

Mieszko Tałasiewicz

Dlaczego nie lubię semantyki światów możliwych¹

Słowa kluczowe: światy możliwe, semantyka, znaczenie, modalność

1. Typowym zastosowaniem teorii światów możliwych w semantyce jest definicja znaczenia zdania (*vel* sądu logicznego) jako zbioru światów możliwych, w których to zdanie jest prawdziwe (lub jako funkcji od światów możliwych do wartości logicznych – co w logice dwuwartościowej na jedno wychodzi). W myśl tej definicji zdania takie jak:

- (1) „George W. Bush w 2007 r. był prezydentem Stanów Zjednoczonych” i
- (2) „W roku akademickim 2007/2008 seminarium ‘Wybrane zagadnienia filozofii języka’ odbywało się w sali nr 25 Instytutu Filozofii UW”

różnią się znaczeniem w tym i tylko w tym, że choć oba są *de facto* prawdziwe, to *mogą* mieć różne wartości logiczne.

Zwolennicy tej koncepcji przyznają, że jest to definicja kłopotliwa ze względu na zdania konieczne (matematyczne czy logiczne), które w myśl tej definicji mają to samo znaczenie (wszystkie są prawdziwe w każdym świecie możliwym), a także ze względu na zdania sądzeniowe o prawdach koniecznych – pojawia się tzw. paradoks wszechwiedzy matematycznej (jeżeli ktoś wie i rozumie, że $2+2=4$, to tym samym zna znaczenie wszystkich prawd matematyki i logiki – bo one znaczą to samo co ‘ $2+2=4$ ’). Zwolennicy owi uważają jednak, że elegancja i trafność, z jaką definicja ta ujmuje znaczenie zdań pozostałych – które przecież stanowią zasadniczy zrąb języka (zdania matematyczne to jednak pewien margines; a ich status poznawczy jest tak czy owak niejasny) – z nawiązką rekompensują ten niedostatek.

Twierdzę, na przekór takiej opinii, że analogiczna trudność obejmuje wszystkie zdania. To prawda, że miło nam się sądzi, że świat rzeczywisty jest w niektórych przynajmniej aspektach przygodny, że mógłby być inny, niż jest. Ja sam tak przy-

¹ Dziękuję dr. Tadeuszowi Ciecierskiemu za inspirujące dyskusje i materiały.

puszczam, a na tym przypuszczeniu funduję swój światopogląd. Jest to bowiem w gruncie rzeczy wiara typu religijnego: wierzę w to, że mam wybór i że za to, jak wybrałem, będę kiedyś sądzony. Ale jako naturalista metodologiczny nie uważam, by było rzeczą słuszną i roztropną mieszanie przekonań religijnych do argumentów naukowych i filozoficznych. Poza tą wiarą – w samym świecie nic by się nie zmieniło, gdyby wszystko, co się tu dzieje, działa się z konieczności. Taki pogląd jest jak najbardziej do pomyślenia: że to, co jest, jest konieczne. Że świat aktualny nie jest przygodny. Że nie ma innych niż nasz światów możliwych. Pogląd taki był niejednokrotnie wyrażany w historii myśli ludzkiej.

W gruncie rzeczy dla problemu znaczenia nie ma różnicy, jak to jest z tym naprawdę. Wystarczy, że świat rzeczywisty ujęty jako konieczny (a zatem jedyny możliwy) jest empirycznie nieodróżnialny od tego samego świata ujętego jako tylko jeden z wielu możliwych. W świecie rzeczywistym dostrzegamy zaś jak najwyraźniej różnicę znaczenia zdań (1) i (2). A skoro tak, to moglibyśmy rozpoznawać równoważne zdania o różnych znaczeniach, nawet gdyby wszystkie zdania były konieczne. A to podważa adekwatność możliwościwiatowej definicji znaczenia w całym spektrum, by tak rzec. Znaczenia to nie są niezrealizowane możliwości.

2. Tak naprawdę ta definicja po prostu nic nie wyjaśnia, bo przecież nie mamy poznawczego dostępu do wszystkich niezrealizowanych możliwości (a różnice znaczeniowe chwytały bardzo łatwo). Jeżeli dbamy choć trochę o psychologiczną realność budowanych przez nas reprezentacji semantycznych, musimy z tej definicji zrezygnować. Budujemy znaczenia z drobnych kawałków świata, a nie z wielu całych światów. Definicja możliwościwiatowa usypia nas, dając nam złudzenie rozwiązania problemu. A problem jest i warto go żmudnie rozwiązywać, badając procedury ostensywne, akwizycję języka itp.

Jest to szczególnie ważne w wypadku zdań niezawierających spójników modalnych. Ale także i modalności – od których cała ta kosmologia się zaczęła – nie wydają się bliższe naszemu rozumieniu z tego tylko powodu, że wyrażono je w terminach światów możliwych. Zgoda, światy możliwe nadają się dobrze do przedstawienia aksjomatyki formalnych systemów modalnych; są one tym dla języków modalnych, czym tabelki prawdziwościowe dla KRZ. Ale to nie jest przecież cała semantyka. Semantyka dostatecznie bogatego języka musi zawierać warunki prawdziwości dla pojedynczego zdania *p*. W wypadku modalnym – musi ostatecznie precyzować, co rozumiemy przez możliwość i kiedy coś jest możliwe, a kiedy nie jest. A tego światy możliwe nie mogą nam dać, chociaż próbują.

Zastanówmy się, co to np. znaczy: „Możliwe jest, że człowiek stanie na Marsie”. Semantyka światów możliwych wyjaśnia nam, że istnieje taki świat możliwy, w którym człowiek stoi na Marsie. *I fail to see* – jak mawiają Anglosasi (to jest nb. wspaniałe wyrażenie) – jakim sposobem wyjaśnienie to mogłoby dać mi

coś więcej niż to, że możliwe jest, że człowiek stanie na Marsie. Pytając o to eksperta, oczekiwałbym raczej wykazania, że:

w chwili t_0 (obecnej) uzasadnione (zgodnie ze standardami metodologicznymi odpowiednich dyscyplin) jest przekonanie, że w pewnej chwili późniejszej t , okres bezpiecznego przebywania człowieka w stanie nieważkości będzie wynosił x (gdzie przez bezpieczny czas będę rozumiał okres taki, że wielu astronautów przed chwilą t spędziło ten okres w stanie nieważkości i nic im się nie stało, a niewielu było takich, którzy okres ten w stanie nieważkości spędzili i coś im się złego stało), a zarazem czas przelotu z Ziemi na Marsa i z Marsa na Ziemię będzie mniejszy niż x i ponadto warunki w pojeździe kosmicznym podczas lądowania na Marsie i startu z tej planety będą nie gorsze niż te, które przed czasem t będą/były testowane na ludziach z dobrym skutkiem itd., itp.

Tych czynników jest bardzo dużo, a stwierdzenie możliwości jest takim ich postrzeżeniem łącznym, podobnym zapewne do postrzeżenia postaci (*gestalt*); jest poniekąd skrótem myślowym. Dlatego, gdy chcemy „na szybko” mieć definicję możliwości – i to definicję ogólną, dostosowaną do dużej klasy zdań modalnych – wyliczanie tych czynników jest bardzo kłopotliwe. Ale nie usprawiedliwia to definicji przez światy możliwe. Gdy przeprowadza się studium wykonalności jakiegoś projektu, to nie zagląda się do światów możliwych, żeby zobaczyć, czy ten projekt tam gdzieś jest zrealizowany, tylko właśnie analizuje się takie drobne czynniki. Takie samo pytanie zadawano sobie przecież najzupełniej na serio – przed lądowaniem na Księżycu. I z całą pewnością autorzy projektu „Apollo” nie zaczytywali się w *Counterfactuals*, tylko zmagali się z mnóstwem faktów, twardych faktów tego świata. Kiedy mówimy o możliwościach, *de facto* mówimy o świecie aktualnym, o tym, że to jego struktura coś dopuszcza, a coś innego wyklucza. Dobra semantyka zdań modalnych powinna koncentrować się na uwidocznieniu tych zależności, tymczasem semantyka światów możliwych właśnie to ignoruje.

Być może złudzenie informatywności semantyki światów możliwych bierze się z pewnego jej podobieństwa (a także, zdaje się, genealogicznego pokrewieństwa) z semantyką indeksową proponowaną dla zdań w szerokim sensie okazjonalnych (w szczególności czasowych i aspektowych). Cóż to np. znaczy, że Ala miała kota? – ano to, że istnieje taki indeks czasowy $t < t_0$ (gdzie t_0 jest czasem wypowiedzenia tego zdania), że Ala ma kota. Co to znaczy, że zgubiłem klucze? To, że $\exists t < t_0$ (autor wypowiedzenia gubi klucze). Itd. Takie indeksy można mnożyć. I należy je mnożyć, jeżeli chcemy postępować na drodze coraz lepszego i dokładniejszego opisu zależności semantycznych języka naturalnego. Jeżeli na przykład chcemy – a cel taki przyświeca bardzo licznym współczesnym badaniom nad językiem – nauczyć komputery mówić ludzkim głosem, to możemy próbować realizować ten zamiar

właśnie przez wyodrębnienie jak największej liczby dobrze określonych indeksów i relacji między nimi. To jest ciężka praca wymagająca współpracy interdyscyplinarnej z lingwistami i psychologami, ale podjęcie tej pracy jest niezbędne w obecnym stadium rozwoju semantyki. Jest to stadium dość już zaawansowane, w którym grube idealizacje pionierskich czasów Fregego i Russella przestają wystarczać. Tak jak w balistyce dawno już nie wystarczają prawa ruchu Galileusza i Newtona i trzeba korzystać z indeksów reprezentujących siłę wiatru, opór powietrza, rotację i rozkład masy pocisku itp. Artylerzystom jednak nie przyjdzie raczej do głowy pomyśleć, że skoro tych indeksów robi się tak dużo, to może warto by je wszystkie odfajkować jakimś symbolicznym *w* i traktować zbiorczo. Zażądają raczej szybszych komputerów. Tymczasem z jakiegoś powodu niektórzy filozofowie po konstatacji, że „indeksy można mnożyć”, zamiast je pracownie mnożyć – koniecznie chcą je zredukować, znajdując wygodną wymówkę właśnie w koncepcji światów możliwych. Światy możliwe są więc czasem ujmowane jako takie worki nieodróżnicowanych indeksów, a semantyka światów możliwych – jako rzekome „rozwiniecie” semantyki indeksowej.

Tylko że to „rozwiniecie” jest dużo gorsze od poprzedniczki. Wyodrębnione indeksy naprawdę są informatywne: uczą nas, że różne aspekty świata mają swój udział w interpretacji wyrażen i pozwalają coraz dokładniej ujmować niuanse znaczeniowe. Indeksy w worku – nie są w stanie tego dokonać. Ale – co najważniejsze – analogia między indeksami a światami możliwymi jest o tyle chybiona, o ile wyodrębnione indeksami aspekty naszego świata należą właśnie do naszego świata, są realne – na przykład jego przekroje czasowe są z przekrojem teraźniejszym powiązane relacjami kauzalnymi (dzięki którym obecnie istnieją świadectwa przeszłości i zadatki na przyszłość); tymczasem światy możliwe są – jeżeli przyjmiemy w poczet ich intuicji założycielskich ideę przestrzeni logicznej Wittgensteina – logicznie niezależne od siebie nawzajem i przede wszystkim niezależne od naszego, rzeczywistego świata.

3. I dlatego właśnie nie lubię możliwych światów. Dałbym im spokój, gdyby były nieszkodliwym dziwactwem. Ale, niestety, światy możliwe nie są wcale tak zupełnie nieszkodliwe. Sprowadziły na złą drogę dobrze zapowiadającą się semantykę indeksową, która mogła stawać się coraz subtelniejszym narzędziem analizy rzeczywistych procesów referencji i przewodniczką po szlakach naszego myślenia, a stała się fantazjowaniem, co by było, gdyby.

Niebezpieczna jest także zwodnicza łatwość, z jaką w semantyce światów możliwych formułować można definicje różnych trudnych a kluczowych pojęć. O definicji znaczenia już mówiliśmy. O tym, jak prosta i elegancka to definicja – i jak perfidnie nieadekwatna.

Podobnie jest z klasyczną już teorią kontrfaktycznych okresów warunkowych. Semantyka ta mówi, że okres typu ‘gdyby zachodziło *p* (które *de facto* nie zacho-

dzi), to zachodziłoby też q jest prawdziwy zawsze i tylko wtedy, gdy w światach możliwych podobnych do naszego, gdzie jednak zachodzi p (które u nas nie zachodzi), zachodzi też q .

Zwolennicy semantyki światów możliwych zgodnie zazwyczaj przyznają, że definicja ta ma dość słaby punkt w miejscu, w którym mówi się o „podobieństwie” owych światów możliwych i naszego świata. Ale zarazem bronią się, że intuicje są tu w każdym razie dość jasne: chodzi o to, że w owych światach mają działać te same prawa i zachodzić te same okoliczności ważne dla danego zdarzenia, poza samym faktem zaistnienia poprzednika (który u nas nie zachodzi). To dzięki tym prawom i okolicznościom (stanowiącym dla tych praw warunki początkowe) jesteśmy w stanie przewidzieć, jak wyglądałyby te możliwe światy, gdyby zaszło w nich p . Bo skąd my to możemy wiedzieć? Tylko stąd, że rozpoznajemy jakieś prawa (astronomii) lub quasi-prawa (historii), które obowiązują u nas i które przy warunkach początkowych p wygenerują q . I tak naprawdę faktyczne istnienie tych praw daje nam poczucie prawdziwości *counterfactuals*².

Zgoda. To wyjaśnia sprawę. A raczej należałoby powiedzieć: to wyjaśniałoby sprawę, gdyby określenie warunków prawdziwości w terminach światów możliwych było pewną skrótową formą wyrażania się (odwoływanie się za każdym razem do praw i okoliczności może być kłopotliwe), ale *de facto* byłaby to wciąż charakterystyka w terminach praw i warunków początkowych. Tymczasem jednak bierze się tryumfalnie możliwościową semantykę kontrfaktycznych okresów warunkowych jako rzecz niezależną i za jej pomocą definiuje się pojęcie praw! Albowiem kolejną „zdobyczą” semantyki światów możliwych jest to, iż rzekomo pozwalają one odróżnić prawa od zdań przygodnie prawdziwych. Przygodnie prawdziwe (załóżmy) zdanie „Wszystkie osoby znajdujące się w tym pokoju to mężczyźni” tym się różni od prawa „Wszystkie planety krążą po elipsach”, że to drugie pozwala uzasadnić kontrfaktyczny okres warunkowy „jeżeli odkrylibyśmy nową planetę, to poruszałaby się ona po elipsie”, a to pierwsze nie pozwala uzasadnić analogicznego zdania „gdyby zwycięzca ogólnopolskiego konkursu recytatorskiego był w tym pokoju, to byłby mężczyzną”. Ale co to znaczy, że gdybyśmy odkryli nową planetę, to poruszałaby się ona po elipsie? To znaczy, w myśl teorii światów możliwych, że w światach podobnych do naszego, w których ta planeta została odkryta, porusza się ona po elipsie. Ale jakie to są światy podobne? Ano takie, w których obowiązuje prawo „Wszystkie planety poruszają się po elipsie” (bo inne światy nie są dość podobne – i tam ten kontrfaktyczny okres warunkowy by nie działał).

² A czasem złudzenie prawdziwości – czy rzeczywiście na przykład historia Europy potoczyłaby się inaczej, gdyby Napoleon wygrał pod Waterloo? A co, jeśli nie? Historycy nie cierpią gdybania; książka *Co by było gdyby* jest wyjątkiem, a poza tym tak naprawdę mówi o realiach, a nie o gdybaniach. Opiera się na wypukleniu pewnych prawidłowości społecznych i ekonomicznych, których istnienie faktyczne domniemujemy.

W ten sposób definicja prawa za pomocą kontrfaktycznych okresów warunkowych prowadzi do regularnego błędnego koła, ukrytego przez teorię światów możliwych, która zwodniczo pozwala nam „zapomnieć”, że w definicji kontrfaktycznych okresów warunkowych uwikłane jest pojęcie prawa. Nie możemy definiować okresu przez prawo, a prawa przez okres³.

* * *

Przykłady można by mnożyć. Skorzystam jednak z „dobrych” wzorów i nie będę wcale mnożył, tylko przejdę od razu do konkluzji tytułowej: nie lubię światów możliwych.

Why I Don't Like Possible Worlds

Key Words: possible worlds, semantics, meaning, modality

The author argues that possible worlds semantics is counterintuitive and fails to achieve its goals. The definition of the meaning of a sentence as the set of possible worlds in which the sentence is true is criticized as inadequate; the semantics of modal expressions is claimed uninformative. Truth-conditions of counterfactuals on one hand, and definitions of law-like expressions (or statements expressing causal connection) on the other hand, are regarded as circular. A further development of indexical semantics – a more fine-grained one, involving interdisciplinary research into the relations among the indices – is advocated as an alternative.

³ *Mutatis mutandis* podobną argumentację należałoby zastosować do możliwościowej definicji przyczynowości. To także błędne koło – z tych samych powodów: przyczynowość może być wykorzystana w ustanowieniu warunków prawdziwości kontrfaktycznego okresu warunkowego.