

Katarzyna Paprzycka

Jak obalić E.J. Lowe’a obalenie teorii identyczności, czyli co by było, gdyby nie wszystkie neurony zachowały się tak, jak się rzeczywiście zachowały?*

Słowa kluczowe: E.J. Lowe, zdania kontrfaktyczne, światy możliwe, alternatywa w poprzedniku, „nie wszystko”, „nie oba”, „byłoby”, „mogłoby”, teoria identyczności, zasada symplifikacji alternatywnych poprzedników

Wyobraźmy sobie, że na wadze umieszczono 42 10-gramowe odważniki (zakładamy, że są to wszystkie dostępne odważniki – więcej ich nie ma). Rzetelna waga wskazała ich łączną masę, tj. 420 g. Co by się stało, gdyby nie wszystkie odważniki zostały umieszczone na wadze?

Sądzę, że nie potrafimy na to pytanie jednoznacznie odpowiedzieć. Wszystko zależy od tego, ile dokładnie odważników zostanie umieszczonych na wadze. Bez tej wiedzy nie mamy podstaw, by wskazać jednoznacznie którąkolwiek z czterdziestu dwu możliwości: mogłoby być tak, że umieszczonych zostaje 41 odważników, 40 odważników... itd., a mogłoby też być tak, że położony zostaje 1 odważnik bądź żaden. Jeżeli tak, to odpowiedź na postawione pytanie będzie brzmiała:

- (1) Gdyby nie wszystkie odważniki zostały umieszczone na wadze, to wskazałaby ona 0 g, 10 g, ... lub 410 g.

* Projekt został sfinansowany ze środków Narodowego Centrum Nauki przyznanych na podstawie decyzji numer DEC–2012/05/B/HS1/02949.

Uznanie tezy 1 wystarcza – jak zobaczymy – aby obalić E.J. Lowe’a obalenie teorii identyczności.

W § 1 przedstawimy argumentację Lowe’a. Gdy pewna osoba unosi rękę, ruch ręki ma przyczynę mentalną (intencję) i fizyczną (neuralną), które – według teorii identyczności – są ze sobą tożsame. Lowe broni dwóch tez. Według pierwszej z nich, niewystąpienie przyczyny mentalnej prowadzioby do niewystąpienia ruchu ręki. Według drugiej, niewystąpienie przyczyny neuralnej prowadzioby i tak do uniesienia ręki, tyle że w nieco inny sposób. Lowe wyciąga stąd wniosek, że przyczyna mentalna nie jest identyczna z przyczyną neuralną, a także że przyczyny mentalne są odpowiedzialne za to, że w ogóle działamy, natomiast przyczyny neuralne determinują już tylko sposób, w jaki działanie jest egzekwowane.

W tekście skoncentrujemy się na Lowe’a uzasadnieniu tezy drugiej¹. Jedy-
nym jego argumentem na rzecz tej tezy jest zastosowanie semantyki światów
możliwych (Lewis 1973). Dlatego też niniejszy tekst nie będzie – wbrew pozorom – dotyczył neurobiologii, ale właśnie zastosowania semantyki światów
możliwych do zdań podobnych do tezy Lowe’a (§ 2). Okazuje się bowiem, że
dla takich zdań semantyka światów możliwych generuje warunki prawdziwości
odmienne od tych, które bylibyśmy skłonni tym zdaniom przypisać (§ 3–4).
Problem ten wydaje się szczególnym przypadkiem tzw. problemu poprzedni-
ków o kształcie alternatywy (§ 5).

W jednym z poprzednich artykułów (Paprzycka 2013) pokazywałam, że
uzasadnienie tezy Lowe’a odwołujące się wyłącznie do werdyktu semantyki
światów możliwych należy podważyć ze względów metodologicznych. Jeżeli
wiemy, że najlepsza teoria daje niewiarygodne wyniki w warunkach typu X ,
to nie należy tej teorii stosować do przewidywania, co się stanie w warunkach
 x_i ($x_i \in X$). W § 6–7 przyjrzymy się dwóm głównym strategiom odpowiedzi
na problem poprzedników o kształcie alternatywy. Zobaczymy, że przyjmując
każdą z tych strategii, można argumentować, że teza Lowe’a jest fałszywa.
Wystarcza to w każdym razie do uchylecia uzasadnienia dla jego kontrower-
syjnej tezy i do odrzucenia jego argumentu przeciw teorii identyczności (§ 8).

1. Lowe’a argument przeciw teorii identyczności

Według zwolenników teorii identyczności (zarówno teorii identyczności jed-
nostkowej, jak i teorii identyczności rodzajowej), zdarzenie (egzemplarz) D

¹ Całość argumentacji Lowe’a w jej szerszym kontekście jest przedstawiona w: Paprzycka 2013. Zarysowany jest tam problem z argumentacją na rzecz tezy o niewystąpieniu przyczyny mentalnej (por. też mój artykuł *Lowe’s Argument Against the Psycho-Neural Token-Identity Thesis*”, w druku). Niniejszy artykuł stanowi rozwinięcie i pogłębienie zarysowanego wcześniej problemu.

(decyzja podmiotu, by podnieść rękę) jest tożsame z pewnym złożonym zdarzeniem (egzemplarzem) neuronalnym N . Argumentacja Lowe'a (2006, 2008) ma wykazać prawdziwość następujących zdań kontrfaktycznych:

(N) Gdyby zdarzenie N nie wystąpiło, to *ręka podmiotu uniosłaby się prawie dokładnie tak samo, jak to rzeczywiście miało miejsce* (Lowe 2008: 105; podkreślenie autora).

(M) Gdyby zdarzenie D nie wystąpiło (...), to *ręka podmiotu w ogóle by się nie uniosła* (Lowe 2008: 105; podkreślenie autora).

Jeżeli tezy N i M są prawdziwe, to zdarzenia D i N nie mogą być tożsame. Lowe w ten sposób przedstawiłby potężny argument przeciwko teoriiom identyczności (w tym przeciwko teoriiom identyczności jednostkowej).

Przyjrzyjmy się argumentacji na rzecz tezy N. Lowe przyjmuje, że zachodzi relacja przyczynowa między pewnym złożonym zdarzeniem (egzemplarzem) neuronalnym N a pewnym zdarzeniem (egzemplarzem) behawioralnym B (tu: uniesieniem ręki). Prawdziwe jest więc następujące zdanie kontrfaktyczne:

Gdyby N nie wystąpiło, to nie wystąpiłoby B .

Według standardowej interpretacji zdań kontrfaktycznych (Lewis 1973) zdanie to jest prawdziwe zawsze i tylko wtedy, gdy we wszystkich najbliższych światach możliwych, gdzie N nie występuje, nie występuje też B . Aby dowieść prawdziwości zdania N, Lowe musiałby pokazać, że we wszystkich tych światach występuje zdarzenie (egzemplarz) podobny do B , tj. również będący uniesieniem ręki, choć charakteryzujący się nieznacznie odmiennymi parametrami.

Lowe uważa, że może tego dowieść, odwołując się do relacji bliskości między światami, narzuconej przez standardową interpretację zdań kontrfaktycznych. Lowe twierdzi, że w najbliższych światach możliwych, gdzie nie występuje N , wystąpi inne zdarzenie neuronalne N^* , które musi być *bardzo* podobne do zdarzenia rzeczywistego N^2 . Przy takim rozumieniu relacji bliskości pozostałe kroki nie budzą większych zastrzeżeń. Ponieważ N różni się od N^* tylko aktywnością co najwyżej paru neuronów, to – przy założeniu masywnego rozproszenia aktywności neuronalnej – N^* będzie się charakteryzowało podobnymi (do N) mocami przyczynowymi, a w związku z tym również doprowadzi do uniesienia ręki B^* , które różnić się będzie od rzeczywistego uniesienia ręki B jedynie w sposób drobiazgowy. Lowe wyprowadza stąd wniosek, że gdyby zdarzenie neuronalne N , które spowodowało uniesienie

² „Dowolny świat możliwy, w którym występuje zdarzenie neuronalne różniące się od N *bardziej* niż w ten minimalny sposób, po prostu nie kwalifikuje się jako *najbliższy* świat możliwy, w którym N nie występuje” (Lowe 2008: 105).

ręki B , nie wystąpiło, to ręka tak czy owak by się uniosła – choć jej uniesienie mogłoby charakteryzować nieco inne tempo, trajektoria, czas itd.

2. Struktura logiczna tezy N

Teza N odnosi się do bardzo złożonego zdarzenia-egzemplara N . Warto ją jednak sformułować tak, aby jej złożoność była widoczna. Zdarzenie N jest „sumą (...) ogromnej liczby pojedynczych zdarzeń neuronalnych” (Lowe 2008: 105). Przyjmijmy, że te pojedyncze zdarzenia neuronalne to jednostkowe zdarzenia N_1, \dots, N_k . Jeżeli tak, to konieczne jest, że N występuje zawsze i tylko wtedy, gdy występują wszystkie zdarzenia składowe N_1, \dots, N_k . W takim razie teza N wydaje się równoważna tezie:

Gdyby nie wszystkie zdarzenia N_1, \dots, N_k wystąpiły (gdyby nie wszystkie neurony zachowywały się tak, jak się rzeczywiście zachowały), to ręka podmiotu uniosłaby się w prawie dokładnie taki sam sposób, jak to rzeczywiście miało miejsce.

Tezę tę można podeprzeć analogicznym rozumowaniem. Najbliższe światy możliwe, w których spełniony jest poprzednik, tj. gdzie nie wszystkie z tych neuronów zachowują się tak, jak się rzeczywiście zachowują, to światy, gdzie *prawie* wszystkie z tych neuronów (z wyjątkiem jednego czy dwóch) zachowują się tak, jak się rzeczywiście zachowują. Można zasadnie sądzić, że odmienna aktywność jednego czy dwóch z tych neuronów nie wpłynie zasadniczo na to, czy ręka się uniesie, a co najwyżej sprawi, że uniesienie ręki będzie się charakteryzowało nieco innymi parametrami.

3. Kontrprzykłady

Zastosujmy semantykę światów możliwych w powyższy sposób do przykładu, od którego zaczęliśmy.

Przykład 1. Na wadze umieszczone zostały wszystkie dostępne (czterdzieści dwa) odważniki 10-gramowe. Co by było, gdyby nie wszystkie odważniki zostały umieszczone na wadze?

Najbliższe światy możliwe, w których nie wszystkie odważniki zostają umieszczone na wadze, to światy, w których prawie wszystkie odważniki (z wyjątkiem jednego czy dwóch) zostają na niej położone. W takim jednak wypadku waga wskaże 410 g lub 400 g, a w każdym razie wartość zbliżoną do 420 g.

(1#) Gdyby nie wszystkie odważniki zostały umieszczone na wadze, to wskazałyby ona 410 g lub 400 g lub wartość zbliżoną do 420 g.

Mamy wszelkie podstawy, by przyjąć twierdzenie słabsze:

(1') Gdyby nie wszystkie odważniki zostały umieszczone na wadze, to waga *mogłaby* wskazać 410 g lub 400 g lub wartość zbliżoną do 420 g.

ale twierdzenie 1# jest zbyt silne.

Przykład 2. Załóżmy, że Lena ma alergię na sierść. Tak się akurat składa, że jej dwie najbliższe przyjaciółki, Asia i Beata, dostają w prezencie zwierzaka: Asia dostaje kota, a Beata – psa. Ponieważ Lena spędza z nimi dużo czasu, występuje u niej silna reakcja alergiczna. Co by się jednak stało, gdyby nie obie przyjaciółki dostały w prezencie swojego zwierzaka (tj. gdyby nie było prawdą, że obie przyjaciółki dostały taki prezent)?

Założmy, że ktoś byłby skłonny odpowiedzieć, że:

(2#) Gdyby nie było prawdą, że obie, Asia i Beata, dostały swojego zwierzaka, to u Leny wystąpiłaby i tak silna reakcja alergiczna.

Wydaje się, że powinniśmy temu komuś przypomnieć, że zdania o kształcie „nieprawda, że zarówno p i q ” dopuszczają sytuację, w której „ani nie p , ani nie q ” jest prawdziwe. Innymi słowy, pytanie dopuszcza sytuację, w której żadna z przyjaciółek nie dostaje swego czworonoga. W takim wypadku, *ceteris paribus*, reakcja alergiczna u Leny by nie wystąpiła. Na pewno mogłoby być tak, że u Leny i tak wystąpi reakcja alergiczna – w sytuacji, w której tylko jedna z koleżanek otrzymuje zwierzę. To znaczy jednak, że przyjmujemy tezę słabszą, a mianowicie:

(2) Gdyby nie było prawdą, że obie, Asia i Beata, dostały swego zwierzaka, to u Leny *mogłaby* i tak *wystąpić* silna reakcja alergiczna.

Przyjmujemy też, że:

(2') Gdyby nie było prawdą, że obie, Asia i Beata, dostały swego zwierzaka, to u Leny *mogłaby nie wystąpić* silna reakcja alergiczna.

Zdanie 2# zatem odrzucamy.

Stosując semantykę światów możliwych³, otrzymamy jednak twierdzenie 2# raczej niż 2 i 2'. *Ceteris paribus*, światy, w których tylko jedna z przyjaciółek otrzymuje swego czworonoga, są bliższe światu rzeczywistego niż światy, w których żadna z przyjaciółek nie dostaje zwierzaka. Stąd w najbliższych światach możliwych, w których zwierzątko jest podarowane jednej z przyjaciółek, u Leny pojawią się objawy alergiczne.

Werdykt semantyki światów możliwych odbiega zatem od naszych intuicji.

Przykłady 3 i 4. Generalnie wydaje się, że takie zastosowanie semantyki światów możliwych prowadzić będzie jak gdyby do systematycznego zniekształcenia logicznej struktury poprzednika o kształcie negacji koniunkcji. To zniekształcenie uchwytują następujące dwa zdania:

(3#) Gdyby nie wszystko było tak, jak rzeczywiście jest, to prawie wszystko byłoby tak, jak rzeczywiście jest.

(4#) Gdyby nie wszystko różniło się od tego, jak rzeczywiście jest, to wszystko byłoby takie, jak rzeczywiście jest.

Zdanie 4# jest równoważne zdaniu:

(4'#) Gdyby coś było takie, jak rzeczywiście jest, to wszystko byłoby takie, jak rzeczywiście jest.

Warunki prawdziwości zdania 3# na gruncie semantyki światów możliwych są analogiczne do warunków prawdziwości generowanych przez tę semantykę dla dotychczas przytaczanych zdań. Najbliższe światy możliwe, które spełniają warunek z poprzednika „nie wszystko jest tak, jak rzeczywiście jest”, to światy, w których prawie wszystko jest tak, jak jest rzeczywiście.

Zdania 4# i 4'# są natomiast prawdziwe w świetle rozważanej semantyki dlatego, że świat rzeczywisty jest wśród światów spełniających poprzednik „nie wszystko różni się od tego, jak rzeczywiście jest”, a także wśród światów spełniających poprzednik „coś jest takie, jakie rzeczywiście jest”. Świat rzeczywisty jest najbliższy światu rzeczywistego. Stąd w najbliższych światach, w których spełniony jest poprzednik, wszystko jest tak, jak w świecie rzeczywistym. Zilustrujmy to jeszcze dwoma przykładami.

³ Dalej zakładam, że stosujemy semantykę światów możliwych w sposób, w który to czyni Lowe. Dopuszczam możliwość, że można byłoby użyć jakiegoś sposobu jej zastosowania (np. poprzez odpowiednie skonstruowanie funkcji selekcji światów), by nie generowała takich rezultatów. Por. też Paprzycka (prace w druku).

Przykład 5. Czarek i Danka zdali egzamin z logiki. Co by było, gdyby nie oboje oblali egzamin z logiki? Innymi słowy, co by było, gdyby Czarek lub Danka zdali egzamin z logiki? Wydaje się dość intuicyjne odpowiedzieć na to pytanie w następujący sposób: gdyby nie oboje oblali logikę, to albo tylko jedno z nich by jej nie oblało (a więc logikę zdało), albo żadne by jej nie oblało (oboje by zdali). Równoważnie: gdyby Czarek lub Danka zdali logikę, to albo tylko jedno z nich by ją zdało, albo oboje by ją zdali.

Jednakże światy, w których Czarek i Danka oboje zdają logikę, jest zdecydowanie najbliższy światu rzeczywistego, w którym przecież oboje logikę zdali. A w takim razie w świetle semantyki światów możliwych na powyższe pytanie powinniśmy odpowiedzieć zgodnie z kształtem twierdzeń 4# i 4'#:

- (5#) Gdyby nie oboje Czarek i Danka oblali egzamin z logiki, to oboje by egzamin zdali.
(5'#) Gdyby Czarek zdał egzamin z logiki lub gdyby Danka egzamin ten zdała, to oboje by go zdali.

Ponownie 5# i 5'# wydają się ignorować pewne z dopuszczonych przez poprzednik możliwości.

Przykład 6. Na wadze położono wszystkie dostępne (czterdzieści dwa) odważniki 10-gramowe. Co by było, gdyby na wadze położono przynajmniej jeden z tych odważników? Jasne jest, że waga wskazałaby 10 g, 20 g, ... lub 420 g.

Jednak najbliższym światem spełniającym poprzednik „przynajmniej jeden odważnik położono na wadze” jest świat rzeczywisty, w którym na wadze położono wszystkie odważniki. Stąd:

- (6'#) Gdyby na wadze położono przynajmniej jeden z 42 odważników, to waga wskazałaby 420 g.

Reasumując, we wszystkich tych zdaniach poprzednik zachęca do rozważenia rozmaitych możliwości, których rozważenie wydaje się być zablokowane przez wymóg uwzględnienia tylko najbliższych światów możliwych.

4. Zdania kontrfaktyczne z zanegowaną koniunkcją w poprzedniku a semantyka światów możliwych

Wspomniane w powyższych przykładach rozbieżności nie są przypadkowe. Rozważaliśmy zdania kontrfaktyczne z zanegowaną koniunkcją w poprzedniku. W niektórych przykładach (1, 2, 3) człony tej koniunkcji były prawdziwe w świecie rzeczywistym, a w innych (4, 5, 6) były one fałszywe. We wszyst-

kich przykładach intuicyjna ocena zdań generowała wyniki odmienne od tych, które wynikały z zastosowania semantyki światów możliwych. Przyjrzyjmy się strukturze tych przykładów.

4.1. Gdy człony koniunkcji są prawdziwe

Założmy, że w świecie rzeczywistym prawdą jest, że p , q i r . Rozważane zdania kontrfaktyczne mają kształt:

(7) Gdyby zaszło $\sim(p \wedge q)$, to zaszłoby $\sim r$.

Otóż negacja koniunkcji „ $\sim(p \wedge q)$ ” będzie prawdziwa w trzech sytuacjach:

(w_{1a}) p jest prawdziwe, a q – fałszywe;

(w_{1b}) p jest fałszywe, a q – prawdziwe;

(w₂) p i q jest fałszywe.

Wydawać by się mogło – a nasze intuicje dotyczące powyższych przykładów zdają się to potwierdzać – że rozważenie wartości logicznej zdań o kształcie 7 wymaga odwołania się do wszystkich trzech możliwych sytuacji. Powyższy sposób zastosowania semantyki światów możliwych jednak nie pozwala w ogóle na rozważenie w₂-światów. *Ceteris paribus*, w₁-światy są bliższe światu rzeczywistego niż w₂-światy. Ponieważ w₂-światy nie są najbliższe światu rzeczywistego, więc to, co się w nich dzieje, nie będzie w ogóle brane pod uwagę.

Takie zastosowanie semantyki światów możliwych będzie się wiązać z systematycznym nieuwzględnieniem pełnej struktury logicznej zdań kształtu 7. Jeżeli traktujemy jako najbliższe światu rzeczywistego tylko w₁-światy, a wykluczamy w₂-światy, to zbiór światów najbliższych, stanowiących podstawę oceny wartości logicznej zdania kontrfaktycznego typu 7, w pełni pokryje się ze światami najbliższymi (i oceną wartości logicznej) dla zdań o formie:

(8) Gdyby zaszło $\sim(p \equiv q)$, to zaszłoby $\sim r$.

Dzieje się tak dlatego, że negacja równoważności „ $\sim(p \equiv q)$ ” jest prawdziwa dokładnie w w₁-światach. Światy najbliższe (w powyższym rozumieniu), spełniające poprzedniki zdań 7 i 8, pokrywają się. To jednak znaczy, że ocena wartości logicznej tych zdań na gruncie rozważanej semantyki zupełnie pomija różnicę między nimi i koncentruje się na części wspólnej.

Warto przy okazji zwrócić uwagę, że negacja równoważności dwóch zdań jest logicznie równoważna *alternatywie rozłącznej* negacji tych zdań, natomiast

negacja koniunkcji dwóch zdań jest logicznie równoważna *alternatywie nierozłącznej* negacji tych zdań. Innymi słowy, forma zdaniowa 7 jest równoważna formie 7', a forma 8 – formie 8'.

(7') Gdyby zaszło $\sim p \vee \sim q$, to zaszłoby $\sim r$.

(8') Gdyby zaszło $(\sim p \vee \sim q) \wedge \sim(\sim p \wedge \sim q)$, to zaszłoby $\sim r$.

Przy omawianym rozumieniu światów najbliższych wartość logiczna zdań o kształcie 7' i 8' będzie taka sama, o ile p i q są prawdziwe w świecie rzeczywistym.

Gdy negowana koniunkcja jest wieloczłonowa (zakładamy ponownie, że wszystkie jej człony są prawdziwe), jeszcze więcej światów będzie ignorowanych.

(9) Gdyby zaszło $\sim(p_1 \wedge \dots \wedge p_k)$, to zaszłoby $\sim r$.

Koniunkcja ta będzie fałszywa w w_1 -światach (gdzie jeden z jej członów jest fałszywy), w w_2 -światach (gdzie dwa z jej członów są fałszywe), ... oraz w w_k -światach (gdzie wszystkie jej człony są fałszywe). Najbliższe światu rzeczywistego są w_1 -światy, a więc tylko one będą stanowić podstawę do określenia wartości logicznej zdania 9 w tej sytuacji.

Problemy powstają, gdy w niektórych z pomijanych sytuacji następnik ma inną wartość logiczną niż w światach wyłonionych jako najbliższe. Tak się dzieje w przykładach 1#, 2#, 3#, a także w przypadku tezy N. Nawet jeżeli w światach, w których tylko jeden czy dwa neurony zachowują się inaczej niż w świecie rzeczywistym, ręka i tak się uniesie, to w światach, w których dostatecznie wiele neuronów zachowuje się inaczej, ręka się nie uniesie.

4.2. Gdy człony koniunkcji są fałszywe

Założmy, że w świecie rzeczywistym prawdą jest, że p , q i r . Rozważmy zdanie kontrfaktyczne:

(10) Gdyby zaszło $\sim(\sim p \wedge \sim q)$, to zaszłoby r .

Jest ono równoważne zdaniu:

(10') Gdyby zaszło $p \vee q$, to zaszłoby r .

Otóż poprzednik tych zdań będzie prawdziwy w trzech sytuacjach:

- (w₁) p i q jest prawdziwe;
(w_{2a}) p jest prawdziwe, a q – fałszywe;
(w_{2b}) p jest fałszywe, a q – prawdziwe.

W tym wypadku będą brane pod uwagę tylko w₁-światy, bo to one są najbliższe świata rzeczywistego, w którym – zgodnie z założeniem – prawdą jest, że p i q . W takim jednak wypadku r również zajdzie, dokładnie tak, jak zachodzi w świecie rzeczywistym. Ignorowane będą w₂-światy. Gdyby r było w tych światach fałszywe, to – wbrew naszym intuicjom – nie wpłynęło to na ocenę wartości logicznej zdań 10 i 10'. Gdy negowana koniunkcja jest wieloczłono-
wa, to jeszcze więcej światów będzie ignorowanych.

W przykładach 4#, 5# i 6# mamy do czynienia ze zdaniami, których poprzedniki wyznaczają jako najbliższe właśnie te światy, w których następnik ma dokładnie tę samą wartość logiczną co w świecie rzeczywistym, a w nieco bardziej odległych światach następnik już różni się wartością logiczną. Uznajemy 4#–6# za zdania fałszywe, ponieważ intuicyjnie wyczuwamy istotność tego, co się dzieje w światach możliwych, w których spełniony jest poprzednik w sposób nieco bardziej odległy od świata rzeczywistego.

Ponownie zastosowanie semantyki światów możliwych wiąże się ze systematycznym nieuwzględnieniem pełnej struktury logicznej zdań tego kształtu.

5. Problem poprzedników o kształcie alternatywy

Werdykt semantyki światów możliwych w stosunku do zdań kontrfaktycznych, których poprzedniki są zanegowanymi koniunkcjami, systematycznie odbiega od naszych intuicji. Wydaje się, że można to rozumieć po prostu jako jeszcze jeden przejaw starego problemu poprzedników o kształcie alternatywy. W końcu negacja koniunkcji dwóch zdań jest równoważna alternatywie negacji tych zdań⁴.

Wcześni krytycy semantyki światów możliwych (Creary, Hill 1975; Ellis i in. 1977; Fine 1975; Nute 1975) zwracali uwagę, że nie oddaje ona dobrze naszych intuicji dotyczących zdań kontrfaktycznych, których poprzednikami są alternatywy. Rozważmy następujące zdanie:

- (11#) Gdyby w czasie II Wojny Światowej Hiszpania była sojusznikiem Ameryki lub Niemiec, to Hitler byłby zadowolony.

⁴ Bodaj pierwszym autorem, który zwrócił uwagę również na zdania kontrfaktyczne z poprzednikami o kształcie negacji koniunkcji, był Donald Nute (1984b). Sugerował on, że zasada symplifikacji poprzedników (por. § 5.1) wydaje się zachodzić również dla tych zdań.

Uznalibyśmy to zdanie za fałszywe. Nawet jeżeli zgodzimy się, że Hitler byłby zadowolony, gdyby Hiszpania, która faktycznie zachowała neutralność, była sojusznikiem Niemiec, to trudno się zgodzić, że Hitler byłby zadowolony, gdyby była ona sojusznikiem Ameryki. Jednakże najbliższe światy możliwe, gdzie spełniona jest alternatywa „Hiszpania jest sojusznikiem Ameryki lub Niemiec”, to światy, w których Hiszpania jest sojusznikiem Niemiec (Hiszpania zachowała co prawda neutralność, ale było jej dużo bliżej do państw Osi niż do Aliantów). Zatem w najbliższych światach możliwych, gdzie spełniona jest ta alternatywa, Hitler jest zadowolony. Tutaj też werdykt semantyki światów możliwych różni się od naszych intuicji.

Problem ten można wyraziście zilustrować w nawiązaniu do naszego przykładu 1:

(12#) Gdyby na wadze umieścić 1 lub 41 odważników 10-gramowych, to waga wskazałaby 410 g.

Zdanie to jest ewidentnie fałszywe – gdyby na wadze umieścić 1 odważnik lub 41 odważników 10-gramowych, to waga wskazywałaby 10 g lub 410 g. Jednakże *najbliższe* światy możliwe, w których spełniona jest alternatywa „na wadze umieszczono 1 lub 41 odważników”, to światy, w których na wagę położono 41 odważników, a w tych światach waga istotnie pokazuje 410 g.

Zwróćmy przy okazji uwagę, że nie ma problemu niezgodności naszych intuicji z werdyktem semantyki światów możliwych w przypadku zdań kontrfaktycznych, których poprzednikami są *człony* wspomnianej alternatywy:

Gdyby na wadze umieścić 1 odważnik 10-gramowy, to waga wskazałaby 10 g.

Gdyby na wadze umieścić 41 odważników 10-gramowych, to waga wskazałaby 410 g.

Podobnie jest w innych przykładach.

Okazuje się, że nie ma łatwego sposobu rozwiązania tego problemu⁵. Powyższe przykłady skłaniają do przyjęcia zasady symplifikacji alternatywnych poprzedników (§ 5.1), która mogłaby również posłużyć do poradzenia sobie z powyższymi kontrprzykładami (§ 5.2). Jednakże przyjęcie tej zasady prowadzi też do poważnych problemów (§ 5.3).

5.1. Zasada symplifikacji alternatywnych poprzedników

Jednym z pierwszych zaproponowanych rozwiązań problemu poprzedników o kształcie alternatywy (Fine 1975; Nute 1975; Ellis i in. 1977) była propo-

⁵ Dobry przegląd proponowanych rozwiązań znajduje się w Nute 1984a oraz Nute, Cross 2001.

zycja dodania do semantyki światów możliwych zasady symplifikacji alternatywnych poprzedników (*Simplification of Disjunctive Antecedents* – SDA):

$$(SDA) [(p \vee q) \Box \rightarrow r] \rightarrow [(p \Box \rightarrow r) \wedge (q \Box \rightarrow r)]$$

Rozważmy zdanie 12#: Gdyby na wadze umieścić 1 odważnik (O_1) lub 41 odważników 10-gramowych (O_{41}), to waga wskazałaby 410 g (W_{410}).

$$(12\#) (O_1 \vee O_{41}) \Box \rightarrow W_{410}$$

Jeżeli przyjmiemy zasadę SDA, to konsekwencją przyjęcia tezy 12# byłoby zdanie ewidentnie fałszywe: $O_1 \Box \rightarrow W_{410}$. Nie jest wszakże prawdą, że gdyby na wadze umieścić jeden odważnik 10-gramowy, to waga wskazałaby 410 g. Jeżeli przyjmiemy SDA, to możemy zastosować *modus tollens*, aby wykazać, że 12# jest zdaniem fałszywym, co jest zgodne ze zgłaszanymi intuicjami.

5.2. Problem poprzedników o kształcie zanegowanej koniunkcji

Dodanie zasady symplifikacji alternatywnych poprzedników rozwiązywałoby też problem poprzedników o kształcie negacji koniunkcji, ale tylko pod warunkiem odpowiedniego przekształcenia zdania wyjściowego. Rozważmy najprostszy przykład 2. Przedstawimy zdanie 2# jako:

$$(2\#) \sim(A \wedge B) \Box \rightarrow L$$

Negacja koniunkcji $\sim(A \wedge B)$ jest logicznie równoważna obu następującym alternatywom:

- (i) $\sim A \vee \sim B$
- (ii) $(A \wedge \sim B) \vee (\sim A \wedge B) \vee (\sim A \wedge \sim B)$

Otóż tylko jeżeli dokonamy podstawienia drugiego z tych zdań, to nasze intuicje dotyczące przykładu 2 zostaną oddane właściwie⁶.

Gdy w poprzedniku zdania 2# podstawimy pierwszą z powyższych alternatyw (i), otrzymujemy:

$$(2^i\#) (\sim A \vee \sim B) \Box \rightarrow L$$

Stosując zasadę SDA do 2ⁱ# otrzymujemy zdania:

⁶ Nute (1984b) rozważa tylko pierwszą z tych możliwości.

$\sim A \Box \rightarrow L$ (Gdyby Asia nie otrzymała kota, to u Leny wystąpiłaby reakcja alergiczna.)

$\sim B \Box \rightarrow L$ (Gdyby Beata nie otrzymała psa, to u Leny wystąpiłaby reakcja alergiczna.)

Można uznać, że w sytuacjach, o których mowa, oba zdania są prawdziwe. Naturalną interpretacją pierwszego z nich jest odwołanie się do światów, w których Beata dostaje psa (tak jak w świecie rzeczywistym), a w których Asia nie otrzymuje kota. W tych światach u Leny wystąpi reakcja alergiczna. Podobnie jest w wypadku zdania drugiego. Takie zastosowanie zasady SDA nie sugeruje, by istniał jakikolwiek problem z $2^{ii\#}$ czy z $2\#$.

Problem się ujawni dopiero, gdy dokonamy podstawienia drugiej z powyższych alternatyw (ii) w poprzedniku $2\#$, otrzymując:

$$(2^{ii\#}) [(A \wedge \sim B) \vee (\sim A \wedge B) \vee (\sim A \wedge \sim B)] \Box \rightarrow L$$

Gdy zastosujemy zasadę SDA do $2^{ii\#}$ otrzymamy m.in. zdanie: $(\sim A \wedge \sim B) \Box \rightarrow L$. Ponieważ uznajemy je za fałszywe (nie jest prawdą, że gdyby ani Asia ani Beata nie dostały zwierzaka, to u Leny powstałaby reakcja alergiczna), to o ile przyjmujemy zasadę SDA, to powinniśmy uznać $2^{ii\#}$, a więc również $2\#$, za zdania fałszywe, co jest zgodne ze zgłaszanymi intuicjami.

5.3. Status zasady SDA

Kłopot polega na tym, że wykazano, iż dodanie zasady symplifikacji alternatywnych poprzedników do semantyki światów możliwych spowodowałoby funktor $\Box \rightarrow$ do funktora implikacji ścisłej (por. np. Loewer 1976), co prowadziłoby do utraty pewnych pożądaných własności zdań kontrfaktycznych, a w szczególności ich niemonotoniczności. Stosując zasadę podstawiania zdań logicznie równoważnych oraz zasadę symplifikacji poprzednika, otrzymamy zasadę wzmocnienia poprzednika $(p \Box \rightarrow r) \rightarrow ((p \wedge q) \Box \rightarrow r)$, którą powszechnie uznaje się za nieobowiązującą dla zdań kontrfaktycznych. (Można jednocześnie utrzymywać zdania: „Gdyby potrzeb zapalkę o pudełko, to by się zapaliła” oraz „Gdyby potrzeb mokrą zapalkę o pudełko, to by się nie zapaliła”). Aby wyprowadzić zasadę wzmocnienia poprzednika, wystarczy w zdaniu $p \Box \rightarrow r$ podstawić w poprzedniku zdanie logicznie równoważne, otrzymując $(p \vee (p \wedge q)) \Box \rightarrow r$. Po zastosowaniu zasady SDA uzyskuje się niepożądany wniosek: $(p \wedge q) \Box \rightarrow r$.

Niemonotoniczność jest utracona przez jednoczesne zastosowanie zasady symplifikacji SDA oraz zasady podstawiania. W tej sytuacji dostępne są trzy główne strategie odpowiedzi. Pierwszą z nich jest przyjęcie zasady podstawiania, natomiast odrzucenie zasady SDA (§ 6). Drugą strategią jest przyjęcie zasady SDA, ale ograniczenie zasady podstawiania (§ 7). Ostatnim rozwiąza-

niem jest bardziej radykalne odejście od semantyki światów możliwych (np. Alonso-Ovalle 2009; Briggs 2012; Fine 2012). Uwzględnienie tych alternatywnych podejść wykracza poza ramy tego tekstu, aczkolwiek warto nadmienić, że na gruncie każdej z tych propozycji daje się uchwycić naszą intuicyjną ocenę wartości logicznej kontrprzykładów z § 3.

6. Odrzucenie SDA i strategia głębokiej symbolizacji

Mimo że zasada SDA wydaje się wysoce intuicyjna, to wskazano na pewne kontrprzykłady dla niej (McKay, van Inwagen 1977). Nawiązując do przytoczonego wyżej przykładu 12#, rozważmy następujące zdanie kontrfaktyczne (pamiętamy, że Hiszpania pozostała państwem neutralnym):

(13) Gdyby w czasie II Wojny Światowej Hiszpania weszła w sojusz czy to z Aliantami czy z państwami Osi, to weszłaby w sojusz z Państwami Osi.

Zdanie to wydaje się nam prawdziwe. Gdybyśmy jednak przyjęli zasadę SDA, to moglibyśmy ze zdania 13 wyprowadzić zdanie:

(14) Gdyby w czasie II Wojny Światowej Hiszpania weszła w sojusz z Aliantami, to weszłaby w sojusz z Państwami Osi.

Zdanie 14 wydaje się fałszywe, a zatem stosując *modus tollens* musieliśmybyśmy uznać zdanie 13 za fałszywe.

McKay i van Inwagen radzą, by odrzucić zasadę SDA.

6.1. Strategia głębokiej symbolizacji

Według niektórych autorów (np. Fine 1975; Loewer 1976; McKay, van Inwagen 1977) intuicyjność zasady SDA można wytłumaczyć tym, że powierzchniowa struktura niektórych zdań kontrfaktycznych nie odpowiada ich strukturze głębokiej. Tak jak nie powinniśmy symbolizować zdania „Nie istnieje żaden jednorożec” za pomocą podwójnej negacji, tak nie powinniśmy oddać zdań takich jak 12# czy 11# jako zdań kontrfaktycznych z poprzednikami o kształcie alternatywy. Głęboką strukturę tych zdań (w przypadku których zasada SDA wydaje się mieć zastosowanie) należy przedstawić jako koniunkcję:

$$(p \Box \rightarrow r) \wedge (q \Box \rightarrow r)$$

Nie znaczy to, że żadnych zdań o podobnej strukturze powierzchniowej nie należy oddawać jako zdań z poprzednikami o kształcie alternatywy:

$$(15) (p \vee q) \Box \rightarrow r$$

Forma 15 przedstawia strukturę zdań takich jak 13.

Jeżeli przyjąć strategię głębokiej symbolizacji, to nie ma konieczności wprowadzenia zasady symplifikacji alternatywnych poprzedników. Powstaje jednak problem, w jaki sposób rozstrzygnąć, kiedy podobną strukturę powierzchniową należy interpretować w jeden, a kiedy w drugi sposób.

Problem ten okaże się jeszcze głębszy, gdy rozważymy przykłady takie, jak 4#–6#. Należałoby bowiem dodać jeszcze trzeci sposób symbolizowania zdań, których struktura powierzchniowa wydaje się mieć strukturę 15, a mianowicie symbolizowanie ich jako jeszcze bardziej rozbudowanej koniunkcji:

$$[(p \wedge q) \Box \rightarrow r] \wedge [(p \wedge \sim q) \Box \rightarrow r] \wedge [(\sim p \wedge q) \Box \rightarrow r]$$

Być może inne przykłady wskazywałyby jeszcze inne możliwości interpretacji takich zdań.

Należy też zwrócić uwagę, że strategia ta nie stanowi rozwiązania problemu zdań z poprzednikami o kształcie negacji koniunkcji (§ 3). Można byłoby oczywiście tę strategię odpowiednio rozszerzyć i wymagać, aby zdania kształtu „Gdyby zaszło $\sim(p \wedge q)$, to zaszłoby r ” symbolizować raz jako:

$$(16) \sim(p \wedge q) \Box \rightarrow r$$

raz jako:

$$(17) (\sim p \Box \rightarrow r) \wedge (\sim q \Box \rightarrow r)$$

a innymi razy jako:

$$(18) [(\sim p \wedge q) \Box \rightarrow r] \wedge [(p \wedge \sim q) \Box \rightarrow r] \wedge [(\sim p \wedge \sim q) \Box \rightarrow r]$$

bądź też zdania o jeszcze innym kształcie.

6.2. Strategia głębokiej symbolizacji a teza Lowe'a

W jaki sposób ta strategia odpowiedzi na problem alternatywnych poprzedników przekłada się na argumentację Lowe'a? Musimy pamiętać, że Lowe stara się obalić teorię identyczności. To na nim spoczywa *onus probandi*. Musi on dowieść, że teza N jest prawdziwa.

Przyjmując strategię głębokiej symbolizacji jako odpowiedź na problem poprzedników o kształcie alternatywy, dopuszczalne są przynajmniej trzy

interpretacje zdania N – jako zdania kształtu analogicznego do 16, 17 lub 18. Widzieliśmy już (§ 5.2), że interpretując to zdanie jako zdanie o kształcie 18 okazuje się ono zdaniem fałszywym, natomiast jest ono prawdziwe przy pozostałych dwóch interpretacjach. Dotyczy to też wszystkich przytoczonych w § 3 kontrprzykładów. Aby *wykazać*, że zdanie N jest prawdziwe, Lowe musiałby najpierw wykazać, że struktura logiczna tego zdania nie ma kształtu analogicznego do 18. Lowe nic takiego nie czyni. A zatem nie wykazuje on, że N jest zdaniem prawdziwym.

Można zasadnie bronić tezy, że strukturę logiczną zdania N należy oddać właśnie jako strukturę analogiczną do 18. Na to wskazują przytaczane w § 3 kontrprzykłady. Aby uznać je za zdania fałszywe, musielibyśmy uznać, że ich struktura logiczna jest analogiczna do 18. W takim jednak razie mamy podstawy, by sądzić, że zdanie N jest zdaniem fałszywym.

Na gruncie omawianej strategii możliwe są odmienne interpretacje zdania N, na gruncie których jest ono zdaniem prawdziwym. Takie interpretacje zdania N wymagałyby jednak dodatkowego wyjaśnienia, dlaczego rozumienie tezy N różni się od rozumienia podanych w § 3 kontrprzykładów. Przyjmując – co nie jest oczywiste – że takie wyjaśnienie można podać, nie udowodniliśmy, że N jest zdaniem fałszywym, a jedynie że istnieją pewne – aczkolwiek bardzo silne – racje, by tak sądzić. Pokazaliśmy w każdym razie, że Lowe *nie wykazał*, iż teza N jest prawdziwa.

7. Ograniczenie zasady podstawiania

Zwolennicy zasady podstawiania wskazywali wcześniej (por. Fine 1975; Nute 1975), że można uniknąć redukcji funktora $\square \rightarrow$ do funktora implikacji ścisłej przez ograniczenie zasady podstawiania. Dopiero jednak ostatnie prace Kita Fine'a (2012 oraz praca w druku) stanowią przełom dla tego sposobu odpowiedzi na problem alternatywnych poprzedników. Po pierwsze, jak zobaczymy (§ 7.1), Fine broni zasady SDA przed kontrprzykładami (McKay, van Inwagen 1977). Po drugie, Fine proponuje rozróżnienie pojęć konsekwencji, które daje mu podstawy dla ograniczenia zasady podstawiania (§ 7.2)⁷. Ograniczenie to pozwala na wykazanie, że zarówno wszystkie przytoczone kontrprzykłady, jak i teza N są zdaniem fałszywymi (§ 7.3).

⁷ Fine (praca w druku) przytacza szereg dodatkowych powodów dla odrzucenia zasady podstawiania.

7.1. Fine'a obrona zasady SDA

Fine odpowiada na kontrargument McKaya i van Inwagena. Przypomnijmy, że wskazują oni, że przyjęcie SDA pozwala nam z prawdziwego zdania 13:

(13) Gdyby w czasie II Wojny Światowej Hiszpania weszła w sojusz czy to z Aliantami, czy z państwami Osi, to weszłaby w sojusz z Państwami Osi.

wyprowadzić fałszywe zdanie 14:

(14) Gdyby w czasie II Wojny Światowej Hiszpania weszła w sojusz z Aliantami, to weszłaby w sojusz z Państwami Osi.

Fine jednak twierdzi, że wrażenie, jakoby drugie zdanie nie wynikało z pierwszego, jest mylne.

Gdy oceniamy wartość logiczną tych zdań powyższy sposób, bezwiednie zmieniamy kontekst ich oceny. Kontekstem oceny zdania 13 są światy, w których Hiszpania jest bliska państwom Osi (tylko wtedy moglibyśmy uznać to zdanie za prawdziwe). Natomiast oceniając wartość logiczną zdania 14 rozszerzamy ten kontekst również o światy, w których Hiszpania wchodzi w sojusz z Aliantami. Fine sugeruje, że poszerzenie to związane jest z pragmatyczną zasadą akomodacji supozycji (*Suppositional Accommodation*), zgodnie z którą zawsze staramy się przyjąć taką interpretację zdania kontrfaktycznego, żeby poprzednik przedstawiał autentyczną możliwość. Aby poprzednik „w czasie II Wojny Światowej Hiszpania weszła w sojusz z Aliantami” przedstawiał autentyczną możliwość, kontekst oceny zdania 14 musi zostać poszerzony o światy, w których Hiszpania wchodzi w sojusz z Aliantami. W tych światach następnik zdania 14 jest fałszywy, stąd uznajemy 14 za zdanie fałszywe.

Jeżeli jednak trzymamy się stałego kontekstu oceny obu zdań, to okaże się, że nie mamy do czynienia z kontrprzykładem dla zasady SDA. W oryginalnym kontekście oceny, w którym Hiszpania jest bliska państwom Osi, zdanie 13 jest prawdziwe, ale zdanie 14 jest wówczas również prawdziwe, bo jest pusto spełnione – nie jest możliwe w tym kontekście, aby Hiszpania weszła w sojusz z Aliantami. Nie mamy tu zatem do czynienia z kontrprzykładem.

Jeżeli rozważymy poszerzony kontekst, składający się ze światów, w których Hiszpania jest bliska państwom Osi, ale także ze światów, w których Hiszpania może wejść w sojusz z Aliantami, to zdanie 14 jest fałszywe, ale fałszywe będzie wówczas również zdanie 13. Przecież jeżeli dopuszczamy jako możliwość to, że Hiszpania wchodzi w sojusz również z Aliantami, to następnik zdania 13 jest fałszywy – wcale nie jest prawdą, że w tych światach

Hiszpania weszłaby w sojusz z Państwami Osi. Ani w jednym, ani w drugim kontekście nie mamy do czynienia z kontrprzykładem dla zasady SDA.

7.2. Ograniczenie zasady podstawiania

Ograniczenie zasady podstawiania zdań równoważnych na pierwszy rzut oka nie wydaje się strategią owocną. Pierwszą przymiarką mogłoby być ograniczenie tej zasady tak, by się nie stosowała do poprzedników zdań kontrfaktycznych. Jest jednak szereg podstawień zupełnie niekłopotliwych, które też byłyby w ten sposób wykluczone (np. podstawienie „ $p \vee q$ ” zamiast „ $q \vee p$ ”).

Można próbować podzielić równoważności na niekłopotliwe, do których zasada się stosuje, i kłopotliwe – do których się nie stosuje. Jednakże taki podział wymagałby jakichś teoretycznych podstaw – tym bardziej w kontekście polemiki z Lowe'em. Inaczej moglibyśmy sobie wyobrazić, że będziemy tupać, że zdanie N jest niekłopotliwe, a on odpowie odmiennym tupaniem.

Takie teoretyczne podstawy przedstawia jednak Fine (2012). Proponuje on odmienną semantykę dla zdań kontrfaktycznych, na gruncie której wyróżnia trzy rodzaje relacji konsekwencji. Wprowadza w szczególności pojęcie implikacji ścisłej, na gruncie której okazuje się, że p nie jest równoważne $p \vee (p \wedge q)$, w sposób ścisły. Uzyskuje się w ten sposób podstawy dla zablokowania wspomnianego wcześniej wyprowadzenia zasady wzmocnienia poprzednika. Wystarczy ograniczyć zasadę podstawiania – przynajmniej w poprzednikach zdań kontrfaktycznych – do zdań ściśle równoważnych.

7.3. Strategia ograniczenia zasady podstawiania a teza Lowe'a

Przy wspomnianym minimalnym ograniczeniu można zastosować zasadę SDA do wszystkich zdań 1#–6#, a także N w sposób sugerowany wyżej (§ 5.2). Okazuje się bowiem, że – w ujęciu Fine'a – tak jak $(p \vee q)$ jest ściśle równoważne $(q \vee p)$, tak również $\sim(p \wedge q)$ jest ściśle równoważne zarówno $\sim p \vee \sim q$, jak i $(p \wedge \sim q) \vee (\sim p \wedge q) \vee (\sim p \wedge \sim q)$.

Zgodnie z zasadą podstawiania poprzedników ściśle równoważnych, zdanie N jest ściśle równoważne:

(N') Gdyby albo tylko jedno ze zdarzeń N_1, \dots, N_k nie wystąpiło, albo ..., albo żadne ze zdarzeń N_1, \dots, N_k nie wystąpiło, to ręka podmiotu uniosłaby się w prawie dokładnie taki sam sposób, jak to rzeczywiście miało miejsce.

Na mocy zasady SDA otrzymujemy:

Gdyby żadne ze zdarzeń N_1, \dots, N_k nie wystąpiło, to ręka podmiotu uniosłaby się w prawie dokładnie taki sam sposób, jak to rzeczywiście miało miejsce.

Jest to zdanie fałszywe, podobnie jak bardzo wiele zdań, które można z N' wyprowadzić za pomocą SDA. Jeżeli tak, to na mocy *modus tollens* możemy wnioskować, że zarówno N' , jak i N są zdaniami fałszywymi.

Zwolennik argumentacji Lowe'a może zaprotestować, że nie podaliśmy racji, aby przyjąć propozycję Fine'a. Istotnie, te racje wyklarują się dopiero w dalszych dyskusjach nad jego nową semantyką. Sytuacja dialektyczna przedstawia się jednak w następujący sposób. Możliwe są takie ograniczenia zasady podstawiania, przy których N jest zdaniem fałszywym, ale też takie, przy których jest ono zdaniem prawdziwym. Nie wykazaliśmy zatem, że przyjmując strategię ograniczenia zasady podstawiania, N jest zdaniem fałszywym. Wykazaliśmy, że istnieją racje, by tak sądzić.

Aby wykazać, że N jest zdaniem prawdziwym, Lowe musiałby wykazać, że przytoczony sposób – i sposoby jemu pokrewne – ograniczenia zasady podstawiania są nieuprawnione. Lowe niczego takiego nie czyni, a to znaczy, że nie wykazał on, iż N jest zdaniem prawdziwym.

8. Wnioski

Podsumujmy powyższe ustalenia. Kontrowersyjna teza Lowe'a (N) ma kształt zdania kontrfaktycznego, którego poprzednikiem jest negacja wieloczłonowej koniunkcji (§ 2). Jedyнным argumentem Lowe'a na rzecz prawdziwości tej tezy jest zastosowanie do niej semantyki światów możliwych. W § 3 przytoczyliśmy szereg przykładów zdań o podobnym kształcie, które oceniamy niezgodnie z werdyktem, jaki uzyskalibyśmy na gruncie semantyki światów możliwych (przynajmniej gdybyśmy stosowali ją w sposób, w jaki stosuje ją Lowe). Widzieliśmy też, że można zrozumieć głębszą strukturę tych przykładów (§ 4), a zatem można ich wygenerować dużo więcej. Jednym z wyjaśnień tego stanu rzeczy jest przywołanie starego problemu, z którym boryka się semantyka światów możliwych, a mianowicie problemu poprzedników o kształcie alternatywy (§ 5). W końcu negacja koniunkcji jest równoważna alternatywie negacji. Widzieliśmy też, że rozwiązania problemu poprzedników o kształcie alternatywy nie dają jednoznacznego werdyktu, jeżeli chodzi o ocenę wartości logicznej tych zdań. Wszystkie jednak dopuszczają taką interpretację zdania takiego jak N , przy której byłoby ono fałszywe. Lowe nie uczynił nic, by taką interpretację wykluczyć. Nie wolno więc zgodzić się z Lowe'em, że wykazał prawdziwość tezy N .

Mamy tym samym podstawy, by odrzucić argumentację Lowe'a przeciw teorii identyczności. Przytoczyliśmy liczne racje za tym, że kluczowa w tej

argumentacji przesłanka N jest fałszywa. Nie ulega w każdym razie żadnej wątpliwości, że uzasadnienie tej przesłanki przez Lowe'a jest wysoce niewystarczające.

Co by więc było, gdyby nie wszystkie neurony zachowywały się tak, jak się rzeczywiście zachowały? Myli się Lowe sugerując, że musimy na to pytanie odpowiedzieć zgodnie z odpowiedzią sugerowaną przez semantykę światów możliwych, a więc, że prawie wszystkie neurony zachowałyby się tak, jak się rzeczywiście zachowały, a w takim razie ręka uniosłaby się prawie tak samo, jak się rzeczywiście uniosła. Myli się on przede wszystkim dlatego, że każe nam uznać za obowiązujący werdykt semantyki światów możliwych w przypadkach, w których semantyka ta często daje odmienne od intuicyjnych rezultaty.

Sądzę, że nie jesteśmy w stanie dać jednoznacznej odpowiedzi na to pytanie, tak samo jak nie byliśmy w stanie jednoznacznie odpowiedzieć na pytanie, co wskazałaby waga, gdyby na jej szali nie położono wszystkich czterdziestu dwóch 10-gramowych odważników, które faktycznie zostały na niej położone. Odpowiedź zależy od tego, ile odważników zostałoby na szali położonych i *per analogiam* od tego, ile neuronów (a być może też od tego, jakich) zachowałyby się inaczej (i w jakim stopniu inaczej), niż się rzeczywiście zachowały. Jeżeli wystarczająco wiele wystarczająco kluczowych neuronów zachowa się wystarczająco odmiennie, to ręka się nie uniesie.

Bibliografia

- Alonso-Ovalle L. (2009), *Counterfactuals, Correlatives, and Disjunction*, „Linguistics and Philosophy” 32, s. 207–244.
- Briggs R. (2012), *Interventionist Counterfactuals*, „Philosophical Studies” 160, s. 139–166.
- Creary L.G., Hill C.S. (1975), *Review of D. Lewis' 'Counterfactuals'*, „Philosophy of Science” 42, s. 341–344.
- Ellis B., Jackson F., Pargetter R. (1977), *An Objection to Possible-World Semantics for Counterfactual Logics*, „Journal of Philosophical Logic” 6, s. 355–357.
- Fine K. (1975), *Critical Notice of D. Lewis' Counterfactuals*, „Mind” 84, s. 451–458.
- Fine K. (2012), *Counterfactuals without Possible Worlds*, „The Journal of Philosophy” 109, s. 221–246.
- Fine K. (w druku), *A Difficulty for the Possible Worlds Analysis of Counterfactuals*, „Synthese”, DOI: 10.1007/s11229-012-0094-y.
- Lewis D. (1973), *Counterfactuals*, Cambridge, MA: Harvard University Press.

- Loewer B. (1976), *Counterfactuals with Disjunctive Antecedents*, „The Journal of Philosophy” 73, s. 531–537.
- Lowe E.J. (2006), *Non-Cartesian Substance Dualism and the Problem of Mental Causation*, „Erkenntnis” 65, s. 5–23.
- Lowe E.J. (2008), *Personal Agency. The Metaphysics of Mind and Action*, Oxford: Oxford University Press.
- McKay T., van Inwagen P. (1977), *Counterfactuals with Disjunctive Antecedents*, „Philosophical Studies” 31, s. 353–356.
- Nute D. (1975), *Counterfactuals and the Similarity of Worlds*, „Journal of Philosophy” 72, s. 773–778.
- Nute D. (1984a), *Conditional Logic*, w: D. Gabbay, F. Guenther (eds), *Handbook of Philosophical Logic*, Vol. 2, s. 387–439, Dordrecht: Reidel.
- Nute D. (1984b), *Topics in Conditional Logic*, Dordrecht: Reidel.
- Nute D., Cross C.B. (2001), *Conditional Logic*, w: D. Gabbay, F. Guenther (eds), *Handbook of Philosophical Logic*, 2nd ed., Vol. 4, s. 1–98, Dordrecht: Kluwer.
- Paprzycka K. (2013), *O argumentacji Lowe'a na rzecz koncepcji przyczynowości intencjonalnej i fizycznej*, „Filozofia Nauki” 1 (81).
- Paprzycka K. (w druku), *Lowe's Argument Against the Psycho-Neural Token-Identity Thesis*, „Pacific Philosophical Quarterly”.
- Paprzycka K. (w druku), *Teza Lowe'a w świetle semantyki Lowe'a*, „Filozofia Nauki”.

Streszczenie

Argumentacja E.J. Lowe'a przeciwko teorii identyczności opiera się na tezie, że gdyby nie wszystkie neurony zachowały się tak, jak się rzeczywiście zachowały, to spowodowałyby takie samo zachowanie, jakie rzeczywiście spowodowały. Uzasadnienie tej przesłanki sprowadza się do zastosowania semantyki światów możliwych. Stosując tę semantykę, można jednak również wykazać prawdziwość ewidentnie fałszywych zdań o podobnej strukturze. Problem ten jest szczególnym przypadkiem tzw. problemu poprzedników o kształcie alternatywy. Przyjmując każdą z dwóch podstawowych strategii odpowiedzi na ten problem, można znaleźć podstawy, by twierdzić, że teza Lowe'a jest fałszywa. Lowe nie wykazał więc, że jego kontrowersyjna teza jest prawdziwa, co wystarcza do odrzucenia jego argumentu przeciw teorii identyczności.

