

# Teza Lowe'a w świetle semantyki Lowe'a<sup>1</sup>

Katarzyna Paprzycka

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

(N<sup>#</sup>) gdyby nie wszystkie neurony zachowały się tak jak się zachowują, to ręka i tak by się uniosła.

Sugerowałam, że Lowe'a uzasadnienie dla tej tezy jest wadliwe. Opiera się on na standardowej semantyce światów możliwych, ale jednym z pierwszych problemów tej semantyki jest problem poprzedników o kształcie alternatywy (por. (2013a)). Dla tego typu zdań, których przykładem jest również zdanie (N<sup>#</sup>), semantyka światów możliwych często generuje wyniki niezgodne z naszymi intuicjami. W (2013a) wskazałam też przykłady zdań o dokładnie tym kształcie, z których każde uzyskuje na gruncie semantyki światów możliwych niezgodną z intuicjami wartość logiczną. Podałam tam też strukturę logiczną, w oparciu o którą można wygenerować wiele tego typu przykładów.

Alternatywnym sposobem rozumienia kontrowersyjności tego typu zdań jest zwrócenie uwagi na to, że argument Lowe'a opiera się na rozstrzygnięciach dotyczących tego, jakie światy uznać za światy bliskie. W §4 zasugeruję odpowiedź Lowe'owi z bliskości światów. Jest ona niekonkluzywna, ale niemniej warto ją sformułować. W dalszej części artykułu postaram się pokazać, że podobną odpowiedź można wystosować 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<sup>#</sup>) są fałszywe (§6).

Wartością semantyki Lowe'a w kontekście tej dyskusji jest to, że wskazuje na pewną metodę rozstrzygnięcia sporu. W §7 zobaczymy, w jaki sposób Lowe mógłby zareagować na tego typu argumentację. To, że może on na taką argumentację odpowiedzieć jest niewątpliwe. Ciekawe jest jednak to, że taka odpowiedź pociąga – na gruncie jego semantyki – pewne dalsze zobowiązania. Zobowiązania te są trudne

---

<sup>1</sup> Chciałabym bardzo serdecznie podziękować Mariuszowi Grygiańcowi oraz anonimowemu recenzentowi za uwagi krytyczne do pierwszej wersji tekstu.

do przyjęcia, a koszt przyjęcia tezy Lowe'a na gruncie semantyki Lowe'a jest dużo większy niż jej koszt na gruncie semantyki światów możliwych.

## 1. Teza Lowe'a

Teza Lowe'a występuje w kontekście jego argumentu z implikacji kontrfaktycznych (2006; 2008) przeciw tezie 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, s. 105, podkreślenie autora)

(M) "Gdyby zdarzenie  $D$  nie wystąpiło [...], to ręka podmiotu *w ogóle by się nie uniosła*." (Lowe 2008, s. 105, podkreślenie autora).

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 niniejszym tekście skoncentrujemy się wyłącznie na argumentacji na rzecz (N). Lowe wychodzi od założenia, że  $N$  powoduje pewne behawioralne zdarzenie (egzemplarz)  $B$  (tu: uniesienie ręki). A zatem:

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, 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) podobny do  $B$ , tj. również będący uniesieniem ręki, choć charakteryzującym się nieznacznie odmiennymi parametrami.

Lowe uważa, że może wykazać, że zdanie (N) jest prawdziwe, odwołując się do relacji bliskości między światami, narzuconej przez standardową interpretację zdań kontrfaktycznych. Twierdzi on, że w najbliższych światach możliwych, gdzie nie występuje  $N$ , wystąpi inne zdarzenie neuronalne  $N^*$ , które musi być *bardzo* podobne

do zdarzenia rzeczywistego  $N$ :  $N$  różni się od  $N^*$  tylko 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 wywodzi, że gdyby zdarzenie neuronalne  $N$ , które spowodowało uniesienie ręki  $B$ , nie wystąpiło, to ręka tak czy owak by się uniosła – choć jej uniesienie miałoby nieco odmienne parametry fizyczne (tempo, trajektoria, czas, etc).

Teza (N) odnosi się do bardzo złożonego zdarzenia egzemplarza  $N$ . Zdarzenie  $N$  jest „sumą [...] ogromnej liczby pojedynczych zdarzeń neuronalnych” (Lowe 2008, s. 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”.

Warto zwrócić uwagę, że argument Lowe’a jest oparty na szeregu założeniach 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 tych neuronów nie odgrywa większej roli od pozostałych. Za Lowe’em założenia te przyjmujemy na potrzeby argumentacji.

## 2. Kontrprzykłady

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

**Przykład A.** 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. 420g. 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. Jeżeli nie wiadomo ile odważników zostałyby położonych zamiast czterdziestu dwóch, to nie ma podstaw, by wyróżnić jednoznacznie którąkolwiek z czterdziestu dwu możliwości: mogłoby być tak, że umieszczonych zostaje 41 odważników, 40 odważników, ..., a mogłoby być tak, że położony zostaje 1 odważnik bądź żaden. Jeżeli tak, to odpowiedź na postawione pytanie będzie brzmiała:

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

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

(A<sup>#</sup>) Gdyby nie wszystkie odważniki zostały umieszczone na wadze, to wskazałaby 410g lub 400g lub wartość zbliżoną do 420g.

Jest ono jednak niezgodne z naszymi intuicjami wyrażonymi w (A). Mamy wszelkie podstawy, by przyjąć twierdzenie słabsze:

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

ale (A<sup>#</sup>) jest 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 obie przyjaciółki dostały w prezencie swojego zwierzaka (tj. gdyby nie było prawdą, że obie przyjaciółki dostały taki prezent)?

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 obie Asia i Beata dostały swego zwierzaka, to u Leny *mogłaby* i tak *wystąpić* silna reakcja alergiczna.

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

*Prima facie* odrzucilibyśmy jednak:

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

Stosując semantykę światów możliwych w powyższy sposób, otrzymamy jednak właśnie zdanie (B<sup>#</sup>) raczej niż (B) i (B'). *Ceteris paribus* światy, w których tylko jedna z przyjaciółek otrzymuje swego czworonoga, są bardziej podobne do świata 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 pojawiają się objawy alergiczne.

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

### **3. Co by było gdyby prawie wszystkie neurony zachowywały się tak, jak się zachowały?**

Można byłoby się zastanowić, czy tezy (N) nie należałoby raczej oddać jako:

(N<sup>†</sup>) 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<sup>†</sup>) jest prawdziwa. Na gruncie standardowej semantyki światów możliwych uzasadnienie dla (N<sup>†</sup>) będzie się dokładnie pokrywało z uzasadnieniem dla (N).<sup>2</sup> Jeżeli jednak tak, to można zasadnie przypuszczać, że życzliwość interpretacji skłonić powinna do przyjęcia, że inkryminowaną tezę Lowe'a (N) należy rozumieć jako prawdziwe zdanie (N<sup>†</sup>) raczej niż jako fałszywe zdanie (N<sup>#</sup>).<sup>3</sup>

Tak jednak nie jest. Kluczowe jest bowiem miejsce 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 występują zdarzenia składowe  $N_1, \dots$  i  $N_k$ . Innymi słowy wystąpienie  $N$  polega na tym, że *wszystkie* zdarzenia składowe  $N_1, \dots$  i  $N_k$  występują. Jeżeli tak, to naturalne wydaje się przypuszczenie, że niewystąpienie  $N$  polega na tym, że nieprawdą jest, iż *wszystkie* zdarzenia składowe  $N_1, \dots$  i  $N_k$  występują, tj. na tym, że *nie wszystkie* zdarzenia składowe  $N_1, \dots$  i  $N_k$  występują.

Ktoś, kto dopuszczałby interpretację (N) jako (N<sup>†</sup>), musiałby przyjąć tezę:

(#) złożone zdarzenie (egzemplarz) nie występuje, zawsze i tylko wtedy gdy występują prawie wszystkie zdarzenia (egzemplarze) nań się składające.

*Prima facie* jednak, że teza (#) jest fałszywa ze względu na fałszywość implikacji:

(##) złożone zdarzenie (egzemplarz) nie występuje, tylko wtedy gdy występują prawie wszystkie zdarzenia (egzemplarze) nań się składające.

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

---

<sup>2</sup> W rzeczy samej dokładnie na tym polega problem. Semantyka światów możliwych wydaje się zlewać ze sobą zdania kształtu (N) ze zdaniami kształtu (N<sup>†</sup>).

<sup>3</sup> Jestem wdzięczna Mariuszowi Grygiańcowi za wyraźne sformułowanie tej możliwości.

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 jednak w takim razie Lowe nie mógłby po prostu zarzucić  $(N)$  na rzecz  $(N^\dagger)$ ? Mógłby to oczywiście zrobić, ale taki manewr wytrąciłby jego oręż w argumentacji przeciw teorii identyczności. W obliczu faktu z jednej strony, że:

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

a z drugiej strony, że:

( $M$ ) “Gdyby zdarzenie  $D$  nie wystąpiło [...], to ręka podmiotu *w ogóle by się nie uniosła*” (Lowe 2008, s. 105, podkreślenie autora)

nie wynika przecież zaprzeczenie tezy o identyczności egzemplarycznej  $D$  i  $N$ . Dlatego też życzliwa interpretacja Lowe’a wymaga przypisania mu tezy  $(N)$ , która jest fałszywa.

#### 4. Argument przeciw tezie Lowe’a z bliskości światów

We wszystkich przytoczonych w §2 przykładach podążaliśmy za sposobem, w jaki Lowe wyznacza zbiór najbliższych światów możliwych. Wydaje się, że trudno zaprzeczyć temu, 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 niżli te, w których prawie żaden odważnik nie zostaje położony na wadze. Najbardziej podobne do świata rzeczywistego wydają się światy, w których 41 (ew. 40) odważników zostaje położonych na wadze. 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 światy, w których wszystkie neurony za wyjątkiem jednego (ew. dwóch) zachowują się tak, jak się rzeczywiście zachowują.

Wydaje się, że trudno zaprzeczyć, że wspomniane światy są najbardziej podobne do świata rzeczywistego. Trudno byłoby temu zaprzeczyć, gdyby Lowe sądził, że poprzednik tezy (N) mówi o *typach* zdarzeń. Lowe zastrzega jednak, że ma on na myśli *egzemplarze*. Formułuje on tezę (N) w odniesieniu do jednego złożonego zdarzenia (egzemplarza) neuronalnego *N*, na które składa się taka a nie inna aktywacja ogromnej liczby neuronów (określmy ten złożony typ aktywacji jako **N**): Gdyby *N* (typu **N**) nie wystąpiło, to ręka by się i tak 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 by się i tak uniosła w *dokładnie taki sam* sposób, jak to rzeczywiście miało miejsce.

Ten wniosek sugerowałby, że przyczyny neuronalne w *ogóle* nie wpływają na ruch ręki.<sup>4</sup>

Nie trzeba jednak się wcale zgodzić 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), David Lewis (1973, s. 29) sugeruje, że w ocenie bliskości światów nie musimy brać pod uwagę drobnych różnic między światami.<sup>5</sup> 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 te same prawidłowości obowiązywały w światach bliskich. W drugiej kolejności ważne jest to, aby fakty dotyczące jednostkowych zdarzeń w ogromnej większości się pokrywały. W trzeciej kolejności istotne jest to, aby wykluczyć nawet małe i lokalne pogwałcenia praw

---

<sup>4</sup> Sądzę – wbrew Lowe'owi – że podobny wniosek można byłoby sformułować po stronie mentalnej.

<sup>5</sup> „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, s. 29)



przyrody. Mało istotne – albo w ogóle nieistotne – są odstępstwa dotyczące pojedynczych faktów.<sup>6</sup>

Porównując ze sobą światy, w których nie wszystkie odważniki zostały położone na wadze, musimy wpierw zadbać, by były to światy rządzone tymi samym prawami. Wykluczamy więc np. światy, w których nie obowiązuje prawo grawitacji i w związku z tym położenie na wagę pewnej liczby odważników będzie miało zupełnie inny niż w świecie rzeczywistym skutek. 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. I tak dalej. I tak dalej. Po trzecie nie ma lokalnych odstępstw od praw natury – wykluczamy więc sytuację, gdy kładąc trzydzieści 10g odważników uzyskujemy nagle wynik 3kg, gdyż akurat wystąpiła anomalia. Biorąc pod uwagę rozważane właśnie możliwe różnice między światami, wydaje się, że kwestia, czy na wadze – działającej w takich samych warunkach tak, jak rzeczywiście działa – położymy 41, 40, 39, ..., 1, czy 0 odważników, może być zakwalifikowana co najwyżej do czwartej kategorii ważności, która może – co dopuszcza Lewis – być nieuwzględniona w ocenie bliskości.<sup>7</sup>

Można bronić tezy, że podobny werdykt co do bliskości światów powinniśmy uzyskać, dokonując refleksji 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 etc.) 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.

---

<sup>6</sup> “(1) It is of the first importance to avoid big, widespread, diverse violations of law. (2) It is of the second importance to maximize the spatio-temporal region throughout which perfect match of particular fact prevails. (3) It is of the third importance to avoid even small, localized, simple violations of law. (4) it is of little or no importance to secure approximate similarity of particular fact even in matters that concern us greatly”. (Lewis 1986, s. 47-48)

<sup>7</sup> 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, ale waha się przed kategorycznym przyjęciem takiego stanowiska.

Za chwilę podniosę kwestię kontrowersyjności takiej oceny bliskości. 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 zostało na wadze położonych (przy zachowaniu tych samych praw przyrody, tych samych warunków, braku anomalii). Najbliższymi światami, w których nie wszystkie  $k$  neurony 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 we wcześniejszych pracach (2013ab), okazują się być – na gruncie semantyki światów możliwych – fałszywe. Przestają więc być kontrprzykładami dla tej semantyki. Podobnie oczywiście fałszywa okazuje się teza Lowe'a. W niektórych z najbliższych światach możliwych – w powyższym 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, ręka 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.

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

Taka odpowiedź na zarzut Lowe'a musi się wydać kontrowersyjna i dialektycznie słaba. Kontrowersyjne bowiem jest określenie bliskości światów. Z taką ripostą wypada się zgodzić, ale ripostującemu należy zwrócić uwagę na dwie kwestie. Po pierwsze, o ile taką odpowiedź ocenimy jako kontrowersyjną 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 wtóre, 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. Konkludując, przytoczona odpowiedź Lowe'owi jest nieślabsza niż Lowe'a argumentacja za tezą (N<sup>#</sup>).

## 5. Semantyka Lowe'a

Lowe opiera swoją argumentację na Lewisiańskiej wersji semantyki światów możliwych. Jednakże Lowe sam zaproponował odmienne rozumienie zdań kontrfaktycznych (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 się tym argumentom oprzeć (§7).

Lowe proponuje rozumienie zdań kontrfaktycznych jako specjalnego przypadku implikacji ścisłej:

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

Warto zwrócić uwagę na dwie kwestie. Po pierwsze – w przeciwieństwie do semantyki Lewisa, na gruncie której zdania o niemożliwych poprzednikach są zawsze prawdziwe, bo są pusto spełnione – tak nie jest 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 niej, 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, bo poprzednik jest pusto spełniony we wszystkich światach możliwych. Semantyka Lowe'a różnicuje wartość logiczną tych zdań. Zdanie (1) jest prawdziwe na gruncie semantyki Lowe'a, ponieważ następnik jest koniecznie prawdziwy: dla każdej liczby naturalnej, istnieje liczba od niej większa. Zdanie (2) natomiast jest fałszywe, ponieważ ani poprzednik nie jest możliwy, ani następnik nie jest konieczny.

Lowe utrzymuje też, że różnice między rodzajami okresów warunkowych sprowadzają się do różnic w zakładanym pojęciu możliwości. Rozważmy przytoczane przez niego przykłady<sup>8</sup>:

(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 nie ma różnicy w warunkach prawdziwości dla zdań (3) i (4), natomiast mogą się one różnić dla zdania (5). Łatwo się jest 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)) i 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 on przykład dwóch osób, które, znajdując się w tym samym 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 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) natomiast, wyraża swą ufność w ostrożność Browna, który w ogóle nie zapaliłby żadnej zapalki, gdyby istniała najmniejsza szansa na wybuch.

## **6. Argument przeciw tezie Lowe'a z semantyki Lowe'a**

Zastosujemy propozycję Lowe'a, by podważyć jego tezę (N). Uczynimy to w trzech krokach. Po pierwsze, wykazemy, że Lowe przyjmuje słabszą wersję zasady symplifikacji alternatywnych poprzedników. Po wtóre, pokażemy, jak zastosować

---

<sup>8</sup> 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.

słabszą wersję tej zasady do wykazania fałszywości (B<sup>#</sup>). Po trzecie, zgeneralizujemy ten argument na przykład A i na tezę Lowe'a.

### 6.1. Zasada symplifikacji alternatywnych poprzedników: mocna i słaba wersja

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

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

W (2013a) pokazałam, w jaki sposób można byłoby wykorzystać zasadę symplifikacji alternatywnych poprzedników (SDA) do obalenia kontrprzykładów i tezy Lowe'a. Istnieje jednak szereg zastrzeżeń wobec przyjęcia tej zasady.

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

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

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) <sup>9</sup> |
| (6) $\Diamond p \vee \Box r$          | (2)              |
| (7) $p \Box \rightarrow r$            | (5), (6)         |

Słabą wersję zasady symplifikacji alternatywnych poprzedników można zastosować do wykazania fałszywości przytoczonych wyżej kontrprzykładów (A<sup>#</sup>), (B<sup>#</sup>), a także tezy Lowe'a (N<sup>#</sup>), opierając się na pewnych dodatkowych acz niekontrowersyjnych założeniach. Prześledzimy to najpierw na najprostszym przykładzie B.

---

<sup>9</sup> Dla dowolnego świata, założmy, 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:

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

$({}^B T_2)$   $\sim B \Box \rightarrow J$  (Analogicznie: Gdyby Beata nie dostała zwierzaka, to u Leny wystąpiłaby alergia; nawet gdyby tylko Asia dostała zwierzaka, to u Leny i tak wystąpiłyby symptomy alergiczne.)

$({}^B T_3)$   $(\sim A \ \& \ \sim B) \Box \rightarrow \sim J$  (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 one być kontrowersyjne. Warto też podkreślić, że są one zupełnie niekontrowersyjne na gruncie Lewisa semantyki światów możliwych. Dla każdej tezy tła, semantyka Lewisa generuje zgodny z intuicjami wynik.

Wyobraźmy sobie, że ktoś stwierdza:

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

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

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

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:

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

Stosując słabą zasadę symplifikacji alternatywnych poprzedników (SDA'), a także tezę  $({}^B M_3)$ , otrzymujemy zdanie kontradykcyjne wobec  $({}^B T_3)$ :

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

Zdanie to jest fałszywe: nie jest prawdą, że gdyby ani Asia ani Beata nie dostały zwierzaka, to u Leny i tak wystąpiłaby alergia. Jeżeli jednak 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 zmuszeni – na mocy deficji (Def) Lowe’a – przez uznanie za prawdziwe twierdzenia  $({}^B T_3)$ ; fałszywe zatem musi 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 podejść do wykazania, że zdanie ( $A^\#$ ) jest fałszywe na gruncie semantyki Lowe'a.

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

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

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

Stosując (Def) Lowe'a, otrzymujemy też odpowiednie twierdzenia, że albo jest możliwe, iż  $n$  odważników zostaje położonych na wadze, albo że jest konieczne, iż waga wskaże  $10n$  g. Ponieważ nie jest konieczne, że waga pokaże jakąkolwiek konkretną wartość, więc 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 410g lub 400g lub wartość zbliżoną do 420g.

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

( $A^{\#r}$ ) Gdyby albo żaden odważnik nie został umieszczony na wadze, albo na wadze umieszczone zostały 1, 2 lub 41 odważniki, to waga wskazałaby 410g lub 400g lub wartość zbliżoną do 420g.

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

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

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

Ponownie werdykt ten jest 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. Przyjmujemy tutaj założenie upraszczające, a mianowicie, że kluczowa jest liczba neuronów, które zachowują się tak, jak w świecie rzeczywistym. Nawet jeżeli rozważamy tylko kwestię tego, czy ręka by się uniosła, czy nie, to może się oczywiście zdarzyć, że w wypadku, gdy rozważamy sytuację, w której  $k-10$  neuronów zachowuje się tak samo, jak w świecie rzeczywistym, to 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  $0 \leq i < j \leq k$ ):

(<sup>N</sup>T<sub>0</sub>) Gdyby żaden neuron nie zachował się tak, jak się rzeczywiście zachował, to ręka by się nie uniosła.

(<sup>N</sup>T<sub>1</sub>) Gdyby jeden neuron zachował się tak, jak się rzeczywiście zachował, to ręka by się nie uniosła.

...

(<sup>N</sup>T <sub>$k-j$</sub> ) Gdyby  $k-j$  neuronów zachowało się tak, jak się rzeczywiście zachowały, to ręka by się nie uniosła

(<sup>N</sup>T <sub>$k-i$</sub> ) Gdyby  $k-i$  neuronów zachowało się tak, jak się rzeczywiście zachowały, to ręka by się uniosła

...

(<sup>N</sup>T <sub>$k-1$</sub> ) Gdyby  $k-1$  neuronów zachowało się tak, jak się rzeczywiście zachowały, to ręka by się uniosła

(<sup>N</sup>T <sub>$k$</sub> ) Gdyby wszystkie neurony zachowały się tak, jak się rzeczywiście zachowały, to ręka by się uniosła



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

Stosując (Def) Lowe’a, otrzymujemy też odpowiednie twierdzenia dotyczące możliwości. Ponieważ nie jest ani konieczne by ręka się uniosła, ani też nie jest konieczne, by się ręka nie uniosła, stąd:

$({}^N M_{k-j})$  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 albo żaden neuron nie zachował się tak, jak się rzeczywiście zachował, albo tak samo jak rzeczywiście się zachowały, zachowałyby się 1, 2 lub  $k$  neurony, to ręka i tak by się uniosła.

Z  $(N^\#')$  oraz  $({}^N M_0)$ , otrzymujemy na mocy (SDA') na przykład 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  $({}^N T_0)$ .

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

## 7. Odpowiedź Lowe’a i jej konsekwencje

Powyższa argumentacja, że kontrprzykłady oraz teza Lowe’a są fałszywe ma następujący kształt.

- |   |             |
|---|-------------|
| (1) (Def)   | (założenie) |
| (2) Teza (SDA') jest prawdziwa.                                     | (1)         |
| (3) Tezy tła $(({}^A T), ({}^B T), ({}^N T))$ są prawdziwe          | (założenie) |
| (4) Tezy o możliwości $(({}^A M), ({}^B M), ({}^N M))$ są prawdziwe | (3), (1)    |
| (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 w oparciu o propozycję (Def) Lowe'a, a z drugiej strony na tezach o możliwości (4).

Jedyną strategią odpowiedzi, która jest dostępna Lowe'owi jest odrzucenie tez o możliwości. Jak zobaczymy, odrzucenie tych tez jest na gruncie semantyki Lowe'a związane z pewnymi konsekwencjami. Będę się starała pokazać, że trudno jest te konsekwencje przyjąć. Prześledźmy to najpierw na przykładzie A.

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

Przyjrzyjmy się, w jaki sposób można byłoby się oprzeć wnioskowi, że zdanie ( $A^\#$ ) jest fałszywe i jakie taka decyzja niesie konsekwencje na gruncie semantyki Lowe'a. Załóżmy, że z jakichś powodów, ktoś chciałby bronić tezy ( $A^\#$ ):

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

Powyższy argument na rzecz fałszywości tego zdania opiera się na zasadzie (SDA'), ale też na tezach o możliwości. Przyjmowaliśmy, że prawdziwe są zdania kształtu:

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

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

( $^A M_{42}$ ) Jest możliwe, że 42 odważniki zostaną położone na wadze,

( $^A M_{41}$ ) Jest możliwe, że 41 odważników zostanie położonych na wadze,

( $^A M_{40}$ ) Jest możliwe, że 40 odważników zostanie położonych na wadze,

jednocześnie przyjmując:

( $^n A M_0$ ) Nie jest możliwe, że 0 odważników zostanie położonych na wadze,

( $^n A M_1$ ) Nie jest możliwe, że 1 odważnik zostanie położony na wadze,

...

( $^n A M_{39}$ ) Nie jest możliwe, że 39 odważników zostanie położonych na wadze.

Mogłoby się wydawać, że przyjęcie założeń o możliwości ( $^A M_{40}$ )-( $^A M_{42}$ ) i niemożliwości ( $^n A M_0$ )-( $^n A M_{39}$ ) jest cokolwiek arbitralne. Lowe mógłby tu jednak

takiego kogoś bronić, sugerując, że możliwe jest, iż ten ktoś takie pojęcie możliwości przyjmuje.

Spór między kimś takim a jego oponentem o tezy o możliwości (na gruncie semantyki Lowe'a) będzie zbliżony do dyskusji o bliskości światów (na gruncie semantyki Lewisa). Wydaje się zatem, że ponownie mamy do czynienia z punktem, gdzie dochodzi o impasu: każda ze stron będzie obstawać przy swoim stanowisku bez możliwości przekonania drugiej. Tak jednak nie jest. Albowiem na gruncie semantyki Lowe'a pojawia się dodatkowy wymiar tej dyskusji, który sprawia, że daje ona szansę na wyjście z impasu.

Otóż – jak pamiętamy – jedna z różnic między semantyką Lowe'a a Lewisa dotyczy wartości logicznej pusto spełnionych zdań kontrfaktycznych. Według Lewisa jeżeli  $\sim\Diamond p$ , to zdanie ' $p \Box \rightarrow q$ ' jest prawdziwe. Według Lowe'a tak jednak nie jest: jeżeli  $\sim\Diamond p$ , to zdanie ' $p \Box \rightarrow q$ ' jest prawdziwe, o ile  $\Box q$ , a fałszywe – w przeciwnym wypadku. To jednak znaczy, że przyjęcie tez ( $n^A M_0$ )-( $n^A M_{39}$ ) o niemożliwości położenia na wadze od 0 do 39 odważników będzie miało dodatkową konsekwencję dla Lowe'a. Ponieważ nie jest prawdą, że konieczne jest, by waga wskazała jakąś konkretną wartość liczbową, więc zdania:

( $A T_0$ ) gdyby na wadze położono 0 odważników, to wskazałaby ona 0 g,

( $A T_1$ ) gdyby na wadze położono 1 odważnik, to wskazałaby ona 10 g,

( $A T_2$ ) 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 ( $n^A M_0$ )-( $n^A M_{39}$ ) – fałszywe.

Zdania kształtu ( $A T_n$ ) 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 rzecz ujmując, argumentacja Lowe'a na rzecz tezy ( $A^\#$ ) będzie przekonująca tylko dla kogoś, kto byłby gotów zaprzeczyć zdaniom ( $A T_0$ )-( $A T_{39}$ ). Będzie nieprzekonująca dla pozostałych z nas. Sprawa przedstawia się analogicznie dla tezy Lowe'a.

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

Fałszywość tezy (N) wykazaliśmy, opierając się na słabej zasadzie symplifikacji alternatywnych poprzedników, którą można wyprowadzić w oparciu o propozycję (Def) Lowe'a, oraz w oparciu o tezy o możliwości  $({}^N M_0)$ - $({}^N M_k)$ . Jediną strategią odpowiedzi, która jest dostępna Lowe'owi jest odrzucenie tez o możliwości. Na gruncie semantyki Lowe'a jednakże odrzucenie tezy o możliwości wiąże się z uznaniem za fałszywe zdań kształtu  $({}^N T_0)$ - $({}^N T_k)$ , a to wydaje się wysoce kontrintuicyjne.

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

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

$(n^N M_1)$  niemożliwe jest, aby tylko jeden z  $k$  neuronów zachował się tak, jak się rzeczywiście zachowały,

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

...

Lowe mógłby tu sugerować, że możemy odwołać się do czynników pragmatycznych – pamiętamy, że pozwala on na rozmaite interpretacje modalności. Możemy sobie wyobrazić, że ktoś istotnie zgadza się, że istnieje taki sens „możliwości” i „niemożliwości”, dla którego zdania  $(n^N M_0)$ ,  $(n^N M_1)$ ,  $(n^N M_2)$ , ... są prawdziwe.

Nie od rzeczy jest tu przypomnieć argument przeciw tezie Lowe'a z bliskości światów. Lowe nie wprowadza co prawda wymogu bliskości, ale spór o to, czy zdania  $(n^N M_0)$ ,  $(n^N M_1)$ ,  $(n^N M_2)$ , ... są prawdziwe czy fałszywe, jest zbliżony do sporu o to, które ze światów są najbliższe światu rzeczywistego. W konkretnym wypadku – takim, z jakim mamy tu do czynienia – jesteśmy trochę bezradni. Nie za bardzo wiadomo, jakich argumentów użyć.

Jednakże przyjęcie tez o niemożliwości  $(n^N M_0)$ ,  $(n^N M_1)$ ,  $(n^N M_2)$ , ... wiąże się z wymiernymi konsekwencjami na gruncie semantyki Lowe'a. Otóż jakkolwiek sens możliwości zostałby pragmatycznie wynegocjowany, to nie będzie to sens, zgodnie z którym konieczne jest, by ręka się uniosła, ani też by się nie uniosła. Jeżeli jednak 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 zachowały się tak, jak się rzeczywiście zachowują. Obrona tezy Lowe'a – na gruncie semantyki Lowe'a – pociąga zatem za sobą ogromny koszt, a mianowicie przyjęcie, że *falszywe* są zdania:

(<sup>N</sup>T<sub>0</sub>) gdyby żaden z  $k$  neuronów nie zachował się tak, jak się rzeczywiście zachował, to ręka by się nie uniosła,

(<sup>N</sup>T<sub>1</sub>) gdyby tylko jeden z  $k$  neuronów zachował się tak, jak się rzeczywiście zachował, to ręka by się nie uniosła,

(<sup>N</sup>T<sub>2</sub>) gdyby tylko dwa z  $k$  neuronów zachowały się tak, jak się rzeczywiście zachowały, to ręka by się nie uniosła,

...

Zdania (<sup>N</sup>T<sub>0</sub>), (<sup>N</sup>T<sub>1</sub>), (<sup>N</sup>T<sub>2</sub>), ... wydają się jednak ewidentnie prawdziwe (przynajmniej na gruncie przyjętych upraszczających założeń). Są to – jak powiedzieliśmy – zdania kontrfaktyczne, które stanowią po prostu część rozumienia tego, jak funkcjonuje układ nerwowy rozważanego podmiotu. Analogię stanowią tu zdania kształtu (<sup>B</sup>T <sub>$n$</sub> ) 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 *prima facie* niezrozumiałe w kontekście rozważanych przykładów.

Być może 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 uniosłaby się podobnie, jak rzeczywiście się uniosła. Lowe stosuje semantykę światów możliwych, by uzasadnić swą tezę. Można tę argumentację kwestionować, wskazując, że opiera się na kontrowersyjnym rozumieniu bliskości światów (§4). Jest to jednak słaba strategia argumentacyjna o tyle, że wydaje się nieuchronnie prowadzić do impasu, choć nie sposób nie zauważyć, iż ukazuje w ten sposób również słabość argumentacji wyjściowej.

Głównym celem tekstu było pokazanie, że choć podobne kontrowersje powstają na gruncie Lowe'a rozumienia zdań kontrfaktycznych, to niewątpliwą zaletą tej semantyki jest to, że pozwala ona na wyjście z impasu. Na gruncie semantyki Lowe'a obowiązuje słaba zasada symplifikacji alternatywnych poprzedników (§6). Można ją w prosty sposób wykorzystać do wykazania fałszywości tezy Lowe'a, opierając się na pewnych przesłankach. Aby bronić tezy Lowe'a (na gruncie jego semantyki), trzeba odrzucić owe przesłanki.

Pierwszy wniosek, który możemy wyprowadzić to ten, że na gruncie semantyki Lowe'a spór o tezę Lowe'a wychodzi z impasu dyskusji o to, co jest możliwe, oraz że pozwala na rozstrzygnięcie sporu w zależności od tego, czy uznamy pewne przesłanki za prawdziwe. Drugi wniosek to ten, że owe przesłanki wydają się być wiarygodne. Trudno sobie w każdym razie wyobrazić, dlaczego ktoś miałby je odrzucić (przy przyjętych założeniach upraszczających). Trzeba byłoby m.in. odrzucić tezę, 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 nieuzasadnione. Co więcej trudno sobie wyobrazić, jakie uzasadnienie dla odrzucenia tej tezy (poza teoretycznym, przytaczanym wyżej) można byłoby zaproponować. Dokładnie tak samo nieuzasadniona wydaje się teza Lowe'a.

Można oczywiście próbować zbagatelizować wspomnianą wyżej zaletę semantyki Lowe'a. Można bowiem zwrócić uwagę, że wcale nie pozbywamy się 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. Odnotowujemy bowiem postęp, pokazując, w jaki sposób nasze zobowiązania teoretyczne przekładają się na inne.

## Cytowana literatura

Jackson, F. (1977) — "A Causal Theory of Counterfactuals", *Australasian Journal of Philosophy* 55, 3-21.

Lewis, D. (1973) — *Counterfactuals*, Cambridge, MA, Harvard University Press.

- Lewis, D. (1986) — “Counterfactual Dependence and Time’s Arrow”, [w:] tegoż *Philosophical Papers*, tom II, Oxford: Oxford University Press, s. 32-66.
- Lowe, E. (1983) — “A Simplification of the Logic of Conditionals”, *Notre Dame Journal of Formal Logic* 24, 357-366.
- Lowe, E. (1995) — “The Truth about Counterfactuals”, *The Philosophical Quarterly* 45, 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 nuerony zachowały się tak, jak się zachowały”, *Przegląd Filozoficzny* (w druku).
- Paprzycka, K. (2013b) — „O argumentacji Lowe’a na rzecz koncepcji przyczynowości intencjonalnej i fizycznej”, *Filozofia Nauki* 21 nr 1 (81), 91-112.
- Tichy, P. (1976) — “A Counterexample to the Stalnaker-Lewis Analysis of Counterfactuals”, *Philosophical Studies* 29(4), 271-273.