

Katarzyna Paprzycka

Teza Lowe'a w świetle semantyki Lowe'a

Filozofia Nauki 22/4, 71-90

2014

Artykuł został opracowany do udostępnienia w internecie przez Muzeum Historii Polski w ramach prac podejmowanych na rzecz zapewnienia otwartego, powszechnego i trwałego dostępu do polskiego dorobku naukowego i kulturalnego. Artykuł jest umieszczony w kolekcji cyfrowej bazhum.muzhp.pl, gromadzącej zawartość polskich czasopism humanistycznych i społecznych.

Tekst jest udostępniony do wykorzystania w ramach
dozwolonego użytku.

Katarzyna Paprzycka

Teza Lowe'a w świetle semantyki Lowe'a¹

W dwóch artykułach (2013ab) zgłaszałam wątpliwości co do zasadności podanego przez Lowe'a argumentu z implikacji kontrfaktycznych (Lowe 2006, 2008). Szczególnie kłopotliwa wydaje się teza Lowe'a, zgodnie z którą:

(N[#]) Gdyby nie wszystkie neurony zachowały się tak, jak się zachowują, to ręka i tak by się uniosła.

Sugerowałam, że podane przez Lowe'a uzasadnienie tej tezy jest wadliwe. Korzysta on ze standardowej semantyki światów możliwych, w której analiza poprzedników o postaci alternatywy stanowi jeden z podstawowych problemów (por. Paprzycka 2013a). W wypadku tego typu zdań, których przykładem jest również zdanie (N[#]), semantyka światów możliwych jest często niezgodna z naszymi intuicjami. W (2013a) wskazałam też przykłady zdań o takiej samej formie, z których każde uzyskuje na gruncie semantyki światów możliwych niezgodną z intuicjami wartość logiczną. Przedstawiłam również strukturę logiczną, na podstawie której można łatwo podać wiele tego rodzaju przykładów.

Innym sposobem rozumienia kontrowersyjności takich zdań jest zwrócenie uwagi na to, że argument Lowe'a opiera się na rozstrzygnięciach dotyczących tego, jakie światy uznać za bliskie. W §4 przedstawię taki zarzut. Nie jest on wprawdzie wiążący, ale warto go sformułować. W dalszej części artykułu spróbuję pokazać, że podobny kontrargument można wytoczyć przeciw Lowe'owi na gruncie jego własnej semantyki zdań kontrfaktycznych. W §5 naszkicuję semantykę Lowe'a, a następnie wykorzystam ją do wykazania, że zdania typu (N[#]) są fałszywe (§6).

¹ Chciałabym bardzo serdecznie podziękować Mariuszowi Grygiałcowi oraz anonimowemu recenzentowi za uwagi krytyczne do pierwszej wersji tekstu.

Wartością semantyki Lowe'a w kontekście tej dyskusji jest fakt, że daje pewną metodę rozstrzygnięcia sporu. W §7 zobaczymy, w jaki sposób Lowe mógłby odpowiedzieć na przedstawioną argumentację. Okazuje się, że taka odpowiedź pociągałaby — na gruncie jego semantyki — pewne dalsze zobowiązania. Zobowiązania te są trudne do przyjęcia, a koszt uznania tezy Lowe'a na gruncie semantyki Lowe'a jest dużo większy niż na gruncie semantyki światów możliwych.

1. TEZA LOWE'A

Teza Lowe'a pojawia się w kontekście jego argumentu z implikacji kontrfaktycznych (Lowe 2006, 2008) przeciw tezie o identyczności egzemplarycznej, przyjmowanej przez wszystkie teorie identyczności. Zgodnie z tezą identyczności egzemplarycznej zdarzenie (egzemplarz) *D* (np. decyzja podmiotu, by podnieść rękę) jest tożsame z pewnym złożonym zdarzeniem (egzemplarzem) neuronalnym *N*. Lowe stara się pokazać, że obowiązują następujące zdania kontrfaktyczne:

- (N) „Gdyby zdarzenie *N* nie wystąpiło, to *ręka podmiotu uniosłaby się prawie dokładnie tak samo, jak to rzeczywiście miało miejsce*” (Lowe 2008: 105).
- (M) „Gdyby zdarzenie *D* nie wystąpiło [...], to *ręka podmiotu w ogóle by się nie uniosła*” (Lowe 2008: 105).

Jeżeli (N) i (M) są prawdziwe, to zdarzenia *D* i *N* nie mogą być tożsame. Okazuje się bowiem, że niewystąpienie *N* ma inne własności kontrfaktyczne niż niewystąpienie *D*.

W artykule skoncentruję się wyłącznie na prawdziwości (N). Lowe wychodzi z założenia, że *N* jest przyczyną pewnego behawioralnego zdarzenia (egzemplarza) *B* (tu: uniesienia ręki). A zatem:

Gdyby *N* nie wystąpiło, to nie wystąpiłoby *B*.

Według standardowej interpretacji kontrfaktycznych zdań warunkowych (Lewis 1973) zdanie to jest prawdziwe zawsze i tylko wtedy, gdy we wszystkich najbliższych światach możliwych, w których *N* nie występuje, nie występuje też *B*. Zdanie (N) jest prawdziwe, jeżeli we wszystkich najbliższych światach, w których nie występuje *N*, występuje zdarzenie (egzemplarz) podobne do *B*, tj. również będące uniesieniem ręki, lecz charakteryzujące się nieznacznie odmiennymi parametrami.

Lowe uważa, że może wykazać, iż (N) jest prawdą, odwołując się do relacji bliskości między światami, narzuconej przez standardową interpretację zdań kontrfaktycznych. Twierdzi, że w najbliższych światach możliwych, w których nie występuje *N*, wystąpi inne zdarzenie neuronalne *N**, które musi być *bardzo* podobne do zdarzenia rzeczywistego *N*: *N* różni się od *N** aktywnością co najwyżej paru neuronów. Jeżeli tak, to *N** będzie się charakteryzowało podobnymi (do *N*) mocami przyczynowymi, a w związku z tym doprowadzi do uniesienia ręki *B**, które nieznacznie

różnić się będzie od rzeczywistego uniesienia ręki B . W ten sposób Lowe przekonuje, że gdyby zdarzenie neuronalne N , które spowodowało uniesienie ręki B , nie wystąpiło, ręka tak czy owak by się uniosła — choć jej uniesienie miałoby nieco odmiennie parametry fizyczne (tempo, trajektoria, czas itd.).

Teza (N) odnosi się do bardzo złożonego zdarzenia egzemplarza N . Zdarzenie N jest „sumą [...] ogromnej liczby pojedynczych zdarzeń neuronalnych” (Lowe 2008: 105). Przyjmijmy, że te pojedyncze zdarzenia neuronalne to zdarzenia (egzemplarze) 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 i N_k . W takim razie teza (N) jest równoważna tezie ($N^\#$):

($N^\#$) Gdyby nie wszystkie neurony zachowały się tak, jak się zachowują, to ręka i tak by się uniosła.

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. Dalej określać będziemy tezę ($N^\#$) mianem „tezy Lowe'a”.

Argument Lowe'a oparty jest na szeregu założeń upraszczających, z których nie wszystkie muszą okazać się prawdziwe. Lowe zakłada, że liczba neuronów przyczynowo odpowiedzialnych za uniesienie ręki, a będących neuronalnym korelatem intencji, jest ogromna. Zakłada też, że wszystkie te neurony są w zbliżonym stopniu przyczynowo odpowiedzialne za uniesienie ręki: żaden z neuronów nie odgrywa większej roli od pozostałych. Na potrzeby argumentacji przyjmijmy te założenia.

2. KONTRPRZYKŁADY

W (2013ab) przytaczałam liczne kontrprzykłady. Przypomnę dwa.

Przykład A. Wyobraźmy sobie, że na wadze umieszczono 42 dziesięciogramowe odważniki (zakładamy, że są to wszystkie dostępne odważniki). Rzetelna waga wskazała ich łączną masę, 420 gramów. 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 znalazłoby się na wadze zamiast czterdziestu dwóch. Jeżeli tego nie wiadomo, to nie ma podstaw, by wyróżnić którąkolwiek z czterdziestu dwu możliwości: mogłoby być tak, że na wadze umieszczono 41 odważników, 40 odważników, ..., a mogłoby być tak, że znalazłby się na niej tylko

jeden odważnik bądź pozostałaby pusta. Odpowiedź na postawione pytanie będzie więc brzmiała:

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

Stosując semantykę światów możliwych, musimy najpierw określić, które światy o spełnionym poprzedniku (światy, w których nie wszystkie odważniki zostały umieszczone na wadze), są najbliższe światu rzeczywistego. Zgodnie z intuicjami Lowe'a są to światy, w których na wadze znalazły się prawie wszystkie odważniki (z wyjątkiem jednego lub dwóch). W takim wypadku waga wskaże jednak 410 g lub 400 g, a w każdym razie wartość zbliżoną do 420 g. Wydaje się, że następujące zdanie jest prawdziwe w świetle semantyki światów możliwych:

- (A[#]) Gdyby nie wszystkie odważniki zostały umieszczone na wadze, to wskazałaby 410 g lub 400 g, a w każdym razie wartość zbliżoną do 420 g.

(A[#]) jest jednak niezgodne z intuicjami wyrażonymi w (A). Mamy wszelkie podstawy po temu, by przyjąć twierdzenie słabsze:

Gdyby nie wszystkie odważniki zostały umieszczone na wadze, to waga *mogłaby* wskazać 410 g lub 400 g, a w każdym razie wartość zbliżoną do 420g,

a (A[#]) odrzucić jako zbyt silne.

Przykład B. Lena ma silną alergię na sierść. Tak się akurat składa, że każda z jej dwóch najbliższych przyjaciółek dostaje 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 było prawdą, że obie przyjaciółki dostały w prezencie zwierzaka?

W takim wypadku, podobnie jak wyżej, nie możemy powiedzieć, co by się stało, ale tylko co mogłoby się stać. Gdyby na przykład tylko jedna przyjaciółka dostała zwierzaka, a druga nie, to u Leny i tak wystąpiłaby silna reakcja alergiczna. Gdyby jednak żadna z przyjaciółek nie otrzymała takiego prezentu, to *ceteris paribus* u Leny taka reakcja by nie wystąpiła. Innymi słowy przyjęlibyśmy:

- (B) Gdyby nie było prawdą, że zarówno Asia, jak i Beata dostały zwierzaka, to u Leny *mogłaby* i tak *wystąpić* silna reakcja alergiczna.

- (B') Gdyby nie było prawdą, że zarówno Asia, jak i Beata dostały zwierzaka, to u Leny *mogłaby nie wystąpić* silna reakcja alergiczna.

Odrzucilibyśmy jednak:

- (B[#]) Gdyby nie było prawdą, że zarówno Asia, jak i Beata dostały zwierzaka, to u Leny wystąpiłaby i tak silna reakcja alergiczna.

Niemniej, stosując semantykę światów możliwych w wyżej omówiony sposób, otrzymamy raczej właśnie zdanie ($B^\#$), a nie (B) i (B'). Światy, w których jedna z przyjaciółek otrzymuje czworonoga, są *ceteris paribus* bardziej podobne do świata rzeczywistego niż światy, w których żadna z przyjaciółek nie dostaje zwierzaka. Dlatego w najbliższych światach możliwych, w których nie wszystkie najbliższe przyjaciółki Leny dostają zwierzaka, u Leny pojawiają się objawy alergiczne.

Werdykt semantyki światów możliwych wydaje się więc odbiegać od naszych intuicji.

3. CO BY BYŁO, GDYBY PRAWIE WSZYSTKIE NEURONY ZACHOWAŁY SIĘ TAK, JAK SIĘ ZACHOWAŁY?

Można by się zastanowić, czy tezy Lowe'a (N) nie należałoby oddać raczej jako:

(N^\dagger) Gdyby prawie wszystkie neurony zachowały się tak, jak się zachowują, to ręka i tak by się uniosła.

Przy założeniach przyjmowanych przez Lowe'a teza (N^\dagger) jest prawdziwa. Na gruncie standardowej semantyki światów możliwych uzasadnienie (N^\dagger) będzie dokładnie pokrywało się z uzasadnieniem (N)². Ale skoro tak, to zgodnie z zasadą życzliwości można zasadnie sugerować, że inkryminowaną tezę Lowe'a (N) należy rozumieć raczej jako prawdziwe zdanie (N^\dagger), a nie jako fałszywe zdanie ($N^\#$)³.

Tak jednak nie jest. Rozstrzygająca jest bowiem rola tezy (N) w argumentacji Lowe'a. Lowe chce pokazać, że niewystąpienie zdarzenia N ma inne własności kontrfaktyczne niż niewystąpienie zdarzenia D . Zdarzenie (egzemplarz) N jest złożone ze wszystkich zdarzeń (egzemplarzy) polegających na określonej aktywności neuronów. Prawdą jest zatem, że N występuje zawsze i tylko wtedy, gdy mają miejsce zdarzenia składowe N_1, \dots i N_k . Innymi słowy, wystąpienie N polega na tym, że zachodzą *wszystkie* zdarzenia składowe N_1, \dots i N_k . Skoro tak, to naturalne wydaje się przypuszczenie, że niewystąpienie N polega na tym, że nieprawdą jest, iż mają miejsce *wszystkie* zdarzenia składowe N_1, \dots i N_k , tj. na tym, że *nie wszystkie* zdarzenia składowe N_1, \dots i N_k występują.

Ktoś, kto dopuszczalby interpretację (N) jako (N^\dagger), musiałby przyjąć twierdzenie:

(#) Do niewystąpienia złożonego zdarzenia (egzemplarza) dochodzi zawsze i tylko wtedy, gdy występują prawie wszystkie zdarzenia (egzemplarze) wchodzące w jego skład.

Teza (#) jest jednak fałszywa ze względu na fałszywość implikacji:

² W rzeczy samej dokładnie na tym polega problem. Wydaje się, że semantyka światów możliwych nie rozróżnia zdań o postaci (N) i o postaci (N^\dagger).

³ Jestem wdzięczna Mariuszowi Grygiałcowi za wyraźne sformułowanie tej możliwości.

- (##) Do niewystąpienia złożonego zdarzenia (egzemplarza) dochodzi tylko wtedy, gdy występują prawie wszystkie zdarzenia (egzemplarze) wchodzące w jego skład.

Niech A będzie złożonym zdarzeniem (egzemplarzem) polegającym na umieszczeniu na wadze wszystkich 42 odważników. Zdarzenie A składa się ze zdarzeń (egzemplarzy) A_1, \dots, A_{42} , gdzie zdarzenie A_i jest zdarzeniem polegającym na umieszczeniu i -tego odważnika na wadze. Oczywiście jest tak, że jeżeli prawie wszystkie z tych odważników zostaną położone na wadze, to nie wystąpi A . Nie jest jednak prawdą, że niewystąpienie A pociąga za sobą położenie prawie wszystkich odważników. Jeżeli wystąpi tylko A_1 , to A również nie wystąpi.

Innymi słowy, (N) nie jest równoważne prawdziwemu (N^\dagger), natomiast jest równoważne fałszywemu ($N^\#$). Czy w takim razie Lowe nie mógłby po prostu zarzucić (N) na rzecz (N^\dagger)? Owszem, mógłby, ale taki manewr wytrąciłby mu oręż z ręki przy ataku na teorię identyczności. Z tego, że z jednej strony:

- (N^\dagger) Gdyby prawie wszystkie neurony zachowały się tak, jak się zachowują, to ręka i tak by się uniosła,

a z drugiej strony:

- (M) „Gdyby zdarzenie D nie wystąpiło [...], to ręka podmiotu *w ogóle by się nie uniosła*” (Lowe 2008: 105),

nie wynika przecież zaprzeczenie tezy o identyczności egzemplarycznej D i N . Dlatego też życzliwa interpretacja Lowe'a wymaga przypisania mu tezy ($N^\#$) — która wydaje się fałszywa.

4. ARGUMENT Z BLISKOŚCI ŚWIATÓW

We wszystkich przykładach przytoczonych w §2 podążaliśmy za sposobem, w jaki Lowe wyznacza zbiór najbliższych światów możliwych. Trudno zaprzeczyć, że wśród światów, w których nie wszystkie odważniki zostają położone na wadze, bardziej podobne do świata rzeczywistego są te światy, w których prawie wszystkie odważniki zostają położone na wadze niż te, w których prawie żaden odważnik nie trafia na wagę. Najbardziej podobne do świata rzeczywistego wydają się światy, w których na wadze jest 41 (ewentualnie 40) odważników. Analogicznie, wśród światów, w których nie wszystkie neurony zachowują się tak, jak się rzeczywiście zachowują, to właśnie światy, w których większość neuronów zachowuje się tak, jak się rzeczywiście zachowuje, są bardziej podobne do świata rzeczywistego niż światy, w których większość neuronów zachowuje się inaczej niż się rzeczywiście zachowuje. Najbardziej podobne wydają się więc światy, w których wszystkie neurony za wyjątkiem jednego (ewentualnie dwóch) zachowują się tak, jak się rzeczywiście zachowują.

Wydaje się, że trudno zaprzeczyć, iż wspomniane światy są najbardziej podobne do świata rzeczywistego. A przynajmniej byłoby to trudne, gdyby Lowe sądził, że poprzednik tezy (N) mówi o *typach* zdarzeń. Lowe zastrzega jednak, że ma na myśli egzemplarze. Formuluje tezę (N) w odniesieniu do jednego złożonego zdarzenia (egzemplarza) neuronalnego *N*, na które składa się taka, a nie inna aktywacja ogromnej liczby neuronów (określimy ten złożony typ aktywacji jako **N**): gdyby *N* (typu **N**) nie wystąpiło, to ręka i tak by się uniosła, choć w nieco inny sposób. Otóż można byłoby utrzymywać, że najbliższymi światami, w których zdarzenie (egzemplarz) *N* typu **N** nie występuje, są te światy, w których występuje inne zdarzenie (egzemplarz) tego samego typu **N**, a więc gdzie wszystkie neurony zachowują się *dokładnie tak samo*, jak zachowują się w świecie rzeczywistym. Światy, w których występują inne egzemplarze tego samego typu zachowania neuronów, będą przecież bardziej podobne do świata rzeczywistego niż światy, w których choćby jeden neuron zachowuje się inaczej. Jeżeli tak, to teza Lowe'a jest fałszywa i powinna być zastąpiona tezą jeszcze ostrzejszą:

Gdyby *N* (typu **N**) nie wystąpiło, to ręka i tak by się uniosła w *dokładnie taki sam* sposób, jak to rzeczywiście miało miejsce.

Ten wniosek sugerowałby, że przyczyny neuronalne w *ogóle* nie wpływają na ruch ręki⁴.

Nie trzeba jednak wcale zgodzić się z twierdzeniem, że wspomniane światy są najbardziej podobne do świata rzeczywistego — jeżeli brać pod uwagę *wszystkie istotne* względy podobieństwa. Odnosząc się do jednego z problemów (problemu prawdziwych poprzedników), Lewis (1973: 29) sugeruje, że w ocenie bliskości światów nie musimy brać pod uwagę drobnych różnic między światami⁵. Innymi słowy, nie wszystkie podobieństwa między światami muszą być uznane za istotne w ocenie bliskości światów.

W późniejszym tekście Lewis (1986) przypisuje różne wagi różnym typom podobieństw między światami. Najistotniejsze jest to, żeby w światach bliskich obowiązywały te same prawidłowości. W drugiej kolejności ważne jest, aby zgadzały się co do przeważającej większości faktów dotyczących zdarzeń jednostkowych. W trzeciej kolejności istotne jest, aby wykluczyć nawet drobne i lokalne pogwałcenia praw przyrody. Mało istotne, albo w ogóle nieistotne, są odstępstwa dotyczące pojedynczych faktów (Lewis 1986: 47-48).

Porównując ze sobą światy, w których nie wszystkie odważniki zostały położone na wadze, musimy wpięć zadbać, by były to światy rządzące się tymi samym prawami. Wykluczamy więc np. światy, w których nie obowiązuje prawo grawitacji

⁴ Sądzę, wbrew Lowe'owi, że podobny wniosek można byłoby sformułować po stronie umysłowej.

⁵ „Perhaps our discriminations of similarity are rather coarse and some worlds different from [world] *i* are enough like *i* so that such small differences as there are fail to register” (Lewis 1973: 29).

i w związku z tym położenie na wagę pewnej liczby odważników będzie miało zupełnie inny skutek niż w świecie rzeczywistym. W drugiej kolejności dbamy o to, by warunki panujące w świecie były podobne. Waga znajduje się w tym miejscu, w którym się znajduje (a nie np. na Księżycu bądź na stacji kosmicznej), dostępne są czterdzieści dwa odważniki, a nie sto czy tylko jeden. Po trzecie, nie ma lokalnych odstępstw od praw natury, wykluczamy więc sytuację, w której kładąc trzydzieści dziesięciogramowych odważników, uzyskujemy nagle wynik 3 kg, ponieważ wystąpiła akurat anomalia. Biorąc pod uwagę możliwe różnice między światami, wydaje się, że to, czy na wadze położymy 41, 40, 39, ..., 1 czy 0 odważników, może być zakwalifikowane co najwyżej do czwartej kategorii ważności, którą wolno pominąć w ocenie bliskości⁶.

Można bronić tezy, że podobny werdykt co do bliskości światów powinniśmy uzyskać, zastanawiając się nad bliskością światów, w których nie wszystkie neurony zachowują się tak, jak się rzeczywiście zachowują. Dalekie są światy, w których obowiązują inne prawa natury. Dalekie są światy, w których struktura neuronalna podmiotu różni się od rzeczywistej (neurony są inaczej ze sobą połączone, mają inne połączenia eferentne itp.) Dalekie są światy, w których występują lokalne anomalie: aktywacja prawie wszystkich neuronów akurat — anomalnie — nie powoduje uniesienia ręki bądź też aktywacja tylko jednego z tych neuronów akurat — anomalnie — powoduje uniesienie ręki. Natomiast dokładna liczba neuronów, które zachowują się tak jak w świecie rzeczywistym, jest już nieistotna do oceny bliskości światów.

Zanim podniosę kwestię kontrowersyjności podobnego ujęcia, przyjmijmy na chwilę takie rozumienie bliskości światów. Najbliższymi światami, w których nie wszystkie odważniki zostały położone na wadze, są światy, w których 41, 40, ..., 1 bądź 0 odważników znalazło się na wadze (przy zachowaniu tych samych praw przyrody, tych samych warunków, braku anomalii). Najbliższymi światami, w których nie wszystkie spośród k neuronów zachowują się tak, jak się rzeczywiście zachowują, są światy, w których $k-1$, $k-2$, ..., 1 bądź 0 neuronów zachowuje się tak, jak się rzeczywiście zachowuje (przy zachowaniu tych samych praw przyrody, tych samych warunków, braku anomalii).

Jeżeli w ten sposób określimy światy najbliższe, to zdania przytaczane jako kontrprzykłady w §2, a także w moich wcześniejszych pracach (2013ab), okazują się — na gruncie semantyki światów możliwych — fałszywe. Przystają więc być kontrprzykładami dla tej semantyki. Fałszywa okazuje się także teza Lowe'a: w niektórych spośród najbliższych światów możliwych — w omawianym rozumieniu — ręka się uniesie, ale nie we wszystkich. W tych z najbliższych światów, w których dostatecznie duża liczba neuronów zachowuje się tak, jak się rzeczywiście zachowuje, rę-

⁶ Powołując się na przykłady przytaczane przez Tichý'ego (1976) oraz Jacksona (1977), Lewis twierdzi, że w pewnych wypadkach różnice dotyczące faktów jednostkowych nie powinny być uwzględniane w ocenie bliskości, lecz unika kategoriycznego przyjęcia uogólnionej wersji takiego stanowiska.

ka się unosi; w innych światach najbliższych, w których dostatecznie duża liczba neuronów zachowuje się inaczej niż w świecie rzeczywistym, ręka się nie unosi. Prawdziwe są zatem zdania:

Gdyby nie wszystkie neurony zachowały się tak, jak się rzeczywiście zachowały, to ręka mogłaby się unieść.

Gdyby nie wszystkie neurony zachowały się tak, jak się rzeczywiście zachowały, to ręka mogłaby się nie unieść.

Falszywa jest jednak teza Lowe'a ($N^\#$).

Taka odpowiedź na zarzut Lowe'a musi wydawać się dyskusyjna i dialektycznie słaba. Kontrowersyjne jest bowiem określenie bliskości światów. Z taką ripostą wypada się zgodzić, ale odpowiadającemu należy zwrócić uwagę na dwie kwestie. Po pierwsze, skoro taką odpowiedź ocenimy jako dyskusyjną i dialektycznie słabą, to *podobnie* należy ocenić argument Lowe'a. Jego argumentacja za tezą ($N^\#$) opiera się przede wszystkim na takim, a nie innym wyznaczeniu najbliższych światów możliwych, a to — zgadzamy się z ripostującym — jest kontrowersyjne. Po drugie, można mieć wątpliwości, czy wyznaczone przez Lowe'a światy najbliższe światu rzeczywistego istotnie są najbardziej podobne do świata rzeczywistego pod wszystkimi względami. Po trzecie, za odmiennym od dokonanego przez Lowe'a wyznaczeniem najbliższych światów możliwych przemawiają przywołane kontrprzykłady. Krótko mówiąc, omówiony zarzut pod adresem Lowe'a nie jest słabszy od argumentacji Lowe'a na rzecz ($N^\#$).

5. SEMANTYKA LOWE'A

Lowe opiera swoją argumentację na Lewisowskiej wersji semantyki światów możliwych. Sam zaproponował jednak odmienne rozumienie kontrfaktycznych zdań warunkowych (1983, 1995), pozbawione wymogu bliskości. W pozostałej części artykułu pokażę, w jaki sposób można użyć semantyki Lowe'a do wykazania, że zarówno kontrprzykłady, jak i teza Lowe'a są fałszywe (§6), a także jak Lowe mógłby te argumenty odeprzeć (§7).

Lowe proponuje rozumienie kontrfaktycznych zdań warunkowych jako specjalnego przypadku implikacji ścisłej:

$$(\text{Def}) (p \Box \rightarrow r) =_{\text{def}} \Box(p \rightarrow r) \ \& \ (\Diamond p \vee \Box r)$$

Warto zwrócić uwagę na dwie kwestie. Po pierwsze, w przeciwieństwie do semantyki Lewisa, w której zdania o niemożliwych poprzednikach są zawsze prawdziwe, ponieważ są pusto spełnione, na gruncie semantyki Lowe'a zdania, których poprzedniki są niemożliwe, będą prawdziwe (gdy ich następniki są konieczne) albo fałszywe (gdy ich następniki nie są konieczne). Przyjrzyjmy się dwóm zdaniom:

- (1) Gdyby istniała największa liczba naturalna n , to istniałaby liczba większa od n , tj. $n + 1$.
- (2) Gdyby istniała największa liczba naturalna n , to liczba n byłaby podzielna przez 2.

Zdania (1) i (2) są prawdziwe na gruncie semantyki Lewisa, ponieważ poprzednik jest pusto spełniony we wszystkich światach możliwych. Semantyka Lowe'a różnicuje zaś ich wartość logiczną. Zdanie (1) jest prawdziwe w semantyce Lowe'a, ponieważ następnik jest koniecznie prawdziwy: dla każdej liczby naturalnej istnieje liczba od niej większa. Natomiast (2) jest fałszywe, ponieważ ani poprzednik nie jest możliwy, ani następnik nie jest konieczny.

Lowe utrzymuje też, że różnice między rodzajami okresów warunkowych prowadzą się do różnic w zakładanym pojęciu możliwości. Rozważmy przytaczane przez niego przykłady⁷:

- (3) Gdyby Oswald nie zabił Kennedy'ego, to zrobiłby to ktoś inny.
- (4) Jeżeli Oswald nie zabije Kennedy'ego, to zrobi to ktoś inny.
- (5) Jeżeli Oswald nie zabił Kennedy'ego, to zrobił to ktoś inny.

Lowe broni tezy, że warunki prawdziwości zdań (3) i (4) są takie same, natomiast mogą się one różnić od warunków prawdziwości zdania (5). Łatwo się zgodzić, że zdanie (5) jest zdaniem prawdziwym, ale można mieć wątpliwości co do prawdziwości zdania (3) lub (4). Lowe sugeruje, że ta różnica jest różnicą modalności — aletycznej w (3) i (4) oraz epistemicznej w (5).

Kolejna kwestia, na którą warto zwrócić uwagę, to fakt, że Lowe opowiada się za tezą, iż w rozumieniu zdań kontrfaktycznych ogromną rolę odgrywają czynniki pragmatyczne. Podaje przykład dwóch osób, które znajdując się w pomieszczeniu, w którym ulatnia się gaz, wypowiadają zdania:

- (6) Gdyby Brown zapalił jedną z tych zapalek, to nastąpiłby wybuch.
- (7) Gdyby Brown zapalił jedną z tych zapalek, to nie nastąpiłby wybuch.

Zdania te wydają się niezgodne, ale Lowe sugeruje, że niezgodność może być tu tak pozorna, jak niezgodność między wypowiedziami „Jestem głodny” i „Nie jestem głodny” dwóch różnych osób. W szczególności Lowe wyobraża sobie, że osoba wypowiadająca (6) chce powiedzieć coś na temat związku przyczynowego między zapaleniem zapalki i wybuchem. Osoba wypowiadająca (7) wyraża natomiast swą ufność w ostrożność Browna, który w ogóle nie zapaliłby żadnej zapalki, gdyby istniała najmniejsza groźba wybuchu.

⁷ Odpowiadają one angielskim zdaniom: (3) If Oswald had not killed Kennedy then someone else would have. (4) If Oswald has not killed Kennedy then someone else will have. (5) If Oswald did not kill Kennedy then someone else did.

6. ARGUMENT PRZECIWIW TEZIE LOWE'A W SEMANTYCE LOWE'A

Zastosuję propozycję Lowe'a, aby podważyć jego tezę (N). Zrobię to w trzech krokach. Po pierwsze, wykażę, że Lowe przyjmuje słabszą wersję zasady symplifikacji alternatywnych poprzedników. Po drugie, pokażę, jak zastosować słabszą wersję tej zasady do wykazania fałszywości (B[#]). Po trzecie, rozciągnę ten argument na przykład A i tezę Lowe'a.

6.1. Zasada symplifikacji alternatywnych poprzedników: mocna i słaba wersja

Zasadę symplifikacji alternatywnych poprzedników (SDA) formułuje się zwykle w następujący sposób:

$$(SDA) \quad ((p \vee q) \Box \rightarrow r) \rightarrow (p \Box \rightarrow r)$$

W (2013a) pokazałam, w jaki sposób można wykorzystać (SDA) do obalenia kontrprzykładów i tezy Lowe'a. Przyjęcie tej zasady wiąże się jednak z szeregiem trudności.

Zasada (SDA) nie obowiązuje na gruncie semantyki Lowe'a, ponieważ $\Diamond(p \vee q) \rightarrow \Diamond p$ nie jest tezą logiki modalnej. Niemniej, zakładając semantykę Lowe'a, można wykazać, że przyjmuje on słabszą wersję zasady symplifikacji alternatywnych poprzedników (SDA')

$$(SDA') \quad (((p \vee q) \Box \rightarrow r) \& \Diamond p) \rightarrow (p \Box \rightarrow r)$$

Wyprowadzenie (SDA') jest proste:

(1)	$(p \vee q) \Box \rightarrow r$	zał.
(2)	$\Diamond p$	zał.
(3)	$\Box((p \vee q) \rightarrow r)$	(Def), (1)
(4)	$\Diamond(p \vee q) \vee \Box r$	(Def), (1)
(5)	$\Box(p \rightarrow r)$	(3) ⁸
(6)	$\Diamond p \vee \Box r$	(2)
(7)	$p \Box \rightarrow r$	(Def), (5), (6)

Słabą wersję zasady symplifikacji alternatywnych poprzedników można zastosować do wykazania fałszywości kontrprzykładów (A[#]), (B[#]), a także tezy Lowe'a (N[#]), korzystając z pewnych dodatkowych, lecz niekontrowersyjnych założeń. Prześledzimy to najpierw na najprostszym przykładzie B.

⁸ Dla dowolnego świata — załóżmy, że p . Stąd wyprowadzamy $p \vee q$, a z (3) wyprowadzamy r .

6.2. Przykład B

Zaczynamy od wyszczególnienia zdań kontrfaktycznych dotyczących tego, co by się stało w poszczególnych sytuacjach kontrfaktycznych:

$({}^B T_1)$ $\sim A \square \rightarrow J$. Gdyby Asia nie dostała zwierzaka, to u Leny wystąpiłaby alergia; gdyby bowiem tylko Beata dostała zwierzaka, to u Leny i tak wystąpiłyby symptomy alergiczne.

$({}^B T_2)$ $\sim B \square \rightarrow J$. Analogicznie do $({}^B T_1)$.

$({}^B T_3)$ $(\sim A \ \& \ \sim B) \square \rightarrow \sim J$. Gdyby ani Asia, ani Beata nie dostały zwierzaka, to u Leny nie wystąpiłaby alergia.

Możemy zdania $({}^B T_1)$ - $({}^B T_3)$ określić mianem „tez tła”. Nie wydają się kontrowersyjne. Warto też podkreślić, że są zupełnie niekontrowersyjne na gruncie semantyki Lewisa. W wypadku każdej tezy tła semantyka Lewisa daje wynik zgodny z intuicjami.

Wyobraźmy sobie teraz, że ktoś stwierdza:

$(B^\#)$ $\sim(A \ \& \ B) \square \rightarrow J$. Gdyby nie było tak, że zarówno Asia, jak i Beata dostały zwierzaka, to u Leny wystąpiłaby alergia.

Podstawiając zdanie równoważne w poprzedniku $(B^\#)$, otrzymujemy:

$(B^\#)$ $[(\sim A \ \& \ B) \vee (B \ \& \ \sim A) \vee (\sim A \ \& \ \sim B)] \square \rightarrow J$

Ze zdania $({}^B T_3)$, (Def), a także prawdziwego w tym wypadku założenia, że nie jest konieczne, że u Leny wystąpi alergia, otrzymujemy zdanie o możliwości:

$({}^B M_3)$ $\diamond(\sim A \ \& \ \sim B)$

Z (SDA') i $({}^B M_3)$ wyprowadzamy zdanie niezgodne z $({}^B T_3)$:

$({}^B T_3^*)$ $(\sim A \ \& \ \sim B) \square \rightarrow J$

Zdanie to jest fałszywe: nie jest prawdą, że gdyby ani Asia, ani Beata nie dostały zwierzaka, to u Leny i tak wystąpiłaby alergia. Ale skoro tak, to albo $(B^\#)$ jest fałszywe, albo fałszywe jest zdanie o możliwości $({}^B M_3)$. Do uznania prawdziwości zdania o możliwości $({}^B M_3)$ jesteśmy zobowiązani — na mocy (Def) — przez uznanie za prawdziwe twierdzenia $({}^B T_3)$. Fałszywe musi więc być $(B^\#)$.

Taki werdykt jest w pełni zgodny z intuicją wyrażoną w §2. W ten sposób można wykazać, że przykład B nie stanowi kontrprzykładu dla semantyki Lowe'a.

6.3. Przykład A

W analogiczny sposób można wykazać, że na gruncie semantyki Lowe'a fałszywe jest zdanie $(A^\#)$.

Zaczynamy od wymienienia tez tła, tj. zdań kontrfaktycznych precyzujących, co by się stało, gdyby na wadze położono określoną liczbę odważników. W tym wypadku możemy ująć te zdania łącznie w następującej formie:

($^A T_n$) Gdyby na wadze położono n odważników, to wskazałaby ona $10n$ g (dla $n = 0, \dots, 42$).

Uznajemy te zdania, zakładając, że waga jest rzetelna. Ich przyjęcie stanowi część rozumienia tego przykładu. Zdania kształtu ($^A T_n$) ponownie nie wydają się kontrolwersyjne.

Stosując (Def), otrzymujemy też odpowiednie twierdzenie: że jest możliwe, iż na wadze znajdzie się n odważników, lub jest konieczne, że waga wskaże $10n$ g. Ponieważ nie jest konieczne, że waga pokaże jakąkolwiek konkretną wartość, prawdziwe są tezy o możliwości:

($^A M_n$) Jest możliwe, że n odważników zostanie położonych na wadze (dla $n = 0, \dots, 42$).

Założmy w takim razie tezę, którą chcemy obalić, tj. kontrprzykład A:

($A^\#$) Gdyby nie wszystkie odważniki zostały umieszczone na wadze, to wskazałaby 410 g lub 400 g, a w każdym razie wartość zbliżoną do 420 g.

Zastępując poprzednik zdania ($A^\#$) zdaniem równoważnym, otrzymujemy:

($A^{\#1}$) Gdyby żaden odważnik nie został umieszczony na wadze lub na wadze znalazły się odważniki w liczbie 1, 2, ... lub 42, to waga wskazałaby 410 g lub 400 g, a w każdym razie wartość zbliżoną do 420 g.

Z ($A^{\#1}$) oraz np. ($^A M_2$) otrzymujemy na mocy (SDA') zdanie:

Gdyby na wadze położono 2 odważniki, to wskazałaby ona 410 g lub 400 g, a w każdym razie wartość zbliżoną do 420 g.

Zdanie to jest jawnie fałszywe: jest niezgodne z ($^A T_2$). Wykazaliśmy tym samym, że fałszywe jest zdanie ($A^\#$).

Werdykt ten jest ponownie w pełni zgodny z intuicją wyrażoną w §2. W ten sposób można wykazać, że przykład A również nie stanowi kontrprzykładu dla semantyki Lowe'a.

6.4. Teza Lowe'a

Wykazanie fałszywości tezy Lowe'a przebiegać będzie analogicznie. Zaczynamy od wymienienia zdań kontrfaktycznych precyzujących, co by się stało w zależności od liczby neuronów, które zachowałyby się tak samo, jak w świecie rzeczywistym. Przyjmujemy, że aktywacja k neuronów stanowi neuronalną przyczynę uniesienia

ręki. Korzystamy tutaj z założenia upraszczającego, a mianowicie, że istotna jest liczba neuronów, które zachowują się tak jak w świecie rzeczywistym. Nawet jeśli rozważamy tylko, czy ręka by się uniosła, to może się oczywiście zdarzyć, że w sytuacji, w której $k-10$ neuronów zachowuje się tak samo jak w świecie rzeczywistym, ręka się uniesie bądź też nie, w zależności od tego, które z neuronów zachowują się podobnie. Nie będziemy tych kwestii szczegółowo rozważać — wystarczy, że wyróżnimy dwie grupy zdań (gdzie $0 \leq i < j \leq k$):

$({}^N T_0)$ Gdyby żaden neuron nie zachował się tak, jak się rzeczywiście zachował, to ręka by się nie uniosła.

$({}^N T_1)$ Gdyby jeden neuron zachował się tak, jak się rzeczywiście zachował, to ręka by się nie uniosła.

...

$({}^N T_{k-j})$ Gdyby $k-j$ neuronów zachowało się tak, jak się rzeczywiście zachowały, to ręka by się nie uniosła.

$({}^N T_{k-i})$ Gdyby $k-i$ neuronów zachowało się tak, jak się rzeczywiście zachowały, to ręka by się uniosła.

...

$({}^N T_{k-1})$ Gdyby $k-1$ neuronów zachowało się tak, jak się rzeczywiście zachowały, to ręka by się uniosła.

$({}^N T_k)$ Gdyby wszystkie neurony zachowały się tak, jak się rzeczywiście zachowały, to ręka by się uniosła.

Również tutaj przyjęcie tych zdań stanowi po prostu część rozumienia „tła”, tj. tego, jak funkcjonuje układ nerwowy rozważanego podmiotu.

Stosując (Def), otrzymujemy też odpowiednie twierdzenia dotyczące możliwości. Skoro nie jest ani konieczne, by ręka się uniosła, ani też nie jest konieczne, by ręka się nie uniosła, to:

$({}^N M_{k-j})$ Jest możliwe, że $k-j$ neuronów zachowa się tak, jak się rzeczywiście zachowały.

$({}^N M_{k-i})$ Jest możliwe, że $k-i$ neuronów zachowa się tak, jak się rzeczywiście zachowały.

Założmy w takim razie tezę Lowe’a:

$(N^\#)$ Gdyby nie wszystkie neurony zachowywały się tak, jak się zachowują, to ręka i tak by się uniosła.

Zastępując poprzednik zdania $(N^\#)$ zdaniem równoważnym, otrzymujemy:

(N[#]) Gdyby żaden neuron nie zachował się tak, jak się rzeczywiście zachował, lub neurony w liczbie 1, 2, ... lub k zachowały się tak samo, jak rzeczywiście się zachowały, to ręka i tak by się uniosła.

Z (N[#]) oraz (N^NM₀) otrzymujemy na mocy (SDA') między innymi zdanie:

Gdyby żaden neuron nie zachował się tak, jak się rzeczywiście zachował, to ręka i tak by się uniosła.

Zdanie to jest jawnie fałszywe — jest niezgodne z (N^T₀).

Wykazaliśmy tym samym, że fałszywe jest zdanie (N), czyli że fałszywa jest teza Lowe'a.

7. ODPOWIEDŹ LOWE'A I JEJ KONSEKWENCJE

Przedstawiona argumentacja za tym, że kontrprzykłady oraz teza Lowe'a są fałszywe, ma następującą postać:

(1)	(Def)	zał.
(2)	Teza (SDA') jest prawdziwa.	(1)
(3)	Tezy tła — (^A T), (^B T), (^N T) — są prawdziwe.	zał.
(4)	Tezy o możliwości — (^A M), (^B M), (^N M) — są prawdziwe.	(1), (3)
(5)	Zatem zdania (A [#]), (B [#]) oraz teza (N [#]) są fałszywe.	(2), (4)

Widzieliśmy wyżej, w jaki sposób poszczególne kroki wynikają z założeń. Wniosek oparty jest z jednej strony na słabej zasadzie symplifikacji alternatywnych poprzedników, którą można wyprowadzić z definicji kontrfaktycznych zdań warunkowych zaproponowanej przez Lowe'a, a z drugiej strony na tezach o możliwości (4).

Jedyną strategią obronną zgodną z semantyką Lowe'a jest odrzucenie tez o możliwości. Jak zobaczymy, odrzucenie ich pociąga na gruncie semantyki Lowe'a pewne konsekwencje. Spróbuję pokazać, że trudno jest je przyjąć. Prześledźmy to najpierw na przykładzie A.

7.1. Konsekwencje przyjęcia kontrprzykładu A na gruncie semantyki Lowe'a

Przyjrzyjmy się, w jaki sposób można odeprzeć wniosek, że zdanie (A[#]) jest fałszywe, i jakie konsekwencje pociąga taka decyzja na gruncie semantyki Lowe'a. Założmy, że z jakichś powodów ktoś chciałby bronić tezy (A[#]):

(A[#]) Gdyby nie wszystkie odważniki zostały umieszczone na wadze, to wskazałyby 410 g lub 400 g, a w każdym razie wartość zbliżoną do 420 g.

Powyższa argumentacja na rzecz fałszywości tego zdania opiera się na zasadzie (SDA¹), ale też na tezach o możliwości. Przyjmowaliśmy, że prawdziwe są zdania kształtu:

(^AM_n) Jest możliwe, że n odważników zostaje położonych na wadze (dla $n = 0, \dots, 42$).

Aby utrzymywać, że teza (^A#) jest prawdziwa, trzeba odrzucić niektóre z tez o możliwości. Zwolennik (^A#) twierdziłby na przykład, że:

(^AM₄₂) Jest możliwe, że 42 odważniki zostaną położone na wadze.

(^AM₄₁) Jest możliwe, że 41 odważników zostanie położonych na wadze.

(^AM₄₀) Jest możliwe, że 40 odważników zostanie położonych na wadze,

jednocześnie przyjmując:

(ⁿA^M₀) Nie jest możliwe, że 0 odważników zostanie położonych na wadze.

(ⁿA^M₁) Nie jest możliwe, że 1 odważnik zostanie położony na wadze.

...

(ⁿA^M₃₉) Nie jest możliwe, że 39 odważników zostanie położonych na wadze.

Mogłoby się wydawać, że przyjęcie przez kogoś założeń o możliwości (^AM₄₀)-(^AM₄₂) i niemożliwości (ⁿA^M₀)-(ⁿA^M₃₉) byłoby cokolwiek arbitralne. Lowe mógłby jednak bronić takiej osoby, sugerując, że przyjmuje ona właśnie takie pojęcie możliwości.

Spór między kimś takim a jego oponentem w sprawie tezy o możliwości (na gruncie semantyki Lowe'a) będzie zbliżony do dyskusji o bliskości światów (na gruncie semantyki Lewisa). Wydaje się więc, że ponownie znajdujemy się w patowej sytuacji: każda ze stron będzie obstawać przy swoim stanowisku bez możliwości przekonania drugiej. Na szczęście na gruncie semantyki Lowe'a pojawia się dodatkowy wymiar tej dyskusji, który pozwala na wyjście z impasu.

Otóż, jak pamiętamy, jedna z różnic między semantykami Lowe'a i Lewisa dotyczy wartości logicznej zdań pusto spełnionych. Według Lewisa jeżeli $\sim\Diamond p$, to zdanie $p \Box \rightarrow q$ jest prawdziwe. Natomiast według Lowe'a, jeżeli $\sim\Diamond p$, to zdanie $p \Box \rightarrow q$ jest prawdziwe, o ile $\Box q$ (a w przeciwnym razie jest fałszywe). To jednak znaczy, że przyjęcie tez (ⁿA^M₀)-(ⁿA^M₃₉) o niemożliwości położenia na wadze od 0 do 39 odważników będzie miało dodatkową konsekwencję dla Lowe'a. Skoro nie jest prawdą, że konieczne jest, by waga wskazała jakąś konkretną wartość liczbową, to zdania:

(^AT₀) Gdyby na wadze położono 0 odważników, to wskazałaby ona 0 g.

(^AT₁) Gdyby na wadze położono 1 odważnik, to wskazałaby ona 10 g.

(^AT₂) Gdyby na wadze położono 2 odważniki, to wskazałaby ona 20 g.

$(^A T_3)$ Gdyby na wadze położono 3 odważniki, to wskazałaby ona 30 g.

....

$(^A T_{39})$ Gdyby na wadze położono 39 odważników, to wskazałaby ona 390 g

są na gruncie semantyki Lowe'a — przy przyjętych założeniach, a w szczególności przy założeniach $(^N M_0)$ - $(^N M_{39})$ — fałszywe.

Zdania o postaci $(^A T_n)$ wydawały się niekontrowersyjne: zdawały jedynie sprawę z rzetelności wagi. Dlatego dziwi uznanie ich za fałszywe. Jest to jednak konsekwencja obrony tezy $(A^\#)$ na gruncie semantyki Lowe'a. Inaczej mówiąc, argumentacja Lowe'a na rzecz tezy $(A^\#)$ będzie przekonująca tylko dla kogoś, kto byłby gotów zaprzeczyć zdaniom $(^A T_0)$ - $(^A T_{39})$. Sprawa przedstawia się analogicznie w wypadku tezy Lowe'a.

7.2. Lowe'a obrona tezy Lowe'a i jej konsekwencje

Fałszywość tezy $(N^\#)$ wykazaliśmy na podstawie słabej zasady symplicyfikacji alternatywnych poprzedników, którą można wyprowadzić z (Def), oraz na podstawie tez o możliwości $(^N M_0)$ - $(^N M_k)$. Jediną strategią obronną zgodną z semantyką Lowe'a jest odrzucenie tez o możliwości. Jednak na gruncie tej semantyki odrzucenie tezy o możliwości wiąże się z uznaniem za fałszywe zdań o postaci $(^N T_0)$ - $(^N T_k)$, a to wydaje się wysoce nieintuicyjne.

Aby obronić prawdziwość tezy Lowe'a, trzeba byłoby przyjąć po pierwsze, że:

$(^N M_0)$ Niemożliwe jest, aby żaden z k neuronów nie zachował się tak, jak się rzeczywiście zachował.

$(^N M_1)$ Niemożliwe jest, aby tylko jeden z k neuronów zachował się tak, jak się rzeczywiście zachował.

$(^N M_2)$ Niemożliwe jest, aby tylko dwa z k neuronów zachowały się tak, jak się rzeczywiście zachowały.

...

Lowe mógłby tu sugerować, że wolno nam odwołać się do czynników pragmatycznych — pamiętamy, że dopuszcza on rozmaite interpretacje modalności. Możemy sobie wyobrazić, że ktoś istotnie zgadza się, że istnieje taki sens możliwości i niemożliwości, przy którym zdania $(^N M_0)$, $(^N M_1)$, $(^N M_2)$, ... są prawdziwe.

Nie od rzeczy jest tu przypomnienie argumentu z bliskości światów. Lowe nie wprowadza co prawda wymogu bliskości, ale spór o to, czy zdania $(^N M_0)$, $(^N M_1)$, $(^N M_2)$, ... są prawdziwe, jest zbliżony do sporu o to, które ze światów są najbliższe światu rzeczywistego. W konkretnym wypadku — a z takim mamy tu do czynienia

— jesteśmy trochę bezradni. Nie za bardzo wiadomo, jakich argumentów użyć, by przekonać oponenta.

Przyjęcie tez o niemożliwości ($n^N M_0$), ($n^N M_1$), ($n^N M_2$), ... wiąże się jednak na gruncie semantyki Lowe'a z wymiernymi konsekwencjami. Otóż jakikolwiek sens możliwości zostałyby pragmatycznie wynegocjowany, to nie będzie to sens, zgodnie z którym konieczne jest, by ręka się uniosła, ani też — by się nie uniosła. Lecz skoro tak, to konsekwencją przyjęcia tez o niemożliwości ($n^N M_0$), ($n^N M_1$), ($n^N M_2$), ... będzie — na gruncie semantyki Lowe'a — odrzucenie tez kontrfaktycznych określających to, co by się stało w sytuacji, gdyby bardzo mało neuronów zachowało się tak, jak się rzeczywiście zachowują. Obrona tezy Lowe'a — w ramach semantyki Lowe'a — wiąże się więc z ogromnymi kosztami, a mianowicie z przyjęciem, że *fałszywe* są zdania:

($^N T_0$) Gdyby żaden z k neuronów nie zachował się tak, jak się rzeczywiście zachował, to ręka by się nie uniosła.

($^N T_1$) Gdyby tylko jeden z k neuronów zachował się tak, jak się rzeczywiście zachował, to ręka by się nie uniosła.

($^N T_2$) Gdyby tylko dwa z k neuronów zachowały się tak, jak się rzeczywiście zachowały, to ręka by się nie uniosła.

...

Zdania ($^N T_0$), ($^N T_1$), ($^N T_2$), ... są jednak ewidentnie prawdziwe (przynajmniej na mocy przyjętych upraszczających założeń). Są to, jak wyżej powiedzieliśmy, zdania kontrfaktyczne, które stanowią po prostu część rozumienia tego, jak funkcjonuje układ nerwowy rozważanego podmiotu. Analogię stanowią tu zdania kształtu ($^B T_n$) z przykładu B: „Gdyby na wadze położono n odważników, to wskazałaby ona $10n$ g (dla $n = 0, \dots, 42$)”. Uznanie takich zdań za fałszywe wydaje się niezrozumiałe w kontekście rozważanych przykładów.

Dopuszczam ewentualność, że ktoś mógłby sobie wyobrazić, dlaczego należałoby uznać takie zdania za fałszywe. Istotne jest jednak to, że *tylko* taka osoba będzie przekonana o prawdziwości tezy Lowe'a (przy założeniu, rzecz jasna, że semantyka Lowe'a jest adekwatna). Pozostali z nas powinni uznać tezę Lowe'a za fałszywą.

8. WNIOSKI

Co by było, gdyby nie wszystkie neurony zachowały się tak, jak się rzeczywiście zachowały? Zgodnie z tezą Lowe'a ręka by się uniosła, tak jak w rzeczywistości. Aby uzasadnić swoje twierdzenie, Lowe stosuje semantykę światów możliwych. Można tę argumentację kwestionować, wskazując, że opiera się na kontrowersyjnym rozumieniu bliskości światów (§4). Jest to jednak strategia argumentacyjna o tyle

niezadowolająca, o ile nieuchronnie prowadzi do impasu; choć nie sposób nie zauważyć, że zarazem ujawnia w ten sposób słabość argumentacji wyjściowej.

Głównym celem artykułu było pokazanie, że choć podobne kontrowersje powstają na gruncie rozumienia zdań kontrfaktycznych zaproponowanego przez samego Lowe'a, to umożliwia ono wyjście z impasu. W semantyce tej obowiązuje słaba zasada symplifikacji alternatywnych poprzedników (§6). Można ją w sposób prosty i na podstawie dodatkowych przesłanek wykorzystać do wykazania fałszywości tezy Lowe'a. Aby bronić tezy Lowe'a (na gruncie jego semantyki), trzeba odrzucić owe przesłanki.

Pierwszy wniosek, który możemy wyciągnąć, jest taki, że dzięki semantyce Lowe'a spór o tezę Lowe'a (i) uwalnia się od jałowej dyskusji o właściwym rozumieniu możliwości oraz (ii) rozstrzyga się w zależności od tego, czy uznamy pewne przesłanki za prawdziwe. Drugi wniosek jest taki, że owe przesłanki wydają się wiarygodne. Trudno sobie w każdym razie wyobrazić, z jakiego powodu ktoś miałby je odrzucić (przy przyjętych przez Lowe'a założeniach upraszczających). Trzeba byłoby m.in. obalić twierdzenie, że gdyby żaden z neuronów, których pobudzenie *de facto* spowodowało uniesienie ręki, nie zachował się tak, jak się rzeczywiście zachował, to ręka by się nie uniosła. Przy przyjętych założeniach upraszczających odrzucenie tej tezy wydaje się po prostu bezzasadne. Co więcej, trudno sobie wyobrazić, jakie uzasadnienie jej odrzucenia można by zaproponować. Dokładnie tak samo nieuzasadniona wydaje się teza Lowe'a.

Można oczywiście zbagatelizować wspomnianą zaletę semantyki Lowe'a i zauważyć, że wcale nie wychodzimy z impasu, a jedynie przenosi się on na inny spór, a mianowicie na kwestię, czy wspomniane tezy tła są prawdziwe, czy nie. Sądzę, że nie musimy tak myśleć. Czynimy bowiem postęp, pokazując, w jaki sposób nasze zobowiązania teoretyczne przekładają się na inne.

BIBLIOGRAFIA

- Jackson F. (1977), *A Causal Theory of Counterfactuals*, „Australasian Journal of Philosophy” 55(1), 3-21.
- Lewis D. (1973), *Counterfactuals*, Cambridge (MA): Harvard University Press.
- Lewis D. (1986), *Counterfactual Dependence and Time's Arrow* [w:] *Philosophical Papers*, t. 2, Oxford: Oxford University Press, 32-66.
- Lowe E. (1983), *A Simplification of the Logic of Conditionals*, „Notre Dame Journal of Formal Logic” 24(3), 357-366.
- Lowe E. (1995), *The Truth about Counterfactuals*, „The Philosophical Quarterly” 45(178), 41-59.
- Lowe E. (2006), *Non-Cartesian Substance Dualism and the Problem of Mental Causation*, „Erkenntnis” 65, 5-23.
- Lowe E. (2008), *Personal Agency. The Metaphysics of Mind and Action*, Oxford: Oxford University Press.

- Paprzycka K. (2013a), *Jak obalić Lowe'a obalenie teorii identyczności jednostkowej, czyli co by się stało, gdyby nie wszystkie neurony zachowały się tak, jak się zachowały*, „Przegląd Filozoficzny — Nowa Seria” 22(2) [86], 471-491.
- Paprzycka K. (2013b), *O argumentacji Lowe'a na rzecz koncepcji przyczynowości intencjonalnej i fizycznej*, „Filozofia Nauki” 21(1) [81], 91-112.
- Tichý P. (1976), *A Counterexample to the Stalnaker-Lewis Analysis of Counterfactuals*, „Philosophical Studies” 29(4), 271-273.