na to pytanie będzie przedstawiała się odmiennie w zależności od tego, którym intuicjom przypiszemy największą wagę. Jeśli będziemy poszukiwali takiej koncepcji,

która pozwoli na oddanie sprawiedliwości naszym intuicjom modalnym, zgodnie z którymi mogłyby istnieć dwa jakościowo nieodróżnialne obiekty, wówczas — obok realizmu modalnego — użyteczne mogą się okazać erzacyzm lingwistyczny oraz pewna zmodyfikowana wersja umiarkowanego realizmu modalnego.

Pierwsze z tych stanowisk wymaga przyjęcia istnienia niezbyt wiarogodnego bytu, jakim jest idealny język. W jego słowniku występuje nieskończenie wiele nazw, do których następnie redukuje się możliwe indywidua. Nie to jednak jest największą trudnością erzacyzmu lingwistycznego. W celu oddania zależności modalnych za pomocą predykatów występujących w słowniku tego języka trzeba będzie przyporządkować im jakieś znaczenia. Gdy mówimy, że człowiek z konieczności jest ssakiem, zwolennik erzacyzmu lingwistycznego będzie oddawał tę modalną zależność poprzez wskazanie, że w każdym zbiorze zdań, który zawiera zdanie atomowe — jeśli mamy tutaj do czynienia ze zdaniem atomowym — o postaci „*a* jest człowiekiem”, zawarte jest również zdanie atomowe o postaci „*a* jest ssakiem”. Alternatywnie erzacysta może uznać, że jeśli żadne z tych zdań nie jest zdaniem atomowym, to w każdym zbiorze zdań, w którym zawarte są zdania *pociągające* za sobą zdanie o postaci „*a* jest człowiekiem”, zawarte są również zdania *textslpociągające* za sobą zdanie o postaci „*a* jest ssakiem”. W obu wypadkach powstaje pytanie, na mocy czego zachodzą takie relacje pomiędzy predykatami. Naturalna odpowiedź brzmi: zachodzą one na mocy znaczeń przyporządkowanych predykatom, którymi najprawdopodobniej będą własności lub relacje. Z przyjęciem takiego stanowiska nie będą mieli problemu filozofowie pozbawieni nominalistycznych skrupułów.

Ci z kolei, którzy wzdragają się przed przyjęciem istnienia bytów ogólnych, mogą zdecydować się na redukcję możliwych indywiduów np. do zbiorów tropów. Jeśli filozofowie ci dopuszczali istnienie tropów nie przysługujących żadnemu z aktualnie istniejących indywiduów,



najlepszym dla nich rozwiązaniem będzie przeprowadzenie takiej rekonstrukcji stanowiska Plantingi, w której zastępuje się uniwersalia i zbiory uniwersaliów tropami i zbiorami tropów. U podstaw tej strategii leży jednak założenie, że źródłem partykularności tropów nie jest bycie własnością tego a tego indywiduum.

W ramach obu koncepcji można reprezentować istnienie świata dwóch elektronów. Erzacysta lingwistyczny może reprezentować świat dwóch elektronów za pomocą zbioru zdań, w których argumentami tych samych predykatów są dwie odmienne nazwy. Z kolei zwolennik zmodyfikowanego umiarkowanego realizmu może reprezentować elektrony za pomocą odmiennych zbiorów tropów takich, że każdemu tropowi będącemu elementem pierwszego zbioru tropów odpowiada podobny trop będący elementem drugiego zbioru tropów i odwrotnie.

Wyliczone wyżej koncepcje uważam jednak za niewiarygodne. Erzacyzm lingwistyczny wymaga wprowadzenia niewiarygodnego bytu, jakim jest idealny język, w którego słowniku występuje nieskończenie wiele nazw. Z kolei jeśli uznać, że źródłem partykularności tropów nie jest bycie własnością tego a tego indywiduum, wówczas nic nie stoi na przeszkodzie, by uznać, iż indywidua istniejące w naszym świecie również są wiązkami takich tropów. Wtedy jednak ich natura i status ontologiczny będą odpowiadały naturze i statusowi ontologicznemu możliwych indywiduów, które również będą utożsamiane z wiązkami tropów. Czy aby nie stanowi to wystarczającej podstawy, by stanowisko to wywołało niedowierzanie, będące standardową reakcją na modalny realizm Lewisa?

Gdy ontologiczne sumienie nie pozwala pogodzić się z istnieniem Lewisowskich światów możliwych, najlepszym wyjściem — przynajmniej dla filozofów pozbawionych nominalistycznych skrupułów — wydaje się akceptacja jakiegoś stanowiska zbliżonego do umiarkowanego realizmu modalnego proponowanego przez Plantingę. Stanowisko to opiera się na aktualistycznie akceptowalnej ontologii indywidualnych istot

oraz stanów rzeczy, które można traktować jako teoriomnogościowe konstrukcje na indywidualnych istotach i uniwersaliach. Niewątpliwym kosztem wybrania tej opcji będzie to, że niektórym intuicyjnie prawdziwym twierdzeniom modalnym nie będą odpowiadały prawdziwe twierdzenia możliwościowe w ramach koncepcji Plantingi.



# Indeks

- Grice P., 101  
Rosen G., 181  
Armstrong D., 16–18, 22–24, 26,  
29–33, 47, 148, 149, 259  
Azzouni J., 237  
Barcan R., 128–131, 140  
Bennett J., 199, 200  
Black M., 125, 289  
Brock S., 153, 155, 157, 158, 161,  
163, 165, 167, 169  
Burgess P., 146  
Carnap R., 68, 70, 71  
Crimmins M., 206  
Dasgupta S., 291, 293, 294  
Davies M., 268  
Dehaene S., 235  
Eklund M., 298–300, 303  
Field H., 57, 58, 60, 64, 67, 99  
Fine K., 51, 54, 308–314, 316, 317,  
319  
Forrest P., 259  
Frege G., 18, 183–185  
Gendler Szabó Z., 69  
Girle R., 128  
Grice P., 101  
Hale B., 177–179, 181, 182, 185,  
187–189, 191–193, 195, 201,  
202  
Hawthorne J., 154, 165, 278, 280  
Humberstone L., 9, 268  
Jackson F., 61–65  
Joyce R., 103, 104  
Kalderon M. E., 60, 78, 146  
Kaplan D., 206  
Kim S., 208, 266, 267, 270  
Kripke S., 5, 7–9, 68, 221, 265  
Kroon F., 60, 61  
Lewis D., 33, 37–39, 41, 44, 47, 58,  
67, 68, 110, 112–123, 125,  
127, 131–135, 137–141, 147,  
149–153, 156, 157, 159, 174–  
177, 180, 182, 187, 189, 196–  
198, 205, 208, 209, 215, 236,  
251–253, 259, 260, 265, 266,  
296, 305, 324, 326

MacFarlane J., 79, 80	Smith M., 61–65
Martin C. B., 38	Stalnaker R., 80, 82, 175, 205
Merricks T., 22, 23, 27, 34, 35, 41, 44, 45	Stanley J., 272–274
	Strawson P. F., 185
Nolan D., 21, 141–143, 154, 165, 254–257, 262–265, 271	Van Fraassen B., 60, 76, 78
Noonan H., 156–159, 167, 170, 180, 181	Van Inwagen P., 58
	Vander Laan D., 268
Oppy G., 61–65	Wójtowicz K., 91
Parsons T., 58, 236	Walton K., 87, 88, 97, 264
Perry J., 206	Williams B., 84
Plantinga A., 41, 44, 123–127, 212, 275, 288, 289, 312, 326, 327	Woodward R., 163, 164, 167
Priest G., 253	Yablo S., 78, 85, 90–92, 94, 97–100, 103, 108, 143, 271, 274
	Zalta E., 58
Quine W.V., 28, 29, 34, 77, 78, 291	
Rosen G., 48, 58–60, 66, 75, 76, 78, 108–112, 141, 142, 146, 151, 153, 155–158, 161, 163, 165, 167, 169, 179–184, 189, 191, 192, 196, 197, 254, 257, 259, 261–263	
Sainsbury R.M., 60, 214, 217, 221, 222, 225, 227, 228, 232, 233, 242	
Shalkowski S., 153	
Shapiro S., 171, 172	
Sider T., 278, 280	

# Bibliografia

- [–] ARMSTRONG D.  
(1989), *A Combinatorial Theory of Possibility*, Cambridge University Press, Cambridge.  
(2004), *Truth and Truthmakers*, Cambridge University Press, Cambridge.
- [–] AZZOUNI J.  
(2010), *Talking about Nothing. Numbers, Hallucinations, and Fictions*, Oxford University Press, Oxford.
- [–] BARCAN R.  
(1946), “A Functional Calculus of First Order Based on Strict Implication,” *Journal of Symbolic Logic* 11, s. 1–16.
- [–] BENNETT J. (2003)  
*A Philosophical Guide to Conditionals*, Oxford University Press, Oxford.
- [–] BLACK M.  
(1954), *The Identity of Indiscernibles*, w: M. Black, *Problems of Analysis*, Cornell University Press, New York, s. 80–92.
- [–] BROCK S.  
(1993), “Modal Fictionalism: A Response to Rosen”, *Mind* 102, s. 147–150.

- [-] BURGESS P. I ROSEN G.  
(1997), *A Subject With No Object. Strategies for Nominalistic Interpretation of Mathematics*, Oxford University Press, Oxford.
- [-] CARNAP R.  
(1956), *Meaning and Necessity. A Study in Semantics and Modal Logic*, The University of Chicago Press, Chicago.
- [-] CRIMMINS M. I PERRY J.  
(1989), *The Principle and the Phone Booth: Reporting Puzzling Beliefs*, *Journal of Philosophy* 86, s. 685–711.
- [-] DASGUPTA S.  
(2009), “*Individuals: an Essay in Revisionary Metaphysics*”, *Philosophical Studies* 145, s. 35–67.
- [-] DAVIES M. I HUMBERSTONE L.  
(1980), “*Two Notions of Necessity*”, *Philosophical Studies* 38, s. 1–30.
- [-] DEHAENE S.  
(1997), *The Number Sense*, Oxford University Press, Oxford.
- [-] EKLUND M.  
(2005), “*Fiction, Indifference and Ontology*”, *Philosophy and Phenomenological Research* LXXI/3, s. 557–579.
- [-] FIELD H.  
(1980), *Science Without Numbers*, Basil Blackwell, Oxford.  
(1989), *Realism, Mathematics and Modality*, Basil Blackwell, Oxford.

(1994/2001), *Deflationist Views of Meaning and Content*, w: H. Field, *Truth and the Absence of Fact*, Oxford University Press, Oxford 2001, s. 104–140.

[–] FINE K.

(1977/2007), *Prior on the Constraction of Possible Worlds and Instants*, w: (Fine 2007), s. 133–175.

(1985/2007), *Plantinga on the Reduction of Possibilist Discourse*, w: (Fine 2007), s. 176–213.

(1994), “*Essence and Modality*”, w: J. Tomberlin (red.), *Philosophical Perspectives* 8/1994, s. 1–16.

(2001), “*The Question of Realism*”, *Philosophers’ Imprint* 1/1, s. 1–30, URL = <http://quod.lib.umich.edu/p/phimp/3521354.0001.002/1?rgn=publications;view=image;q1=dlps> .

(2003/2007), *The Problem of Possibilia*, w: (Fine 2007), s. 214–231.

(2007), *Modality and Tense*, Oxford University Press, Oxford.

[–] FORREST P. I ARMSTRONG D.

(1984), “*An Argument Against David Lewis’ Theory of Possible Worlds*”, *Australian Journal of Philosophy* 62, s. 164–168.

[–] FREGE G.

(1892/1977), *Sens i znaczenie*, przeł. B. Wolniewicz, w: G. Frege, *Pisma semantyczne*, PWN, Warszawa, s. 60–88.

[–] GENDLER SZABÓ Z.

(2000), *Problems of Compositionality*, Gerland Publishing.

[–] GIRLE R.

(2003), *Possible Worlds*, Acumen, Chesham.



- [-] GRICE P.  
(1989), *Studies in the Way of Words*, Harvard University Press, Cambridge (Massachusetts).
- [-] HALE B.  
(1995), “Modal Fictionalism: A Simple Dilemma”, *Analysis* 55/2, s. 63–67.  
(1995b), “A Desperate Fix”, *Analysis* 55/2, s. 74–81.
- [-] HALE B. I HOFFMAN A.  
(2010) (red.), *Modality. Metaphysics, Logic, and Epistemology*, Oxford University Press, Oxford.
- [-] HAWTHORNE J. I SIDER T.  
(2006), *Locations*, w: J. Hawthorne, *Metaphysical Essays*, Oxford University Press, Oxford.
- [-] HUMBERSTONE L.  
(1981), “From Worlds to Possibilities”, *Journal of Philosophical Logic* 10, s. 313–339.
- [-] JACKSON F., OPPY G. I SMITH M.  
(1994), “Minimalism and Truth-Aptness”, *Mind* 103, s. 287–302.
- [-] JOYCE R.  
(2005), *Moral Fictionalism*, w: (Kalderon 2005), s. 287–313.
- [-] KALDERON M. E.  
(2005) (red.), *Fictionalism in Metaphysics*, Oxford University Press, Oxford.  
(2005b), *Moral Fictionalism*, Oxford University Press, Oxford.

- [-] KAPLAN D.  
(1977/1989), Demonstratives, w: J. Almog, J. Perry i H. Wettstein (red.) Themes from Kaplan, Oxford University Press, New York, s. 481–504.
- [-] KIM S.  
(2005), Modal Fictionalism and Analysis, w: (Kaderon 2005), s. 116–133.
- [-] KRIPKE S.  
(1959), “A Completeness Theorem in Modal Logic”, Journal of Symbolic Logic 24/1, s. 1–14.  
1980/2001), Nazywanie a konieczność, przeł. B. Chwedeńczuk, Aletheia, Warszawa 2001.
- [-] KROON F.  
(2011), “Fictionalism in Metaphysics”, Philosophical Compass 6/11, s. 786–803.
- [-] LEWIS D.  
(1968/1983), Counterpart Theory and Quantified Modal Logic, w: Lewis 1983, s. 26–46.  
(1970/1983), General Semantics, w: (Lewis 1983), s. 189–232.  
(1973), Counterfactuals, Basil Blackwell, Oxford.  
(1973/1986), Counterfactuals and Comparative Possibility, w: D. Lewis, Philosophical Papers, t. II, Oxford University Press, Oxford 1986, s. 3–31.  
(1973/1995), Światy możliwe, przeł. U. Żegleń, w: (Szubka 1995), s. 127–135.  
(1978/1991), Prawda w fikcji, przeł. M. Zięba, w: (Paśniczek 1991), s. 161–189.

(1983), *Philosophical Papers*, t. 1, Oxford University Press, Oxford.

(1983b), “New Work for a Theory of Universals”, *Australian Journal of Philosophy* 64/4, s. 343–377.

(1986), *On the Plurality of Worlds*, Blackwell, Oxford.

(1998/1999), *A World of Truthmakers?*, w: (Lewis 1999), s. 215–220.

(1999), *Papers in Metaphysics and Epistemology*, Cambridge University Press.

(2001), “Truthmaking and Difference-Making”, *Nous* 35/4, s. 602–615.

[–] MACFARLANE J.

(2011), *What is Assertion?*, w: Brown J. i Cappelen H. (red.), *Assertion*, Oxford University Press, Oxford, s. 79–96.

[–] MARTIN C. B.

(1996), “How It Is: Entities, Absence and Voids”, *Australian Journal of Philosophy* 74, s. 57–65.

[–] MERRICKS T.

(2007), *Truth and Ontology*, Oxford University Press, Oxford.

[–] NOLAN D.

(1997), “Three Problems for ‘Strong’ Modal Fictionalism”, *Philosophical Studies* 87, s. 259–275.

(2011), *Modal Fictionalism* (Winter 2011 Edition), E. N. Zalta (red.), URL = <http://plato.stanford.edu/entries/fictionalism-modal/#Inc> .

(2011b), Rosen's Incompleteness Worry, w: (Nolan 2011),  
URL = <http://plato.stanford.edu/entries/fictionalism-modal/supplement2.html> .

- [–] NOLAN D., O'LEARY-HAWTHORNE J.  
(1996), "Reflexive Fictionalism", *Analysis* 56/1, s. 23–32.
- [–] NOONAN H.  
(1994), "In Defence of the Letter of Fictionalism", *Analysis* 54/3, s. 133–139.
- [–] PARSONS T.  
(1991), Meinongowska analiza przedmiotów fikcyjnych, przeł. J. Paśniczek, w: (Paśniczek 1991), s. 137–152.
- [–] PAŚNICZEK J.  
(1991) (red.), *Ontologia fikcji*, Biblioteka Myśli Semiotycznej, Warszawa.
- [–] PLANTINGA A.  
(1974), *The Nature of Necessity*, Clarendon Press, Oxford.
- [–] PRIEST G.  
(2005), *Towards Non-Being. The Semantics and Metaphysics of Intentionality*, Oxford University Press, Oxford.
- [–] QUINE W.V.  
(1948/2000), O tym, co istnieje, przeł. B. Stanosz, w: W.V. Quine, *Z punktu widzenia logiki*, Aletheia, Warszawa 2000.  
(1960), "Variables Explained Away", *Proceedings of the American Philosophical Society* 104/3, s. 343–347.

(1969), *Existence and Quantification*, w: W.V. Quine, *Ontological Relativity and Other Essays*, Columbia University Press, New York 1969, s. 91–113.

(1978/1981), *Postscript on Metaphor*, w: W.V. Quine, *Theories and Things*, Harvard University Press, Cambridge (Massachusetts) 1981, s. 187–189.

[–] ROSEN G.

(1990), “Modal Fictionalism”, *Mind* 99, s. 327–354.

(1993), “A Problem for Fictionalism about Possible Worlds”, *Analysis* 53/2, s. 71–81.

(1995), “Modal Fictionalism Fixed”, *Analysis* 55, s. 67–73.

(2005), *Problems in the History of Fictionalism*, w: (Kalderon 2005), s. 14–64.

(2010), *Metaphysical Dependence*, w: (Hale and Hoffman 2010), s. 109–135.

[–] SAINSBURY R.M.

(2005), *Reference Without Referents*, Clarendon Press, Oxford.

(2010), *Fiction and Fictionalism*, Routledge, London.

[–] SHALKOWSKI S.

(2010), *IBE, GMR, and metaphysical projects*, w: (Hale and Hoffman 2010), s. 169–187.

[–] SHAPIRO S.

(2005b), *Logical Consequence, Proof Theory, and Model Theory*, w: S. Shapiro (red.), *The Oxford Handbook of Philosophy of Mathematics and Logic*, Oxford University Press, Oxford 2005, s. 651–670.

- [–] STALNAKER R.
- (1975/1999), Indicative Conditionals, w: (Stalnaker 1999), s. 63–77.
- (1976), Propositions, w: A. F. MacKay i D.D. Merrill (red.), Issues in the Philosophy of Language, Yale University Press, New Haven, s. 79–91.
- (1976/1995), Światy możliwe, przeł. T. Szubka i U. Żegleń, w: (Szubka 1995), s. 137–154.
- (1978/1999), Assertion, w: (Stalnaker 1999), s. 78–95.
- (1999), Context and Content, Oxford University Press, Oxford.
- [–] STANLEY J.
- (2001), “Hermeneutic Fictionalism”, Midwest Studies of Philosophy XXV, s. 36–71.
- [–] STRAWSON P. F.
- (1950/1967), O odnoszeniu się użycia wyrażeń do przedmiotów, przeł. J. Pelc, w: J. Pelc (red.), Logika i język. Studia z semiotyki logicznej, PWN, Warszawa.
- [–] SZUBKA T.
- (1995) (red.), Metafizyka w filozofii analitycznej, TN KUL, Lublin.
- [–] VAN FRAASSEN B.
- (1980), The Scientific Image, Clarendon Press, Oxford.
- [–] VAN INWAGEN P.
- (1977), “Creatures of Fiction”, American Philosophical Quarterly 14/4, s. 299–308.

- [-] VANDER LAAN D.  
(2004), Counterpossibles and Similarity, w: F. Jackson i G. Priest (red.), *Lewisian Themes. The Philosophy of David K. Lewis*, Oxford University Press, Oxford, s. 258–275.
  
- [-] WALTON K.  
(1990), *Mimesis as Make-Believe. On Foundations of the Representational Arts*, Harvard University Press, Cambridge (Massachusetts).  
(1993/2005), Metaphor and Prop Oriented Make-Believe, w: (Kalderson 2005), s. 67–87.
  
- [-] WARZOSZCZAK P.  
(2009), „Dwa typy modalnego fikcjonalizmu”, *Filozofia Nauki* 65, s. 41–67.  
(2011), „Antyrealizm modalny: instrumentalizm vs. fikcjonalizm”, *Studia Philosophica Wratislaviensia* 1/2011, s. 31–49.  
(2012), „Carnap a współczesne spory ontologiczne, cz. II. Czy późny Carnap był zwolennikiem epistemizmu?”, *Filozofia Nauki*, w druku.
  
- [-] WILLIAMS B.  
(2002), *Truth and Truthfulness. An Essay in Genealogy*, Princeton University Press, Princeton.
  
- [-] WOODWARD R.  
(2008), “Why Modal Fictionalism is not Self-Defeating”, *Philosophical Studies* 139, s. 273–288.
  
- [-] WÓJTOWICZ K.  
(2003), *Spór o istnienie w matematyce*, Semper, Warszawa.

[–] YABLO S.

(1996), “How in the World?”, *Philosophical Topics* 24, s. 255–286.

(2001), “Appendix on Jason’s Critique”, *Midwest Studies in Philosophy* XXV, s. 93–102.

(2001/2010), *Go Figure: A Path through Fictionalism*, w: Yablo 2010, s. 177–199.

(2005), *The Myth of the Seven*, w: (Kalderon 2005), s. 88–115.

(2010), *Things*, Oxford University Press, Oxford.

[–] ZALTA E.

(2000), *Pretense Theory and Abstract Object Theory*, w: Everett A. i Hofweber T. (red.), *Empty Names, Fiction and the Puzzle of Non-Existence*, CSLI Publications, Stanford, s. 117–147.