Marian Przełęcki

Studia z metodologii formalnej

Filozofia Nauki 1/2/3, 23-399

1993

Artykuł został zdigitalizowany i 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.

MUZEUM HISTORII POLSKI

Filozofia Nauki
Rok I, 1993, Nr 2-3

Operacjonizm

I

1. Drugie ćwierćwiecze naszego stulecia charakteryzuje wyraźne ożywienie tendencji pozytywistycznych. Pozytywizm wypierany w początkowych latach XX wieku przez kierunki metafizyczne przechodzi w latach trzydziestych do nowej ofensywy. Występując w postaci nadzwyczaj radykalnej, skupia w swoich szeregach zarówno filozofów, jak i uczonych specjalistów. Jego najbardziej reprezentatywną odmianę stanowi powstały w Europie neopozytywizm. W innej nieco postaci pojawiły się tendencje pozytywistyczne na terenie Ameryki. W roku 1927 ukazało się tam dzieło znanego fizyka P. W. Bridgmana pt. *The Logic of Modern Physics*. Dzieło to zainicjowało nowy kierunek w metodologii nauk, znany pod nazwą „operacjonizmu". Kierunek ten stanowi dzisiaj jeden z ważniejszych nurtów metodologii zachodniej. Szereg filozofów i przedstawicieli nauk szczegółowych, interesujących się metodologią swych specjalności, uważa się za jego zwolenników. Jego przedstawieniu, krytyce i obronie poświęcono wiele publikacji. Operacjonizm należy niewątpliwie do doktryn pozytywistycznych. Szczególnie bliskie pokrewieństwo łączy go z neopozytywizmem. Kierunki te, mimo iż powstały niezależnie od siebie, wykazywały od początku daleko idącą zbieżność. Zbieżność ta pogłębiła się jeszcze w wyniku ich wzajemnego wpływu oraz rezygnacji z początkowego radykalizmu, cechującego ich stanowiska wyjściowe.

Cechą charakterystyczną operacjonizmu jest jego ścisłe powiązanie z problematyką fizyki współczesnej. Operacjonizm stanowi próbę uporania się z tymi zagadnieniami i trudnościami, przed którymi postawiły uczonych współczesne odkrycia i teorie fizykalne. W pierwszym rzędzie — teoria względności i mechanika kwantowa. Teorie te podważyły podstawy naszej dotychczasowej wiedzy fizykalnej. Zakwestionowały nie tylko jej podstawowe prawa, lecz i jej najogólniejsze pojęcia. Pojęcia te okazały się nieprzydatne do opisu nowych faktów. Aparat pojęciowy fizyki klasycznej prowadził na ich gruncie do nieprzezwyciężalnych trudności. Powstała tym samym konieczność rewizji szeregu podstawowych pojęć fizykalnych: przestrzeni, czasu, masy czy energii. Przykładem takiej rewizji jest analiza pojęcia równoczesności, przeprowadzona przez Einsteina w szczegółowej teorii względności. Ta właśnie analiza stała się punktem wyjścia dla twórcy operacjonizmu. Upatrywał w niej Bridgman wyraz nowej, właściwej postawy wobec problemu znaczenia i definiowania pojęć fizykalnych. Takie ujęcie sensu pojęć fizykalnych i sposobu jego określania pozwala, zdaniem Bridgmana, nie tylko na uporanie się z tymi trudnościami, które wyłoniły się na gruncie najnowszych teorii fizykalnych, ale również na uniknięcie podobnego kryzysu w przyszłości.

Owej «operacyjnej» koncepcji znaczenia i definiowania pojęć fizykalnych nie ogranicza Bridgman do pojęć teorii względności czy mechaniki kwantowej. Rozciąga ją na wszelkie pojęcia naukowe. A więc nie tylko na pojęcia fizyki czy chemii, ale i na pojęcia innych nauk przyrodniczych, a nawet na pojęcia nauk filozoficznych i społecznych. Podobnie postępują i inni operacjoniści. Szczególnie wielu zwolenników znalazł operacjonizm wśród psychologów i socjologów. Problem zakresu stosowalności operacjonistycznego punktu widzenia jest jednak sprawą sporną. Główna różnica zdań dotyczy pojęć matematycznych. Bridgman rozszerza swój punkt widzenia i na te pojęcia, inni protestują przeciwko temu upatrując tu nieporozumienie. Wśród przedstawicieli operacjonizmu panuje jednak zgoda co do tego, że koncepcja ta znajduje zastosowanie wobec wszelkich pojęć empirycznych. Tak też ujęty pogląd operacjonistyczny będzie przedmiotem dalszych rozważań. Warto zaznaczyć, że pogląd ten ma w tej postaci charakter normatywny, a nie sprawozdawczy. Nie jest to opis tego, jak się faktycznie pojęcia empiryczne definiuje, lecz program mówiący, jak się je definiować powinno. Operacjoniści zdają sobie z tego na ogół sprawę, wskazując przykłady nieoperacjonistycznej postawy uczonych specjalistów.

2. Przed przystąpieniem do przedstawienia i krytycznej analizy poglądu operacjonistycznego chciałbym wyraźnie zdać sprawę z ograniczeń, jakim to zadanie uległo w obecnej pracy. Nie zajmuję się w niej zagadnieniem społecznej genezy i funkcji poglądów operacjonistycznych. Poglądy te, jak już wspomniałem, stanowią pewną odmianę metodologii pozytywistycznej. Wspólne też są ich źródła społeczne i ich rola. A te sprawy były już wielokrotnie w przypadku neopozytywizmu przedstawiane i dyskutowane. Pokrewne wydają się też źródła filozoficzne obu tych kierunków. Związany z fizyką operacjonizm stanowi kontynuację linii rozwojowej reprezentowanej przez Macha i Poincarégo, do której m. in. nawiązywał również neopozytywizm. Na powstanie operacjonizmu nie pozostał poza tym bez wpływu pragmatyzm, jeden z dominujących w Ameryce prądów filozoficznych. Koncepcja operacjonistyczna wykazuje wyraźne pokrewieństwo z pragmatyczną metodą analizy znaczeniowej, a zwłaszcza z tą jej wersją, która znalazła wyraz w rozprawie Peirce'a: *How to make our ideas clear* (1878).

Omawiając poglądy operacjonistyczne ograniczyć chcę się do tego, co stanowi ich zasadniczy trzon: do operacjonistycznej koncepcji definicji pojęć naukowych. Poglądy te zawierają bowiem, prócz owej koncepcji, rozważania dotyczące i innych zagadnień metodologicznych. Rozważania te, raczej marginesowe i luźnie z tamtą koncepcją związane, obracają się wokół takich problemów z zakresu metodologii nauk fizykalnych, jak istota teorii fizykalnej, wyjaśniania, modelu, rola matematyki w fizyce itp. Ale i sama operacjonistyczna koncepcja definicji nastręcza przy próbach jej przedstawienia pewne trudności. Przede wszystkim z powodu ogólnikowości, niejasności, a często i niekonsekwencji operacjonistycznych wywodów. Taki charakter mają zarówno rozważania samego Bridgmana, jak i — w większym jeszcze stopniu — jego zwolenników. Rozważania te dopuszczają wskutek tego szereg różnych, nierównoważnych wzajem interpretacji. Poza tym zachodzą wyraźne rozbieżności pomiędzy poglądami poszczególnych reprezentantów tego kierunku. W tej sytuacji trudno mówić o jednej operacjonistycznej koncepcji definicji. Mamy tu raczej do czynienia z długim ich szeregiem, wśród którego dokonać musimy dopiero pewnego wyboru. Omawiana w obecnej pracy koncepcja definicji pojęć naukowych jest — co za tym idzie — jedną z wielu możliwych interpretacji operacjonistycznego stanowiska. Jest to jednak, jak sądzę, interpretacja naczelna, odróżniająca operacjonistyczny punkt widzenia od poglądów pokrewnych.

II

1. Operacjonistyczna koncepcja definicji stanowić miała — jak widzieliśmy — uogólnienie tego punktu widzenia, który znalazł wyraz w einsteinowskiej definicji równoczesności. To, co w postępowaniu Einsteina było, zdaniem Bridgmana, istotne, polegało na powiązaniu dwóch zagadnień: znaczenia pojęcia fizykalnego i czynności prowadzących do zastosowania tego pojęcia w konkretnym przypadku. Einstein szukając odpowiedzi na pytanie, jaki sens ma pojęcie równoczesności dwóch zdarzeń, badał czynności, za pomocą których dochodzimy do stwierdzenia takiej równoczesności. Taki sposób określania sensu pojęć nazywa Bridgman definicją operacyjną. Zasadniczy postulat operacjonizmu głosi, że definicje pojęć naukowych winny być definicjami operacyjnymi. Postulat ten, dotyczący wszelkich pojęć empirycznych, został przez Bridgmana szerzej rozwinięty i poparty szeregiem przykładów tylko w zastosowaniu do pojęć pewnego typu: wielkości fizykalnych, czyli pojęć odnoszących się do mierzalnych fizycznych własności przedmiotów. Postulat operacjonistyczny ulega w tym ujęciu swoistej konkretyzacji. Zaznajomienie się z nim wydaje się przydatne dla zrozumienia koncepcji operacjonistycznej w jej szerszym ujęciu.

Na czym polega definicja operacyjna wielkości fizykalnej? Wielkość fizykalną definiujemy operacyjnie wyszczególniając czynności, za pomocą których tę wielkość mierzymy, czyli operacje pomiarowe. Najczęściej przez Bridgmana cytowanym przykładem definicji operacyjnej jest definicja długości. Jaki sens ma to pojęcie? Co to znaczy, iż dany przedmiot jest tak a tak długi? Wedle operacjonisty, odpowiedź na to pytanie może dać nam tylko analiza sposobu pomiaru długości owego przedmiotu. Przypuśćmy, iż przedmiotem tym jest ściana pewnego budynku. Jej długość mierzymy przykładając wzdłuż niej sztabę mierniczą. Zespół tych operacji pomiarowych stanowi, według operacjonisty, znaczenie owej długości. A zatem jej definicja operacyjna winna podawać dokładny opis tych operacji.

Taka koncepcja definicji pociąga za sobą pewne osobliwe konsekwencje. Nie każdą długość mierzymy za pomocą opisanych operacji. Operacje te są na ogół wykonalne wobec przedmiotów naszego codziennego otoczenia, takich jak budynki, meble, czy ulice. Istnieją jednak przedmioty, w stosunku do których operacje te są niewykonalne. Czyż możemy mierzyć w ten sposób odległość gwiazd, średnice atomów, długość przedmiotów poruszających się z prędkością bliską prędkości światła? A skoro to jest niemożliwe, to pojęcie długości zdefiniowane za pomocą wspomnianych operacji pozbawione jest w stosunku do tych przedmiotów jakiegokolwiek sensu. Pojęcie to ma zastosowanie tylko w tej dziedzinie, w której wykonalne są operacje wyszczególnione w jego definicji.

Pojęcie długości stosujemy jednakże i do przedmiotów spoza naszego codziennego otoczenia. Jaki więc sens ma to pojęcie w zastosowaniu do przedmiotów takich, jak gwiazdy czy atomy? Aby odpowiedzieć na to pytanie musimy zbadać, za pomocą jakich operacji mierzymy długość tych przedmiotów. Okaże się wtedy, że są to operacje zupełnie inne niż te, za pomocą których mierzymy przedmioty codziennego otoczenia. Zespół tych operacji stanowi sens analizowanego pojęcia. Definicja operacyjna długości tego rodzaju przedmiotów winna zatem podawać dokładny opis tych właśnie operacji pomiarowych. Tylko tak zdefiniowane pojęcie długości ma zastosowanie do rozważanych obiektów. Jest to jednak pojęcie różne od poprzedniego. Pojęcia zdefiniowane za pomocą różnych operacji pomiarowych są pojęciami znaczeniowo różnymi. Pojęcia takie symbolizujemy zazwyczaj za pomocą tego samego terminu. Postępowanie takie może mieć miejsce wtedy, gdy różne wyszczególnione w definicjach tych pojęć operacje pomiarowe dają w przypadkach, w których są wspólnie wykonalne, te same wyniki. I tak, w stosunku do wielu przedmiotów znajduje zastosowanie zarówno opisana wyżej metoda pomiaru długości za pomocą sztaby mierniczej, jak i różna do niej metoda pomiaru długości za pomocą teodolitu. Metody te, jako odwołujące się do różnych operacji pomiarowych, definiują różne pojęcia długości: «długość dotykową» i «długość optyczną». Pojęcia te symbolizujemy zazwyczaj za pomocą tego samego terminu, gdyż w tych przypadkach, w których obie te metody znajdują zastosowanie, dają one — w granicach błędu pomiaru — identyczne wyniki. To stwierdzenie ma jednak charakter pewnego empirycznego uogólnienia i jako takie narażone jest zawsze na obalenie. Przy zwiększeniu dokładności naszych pomiarów lub przy przejściu do nowych dziedzin doświadczenia otrzymać możemy w wyniku obu pomiarów różne rezultaty.

Na te osobliwe konsekwencje operacjonistycznej doktryny kładą sami operacjoniści duży nacisk i wyraźnie przeciwstawiają ją poglądowi tradycyjnemu. Według tego poglądu, znaczenie wielkości fizykalnych jest niezależne od sposobu ich mierzenia. Tę samą wielkość fizykalną możemy mierzyć na wiele różnych sposobów. W wyrażeniach

„długość ściany budynku" i „długość średnicy atomu" termin „długość" ma to samo znaczenie, mimo że długość tę mierzymy w obu przypadkach za pomocą tak różnych metod pomiaru. Operacjoniści przeciwstawiając się temu poglądowi podkreślają, iż znaczenie wielkości fizykalnych określa sposób ich pomiaru. Istnieje tyle różnych pojęć długości, ile jest różnych metod jej pomiaru. Bridgman wyróżnia w swej analizie tego pojęcia co najmniej siedem różnych sposobów jego pomiaru i — co za tym idzie — tyleż różnych znaczeń, w jakich używamy tego pojęcia w fizyce współczesnej. W pracach tego autora znajdujemy podobne analizy szeregu pojęć z różnych działów fizyki.

Operacjoniści przyznają, iż różne z ich punktu widzenia wielkości fizykalne grają często tę samą rolę w równaniach matematycznych, stanowiących podstawowe założenia lub twierdzenia teorii fizykalnych. Z tego też względu bywają symbolizowane przez te same terminy. Teoria fizykalna nie jest jednak częścią matematyki. Jest nauką doświadczalną, opisującą otaczający nas świat fizyczny. Jej pojęcia wyposażone muszą być zatem w sens fizykalny, sens który by pozwalał nam stosować je do przedmiotów fizycznych. Sens zaś taki posiadają tylko o tyle, o ile istnieje sposób ich pomiaru. On to decyduje o tym, czy i jakie znaczenie przysługuje danej wielkości fizykalnej. To stanowisko usunąć ma trudności, jakie powstały w wyniku najnowszych odkryć fizykalnych. Zakwestionowanie podstawowych praw fizyki klasycznej, dotyczących np. masy, staje się na gruncie tego stanowiska zrozumiałe. Skoro masę przedmiotów naszego codziennego otoczenia mierzymy w inny sposób niż masę przedmiotów poruszających się z prędkością bliską prędkości światła, mamy tu do czynienia nie z jedną, lecz z dwiema różnymi — jeśli chodzi o ich sens fizykalny — wielkościami, i nic dziwnego, że inne rządzą nimi prawidłowości. Nieuprawnione jest zatem, według operacjonistów, przenoszenie zależności stwierdzonych w pewnej dziedzinie doświadczenia na dziedziny nowe, w których zmienia się sposób pomiaru wchodzących w grę wielkości. Przestrzeganie tego zakazu winno, ich zdaniem, zapobiec na przyszłość konieczności takiej rewizji naszych poglądów, jaką pociągnęły za sobą teoria względności czy mechanika kwantowa.

2. Taka jest — w skrócie i uproszczeniu — treść postulatów operacjonistycznych, ograniczonych do definicji wielkości fizykalnych. Postulaty te jednak, według operacjonistów, obejmować mają — jak wiemy — wszelkie pojęcia empiryczne. Definicja każdego pojęcia empirycznego winna być definicją operacyjną. Co rozumie się tutaj przez definicję operacyjną? Definicja operacyjna wielkości fizykalnej polega na wyszczególnieniu czynności, zą pomocą których tę wielkość mierzymy. Nie każde jednak pojęcie ma charakter ilościowy. Nie każde więc pojęcie można zdefiniować przez podanie operacji pomiarowych. W jaki sposób uogólnić to postępowanie, tak aby obejmowało ono pojęcia odnoszące się zarówno do mierzalnych, jak i do niemierzalnych własności przedmiotów? W przypadku własności mierzalnych, ilościowych możemy zawsze stwierdzić, w jakim stopniu dana własność występuje w pewnym konkretnym przypadku. Na tym polega właśnie pomiar tej własności. W przypadku

własności niemierzalnych, jakościowych takie postępowanie jest niemożliwe. Tutaj możemy stwierdzić tylko to, czy w pewnym konkretnym przypadku dana własność występuje, czy nie. Nazwijmy czynności, za pomocą których stwierdzamy obecność lub nieobecność pewnej własności — operacjami sprawdzającymi. Definicja operacyjna własności niemierzalnej polegać ma zatem na wyszczególnieniu operacji sprawdzających. Zilustrujmy to na prostym przykładzie. Przypuśćmy, iż chcemy zbudować operacyjną definicję magnesu. Zgodnie więc z zaleceniami operacjonistów badamy czynności, za pomocą których stwierdzamy, czy ciało jakieś jest magnesem, czy nie. Czynności te polegać mogą np. na umieszczeniu w pobliżu danego ciała niewielkiego kawałka żelaza i na obserwowaniu, czy ciało to przyciąga ów kawałek. Definicja operacyjna magnesu zawierać winna dokładny opis tych czynności.

I w tym przypadku spotykamy się oczywiście z rozważanymi wyżej konsekwencjami koncepcji operacjonistycznej. Pojęcie zdefiniowane operacyjnie ma sens tylko w tej dziedzinie, w której wykonalne są operacje sprawdzające wyszczególnione w jego definicji. Pojęcia zdefiniowane za pomocą różnych operacji sprawdzających są pojęciami znaczeniowo różnymi. To, czy dane ciało posiada własności magnetyczne, sprawdzać możemy na wiele różnych sposobów. Nie tylko przez zbliżanie do niego żelaznych opiłków, lecz i przez zmianę jego położenia względem jakiegoś zamkniętego obwodu. Te dwie metody sprawdzania definiują jednak dwa różne pojęcia magnesu, które tylko symbolizowane są za pomocą tego samego terminu.

Przedstawiona interpretacja postulatów operacjonistycznych, obejmujących wszelkie pojęcia empiryczne, daleka jest oczywiście od precyzji i wymaga dalszych wyjaśnień. Wydaje się jednak, iż chwyta ona w przybliżeniu to, o co operacjonistom idzie. A w każdym razie to, o co idzie twórcy tego kierunku i głównym jego zwolennikom. Spotykamy bowiem i inne interpretacje operacjonistycznych postulatów, bardziej liberalne od przedstawionej. Interpretacje takie reprezentują przede wszystkim opowiadający się za operacjonizmem psychologowie i socjologowie. Do definicji operacyjnych zaliczają oni nie tylko definicje, które podają czynności służące do *sprawdzenia*, czy przedmiot definiowany występuje w pewnym konkretnym przypadku. Charakter operacyjny przypisują również definicjom, które wyszczególniają czynności służące do *wytworzenia* przedmiotu definiowanego. Definicje operacyjne szeregu terminów psychologicznych należą do tego właśnie rodzaju. Oczywiście i na terenie psychologii natrafiamy na definicje operacyjne w ścisłym tego słowa znaczeniu. Są nimi np. definicje dyspozycji psychicznych, odwołujące się do ich pomiaru za pomocą różnego rodzaju testów. Jako przykład ich służyć może znane hasło: „inteligencja jest to to, co mierzy test inteligencji". Oprócz nich spotykamy jednak szereg definicji, którym nie z tego względu przypisuje się charakter operacyjny. Należą tu definicje zjawisk psychicznych, odwołujące się do czynności, które mają na celu wywołanie tych zjawisk. Nie chodzi tu — jak w definicjach poprzednich — o opis operacji, za pomocą których sprawdzamy, czy dane zjawisko, wywołane przez jakieś inne fakty, zachodzi, lecz o opis operacji, za pomocą których to zjawisko wywołujemy, za pomocą których stwarzamy jego przyczynę, jego bodziec, a nawet najczęściej po prostu o opis samej owej przyczyny, samego owego bodźca. Taki charakter mają np. definicje wrażeń zmysłowych jako czegoś, co wywołane jest przez określone podniety fizyczne czy fizjologiczne. Definicje takie przytaczają niektórzy psychologowie jako przykład definicji operacyjnych. Jest rzeczą oczywistą, że mamy tu do czynienia z wyraźną modyfikacją tego pojęcia.

Jeszcze dalej idące modyfikacje napotykamy na terenie operacjonistycznej socjologii. Jej przedstawiciele — podobnie jak psychologowie — dopuszczają w charakterze operacji definiujących, czynności prowadzące do wytworzenia przedmiotu definiowanego. Sądzą przy tym, iż definicja operacyjna nie musi mieć charakteru czysto słownego. Zamiast opisu tych czynności może zawierać po prostu ich wykonanie. Definiuję operacyjnie słowo „skok", gdy wypowiadając je wykonuję skok i wytwarzam w ten sposób przedmiot definiowany. Co więcej, czynności definiujące mogą polegać nie tylko na wytworzeniu przedmiotu definiowanego, lecz i na jego wskazaniu. Definicją operacyjną słowa „kot" byłby więc zabieg polegający na wypowiedzeniu tego słowa i jednoczesnym wskazaniu kota. Tak pojmowane definicje operacyjne obejmowałyby zatem wszelkie tzw. definicje dejktyczne. I ta koncepcja wydawała się jednak niektórym operacjonistom zbyt wąska. Charakter operacyjny skłonni byli oni upatrywać nawet w definicji określającej wodę jako H_2O. Trudno doprawdy domyślić się, na jakiej podstawie. Przy takiej koncepcji definicji operacyjnych nie sposób byłoby chyba znaleźć definicję, której można by tego miana odmówić. Ale i poprzednie koncepcje rozciągają zakres definicji operacyjnych w sposób pozbawiający to pojęcie użyteczności. Tak szerokie pojmowanie definicji operacyjnych zaciera granicę pomiędzy nimi a pozostałymi rodzajami definicji, pozbawia operacjonistyczny punkt widzenia jego charakterystycznych, odróżniających go od innych stanowisk, właściwości. W dalszych rozważaniach pominiemy zatem te interpretacje. Tym bardziej, że są to najwyraźniej interpretacje uboczne, nie odróżniane nawet często od scharakteryzowanej poprzednio interpretacji naczelnej.

3. Tę naczelną interpretację doktryny operacjonistycznej przedstawiliśmy w sposób bardzo ogólnikowy i nieścisły, idąc w tym zresztą za praktyką samych operacjonistów. W żadnej z dostępnych mi prac operacjonistycznych nie znalazłem przykładu dosłownie sformułowanej definicji operacyjnej. Mówi się w nich, że takie a takie pojęcie jest równoznaczne z zespołem takich to a takich operacji, że definiujemy je przez wyszczególnienie tych operacji itp. Nigdzie jednak nie przytacza się takich definicji, ani tym bardziej ich ogólnego schematu. Próba zaś odpowiedzi na te pytania natrafia od razu na poważne trudności.

Definicje rozważanych przez nas pojęć, w rodzaju pojęcia magnesu, przybierają zwykle w nauce postać równoważności.

$$Qx \equiv ...x...,$$

gdzie Q jest terminem definiowanym, a na miejscu $...x...$ stoi funkcja zdaniowa o jednej zmiennej wolnej, stanowiąca człon definiujący. Taka definicja równoważnościowa posiada — jak każda definicja w ścisłym tego słowa znaczeniu — pewną charakterystyczną właściwość: ustala znaczenie terminu definiowanego dla każdego przypadku jego zastosowania, dla wszelkich przedmiotów, dla każdego x. Jest to zresztą niezbędne, jeśli definicja taka ma być definicją formalnie poprawną, a więc niesprzeczną i przekładalną. Ale to właśnie sprawia, iż tego rodzaju definicja nie może być definicją operacyjną. Skoro ustala ona znaczenie terminu definiowanego dla wszelkich przedmiotów, ustala je też dla przedmiotów, wobec których niewykonalne są wyszczególnione w niej operacje sprawdzające. A więc i dla przedmiotów, dla których nie możemy przyłożyć sztaby mierniczej lub przybliżyć żelaznych opiłków. A przecież w stosunku do tych przedmiotów zdefiniowane operacyjnie pojęcia długości czy magnesu pozbawione mają być jakiegokolwiek znaczenia. Definicja operacyjna nadawać ma terminowi definiowanemu znaczenie tylko zastosowaniu do tych przedmiotów, wobec których wykonalne są wyszczególnione w niej operacje sprawdzające. Nie może to być zatem żadna definicja w ścisłym tego słowa znaczeniu, żadna definicja zupełna. Definicja operacyjna musi mieć charakter definicji częściowej, ustalającej znaczenie terminu definiowanego w pewnym tylko stopniu.

Taki właśnie charakter mają wprowadzone przez Carnapa definicje częściowe lub warunkowe, zwane inaczej redukcjami. Definicje operacyjne uważać możemy za ich poszczególny przypadek. Przybierają one zatem postać następującą:

$$P_1x \supset (Qx \equiv P_2x)\,,$$

gdzie Q jest terminem definiowanym, P_1 — opisem czynności, które musimy wykonać, aby stwierdzić, czy jakiś przedmiot posiada własność Q, a P_2 — opisem pewnego rezultatu tych czynności. Schemat powyższy głosi:

Jeżeli wykonamy operację P_1, to x ma własność Q zawsze i tylko wtedy, gdy otrzymamy rezultat P_2.

Nasza operacyjna definicja magnesu brzmiałaby więc tak:

Jeżeli w pobliżu x umieścimy niewielki kawałek żelaza, to x jest magnesem zawsze i tylko wtedy, gdy x przyciąga ten kawałek.

Takie ujęcie definicji operacyjnych pociąga za sobą omawianą konsekwencję. Definicja częściowa ustala wszak znaczenie terminu Q tylko dla tych przedmiotów, które spełniają warunek P_1. Te przy tym spośród nich, które okazują się P_2, posiadają własność definiowaną, tym zaś, które nie są P_2, własność definiowana nie przysługuje. Znaczenie terminu Q dla przedmiotów nie spełniających warunku P_1 pozostaje nie ustalone. Jeśli więc opisane w P_1 operacje sprawdzające są w stosunku do pewnego przedmiotu niewykonalne, przedmiot ten nie spełnia tym samym warunku P_1, a definicja operacyjna terminu Q nie wyposaża go w stosunku do tego przedmiotu w żaden określony sens. Jeżeli umieszczenie w pobliżu pewnego przedmiotu niewielkiego kawałka żelaza jest z jakichś powodów niewykonalne, wypowiedź orzekająca o tym przedmiocie termin „magnes" wprowadzony za pomocą przytoczonej definicji operacyjnej nie jest w ogóle zdaniem, lecz nonsensem pozbawionym wartości logicznej.

Podobne konsekwencje otrzymujemy w przypadku każdej definicji operacyjnej traktowanej jako pewna definicja częściowa. Takie ujęcie definicji operacyjnych odpowiada więc, jak się zdaje, intencjom operacjonistów. Musimy jednak wyraźnie zdać sobie sprawę z charakteru tej interpretacji. Tak zinterpretowane definicje operacyjne nie są — podobnie jak wszelkie definicje częściowe — definicjami w ścisłym tego słowa znaczeniu. Nie spełniają one warunku przekładalności i nie pozwalają — co za tym idzie — na wyrugowanie terminu definiowanego. Ustalają jednak w pewnym stopniu jego znaczenie, umożliwiając rozstrzyganie szeregu twierdzeń, w których ten termin występuje.

Prócz zagadnienia formalnej struktury definicji operacyjnych bliższych wyjaśnień wymaga problem charakteru operacji definiujących. W naszym schemacie definicji operacyjnej termin P_1 symbolizował pewną wykonywaną przez nas czynność sprawdzającą, a P_2 — pewien stan rzeczy, który był tej czynności wynikiem. W przypadku gdy termin definiowany ma charakter ilościowy, czynności te sprowadzają się do manipulacji przyrządami pomiarowymi i odczytywania ich wskazań. Muszą to być — jak wiemy — czynności wykonalne. I to nie tylko w logicznym czy fizykalnym, lecz i w technicznym sensie tego terminu. Większość operacjonistów uważa bowiem, iż czynności, których nie jesteśmy w stanie wykonać wskutek pewnych trudności technicznych, nie nadają się na operacje definiujące, nawet jeśli nie stoją w sprzeczności z twierdzeniami logiki czy prawami przyrody.

Pewne rodzaje operacji sprawdzających stały się przedmiotem szczególnego zainteresowania i dyskusji w gronie operacjonistów. Idzie tu np. o czynności polegające na prostej obserwacji zjawisk. Czy czynność taka zasługuje na miano „operacji"? Czy nadaje się na to, aby figurować w definicji operacyjnej? Na ogół odpowiada się na te pytania twierdząco. Żywszą dyskusję wzbudziło również zagadnienie czynności umysłowych, logicznych i matematycznych, takich jak liczenie, mnożenie itp. Czy czynności te mogą wchodzić w skład operacji definiujących, czy też operacjami takimi mogą być tylko czynności fizyczne, takie jak przykładanie sztaby mierniczej? Bridgman dopuszcza i ten rodzaj czynności, zwanych przez niego *paper-and-pencil operations*. Inni jednak protestują przeciwko temu. Sprawa ta — jak mi się wydaje — wiąże się ściśle z zagadnieniem zakresu operacjonistycznych postulatów. Jeśli postulaty te obejmować mają również pojęcia logiczne i matematyczne, w skład operacji sprawdzających muszą wchodzić i owe czynności umysłowe. Jeśli natomiast — tak jak to czyni większość operacjonistów — ograniczymy te postulaty do pojęć empirycznych, wprowadzenie owych dość problematycznych «operacji» stanie się zbyteczne.

III

1. Rozważania dotychczasowe poświęcone były głównie próbom zdania sprawy z treści poglądów operacjonistycznych. Starałem się nieco dokładniej niż to czynią sami operacjoniści przedstawić zasadniczy trzon tych poglądów: operacjonistyczną koncepcję definicji pojęć empirycznych. Koncepcja ta ma charakter pewnego programu, żądającego definiowania wszelkich pojęć empirycznych na drodze operacyjnej. Wspominaliśmy pokrótce o motywach, które skłaniały twórcę operacjonizmu do wysuwania takiego programu. Czy jest to program słuszny, możliwy do realizacji? I czy

istotnie motywy, na które powołują się operacjoniści, uzasadniają go w sposób przekonujący? Próbom odpowiedzi na te pytania poświęcić chciałbym dalsze rozważania.

Już sama charakterystyka koncepcji operacjonistycznej zawiera momenty, które budzą szereg wątpliwości i zastrzeżeń. Mam na myśli przede wszystkim ów wielokrotnie przez operacjonistów podkreślany fakt definiowania przez różne operacje sprawdzające różnych znaczeniowo terminów. Jeśli dwie operacje, pozwalające nam na podstawie doświadczenia rozstrzygnąć, czy przedmiot jakiś podpada pod dane pojęcie czy nie podpada, różnią się czymkolwiek pomiędzy sobą, to definiują one różne pojęcia. Tak jest w przypadku wszelkich, najdrobniejszych nawet różnic pomiędzy operacjami sprawdzającymi. Widzieliśmy, iż operacja pomiaru długości pewnego przedmiotu polegająca na przykładaniu wzdłuż niego sztaby mierniczej określa inne pojęcie długości niż operacja polegająca na nastawianiu teodolitu. Ale tak jest nie tylko w przypadku różnic tak jaskrawych. Operacja polegająca na kolejnym przykładaniu sztaby mierniczej od punktu *A* do punktu *B* różni się od operacji polegającej na przykładaniu tej sztaby od punktu *B* do punktu *A*. I te zatem operacje dają początek różnym pojęciom długości. Gdybyśmy jednak w konsekwencji uwzględnić chcieli wszelkie różnice zachodzące pomiędzy operacjami pomiarowymi, postulaty operacjonistyczne stawałyby się niemożliwe do realizacji. Otrzymywalibyśmy niezliczoną ilość pojęć odpowiadających jednemu tradycyjnemu pojęciu długości, a ich definicje podawać by musiały niezliczoną ilość szczegółów: kierunek, w którym odkładamy sztabę mierniczą, jej drogę, prędkość i przyspieszenie przy przenoszeniu od jednej pozycji do drugiej itd. Analogiczne trudności napotykamy i w przypadku innych pojęć naukowych, i to zarówno ilościowych, jak i jakościowych. Toteż operacjoniści nie postępują w praktyce tak rygorystycznie i spośród różnic zachodzących pomiędzy operacjami uwzględniają tylko niektóre. Bridgman, poddając analizie operacyjnej pojęcie długości, wyróżnia parę tylko znaczeń tego terminu, odpowiadających paru typom operacji pomiarowych, które poza tym, w obrębie tego samego typu, mogą się znacznie nawet różnić pomiędzy sobą.

Stanowisko takie wydaje się jednak niekonsekwentne. Dlaczego mamy uwzględniać różnicę między operacją przykładania sztaby mierniczej a operacją nastawiania teodolitu, a lekceważyć różnicę między operacją przykładania sztaby w pewnym kierunku a operacją przykładania jej w kierunku przeciwnym? I w pierwszym, i w drugim przypadku oba rodzaje operacji dają w granicach naszego dotychczasowego doświadczenia wyniki te same. I w pierwszym, i w drugim przypadku przy przejściu do nowych dziedzin doświadczenia lub przy zwiększeniu dokładności naszych pomiarów wyniki te mogą okazać się różne. Jeżeli fakt ten skłania operacjonistów do wyróżnienia dwóch pojęć długości w przypadku pierwszym, winien prowadzić ich do takiego samego wniosku w przypadku drugim. Nie widać powodu, dla którego mogliby oni pewne różnice pomiędzy operacjami pomiarowymi czy sprawdzającymi traktować inaczej niż pozostałe. Jeśli tak czynią, postępują niekonsekwentnie.

Stanowisko konsekwentne a zarazem możliwe do realizacji można zachować tylko za cenę rezygnacji z tezy głoszącej, iż różne operacje sprawdzające definiują różne pojęcia. Można wówczas przyjąć, iż definicje podające różne metody pomiaru długości stanowią uzupełniające się wzajemnie definicje tego samego pojęcia. Stanowisko takie pociągałoby jednak za sobą dość osobliwą konsekwencję. Układ owych definicji nie byłby zwykłą umową terminologiczną, lecz pewnym twierdzeniem doświadczalnym. Przypuśćmy, iż wprowadzamy do naszego języka termin Q za pomocą dwóch podających różne metody sprawdzania definicji operacyjnych:

(1) $$P_1x \supset (Qx \equiv P_2x)\,,$$

(2) $$P_3x \supset (Qx \equiv P_4x)\,.$$

Z definicji tych wynika logicznie twierdzenie, które nie zawiera już terminu definiowanego i które ma w zasadzie charakter doświadczalny:

$$P_1x \cdot P_3x \supset (P_2x \equiv P_4x)\,.$$

Głosi ono, że w stosunku do przedmiotów dostępnych obu metodom sprawdzania metody te dają zawsze zgodne wyniki, a więc bądź oba pozytywne, bądź oba negatywne. Jest to pewne empiryczne uogólnienie i jako takie narażone jest zawsze na obalenie. Mimo że całe dotychczasowe doświadczenie może je potwierdzać, musimy być przygotowani na to, że natrafimy w przyszłości na przypadek obalający tę zależność, np. na przedmiot, który okaże się zarazem $P_1 \cdot P_2 \cdot P_3 \cdot {\sim}P_4$. Przypadek taki pociągałby za sobą konieczność odrzucenia którejś z przyjętych przez nas definicji operacyjnych terminu Q. Definicje te byłyby więc twierdzeniami empirycznymi, podległymi kontroli doświadczenia.

Zilustrujmy to na przykładzie omawianego już przez nas pojęcia magnesu. Przyjmijmy dwie definicje operacyjne tego terminu, odwołujące się do różnych sposobów stwierdzania własności magnetycznych ciał:

(1) Jeżeli w pobliżu x umieścimy niewielki kawałek żelaza, to x jest magnesem zawsze i tylko wtedy, gdy x przyciąga ten kawałek.

(2) Jeżeli zmieniamy położenie x względem zamkniętego obwodu, to x jest magnesem zawsze i tylko wtedy, gdy w obwodzie tym powstaje prąd elektryczny.

Definicje te zakładają prawdziwość następującego twierdzenia doświadczalnego:

Jeżeli w pobliżu x umieścimy niewielki kawałek żelaza, a zarazem x poruszamy względem zamkniętego obwodu, to x przyciąga ów kawałek zawsze i tylko wtedy, gdy w obwodzie tym powstaje prąd elektryczny.

Dotychczasowe dane potwierdzają taką zależność. Nie jest jednak wyłączone, że znajdziemy w przyszłości takie ciało, które poddane obu procedurom przyciągać będzie opiłki żelaza, lecz nie wzbudzi prądu elektrycznego. W takim przypadku zmuszeni bylibyśmy zmodyfikować nasz układ definicji. Tym samym zmienilibyśmy znaczenie definiowanego terminu. A zatem przyjęcie możliwości definiowania przez różne operacje sprawdzające tego samego pojęcia pociąga za sobą konieczność rewizji definicji operacyjnych, a co za tym idzie i znaczenia pojęć naukowych w miarę zapoznawania się z nowymi, nieprzewidzianymi faktami.

Można by uniknąć tej konsekwencji przez inne nieco formułowanie dodatkowych definicji operacyjnych tego samego pojęcia. Załóżmy, że podobnie jak poprzednio wprowadziliśmy termin Q za pomocą definicji:

(1) $$P_1x \supset (Qx \equiv P_2x),$$

Definicja ta, jak wiemy, ustala znaczenie terminu Q dla tych wszystkich przedmiotów, które spełniają warunek P_1. Następną definicję operacyjną odwołującą się do innej metody sprawdzania można by sformułować tak, aby nadawała ona terminowi Q znaczenie tylko w tej dziedzinie, w której do tej pory znaczenia był pozbawiony, a zatem tylko w stosunku do tych przedmiotów, które nie spełniają warunku P_1:

(2) $$\sim P_1x \supset (P_3x \supset Qx \equiv P_4x).$$

Z układu tak sformułowanych definicji operacyjnych terminu Q nie wynika już żadne twierdzenie doświadczalne. Układ takich definicji ma więc — podobnie jak wszelka definicja *sensu stricto* — charakter zwykłej umowy terminologicznej. Ten rezultat okupuje się pewnym skomplikowaniem formuł definicyjnych. Nasza uzupełniająca definicja magnesu przyjmuje postać następującą:

(2a) Jeżeli w pobliżu x nie umieścimy żadnego kawałka żelaza, a x poruszamy względem zamkniętego obwodu, to x jest magnesem zawsze i tylko wtedy, gdy w obwodzie tym powstaje prąd elektryczny.

Na tej drodze uniknąć więc można jednej z trudności operacjonistycznego stanowiska. Należy jednak podkreślić, iż zdecydowana większość operacjonistów stanowczo by tego rodzaju rozwiązanie odrzuciła, nie godzą się na rezygnację z tezy, do której przywiązują tak wielką wagę: tezy o definiowaniu przez różne zespoły operacji różnych znaczeniowo terminów.

2. Trudność powyższa nie jest jedyną wadą operacjonistycznej koncepcji. Zwróćmy uwagę na inną osobliwą konsekwencję tego stanowiska. Wszelkie pojęcie naukowe definiujemy tu za pomocą terminów odnoszących się do naszych czynności i wytworów tych czynności. Każde więc twierdzenie, w skład którego pojęcie takie wchodzi, mówi coś pośrednio o pewnych swoistych zachowaniach się ludzi i wytworach tych zachowań. Ściśle biorąc, co prawda, definicje operacyjne będąc definicjami częściowymi nie pozwalają na wyrugowanie terminu definiowanego i zastąpienie go terminami, za pomocą których został wprowadzony. W pewnym stopniu jednakże ustalają jego znaczenie i sprowadzają je tym samym do opisu ludzkich czynności. Można by zatem powiedzieć, iż fizyka posługująca się terminami zdefiniowanymi operacyjnie jest w gruncie rzeczy nauką o pewnych zachowaniach się ludzi. Jej prawa mówiące z pozoru o atomach czy falach elektromagnetycznych dotyczą w istocie zależności między naszymi czynnościami pomiarowymi a ich rezultatami. Jest to konsekwencja podważająca słuszność operacjonistycznego programu.

Dałoby się może tej konsekwencji uniknąć interpretując — zgodnie zresztą z sugestiami pewnych operacjonistycznych sformułowań — nieco inaczej terminy występujące w definicji operacyjnej. Nie jako opis ludzkich czynności czy ich wytworów, lecz jako opis pewnych stanów rzeczy nie będących w zasadzie rezultatem czyichś zabiegów. A w każdym razie o tych zabiegach nie mówiłoby się *explicite* w samym sformułowaniu definicji operacyjnych. Przytaczana przez nas wielokrotnie definicja operacyjna magnesu brzmiałaby w tym sformułowaniu mniej więcej tak:

Jeżeli w pobliżu *x* znajduje się niewielki kawałek żelaza, to *x* jest magnesem zawsze i tylko wtedy, gdy *x* przyciąga ten kawałek.

Mam wątpliwości czy interpretacja taka dałaby się przeprowadzić całkowicie konsekwentnie. Nie odpowiada ona poza tym stanowczo głównej intencji operacjonistów. Intencja ta polega na żądaniu, aby pojęcia naukowe definiować właśnie za pomocą naszych *operacji*. To zresztą stanowi charakterystyczną cechę operacjonizmu, odróżniającą ten kierunek od innych pokrewnych stanowisk, a zwłaszcza od tak bliskiego operacjonizmowi neopozytywizmu. W tym jednak punkcie racja nie leży — jak mi się wydaje — po stronie operacjonizmu. Nie tylko z tych powodów, które starałem się pokrótce przedstawić. Ów charakterystyczny dla operacjonizmu postulat wydaje się żądaniem nieuzasadnionym. Cele, jakim program operacjonistyczny ma służyć, postulatu tego bynajmniej nie narzucają.

Naczelnym celem operacjonizmu jest zapewnienie empirycznego sensu pojęciom naukowym. Operacjonizm zajmuje tutaj wspólny front z wszelkimi odmianami pozytywizmu w walce przeciwko metafizycznym tendencjom. Jeśli nauka ma opisywać otaczający nas świat materialny, jej pojęcia muszą posiadać sens empiryczny, pozwalający nam na stosowanie ich do przedmiotów danych nam w doświadczeniu. Pojęcie naukowe musi mieć takie znaczenie, aby można było o przedmiocie danym nam w doświadczeniu rozstrzygnąć, czy pod pojęcie to podpada, czy nie. Tylko wtedy twierdzenia naukowe podlegać mogą kontroli doświadczenia. Tylko wtedy doświadczenie decydować może o ich prawdziwości czy fałszywości. Ten postulat empiryczności pojęć naukowych operacjonizm usiłuje zrealizować przez wprowadzanie tych pojęć do języka nauki na drodze definicji operacyjnych. Próby tej, jak widzieliśmy, nie można uznać za udaną. Ale wbrew temu, co sugerują operacjoniści, nie jest to jedyna droga do realizacji postulatu empiryczności. Postulat ten bynajmniej nie pociąga za sobą operacjonistycznej koncepcji definicji. Aby pojęcie miało charakter empiryczny, musi istnieć sposób sprawdzania, czy dany przedmiot podpada pod to pojęcie, czy nie podpada. Nie wynika z tego jednakże, że pojęcie takie definiować musimy opisując ów właśnie sposób, wyszczególniając składające się nań czynności sprawdzające. A tego właśnie żądają operacjoniści.

Trafniejszą próbę realizacji postulatu empiryczności pojęć naukowych reprezentuje, jak mi się wydaje, neopozytywizm. Rozwiązanie neopozytywistyczne nie narażone jest w każdym razie na te zarzuty, jakim podlega koncepcja operacjonistyczna. To stwierdzenie wymaga pewnych zastrzeżeń. Neopozytywizm nie jest bynajmniej jakimś poglądem jednolitym. U różnych autorów przybierał postać różną. Poza tym jako całość ulegał wyraźnej ewolucji. Koncepcja, którą mam na myśli, opracowana została w późniejszych pracach Carnapa, w pierwszym rzędzie w rozprawie *Testability and Meaning*. W porównaniu z koncepcją operacjonistyczną — nawet z jej zmodyfikowaną już przez nas wersją — jest to koncepcja bardziej liberalna i lepiej odpowiadająca faktycznej praktyce naukowej. Przedstawienie jej wykracza poza zadania obecnej pracy. Chciałbym tylko w paru słowach scharakteryzować jej stosunek do poglądu operacjonistycznego.

Cel obu tych tendencji jest wspólny: zagwarantowanie empirycznego sensu terminom naukowym. Wedle neopozytywizmu jednak nie trzeba dla osiągnięcia tego celu definiować tych terminów, opisując, w jaki sposób stosujemy je do przedmiotów danych nam w doświadczeniu. Wystarczy jeśli definiujemy je tak, aby można je było do takich przedmiotów stosować. A tak jest zawsze wtedy, gdy definiujemy je za pomocą terminów spostrzeżeniowych, terminów odnoszących się do obserwowalnych własności i stosunków. Terminy odnoszące się do naszych czynności sprawdzających czy pomiarowych są tylko ich poszczególnym przypadkiem i nie widać powodu, dla którego trzeba by terminy definiujące do tych tylko terminów spostrzeżeniowych ograniczać. To rozszerzenie postulatu operacjonistycznego nie jest jedyną jego modyfikacją. Neopozytywiści podkreślają, iż sens empiryczny posiada nie tylko termin zdefiniowany bezpośrednio za pomocą terminów spostrzeżeniowych, ale i termin określony przez nie pośrednio. Jeśli terminy Q_1 i Q_2 zdefiniowałem za pomocą terminów spostrzeżeniowych, a termin Q_3 — za pomocą terminów Q_1 i Q_2, to zapewniłem tym samym sens empiryczny również temu ostatniemu. Stanowisko neopozytywizmu sprowadza się zatem do żądania, aby wszelkie terminy nauk empirycznych były w ostatecznej instancji definiowane — i to zarówno na drodze definicji zupełnych jak i częściowych — za pomocą terminów spostrzeżeniowych. Gwarantuje to istotnie empiryczny charakter pojęć naukowych, a usuwa te trudności, na które zwracaliśmy uwagę przy analizie koncepcji operacjonistycznej. Warto może jednak zaznaczyć, że i to stanowisko w stosunku do pewnych pojęć fizykalnych, zwłaszcza w fizyce teoretycznej, nasuwa pewne trudności i wymaga dalszych korektur.

3. Jako naczelny cel programu operacjonistycznego traktowaliśmy w rozważaniach dotychczasowych zapewnienie pojęciom naukowym sensu empirycznego. Tak też jest istotnie w przypadku nauk przyrodniczych. Inaczej nieco jednak przedstawia się ta sprawa na terenie nauk humanistycznych czy społecznych. I tutaj oczywiście operacjonizm ma zapewnić empiryczność pojęć, wyrugować terminy metafizyczne. Ale nie na to zadanie kładzie się tutaj główny nacisk. Jako naczelne zadanie wysuwa się w tej dziedzinie uściślenie aparatu pojęciowego. Definicje operacyjne mają, według operacjonistów, zapewnić ścisłość (lub jak się nieraz mówi *reliability*) terminom definiowanym. Przede wszystkim chodzi tu o zastąpienie nieprecyzyjnych terminów języka potocznego, którymi posługują się nauki psychologiczne czy społeczne, ostrymi terminami zdefiniowanymi operacyjnie. Znaczenie terminów potocznych nie wyposaża ich na ogół w wyraźne kryteria stosowalności. Stąd trudności i rozbieżności w konkretnych przypadkach ich zastosowań. Definicje operacyjne, formułujące kryteria stosowalności terminu definiowanego, mają usunąć te wadliwości. Wydaje się, że operacyjny sposób definiowania terminów naukowych istotnie to zadanie spełnia. W znacznym stopniu

zmniejsza nieostrość pojęć naukowych, co w naukach takich jak psychologia czy socjologia jest zadaniem niezmiernie ważnym. Widzieliśmy jednak, jakie trudności napotyka realizacja tego zadania na drodze proponowanej przez operacjonizm. Omawiane przez nas konsekwencje nie wyczerpują przy tym ogółu trudności, jakie pociąga za sobą ta doktryna. Na jedną z nich chciałbym jeszcze na zakończenie zwrócić uwagę.

Postulaty operacjonistyczne pozostają — jak mi się wydaje — w konflikcie z innymi żądaniami wysuwanymi pod adresem pojęć naukowych. Od pojęć tych żądamy, aby nadawały się do budowania teorii naukowych, aby pozwalały na ustalanie ogólnych, uzasadnionych i prostych praw naukowych. Nie każde pojęcie w równym stopniu do celu takiego się nadaje. Zależy to również od tego, za pomocą jakich cech zostało zdefiniowane. Za najlepiej do tego celu służące uważa się tzw. cechy istotne. Idzie tu o takie cechy danego przedmiotu, z których na mocy praw przyrodzonych wynikają inne, ważne z pewnego punktu widzenia, jego własności. Znajomość cech istotnych przedmiotu pozwala nam więc bez trudu przewidzieć i wyjaśnić rządzące nim prawidłowości. Stąd też żądanie definiowania terminów za pomocą cech istotnych wydaje się słuszne i zgodne z faktyczną praktyką badawczą. Postulat operacjonistyczny pozostaje najczęściej z tym żądaniem w konflikcie. Sposób sprawdzania, czy dany przedmiot podpada pod termin definiowany, nie stanowi na ogół cechy istotnej tego przedmiotu. Prawa, które nim rządzą, dotyczą z reguły innych jego własności. Nic też dziwnego, że rozwój pojęć naukowych przemawia wyraźnie przeciwko koncepcji operacjonistycznej. Kolejne definicje danego terminu odwołują się do coraz to «istotniejszych» własności, nie wiążąc jego sensu z ściśle określoną metodą sprawdzania czy pomiaru.

Zbierając wyniki naszych rozważań musimy zatem stwierdzić, iż zadanie zapewnienia terminom naukowym empirycznego, ścisłego, a zarazem przydatnego pod względem teoretycznym znaczenia realizować musimy na innej drodze niż ta, którą zarysowuje operacjonizm. Operacjonistyczna wersja pozytywistycznej metodologii budzi zbyt wiele wątpliwości i sprzeciwów.

UWAGI BIBLIOGRAFICZNE

Stanowisko operacjonistyczne, zwłaszcza w zastosowaniu do pojęć fizykalnych, przedstawione zostało w sposób najbardziej wyczerpujący w cytowanej książce P. W. Bridgmana *The Logic of Modern Physics* (1927). Poglądy tam wyłożone referował i rozwijał Bridgman w szeregu artykułów: „Operational Analysis" (*Philosophy of Science*, V, 1938), „Some Implications of Recent Points of View in Physics" (*Revue Internationale de Philosophie*, III, 1949), „The Operational Aspect of Meaning" (*Synthese*, VIII, 1950–51), „The Nature of Some of Our Physical Concepts" (*British Journal for the Philosophy of Science*, I–II, 1951), oraz w książce: *The Nature of Physical Theory* (1936). Analizie koncepcji operacjonistycznej dotyczącej pojęć fizykalnych poświęcone są m. in. prace następujących autorów: R. Lindsay „A Critique of Operationalism in Physics" (*Philosophy of Science*, IV, 1937), Ph. Frank *Foundations of Physics*

(1946), Bernstein „P.W. Bridgman, In Revolt against Formalism" (*Synthese*, VIII, 1950–51), Hesse „Operational Definition and Analogy in Physical Theories" (*British Journal for the Philosophy of Science*, II, 1951). Operacjonizm na terenie psychologii omawia m. in. artykuł H. Israela i B. Goldsteina „Operationism in Psychology" (*Psychological Review*, LI, 1944). Dyskusji nad operacjonizmem poświęcono ponadto cały numer *Psychological Review* (LII, 5, 1945). Zagadnienie definicji operacyjnych w socjologii poruszają m. in. prace następujące: G. Lundberg „Social Research" oraz „Operational Definitions in the Social Sciences" (*American Journal of Sociology*, XLVII, 1942), S. Dodd „Operational Definitions Operationally Defined" (*American Journal of Sociology*, XLVIII, 1943), Adler „Operational Definitions in Sociology" (*American Journal of Sociology*, LII, 1947). W literaturze polskiej analizie operacjonizmu poświęcone są artykuły: E. Poznańskiego „Analiza operacyjna pojęć fizyki" (*Przegląd Filozoficzny*, XXXV, 1932) oraz M. Przełęckiego „O tzw. definicjach operacyjnych" (*Studia Logica*, III, 1955). Ten ostatni wykorzystuje w znacznym stopniu praca obecna. Oprócz tego istnieje w naszej literaturze filozoficznej próba zastosowania operacjonistycznego punktu widzenia do zagadnienia prawdziwości, zawarta w pracy E. Poznańskiego i A. Wundeheilera *Pojęcie prawdy na terenie fizyki* (*Fragmenty Filozoficzne* I, 1934). Próbę tę poddał analizie krytycznej A. Schaff w książce *Z zagadnień marksistowskiej teorii prawdy* (1951).

Filozofia Nauki
Rok I, 1993, Nr 2-3

O definiowaniu terminów spostrzeżeniowych

1. Problem, któremu w głównej mierze poświęcone będą te rozważania, to problem tzw. interpretacji bezpośredniej[1]. Procedurę nadawania terminom deskryptywnym określonych denotacji — jak również wynik takiej procedury — nazywa się często ich interpretacją. Wyróżnić możemy dwa podstawowe rodzaje interpretacji terminów deskryptywnych: interpretację pośrednią i bezpośrednią. Interpretując dany termin deskryptywny pośrednio, korzystamy z innych terminów deskryptywnych zinterpretowanych już poprzednio; przyporządkowujemy takiemu terminowi pewien przedmiot jako denotację dzięki temu, iż inne terminy już denotację posiadają. Jest rzeczą oczywistą, że interpretacja pośrednia nie może stanowić jedynego sposobu nadawania terminom deskryptywnym określonych denotacji. Musi istnieć prócz niej sposób taki, który nie odwołuje się do innych terminów deskryptywnych o już ustalonych denotacjach. Sposób taki nazywamy interpretacją bezpośrednią. Interpretację tę utożsamia się często z tzw. definicją dejktyczną lub ostensywną, a terminy zinterpretowane bezpośrednio — z terminami spostrzeżeniowymi.

Problem interpretacji terminów spostrzeżeniowych bywał już przedmiotem analizy[2]. W dalszym ciągu jednak pozostaje tu wiele niejasności i problemów spornych. Toteż warto, jak mi się wydaje, przyjrzeć się temu zagadnieniu nieco bliżej, chociażby po to, by wyraźniej wskazać trudności związane z różnymi nasuwającymi się możliwościami rozwiązań. Przedmiotem naszych rozważań będą zresztą nie tylko problemy interpretacji bezpośredniej. Biorąc pod uwagę te terminy, które tradycyjnie uważa się za terminy spostrzeżeniowe, rozpatrzymy sposoby ich interpretacji niezależnie od tego, czy sposoby te zasługują istotnie na miano interpretacji bezpośredniej. Do owych

1) Por. M. Przełęcki, „Interpretacja systemów aksjomatycznych", *Studia Filozoficzne*, 1960, nr 6 (21)

2) Najpoważniejszą ze znanych mi prób takiej analizy stanowi praca J. Kotarbińskiej, „Tak zwana definicja dejktyczna", *Fragmenty Filozoficzne — Seria Druga*, Warszawa 1959.

terminów spostrzeżeniowych zalicza się zwykle terminy dwóch rodzajów: nazwy indywiduowe oraz jedno i wieloargumentowe predykaty pierwszego rzędu. Pierwsze z nich pojmowane bywają jako terminy denotujące konkretne rzeczy materialne, drugie — jako terminy denotujące klasy i stosunki zachodzące pomiędzy takimi rzeczami. Tak bywa w każdym razie na terenie tzw. języków fizykalistycznych, do których ograniczymy tutaj nasze rozważania. Rozważania te ograniczymy zarazem do języków ekstensjonalnych, utożsamiając tym samym — o ile wyraźnie nie stwierdzimy czegoś przeciwnego — własności rzeczy z odpowiadającymi im klasami. Podział terminów na nazwy i predykaty jest ze względu na możliwość ich bezpośredniej interpretacji podziałem niezmiernie doniosłym. Inaczej wygląda ta procedura w każdym z tych przypadków i w związku z tym wymaga oddzielnego potraktowania.

Pośrednia interpretacja terminu *T* przybrać może postać zwykłej, werbalnej definicji tego terminu. Przedmiot, który stanowić ma denotację terminu *T*, zostaje w nich scharakteryzowany za pomocą słownego opisu. Bezpośrednia interpretacja terminu *T*, czyli, przy pewnym rozumieniu, definicja ostensywna terminu *T*, przyporządkowuje temu terminowi przedmiot, który ma stanowić jego denotację, nie przez słowny opis tego przedmiotu, lecz — jak się na ogół twierdzi — po prostu przez jego wskazanie. Aby procedura taka mogła być skuteczna, ów przedmiot wskazywany musi być przedmiotem spostrzegalnym. Otóż wydaje się, że przedmiotami spostrzegalnymi mogą być, w dosłownym znaczeniu, tylko przedmioty konkretne. Bowiem tylko takie przedmioty mogą być bodźcami dla naszych narządów zmysłowych, a to wydaje się niezbędnym warunkiem dosłownie rozumianej spostrzegalności. O «spostrzegalności» przedmiotów abstrakcyjnych: klas (własności) czy stosunków nie można mówić w tym samym, literalnym sensie. Czy i jaki sens ma mówienie o «spostrzegalności» przedmiotów takich, jak czerwień lub wyższość — do tego pytania powrócimy w dalszym toku rozważań. Tutaj poprzestaniemy na stwierdzeniu, iż — dosłownie biorąc — wskazywać i spostrzegać możemy tylko konkretne indywidua. Stąd różnica pomiędzy definicją ostensywną nazwy indywiduowej a definicją ostensywną predykatu. Definicja ostensywna nazwy indywiduowej jest zabiegiem nastręczającym bez porównania mniej wątpliwości i problemów. Przyjrzyjmy się jej obecnie nieco bliżej.

2. Niech terminem definiowanym będzie nazwa indywiduowa *a*. Jej definicję ostensywną przedstawia się zwykle jako wypowiedź:

(1) $$\text{to}=a,$$

gdzie „to" jest wyrażeniem okazjonalnym, które w połączeniu z odpowiednim wskazującym gestem stanowi nazwę indywiduową określonego przedmiotu. Ten właśnie przedmiot wypowiedź (1) przyporządkowuje nazwie *a* jako jej denotację. W tej sytuacji wypowiedź ta istotnie nadaje interpretację symbolowi *a*. Czy można jednak utrzymywać, iż jest to interpretacja bezpośrednia? Wszak w wypowiedzi (1) oprócz nazwy *a* i stałych logicznych występuje termin deskryptywny „to". Przedmiot przyporządkowany nazwie *a* wyróżniony zostaje nie przez proste wskazanie, lecz przez wskazanie połączone z użyciem innego terminu deskryptywnego. A więc pośrednio raczej niż bezpośrednio.

Stawia się niekiedy sprawę tak, jak gdyby użycie owej nazwy okazjonalnej „to" było tutaj czymś nieistotnym. Można nazwie *a* przyporządkować jako denotację pewien przedmiot po prostu wskazując go (lub w jakiś inny, czysto fizyczny sposób wyodrębniając go z otoczenia) bez posługiwania się jakimkolwiek terminem różnym od nazwy *a*. Zamiast wypowiedzi (1) mielibyśmy zatem coś w rodzaju:

(2) $\underset{\downarrow}{a},$

gdzie ↓ symbolizuje ów wskazujący gest. Czyż w ten sposób nie można symbolowi *a* nadać zamierzonej interpretacji?

Odpowiedź na to pytanie zależy od tego, jak się skuteczność takiego zabiegu pojmuje. Można ją bowiem pojmować szerzej lub węziej. Załóżmy, iż termin *T* chcemy zinterpretować tak, aby denotował przedmiot *P*. Ze względu na ten cel daną procedurę interpretacyjną można uważać za skuteczną, jeśli u osoby, do której jest skierowana, wywołuje skutek polegający na rozumieniu *T* jako nazwy przedmiotu *P*. Skuteczność takiej procedury można jednak rozumieć węziej. Zgodnie z tym rozumieniem dana procedura interpretacyjna jest skuteczna dopiero wtedy, gdy dostarcza uzasadnienia dla stwierdzenia, iż termin *T* jest nazwą przedmiotu *P*. Tak jest na przykład wtedy, gdy dana procedura polega na podaniu odpowiedniej definicji terminu *T*. W przypadku pierwszym wchodzą w grę zależności faktyczne o charakterze przyczynowym (na przykład pewne prawidłowości psychologiczne). Ze względu na nie dana procedura prowadzi do rozumienia terminu *T* jako terminu denotującego przedmiot *P*. W przypadku drugim wchodzą w grę zależności logiczne. Ze względu na nie dana procedura dostarcza racji logicznej dla twierdzenia, iż *T* denotuje przedmiot *P*. Do rozróżnienia tego powrócimy jeszcze w związku z definicją ostensywną predyktatów, gdzie odgrywa ona doniosłą rolę. Obecnie chciałbym się do niego odwołać dla zdania sprawy z różnicy pomiędzy procedurami (1) i (2).

Tylko pierwsza z nich wyznacza denotację terminu *a* w sensie ściślejszym. Wypowiedź (1) stanowi istotnie rację logiczną dla stwierdzenia, iż termin *a* denotuje przedmiot wskazywany. Procedura druga denotacji terminu *a* w tym sensie nie wyznacza. Chyba że zwrot (2) traktowany jest jako wypowiedź skrótowa, która wyrażona w sposób wyczerpujący winna brzmieć tak samo, jak wypowiedź (1). Jeśli jednak nie traktujemy go jako wypowiedzi zawierającej w sposób domyślny termin okazjonalny „to", zwrot ten nie stanowi racji logicznej dla stwierdzenia, iż termin *a* denotuje przedmiot wskazywany. Chociażby dlatego, że nie ma w ogóle charakteru zdaniowego. Wypowiedzenie terminu *a* wraz ze wskazaniem pewnego przedmiotu może nasunąć przypuszczenie, iż termin *a* ten właśnie przedmiot denotuje; nie ma jednak sprzeczności w przypuszczeniu, iż tak nie jest, gdyż sam fakt wypowiedzenia jakiegoś słowa i wykonania jednocześnie pewnego wskazującego gestu nie musi być wyrazem decyzji przyporządkowania temu słowu jako denotacji przedmiotu wskazanego, lecz może wyrażać jakąś inną intencję. Można się co najwyżej zgodzić, iż zabieg taki «psychologicznie» pociąga za sobą przekonanie, iż przedmiot wskazany stanowi denotację terminu *a*, gdyż w takiej sytuacji bywa z reguły stosowany, jak się tego skądinąd możemy domyślać. W tym sensie jest to więc zabieg skuteczny. Do tego, aby mógł być uznany za zabieg skuteczny w sensie ściślejszym, musi zakładać wypowiedź (1) posługującą się terminem okazjonalnym „to".

Czy mamy wobec tego przyznać, iż jest to pośredni sposób interpretacji nazwy indywiduowej? Myślę, że na dwa sposoby można zdać sprawę z tej sytuacji. Traktując sprawę rygorystycznie, definicję ostensywną postaci (1) uważać musimy za pośredni sposób interpretacji terminu *a*; sposób, który przyporządkowuje mu jako denotację przedmiot stanowiący denotację innego terminu deskryptywnego, mianowicie terminu „to". Ów gest wskazujący potrzebny jest do nadania interpretacji temu ostatniemu. A zatem tylko terminy okazjonalne w rodzaju terminu „to" mogą być interpretowane bezpośrednio. Sposób ich interpretacji polega po prostu na użyciu takiego terminu w określonej sytuacji, na przykład łącznie z pewnym wskazującym gestem. Takie użycie terminu „to" przyporządkowuje mu istotnie jako denotację ów wskazywany przedmiot. Pod tym względem terminy okazjonalne różnią się więc zasadniczo od wszelkich nazw pozostałych. Traktując sprawę mniej rygorystycznie, możemy jednak — jak mi się zdaje — uważać definicję ostensywną postaci (1) za bezpośredni sposób interpretacji terminu *a*. Zwróćmy bowiem uwagę na fakt, iż termin „to" pozbawiony jest przed wygłoszeniem tej definicji jakiejkolwiek interpretacji. Dopiero za pomocą definicji ostensywnej terminu *a*, w której skład wchodzi odpowiedni gest wskazujący, nadajemy interpretację terminowi „to" — a tym samym i terminowi *a*. Można zatem utrzymywać, iż nie posługujemy się tutaj dla interpretacji tego ostatniego żadnym terminem deskryptywnym wyposażonym z góry, przed wygłoszeniem definicji, w jakąś interpretację. W tej sytuacji dopuszczalne wydaje się stanowisko upatrujące w definicji ostensywnej nazwy indywiduowej *a* postaci (1) bezpośredni sposób jej interpretacji.

Stwierdziliśmy, iż definicja taka istotnie wyznacza denotację terminu definiowanego. Czy w sposób jednoznaczny? Czy nazwie *a* przyporządkowany zostaje tą drogą dokładnie jeden przedmiot? Wydaje się, iż na pytanie to odpowiedzieć trzeba przecząco. Charakterystyka jakiegoś przedmiotu za pomocą terminu okazjonalnego „to" i wskazującego gestu nie jest nigdy charakterystyką jednoznaczną. Nie jesteśmy w stanie wyznaczyć w ten sposób — jak również w żaden inny sposób — dokładnych granic czasowo–przestrzennych wskazywanego obiektu. A zatem charakterystykę taką spełnia szereg obiektów, które co prawda zachodzą na siebie, lecz różnią się nieco pod względem granic w czasie i przestrzeni. Każdy z tych obiektów uważany być może na mocy definicji ostensywnej nazwy *a* za jej denotację. Denotacja ta nie jest więc wyznaczona w sposób jednoznaczny. Sytuację tę pogarsza fakt, iż ogromna większość nazw indywiduowych odnosi się nie do przedmiotów «chwilowych», lecz do przedmiotów względnie trwałych, takich jak ten oto stół (od momentu produkcji aż do momentu zniszczenia) lub określony człowiek (od chwili narodzin do chwili śmierci). Przedmioty te trwają znacznie dłużej niż spostrzeżenia, do których się przy ostensywnym definiowaniu ich nazw odwołujemy. Nie możemy przedmiotu takiego spostrzec w całej jego czasowej rozciągłości. Możemy spostrzegać tylko pewne jego części, pewne czasowe «fazy». A przecież jako denotację definiowanej ostensywnie nazwy przyporządkowujemy jej nie poszczególną, spostrzeganą właśnie «fazę», lecz całość obejmującą również nie spostrzegane w danej chwili «fazy». Całość taka jest z tej racji wyznaczona pod względem swoich czasowych granic w sposób bardzo niejednoznaczny. Tak samo więc niejednoznacznie wyznaczona jest denotacja definiowanej tą drogą nazwy. Nazwę indywiduową o niejednoznacznie przyporządkowanej denotacji uważać możemy za termin nieostry, rozszerzając odpowiednio użytek tego słowa stosowanego na ogół tylko do predykatów. Każda zdefiniowana ostensywnie nazwa indywiduowa jest terminem nieostrym. Każda taka nazwa pozwala zatem na formułowanie twierdzeń, które z tego powodu są twierdzeniami zasadniczo nierozstrzygalnymi[3]. W przypadku zdefiniowanych ostensywnie predykatów sprawa przedstawia się — jak zobaczymy — podobnie. Zamykając tym uwagi dotyczące definicji ostensywnej nazw indywiduowych, przejdźmy obecnie do głównego dla naszych rozważań problemu definicji ostensywnej predykatów.

3. W przypadku nazwy indywiduowej przedmiotem przyporządkowywanym jej przez definicję ostensywną jest pewien przedmiot konkretny: rzecz materialna. Rzecz taką można, w dosłownym znaczeniu, spostrzec i pokazać. W przypadku predykatu przedmiotem tym jest pewien przedmiot abstrakcyjny: klasa (własność) lub stosunek. Ale klasy czy stosunku, dosłownie biorąc, ani spostrzec, ani pokazać nie potrafimy. Możemy spostrzec i pokazać jedynie konkretne rzeczy — bądź takie, które należą do danej klasy lub które łączy dany stosunek, bądź takie, które tego warunku nie spełniają. Staje wobec tego przed nami problem następujący: w jaki sposób przez wskazanie konkretnych rzeczy można wyznaczyć denotację predykatu? W szczególności, czy wystarczy się w tym celu posłużyć jako jedynymi terminami deskryptywnymi nazwami indywiduowymi owych wskazywanych rzeczy, czy też trzeba się prócz tego odwołać do innych terminów deksryptywnych, o charakterze predykatów? Procedurę odpowiadającą tej pierwszej ewentualności chciałbym nazywać bezpośrednim sposobem interpretacji danego predykatu. Odwołuje się ona co prawda do nazw indywiduowych wskazywanych przedmiotów, ale te nazwy traktować można zawsze jako terminy okazjonalne, co sprawia, iż procedura ta jest w takiej samej sytuacji, jak opisana poprzednio procedura ostensywnego definiowania nazw indywiduowych, której przyznaliśmy charakter interpretacji bezpośredniej. Natomiast procedura odpowiadająca drugiej z wymienionych ewentualności należy najwyraźniej do pośrednich sposobów interpretacji, odwołujących się przy interpretowaniu danego predykatu do innych predykatów deskryptywnych o ustalonej z góry definicji.

3) Na ów fakt nieostrości nazw indywiduowych zwraca uwagę H. Mehlberg w książce *The Reach of Science*, Toronto 1958.

Na czym polega więc definicja ostensywna predykatu pojmowana jako bezpośredni sposób jego interpretacji? Problem ten rozważymy szczegółowo na przykładzie wybranego predykatu, uważanego za typowy termin spostrzeżeniowy, a więc termin, który tą właśnie drogą uzyskiwać ma swoją interpretację. Niech predykatem tym będzie — zgodnie z tradycją dotychczasowych dyskusji — predykat „żółty". Predykatu tego nie możemy zdefiniować ostensywnie przez pokazanie wprost jego denotacji, gdyż jest nią pewien przedmiot abstrakcyjny: klasa przedmiotów żółtych (czy też odpowiadająca jej własność: barwa żółta). Definiujemy go zatem pokazując pewne przedmioty konkretne, i to — jak się na ogół przyjmuje — zarówno takie, które do tej denotacji chcemy zaliczyć, tj. przedmioty żółte, jak i takie, których do niej zaliczyć nie chcemy, tj. przedmioty nieżółte. O pierwszych orzekamy definiowany predykat, o drugich — jego negację. Najprostszym sposobem ujęcia tak rozumianej definicji ostensywnej predykatu „żółty" jest przedstawienie jej w postaci następującego układu postulatów:

(3)
$$
\begin{array}{ll}
a_1 \in \dot{Z} & \sim b_1 \in \dot{Z}, \\
\dots & \dots \\
a_n \in \dot{Z} & \sim b_m \in \dot{Z},
\end{array}
$$

gdzie $\dot{Z}$ jest skrótem predyktu „żółty", a $a_1, \ldots, a_n, b_1, \ldots, b_m$ — nazwami indywiduowymi wskazywanych przedmiotów. Ich rolę pełnić może termin okazjonalny „to" łącznie z gestami wskazującymi owe przedmioty.

Tak pojęta definicja ostensywna predykatu $\dot{Z}$ przyporządkowuje mu jako denotację klasę spełniającą podany układ postulatów. Natychmiast jednak nasuwa się spostrzeżenie, iż denotacja ta wyznaczona jest w sposób niesłychanie wieloznaczny. Może nią być dowolna klasa zawierająca jako swoje elementy przedmioty $a_1, \ldots, a_n$, a nie zawierająca przedmiotów $b_1, \ldots, b_m$; lub w języku własności: dowolna własność przysługująca przedmiotom $a_1, \ldots, a_n$, a nie przysługująca przedmiotom $b_1, \ldots, b_m$. Wszelka taka klasa czy własność spełnia postulaty (3). Oczywiście, w przypadku gdyby wskazane przedmioty $a_1, \ldots, a_n$ wyczerpywały ogół przedmiotów należących do denotacji predykatu $\dot{Z}$, można by ją na tej drodze wyznaczyć jednoznacznie, definiując wprost:

(4)
$$\dot{Z} = \{a_1, \ldots, a_n\}.$$

Tak samo byłoby w przypadku, gdyby przedmioty $b_1, \ldots, b_m$ stanowiły ogół przedmiotów, które nie podpadają pod termin $\dot{Z}$. Ten ostatni można by zdefiniować *explicite*:

(5)
$$\dot{Z} = - \{b_1, \ldots, b_m\}.$$

Tak jednak z reguły nie bywa. Predykaty spostrzeżeniowe w rodzaju predykatu $\dot{Z}$ oraz ich negacje odnoszą się do nieograniczonej liczby przedmiotów. Jest rzeczą oczywistą, iż wszystkich rzeczy żółtych ani nieżółtych wymienić, a tym bardziej pokazać nie możemy. Wobec czego postulaty (3) determinują denotację predykatu $\dot{Z}$ w stopniu jedynie nieznacznym.

Toteż zwraca się niekiedy uwagę, iż postulaty (3) nie oddają tego, o co chodzi w procedurze ostensywnego definiowania predykatów. Wskazując a_1 jako przedmiot należący do denotacji predykatu $\dot{Z}$, chcemy stwierdzić nie to, że klasa $\dot{Z}$ zawiera przedmiot a_1, lecz to, że klasa $\dot{Z}$ zawiera przedmioty takie, jak a_1. Intencję tę oddaje wypowiedź o charakterze definicji równoważnościowej:

(6) $$\bigwedge_x (x \in \dot{Z} \equiv x \text{ jest taki, jak } a_1).$$

Wypowiedź ta zawiera jednak wysoce zagadkowy zwrot „taki, jak", który wymaga wyjaśnienia. Jeśli przy tym wypowiedź (6) stanowić ma bezpośrednią interpretację terminu $\dot{Z}$, zwrot ten musi być wyjaśniony bez odwoływania się do jakichkolwiek terminów deskryptywnych. Niestety, wszystkie znane lub nasuwające się propozycje są wyraźnie nieadekwatne.

Taki charakter mają przede wszystkim dwie interpretacje skrajne:

1. x jest taki, jak y — to tyle, co — x i y mają pewną własność wspólną;

2. x jest taki, jak y — to tyle, co x i y mają wszystkie własności wspólne.

Przy interpretacji pierwszej definicja (6) przybiera postać:

(7) $$\bigwedge_x (x \in \dot{Z} \equiv \bigvee_Y (x \in Y \wedge a_1 \in Y)),$$

przy interpretacji drugiej — postać:

(8) $$\bigwedge_x (x \in \dot{Z} \equiv \bigwedge_Y (a_1 \in Y \rightarrow x \in Y)).$$

Oba sformułowania są wyraźnie nietrafne. Każdy przedmiot ma jakąś własność wspólną z a_1 — chociażby własność bycia danym przedmiotem lub a_1. W języku klas staje się to jeszcze wyraźniejsze: istnieje zawsze klasa zawierająca zarówno x, jak i a_1; na przykład klasa $\{x, a_1\}$. A zatem zgodnie z definicją (7) klasa $\dot{Z}$ byłaby identyczna z klasą pełną! Z drugiej strony, przedmiot mający wszystkie własności a_1 — to po prostu przedmiot identyczny z a_1. A więc wedle definicji (8) klasa $\dot{Z}$ — to klasa $\{a_1\}$!

Nie ratuje sytuacji fakt odwołania się — jak to miało miejsce poprzednio — do dalszych przedmiotów wzorcowych: pozytywnych $a_2, ..., a_n$ i negatywnych $b_1, ..., b_m$. Zamiast definicji (7) i (8) otrzymujemy wówczas definicje[4]:

(9) $$\bigwedge_x (x \in \dot{Z} \equiv \bigvee_Y (x \in Y \wedge a_1 \in Y \wedge ... \wedge a_n \in Y \wedge \sim b_1 \in Y \wedge ... \wedge \sim b_m \in Y)),$$

(10) $$\bigwedge_x (x \in \dot{Z} \equiv \bigwedge_Y (a_1 \in Y \wedge ... \wedge a_n \in Y \wedge \sim b_1 \in Y \wedge ... \wedge \sim b_m \in Y \rightarrow x \in Y)).$$

Prowadzą one do podobnych konsekwencji, co poprzednie. Definicja (9) utożsamia klasę $\dot{Z}$ z klasą obejmującą wszystkie przedmioty z wyjątkiem wzorców negatywnych. Dla każdego bowiem x różnego od $b_1, ..., b_m$ istnieje klasa zawierająca x oraz $a_1, ..., a_n$, a nie zawierająca $b_1, ..., b_m$; na przykład klasa $\{x, a_1, ..., a_n\}$. To samo wyrazić można w języku własności: istnieje własność przysługująca wspólnie przedmiotom $x, a_1, ..., a_n$, a nie przysługująca żadnemu z przedmiotów $b_1, ..., b_m$; na przykład własność bycia tożsamym z x lub z a_1, ... lub z a_n. Natomiast definicja (10) wyznacza — jak się łatwo

4) Prostszą postać definicji (10) przedstawia wypowiedź:

$$\bigwedge_x (x \in \dot{Z} \equiv \bigwedge_Y (a_1 \in Y \wedge ... \wedge a_n \in Y \rightarrow x \in Y)).$$

Określa ona $\dot{Z}$ tak samo, jak tamta, przy założeniu, iż żaden z wzorców negatywnych nie jest identyczny z jakimś wzorcem pozytywnym.

przekonać — jako denotację predykatu Ż klasę składającą się wyłącznie z pozytywnych przedmiotów wzorcowych: $\{a_1, ..., a_n\}$[5].

Wniosek z naszych dotychczasowych rozważań jest zdecydowanie negatywny. Uwzględnione przez nas definicje równoważnościowe predykatu Ż [na przykład (9) czy (10)] mają charakter paradoksalny. Przyporządkowują predykatowi Ż denotacje najzupełniej różne od zamierzonej. Natomiast wypowiedzi definicyjne o charakterze postulatów [na przykład typu (3)] determinują tę denotację tylko częściowo, i to w stopniu bez porównania mniejszym, niż to ma miejsce w rzeczywistej praktyce ostensywnego definiowania terminów. Fakt ten jest konsekwencją ograniczenia się przy omawianych definicjach ostensywnych do nazw indywiduowych poszczególnych wskazywanych rzeczy jako do jedynych terminów pozalogicznych. Wyłącznie za pośrednictwem tych nazw (na przykład terminu okazjonalnego „to") usiłuje się scharakteryzować klasę (własność), która stanowić ma denotację predykatu Ż. Stąd nieuchronna nieadekwatność lub wieloznaczność takiej charakterystyki.

Czy konsekwencji tych nie można uniknąć — choćby częściowo — bez rezygnacji z bezpośredniego charakteru takiej procedury? Stawia się nieraz sprawę tak, jak gdyby można to było osiągnąć przez odpowiedni dobór przedmiotów wzorcowych. Weźmy dla przykładu nasz spostrzeżeniowy predykat Ż. Można — twierdzi się — za pomocą postulatów (3) wyznaczyć jego denotację w sposób o wiele ściślejszy, dobierając odpowiednio przedmioty wzorcowe $a_1, ..., a_n, b_1, ..., b_m$. Jako wzorce pozytywne należy wskazywać takie przedmioty żółte, które różnią się pomiędzy sobą możliwie znacznie, jako wzorce negatywne — analogiczne przedmioty nieżółte, przy czym wzorce negatywne winny się od wzorców pozytywnych różnić możliwie nieznacznie. W przypadku idealnym przedmioty wzorcowe $a_1, ..., a_n, b_1, ..., b_m$ winny być takie, aby barwa żółta była jedyną własnością przysługującą wspólnie przedmiotom $a_1, ..., a_n$, a nie

5) W cytowanej pracy J. Kotarbińskiej omawiany obecnie zabieg ostensywnego definiowania predykatu Ż ujmuje się w sposób nieco odmienny. Uważa się mianowicie, iż wskazując pozytywne i negatywne przykłady własności Ż, zakładamy, że istnieje własność taka, którą jakiś przedmiot posiada wtedy i tylko wtedy, gdy ma własność Ż, i która przysługuje wszystkim wskazanym wzorcom pozytywnym, a nie przysługuje żadnemu z wzorców negatywnych. Jest to tzw. założenie definicyjne. Ujęte symbolicznie głosi ono:

$$\bigvee_Y [\bigwedge_x (x \in \dot{Z} \equiv x \in Y) \wedge a_1 \in Y \wedge ... \wedge a_n \in Y \wedge \sim b_1 \in Y \wedge ... \wedge \sim b_m \in Y].$$

Sformułowanie powyższe okazuje się jednak logicznie równoważne układowi postulatów (3). Nie głosi ono nic ponad to, co głoszą te postulaty, i tak samo wieloznacznie jak one determinuje klasę Ż. Może nią być wszelka klasa zawierająca $a_1, ..., a_n$, a nie zawierająca $b_1, ..., b_m$. Ściśle biorąc, we wspomnianej pracy występuje sformułowanie różniące się nieco od podanego:

$$\bigvee_Y \bigwedge_x (x \in \dot{Z} \equiv x \in Y \wedge a_1 \in Y \wedge ... \wedge a_n \in Y \wedge \sim b_1 \in Y \wedge ... \wedge \sim b_m \in Y).$$

Determinuje ono denotację Ż w stopniu jeszcze mniejszym niż poprzednie. Ż może być bądź dowolną klasą zawierającą $a_1, ..., a_n$ i nie zawierającą $b_1, ..., b_m$, bądź klasą pustą. Wydaje się więc, iż sformułowanie poprzednie lepiej oddaje istotę omawianej propozycji. Należy dodać, iż we wspomnianej pracy oprócz sformułowania przytoczonego występuje — i to jako propozycja naczelna — sformułowanie inne, do którego powrócimy w dalszym ciągu pracy.

przysługującą żadnemu z przedmiotów $b_1, ..., b_m$. Wskazanie takich przedmiotów wzorcowych miałoby wystarczać do jednoznacznej charakterystyki denotacji terminu Ż. Tak też byłoby istotnie, gdyby przypadek taki mógł mieć miejsce. Widzieliśmy jednak, iż nie ma i nie może być przedmiotów, które by spełniały żądany warunek. Jakkolwiek liczne i różnorodne dobralibyśmy przedmioty wzorcowe, zawsze istnieć będzie poza barwą żółtą nieskończona liczba własności, które odznaczają się tym, iż przysługują wszystkim wzorcom pozytywnym, a nie przysługują żadnemu z wzorców negatywnych. Jedną z takich własności będzie po prostu własność bycia wzorcem pozytywnym, a więc a_1 lub a_2 ..., lub a_n. To samo wyrazić można w języku klas: istnieje prócz klasy przedmiotów żółtych nieskończona liczba klas zawierających wszystkie wzorce pozytywne, a nie zawierających żadnego z wzorców negatywnych, na przykład klasa $\{a_1, ..., a_n\}$. Toteż niezależnie od tego, jakie byłyby przedmioty wzorcowe, o żadnym przedmiocie nieidentycznym z którymś z nich nie potrafimy rozstrzygnąć, czy należy, czy też nie należy do scharakteryzowanej przez postulaty (3) klasy Ż. Nawet jeśli ów przedmiot c ma barwę taką samą, jak któryś z wzorców pozytywnych, nie mamy prawa orzec, iż należy on do Ż, bo przecież Ż może być klasą $\{a_1, ..., a_n\}$. Z drugiej strony, mimo iż c może być barwy takiej samej, jak pewien wzorzec negatywny, nie możemy go zasadnie z Ż wykluczyć, skoro klasa $\{a_1, ..., a_n, c\}$ również spełnia nasze postulaty. A zatem żaden dobór przedmiotów wzorcowych nie jest w stanie uczynić charakterystyki klasy Ż bardziej jednoznaczną [lub — w przypadku definicji (9) czy (10): bardziej adekwatną]. Taki czy inny dobór przedmiotów wzorcowych może, jak się wydaje, prowadzić do tego celu tylko przy przyjęciu poza omawianymi wypowiedziami pewnych założeń dodatkowych. Założenia te z góry ograniczać muszą wchodzące w grę własności (klasy) do własności pewnego określonego rodzaju. Nie możemy brać pod uwagę przy ostensywnym definiowaniu terminów wszelkich własności przysługujących wskazywanym przedmiotom, lecz tylko pewne spośród nich. Wtedy tylko odpowiedni dobór przedmiotów wzorcowych prowadzić może do wyeliminowania spośród owych własności uwzględnianych — wszystkich prócz jednej mającej stanowić denotację definiowanego predykatu.

4. Co musimy założyć o branych pod uwagę własnościach przedmiotów wskazywanych, aby drogą definicji ostensywnej wyodrębnić własność zamierzoną? Na podstawie tego, co się na ten temat zwykło mówić, nasuwa się szereg możliwych odpowiedzi. Narzuca się przede wszystkim odpowiedź, iż własność definiowana ostensywnie — to w każdym razie własność «spostrzegalna». Istotnie, procedura ostensywnego definiowania terminów zakłada, jak się wydaje, iż własność, którą w ten sposób usiłuje się przyporządkować danemu terminowi, jest własnością «dostępną bezpośredniej obserwacji», własnością, którą «można spostrzec». Założenie takie, choć narzucające się w sposób oczywisty, ma jednak sens wysoce zagadkowy. Pojęcie „własności spostrzegalnej" jest, mimo swego obiegowego charakteru, pojęciem ogromnie niejasnym. Wspominałem już, iż o «spostrzegalności» można mówić w sensie dosłownym tylko w stosunku do jednostkowych, konkretnych rzeczy. Nie można w tym samym, literalnym

sensie mówić o «spostrzegalności» takich ogólnych, abstrakcyjnych przedmiotów, jak własność czy stosunek. Spostrzegam w tym sensie ten czy ów żółty przedmiot, ale nie — żółtą barwę. W jakim więc znaczeniu mówić można o «spostrzegalnych» własnościach? Nie umiem dać odpowiedzi na to pytanie. Co więcej, nie wydaje mi się, aby pojęcie to miało w ogóle jakiś ściślej określony sens. W każdym razie jego zakres pozostaje w najwyższym stopniu nieustalony. Zgodzić się można jeszcze na to, iż określony odcień barwny jest własnością spostrzegalną, a na przykład określony stan namagnesowania taką własnością nie jest. Ale już przy własności takiej, jak alternatywa dwóch odcieni barwnych, powstają wątpliwości. Czy można istotnie spostrzec, iż dany przedmiot ma odcień taki lub taki?[6] A jeśli to jest wątpliwe, to wątpliwości budzi również spostrzegalny charakter własności takich, jak barwa żółta czy czerwona. Nie są to przecież określone odcienie barwne, lecz alternatywy szeregu takich odcieni. A więc może spostrzegamy nie to, że dany przedmiot ma barwę żółtą, lecz to, że ma pewien odcień takiej barwy. To samo dotyczy *a fortiori* takich własności barwnych, jak żółty lub czerwony. Czy można powiedzieć: widzę, że to jest żółte lub czerwone? Czy raczej: widzę, że to jest żółte (*resp.* czerwone), i z tego wnioskuję, że jest żółte lub czerwone? Podobny charakter miałaby własność taka, jak barwa «ciepła», oraz wszelkie własności barwne «złożone» — nie tylko alternatywnie — z określonych odcieni barwnych. Opuszczając sferę własności barwnych, przechodzimy na teren jeszcze bardziej niepewny. Czy jest własnością spostrzegalną własność bycia tym oto przedmiotem? A własność bycia tym lub tamtym przedmiotem, na przykład wspomniana przez nas własność bycia wzorcem pozytywnym: a_1 lub a_2..., lub a_n? Nie potrafię odpowiedzieć na te pytania. Dość wyraźnie natomiast zarysowują się pewne wnioski ogólne.

I tak, sądzę, iż jakiekolwiek rozróżnienie «spostrzegalnych» i «niespostrzegalnych» przedmiotów abstrakcyjnych dotyczyć może tylko własności pojmowanych jako coś różnego od odpowiadających im klas. Przeprowadzone więc być może tylko w języku intensjonalnym. Klasa przedmiotów żółtych jest identyczna z klasą przedmiotów wysyłających fale elektromagnetyczne o określonej długości. Tymczasem zgodnie z panującym zwyczajem językowym barwę żółtą zalicza się do własności spostrzegalnych, a własność wysyłania fal elektromagnetycznych o określonej długości — do własności niespostrzegalnych. Ten fakt, jak się wydaje, przemawia za tym, iż owo rozróżnienie ma charakter nie tyle «ontologiczny», co «językowy». To nie dana klasa, lecz pewna jej charakterystyka może być «spostrzegalna» (lub raczej «spostrzeżeniowa»). Istotnie, w przeciwieństwie do zagadkowego pojęcia „własności spostrzegalnej", pojęcie „terminu spostrzeżeniowego" ma, jak się wydaje, sens wyraźniejszy.

Pojęcie to określa się często w sposób następujący[7]:

6) Negatywną odpowiedź na to pytanie daje np. H. Mehlberg, op. cit.

7) Por. R. Carnap, „Testability and Meaning", *Philosophy of Science*, 1936–1937, nr 3–4.

Predykat P jest terminem spostrzeżeniowym dla osoby O, jeżeli dla pewnego x: osoba O może rozstrzygnąć bezpośrednio na podstawie spostrzeżenia, czy x jest przedmiotem należącym do denotacji predykatu P.

Określenie to wymaga oczywiście dalszej eksplikacji, która tutaj nie nastręcza poważniejszych trudności. W tej chwili musimy z niej jednak zrezygnować i zadowolić się paru ogólnikowymi uwagami. W przytoczonym określeniu nie żądamy, aby o każdym przedmiocie można było rozstrzygnąć bezpośrednio na podstawie spostrzeżenia, czy przedmiotowi temu predykat P przysługuje. Gdyby to bowiem był przedmiot niespostrzegalny, rozstrzygnięcie takie byłoby niemożliwe, niezależnie od spostrzeżeniowego czy niespostrzeżeniowego charakteru predykatu P. Dlatego też ograniczamy się do żądania, by w stosunku do pewnych przedmiotów — przedmiotów spostrzegalnych — rozstrzygnięcie takie było możliwe. Pojęcie terminu spostrzeżeniowego jest tutaj zrelatywizowane do osoby O. Wydaje się, iż relatywizację tę zastąpić możemy relatywizacją do języka J. Określenie nasze żądać będzie wówczas, aby każda «normalna» osoba mówiąca językiem J mogła o pewnym przedmiocie rozstrzygnąć bezpośrednio na podstawie spostrzeżenia, czy przedmiotowi temu predykat P przysługuje. Charakter spostrzeżeniowy danego terminu zależny jest w ten sposób od określonego języka, w szczególności — od sposobu, w jaki dany termin do określonego języka został wprowadzony. Można, jak się zdaje, przyjąć, iż podział na tak rozumiane terminy spostrzeżeniowe i niespostrzeżeniowe pokrywa się z podziałem na terminy zinterpretowane bezpośrednio i terminy zinterpretowane pośrednio. Termin, któremu nadano interpretację przez proste wskazanie pewnych przedmiotów, może być stosowany bezpośrednio na podstawie spostrzeżenia. Natomiast zastosowanie terminu zinterpretowanego przy pomocy innych terminów sprowadza się bezpośrednio do zastosowania owych terminów, a tylko pośrednio — do spostrzeżeń. Ten sam termin może być przy tym do jednego języka wprowadzony drogą interpretacji bezpośredniej, a do drugiego — drogą interpretacji pośredniej. Mówiąc o „tym samym" terminie, mam oczywiście na myśli nie tylko jego równokształtność, ale i fakt posiadania w obu językach tej samej denotacji.

Jeżeli terminy zinterpretowane bezpośrednio uważać będziemy za terminy spostrzeżeniowe, a definicję ostensywną pojmować będziemy jako metodę interpretacji bezpośredniej, to wszelki termin wprowadzony do danego języka drogą definicji ostensywnej będzie miał tym samym w tym języku charakter terminu spostrzeżeniowego. Innymi słowy: wszelka własność scharakteryzowana ostensywnie będzie własnością «spostrzegalną». A zatem założenie ograniczające brane pod uwagę własności przedmiotów wzorcowych dla ich własności «spostrzegalnych», nie stanowi w istocie żadnego ograniczenia. I tak, wspomniana własność bycia wzorcem pozytywnym pozostaje nadal jedną z możliwych interpretacji definiowanego ostensywnie predykatu $Ż$. Założenie, które by czyniło tę interpretację bardziej jednoznaczną, musi nakładać na własność $Ż$ jakiś warunek węższy.

Jako warunek taki wymienia się niekiedy założenie ograniczające brane pod uwagę własności przedmiotów wzorcowych do własności spostrzegalnych określonej sfery zmysłowej, na przykład w przypadku predykatu takiego, jak $\dot{Z}$ — do własności «wzrokowych». Sam proces ostensywnego definiowania terminu $\dot{Z}$, na który składają się m.in. pewne gesty wskazujące, sugeruje takie założenie. Założenie to jest jednak najwyraźniej niewystarczające. Samo pojęcie „własności spostrzegalnej wzrokowo" nasuwa podobne uwagi, co pojęcie „własności spostrzegalnej". Poza tym — choć węższe od tamtego — jest dla naszych celów nadal o wiele za szerokie. Wystarczy zwrócić uwagę na to, iż niezależnie od doboru wzorców pozytywnych $a_1, ..., a_n$ i negatywnych $b_1, ..., b_m$, definiowany ostensywnie z ich pomocą przez postulaty (3) predykat $\dot{Z}$ może być interpretowany jako denotujący taką «spostrzegalną wzrokowo» własność, jak kształt a_1 lub kształt a_2..., lub kształt a_n!

5. Rozważania dotychczasowe, jak się wydaje, nieodparcie nasuwają wniosek, iż założeniem niezbędnym do jednoznacznej i adekwatnej interpretacji predykatu $\dot{Z}$ jest założenie stwierdzające wyraźnie, że idzie tu o barwę. Bez takiego założenia co do rodzaju własności $\dot{Z}$ żaden zestaw przedmiotów wzorcowych nie jest w stanie wyeliminować interpretacji niewłaściwych. Musimy wiedzieć z góry, iż własności przedmiotów wzorcowych wchodzące w grę jako możliwe interpretacje definiowanego ostensywnie predykatu $\dot{Z}$ — to barwy tych przedmiotów. Chcąc wyraźnie sformułować takie założenie, musimy dołączyć do definicji ostensywnej w postaci postulatów (3) dodatkowy postulat głoszący, iż $\dot{Z}$ jest barwą. Otrzymamy w ten sposób układ postulatów:

(11)
$$\begin{array}{lll} a_1 \in \dot{Z} & \sim b_1 \in \dot{Z} & \dot{Z} \in \mathbb{B} \\ \dots & \dots & \\ a_n \in \dot{Z} & \sim b_m \in \dot{Z} & \end{array}$$

W podobny sposób możemy uzupełnić definicje ostensywne typu równoważnościowego. Definicje te traktowaliśmy jako próby precyzacji ogólnikowej wypowiedzi (6). Wypowiedź ta, uzupełniona informacją o rodzaju własności definiowanej, brzmi obecnie:

(12)
$$\bigwedge_x (x \in \dot{Z} \equiv x \text{ jest pod względem barwy taki, jak } a_1)$$

a definicja (9), stanowiąca pewną interpretację wypowiedzi (6), odwołującą się do szeregu przedmiotów wzorcowych, przybiera postać następującą:

(13)
$$\bigwedge_x (x \in \dot{Z} \equiv \bigvee_{Y \in \mathbb{B}} (x \in Y \wedge a_1 \in Y \wedge ... \wedge a_n \in Y \wedge \sim b_1 \in Y \wedge ... \wedge \sim b_m \in Y)).$$

Analogicznie uzupełnić można pozostałe rodzaje definicji.

Natychmiast powstają oczywiście pytania dotyczące sensu wprowadzonego tutaj terminu $\mathbb{B}$. Jak termin ten należy rozumieć? W szczególności, jakie musi być jego znaczenie, aby dołączony do przytoczonych wypowiedzi definicyjnych pozwalał istotnie na jednoznaczną i adekwatną interpretację definiowanego predykatu $\dot{Z}$? Jaką drogą wreszcie znaczenie to może być ustalone? Spróbujemy w dalszym ciągu udzielić odpowiedzi na te pytania. Zanim to jednak uczynimy, trzeba wyraźnie stwierdzić fakt następujący. Niezależnie od tego, w jaki sposób sprecyzujemy sens terminu $\mathbb{B}$, termin ten ma niewątpliwie charakter terminu deskryptywnego. Reprezentuje on deksryptywny predykat drugiego rzędu, denotujący pewien rodzaj cech przysługujących przedmiotom materialnym. A zatem fakt odwołania się przy definicji ostensywnej predykatu Ż do deskryptywnego predykatu B sprawia, iż definicja taka przestaje być bezpośrednim sposobem interpretacji predykatu Ż. Tutaj jednak trzeba dodać podobne zastrzeżenie, jak przy definicji ostensywnej nazw indywiduowych. Definicja ostensywna predykatu Ż traktowana jako sposób bezpośredni nie może być uznana za zabieg skuteczny, jeśli tę skuteczność rozumieć w sensie ściślejszym. Wypowiedzi wchodzące w skład takiej procedury nie pociągają logicznie twierdzenia, iż denotacją predykatu Ż jest klasa przedmiotów żółtych. Nie prowadząc do takiego twierdzenia logicznie, procedura taka może jednak prowadzić do niego faktycznie. I tak, zabieg polegający na wygłoszeniu postulatów (3) oraz wskazaniu odpowiednich przedmiotów może wywoływać u osoby, do której jest skierowany, skutek polegający na rozumieniu predykatu Ż jako nazwy przedmiotów żółtych. Może więc uczynić tę osobę zdolną do właściwego posługiwania się owym predykatem. I w tym luźniejszym sensie definicja ostensywna predykatu Ż może być, nawet jako sposób bezpośredni, zabiegiem skutecznym. Inna rzecz, czy tak jest istotnie i jak do tego dochodzi. Ale to nie tyle sprawa logiki, co psychologii. Skuteczności takiego zabiegu nie gwarantują zależności logiczne, lecz psychologiczne. Interesując się tutaj interpretacją terminów spostrzeżeniowych z logicznego punktu widzenia, przejdziemy do analizy definicji ostensywnej predykatu Ż ujmującej tę procedurę jako sposób interpretacji pośredniej.

Jako sposób taki definicja ta dopuszcza ujęcia różne. Różnice te wiążą się z różnicami w sposobie pojmowania terminu „barwa", do którego się wszystkie tego rodzaju definicje bezpośrednio lub pośrednio odwołują. Termin to bez wątpienia wieloznaczny. W języku potocznym używany bywa najczęściej w znaczeniu bardzo szerokim. Trudno podać definicję tak rozumianej barwy. Chciałbym tylko zwrócić uwagę na to, jak szeroki jest zakres tego pojęcia. Oto przy rozumieniu tym barwą jest nie tylko określony odcień barwy, na przykład barwa tego oto przedmiotu, ale i własności takie, jak barwa żółta, barwa «ciepła», barwa tego lub tamtego przedmiotu, barwa żółta lub czerwona, barwa nieżółta itp. Oznaczmy owo szerokie pojęcie barwy przez B′. Mimo iż B′ odpowiada najlepiej znaczeniu potocznemu, jest to termin dla naszych celów nieprzydatny. Założenie ograniczające możliwe interpretacje predykatu Ż do barw tak rozumianych uściśla interpretację tego predykatu w stopniu niewielkim.

Przypuśćmy, że figurujący w postulatach (11) termin B rozumiany jest jako B′. Jakie interpretacje dopuszcza zdefiniowany za pomocą takich postulatów predykat Ż? Zgodnie z pierwotnymi postulatami (3), klasą Ż mogła być, jak pamiętamy, jakakolwiek klasa zawierająca $a_1, ..., a_n$, a nie zawierająca $b_1, ..., b_m$. W stosunku do każdego przedmiotu różnego od przedmiotów wzorcowych mieliśmy całkowitą swobodę decydowania, czy go do Ż zaliczyć, czy nie. Obecnie sytuacja zmienia się o tyle, że swoboda decyzji przysługuje nam tylko w stosunku do tych przedmiotów, które są barwy różnej od barwy każdego z przedmiotów wzorcowych. Jeśli c jest przedmiotem nieidentycznym z żadnym przedmiotem wzorcowym, ale równobarwnym z jakimś wzorcem pozytywnym, na przykład a_1, klasa $\dot{Z}$ musi przedmiot c obejmować. Jeśli natomiast przedmiot c jest różny od przedmiotów wzorcowych, lecz równobarwny z jakimś wzorcem negatywnym, na przykład b_1, musimy go z klasy $\dot{Z}$ wykluczyć. Ten stan rzeczy jest konsekwencją pewnej zależności ogólnej:

(14) $$\bigwedge_{x,y} (x \in \dot{Z} \rightarrow (y \,\mathrm{Rb}\, x \rightarrow y \in \dot{Z})),$$

gdzie Rb symbolizuje stosunek równobarwności dwóch przedmiotów. Zależność (14) z kolei wydaje się konsekwencją postulatu:

(15) $$\dot{Z} \in \mathcal{B}',$$

dołączonego do pierwotnych postulatów (3). Wydaje się bowiem, iż jakkolwiek byśmy zdefiniowali termin $\mathcal{B}'$, jego definicja musiałaby implikować twierdzenie (14) oraz analogiczne twierdzenia dla pozostałych własności należących do $\mathcal{B}'$. Wymaga tego znaczenie terminu „barwa" — niezależnie od takiej czy innej jego precyzacji.

A zatem przy rozumieniu terminu $\mathcal{B}$ jako $\mathcal{B}'$ postulaty (11) dopuszczają jako denotację predykatu $\dot{Z}$ tylko taką klasę, która zawiera wszystkie przedmioty równobarwne z jakimś wzorcem pozytywnym, a nie zawiera żadnego przedmiotu równobarwnego z którymś z wzorców negatywnych. Ale też i wszelką taką klasę! A tym samym denotację tę określają w sposób bardzo wieloznaczny. O żadnym przedmiocie nierównobarwnym z przedmiotami wzorcowymi nie potrafimy powiedzieć, czy pod termin $\dot{Z}$ podpada, czy nie. Niech d będzie takim przedmiotem. Jako klasę $\dot{Z}$ możemy przyjąć po prostu klasę przedmiotów równobarwnych z $a_1, ..., a_n$. Wówczas przedmiot d do klasy $\dot{Z}$ należeć nie będzie, mimo iż może to być przedmiot o odcieniu barwy żółtej, leżącym pomiędzy odcieniami przysługującymi przedmiotom $a_1, ..., a_n$. Ale jako klasę $\dot{Z}$ możemy przyjąć równie dobrze klasę przedmiotów równobarwnych z przedmiotami $a_1, ..., a_n, d$. Wówczas klasa $\dot{Z}$ obejmować będzie przedmiot d, mimo iż może to być przedmiot barwy czerwonej. A zatem klasa $\dot{Z}$ nie tylko nie jest zdeterminowana całkowicie, ale jest zdeterminowana w stopniu bez porównania mniejszym niż — zdeterminowana również nie całkowicie — denotacja predykatu „żółty" w jego znaczeniu potocznym. Natomiast definicja (13), w której termin $\mathcal{B}$ rozumiany jest jako $\mathcal{B}'$, utożsamia — jak się łatwo przekonać — klasę $\dot{Z}$ z klasą obejmującą wszystkie przedmioty z wyjątkiem tych, które są równobarwne z jakimś wzorcem negatywnym. A zatem jest definicją wyraźnie nieadekwatną. Są to konsekwencje przyjętego rozumienia terminu $\mathcal{B}'$, przy którym pod termin ten podpada m.in. klasa taka, jak wspomniana wyżej klasa przedmiotów równobarwnych z $a_1, ..., a_n, d$. Mówiąc językiem własności: za barwę uważa się przy tym rozumieniu m.in. własność posiadania takiego odcienia barwnego, jak przedmiot a_1, lub takiego, jak a_2..., lub takiego, jak a_n. Konsekwencji tych przeto uniknąć można przez przyjęcie w charakterze terminu $\mathcal{B}$ predykatu rozumianego znacznie węziej niż predykat $\mathcal{B}'$.

6. Spośród owych węższych pojęć barwy jedno wyróżnia się stanowczo jako stosunkowo wyraźne, a zarazem przydatne do naszych celów. Mam na myśli znaczenie

najwęższe: barwę ograniczoną do klasy określonych odcieni barwnych. Oznaczmy pojęcie to przez B''. Scharakteryzować je można wskazując na to, iż z założenia:

(16) $$\dot{Z} \in B''$$

wynika nie tylko twierdzenie (14), które przytaczaliśmy jako konsekwencję założenia (15), ale i twierdzenie odwrotne:

(17) $$\bigwedge_{x,y} (x \in \dot{Z} \rightarrow (y \in \dot{Z} \rightarrow y \operatorname{Rb} x)),$$

które z założenia tamtego bynajmniej nie wynikało. Analogiczne zależności zachodzą oczywiście i dla pozostałych własności klasy B''. Zależności te uwidoczniają ścisły związek między tak rozumianą barwą B'' a stosunkiem równobarwności dwóch przedmiotów Rb. Związek ten pozwala na zdefiniowanie terminu B'' za pomocą terminu Rb. Przy założeniu zwrotności, symetryczności i przechodniości stosunku Rb, definicja taka przybiera postać definicji przez abstrakcję:

(18) $$\bigwedge_{Y} [Y \in B'' \equiv \bigwedge_{x,y} (x \in Y \rightarrow (y \in Y \equiv y \operatorname{Rb} x))].$$

Barwa B'' utożsamiona zostaje z rodziną klas abstrakcji stosunku równobarwności Rb.

Odkładając na później problemy dotyczące terminu Rb, zastanówmy się obecnie nad konsekwencjami ograniczenia możliwych interpretacji definiowanego ostensywnie predykatu $\dot{Z}$ do barw traktowanych jako elementy B''. Weźmy pod uwagę definicję ostensywną w postaci postulatów typu (11). Jest rzeczą widoczną, iż przy rozumieniu B jako B'' postulaty takie wyznaczają klasę $\dot{Z}$ w sposób jednoznaczny. Wystarczy przy tym do tego celu wskazanie jednego wzorca pozytywnego a_1. Postulaty:

(19) $$a_1 \in \dot{Z} \quad \dot{Z} \in B'',$$

implikują na gruncie definicji (18) terminu B'' definicję równoważnościową predykatu $\dot{Z}$:

(20) $$\bigwedge_{x} (x \in \dot{Z} \equiv x \operatorname{Rb} a_1).$$

$\dot{Z}$ jest na mocy tej definicji klasą obejmującą wszystkie, i tylko takie przedmioty, które są równobarwne z wzorcem a_1. Wskazywanie dalszych przedmiotów wzorcowych jest w najlepszym razie czynnością zbędną. Jeśli są to przedmioty równobarwne z a_1, klasa w ten sposób wyznaczona jest identyczna z klasą pierwotną. Jeśli zaś któryś z nich nie jest równobarwny z a_1, klasa tak zdefiniowana jako przedmiot sprzeczny nie istnieje. Do podobnej konkluzji dochodzimy w przypadku definicji ostensywnej typu równoważnościowego (13). I tu wystarczy ograniczyć się do jednego wzorca pozytywnego. Definicja:

(21) $$\bigwedge_{x} (x \in \dot{Z} \equiv \bigvee_{Y \in B''} (x \in Y \wedge a_1 \in Y)),$$

jest równoważna logicznie definicji (20) i wyznacza taką samą jak tamta klasę $\dot{Z}$. Podawanie dalszych wzorców równobarwnych z a_1 klasy tej nie zmienia, natomiast wskazanie wzorca nierównobarwnego z a_1 czyni z $\dot{Z}$ klasę pustą.

Tak więc definicja ostensywna predykatu $\dot{Z}$ odwołująca się do deskryptywnego terminu B'' przyporządkowuje predykatowi $\dot{Z}$ denotację w sposób jednoznaczny. Czy również w sposób adekwatny? W myśl definicji (20) $\dot{Z}$ — to po prostu barwa przedmiotu a_1, będąca jedną z klas abstrakcji stosunku równobarwności. Barwa ta jako klasa przedmiotów równobarwnych z a_1 czy też własność przysługująca swoiście takim przedmiotom utożsamiona być może z określonym odcieniem barwnym przysługującym przedmiotowi a_1. Nie może być zatem uważana za barwę żółtą, gdyż ta obejmuje szereg odcieni barwnych. Jako definicja barwy żółtej wypowiedź (20) nie jest więc definicją adekwatną. Wypowiedź tę natomiast traktować można jako definicję określonego odcienia barwy żółtej. Lepiej wobec tego użyć w niej zamiast terminu $\dot{Z}$ terminu $\dot{Z}_1$, który by ów odcień symbolizował. Nie osiągamy, jak widać, przedstawioną drogą zadowalającej definicji ostensywnej predykatu $\dot{Z}$ odpowiadającego potocznemu wyrazowi „żółty". Uzyskujemy za to sposób ostensywnego definiowania predykatów typu $\dot{Z}_1$, które odpowiadają takim wyrażeniom potocznym, jak „barwa tego przedmiotu". Rezultat ten stanowi zarazem krok w kierunku poszukiwanej definicji predykatu $\dot{Z}$.

6a. Zanim przyjrzymy się, w jaki sposób dochodzi się do tamtej definicji, musimy wyjaśnić pewne punkty dotyczące definicji predykatu $\dot{Z}_1$. Idzie tu o leżący u podstawy tej definicji termin Rb. Jaki jest sens tego terminu? I jaką drogą ten sens ustalamy? Termin Rb jest, jak wiemy, dwuargumentowym predykatem pierwszego rzędu, denotującym stosunek równobarwności pomiędzy przedmiotami. Jest to więc pewien predykat deskryptywny, podobnie jak $\dot{Z}_1$. Czy jest to również predykat spostrzeżeniowy? Czy o tym, że dwa przedmioty spostrzegalne są równobarwne, można się przekonać bezpośrednio na podstawie spostrzeżenia, przyglądając się po prostu tym przedmiotom? Na pytanie to można odpowiedzieć twierdząco, jeśli się równobarwność utożsami z nieodróżnialnością pod względem barwy. Przy tym rozumieniu dwa przedmioty są równobarwne wtedy, i tylko wtedy, gdy są nieodróżnialne pod względem barwy przy bezpośrednim porównaniu. Oznaczmy ów stosunek nieodróżnialności pod względem barwy przez Nb. Czy można istotnie wprowadzony wyżej stosunek Rb uważać za identyczny z Nb? Zgodnie z uczynionym przez nas założeniem, Rb ma być stosunkiem równościowym. Czy jest nim również Nb? Na pewno nie. Jest to stosunek zwrotny i symetryczny, lecz nieprzechodni. Przedmiot a może być nieodróżnialny od b, przedmiot b nieodróżnialny od c, a mimo to przedmiot a może się od c różnić. Gdybyśmy się w definicji barwy $\mathcal{B}''$ posłużyli predykatem Rb rozumianym jako Nb, otrzymalibyśmy konsekwencję następującą: barwa przedmiotu a byłaby identyczna z barwą przedmiotu b, ta z kolei — z barwą przedmiotu c, a mimo to barwa przedmiotu a nie byłaby identyczna z barwą przedmiotu c! Konsekwencja najwyraźniej sprzeczna z uwagi na przechodniość stosunku identyczności. Stosunek Rb wypadnie zatem uznać za różny od stosunku Nb.

Nieodróżnialność pod względem barwy a od b jest warunkiem niezbędnym, ale nie wystarczającym do tego, aby a uznać za równobarwne z b, czyli za posiadające tę samą co b barwę. Związek pomiędzy Rb a Nb ma charakter bardziej złożony. Aby a było równobarwne z b, trzeba nie tylko tego, by a było nieodróżnialne pod względem barwy od b, ale i tego, by każdy przedmiot nieodróżnialny pod względem barwy od a był zarazem nieodróżnialny pod względem barwy od b, i na odwrót. Gdyby istniał przedmiot nieodróżnialny od a, lecz odróżnialny od b, lub na odwrót, nie moglibyśmy przedmiotów a i b uważać za równobarwne. Zależność ta pozwala na zdefiniowanie predykatu Rb za pomocą predykatu Nb:

(22) $$\bigwedge_{x,y} (x \,\mathrm{Rb}\, y \equiv \bigwedge_{z}(z \,\mathrm{Nb}\, x \equiv z \,\mathrm{Nb}\, y)).$$

Tak zdefiniowany stosunek Rb ma zagwarantowaną zwrotność, symetryczność i przechodniość. Termin Rb nie jest oczywiście przy tym ujęciu terminem spostrzeżeniowym. Nie wystarczy przyjrzeć się wyłącznie przedmiotom a i b, aby przekonać się, czy są równobarwne. Trzeba by było w tym celu porównać te przedmioty z wszystkimi pozostałymi, co jest oczywiście rzeczą niewykonalną. Dlatego też doświadczenie nie pozwala nam nigdy osiągnąć w tej sprawie całkowitej pewności[8].

Definicja (22) przesuwa problem zakładanego przez definicję ostensywną terminu deskryptywnego z predykatu Rb na Nb. W jaki sposób ustalamy sens tego ostatniego? Predykat Nb możemy już, w przeciwieństwie do Rb, uważać za termin spostrzeżeniowy. Możemy, innymi słowy, traktować go jako termin interpretowany bezpośrednio. Próby zdania sprawy ze sposobu bezpośredniej interpretacji predykatu Nb doprowadzają nas jednak natychmiast do takich samych trudności, na jakie natknęliśmy się w związku z naszym wyjściowym predykatem Ż. Predykat Nb definiujemy ostensywnie przez wskazanie w charakterze wzorców pozytywnych pewnej liczby par przedmiotów nieodróżnialnych pod względem barwy, w charakterze wzorców negatywnych — pewnej liczby par przedmiotów odróżnialnych. Definicja ostensywna predykatu Nb może przybierać w tej sytuacji postać układu postulatów analogicznego do układu postulatów (3) dla predykatu Ż:

(23) $$\begin{array}{ll} a_1 \,\mathrm{Nb}\, b_1 & \sim c_1 \,\mathrm{Nb}\, d_1 \\ \dots & \dots \\ a_n \,\mathrm{Nb}\, b_n & \sim c_m \,\mathrm{Nb}\, d_m. \end{array}$$

Tak samo jak tamten, układ powyższy wyznacza denotację predykatu Nb w sposób niezmiernie wieloznaczny. Nb może być jakimkolwiek stosunkiem, który zachodzi między elementami par $(a_1, b_1), \dots, (a_n, b_n)$, a nie zachodzi między elementami par $(c_1, d_1), \dots, (c_m, d_m)$; w szczególności może być identyczny z klasą par stanowiących wzorce pozytywne $\{(a_1, b_1), \dots, (a_n, b_n)\}$. Podobnie jak w przypadku predykatu Ż, nie ratuje sytuacji odpowiedni dobór przedmiotów wzorcowych. Żaden zestaw wzorców nie daje podstawy do rozstrzygnięcia, czy między elementami pary nieidentycznej z żadną parą wzorcową stosunek Nb zachodzi, czy też nie zachodzi. Również definicja ostensywna predykatu Nb o postaci równoważnościowej, na przykład definicja:

(24) $$\bigwedge_{x,y}(x \,\mathrm{Nb}\, y \equiv \bigvee_{R}(xRy \wedge a_1Rb_1 \wedge \dots \wedge a_nRb_n \wedge \sim c_1Rd_1 \wedge \dots \wedge \sim c_mRd_m)),$$

8) Związek między pojęciem równobarwności a pojęciem nieodróżnialności przedstawiają w podobny sposób: B.Russell, *An Inquiry into Meaning and Truth*, New York 1940; H. Putnam, „Reds, Greens and Logical Analysis", *Philosophical Review*, 1956.

dzieli wady analogicznej definicji (9) predykatu Ż, wyznaczając denotację predykatu Nb w sposób jawnie za szeroki.

Podobnie jak poprzednio, jedynym sposobem uniknięcia — choćby częściowego — tych wadliwości wydaje się nałożenie na predykat Nb pewnego dodatkowego warunku, ograniczającego możliwe interpretacje tego predykatu do określonej klasy stosunków R:

(25) $$\mathrm{Nb} \in \mathbf{R}$$

Nie nasuwa się jednak tutaj tak intuicyjna interpretacja owej klasy, jak w przypadku rozważanym poprzednio. Tam nie ulegało wątpliwości, iż B — to tak czy inaczej rozumiana klasa barw. Tutaj nie bardzo wiadomo, jaką klasę stosunków miałoby owo R reprezentować. Niezależnie jednak od różnych możliwości interpretacyjnych terminu R można, jak się wydaje, założyć, iż musi to być, podobnie jak B, termin deskryptywny. Znów zatem powstaje pytanie, w jaki sposób ustalić można jego sens, i raz jeszcze stajemy wobec trudności, z którymi usiłowaliśmy uporać się poprzednio. Nie pozostaje nam przeto wobec groźby *regressus ad infinitum* nic innego, niż przyjąć, iż termin Nb zinterpretowany jest bezpośrednio, bez odwoływania się do innych terminów deskryptywnych. A jeśli chcemy jednocześnie przyjąć, iż termin Nb uzyskuje w ten sposób interpretację choć w przybliżeniu jednoznaczną i adekwatną, musimy zrezygnować z traktowania tej procedury interpretacyjnej jako zabiegu, który prowadzi do takiej konkluzji logicznie. Skuteczność tej procedury musi być rozumiana w sposób luźniejszy. Bezpośrednia interpretacja terminu Nb jest zabiegiem, który może prowadzić faktycznie do właściwego rozumienia tego terminu, mimo iż zależności logiczne, jakie tu wchodzą w grę, bynajmniej tego nie gwarantują.

6b. Sytuacja zatem przedstawia się, jak następuje. Jeśli założymy, iż drogą pewnej językowej «tresury» potrafimy uzyskać właściwą interpretację stosunku nieodróżnialności pod względem barwy Nb, możemy za pomocą Nb zdefiniować drogą definicji (22) stosunek równobarwności Rb, a za pomocą tego ostatniego — określony odcień barwy żółtej $Ż_1$, bądź bezpośrednio — drogą definicji (20), bądź pośrednio — poprzez definicję barwy B″ (18) i układ postulatów (19) lub definicję (21). Powstaje pytanie, jak przedstawia się krok końcowy: przejście od określonego odcienia żółtego $Ż_1$ do barwy żółtej Ż. Jako najprostsze nasuwa się rozwiązanie następujące. Żółta barwa — to suma wszystkich żółtych odcieni. Definiujemy zatem szereg żółtych odcieni $Ż_1, ..., Ż_n$ w sposób taki, w jaki zdefiniowaliśmy $Ż_1$, a następnie żółtą barwę Ż — jako ich sumę:

(26) $$Ż = Ż_1 \cup \ldots \cup Ż_n \,.$$

Nie potrzebujemy oczywiście wprowadzać odrębnych nazw dla poszczególnych odcieni. Możemy je scharakteryzować po prostu jako barwy odpowiednich przedmiotów wzorcowych $B(a_1), ..., B(a_n)$. Definicja powyższa przybiera wówczas formę:

(27) $$Ż = B(a_1) \cup \ldots \cup B(a_n) \,.$$

spotykaną w codziennej praktyce określania barw. Wszak wyjaśnia się nieraz, co to barwa żółta, mówiąc, że to barwa taka, jak barwa tego przedmiotu, lub taka, jak barwa tamtego, itp. Ale nie jest to procedura zadowalająca. Przecież na barwę żółtą składa się ogromna ilość odcieni, którą trudno przy takiej procedurze wyczerpać. A nieuwzględnienie jakiegoś z nich — i to odcienia leżącego pomiędzy wymienionymi — sprawia, iż przedmioty o tym odcieniu musimy z klasy $\dot{Z}$ wykluczyć. Ze względu na ten fakt oraz ze względu na fakt nieostrości potocznego predykatu „żółty" lepiej ograniczyć się do definicji cząstkowej, formułującej tylko częściowe kryteria definicyjne dla predykatu $\dot{Z}$:

(28)
$$B(a_1) \cup \ldots \cup B(a_n) \subset \dot{Z}$$
$$B(b_1) \cup \ldots \cup B(b_m) \subset -\dot{Z}.$$

Tutaj znów otrzymujemy konsekwencję znaną z rozważań poprzednich. Jeżeli przedmiot c ma odcień nie uwzględniony w definicji (28), nie mamy żadnych podstaw do rozstrzygnięcia, czy przedmiot ten jest, czy też nie jest $\dot{Z}$ — nawet jeśli w grę wchodzi odcień leżący pomiędzy odcieniami wymienionymi w definicji. Ta droga zatem nie prowadzi do zamierzonego celu[9].

Nie prowadzi do niego, jak się wydaje, dlatego, że nie uwzględnia podstawowego w tej dziedzinie faktu: uporządkowania odcieni barwnych. Przyjmuje się powszechnie, że odcienie barwne tworzą zbiór uporządkowany, przedstawiany tradycyjnie w postaci tzw. bryły barw. Jest to uporządkowanie «trójwymiarowe», gdyż dwa odcienie barwne mogą się, jak wiadomo, różnić między sobą pod względem jakości, jasności i nasycenia. Dopiero uwzględnienie tego porządku pozwala na właściwe scharakteryzowanie barw takich, jak barwa żółta czy czerwona. W jaki sposób możemy się w definicjach naszych do uporządkowania takiego odwołać? Czy trzeba, w szczególności, wprowadzać w tym celu dodatkowe predykaty deskryptywne? Uporządkowanie odcieni barwnych przeprowadzić można na podstawie stosunków różnorodnych. Jedną z takich możliwości stanowi uporządkowanie oparte na uwzględnionym już przez nas stosunku nieodróżnialności pod względem barwy Nb. Fakt ten pozwala na określenie owego porządku bez konieczności wprowadzania w charakterze terminów pierwotnych dodatkowych predykatów deskryptywnych.

Konstrukcja porządku barw oparta na stosunku Nb nie jest zadaniem prostym. Nie mam też zamiaru przedstawiać jej tutaj, nawet w postaci szkicowej. Poprzestać muszę na odwołaniu się do znanych konstrukcji Carnapa czy Goodmana. Ten ostatni zwłaszcza rozpatruje sytuację bardzo zbliżoną do obecnej, gdyż odwołującą się do

9) Warto może dodać, że do podobnych rezultatów jak powyższe dałoby się chyba dojść i na gruncie poprzedniego rozumienia barwy $\mathcal{B}'$. Trzeba się tylko wtedy posłużyć definicjami typu (8) i (10).
Wydaje się, iż definicja:
$$\bigwedge_x (x \in \dot{Z}_1 \equiv \bigwedge_{Y \in \mathcal{B}'} (a_1 \in Y \rightarrow x \in Y)),$$
określa $\dot{Z}_1$ podobnie, jak to czyniły definicje (20) i (21), tj. jako określony odcień barwy żółtej, a definicja:
$$\bigwedge_x (x \in \dot{Z}_1 \equiv \bigwedge_{Y \in \mathcal{B}'} (a_1 \in Y \wedge \ldots \wedge a_n \in Y \rightarrow x \in Y)),$$
charakteryzuje barwę żółtą $\dot{Z}$ w sposób analogiczny do definicji (26) czy (27) — jako sumę określonych odcieni barwnych przysługujących przedmiotom wzorcowym $a_1, .., a_n$. Definicje te dzielą oczywiście wady definicji poprzednich, odwołując się w dodatku do znacznie bardziej niejasnego, niż $\mathcal{B}''$, terminu $\mathcal{B}'$.

predykatu identycznego pod istotnymi względami z predykatem Nb[10]. Mówiąc ogólnikowo, konstrukcja ta polega na zdefiniowaniu, w ostatecznej instancji za pomocą predykatu Nb, pewnego stosunku porządkującego, stanowiącego bezpośrednią już podstawę owego «trójwymiarowego» porządku barw. Pierwszy krok ku temu celowi stanowi definicja stosunku łączącego przedmioty *a*, *b*, *c* takie, iż *b* „leży ze względu na barwę bezpośrednio pomiędzy" *a* i *c*. Stosunek ten utożsamić można ze stosunkiem, który zachodzi między przedmiotami *a*, *b*, *c* wtedy, i tylko wtedy, gdy: $b \text{ Nb } a \wedge b \text{ Nb } c \wedge {\sim} a \text{ Nb } c$. Opierając się na tak zdefiniowanym stosunku, określić można z kolei stosunek „leżenia pomiędzy" w zastosowaniu do odcieni barwnych. Nasuwa się wówczas następująca sugestia co do definicji predykatów typu Ż. Nie definiujemy Ż przez wyliczenie wszystkich odcieni barwnych, jakie się na Ż składają. Wymieniamy dwa odcienie skrajne, $Ż_1$ i $Ż_2$, i określamy Ż jako sumę tych wszystkich odcieni, które «leżą pomiędzy» $Ż_1$ i $Ż_2$. Oczywiście i tutaj ograniczyć się można do odpowiedniej definicji cząstkowej. W każdym razie definicje obecne nie grzeszą już tymi wadami, na które zwracaliśmy uwagę poprzednio.

W ten sposób osiągnęliśmy wreszcie cel naszej analizy: definicję predykatu Ż. Definicja ta nie stanowi jednak bezpośredniej interpretacji tego terminu. Odwołuje się do innych terminów deskryptywnych, które poprzez szereg ogniw definicyjnych sprowadzić się dają do predykatu Nb interpretowanego już bezpośrednio. Taki stan rzeczy wywołany został przez trudności, na jakie natrafiała bezpośrednia interpretacja predykatu Ż. Ale na analogiczne trudności natrafia, jak widzieliśmy, bezpośrednia interpretacja predykatu Nb. Jaką zatem korzyść daje przesunięcie tych trudności z Ż na Nb? Warto tutaj, jak się zdaje, zwrócić uwagę na punkty następujące. Przede wszystkim nasuwa się spostrzeżenie, iż wspomniane trudności są w przypadku predykatu Nb mniejsze niż w przypadku predykatu Ż. Choćby ze względu na to, iż spostrzeżeniowy charakter predykatu Nb budzi mniej wątpliwości niż spostrzeżeniowy charakter predykatu Ż. Poza tym jest to termin bez porównania ostrzejszy. Łatwiej z tych względów za pomocą odpowiednio dobranych przykładów naprowadzić na właściwe rozumienie tego terminu. Ważniejszy wydaje się jednak wzgląd inny. Predykat Nb jest terminem «logicznie wcześniejszym» od predykatu Ż. Za pomocą terminu Nb potrafimy zdefiniować wszelkie predykaty barwne, zarówno typu $Ż_1$, jak i Ż, gdy tymczasem postępowanie odwrotne nie wydaje się wykonalne. A nawet gdyby w zasadzie było możliwe, byłoby postępowaniem niesłychanie nieekonomicznym. Zamiast jedynego terminu interpretowanego bezpośrednio, mielibyśmy terminów takich ilość niezliczoną. Wydaje się wreszcie, iż przedstawiona procedura odpowiada dość dobrze faktycznej praktyce definiowania predykatów barwnych. Dotyczy to przede wszystkim predykatów typu $Ż_1$, tj. predykatów denotujących określone odcienie barwne. Tak je chyba istotnie określamy i określać musimy. Jakże inaczej zdefiniować określony odcień

10) Por. *The Structure of Appearance*, Cambridge, Mass. 1951. Por. również uwagi H. Putnama, op. cit.

barwy żółtej, niż przez wskazanie odpowiedniego przedmiotu i stwierdzenie, że dany odcień — to barwa (w sensie B'') tego oto przedmiotu? Gorzej natomiast przedstawiona procedura odpowiada, jak się wydaje, faktycznemu sposobowi definiowania predykatów typu $\dot{Z}$, tj. predykatów takich, jak „żółty" czy „czerwony". Predykaty te chyba określamy w rzeczywistości w sposób prostszy, bez pośrednictwa tylu ogniw definicyjnych. Opierać się zatem musimy — jako na predykatach zinterpretowanych bezpośrednio — na terminach różnych od Nb. Jakie terminy mogą tu wchodzić w grę? I jak dochodzimy od nich do predykatów typu $\dot{Z}$? Rozpatrzymy pokrótce niektóre spośród nasuwających się tutaj możliwości.

7. Najczęściej bierze się pod uwagę sytuację następującą. Stosunek, na którym się przy definicji predykatów typu $\dot{Z}$ opieramy, to nie stosunek nieodróżnialności, lecz podobieństwa pod względem barwy. Oznaczmy go terminem Pb. Predykat Pb uważa się, tak jak predykat Nb, za termin spostrzeżeniowy. Podobnie jak można spostrzec, czy a nie różni się pod względem barwy od b, można również bezpośrednio na podstawie spostrzeżenia przekonać się, czy a jest podobne pod względem barwy do b. Zakłada się zatem, że termin Pb jest zinterpretowany bezpośrednio. Interpretację tę pojmować można tak, jak interpretację terminu Nb. Wskazujemy jako wzorce pozytywne pary przedmiotów podobne pod względem barwy, na przykład dwa przedmioty o różnych odcieniach barwy żółtej, a jako wzorce negatywne — pary przedmiotów niepodobnych, na przykład przedmiot żółty i czerwony:

(29)
$$\begin{array}{ll} a_1 \,\mathrm{Pb}\, b_1 & \sim c_1 \,\mathrm{Pb}\, d_1 \\ \ldots & \ldots \\ a_n \,\mathrm{Pb}\, b_n & \sim c_m \,\mathrm{Pb}\, d_m \end{array}$$

Podobnie jak w przypadku terminu Nb, przyjmujemy, iż można tą drogą doprowadzić kogoś do właściwego rozumienia terminu Pb, mimo iż układ postulatów (29) wyznacza interpretację tego terminu w sposób niezmiernie wieloznaczny.

Z chwilą, gdy dysponujemy już zinterpretowanym predykatem Pb, definicja predykatu $\dot{Z}$ nie przedstawia większych trudności. Można ją zresztą ująć rozmaicie. Najprościej chyba w sposób następujący:

(30)
$$\bigwedge_x (x \in \dot{Z} \equiv x \,\mathrm{Pb}\, a_1),$$

gdzie a_1 stanowi jedyny przedmiot wskazywany w charakterze wzorca. Można powiedzieć, iż reprezentuje on «idealny» wzorzec barwy żółtej. Definicja (30) stanowi dokładny odpowiednik definicji (20). Tylko gdy tam definiowaliśmy $\dot{Z}$ jako klasę przedmiotów równobarwnych z przedmiotem wskazywanym, tutaj określamy ją jako klasę przedmiotów podobnych pod względem barwy do przedmiotu wskazywanego. Toteż tam interpretowaliśmy $\dot{Z}$ jako określony odcień żółty, tutaj — jako barwę żółtą obejmującą szereg odcieni.

Przeciwko definicji (30) można jednak podnieść zarzut następujący. Pb nie jest, w przeciwieństwie do Rb, stosunkiem równościowym. Jest to, podobnie jak Nb, stosunek zwrotny i symetryczny, lecz nieprzechodni. Nieprzechodniość stosunku Pb sprawia, iż klasa $\dot{Z}$ określona tak jak wyżej może zawierać przedmioty, które nie pozostają do siebie w stosunku Pb, mimo iż pozostają w tym stosunku do przedmiotu a_1. Jeśli fakt ten traktujemy jako dowód nieadekwatności definicji (30), możemy określić $\dot{Z}$ w sposób nieco węższy. Możemy mianowicie utożsamić klasę $\dot{Z}$ z tzw. kręgiem podobieństwa stosunku Pb[11]. Klasę $\dot{Z}$ nazywamy kręgiem podobieństwa stosunku Pb, jeśli spełnia warunek następujący:

(31) $$\bigwedge_x (x \in \dot{Z} \equiv \bigwedge_y (y \in \dot{Z} \rightarrow x \,\mathrm{Pb}\, y)),$$

tj. jeśli jest klasą, której każdy element jest podobny pod względem barwy do wszystkich jej elementów i której elementem jest każdy przedmiot podobny pod względem barwy do wszystkich jej elementów. Żądając od klasy $\dot{Z}$, aby stanowiła krąg podobieństwa stosunku Pb, charakteryzujemy ją węziej niż to czyniła definicja (30), i przez to gwarantujemy, że przedmioty, które do niej należą, podobne są pod względem barwy nie tylko do pewnego przedmiotu wzorcowego, ale i do siebie nawzajem. Tym samym uzyskujemy definicję predykatu $\dot{Z}$ nie narażoną już na poprzedni zarzut.

Jak widać, przyjęcie w charakterze terminu wyjściowego predykatu Pb pozwala dojść do predykatu $\dot{Z}$ w sposób prostszy niż oparcie się na predykacie Nb. Dodajmy przy tym, iż stosunek Pb stanowi, podobnie jak Nb, wystarczającą podstawę do uporządkowania ogółu barw. Mimo to procedurę powyższą trudno uznać za zadowalającą, a to ze względu na dość zagadkowy charakter owego terminu wyjściowego. Czy można z sensem mówić o tym, iż pewne przedmioty są do siebie podobne pod względem barwy, a inne — nie? Czyż nie jest raczej tak, iż wszelkie przedmioty są do siebie podobne pod względem barwy, a tylko jedne — w stopniu większym, a inne — w mniejszym? Jeśli tak, to z sensem można mówić tylko o podobieństwie pod względem barwy w stopniu takim a takim, a nie o podobieństwie pod względem barwy *tout court*! A zatem w przypadku predykatu Pb idzie również o podobieństwo w określonym stopniu. Ma to być podobieństwo w stopniu co najmniej takim, w jakim podobne są do siebie przedmioty wzorcowe. Czy można jednak istotnie wyznaczyć ów stopień podobieństwa drogą interpretacji bezpośredniej? A więc tak, aby samo spostrzeżenie dwóch przedmiotów wystarczało do rozstrzygnięcia, czy są one do siebie w owym stopniu podobne, czy nie? Mógłby ktoś odpowiedzieć, iż pewne stopnie podobieństwa pod względem barwy taki właśnie charakter posiadają. Świadczyć ma o tym na przykład uporczywa praktyka odróżniania w widmie słonecznym pięciu różnych barw, a nie dwóch czy dziesięciu. Patrząc na takie widmo, spostrzegamy, iż pewne odcienie są do siebie podobne, a inne — nie. Taki właśnie «spostrzegalny» stosunek podobieństwa usiłuje się przyporządkować drogą bezpośredniej interpretacji predykatowi Pb. Argumentacja ta nie wydaje się jednak w pełni przekonywająca. W widmie słonecznym wyróżnia są równie dobrze pięć, jak i siedem różnych barw. Jeszcze większe wahania występują przy ocenie podobieństwa odcieni barwnych spoza widma słonecznego, a więc odcieni o różnej jasności i nasyceniu. Fakty te w każdym razie świadczą o tym, że

11) Por. R. Carnap, *Der Logische Aufbau der Welt*, Berlin 1928.

pojęcie podobieństwa pod względem barwy, które ma się tutaj na uwadze, jest pojęciem niezmiernie nieostrym. Predykat Pb jest pod tym względem w sytuacji gorszej nie tylko od predykatu Nb, ale również od predykatu Ż. Nie bardzo, co za tym idzie, nadaje się do roli terminu wyjściowego przy definiowaniu tego ostatniego.

Znacznie ostrzejsze od owego «absolutnego» pojęcia podobieństwa pod względem barwy jest odpowiednie pojęcie «porównawcze». Wyrażamy je w sposób najprostszy za pomocą trójargumentowego predykatu pierwszego rzędu stwierdzającego, iż przedmiot a jest bardziej podobny pod względem barwy do przedmiotu b niż do przedmiotu c — w skrócie: Pbb (a, b, c). Można, jak się zdaje, założyć, że Pbb jest terminem spostrzeżeniowym, zinterpretowanym, podobnie jak Nb czy Pb, bezpośrednio przez wskazanie w charakterze wzorców pozytywnych i negatywnych odpowiednich «trójek» przedmiotów. Stosują się też do takiej interpretacji uwagi wypowiedziane pod adresem tamtych predykatów — z tym, że predykat Pbb jest pod tym względem w sytuacji pośredniej pomiędzy predykatem Nb a predykatem Pb. Wydaje się, iż na ogół można bez większych trudności rozstrzygnąć drogą spostrzeżeniową, które spośród trzech wskazanych przedmiotów są do siebie bardziej podobne pod względem barwy niż pozostałe. Predykat Pbb stanowi wygodną podstawę zarówno do definicji poszczególnych barw, jak i do ich uporządkowania.

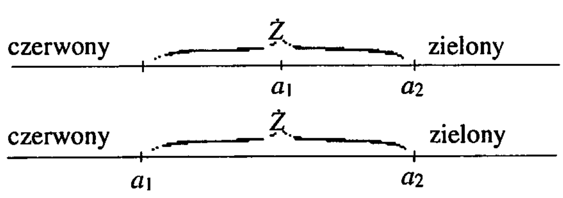
Interesującą nas definicję predykatu Ż można sformułować przy pomocy predykatu Pbb na wiele różnych sposobów. I tak, definicja:

(32) $$\bigwedge_x (x \in \dot{Z} \equiv \mathrm{Pbb}(a_1, x, a_2)),$$

odwołuje się do dwóch przedmiotów wzorcowych a_1, a_2 i zalicza do klasy Ż jakiś przedmiot wtedy, gdy wzorzec a_1 podobny jest do niego pod względem barwy bardziej niż do przedmiotu a_2. Natomiast definicja:

(33) $$\bigwedge_x (x \in \dot{Z} \equiv \mathrm{Pbb}(a_1, x, a_2) \wedge \mathrm{Pbb}(a_2, x, a_1)),$$

odwołująca się również do dwóch przedmiotów wzorcowych, utożsamia Ż z klasą tych przedmiotów, do których każdy z wzorców podobny jest pod względem barwy bardziej niż do wzorca pozostałego. Przyjmując dla uproszczenia liniowe uporządkowanie odcieni barwnych (na przykład odcieni widma słonecznego), różnicę powyższą unaocznić możemy za pomocą następujących wykresów



z których pierwszy odpowiada definicji (32), drugi — definicji (33). Wypowiedzi te nie wyczerpują wszystkich form, jakie przybierać może definicja predykatu Ż za pomocą predykatu Pbb. Poprzestaniemy jednak tutaj na tych przykładach, gdyż rozważanie pozostałych możliwości nie wnosi poza komplikacjami formalnymi niczego nowego.

8. Omówiliśmy powyżej procedury definiowania predykatu Ż odwołujące się kolejno do predykatów Nb, Pb i Pbb. Jedna jeszcze wersja takiej procedury wymaga krótkiego chociażby omówienia. Mam na myśli procedurę przedstawioną w cytowanej już przez nas pracy J. Kotarbińskiej. Nie sposób przytaczać tutaj jej szczegółowej charakterystyki. Ograniczę się zatem do przedstawienia tylko tego, co bezpośrednio się wiąże z omawianymi przez nas sprawami, i to przedstawienia odbiegającego od sformułowań oryginalnych, ale ułatwiającego za to porównanie z procedurami poprzednimi.

Punktem wyjścia jest tutaj, podobnie jak to miało miejsce poprzednio, wypowiedź (6). Wypowiedź ta zawiera, jak pamiętamy, zagadkowy zwrot „taki, jak", dopuszczający różne możliwości interpretacyjne. Interpretacja, którą nadaje mu się obecnie, polega na traktowaniu go jako równoznacznego z wyrażeniem „podobny pod takim a takim względem i w takim a takim stopniu", a całość procedury definicyjnej wyznaczać ma, o jaki wzgląd i stopień chodzi. Tak rozumiana procedura definicyjna odwołuje się zatem jako do terminu wyjściowego do czteroargumentowego predykatu: *a* jest podobny do *b* pod względem *W* i w stopniu (co najmniej) *S*; w skrócie: P (a, b, W, S). Definicję terminu Ż za pomocą predykatu P przedstawić można w sposób następujący. Wypowiedź:

(34) $$\bigwedge_x (x \in \dot{Z} \equiv \mathrm{P}(x, a_1, W_1, S_1)),$$

określa Ż jako klasę tych przedmiotów, które są podobne do przedmiotu wzorcowego a_1 pod względem W_1 i w stopniu S_1, a układ postulatów:

(35) $$\begin{array}{ll} \mathrm{P}(a_2, a_1, W_1, S_1) & \sim \mathrm{P}(b_1, a_1, W_1, S_1) \\ \ldots & \ldots \\ \mathrm{P}(a_n, a_1, W_1, S_1) & \sim \mathrm{P}(b_m, a_1, W_1, S_1), \end{array}$$

wyznacza ów wzgląd W_1 i stopień S_1 przez wskazanie takich przedmiotów wzorcowych $a_2, \ldots, a_n$, które są podobne do przedmiotu a_1 pod względem W_1 i w stopniu S_1, i takich $b_1, \ldots, b_m$, które tego warunku nie spełniają[12].

Procedura powyższa wyznacza — jak to w cytowanej pracy zostało przekonująco pokazane — W_1 i S_1, a tym samym i Ż, w sposób o wiele ściślejszy niż rozważane poprzednio układy postulatów typu (3). Tak jest jednak tylko wtedy, gdy predykat P spełnia określone warunki. Predykat ten musi być w każdym razie rozumiany tak, aby gwarantował zachodzenie następujących zależności:

(36) P $(x, y, W, S) \wedge y$ jest nieodróżnialne pod względem W od $z \rightarrow \mathrm{P}(x, z, W, S)$

12) Wypowiedziom (34) i (35) odpowiada w cytowanej pracy tzw. założenie definicyjne:

$$\bigvee_{W,S} [\bigwedge_x (x \in \dot{Z} \equiv \mathrm{P}(x, a_1, W, S)) \wedge \mathrm{P}(a_2, a_1, W, S) \wedge \ldots$$
$$\wedge\, \mathrm{P}(a_n, a_1, W, S) \wedge \sim\mathrm{P}(b_1, a_1, W, S) \wedge \ldots \wedge \sim\mathrm{P}(b_m, a_1, W, S)].$$

Ściśle biorąc, założenie to przybiera tam postać różniącą się nieco od powyższej:

$$\bigvee_{W,S} [\bigwedge_x (x \in \dot{Z} \equiv \mathrm{P}(x, a_1, W, S) \wedge \ldots \wedge \mathrm{P}(x, a_n, W, S) \wedge$$
$$\wedge \sim\mathrm{P}(x, b_1, W, S) \wedge \ldots \wedge \sim\mathrm{P}(x, b_m, W, S))].$$

Wydaje się jednak, iż sformułowanie poprzednie, zachowując istotę omawianej propozycji, unika pewnych zarzutów, na które narażone jest sformułowanie oryginalne

(37) $\sim P(x, y, W, S) \wedge y$ jest nieodróżnialne pod względem W od $z \rightarrow \sim P(x, z, W, S)$

(38) $P(x, y, W, S) \wedge x$ jest bardziej podobne pod względem W do z niż do $y \rightarrow$
$\rightarrow P(x, z, W, S)$

(39) $\sim P(x, y, W, S) \wedge x$ jest bardziej podobne pod względem W do y niż do $z \rightarrow$
$\rightarrow \sim P(x, z, W, S)$
(dla wszelkich x, y, z, W i S).

W stosunku do tak rozumianego predykatu P powstaje pytanie, które dla naszych rozważań ma znaczenie decydujące: czy P jest terminem logicznym, czy też terminem deskryptywnym? Pytanie to trudno rozstrzygnąć w sposób bezapelacyjny. Wszystko jednak wydaje się przemawiać za deskryptywnym charakterem tego predykatu. W każdym razie wszelkie próby zdefiniowania predykatu P wyłącznie za pomocą terminów logicznych okazują się bezskuteczne. Przy tym próby te okazują, iż intuicyjny sens predykatu P nie jest całkowicie jasny. Co mają, w szczególności, reprezentować dwa spośród jego argumentów: W i S? «Wzgląd» W utożsamić można chyba z pewną rodziną własności; występujący w definicji predykatu Ż wzgląd W_1 — to po prostu barwa B w którymś z wyróżnionych przez nas znaczeń. A «stopień» S? Tutaj trudniej znaleźć zadowalającą odpowiedź. Nie bardzo wiadomo w szczególności, jak by miały wyglądać konkretne podstawienia tej zmiennej, na przykład stopień S_1 z definicji predykatu Ż. W pewnych przypadkach specjalnych sprawa staje się wyraźniejsza. Tak na przykład, gdy wzgląd W jest pewną wielkością, stopień S może być wyrażony liczbowo[13]. Interpretacja taka nie daje się jednak zastosować w przypadku ogólnym; zwłaszcza nie nadaje się do omawianych przez nas przypadków definiowania predykatów typu Ż. Nasuwają się co prawda i tutaj pewne interpretacje, zarówno owego S, jak i całego wyrażenia $P(a, b, W, S)$, ale żadna z tych interpretacji, sformułowana wyłącznie w terminach logicznych, nie wyposaża predykatu P w żądany sens[14].

Uzasadnione wydaje się w tej sytuacji przypuszczenie, iż predykat P spełniający wyszczególnione postulaty jest predykatem deskryptywnym. A skoro tak, to omawiana obecnie procedura jest w sytuacji takiej samej, jak wszystkie procedury poprzednie: jest pośrednim sposobem interpretacji terminu Ż, odwołującym się do innych terminów

---

13) W takim przypadku można bez większych trudności określić sens wyrażenia $P(a, b, W, S)$. Niech $W(a)$, $W(b)$ reprezentują wielkości przedmiotów a, b, należące do rodzaju W. Zwrot $P(a, b, W, S)$ możemy uważać w tej sytuacji za skrót wyrażenia $|W(a)—W(b)| \leq S$, gdzie S jest po prostu zmienną liczbową.

14) Nasuwa się np. interpretacja następująca. S — to po prostu pewna podklasa klasy W, a więc klasa zawierająca niektóre spośród własności należących do W. Dwa przedmioty są do siebie podobne pod względem W i w stopniu S, jeśli obu tym przedmiotom przysługują wszystkie własności należące do klasy S:

$$P(a, b, W, S) \equiv \bigwedge_{Y \in S} (Y \in W \wedge a \in Y \wedge b \in Y).$$

Ta dość intuicyjna interpretacja przedstawia pewne pośrednie ogniwo między dwiema skrajnymi interpretacjami zwrotu „taki, jak", wyróżnionymi przez nas poprzednio. Nie nadaje ona jednak, jak się łatwo przekonać, predykatowi P sensu takiego, który by pociągał za sobą sformułowane wyżej zależności (38) i (39).

deskryptywnych. W porównaniu z tamtymi procedurami ma ona zarówno pewne zalety, jak i pewne wady, płynące ze swoistego charakteru terminu P. Jest to, z jednej strony, termin na tyle ogólny, że nadaje się do ostensywnego definiowania nie tylko własności barwnych, ale i innych własności «spostrzegalnych», co różni go korzystnie od terminów Nb, Pb i Pbb. Z drugiej strony jednakże, jest to, z tego właśnie względu, termin o dość nieokreślonym, jak widzieliśmy, znaczeniu. Przy tym sposób ustalenia tego znaczenia pozostaje niewyjaśniony. Nie jest to w każdym razie termin interpretowany bezpośrednio, tak jak Nb, Pb czy Pbb, choćby z tego powodu, iż pewne jego argumenty — to nie przedmioty konkretne, lecz abstrakcyjne, których ani spostrzec, ani pokazać nie można. A zatem musi to być termin interpretowany pośrednio, przez powołanie się na inne terminy deskryptywne. Jakie terminy mogą tu wchodzić w grę? Czy rozumienie predykatu P nie zakłada na przykład rozumienia predykatów stanowiących pewne jego konkretyzacje, tj. predykatów typu P (a, b, W_1, S_1), P (a, b, W_2, S_2) itp.? (Jednym z nich byłby nasz predykat Pb!) Nie umiem znaleźć odpowiedzi na te pytania. A dopiero odpowiedź na nie mogłaby pozwolić na właściwą ocenę przedstawionej procedury z interesującego nas tutaj punktu widzenia.

*

Dobiegliśmy w ten sposób do końca naszych rozważań. Jakiż jest ich rezultat? W stosunku do naczelnego zagadnienia interpretacji bezpośredniej — wyraźnie negatywny, zwłaszcza jeśli idzie o terminy o charakterze predykatów. Interpretacja tych terminów nie odwołująca się do innych predykatów deskryptywnych nie przyporządkowuje im denotacji w sposób choćby w przybliżeniu jednoznaczny i adekwatny. Jeśli zatem terminy te uzyskują taką drogą interpretację zamierzoną, odpowiadającą tej, jaka przysługuje im w języku potocznym, to dzieje się tak nie dlatego, że gwarantują to wchodzące w grę zależności logiczne. Jak więc do tego dochodzi? Zagadnienie to wykracza poza zadania obecnej pracy. Toteż ograniczę się tylko do paru ogólnikowych sugestii. Pierwsze, co się nasuwa, to spostrzeżenie, iż w faktycznej praktyce ustalania sensu terminów takich, jak Ż czy Nb, odwołujemy się do bardzo wielkiej liczby przedmiotów wzorcowych. Każde zastosowanie takiego predykatu lub jego negacji stanowi z reguły wzbogacenie repertuaru przedmiotów wzorcowych, co przyczynia się do jego precyzacji. Po drugie wydaje się, iż interpretacja takich terminów nie ma charakteru jednorodnego. Interpretuje się je początkowo w sposób bezpośredni, i to nie tylko predykaty Nb, Pb czy Pbb, ale i predykaty takie, jak Ż, czerwony, zielony itp., następnie zaś ustala się między nimi związki definicyjne, które wpływają na dalszą precyzację tych terminów. Mogą to być na przykład definicje w rodzaju przytaczanej definicji predykatu Ż za pomocą predykatu Nb lub pewne postulaty luźniejsze, w rodzaju postulatu głoszącego, iż żaden przedmiot, który jest Ż, nie jest zarazem czerwony, zielony itp. Wszystko to jednak jest na pewno niewystarczające do nadania tym terminom sensu takiego, jaki przysługuje im istotnie w języku potocznym. Dołącza się tutaj chyba zawsze coś, co nazwać można „aktem abstrakcji" i co wymyka się próbom przedstawienia w postaci operacji logicznej, jaką jest na przykład definicja. Na podstawie spostrzeżenia szeregu przedmiotów wzorcowych uprzytamniamy sobie jakąś ową własność (czy stosunek), o którą chodzi w przypadku definiowanego predykatu, mimo iż nie mamy żadnych racji logicznych do tego, aby tę właśnie własność (czy stosunek) spośród niezliczonej liczby innych możliwości wyróżnić. Fakt ten wyrazić możemy stwierdzając, iż bezpośredni sposób interpretacji predykatów deskryptywnych nie stanowi procedury typu definicyjnego.

Na koniec — pewne uwagi terminologiczne. Wspominałem na wstępie o praktyce utożsamiania interpretacji bezpośredniej z definicją ostensywną, a terminów tak zinterpretowanych — z terminami spostrzeżeniowymi. Praktyka ta budzić może pewne wątpliwości. Nie idzie w tej chwili o to, iż przyznaje się tu miano „definicji" czemuś, co, jak przed chwilą stwierdziliśmy, definicją nie jest. Termin „definicja ostensywna" można przecież traktować tak, jak traktuje się wyrażenie „kula spłaszczona". Tak też na ogół traktowaliśmy go na tych stronach. Idzie raczej o to, iż utożsamiając definicję ostensywną z interpretacją bezpośrednią, ujmujemy tę procedurę w sposób dość wąski. W trakcie naszych rozważań doszliśmy do konkluzji, iż interpretacja predykatu Ż może przybierać formę definicji odwołującej się do predykatów takich, jak Nb, Pb czy Pbb. Definicja taka reprezentuje pośredni sposób interpretacji predykatu Ż. Czy mamy jej z tej racji odmówić miana definicji ostensywnej? Moglibyśmy jej to miano przyznać, gdybyśmy definicję ostensywną pojęli w sposób znacznie szerszy. Przy takim rozumieniu o ostensywnym charakterze procedury definicyjnej decydowałoby to, czy odwołujemy się w niej do wskazywania i spostrzegania pewnych przedmiotów wzorcowych. A ten warunek spełniają wszystkie uwzględnione przez nas definicje predykatu Ż. Podobne uwagi stosują się do pojęcia terminu spostrzeżeniowego. Utożsamiając taki termin z terminem zinterpretowanym bezpośrednio uniemożliwiamy podciągnięcie pod to pojęcie predykatu takiego, jak Ż, co wydaje się niezgodne z powszechnie przyjętą praktyką. Ale tutaj nie nasuwa się tak proste rozszerzenie tego pojęcia, jak w przypadku definicji ostensywnej. W szczególności nie można traktować jako terminu spostrzeżeniowego wszelkiego terminu wprowadzonego za pomocą tak szeroko rozumianej definicji ostensywnej. Trzeba by było wówczas uznać za termin spostrzeżeniowy termin taki, jak „masa 1 kg", wprowadzony za pomocą następującej definicji ostensywnej:

(40) $\bigwedge_x$ (x ma masę 1 kg $\equiv$ x jest równe pod względem masy temu oto przedmiotowi).

Konsekwencja jaskrawo niezgodna z obiegowym pojęciem terminu spostrzeżeniowego. Nie widzę takiego sposobu określenia tego pojęcia, które by istotnie chwytało wszelkie związane z nim intuicje. Pewne tego powody starałem się podać w toku naszych rozważań. Nie wydaje mi się zresztą ta sprawa — dotycząca ostatecznie pewnych terminologicznych konwencji — problemem doniosłym. Wpływa ona jednak na kwestie takie, jak wybór tytułu niniejszej rozprawy, który w świetle tych wywodów okazuje się tytułem nieadekwatnym.

Filozofia Nauki
Rok I, 1993, Nr 2-3

W sprawie obserwacyjnej definiowalności terminów teoretycznych

Jednym z klasycznych problemów empirystycznej filozofii nauki jest problem definiowalności terminów teoretycznych przez terminy obserwacyjne. Szczególnego znaczenia problem ten nabiera na gruncie założeń neopozytywistycznych. Nic też dziwnego, że od samego powstania tego kierunku znajdował się w centrum zainteresowań jego przedstawicieli. Ale choć poświęcono mu sporo uwagi i wysiłku, wysunięto wiele sugestii i propozycji, trudno uznać go za problem definitywnie rozwiązany. W znanym artykule „The Theoretician's Dilemma" z r. 1958, zawierającym rodzaj rekapitulacji dotychczasowych badań nad tym zagadnieniem, Hempel podsumowuje przegląd omawianych wyników sceptycznym wnioskiem stwierdzającym brak konkluzywnych argumentów przemawiających za lub przeciw definiowalności terminów teoretycznych nauk empirycznych za pomocą terminów czysto obserwacyjnych. Co więcej, wyraża wątpliwość, czy jakikolwiek argument może tę kwestię przesądzić w sposób ostateczny[1]. W ciągu piętnastu lat, jakie upłynęły od tego stwierdzenia, sytuacja nie zmieniła się w sposób zasadniczy. W opublikowanym w r. 1973 artykule „On the Different Ingredients of an Empirical Theory" Hintikka powtarza niemal dosłownie konkluzję Hempla. „Filozofowie nauki dyskutowali często nad tym, czy terminy teoretyczne mogą być zdefiniowane przez pojęcia obserwacyjne. Wydaje się, że istnieje zgoda co do tego, że nie mogą, choć trudno znaleźć przekonujące racje, dlaczego by tak miało być"[2]. Diagnoza ta wydaje się trafna. Wyniki zawarte w licznych pracach poświęconych problemo-

1) *Minnesota Studies in the Philosophy of Science*, vol. II, 1958, s. 70.
2) *Logic, Methodology and Philosophy of Science IV, 1973, s. 319*

wi obserwacyjnej definiowalności terminów teoretycznych mają z reguły charakter rozwiązań cząstkowych. W sposób konkluzywny okazują jedynie, że pewne definicje obserwacyjne (ewentualnie — definicje obserwacyjne pewnych typów) nie nadają się na definicje określonych terminów teoretycznych (ewentualnie — terminów teoretycznych określonych typów)[3]. Brak jednak argumentów okazujących w sposób przekonujący, że terminów tych nie można zdefiniować za pomocą żadnych definicji obserwacyjnych — a tym bardziej, że nie można w ten sposób zdefiniować żadnych terminów teoretycznych. Jeśli się wysuwa tego rodzaju przypuszczenia, to popiera je się wywodami mającymi świadczyć nie tyle o tym, że takie procedury definicyjne są niemożliwe, ile raczej o tym, że są to procedury nie wskazane z pewnych względów pragmatycznych[4]. Z drugiej strony, przytacza się niekiedy przykłady terminów teoretycznych zdefiniowanych *explicite* przez terminy obserwacyjne[5]. Jeśli są to przykłady trafne, ogólna teza o obserwacyjnej niedefiniowalności terminów teoretycznych utrzymać się nie daje.

Problem jest więc nadal otwarty. Czy istnieje jakakolwiek możliwość rozstrzygnięcia go w sposób ogólny? Czy można w szczególności podać konkluzywne argumenty przemawiające przeciw obserwacyjnej definiowalności terminów teoretycznych — jeśli nie wszystkich, to przynajmniej niektórych wyróżnionych ich rodzajów? Jedno jest niewątpliwe. Niezbędnym warunkiem sformułowania takich argumentów jest uprzednia precyzacja samego problemu. To, czy terminy teoretyczne są, czy też nie są definiowalne przez terminy obserwacyjne, zależy w sposób oczywisty od tego, co przez pojęcia terminu obserwacyjnego i teoretycznego rozumiemy. Tylko określona, dostatecznie wyraźna interpretacja tych wieloznacznych i niejasnych pojęć może pozwolić na uchwycenie ich wzajemnych związków. Taka też interpretacja leży u podstawy przedstawionej niżej argumentacji. Wskazuję w niej pewne konsekwencje płynące na gruncie owej interpretacji z założenia definiowalności terminów teoretycznych przez terminy obserwacyjne — konsekwencje przemawiające, w moim przekonaniu, przeciwko takiemu założeniu. Dostarczają one tym samym częściowego przynajmniej uzasadnienia dla owego dominującego w literaturze przedmiotu poglądu o obserwacyjnej niedefiniowalności terminów teoretycznych. Argumenty poniższe odwołują się do pewnych środków formalnych i zakładają, co za tym idzie, określoną formalizację rozważanego problemu. Jest to formalizacja dokonana za pomocą pojęć teroriomodelowych i umożliwiająca dzięki temu wykorzystanie pewnych ogólnych teoriomodelowych zależności. Rozważania te stanowią tym samym przykład zastosowania pewnych metod formalnych do rozstrzygania problemów filozoficznych.

3) Taki charakter ma np. słynny wywód Carnapa w „Testability and Meaning", *Philosophy of Science*, 1936-1937.
4) Taki rodzaj argumentacji zawiera m. in. *Scientific Explanation* Braithwaite'a, 1953.
5) Przykład taki podaje np. R. Wójcicki w publikacji zbiorowej *Teoria i doświadczenie*, 1966, s. 67

Eksplikacja pojęć terminu obserwacyjnego i teoretycznego, jaką tu zakładam, przedstawiona została szerzej w innych moich publikacjach, w szczególności w monografii *The Logic of Empirical Theories*. Tutaj ograniczę się do podkreślenia tych tylko jej rysów, które mają charakter decydujący dla dyskutowanej przez nas sprawy. Można je ująć najkrócej, jak następuje:

Terminy obserwacyjne to terminy wyposażone w określoną interpretację wyłącznie w dziedzinie przedmiotów spostrzegalnych, a *terminy teoretyczne* to terminy odnoszące się (między innymi lub wyłącznie) do przedmiotów niespostrzegalnych.

Charakterystyka ta wskazuje zarazem źródło niedefiniowalności terminów teoretycznych przez terminy obserwacyjne. Jakie intuicje kryją się za powyższymi założeniami? Stwierdzić trzeba przede wszystkim wyraźnie, że przyjęte tu rozumienie terminów obserwacyjnych reprezentuje tylko jedną z możliwych eksplikacji tego pojęcia: eksplikację rygorystyczną i wąską, ale za to dostatecznie wyraźną i wyróżniającą, jak sądzę, ważną metodologicznie kategorię wyrażeń. Terminy obserwacyjne tak rozumiane utożsamić można z terminami zinterpretowanymi w sposób bezpośredni. Najprostszym rodzajem takiej interpretacji jest tzw. *definicja ostensywna*. A terminy zdefiniowane ostensywnie to terminy, których jedynym kryterium stosowalności jest wygląd danego przedmiotu (szeroko rozumiany, a więc nie tylko wizualny). Kryterium takie z natury rzeczy ograniczone jest do przedmiotów bezpośrednio spostrzegalnych (a więc pewnych makroobiektów). W stosunku do przedmiotów bezpośrednio niespostrzegalnych (w szczególności wszelkich mikroobiektów) terminy zdefiniowane ostensywnie pozbawione są jakichkolwiek kryteriów stosowalności. W rezultacie tylko w dziedzinie przedmiotów spostrzegalnych interpretację terminów obserwacyjnych można uważać za ustaloną. We wszelkiej dziedzinie przedmiotów niespostrzegalnych terminy te pozostają nie określone; mogą zatem być interpretowane dowolnie.

Formalnym odpowiednikiem tych założeń może być następująca konstrukcja teoriomodelowa. Językiem obserwacyjnym J_o niech będzie język rachunku predykatów z identycznością, zawierający prócz stałych logicznych (spójników zdaniowych, kwantyfikatorów i predykatu identyczności) oraz zmiennych indywiduowych, predykaty obserwacyjne: $o_1, ..., o_n$. Niech U_o będzie zbiorem przedmiotów spostrzegalnych. W zbiorze tym bezpośrednia interpretacja predykatów $o_1, ..., o_n$ przyporządkowuje im jako denotacje określone relacje $O_1, ..., O_n$, wyznaczając tym samym model właściwy języka J_o, $\mathfrak{m}_o = \langle U_o, O_1, ..., O_n \rangle$. Jest to założenie upraszczające stan faktyczny, gdyż w rzeczywistości nawet w zbiorze przedmiotów spostrzegalnych interpretacja predykatów obserwacyjnych nie bywa jednoznaczna. Wyznacza ona w rezultacie nie jeden model właściwy języka obserwacyjnego, lecz pewną ich rodzinę. Przyjmujemy tu jednak owo założenie upraszczające, ponieważ przy założeniach bardziej realistycznych przedstawiona niżej argumentacja zachowuje walor *a fortiori*.

W sytuacji, gdy język obserwacyjny występuje jako fragment języka teoretycznego, jego interpretacja nie może być ograniczona do zbioru przedmiotów spostrzegalnych. Typowe teorie empiryczne (np. teorie fizyczne) mówią z reguły o pewnych przedmiotach niespostrzegalnych (atomach, elektronach itp.) i takie też przedmioty wchodzić muszą w skład universum języka obserwacyjnego stanowiącego część języka teorii. Ale w zbiorze przedmiotów niespostrzegalnych interpretacja predykatów obserwacyjnych pozostaje, jak widzieliśmy, nie określona. Fakt ten znajduje wyraz w założeniu głoszącym, iż modelem właściwym języka J_o (traktowanego jako część języka teoretycznego) jest dowolne rozszerzenie modelu $\mathfrak{m}_o$, a więc wszelki model $\mathfrak{m}'_o = \langle U'_o, O'_1, \ldots, O'_n \rangle$, w którym uniwersum U'_o zawiera uniwersum U_o: $U'_o \supset U_o$, a relacje $O'_1, \ldots, O'_n$ ograniczone do universum U_o pokrywają się z relacjami $O_1, \ldots, O_n$: $O'_i | U_o = O_i$ dla $i = 1, \ldots, n$. Rodzinę takich modeli oznaczymy przez M_o.

Terminy teoretyczne ograniczyliśmy wyżej do terminów desygnujących (między innymi przynajmniej) pewne przedmioty niespostrzegalne. Jest to ograniczenie istotne dla dalszych rozważań. Wyłącza ono z kategorii terminów teoretycznych pewne terminy tradycyjnie do niej zaliczane, takie chociażby, jak termin *inteligentny*. Typowe jednak terminy teoretyczne występujące w teoriach empirycznych (m. in. wszystkie terminy fizyczne) warunek powyższy spełniają. Możemy go traktować jako warunek wyróżniający kategorię terminów teoretycznych *sensu stricto*. W przeciwieństwie do terminów obserwacyjnych, terminy teoretyczne najszerzej nawet rozumiane mogą być interpretowane jedynie pośrednio. Takim pośrednim sposobem interpretacji danego terminu teoretycznego jest charakterystyka jego denotacji za pomocą zinterpretowanych uprzednio terminów obserwacyjnych. Załóżmy, iż opisany poprzednio język obserwacyjny J_o wzbogacamy o pewien predykat teoretyczny t, przechodząc w ten sposób do języka teoretycznego J o predykatach $o_1, \ldots, o_n, t$. Interpretacja pośrednia predykatu t polega na przyjęciu pewnego zdania δ języka J (zawierającego predykat t i niektóre przynajmniej z predykatów $o_1, \ldots, o_n$) w roli postulatu charakteryzującego denotację predykatu t jako relację spełniającą warunki sformułowane w zdaniu δ przy założeniu, iż predykaty $o_1, \ldots, o_n$ interpretowane są w sposób wyznaczony przez rodzinę modeli M_o. Innymi słowy, modelem właściwym tak zinterpretowanego języka J jest każde przedłużenie któregoś z modeli należących do rodziny M_o, będące modelem zdania δ, a więc, każdy model $\mathfrak{m} = \langle U, O_1, \ldots, O_n, T \rangle$, którego fragment odpowiadający językowi J_o, $\mathfrak{m}|_o = \langle U, O_1, \ldots, O_n \rangle$, pokrywa się z pewnym modelem rodziny M_o, i w którym prawdziwe jest zdanie δ. Zdanie δ nosi nazwę *postulatu znaczeniowego* dla predykatu t. W skrajnym przypadku postulat taki może mieć postać definicji równoważnościowej predykatu t za pomocą predykatów $o_1, \ldots, o_n$. Ten właśnie przypadek będzie przedmiotem dalszych rozważań.

Załóżmy dla uproszczenia, że predykat t jest predykatem jednoargumentowym (uogólnienie na predykaty k-argumentowe nie przedstawia większych trudności). Definicja δ przybiera w tym wypadku postać:

$$t(x) \equiv \alpha(x), \tag{1}$$

gdzie $\alpha(x)$ jest formułą zdaniową języka obserwacyjnego J_o o jednej zmiennej wolnej. Rozważany przez nas problem definiowalności predykatu teoretycznego t w języku obserwacyjnym J_0 zakłada pewną zamierzoną interpretację predykatu t i sprowadza się do pytania, czy istnieje taka obserwacyjna formuła $\alpha(x)$, której interpretacja właściwa pokrywa się z ową zamierzoną interpretacją predykatu t. Ta ostatnia nie musi być przy tym interpretacją jednoznaczną. Może przyporządkowywać predykatowi t jako jego denotację nie jeden zbiór przedmiotów, lecz rodzinę takich zbiorów. Idzie wówczas o to, czy ta sama rodzina zbiorów przyporządkowana być może pewnej obserwacyjnej formule $\alpha(x)$ zgodnie z jej właściwą interpretacją.

To, jaka jest zamierzona interpretacja danego predykatu teoretycznego, zależy oczywiście od tego, co to jest za predykat. Wydaje się jednak, że istnieją pewne warunki niezbędne, które spełniać musi interpretacja każdego w zasadzie predykatu teoretycznego. Interpretacja nie spełniająca któregoś z tych warunków nie może być uznana za interpretację adekwatną. Otóż argumentacja nasza zmierza do okazania, że żadna definicja obserwacyjna typu (1) nie może wyposażyć predykatu teoretycznego t w interpretację adekwatną. Dotyczy to wszystkich predykatów teoretycznych *sensu stricto*, a więc predykatów, które z zamierzenia odnosić się mają do pewnych przedmiotów niespostrzegalnych, tj. nie należących do universum U_0. Okazuje się mianowicie, iż interpretacja takiego predykatu określona przez definicję (1) nie jest w stanie wyznaczyć żadnego ustalonego podziału owych przedmiotów niespostrzegalnych na przedmioty desygnowane przez predykat t i pozostałe. Dopuszcza zawsze traktowanie pod tym względem wszystkich przedmiotów niespostrzegalnych *en bloc*.

Fakt ten jest konsekwencją pewnych ogólnych zależności teoriomodelowych. Bezpośrednie zastosowanie znajduje tu twierdzenie następujące (szkic jego dowodu podaję w załączonym Dodatku):

(T1) *Niech $\alpha(x)$ będzie formułą języka J_0, a $\mathfrak{m}_0$ modelem tego języka o universum U_0. Jeżeli pewien element universum U_0 spełnia formułę $\alpha(x)$ w każdym rozszerzeniu modelu $\mathfrak{m}_0$, to dla dowolnego zbioru U zawierającego zbiór U_0 istnieje rozszerzenie $\mathfrak{m}$ modelu $\mathfrak{m}_0$ o universum U takie, iż każdy element universum U nie należący do U_0 spełnia formułę $\alpha(x)$ w modelu $\mathfrak{m}$.*

Innym nieco sformułowaniem twierdzenia (T1) jest jego wersja następująca:

(T2) *Jeżeli pewien element universum U_0 nie spełnia formuły $\alpha(x)$ w żadnym rozszerzeniu modelu $\mathfrak{m}_0$, to dla dowolnego zbioru U zawierającego zbiór U_0 istnieje rozszerzenie $\mathfrak{m}$ modelu $\mathfrak{m}_0$ o universum U takie, iż żaden element universum U nie należący do U_0 nie spełnia formuły $\alpha(x)$ w modelu $\mathfrak{m}$.*

Przyjrzyjmy się konsekwencjom tych twierdzeń dla rozważanej przez nas sytuacji. Co mówią nam one o interpretacji predykatu teoretycznego t wyznaczonej przez definicję (1)? Przypuśćmy, że jest to predykat, który prócz pewnych przedmiotów niespostrzegalnych desygnować ma również pewne przedmioty spostrzegalne. Definiująca go formuła $\alpha(x)$ spełniać musi wówczas poprzednik twierdzenia (T1). Tym samym jedną z jej dopuszczalnych interpretacji okazuje się, na mocy tego twierdzenia, interpretacja taka, przy której formułę tę spełnia każdy przedmiot niespostrzegalny — niezależnie od tego, jaki zbiór przedmiotów niespostrzegalnych weźmiemy pod uwagę.

W konsekwencji, jedną z dopuszczalnych interpretacji predykatu t okazuje się zawsze interpretacja zaliczająca do jego desygnatów wszelkie przedmioty niespostrzegalne. A to wydaje się z reguły niezgodne z jego zamierzonym sensem. Z drugiej strony, jeśli istnieją przedmioty spostrzegalne, których predykat t nie desygnuje, jedną z jego dopuszczalnych interpretacji będzie, na mocy twierdzenia (T2), interpretacja wykluczająca spośród jego desygnatów wszelkie przedmioty niespostrzegalne — niezależnie od tego, jakie przedmioty tego typu obejmuje universum danego języka. To również wydaje się świadczyć o nieadekwatności takiej interpretacji. Tak więc, w przypadku predykatów teoretycznych, desygnujących prócz przedmiotów niespostrzegalnych niektóre (i tylko niektóre) przedmioty spostrzegalne, mamy do czynienia z sytuacją następującą. Predykat taki wyposażony zostaje przez definicję obserwacyjną w interpretację niezmiernie wieloznaczną. Wśród jego denotacji znajdujemy, z jednej strony, zbiór nie obejmujący żadnych przedmiotów niespostrzegalnych, z drugiej strony, zbiór obejmujący wszystkie takie przedmioty. Sądzę, że żadnej z tych denotacji nie można uznać za należącą do denotacji zamierzonych.

W przypadku predykatów teoretycznych desygnujących wyłącznie przedmioty niespostrzegalne, ich definicje obserwacyjne pociągają konsekwencje bardziej jeszcze paradoksalne. Bezpośrednie zastosowanie znajduje tu zależność będąca bezpośrednim wnioskiem z twierdzenia (T2):

(T3) *Jeżeli żaden element universum U_0 nie spełnia formuły $\alpha(x)$ w żadnym rozszerzeniu modelu $\mathfrak{m}_0$, to dla dowolnego zbioru U zawierającego zbiór U_0 istnieje rozszerzenie $\mathfrak{m}$ modelu $\mathfrak{m}_0$ o universum U takie, iż żaden element universum U nie spełnia formuły $\alpha(x)$ w modelu $\mathfrak{m}$.*

Poprzednik tego twierdzenia zachodzić musi dla każdej formuły $\alpha(x)$ definiującej predykat teoretyczny omawianego typu. W konsekwencji, formuła ta dopuszcza zawsze interpretację taką, przy której nie spełnia jej żaden w ogóle element universum — niezależnie od tego, jakie universum dany język zakłada. Tak więc, jedną z denotacji predykatu tak zdefiniowanego okazuje się zawsze zbiór pusty. Interpretacja, zgodnie z którą dany predykat teoretyczny nie desygnuje niczego, nie jest na pewno jego interpretacją zamierzoną. A takiej interpretacji żadna definicja obserwacyjna wykluczyć nie jest w stanie.

Fakty te przemawiają, w moim przekonaniu, za niedefiniowalnością terminów teoretycznych przez terminy obserwacyjne, rozstrzygając w ten sposób ów tradycyjny problem. Trzeba jednak podkreślić raz jeszcze, że wynik ten jest konsekwencją ściśle określonych założeń. Opiera się w szczególności na swoistym, dość rygorystycznym rozumieniu obu podstawowych pojęć: terminu obserwacyjnego i teoretycznego. To dlatego tylko, że interpretacja terminów obserwacyjnych jest w stosunku do przedmiotów niespostrzegalnych całkowicie nie określona, a terminy teoretyczne takie właśnie przedmioty mają desygnować — nie sposób za pomocą tych pierwszych scharakteryzować wystarczająco ściśle interpretacji tych drugich. Przy innym, bardziej liberalnym rozumieniu obu pojęć można dojść do konkluzji odmiennej. To tłumaczy fakt powoływania się na przykłady definiowalności niektórych terminów teoretycznych przez terminy obserwacyjne. Przyjrzyjmy się jednemu z takich przykładów, wspomnianemu na wstępie naszych rozważań. Terminy obserwacyjne, jakie w tym przypadku wchodzą w grę — to dwuargumentowy predykat „jest dwa razy krótszy od" i nazwa (a raczej deskrypcja) „wzorzec metra z Sèvres". Łatwo można okazać, że za pomocą owego predykatu obserwacyjnego zdefiniować się daje (przez odpowiedni ciąg definicji równoważnościowych) dwuargumentowy predykat „jest dziesięć miliardów razy krótszy od", a za pomocą tego ostatniego i wspomnianej nazwy obserwacyjnej jednoargumentowy predykat „jest dziesięć miliardów razy krótszy od wzorca metra z Sèvres" (inaczej — „ma długość 1 Å"), a więc predykat teoretyczny w ścisłym słowa tego znaczeniu, desygnujący wyłącznie przedmioty niespostrzegalne (mikroobiekty o rozmiarach atomu). Czy definicja nasza nadaje w istocie predykatowi temu interpretację zamierzoną? Odpowiedź zależy najwyraźniej od tego, jak pojmujemy ów wyjściowy predykat „jest dwa razy krótszy od". Jeśli traktujemy go jako predykat obserwacyjny w przyjętym przez nas sensie, a więc zdefiniowany ostensywnie i pozbawiony wobec tego określonej interpretacji w dziedzinie przedmiotów niespostrzegalnych, nie możemy uznać interpretacji zdefiniowanego za jego pomocą predykatu teoretycznego za interpretację adekwatną, choćby dlatego że dopuszcza ona wówczas jako jedną z jego denotacji zbiór pusty. Może to być interpretacja adekwatna tylko przy założeniu, że predykat „jest dwa razy krótszy od" ma interpretację ściśle określoną również w zbiorze przedmiotów niespostrzegalnych. Ale wówczas nie uznamy go, zgodnie z naszą terminologią, za predykat obserwacyjny. Będzie to po prostu jeden z predykatów teoretycznych — i to predykatów teoretycznych *sensu stricto*. Przykład powyższy nie stoi zatem w sprzeczności z tezą o obserwacyjnej niedefiniowalności terminów teoretycznych, rozumianą w przyjęty w tej pracy sposób.

Teza to, jak widzieliśmy, nienowa, ale uzasadniana zwykle w sposób diametralnie różny od przedstawionego powyżej. Tradycyjne argumenty przeciwko obserwacyjnej definiowalności terminów teoretycznych zwracają uwagę na fakt, że definiowanie takie jest procedurą zbyt rygorystyczną. Ma ono wyznaczać interpretację terminu teoretycznego w sposób zbyt sztywny, wiążąc ją zbyt ściśle ze specyficznymi kryteriami jego stosowalności. Argumentacja nasza podkreśla fakt wręcz przeciwny. Definiowanie terminów teoretycznych przez terminy obserwacyjne jest procedurą w pewnym sensie zbyt liberalną. Wyznacza interpretację danego terminu w sposób zbyt luźny, dopuszczając, z konieczności, interpretacje jawnie niezamierzone. Sądzę, że mimo tych przeciwieństw oba rodzaje argumentacji zawierają spostrzeżenia słuszne, tylko dotyczące sytuacji różnych. Nasze wywody odnosiły się do terminów teoretycznych w ścisłym tego słowa znaczeniu, a więc terminów desygnujących pewne przedmioty niespostrzegalne. Argumentacja tradycyjna natomiast zachowuje walor w stosunku do terminów teoretycznych desygnujących wyłącznie przedmioty spostrzegalne. Skoro przy wprowadzaniu tego rodzju terminów do języka obserwacyjnego nie rozszerzamy jego universum poza dziedzinę przedmiotów spostrzegalnych, definicja obserwacyjna takiego terminu determinuje jego interpretację w sposób jednoznaczny, a to w wielu przypadkach wydaje się istotnie ograniczeniem niepożądanym[6]. W ten sposób pogląd głoszący obserwacyjną niedefiniowalność terminów teoretycznych znajduje potwierdzenie w różnych stron i za pomocą różnych argumentów — w zależności od takiej czy innej eksplikacji jego zakresu i treści.

DODATEK

Szkic dowodu przytaczanego w artykule twierdzenia (T1) ograniczymy do przypadku możliwie najprostszego, ale zarazem dostatecznie ogólnego. Niech J_o będzie językiem rachunku predykatów z identycznością, zawierającym jeden dwuargumentowy predykat pozalogiczny o. *Modelami* języka J_0 nazywać będziemy wyłącznie modele z absolutnym pojęciem identyczności. Jeśli nie ograniczamy się do modeli tego typu, dowód twierdzenia (T1) jest natychmiastowy. Przy ograniczeniu się do modeli z absolutnym pojęciem identyczności dowodzimy twierdzenia (T1), opierając się na następującym lemacie:

(L) *Niech $\alpha(x)$ będzie formułą języka J_o o jednej zmiennej wolnej, a $\mathfrak{m}_0 = \langle U_o, O_o\rangle$ modelem tego języka. Jeżeli pewien element universum U_o spełnia formułę $\alpha(x)$ w każdym rozszerzeniu modelu $\mathfrak{m}_o$, to istnieje formuła czysto egzystencjalna $\beta(x)$ spełniona przez pewien element universum U_o w modelu $\mathfrak{m}_0$, taka iż*

$$\bigwedge x\,(\beta(x) \rightarrow \alpha(x)) \in Cn(O).$$

Lemat ten stanowi pewną odmianę znanych zależności teoriomodelowych i jego dowód przebiega w analogiczny sposób. Niech J_o' będzie językiem J_o wzbogaconym o nazwy wszystkich elementów universum U_o, $\{s_u\}_{u\in U_o}$, Cn' konsekwencją w J_o', a $\mathfrak{m}_o' = \langle U_o, O_o, \{a_u\}_{u\in U_o}\rangle$, gdzie $a_u = u$, modelem J_o', będącym przedłużeniem modelu $\mathfrak{m}_o$. Załóżmy, że spełniony jest poprzednik lematu (L). Wówczas, dla pewnej nazwy s_u, zdanie $\alpha(s_u)$ jest prawdziwe w każdym rozszerzeniu modelu $\mathfrak{m}_o'$. Okażemy, że zdanie to musi być konsekwencją diagramu modelu $\mathfrak{m}_o$, $D(\mathfrak{m}_o)$. Przypuśćmy, że tak nie jest: $\alpha(s_u) \notin Cn'(D(\mathfrak{m}_o))$. Istnieje wobec tego model $\mathfrak{m}'$ języka J_o', który jest modelem diagramu $D(\mathfrak{m}_o)$, a nie jest modelem zdania $\alpha(s_u)$. Na mocy znanych twierdzeń teorii modeli[7], model $\mathfrak{m}'$ jest rozszerzeniem *sensu largo* modelu $\mathfrak{m}_o'$, czyli jest rozszerzeniem zwykłym (*sensu stricto*) pewnego modelu izomorficznego z $\mathfrak{m}_o'$. Zdanie $\alpha(s_u)$ jest fałszywe w modelu $\mathfrak{m}'$. Ponieważ każde rozszerzenie *sensu largo* modelu $\mathfrak{m}_o'$ jest izomorficzne z pewnym rozszerzeniem zwykłym tegoż modelu, zdanie $\alpha(s_u)$ okazuje się fałszywe w pewnym rozszerzeniu zwykłym modelu $\mathfrak{m}_o'$, a to jest sprzeczne z przyjętym założeniem. Wnosimy stąd, iż $\alpha(s_u) \in Cn'(D_o(\mathfrak{m}_o))$. Istnieje wobec tego skończony podzbiór $D_o(\mathfrak{m}_o)$ diagramu $D(\mathfrak{m}_o)$ taki, iż $\alpha(s_u) \in Cn'(D(\mathfrak{m}_o))$. Niech $\gamma(s_u, s_{u_1}, \ldots, s_{u_n})$ będzie koniunkcją

6) Ujemne konsekwencje takiego ograniczenia starałem się przedstawić m. in. w artykule „Pojęcia teoretyczne a doświadczenie", *Studia Logica*, 1961.

7) Por. np. J. Łoś, „On the Extending of Models I", *Fundamenta Mathematicae*, 1955.

zdań zbioru $D_o(\mathfrak{m}_o)$, a $s_u, s_{u_1}, ..., s_{u_n}$ wszystkimi nazwami występującymi w tych zdaniach. Skoro $\alpha(s_u)$ jest konsekwencją $\gamma(s_u, s_{u_1}, ..., s_{u_n})$, a nazwy $s_{u_1}, ..., s_{u_n}$ nie występują w $\alpha(s_u)$, $\alpha(s_u) \in Cn'(\bigvee x_1 \ldots \bigvee x_n \gamma(s_u, x_1, ..., x_n))$. To z kolei pociąga zależność:

$$\bigwedge x\,(\bigvee x_1 \ldots \bigvee x_n \gamma(x, x_1, ..., x_n) \rightarrow \alpha(x)) \in Cn(O).$$

Ponieważ $\gamma(s_u, s_{u_1}, ..., s_{u_n})$ jest koniunkcją zdań z diagramu modelu $\mathfrak{m}_o$, formuła $\beta(x) = \bigvee x_1 \ldots \bigvee x_n \gamma(x, x_1, ..., x_n)$ jest formułą czysto egzystencjalną, spełnioną przez pewien element universum U_o w modelu $\mathfrak{m}_o$. Tym samym spełnia warunki wymienione w następniku lematu (L).

Lemat ten posłuży nam do dowodu twierdzenia T1:

(T1) *Jeżeli pewien element universum U_o spełnia formułę $\alpha(x)$ w każdym rozszerzeniu modelu $\mathfrak{m}_o$, to dla dowolnego zbioru U zawierającego zbiór U_o istnieje rozszerzenie $\mathfrak{m}$ modelu $\mathfrak{m}_o$ o universum U takie, iż każdy element universum U nie należący do U_o spełnia formułę $\alpha(x)$ w modelu $\mathfrak{m}$.*

Zakładamy prawdziwość poprzednika twierdzenia (T1) i korzystamy z lematu (L). Niech $\beta(x)$ będzie formułą, której istnienie gwarantuje, przy powyższym założeniu, lemat (L). Aby okazać prawdziwość następnika twierdzenia (T1), wystarczy okazać, że warunki sformułowane w tym następniku spełnia formuła $\beta(x)$. Formuła ta ma, jak wiemy, postać $\bigvee x_1 \ldots \bigvee x_n \gamma(x, x_1, ..., x_n)$, gdzie $\gamma(x, x_1, ..., x_n)$ jest koniunkcją pewnych formuł atomowych (m. in. identycznościowych) oraz negacji takich formuł. Jest to formuła spełniona przez pewien element universum U_o — niech będzie nim u — w modelu $\mathfrak{m}_o$, a w konsekwencji (jako formuła czysto egzystencjalna) i w każdym rozszerzeniu modelu $\mathfrak{m}_o$. Niech U będzie nadzbiorem U_o. Musimy zdefiniować takie rozszerzenie $\mathfrak{m}$ modelu $\mathfrak{m}_o$ o universum U, w którym każdy element zbioru $U—U_o$ spełnia formułę $\beta(x)$. Rozszerzenie to konstruujemy definiując relację O w sposób następujący: dla dowolnych $u_i', u_j' \in U$, $O(u_i', u_j') \equiv O_o(u_i, u_j)$, gdzie $u_k = u_k'$, gdy $u_k' \in U_o$, oraz $u_k = u$, gdy $u_k' \in U—U_o$. Tak określona relacja O pokrywa się z relacją O_o w zbiorze U_o. Natomiast wszystkie elementy zbioru $U—U_o$ zachowują się ze względu na relację O tak, jak element u ze względu na relację O_o. Gdyby formuła $\beta(x)$ nie zawierała formuł identycznościowych, widoczne byłoby natychmiast, że w tak zdefiniowanym rozszerzeniu $\mathfrak{m}$ każdy element zbioru $U—U_o$ spełnia formułę $\beta(x)$ (skoro spełnia ją element u). Pozostaje nam wobec tego jedynie rozpatrzenie ewentualnych formuł identycznościowych występujących w formule $\beta(x) = \bigvee x_1 \ldots \bigvee x_n \gamma(x, x_1, ..., x_n)$. W grę wchodzą tylko formuły zawierające zmienną x. Mogą być one, ogólnie biorąc, czterech rodzajów: (1) $x = x$, (2) $x \neq x$, (3) $x = x_i$, (4) $x \neq x_i$. Formuła (1) jest spełniona przez każdy przedmiot, a formuła (2) przez żaden, a więc w $\beta(x)$ występować nie może. Formuła (3) występuje w sposób nieistotny, bo każde wyrażenie typu $\bigvee x_i\,(x = x_i \wedge \phi(x_i))$ jest równoważne logicznie wyrażeniu $\phi(x)$. Pozostaje formuła (4), a ta spełniona jest przez każdy element zbioru $U—U_o$, skoro w formule $\beta(x)$ spełnionej przez u w modelu $\mathfrak{m}_o$ x_i reprezentuje pewien element zbioru U_o. W rezultacie każdy element zbioru $U—U_o$

spełnia formułę $\beta(x)$, a w konsekwencji formułę $\alpha(x)$, w rozszerzeniu $\mathfrak{m}$, co należało okazać.

Przytaczane w artykule twierdzenia (T2) i (T3) są natychmiastowymi konsekwencjami twierdzenia (T1).

SEMANTYKA

Filozofia Nauki
Rok I, 1993, Nr 2-3

W sprawie terminów nieostrych

Sprawa terminów nieostrych stanowi jeden z niepokojących problemów logicznych. Istnieją przedmioty, co do których nigdy nie potrafimy rozstrzygnąć, czy podpadają one, czy też nie podpadają pod taki termin. I to nie z powodu niedostatecznej znajomości tych przedmiotów. Czy dlatego więc, że przedmioty te istotnie nie są ani *A*, ani nie-*A*? Tak właśnie odpowiadają niektórzy, odrzucając tym samym jedno z fundamentalnych praw logicznych, głoszące iż każdy przedmiot posiada bądź cechę *A*, bądź jej negację. Inni, chcąc uniknąć tej drastycznej konsekwencji, przyjmują, że wchodzące tu w grę przedmioty są *A* lub nie-*A*, a tylko rozstrzygnięcie, która z tych ewentualności zachodzi, jest rzeczą niemożliwą. Uznają tym samym istnienie w języku nauki twierdzeń zasadniczo nierozstrzygalnych, co również nie może być uznane za rozwiązanie zadowalające. Reprezentowane bywa wreszcie i trzecie stanowisko. Zgodnie z nim, twierdzenia, o których mowa — to w ogóle nie zdania w sensie logicznym, lecz wypowiedzi pozbawione znaczenia, «nonsensy». Odpadają tym samym na gruncie tego poglądu kłopoty z ograniczeniem powszechnego waloru praw logicznych, czy z istnieniem zasadniczo nierozstrzygalnych problemów. Uwagi poniższe mają na celu dostarczenie pewnej dodatkowej argumentacji na poparcie tego stanowiska. Polega ono na powiązaniu nieostrości terminów z tym sposobem wprowadzania wyrażeń do języka nauki, który znany jest pod nazwą redukcji lub definicji cząstkowej. Idę tu za sugestiami Mehlberga, który pierwszy zwrócił uwagę na ten związek, nie wyciągając stąd jednak podobnych konsekwencji.[1]

1) „Positivisme et science I". *Studia Philosophica* III, 1948.

Teorię definicji cząstkowych opracował Carnap w pracy „Testability and Meaning"[2]. Była ona potem niejednokrotnie — również w polskiej literaturze logicznej — referowana i stosowana przy rozważaniu pewnych zagadnień[3]. Dlatego też ograniczę się tutaj do skrótowego i schematycznego jej przypomnienia. Najogólniejsza postać definicji cząstkowej terminu Q składa się, wedle Carnapa, z następujących wypowiedzi:

(1) $\bigwedge_x [P_1(x) \rightarrow (P_2(x) \rightarrow Q(x))]$.

(2) $\bigwedge_x [R_1(x) \rightarrow (R_2(x) \rightarrow \sim Q(x))]$.

Zgodnie z wyjaśnieniami autora wypowiedź (1) ustala znaczenie terminu Q dla przedmiotów spełniających warunki P_1 i P_2, wypowiedź (2) — dla przedmiotów spełniających warunki R_1 i R_2; pierwsze z tych przedmiotów zostaną określone jako Q, drugie — jako $\sim Q$. W stosunku do przedmiotów, które nie są ani P_1 i P_2, ani R_1 i R_2, definicja ta nie formułuje żadnych kryteriów stosowalności terminu Q.

Carnapowska postać definicji cząstkowych podyktowana została przez wzgląd na to, iż definicje te miały — w myśl jego intencji — służyć do wprowadzania terminów dyspozycyjnych. Jeśli zadania definicji cząstkowych nie chcemy zawężać do tego jedynie celu, możemy je z łatwością uogólnić, co też uczynił Mehlberg w cytowanej wyżej pracy. Uogólnienie takie narzuca się od razu, gdy zważymy, iż wypowiedzi (1) i (2) równoważne są następującym formułom:

(3) $\bigwedge_x (P_1(x) \cdot P_2(x) \rightarrow Q(x))$

(4) $\bigwedge_x (R_1(x) \cdot R_2(x) \rightarrow \sim Q(x))$.

Postać ogólniejszą otrzymujemy przechodząc po prostu do wypowiedzi:

(5) $\bigwedge_x (P(x) \rightarrow Q(x))$

(6) $\bigwedge_x (R(x) \rightarrow \sim Q(x))$.

Wypowiedzi te — podobnie jak formuły (1) i (2) — ustalają znaczenie terminu definiowanego Q tylko częściowo: dla przedmiotów spełniających bądź warunek P, bądź R. Przedmioty, które są P, podpadają pod termin Q, przedmioty, które są R — pod jego negację. Oprócz nich mogą jednak tutaj istnieć zawsze przedmioty, które nie są ani P, ani R. W stosunku do nich znaczenie terminu Q jest nieustalone. Wypowiedzi (5) i (6) nie formułują w tych przypadkach żadnych kryteriów stosowalności. Gdy $R = \sim P$, definicja cząstkowa przechodzi w definicję zupełną. Wypowiedzi (5) i (6) stają się równoważne zwykłej formule definicyjnej:

(7) $\bigwedge_x (Q(x) \equiv P(x))$.

Ta ustala znaczenie terminu Q dla wszystkich — w zasadzie — przedmiotów, gdyż każdy przedmiot jest bądź P, bądź $\sim P$.

2) *Philosophy and Science* III/IV, 1936/37.

3) M.in. J. Kamińska: „Ewolucja Koła Wiedeńskiego". *Myśl Współczesna* 2/9, 1947; M. Przełęcki: „O tzw. definicjach operacyjnych", *Studia Logica* III, 1955.

W świetle powyższej charakterystyki definicji cząstkowych wyłania się możliwość ujęcia terminów nieostrych jako terminów cząstkowo definiowalnych. Znaczenie terminu nieostrego wydaje się takie, iż oddać je może jedynie definicja cząstkowa. Istnieją bowiem zawsze przedmioty, wobec których znaczenie to jest nieokreślone, wobec których nie dysponujemy żadnym kryterium stosowalności danego terminu. Każdy, kto ma mniej niż 18 lat, jest jeszcze na pewno młodzieńcem; nikt, kto ma więcej niż 30, młodzieńcem już nie jest. A czy jest nim ktoś, kto liczy lat 25? Wydaje się, iż nie posiadamy w tym przypadku żadnych kryteriów stosowalności tego terminu, które by mogły umożliwić nam decyzję. Jego definicję stanowić by więc mogła następująca definicja cząstkowa:

(8) $\bigwedge_{x}$ (x ma mniej niż 18 lat $\rightarrow$ x jest młodzieńcem).

(9) $\bigwedge_{x}$ (x ma więcej niż 30 lat $\rightarrow$ x nie jest młodzieńcem).

Można się oczywiście spierać o to, czy trafnie zostały tutaj dobrane owe graniczne lata. Spór to raczej bezprzedmiotowy, gdyż i tutaj mamy do czynienia ze zjawiskiem nieostrości: nieostrości drugiego — rzec można — stopnia. Nie biorąc jej pod uwagę w obecnych rozważaniach, upraszczam oczywiście całe zagadnienie. Jedno stwierdzić można jednak z całą pewnością. Jakkolwiekbyśmy przesunęli te granice, zawsze pozostać muszą przedmioty, co do których wypowiedzi definicyjne nie mogą ustalać żadnych kryteriów stosowalności terminu „młodzieniec", jeśli tylko znaczenie tego terminu odpowiadać ma jego potocznemu sensowi. Tak samo jest w przypadku pozostałych terminów nieostrych.

Pewne wątpliwości może budzić obszerna klasa terminów nieostrych, obejmująca nazwy wszelkich jakości zmysłowych. Klasycznym przykładem terminu nieostrego jest przecież obok nazwy „młodzieniec" orzecznik „czerwony". Wątpliwości te płyną stąd, iż mamy tu do czynienia z terminami, które uchodzą za terminy pierwotne. Trudno więc traktować je jako terminy cząstkowo definiowalne. Mówi się jednakże i tutaj o ustalaniu znaczenia tych terminów za pomocą tzw. definicji dejktycznych. Definicje te, które odróżnia się na ogół od definicji *sensu stricto*, polegać mają, w myśl ogólnikowych wyjaśnień, na wskazaniu takich przedmiotów, które pod dany termin podpadają, i takich, które do jego zakresu nie należą. Wydaje się, iż zabieg ten pozostawia zawsze pewną klasę przedmiotów, co do których żadnego kryterium stosowalności danego terminu się nie ustaliło. Przedmioty takie, jak te oto, są czerwone; przedmioty takie, jak tamte, czerwone nie są. Określenia te pojmowane bywają zwykle tak, że suma klas przedmiotów takich, jak te, i takich, jak tamte, nie wyczerpuje bynajmniej ogółu przedmiotów. Istnieją więc i tu przypadki, wobec których znaczenie terminu „czerwony" pozostaje niezdeterminowane. Owe pseudo-definicje dejktyczne traktować można zatem jako pewne pseudo-definicje cząstkowe.

Nieostre mogą być oczywiście nie tylko terminy zdefiniowane cząstkowo. Mogą to być również terminy wprowadzone za pomocą definicji zupełnych, jeżeli w ich definiensie figurują już jakieś terminy nieostre. I tak, jeśli w przytoczonej wyżej definicji

(7) występujący w definiensie termin P jest terminem nieostrym, taki sam charakter będzie miał oczywiście definiowany termin Q. Warto jednak zaznaczyć, iż nie musi być tak w przypadku każdej definicji zupełnej. Zależy to od rodzaju takiej definicji. Mogą istnieć definicje zupełne, w których nieostrość terminów pierwotnych nie przenosi się na termin definiowany.

Po tych zastrzeżeniach możemy powiedzieć ogólnie, że terminy nieostre — to terminy, które «w ostatecznej instancji» definiowalne są jedynie cząstkowo. Jaki charakter mają przy tym założeniu twierdzenia budzące w logikach tyle teoretycznego niepokoju: twierdzenia orzekające terminy nieostre o przedmiotach należących do tzw. zakresu nieostrości? Jeśli definicję terminu nieostrego sformułujemy tak jak wyżej, w postaci wypowiedzi (5) i (6), twierdzeniem takim będzie zdanie przypisujące orzecznik Q przedmiotowi, który nie jest ani P, ani R. Ale znaczenie terminu Q zostało określone tylko dla tych przedmiotów, które jeden z tych warunków spełniają. Nasuwa się wobec tego przypuszczenie, iż omawiane twierdzenie jest w gruncie rzeczy wypowiedzią pozbawioną sensu, gdyż zawiera termin, któremu w tym kontekście żadne określone znaczenie nie przysługuje. Termin nieostry nie posiada — jako termin cząstkowo definiowalny — uniwersalnych kryteriów stosowalności. W twierdzeniach, o których mowa, wykraczamy poza ustalone dla danego terminu kryteria. Używamy go w takiej sytuacji, dla której żadnych kryteriów stosowalności nie sformułowaliśmy. Jeśli odpowiednia definicja cząstkowa stanowi jedyny zabieg wprowadzający dany termin do języka, termin ten w omawianej sytuacji pozbawiony jest jakiegokolwiek znaczenia. Tym samym też pozbawione jest znaczenia całe zawierające go wyrażenie.

Terminami zdefiniowanymi cząstkowo — a więc i terminami nieostrymi — operować musimy ostrożniej, niż terminami wprowadzonymi za pomocą definicji zupełnych. Nie każde poprawnie zbudowane wyrażenie, w którym figuruje termin zbudowany cząstkowo, jest wyrażeniem sensownym. Sensowność wyrażeń jest tutaj uzależniona od pewnych faktów pozajęzykowych. To, czy wyrażenie $Q(a)$ jest zdaniem, czy nonsensem, zależne jest od tego, czy przedmiot a jest, czy też nie jest P lub R. O tym zaś może nas w zasadzie przekonać tylko doświadczenie. Jest to niewątpliwie konsekwencja kłopotliwa. Wydaje się jednak, że nie jest to jedyna sytuacja, w której sensowność jakiegoś wyrażenia uzależniona jest od pewnych faktów empirycznych. Mam tu na myśli m.in. zwroty okazjonalne. Wyrażenie okazjonalne „ten" posiada określony sens w zależności od tego, czy towarzyszy mu, czy też nie, pewien gest wskazujący. Poprawność budowy nie wystarcza więc i tutaj dla zapewnienia sensowności zwrotom zawierającym to wyrażenie. Można chyba stwierdzić, że zarówno w języku potocznym, jak i w języku nauk empirycznych, którym w faktycznej praktyce badawczej jest w dużym stopniu język potoczny, względy czysto formalne nie wystarczają do zagwarantowania sensowności wyrażeń. Mają tutaj walor takie kryteria sensowności, które odwołują się do faktów pozajęzykowych. Takie m.in. kryteria decydują o sensowności wyrażeń zawierających terminy cząstkowo definiowalne, których poszczególny przypadek stanowią nazwy nieostre.

Przyjęcie takiego stanowiska usuwa — za cenę wspomnianych komplikacji — te kłopoty płynące z posługiwania się terminami nieostrymi, które sygnalizowałem na wstępie tych rozważań. Twierdzenia, które — zdaniem niektórych — miały stać w sprzeczności z prawem wyłączonego środka czy z prawem sprzeczności, okazują się wyrażeniami pozbawionymi znaczenia i jako takie nie zagrażają oczywiście ich prawdziwości. Co więcej, nie wprowadzamy tu również zasadniczo nierozstrzygalnych twierdzeń. Wyrażenie orzekające termin nieostry o przedmiocie z tzw. zakresu nieostrości jest rezultatem nieuprawnionego użycia tego terminu poza dziedziną jego stosowalności i jako takie pozbawione jest sensu. Znika zatem niepokojąca możliwość formułowania przy pomocy terminów nieostrych zdań, które mają określone znaczenie, a których mimo to nigdy nie potrafimy rozstrzygnąć.

Oto główne — ujemne i dodatnie — konsekwencje zarysowanego stanowiska. Zdając sobie w pełni sprawę ze szkicowego i wysoce problematycznego charakteru powyższych uwag, traktować je chciałbym jako głos w dyskusji nad jednym z wciąż aktualnych problemów logicznych.

Filozofia Nauki
Rok I, 1993, Nr 2-3

Z semantyki pojęć otwartych

I

1. Główny problem, któremu będą poświęcone te rozważania, ma bogatą tradycję filozoficzną. Wyłonił się on po raz pierwszy w związku z refleksją nad nazwami nieostrymi języka potocznego. Klasyczny ich przykład stanowić może nazwa „młodzieniec". Nie ulega wątpliwości, iż każdy, kto ma mniej niż 18 lat, jest jeszcze młodzieńcem; nikt, kto ma więcej niż 30 lat, młodzieńcem już nie jest. A czy jest nim ktoś, kto liczy, dajmy na to, 25 lat? Znaczenie, jakie przysługuje tej nazwie w języku potocznym, jest takie, iż pytania tego rozstrzygnąć nam nie pozwala. Jaki charakter ma wobec tego wypowiedź stwierdzająca, iż ów 25-letni osobnik jest młodzieńcem? Czy wypowiedź ta jest zdaniem prawdziwym lub fałszywym, lecz zasadniczo nierozstrzygalnym — jak mówią jedni? Czy też jest pozbawiona wartości logicznej — jak chcą inni? A może zarówno ona, jak i jej negacja, są zdaniami fałszywymi — jak głoszą niektórzy, odrzucając tym samym jedno z fundamentalnych praw logicznych: zasadę wyłączonego środka? Pytania te pociągają za sobą dalsze. O czym się właściwie w przytoczonej wypowiedzi mówi? Czy nazwa „młodzieniec" denotuje jakiś określony zbiór przedmiotów? A jeśli tak, to jaki?

Te problemy natury semantycznej dotyczące nazw nieostrych języka potocznego nabrały ostatnio aktualności w związku z badaniami logicznymi nad językiem empirycznych teorii naukowych. Podobny charakter jak nazwy nieostre języka potocznego wykazuje jeden z podstawowych rodzajów terminów występujących w teoriach empirycznych: tzw. terminy teoretyczne. I one zatem pozwalają na formułowanie wypowiedzi nasuwających te same pytania, które powstają w związku z nazwami nieostrymi. Bo też wspólne wydaje się w obu przypadkach źródło tych trudności. Nazwy nieostre i terminy teoretyczne odznaczają się tymi samymi logicznymi właściwościami. Zarówno jedne, jak i drugie określić możemy — upraszczając nieco sprawę — jako terminy definiowalne warunkowo (lub — w innej terminologii — cząstkowo). Adekwatna definicja takiego terminu przybierać musi postać definicji warunkowej. Niechaj terminem tym będzie predykat Q. Definicja warunkowa predykatu Q — to wypowiedź o postaci

(1) $$\bigwedge_x [\Psi x \rightarrow (Qx \equiv \Phi x)],$$

gdzie Ψ i Φ — to wyrażenia o ustalonym uprzednio znaczeniu. Jak widać, definicja (1) ustala znaczenie predykatu Q tylko częściowo: dla przedmiotów spełniających warunek Ψ. Te z nich, które są Φ, są Q; te, które nie są Φ, nie są Q. W stosunku do przedmiotów nie spełniających warunku Ψ, definicja (1) żadnych kryteriów stosowalności predykatu Q nie ustanawia.

Pod schemat definicji warunkowej (1) dają się podciągnąć i inne spotykane postaci definicji cząstkowych. I tak, tzw. definicja redukcyjna predykatu Q przy pomocy predykatów P_1, P_2

(2) $$\bigwedge_x [P_1x \rightarrow (Qx \equiv P_2x)]$$

podpada bezpośrednio pod schemat (1). Ale pod schemat ten można też podciągnąć definicję cząstkową predykatu Q w jej postaci ogólnej

(3) $$\bigwedge_x [(P_1x \rightarrow Qx) \wedge (P_2x \rightarrow \sim Qx)].$$

Definicja ta pociąga twierdzenie

$$\sim \bigvee_x (P_1x \wedge P_2x);$$

przy założeniu jego prawdziwości jest równoważna wypowiedzi

$$\bigwedge_x [(P_1x \vee P_2x) \rightarrow (Qx \equiv P_1x)],$$

a więc wypowiedzi podpadającej już bezpośrednio pod schemat (1)[1]. Pod schemat ten podpadają również tzw. jednostronne definicje cząstkowe predykatu Q:

(4) $$\bigwedge_x (P_1x \rightarrow Qx),$$

oraz

(5) $$\bigwedge_x (P_2x \rightarrow \sim Qx).$$

Pierwsza z nich jest równoważna logicznie wypowiedzi

$$\bigwedge_x [P_1x \rightarrow (Qx \equiv P_1x)].$$

druga — wypowiedzi

$$\bigwedge_x [P_2x \rightarrow (Qx \equiv \sim P_2x)].$$

1) Definicja cząstkowa predykatu Q w jej postaci ogólnej może być sformułowana tak, aby była definicją nietwórczą:

$$\bigwedge_x [(P_1x \wedge \sim P_2x \rightarrow Qx) \wedge (P_2x \wedge \sim P_1x \rightarrow \sim Qx)].$$

Definicja ta jest równoważna logicznie wypowiedzi typu (1)

$$\bigwedge_x [(P_1x \equiv \sim P_2x) \rightarrow (Qx \equiv P_1x)].$$

Otóż znaczenie nazwy nieostrej wydaje się takie, iż jej adekwatna definicja, odwołująca się wyłącznie do nazw ostrych, musi być definicją warunkową. Taką definicję nazwy „młodzieniec" mogłaby stanowić następująca definicja cząstkowa typu (3):

$$\bigwedge_x [(x \text{ ma mniej niż 18 lat} \rightarrow x \text{ jest młodzieńcem}) \wedge (x \text{ ma więcej niż 30 lat} \rightarrow x \text{ nie jest młodzieńcem})].$$

Widać zarazem na tym przykładzie, na czym polega tu uproszczenie zagadnienia. Owe granice wieku mają w pewnej mierze charakter arbitralny i niezależnie od tego, jak je dobierzemy, w jakimś stopniu arbitralne pozostaną. Mamy tu do czynienia ze zjawiskiem nieostrości drugiego — rzec można — stopnia, którego w naszych rozważaniach nie bierzemy w ogóle pod uwagę.

Uproszczenie w przypadku terminów teoretycznych polega na czymś innym. Badania logiczne nad językiem empirycznych teorii naukowych prowadzą do wniosku, iż związki logiczne, jakie teoria empiryczna ustanawia pomiędzy terminami teoretycznymi a terminami spostrzeżeniowymi, sprowadzają się na ogół do definicji warunkowych terminów teoretycznych przez terminy spostrzeżeniowe, przybierających którąś z uwzględnionych przez nas postaci. Taki charakter ma np. na gruncie genetyki klasycznej definicja terminu teoretycznego „genotyp" odwołująca się wyłącznie do takich terminów w szerokim sensie spostrzeżeniowych, jak „fenotyp", czy „potomek". Można ją sformułować w postaci definicji cząstkowej typu (3). Pozostaje jednak zagadnieniem otwartym, czy ten rodzaj związku logicznego z terminami spostrzeżeniowymi charakteryzuje wszystkie terminy teoretyczne. Wysuwa się niekiedy możliwości innych jeszcze, luźniejszych niż definicja warunkowa, związków logicznych pomiędzy terminami teoretycznymi a terminami spostrzeżeniowymi. Nie jest jednak rzeczą stwierdzoną, czy możliwości te znajdują realizację w jakiejś istniejącej teorii naukowej[2].

Nie przesądzając zatem sprawy, czy rozważaniami swymi obejmiemy istotnie wszelkie nazwy nieostre lub wszelkie terminy teoretyczne, rozważania te ograniczymy wyraźnie do terminów definiowalnych warunkowo. Pojęcia odpowiadające takim terminom obejmuje się niekiedy nazwą pojęć „otwartych". One to właśnie nasuwają przytoczone na wstępie problemy semantyczne. Niech a_1 będzie przedmiotem nie spełniającym sformułowanego w definicji warunkowej (1) predykatu Q warunku Ψ. Zdanie przypisujące przedmiotowi a_1 predykat Q: Qa_1 stanowić może przykład owych problematycznych pod względem semantycznym wypowiedzi. Zadaniem naszym będzie przede wszystkim scharakteryzowanie w sposób ogólny i ścisły klasy owych problematycznych wypowiedzi. Zdanie Qa_1 stanowi tylko jeden z ich przykładów. Jest zaś z drugiej strony rzeczą jasną, iż wypowiedzi te nie obejmują wszystkich zdań zawierających termin Q. Postaramy się następnie przedstawić za pomocą możliwie precyzyjnej i jednolitej aparatury pojęciowej zarówno podstawowe problemy semanty-

2) Sprawy te omawiam bliżej w pracach: „Pojęcia teoretyczne a doświadczenie", *Studia Logica* 11 (1961); „O pojęciu genotypu", *Studia Filozoficzne* 26 (1961).

czne, jakie klasa owych wypowiedzi nasuwa, jak i główne, znane z literatury, próby ich rozwiązania. Aparatury takiej dostarczy nam współczesna semantyka logiczna, pojęta jako teoria modeli języków sformalizowanych. Zastosowanie tej aparatury do wspomnianych zagadnień wzoruje się na tym sposobie wykorzystania współczesnej semantyki logicznej w problematyce filozoficznej, który znalazł wyraz w pracy R. Suszki „Logika formalna a niektóre zagadnienia teorii poznania"[3].

Rozważania swe ograniczymy zatem do języków sformalizowanych, i to języków możliwie prostych pod względem syntaktycznym. Będą to wyłącznie języki elementarne, oparte na węższym rachunku predykatów z identycznością. Przykład ich stanowić będzie język J, który prócz *zmiennych indywiduowych*: x, y, ... oraz *stałych logicznych* $\sim$, $\wedge$, $\vee$, $\rightarrow$, $\equiv$, $\bigwedge_x$, $\bigvee_x$, $=$, zawiera proste wyrażenia logiczne dwóch rodzajów — *nazwy indywiduowe* a_1, a_2, ..., a_n oraz *predykaty* (o dowolnej liczbie argumentów): P_1, P_2, ... , P_m[4]. *Reguły konstrukcji* języka J, których tu przytaczać nie będziemy, charakteryzują, w jaki sposób wyrażenia języka J, w szczególności zdania tego języka, zbudowane są z wyrażeń prostych. Tak pojęty język J jest tworem formalnym (lub raczej na wpół formalnym). Poza stałymi logicznymi, co do których zakładamy, iż wyposażone są w klasyczne własności semantyczne, pozostałe wyrażenia języka J scharakteryzowane są tylko pod względem syntaktycznym. Wyrażenia te pełnią określone funkcje semantyczne — do czegoś się odnoszą i coś stwierdzają — dopiero wtedy, gdy obok języka J dany jest pewien jego model, który wyznacza określoną interpretację języka J.

Pojęcia modelu języka sformalizowanego wyjaśniać tu w sposób systematyczny nie możemy. Musimy ograniczyć się do paru ogólnikowych uwag. *Modelem* danego języka J jest każda dziedzina rzeczywistości, o której można mówić w języku J. Model taki, $\mathfrak{M}$, jest układem

$$\langle U, C \rangle$$

składającym się z dwóch członów. Człon pierwszy U jest dowolnym niepustym zbiorem przedmiotów, zwanym *universum modelu* $\mathfrak{M}$. Universum to stanowi zakres zmienności zmiennych języka J. Drugi człon C, zwany niekiedy *charakterystyką modelu* $\mathfrak{M}$, obejmuje niektóre wybrane elementy zbioru U, oraz niektóre wyróżnione podzbiory zbioru U, lub relacje zachodzące między jego elementami. Każdy z członów charakterystyki C stanowi denotację pewnego prostego wyrażenia pozalogicznego języka J, oraz każde takie wyrażenie denotuje pewien człon charakterystyki C. Istnieje więc określona zależność pomiędzy strukturą syntaktyczną danego języka a typem jego modelu. I tak, każdy model $\mathfrak{M}$ opisanego wyżej języka J jest układem

$$\langle U, x_1, ..., x_n, X_1, ..., X_m \rangle,$$

3) *Myśl Filozoficzna* 28, 29 (1957).

4) Pomijamy więc dla uproszczenia wyrażenia pozalogiczne o charakterze terminów funkcyjnych. Podobnie w celu dalszego uproszczenia wywodów występujące w przykładach predykaty traktować będziemy zawsze jako predykaty jednoargumentowe. Wywody te dają się jednak bez trudu uogólnić na predykaty n-argumentowe.

który składa się z pewnego niepustego zbioru U, n wybranych elementów tego zbioru: $x_1, ..., x_n$ oraz m relacji zachodzących między jego elementami: $X_1, ..., X_m$ (przy czym relacja X_i jest tyluczłonowa, iluargumentowy jest predykat P_i)[5]. Każdy taki model $\mathfrak{M}$ wyznacza pewną *interpretację* języka J. Zmienne tego języka: $x, y, ...$ przebiegają zbiór U, nazwy indywiduowe: $a_1,..., a_n$ denotują odpowiednio przedmioty: $x_1, ..., x_n$, a predykaty: $P_1, ..., P_m$ — relacje: $X_1, ..., X_m$.

Gdy mamy dany język J oraz pewien jego model $\mathfrak{M}$, to można w sposób ścisły zdefiniować pojęcie zdania języka J prawdziwego w modelu $\mathfrak{M}$. Definicji tego podstawowego we współczesnej semantyce logicznej pojęcia przytaczać tu nie będziemy. Poprzestaniemy na prostym przykładzie. Mówiąc ogólnie i ogólnikowo zarazem, zdanie Z języka J jest *prawdziwe w modelu* $\mathfrak{M}$ wtedy i tylko wtedy, gdy rzeczy mają się tak właśnie, jak głosi zdanie Z w interpretacji języka J, wyznaczonej przez model $\mathfrak{M}$. Niechaj modelem $\mathfrak{M}$ naszego języka J będzie następujący, określony układ przedmiotów:

$$\langle \boldsymbol{U}, \boldsymbol{a}_1, ..., \boldsymbol{a}_n, \boldsymbol{P}_1, ..., \boldsymbol{P}_m \rangle.$$

Zdanie $P_1 a_1$ jest prawdziwe w modelu $\mathfrak{M}$ wtedy i tylko wtedy, gdy przedmiot stanowiący w modelu $\mathfrak{M}$ denotację nazwy a_1 należy do zbioru stanowiącego w modelu $\mathfrak{M}$ denotację predykatu P_1, a więc gdy $\boldsymbol{a}_1 \in \boldsymbol{P}_1$. Omówione obecnie pojęcia modelu języka sformalizowanego i zdania prawdziwego w modelu reprezentują te pojęcia semantyki logicznej, na których oprzemy naszą analizę semantyczną pojęć otwartych[6].

2. Załóżmy, iż przedstawiony uprzednio język J, o stałych pozalogicznych $a_1, ..., a_n$, $P_1, ..., P_m$, wzbogacamy o jednoargumentowy predykat Q wprowadzony za pomocą definicji warunkowej $D(Q)$

$$\bigwedge_x [\Psi x \rightarrow (Qx \equiv \Phi x)],$$

w której predykaty Ψ i Φ należą do (prostych lub złożonych) wyrażeń języka J. Przechodzimy w ten sposób do języka J' stanowiącego rozszerzenie języka J. Modelami $\mathfrak{M}'$ języka J' są układy przedmiotów typu

$$\langle U, x_1, ..., x_n, X_1, ..., X_m, Y \rangle.$$

Niech $Z(Q)$ reprezentuje dowolne zdanie języka J' zawierające termin Q. W stosunku do każdego takiego zdania postawić można pytanie następujące: Czy zdanie to jest na gruncie definicji $D(Q)$ równoważne jakiemuś zdaniu języka J, a więc zdaniu nie zawierającemu terminu Q? Gdyby definicja terminu Q była zwykłą definicją równoważnościową, odpowiedź na takie pytanie byłaby zawsze twierdząca. Definicja taka byłaby bowiem definicją przekładalną. Termin Q byłby na jej podstawie eliminowalny z każdego zdania $Z(Q)$. Mówiąc dokładniej, dla każdego zdania $Z(Q)$ istniałoby zdanie Z nie zawierające terminu Q, takie iż zdanie

$$D(Q) \rightarrow (Z(Q) \equiv Z)$$

5) Relacja jednoczłonowa jest identyczna z podzbiorem zbioru U.

6) W przedstawieniu powyższych pojęć semantycznych korzystałem z cytowanej pracy R. Suszki.

byłoby tautologią języka J', czyli zdaniem prawdziwym w każdym modelu $\mathfrak{M}'$ języka J'. Definicja $D(Q)$ mogłaby być definicją równoważnościową tylko wtedy, gdyby zdanie $\bigwedge_x \Psi x$ było tautologią. Zakładamy tutaj, iż tak nie jest, i że definicja $D(Q)$ nie jest równoważna logicznie żadnej definicji równoważnościowej. Jako taka nie jest definicją przekładalną. Istnieją jednak zdania języka J' zawierające termin Q, z których termin ten na gruncie definicji $D(Q)$ wyeliminować się daje. Ogół zdań języka J' zawierających termin Q możemy zatem podzielić na dwie klasy:

(1) zdania, z których Q jest eliminowalne na podstawie definicji $D(Q)$, oraz

(2) zdania, które tego warunku nie spełniają.

Warunek ów — oznaczmy go dla skrótu (EL) — definiujemy, jak następuje

(EL) *Q jest eliminowalne z $Z(Q)$ na podstawie definicji $D(Q)$ wtedy i tylko wtedy, gdy istnieje zdanie Z nie zawierające Q, takie iż zdanie*

$$D(Q) \rightarrow (Z(Q) \equiv Z)$$

jest prawdziwe w każdym modelu $\mathfrak{M}'$ (języka J')[7].

Jakież to zdania $Z(Q)$ spełniają warunek (EL)? Należą tu oczywiście te wszystkie zdania $Z(Q)$, w których Q występuje w sposób nieistotny, a więc które są logicznie równoważne zdaniom nie zawierającym Q. Będą to przede wszystkim wszelkie tautologie, np.:

$$Qa_1 \vee \sim Qa_1,$$

oraz ich negacje, np.:

$$Qa_1 \wedge \sim Qa_1;$$

prócz nich zdania, których przykładem może być wypowiedź

$$P_1a_1 \wedge (\sim P_1a_1 \rightarrow Qa_1)$$

logicznie równoważna wypowiedzi P_1a_1. Ale warunek (EL) spełniają i takie zdania $Z(Q)$, w których Q występuje w sposób istotny. Są to zdania, które, nie będąc równoważne logicznie zdaniom nie zawierającym Q, są im równoważne na gruncie definicji $D(Q)$. Oto parę przykładów:

$$\Psi a_1 \wedge Qa_1,\ \Psi a_1 \rightarrow Qa_1,\ \bigwedge_x (\Psi x \rightarrow Qx),$$

$$\bigvee_x (\Psi x \wedge Qx),\ \Psi a_1 \rightarrow \sim Qa_1,\ \sim \bigvee_x (\Psi x \wedge Qx).$$

Można z łatwością okazać, iż przy założeniu prawdziwości definicji $D(Q)$ zdania te są równoważne m.in. następującym zdaniom nie zawierającym Q:

---

7) W dalszych definicjach ową relatywizację do języka pomijamy, chcąc uprościć w miarę możności ich sformułowania. $\mathfrak{M}$ i $\mathfrak{M}'$ będziemy traktowali jako schematy reprezentujące odpowiednio modele języków J i J'.

Występujący w podanej definicji zwrot:

$D(Q) \rightarrow (Z(Q) \equiv Z)$ *jest prawdziwe w każdym modelu* $\mathfrak{M}'$

jest równoważny oczywiście stwierdzeniu:

dla dowolnego modelu $\mathfrak{M}'$: jeżeli $D(Q)$ jest prawdziwe w $\mathfrak{M}'$, to $Z(Q)$ jest prawdziwe w $\mathfrak{M}'$ wtedy i tylko wtedy, gdy Z jest prawdziwe w $\mathfrak{M}'$.

Do tego ostatniego sformułowania będziemy się w dalszym ciągu niekiedy odwoływali.

$$\Psi a_1 \wedge \Phi a_1, \Psi a_1 \rightarrow \Phi a_1, \bigwedge_x (\Psi x \rightarrow \Phi x),$$
$$\bigvee_x (\Psi x \wedge \Phi x), \Psi a_1 \rightarrow \sim \Phi a_1, \sim \bigvee_x (\Psi x \wedge \Phi x).$$

A oto z kolei przykłady zdań $Z(Q)$ nie spełniających warunku (EL):

$$Qa_1, \Psi a_1 \vee Qa_1, \ \Psi a_1 \equiv Qa_1, \ \sim \Psi a_1 \wedge Qa_1,$$
$$\sim \Psi a_1 \rightarrow Qa_1, \bigwedge_x Qx, \bigvee_x Qx, \bigwedge_x (Qx \rightarrow \Psi x).$$

Nie ma takich zdań nie zawierających terminu Q, którym by zdania powyższe były równoważne — nawet przy założeniu prawdziwości definicji $D(Q)$.

Klasę zdać $Z(Q)$ spełniających warunek (EL) scharakteryzować można na wiele sposobów. Przytoczymy tutaj jeden z nich, rzucający światło na typ kontekstu, w jakim termin Q występuje w zdaniach tej klasy. Niechaj $Z(\Psi \wedge Q)$ będzie zdaniem, które powstaje z $Z(Q)$ przez zastąpienie każdego wyrażenia typu Qx wyrażeniem typu $\Psi x \wedge Qx$. Łatwo można okazać (dowód dla uproszczenia pomijamy) zależność następującą:

Q jest eliminowalne z $Z(Q)$ na podstawie definicji $D(Q)$ wtedy i tylko wtedy, gdy zdanie

$$D(Q) \rightarrow (Z(Q) \equiv Z(\Psi \wedge Q))$$

jest prawdziwe w każdym modelu $\mathfrak{M}'$.

Zdania $Z(Q)$, które spełniają powyższy warunek, to przede wszystkim takie zdania $Z(Q)$, które są równoważne logicznie zdaniom $Z(\Psi \wedge Q)$. Wszystkie przytoczone poprzednio przykłady zdań $Z(Q)$ spełniających warunek (EL) taki właśnie mają charakter, np. zdanie $\Psi a_1 \rightarrow Qa_1$, równoważne logicznie zdaniu $\Psi a_1 \rightarrow \Psi a_1 \wedge Qa_1$. Można powiedzieć, iż są to zdania, które jeśli w ogóle mówią o Q w sposób istotny, mówią tak tylko o Ψ-owej części Q; nakładają pewne warunki co najwyżej na te przedmioty będące Q, które są Ψ zarazem. Zdania zaś, które spełniają powyższy warunek, a które nie są równoważne zdaniom $Z(\Psi \wedge Q)$ logicznie, to — zdania równoważne tym ostatnim przy założeniu prawdziwości definicji $D(Q)$. Przy tym zatem założeniu i one nakładają pewne warunki co najwyżej na te przedmioty będące Q, które są Ψ zarazem. Ich z kolei przykładem mogą być wypowiedzi:

$$\bigvee_x (\Psi x \wedge \sim Qx \wedge \Phi x) \wedge Qa_1, \ \bigwedge_x [\Psi x \rightarrow (Qx \equiv \Phi x)] \vee Qa_1.$$

Omawiana zależność staje się zrozumiała, jeśli zważymy, że definicja $D(Q)$ ustala równoważność wyrażeń Qx z wyrażeniami Φx nie zawierającymi Q — tylko dla tych przedmiotów x, które są Ψ[8].

W rozważaniach naszych ważniejszą rolę odgrywa jednak pewna własność natury semantycznej, jaka przysługuje zdaniom $Z(Q)$ spełniającym warunek (EL). Definicja $D(Q)$ dopuszcza przy ustalonej interpretacji wyrażeń języka J różne interpretacje terminu Q. Wśród ogółu zdań $Z(Q)$ wyróżnić możemy zdania, które odznaczają się tym, iż ich wartość logiczna jest niezależna od tego, którą z owych dopuszczalnych interpreta-

8) Analogiczną zależność otrzymujemy przyjmując zamiast $\Psi \wedge Q$ implikację $\Psi \rightarrow Q$.

cji terminu Q wybierzemy. Warunek ten — oznaczmy go dla skrótu (OL) — zdefiniować można jak następuje:

(OL) *$Z(Q)$ ma określoną wartość logiczną ze względu na Q na podstawie definicji $D(Q)$ wtedy i tylko wtedy, gdy dla dowolnych modeli $\mathfrak{M}'_1$ i $\mathfrak{M}'_2$ różniących się co najwyżej denotacją terminu Q zachodzi zależność następująca: jeżeli $D(Q)$ jest prawdziwe w $\mathfrak{M}'_1$ oraz $D(Q)$ jest prawdziwe w $\mathfrak{M}'_2$, to $Z(Q)$ jest prawdziwe w $\mathfrak{M}'_1$ wtedy i tylko wtedy, gdy $Z(Q)$ jest prawdziwe w $\mathfrak{M}'_2$.*

A zatem, jeśli $Z(Q)$ jest prawdziwe (resp. fałszywe) w pewnym modelu $\mathfrak{M}'$, to pozostanie też prawdziwe (resp. fałszywe) w każdym modelu, który różni się od $\mathfrak{M}'$ jedynie denotacją terminu Q — o ile tylko w obu tych modelach denotacje terminu Q spełniają warunek sformułowany w jego definicji. Otóż okazuje się, iż określoną wartość logiczną ze względu na Q posiadają wszystkie i tylko te zdania $Z(Q)$, z których Q jest eliminowane. *Warunki* (EL) *i* (OL) *są więc wzajemnie równoważne.*

Oto szkic dowodu tej zależności: Wynikanie warunku (OL) z warunku (EL) jest oczywiste. Warunek (EL) stwierdza, że istnieje takie zdanie Z nie zawierające Q, iż w dowolnym modelu $\mathfrak{M}'$, w którym prawdziwe jest $D(Q)$, zdanie $Z(Q)$ ma tę samą wartość logiczną, co zdanie Z. A zatem, jeżeli weźmiemy dowolne modele $\mathfrak{M}'_1$ i $\mathfrak{M}'_2$, które różnią się tylko denotacją Q, a więc w których zdanie Z ma tę samą wartość logiczną, i w których ponadto prawdziwe jest $D(Q)$, to zdanie $Z(Q)$ musi mieć w obu tych modelach identyczną wartość logiczną: tę samą, którą ma Z. A to właśnie stwierdza warunek (OL). Chcąc okazać, iż z warunku (OL) wynika warunek (EL), weźmy jako model $\mathfrak{M}'_2$ wymieniony w warunku (OL), model, w którym denotacją terminu Q byłby zbiór identyczny z tym, który w modelach $\mathfrak{M}'_1$ i $\mathfrak{M}'_2$ stanowi denotację predykatu Φ. Definicja $D(Q)$ musi być zdaniem prawdziwym w tak określonym modelu $\mathfrak{M}'_2$, wobec czego odpowiednie założenie w sformułowaniu warunku (OL) może zostać pominięte. Jednocześnie stwierdzenie, iż $Z(Q)$ jest prawdziwe w tak określonym modelu $\mathfrak{M}'_2$, jest równoważne stwierdzeniu, iż $Z(\Phi)$ jest prawdziwe w modelu $\mathfrak{M}'_1$, i może być przez to ostatnie zastąpione[9]. Otrzymujemy w ten sposób następującą konsekwencję — oznaczamy ją dla skrótu (EL*):

---

9) Można okazać to wyraźniej dla przypadku takiego, w którym Ψ i Φ są prostymi predykatami języka J, np. P_1 i P_2, a więc w którym definicja $D(Q)$ przybiera postać $\bigwedge_x [P_1x \rightarrow (Qx \equiv P_2x)]$. Jeżeli modelem $\mathfrak{M}'_1$ jest układ $\langle U, x_1, ..., x_n, X_1, ..., X_m, Y\rangle$, jako model $\mathfrak{M}'_2$ przyjmiemy układ $\langle U, x_1, ..., x_n, X_1, ..., X_m, X_2\rangle$. Wówczas założenie, iż $D(Q)$ jest prawdziwe w $\mathfrak{M}'_2$, jest równoważne tautologii $\bigwedge_{x \in U} [x \in X_1 \rightarrow (x \in X_2 \equiv x \in X_2)]$, i jako takie może być pominięte. Natomiast stwierdzenie, iż $Z(Q)$ jest prawdziwe w $\mathfrak{M}'_2$, jest równoważne stwierdzeniu, iż $Z(P_2)$ jest prawdziwe w $\mathfrak{M}'_1$ i może być przez nie zastąpione. Weźmy dla przykładu jako $Z(Q)$ zdanie $P_1a_1 \wedge Qa_1$. $P_1a_1 \wedge Qa_1$ jest prawdziwe w $\mathfrak{M}'_2$ wtedy i tylko wtedy, gdy $x_1 \in X_1 \wedge x_1 \in X_2$; a zatem — wtedy i tylko wtedy, gdy $P_1a_1 \wedge P_2a_2$ jest prawdziwe w $\mathfrak{M}'_1$.

(EL*) *Dla dowolnego modelu $\mathfrak{M}'$: jeżeli $D(Q)$ jest prawdziwe w $\mathfrak{M}'$, to $Z(Q)$ jest prawdziwe w $\mathfrak{M}'$ wtedy i tylko wtedy, gdy $Z(\Phi)$ jest prawdziwe w $\mathfrak{M}'$; lub krócej:*

$$D(Q) \rightarrow (Z(Q) \equiv Z(\Phi))$$

jest prawdziwe w każdym modelu $\mathfrak{M}'$.

Ponieważ $Z(\Phi)$ jest zdaniem nie zawierającym terminu Q, warunek (EL*) pociąga za sobą warunek (EL). Tak więc wszystkie i tylko te zdania $Z(Q)$, które spełniają warunek (EL), spełniają warunek (OL).

Pojęcie zdania $Z(Q)$ spełniającego warunek (OL) zbliża się w pewnym stopniu do pojęcia zdania „zdeterminowanego" wprowadzonego do analogicznych rozważań przez H. Mehlberga[10]. Okazuje się więc, iż to ostatnie pokrywa się w przybliżeniu z pojęciem zdania $Z(Q)$, z którego Q daje się wyeliminować na podstawie definicji $D(Q)$. Jednocześnie przeprowadzony dowód pokazuje, iż warunek „eliminowalności" (EL) możemy zastąpić pozornie mocniejszym warunkiem (EL*), Nie tylko bowiem — co jest rzeczą oczywistą — warunek (EL*) pociąga warunek (EL), ale i na odwrót: (EL) pociąga (EL*), gdyż — jak wykazaliśmy przed chwilą — warunek (EL) pociąga warunek (OL), a ten ostatni pociąga z kolei warunek (EL*). Jeśli więc $D(Q)$ ma, jak zakładaliśmy, postać

$$\bigwedge_x [\Psi x \rightarrow (Qx \equiv \Phi x)],$$

to każde zdanie $Z(Q)$, z którego Q jest na podstawie tej definicji eliminowalne, jest na jej gruncie równoważne zdaniu $Z(\Phi)$, powstającemu z poprzedniego przez zastąpienie predykatu Q predykatem Φ należącym do języka J.

3. Wprowadzone dotychczas pojęcia „eliminowalności" (EL) i „określoności" (OL) nazwać można pojęciami „absolutnymi". Prócz nich zdefiniować można odpowiednie pojęcia o charakterze „relatywnym" — i to o dwojakiej co najmniej relatywizacji. Przede wszystkim — pojęcia „eliminowalności" (ET) i „określoności" (OT) zrelatywizowane do zbioru zdań T języka J. Ten zbiór zdań utożsamić możemy z pewną teorią sformułowaną w języku J, jak to się najczęściej przy zastosowaniach tego rodzaju pojęć czyni. Powiedzmy zatem, iż

(ET) *Q jest e l i m i n o w a l n e z $Z(Q)$ n a p o d s t a w i e d e f i n i c j i $D(Q)$ w t e o r i i T wtedy i tylko wtedy, gdy istnieje zdanie Z nie zawierające Q, takie iż zdanie*

$$D(Q) \rightarrow (Z(Q) \equiv Z)$$

jest prawdziwe w każdym modelu $\mathfrak{M}'$, w którym prawdziwa jest teoria T.

Jest rzeczą widoczną, iż warunek powyższy jest pewnym uogólnieniem warunku poprzedniego. Gdy zbiór T składa się wyłącznie z tautologii języka J, warunek (ET) sprowadza się do warunku (EL). W przypadku ogólnym jednak, gdy T obejmuje zdania nie będące tautologiami, klasa zdań $Z(Q)$ spełniających warunek (ET) obejmuje nie tylko wszystkie zdania $Z(Q)$ spełniające warunek (EL), ale i pewne zdania $Z(Q)$, które

10) W książce: *The Reach of Science*, Toronto 1958. Dokładny odpowiednik pojęcia zdania „zdeterminowanego" wprowadzam na dalszych stronach.

tamtego warunku nie spełniają. I tak np. zdanie Qa_1 spełnia warunek (ET), jeśli tylko zdanie Ψa_1 należy do teorii T; podobnie jest ze zdaniem $\bigvee_x(P_1x \wedge Qx)$ pod warunkiem, iż twierdzeniem teorii T jest zdanie $\bigwedge_x(P_1x \rightarrow \Psi x)$. W wypadku zaś, gdy teoria T obejmuje zdanie $\bigwedge_x \Psi x$ termin Q jest eliminowany z każdego zdania $Z(Q)$ w tak scharakteryzowanej teorii, gdyż na jej gruncie definicja $D(Q)$ staje się równoważna zwykłej definicji równoważnościowej.

Pojęciu „eliminowalności" (ET) odpowiada tak samo zrelatywizowane pojęcie „określoności" (OT).

(OT) *$Z(Q)$ ma określoną wartość logiczną ze względu na Q na podstawie definicji $D(Q)$ w teorii T wtedy i tylko wtedy, gdy dla dowolnych modeli $\mathfrak{M}_1'$ i $\mathfrak{M}_2'$, które różnią się co najwyżej denotacją terminu Q i w których prawdziwa jest teoria T, zachodzi zależność następująca: jeżeli $D(Q)$ jest prawdziwe w $\mathfrak{M}_1'$ oraz $D(Q)$ jest prawdziwe w $\mathfrak{M}_2'$, to $Z(Q)$ jest prawdziwe w $\mathfrak{M}_1'$ wtedy i tylko wtedy, gdy $Z(Q)$ jest prawdziwe w $\mathfrak{M}_2'$.*

Podobnie, jak warunki (EL) i (OL), warunku (ET) i (OT) są wzajemnie równoważne. Dowód ich równoważności przebiega analogicznie do poprzedniego.

Ważniejszą jednak rolę niż pojęcia „eliminowalności" i „określoności" zrelatywizowane do zbioru zdań T języka J odgrywają w dalszych rozważaniach analogiczne pojęcia — oznaczmy je (EM) i (OM) — zrelatywizowane do modelu $\mathfrak{M}$ języka J. Formułując ich definicje, posłużymy się pojęciem „wzbogacenia" danego modelu. Ograniczając się do opisanych języków J i J', model $\mathfrak{M}'$ języka J' nazywać możemy *wzbogaceniem* modelu $\mathfrak{M}$ języka J, jeśli universa i denotacje terminów wspólnych dla obu języków J i J' są w obu modelach $\mathfrak{M}$ i $\mathfrak{M}'$ identyczne. I tak, jeśli

$$\mathfrak{M} = \langle U, x_1, ..., x_n, X_1, ..., X_m \rangle$$

stanowi model języka J, model

$$\mathfrak{M}' = \langle U, x_1, ..., x_n, X_1, ..., X_m, Y \rangle$$

języka J' jest wzbogaceniem modelu $\mathfrak{M}$. Definicja warunku (EM) głosi obecnie, iż

(EM) *Q jest eliminowalne z $Z(Q)$ na podstawie definicji $D(Q)$ w modelu $\mathfrak{M}$ wtedy i tylko wtedy, gdy istnieje zdanie Z nie zawierające Q, takie iż zdanie*

$$D(Q) \rightarrow (Z(Q) \equiv Z)$$

jest prawdziwe w każdym modelu $\mathfrak{M}'$, stanowiącym wzbogacenie modelu $\mathfrak{M}$.

Podobnie, jak w przypadku warunku (ET), klasa zdań $Z(Q)$ spełniających warunek (EM) obejmuje oprócz zdań spełniających warunek (EL) pewne zdania, które tamtego warunku nie spełniają. Przykłady można tu podać analogiczne do poprzednich. Przyjmijmy jako model $\mathfrak{M}$ języka J określony układ

$$\langle \boldsymbol{U}, \boldsymbol{a}_1, ..., \boldsymbol{a}_n, \boldsymbol{P}_1, ..., \boldsymbol{P}_m \rangle.$$

Niech $\overline{\Psi}$ będzie denotacją w modelu $\mathfrak{M}$ predykatu Ψ. Zdanie Qa_1 spełnia warunek (EM), jeśli $\boldsymbol{a}_1 \in \overline{\Psi}$; zdanie $\bigvee_x (P_1x \wedge Qx)$ — jeśli $\bigwedge_{x \in U} (x \in \boldsymbol{P}_1 \rightarrow x \in \overline{\Psi})$. Gdyby zaś prawdą było, iż $\bigwedge_{x \in U}(x \in \overline{\Psi})$, termin Q byłby eliminowany z każdego zdania $Z(Q)$ w modelu $\mathfrak{M}$, gdyż w modelu tym definicja $D(Q)$ równoważna byłaby zwykłej definicji równoważnościowej.

Pojęciu (EM) odpowiada równoważne mu zakresowo pojęcie (OM).

(OM) *$Z(Q)$ ma określoną wartość logiczną ze względu na Q na podstawie definicji $D(Q)$ w modelu $\mathfrak{M}$ wtedy i tylko wtedy, gdy dla dowolnych modeli $\mathfrak{M}_1'$ i $\mathfrak{M}_2'$ stanowiących wzbogacenia modelu $\mathfrak{M}$ zachodzi zależność następująca: jeżeli $D(Q)$ jest prawdziwe w $\mathfrak{M}_1'$, oraz $D(Q)$ jest prawdziwe w $\mathfrak{M}_2'$ to $Z(Q)$ jest prawdziwe w $\mathfrak{M}_1'$ wtedy i tylko wtedy, gdy $Z(Q)$ jest prawdziwe w $\mathfrak{M}_2'$.*

Wzajemnej równoważności warunków (EM) i (OM) dowieść można w sposób analogiczny do dowodu równoważności warunków (EL) i (OL). Warunek (OM) jest tym pojęciem, które pokrywa się ściśle z pojęciem zdania „zdeterminowanego" u Mehlberga. Jako przykład zdania „niezdeterminowanego" przytacza się tam zdanie typu Qa_1, gdy nieprawdą jest, iż a_1 należy do $\overline{\Psi}$. Wtedy to bowiem zależnie od takiej lub innej — zgodnej z definicją $D(Q)$ — interpretacji Q, zdanie to staje się prawdziwe lub fałszywe. Mówiąc ogólnie, zdanie $Z(Q)$ jest „niezdeterminowane", gdy przy danej interpretacji terminów pozostałych zdanie to zmienia swoją wartość logiczną w zależności od tego, jaką interpretację — zgodną z definicją $D(Q)$ — nadamy terminowi Q. A to właśnie jest cechą charakterystyczną zdań $Z(Q)$, które nie spełniają warunku (OM).

Chciałbym w tym miejscu zwrócić uwagę na ważną różnicę, jaka dzieli ostatnią parę pojęć (EM) i (OM) od pojęć poprzednich. Aby rozstrzygnąć, czy dane zdanie $Z(Q)$ spełnia warunek (EL), trzeba rozstrzygnąć, czy zdanie

(1) $$D(Q) \rightarrow (Z(Q) \equiv Z(\Phi))$$

wynika logicznie z pustej klasy zdań[11]. W przypadku, dajmy na to, zdania Qa_1 okazuje się, iż tak nie jest. Aby rozstrzygnąć, czy dane zdanie $Z(Q)$ spełnia warunek (ET), trzeba rozstrzygnąć, czy zdanie (1) wynika logicznie ze zdań zbioru T. W przypadku zdania Qa_1 tak jest wtedy, gdy ze zdań zbioru T wynika logicznie zdanie Ψa_1. Rozstrzygnięcie powyższych pytań nie wymaga odwołania się do doświadczenia. W rezultacie stwierdzenie, iż dane zdanie $Z(Q)$ spełnia warunek (EL) czy (ET), ma zawsze charakter zdania analitycznego. Inaczej przedstawia się sprawa warunku (EM). Aby rozstrzygnąć, czy dane zdanie $Z(Q)$ spełnia warunek (EM), trzeba rozstrzygnąć, czy zdanie (1) jest prawdziwe przy dowolnej interpretacji Q i przy tej interpretacji pozostałych terminów, które wyznacza dany model $\mathfrak{M}$. W przypadku zdania Qa_1 jest tak wtedy, gdy przedmiot denotowany w modelu $\mathfrak{M}$ przez a_1 należy do zbioru denotowanego w modelu $\mathfrak{M}$ przez Ψ. Ale to, czy tak jest, czy nie, może być sprawą doświad-

11) Jest to sformułowanie równoważne sformułowaniu warunku (EL) przytoczonemu w tekście. Podobna uwaga dotyczy warunków (ET) i (EM).

czenia. A zatem stwierdzenie, iż dane zdanie $Z(Q)$ spełnia warunek (EM), może mieć charakter zdania syntetycznego, wymagającego dla swego uzasadnienia czy obalenia odwołania się do doświadczenia. Taka sama różnica zachodzi oczywiście pomiędzy warunkami (OL) i (OT) a (OM). Powrócimy do niej jeszcze w dalszym toku rozważań.

Wprowadzone przez nas warunki „eliminowalności" i „określoności" pozostają w ścisłym związku z warunkami rozstrzygalności zdań typu $Z(Q)$. Rozstrzygnąć bowiem dane zdanie — to uzasadnić bądź samo to zdanie, bądź jego negację, czyli — swobodnie mówiąc — okazać, iż zdanie to jest prawdą, lub okazać, iż zdanie to jest fałszem. Jeśli jednak zdanie $Z(Q)$ nie ma określonej wartości logicznej, a więc zależnie od takiej czy innej dopuszczalnej interpretacji terminu Q staje się raz prawdą, raz fałszem, to bez jakichś dodatkowych założeń tego właśnie uczynić nie możemy. A zatem jeśli zdanie $Z(Q)$ ma być rozstrzygalne, musi spełniać warunek „określoności" (a tym samym i „eliminowalności"). Zależność odwrotna zachodzi przy założeniu, iż zdania języka J, a więc zdania nie zawierające terminu Q, są rozstrzygalne. Jeśli wówczas zdanie $Z(Q)$ spełnia warunek „eliminowalności", a więc jest równoważne pewnemu zdaniu języka J, musi być zdaniem rozstrzygalnym. Różnym rodzajom pojęć „eliminowalności" („określoności") odpowiadają różne pojęcia rozstrzygalności: L-rozstrzygalność, T-rozstrzygalność, M-rozstrzygalność. Idzie tu, ściślej mówiąc, o różne rodzaje sprowadzalności danego zdania do zdań rozstrzygalnych. I tak, zdanie $Z(Q)$ jest *L-rostrzygalne*, gdy zdanie to jest sprowadzalne do pewnego zdania rozstrzygalnego na podstawie samej definicji Q; zdanie $Z(Q)$ jest *T-rozstrzygalne*, gdy jest sprowadzalne do pewnego zdania rozstrzygalnego na podstawie definicji $D(Q)$ w teorii T; wreszcie zdanie $Z(Q)$ jest *M-rozstrzygalne*, gdy jest sprowadzalne do pewnego zdania rozstrzygalnego na podstawie definicji $D(Q)$ w modelu $\mathfrak{M}$, a więc gdy jest równoważne pewnemu zdaniu rozstrzygalnemu przy dowolnej — zgodnej z definicją $D(Q)$ — interpretacji terminu Q i przy tej interpretacji terminów pozostałych, którą wyznacza model $\mathfrak{M}$. Do tego ostatniego pojęcia M-rozstrzygalności odwoływać się będziemy głównie w dalszej dyskusji. Wspomniane wyżej zdania „niezdeterminowane" — to zdania nierozstrzygalne w tym właśnie sensie. Jako przykład zdania $Z(Q)$ rozstrzygalnego w każdym z powyższych znaczeń przytoczyć można zdanie $\Psi a_1 \wedge Qa_1$. Natomiast zdanie Qa_1, nie będąc L-rozstrzygalne, jest T-rozstrzygalne, gdy zdanie Ψa_1 jest twierdzeniem teorii T, a — M-rozstrzygalne, gdy zdanie Ψa_1 jest prawdziwe w modelu $\mathfrak{M}$.

II

4. Uzyskaliśmy w ten sposób odpowiedź na pierwsze z postawionych na wstępie pytań — charakterystykę klasy owych problematycznych pod względem semantycznym wypowiedzi zawierających pewien zdefiniowany warunkowo termin. Jeśli terminem tym jest predykat Q, a jego definicją definicja warunkowa $D(Q)$, to klasa owa jest identyczna z klasą tych wszystkich zdań języka J' zawierających termin Q, które nie spełniają warunku (OM) (lub, co na jedno wychodzi, warunku (EM)). Zdania te odznaczają się, jak wiemy, tym, iż przy danej, wyznaczonej przez określony model $\mathfrak{M}$,

interpretacji języka J zmieniają swą wartość logiczną w zależności od takiej, czy innej, zgodnej z definicją $D(Q)$, interpretacji terminu Q. Powstaje w tej sytuacji pytanie, czy zdania takie mamy prawo uważać za wypowiedzi prawdziwe lub fałszywe. A jeśli tak, to w jakim mianowicie sensie? W związku z tym pozostają pytania dalsze. O czym się właściwie w zdaniach tych mówi? Czy termin Q posiada jakąś denotację? A jeśli tak, to jaką?

Rozważmy przede wszystkim zagadnienie prawdziwości owych zdań $Z(Q)$, nie spełniających warunku (OM). Należy na wstępie zwrócić uwagę na fakt, iż w naszych dotychczasowych rozważaniach nie posługiwaliśmy się w ogóle „absolutnym" pojęciem prawdziwości. Jedyne pojęcie prawdy, z jakiego do tej pory czyniliśmy użytek, to — „relatywne" pojęcie prawdziwości w modelu. Nie posługiwaliśmy się w ogóle zwrotem: zdanie Z jest prawdziwe, lecz wyłącznie zwrotem: zdanie Z jest prawdziwe w modelu $\mathfrak{M}$[12]. Ten ostatni — przypominam — rozumiany był tak, iż pociągał m.in. równoważność następującą:

Zdanie P_1a_1 jest prawdziwe w modelu
$$\mathfrak{M} = \langle U, a_1, ..., a_n, P_1, ..., P_m\rangle$$
wtedy i tylko wtedy, gdy $a_1 \in P_1$,

oraz analogiczne równoważności dla pozostałych zdań języka J. Przejście do „absolutnego" pojęcia prawdziwości polega na wyborze określonego modelu $\mathfrak{M}^*$ języka J i zdefiniowaniu zdania prawdziwego jako zdania prawdziwego w modelu $\mathfrak{M}^*$. Rodzina modeli $\mathfrak{M}$ języka J obejmuje wszystkie układy przedmiotów, o których można mówić w języku J. Zakłada się, iż jeden z tych układów jest tym, o którym faktycznie mówi się w języku J. Model ten, $\mathfrak{M}^*$, stanowi tzw. *model właściwy* języka J. Jest to model, który dostarcza przekładu wyrażeń języka J na metajęzyk MJ. Interpretacja wyrażeń języka J wyznaczona przez model właściwy $\mathfrak{M}^*$ stanowi przekład tych wyrażeń na język, którym mówimy sami opisując język J. „Absolutne" pojęcie prawdziwości — w przeciwieństwie do „relatywnego" pojęcia prawdziwości w modelu — ma zastosowanie tylko do języka zinterpretowanego, a więc do takiego języka J, dla którego dany jest jego model właściwy $\mathfrak{M}^*$. Jeśli dla rozważanego przez nas języka J dany jest jako jego model właściwy $\mathfrak{M}^*$ podany wyżej układ

$$\langle U, a_1, ..., a_n, P_1, ..., P_m\rangle,$$

możemy stwierdzić po prostu, iż

zdanie P_1a_1 jest prawdziwe wtedy i tylko wtedy, gdy $a_1 \in P_1$.

Analogiczne równoważności mają walor dla pozostałych zdań języka J[13].

W rezultacie pojęcia semantyki logicznej w rodzaju „absolutnego" pojęcia prawdziwości stosowane były w dotychczasowych badaniach wyłącznie do języków posiadających określoną, jednoznaczną interpretację, języków całkowicie zinterpretowanych

12) Pomijamy tu dla uproszczenia — i w dalszym ciągu pomijać będziemy — konieczną relatywizację do języka: zdanie Z jest prawdziwe w języku J, *resp.* — w modelu $\mathfrak{M}$ języka J.

13) Tak rozumiane pojęcie modelu właściwego i zdania prawdziwego występuje w cytowanej pracy R. Suszki oraz w pracy J. Kemeny'ego: „A New Approach to Semantics", *Journal of Symbolic Logic* 21 (1956).

(lub — w terminologii Kemeny'ego — semantycznie zdeterminowanych). Jak pod tym względem przedstawia się charakter rozważanych przez nas języków J i J'? Próba odpowiedzi na to pytanie opierać się będzie na pewnych założeniach stanowiących idealizację faktycznie istniejącej sytuacji. Zakładamy przede wszystkim, iż język J nie odbiega od typu języków uwzględnianych do tej pory w badaniach semantycznych. Jest to język całkowicie zinterpretowany. Dla języka J dany jest więc jednoznacznie określony model właściwy $\mathfrak{M}^*$. Niechaj będzie nim, jak poprzednio, układ

$$\langle U, a_1, \ldots, a_n, P_1, \ldots, P_m \rangle .$$

Możemy przyjąć wobec tego, że dla dowolnego zdania Z języka J:

Z jest prawdziwe wtedy i tylko wtedy, gdy
Z jest prawdziwe w modelu $\mathfrak{M}^*$.

Warto w tym miejscu dodać pewne wyjaśnienie. Gdy w toku rozważań poprzednich mówiliśmy po prostu o tym, iż dane zdanie $Z(Q)$ spełnia warunek (OM) (czy (EM)) lub go nie spełnia — model $\mathfrak{M}$ języka J, do którego milcząco relatywizowaliśmy ów warunek, utożsamialiśmy z modelem właściwym $\mathfrak{M}^*$ języka J.

Nasuwa się pytanie, w jaki sposób dla języka J wyznaczony zostaje jego model właściwy $\mathfrak{M}^*$. I dlaczego w stosunku do języka J takie założenie przyjmujemy? Otóż, jak wspominaliśmy na wstępie, wzorcem analizowanych przez nas języków są języki empirycznych teorii naukowych. Terminy specyficzne takich teorii dzieli się zwykle na dwa rodzaje: na terminy spostrzeżeniowe i teoretyczne. Jeśli język J' utożsamimy z językiem pewnej teorii empirycznej, to język J stanowić będzie jego część spostrzeżeniową. Terminy: $a_1, \ldots, a_n, P_1, \ldots, P_m$, należeć będą do terminów spostrzeżeniowych, a termin Q — do teoretycznych. Terminy języka J jako terminy spostrzeżeniowe, a więc odnoszące się do przedmiotów spostrzegalnych, dopuszczają — jak się na ogół przyjmuje — interpretację bezpośrednią. Przedmioty mające stanowić ich denotacje można przyporządkować im przez bezpośrednie ich wskazanie, czyli za pomocą reguł semantycznych o charakterze definicji ostensywnych. Założenie, iż jest to przyporządkowanie jednoznaczne, stanowi oczywiście idealizację stanu faktycznego. Tutaj jednakże możemy takie upraszczające założenie uczynić i przyjąć, że na tej drodze jednoznaczne wyznaczenie modelu właściwego $\mathfrak{M}^*$ języka J jest możliwe[14].

Zgoła inaczej przedstawia się charakter języka J'. Termin Q jako termin teoretyczny, a więc odnoszący się do własności niespostrzegalnej, bezpośredniej interpretacji nie dopuszcza. Interpretacja terminu Q zdeterminowana jest jedynie przez jego definicję $D(Q)$, odwołującą się do terminów spostrzeżeniowych o interpretacji wyznaczonej przez model właściwy $\mathfrak{M}^*$ języka J. W tej sytuacji model właściwy $\mathfrak{M}'^*$ języka J' scharakteryzowany zostaje wyłącznie przez następujące dwa warunki:

14) Język teorii empirycznych omawiam w pracy: „Pojęcia teoretyczne a doświadczenie", cyt. wyd.; sposoby interpretacji terminów specyficznych teorii empirycznych — w pracy: „Interpretacja systemów aksjomatycznych", *Studia Filozoficzne* 21 (1960); interpretację terminów spostrzeżeniowych — w pracy: „O definiowaniu terminów spostrzeżeniowych", *Rozprawy logiczne: Księga pamiątkowa ku czci profesora K. Ajdukiewicza*, Warszawa 1964.

1) model $\mathfrak{M}'^*$ stanowi wzbogacenie modelu $\mathfrak{M}^*$;

2) definicja $D(Q)$ jest prawdziwa w modelu $\mathfrak{M}'^*$.

Model właściwy $\mathfrak{M}'^*$ języka J' musi być więc modelem takim, w którym universum i denotacje terminów należących do języka J są takie same, jak w modelu właściwym $\mathfrak{M}^*$ języka J, a denotacja terminu Q czyni zadość warunkowi sformułowanemu w definicji $D(Q)$. Czy model $\mathfrak{M}'^*$ zostaje w ten sposób wyznaczony jednoznacznie? Tak byłoby tylko wtedy, gdyby definicja $D(Q)$:

$$\bigwedge_x [\Psi x \to (Qx \equiv \Phi x)]$$

była równoważna w modelu $\mathfrak{M}^*$ zwykłej definicji równoważnościowej. To zaś z kolei byłoby możliwe tylko wtedy, gdyby zbiór stanowiący w modelu $\mathfrak{M}^*$ denotację predykatu Ψ pokrywał się ze zbiorem U stanowiącym universum modelu $\mathfrak{M}^*$. Ponieważ interesuje nas tutaj sytuacja charakterystyczna dla pojęć otwartych — takie właśnie pojęcia odpowiadają terminom teoretycznym teorii empirycznych — zakładamy, iż tak nie jest. Przy tym założeniu model $\mathfrak{M}'^*$ wyznaczony zostaje w sposób niejednoznaczny. Warunki (1) i (2) definiują pewną rodzinę $\mathfrak{R}\mathfrak{M}'$ modeli języka J', obejmującą więcej niż jeden model tego języka. Model właściwy $\mathfrak{M}'^*$ jest jednym z modeli należących do tej rodziny:

$$\mathfrak{M}'^* \in \mathfrak{R}\mathfrak{M}'.$$

Mamy tu zatem do czynienia z sytuacją odmienną od tych, jakie uwzględniano w dotychczasowych badaniach semantycznych. Język J' jest językiem częściowo zinterpretowanym. Jaki więc sens możemy wiązać tu ze zwrotem orzekającym o dowolnym zdaniu Z języka J', iż jest to zdanie prawdziwe?

Zanim przejdziemy do przedstawienia różnych prób odpowiedzi na to pytanie, rozważmy dla przykładu charakter semantyczny paru prostych zdań języka J'. Przyjmijmy w tym celu szereg upraszczających założeń, z których i w dalszych przykładach będziemy stale korzystali. Niech definicja $D(Q)$ terminu Q przybierze postać definicji cząstkowej typu (3) — postać typową zresztą dla definicji terminów teoretycznych:

$$\bigwedge_x [(P_1 x \to Qx) \wedge (P_2 x \to \sim Qx)]$$

Załóżmy, iż zbiory P_1 i P_2 wyłączają się wzajemnie, a ich suma nie wyczerpuje universum U. Wówczas rodzina modeli języka J', $\mathfrak{R}\mathfrak{M}'$, obejmuje wszystkie i tylko takie modele

$$\langle U, a_1, ..., a_n, P_1, ..., P_m, Y\rangle,$$

w których

$$P_1 \subset Y \subset -P_2{}^{15}.$$

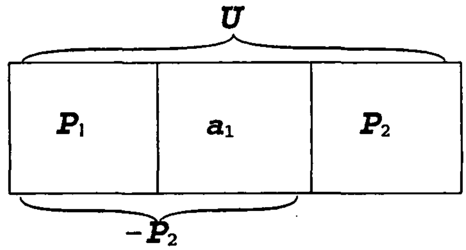
Jeśli założymy ponadto dla dalszego uproszczenia, iż istnieje tylko jeden element universum U nie należący ani do zbioru P_1, ani do zbioru P_2, rodzina $\mathfrak{R}\mathfrak{M}'$ obejmie dokładnie dwa modele:

$$\mathfrak{M}_1' = \langle U, a_1, ..., a_n, P_1, ..., P_m, P_1\rangle,$$

15) Zbiór $-P_1$ jest dopełnieniem zbioru P_2 (do universum U).

$$\mathfrak{M}_2' = \langle U, a_1, ..., a_n, P_1, ..., P_m, -P_2\rangle.$$

Model właściwy $\mathfrak{M}'^*$ jest zatem identyczny z jednym z nich: $\mathfrak{M}'^* = \mathfrak{M}_1'$ lub $\mathfrak{M}'^* = \mathfrak{M}_2'$. Umówmy się jeszcze, że a_1 będzie tym przedmiotem, który nie należy ani do P_1, ani do P_2, a a_2 — przedmiotem różnym od a_1, a więc należącym bądź do P_1, bądź do P_2. Stosunki te ilustruje następujący diagram:



Weźmy pod uwagę trzy zdania języka J' zawierające termin Q:

(a) $$Qa_1 \vee \sim Qa_1$$

(b) $$Qa_2$$

(c) $$Qa_1$$

i określmy ich wartość logiczną w modelach rodziny $\mathfrak{R}\mathfrak{M}'$. Dwa pierwsze reprezentują klasę zdań $Z(Q)$ spełniających warunek (OM), trzecie — klasę zdań $Z(Q)$ nie spełniających tego warunku. Zdanie (a) jako tautologia jest prawdziwe zarówno w modelu $\mathfrak{M}_1'$, jak i $\mathfrak{M}_2'$. Zdanie (b) ma tę samą wartość logiczną w obu modelach niezależną od tego, czy Q denotuje zbiór P_1, czy $-P_2$. Jeśli $a_2 \in P_1$, zdanie (b) jest prawdziwe zarówno w modelu $\mathfrak{M}_1'$, jak i $\mathfrak{M}_2'$.; jeśli zaś $a_2 \in P_2$, zdanie (b) jest fałszywe zarówno w modelu $\mathfrak{M}_1'$, jak i $\mathfrak{M}_2'$. Natomiast zdanie (c) przybiera różną wartość logiczną w zależności od tego, czy denotacją Q jest zbiór P_1, czy $-P_2$. Zdanie (c) jest fałszywe w modelu $\mathfrak{M}_1'$, a prawdziwe w modelu $\mathfrak{M}_2'$.

5. Istnienie w języku J' zdań tego ostatniego typu, czyli zdań $Z(Q)$ nie spełniających warunku (OM), sprawia zasadniczą trudność przy próbach zdefiniowania w zastosowaniu do języka J' „absolutnego" pojęcia prawdziwości. Przedstawimy obecnie szereg możliwości, jakie się w tej sprawie zarysowują, starając się uwzględnić główne typy rozwiązań. Będą to na ogół stanowiska, które są w dyskusjach nad tym problemem faktycznie reprezentowane, choć z reguły w szkicowej postaci i w sformułowaniach odbiegających od proponowanych. Idzie tu jednak o to, aby owe rozwiązania umieścić w ramach precyzyjnej i jednolitej aparatury pojęciowej współczesnej semantyki logicznej. Pięć uwzględnionych przez nas stanowisk reprezentuje pięć głównych typów odpowiedzi na pytanie, na czym polega prawdziwość czy fałszywość zdań języka J', dla którego dana jest jedynie rodzina modeli $\mathfrak{R}\mathfrak{M}'$. Niech Z będzie dowolnym zdaniem języka J'. Wspomniane stanowiska głoszą:

(I) *Z jest prawdziwe wtedy i tylko wtedy, gdy Z jest prawdziwe w każdym modelu* $\mathfrak{M}'$ *należącym do rodziny* $\mathfrak{R}\mathfrak{M}'$;

Z jest fałszywe wtedy i tylko wtedy, gdy Z jest fałszywe w każdym modelu $\mathfrak{M}'$ *należącym do rodziny* $\mathfrak{R}\mathfrak{M}'$.

(II) *Z jest prawdziwe wtedy i tylko wtedy, gdy Z jest prawdziwe w pewnym modelu* $\mathfrak{M}'$ *należącym do rodziny* $\mathfrak{R}\mathfrak{M}'$;

Z jest fałszywe wtedy i tylko wtedy, gdy Z jest fałszywe w każdym modelu $\mathfrak{M}'$ *należącym do rodziny* $\mathfrak{R}\mathfrak{M}'$.

(III) *Z jest prawdziwe wtedy i tylko wtedy, gdy Z jest prawdziwe w każdym modelu* $\mathfrak{M}'$ *należącym do rodziny* $\mathfrak{R}\mathfrak{M}'$;

Z jest fałszywe wtedy i tylko wtedy, gdy Z jest fałszywe w pewnym modelu $\mathfrak{M}'$ *należącym do rodziny* $\mathfrak{R}\mathfrak{M}'$.

(IV) *Z jest prawdziwe wtedy i tylko wtedy, gdy Z jest prawdziwe w modelu* $\mathfrak{M}'_i$;

Z jest fałszywe wtedy i tylko wtedy, gdy Z jest fałszywe w modelu $\mathfrak{M}'_i$;

gdzie $\mathfrak{M}'_i$ jest określonym modelem wybranym — na podstawie dodatkowych, omówionych w dalszym ciągu pracy, założeń — z rodziny $\mathfrak{R}\mathfrak{M}'$.

(V) *Z jest prawdziwe wtedy i tylko wtedy, gdy Z jest prawdziwe w modelu* $\mathfrak{M}'^*$;

Z jest fałszywe wtedy i tylko wtedy, gdy Z jest fałszywe w modelu $\mathfrak{M}'^*$;

gdzie $\mathfrak{M}'^*$ jest modelem właściwym języka J' scharakteryzowanym wyłącznie przez warunek: $\mathfrak{M}'^* \in \mathfrak{R}\mathfrak{M}'$.

Przechodząc do omówienia powyższych stanowisk, rozpatrzmy na wstępie, jak wypadnie nam zakwalifikować zgodnie z każdym z nich wybrane przykładowo zdania (a), (b) i (c) co do ich „absolutnej" wartości logicznej. Kwalifikacja ta w stosunku do zdań (a) i (b) jest zgodna. Wedle wszystkich pięciu stanowisk zdanie (a) jest prawdziwe, zdanie (b) jest prawdziwe lub fałszywe. Różnice tych stanowisk uwidoczniają się dopiero w stosunku do zdania (c). Wedle (I) zdanie (c) nie jest ani prawdziwe, ani fałszywe; wedle (II) jest prawdą; wedle (III) — fałszem. Zgodnie z (IV) zdanie (c) ma określoną wartość logiczną, zależną od tego, który to model z rodziny $\mathfrak{R}\mathfrak{M}'$ jest modelem $\mathfrak{M}'_i$; w dalszym toku rozważań okaże się, że w omawianym przykładzie jest to model taki, iż zdanie (c) jest fałszem. Wreszcie zgodnie z (V) zdanie (c) jest prawdziwe lub fałszywe, ale — wobec niejednoznacznej charakterystyki modelu $\mathfrak{M}'^*$ — rozstrzygnięcie, która z tych ewentualności zachodzi, jest rzeczą niemożliwą. Powyższa ocena „absolutnej" wartości logicznej wybranych zdań języka J' nasuwa pewne ogólne wnioski co do charakteru każdego z wyszczególnionych stanowisk.

Oto zgodnie z każdym z tych stanowisk wszystkie zdania języka J' nie zawierające Q oraz te zdania $Z(Q)$, które spełniają warunek (OM), posiadają określoną wartość logiczną: są prawdziwe lub fałszywe, a przy tym w zasadzie rozstrzygalne. Również zgodnie z każdym z powyższych stanowisk wszystkie tautologie języka J', w tym zdania typu $Z(Q)$, są prawdziwe. Niezależnie więc od różnic w pojmowaniu prawdziwości i fałszywości zdań języka J' na gruncie każdego z tych stanowisk zachowują walor wszystkie prawa klasycznego rachunku logicznego, obejmującego klasyczny rachunek zdań i rachunek kwantyfikatorów. W szczególności pozostaje prawdą logiczne prawo wyłączonego środka dla terminu Q: $\bigwedge_x (Qx \vee \sim Qx)$, a tym samym następujące jego podstawienie: $Qa_1 \vee \sim Qa_1$. To samo oczywiście dotyczy logicznego prawa sprzeczności: $\bigwedge_x \sim (Qx \wedge \sim Qx)$ i jego podstawienia: $\sim (Qa_1 \wedge \sim Qa_1)$. Różnie natomiast potraktowane są wedle powyższych stanowisk zdania $Z(Q)$ nie spełniające warunku (OM), tj. zdania „niezdeterminowane". Rozpatrzmy kolejno poszczególne rozwiązania i ich konsekwencje.

Stanowisko (I) odróżnia się od wszystkich pozostałych tym, iż owym zdaniom „niezdeterminowanym" odmawia wartości logicznej. Zdania te, jako prawdziwe w pewnych modelach rodziny $R(T)'$, a fałszywe w innych, nie są ani prawdziwe, ani fałszywe w sensie „absolutnym". Wedle stanowiska (I) — w przeciwieństwie do wszystkich pozostałych — zbiór zdań prawdziwych i fałszywych nie wyczerpuje ogółu zdań języka J'. Jak zakwalifikować więc owe wypowiedzi posiadające syntaktyczny charakter zdań, a pozbawione wartości logicznej? Reprezentowane są tutaj dwie możliwości. Wersja pierwsza (I.1) traktuje owe wypowiedzi jako wyrażenia bezsensowne. Nie należą one w ogóle do wyrażeń języka J'. Konsekwencje takiego rozwiązania są jednak trudne do przyjęcia. Pokazywaliśmy poprzednio, iż to, czy dane zdanie $Z(Q)$ spełnia warunek (OM), zależeć może od doświadczenia. W konsekwencji więc to, czy dane wyrażenie należy do języka J', czy też jest wyrażeniem bezsensownym, uzależnione zostaje od doświadczenia, a to jest czymś, czego przy konstrukcji języka usiłuje się za wszelką cenę uniknąć. Inną, rażącą nasze intuicje, konsekwencją jest fakt, iż koniunkcja, czy alternatywa wyrażeń bezsensownych może być wyrażeniem sensownym. A tak jest właśnie w przypadku wyrażeń Qa_1, $\sim Qa_1$. Oba — jako zdania „niezdeterminowane" — są bezsensowne, a ich koniunkcja czy alternatywa są wyrażeniami sensownymi; pierwsza — fałszywym, druga — prawdziwym zdaniem języka J'[16].

Wersja druga omawianego stanowiska (I.2) traktuje wypowiedzi posiadające syntaktyczny charakter zdań, a pozbawione wartości logicznej — jako sensowne zdania języka J'. Dopuszcza się więc istnienie zdań języka J' nie będących ani prawdą, ani fałszem. Odpadają przy tym ujęciu te kłopoty, o których wspominaliśmy przed chwilą. Ale powstają inne. Samo pojęcie zdania sensownego, a zatem coś stwierdzającego, a przy tym ani prawdziwego, ani fałszywego, wydaje się nieco zagadkowe. Ale bardziej kłopotliwe są konsekwencje inne. Choć stanowisko obecne zachowuje w całości klasyczny rachunek logiczny, zmusza jednak do odrzucenia pewnych klasycznych zależności metalogicznych. Logiczne prawo wyłączonego środka pozostaje prawdą, ale metalogiczne prawo wyłączonego środka traci swój walor. Istnieją zdania sprzeczne: Qa_1, $\sim Qa_1$, z których żadne nie jest prawdziwe. A ponieważ jednocześnie zdanie

16) Stanowisko (I. 1) odpowiada, z grubsza biorąc, stanowisku, jakie zajmowałem w artykule: „W sprawie terminów nieostrych", *Studia Logica* 8 (1958).

$Qa_1 \vee \sim Qa_1$ pozostaje prawdziwe, traci również swój walor klasyczna matryca alternatywy; okazuje się bowiem prawdą alternatywa dwóch zdań, z których żadne nie jest prawdziwe. Podobne konsekwencje otrzymujemy w przypadku pewnych innych metalogicznych zależności. A są to konsekwencje, których zlekceważyć nie podobna[17].

Wedle stanowiska (II) — i wszystkich dalszych — zbiór zdań prawdziwych i fałszywych języka J' wyczerpuje ogół zdań tego języka. Wszystkie więc zdania $Z(Q)$ nie spełniające warunku (OM) są prawdziwe lub fałszywe. Co więcej, zdania te nie zasługują tutaj na miano zdań „niezdeterminowanych". Ich wartość logiczna zostaje jednoznacznie określona. I tak zarówno zdanie Qa_1, jak i zdanie $\sim Qa_1$, zakwalifikowane zostają jako zdania prawdziwe. Widać z tego od razu, iż mimo zachowania klasycznego rachunku logicznego odrzuca się tu — i to w sposób radykalniejszy niż poprzednio — pewne klasyczne zależności metalogiczne. Wobec istnienia pary zdań sprzecznych, z których oba są prawdziwe, traci walor metalogiczne prawo sprzeczności. A wobec fałszywości zdania $Qa_1 \wedge \sim Qa_1$ traci walor klasyczna matryca koniunkcji, gdyż ta okazuje się fałszywa, mimo iż oba jej czynniki są prawdziwe[18]!

„Dualne" w stosunku do (II) stanowisko (III) charakteryzuje się analogicznymi właściwościami. Zdania $Z(Q)$, które nie spełniają warunku (OM), są prawdziwe lub fałszywe. Ich wartość logiczna jest przy tym jednoznacznie określona. Tutaj jednak zarówno zdanie Qa_1, jak i jego negacja $\sim Qa_1$, zaliczone zostają do zdań fałszywych. Odpadają zatem — przy zachowaniu klasycznego rachunku logicznego — takie klasyczne zależności metalogiczne, jak metalogiczne prawo wyłączonego środka czy klasyczna matryca alternatywy, skoro alternatywa $Qa_1 \vee \sim Qa_1$ pozostaje prawdą mimo fałszywości obu jej składników[19]!

Charakterystyka stanowiska (IV) musi pozostać ogólnikowa, dopóki w dalszym toku rozważań nie omówimy bliżej modelu $\mathfrak{M}'_i$, do którego się sformułowanie tego stanowiska odwołuje. W każdym razie już teraz możemy stwierdzić, że i tutaj wszystkie zdania $Z(Q)$ nie spełniające warunku (OM) muszą być prawdą lub fałszem. Co więcej, i tutaj zdania te tracą charakter zdań „niezdeterminowanych". Ich wartość logiczna zostaje jednoznacznie określona. W świetle dalszej charakterystyki modelu $\mathfrak{M}'_i$ okaże się, że zdanie Qa_1 jest zdaniem fałszywym, a zdanie $\sim Qa_1$ — prawdziwym. W przeciwieństwie jednak do stanowisk poprzednich, na gruncie stanowiska (IV) zachowują walor nie tylko wszelkie prawa klasycznego rachunku zdań i kwantyfikatorów, ale i wszelkie klasyczne zależności metalogiczne — choć pewne odstępstwo od klasycznych

17) Stanowisko (I. 2) pokrywa się ze stanowiskiem reprezentowanym przez H. Mehlberga w książce: *The Reach of Science*.

18) Stanowisko (II) odpowiada w przybliżeniu stanowisku reprezentowanemu przez W. Rozebooma w pracy: „The Factual Content of Theoretical Concepts", *Minnesota Studies*... Vol. 3, 1962.

19) Nie jest mi znany w literaturze przedmiotu żaden reprezentant stanowiska (III).

stosunków logicznych ma miejsce i tutaj — w punkcie, który wyjaśnimy w dalszym ciągu pracy[20].

W tym miejscu podkreślić chciałbym raz jeszcze wspólną konsekwencję omówionych ostatnio stanowisk (II), (III) i (IV), która budzić może uzasadnione wątpliwości czy sprzeciwy. Wszystkie te stanowiska określają, jak widzieliśmy, w sposób jednoznaczny wartość logiczną zdań „niezdeterminowanych". Nie idzie mi tutaj o to, że określenia te są niezgodne, że to samo zdanie wedle jednego stanowiska jest prawdą, wedle innego — fałszem. Idzie raczej o sam fakt takiego określenia. Wydaje się, iż przez to pierwotny charakter semantyczny owych zdań „niezdeterminowanych" ulega zmianie. Ich wartość logiczna zostaje przesądzona. Język *J'* traci wskutek tego w pewnym stopniu charakter języka otwartego. Język o pojęciach otwartych winien pozwalać na ich stopniowe uściślanie. Procedura ta, którą omawiałem na innym miejscu[21], polega na zaopatrywaniu terminu wprowadzonego za pomocą definicji warunkowej w dalsze definicje warunkowe. Definicje te, rozszerzając zakres stosowalności danego terminu, umożliwiają tym samym rozstrzyganie pewnych nierozstrzygalnych dotąd zdań ten termin zawierających. Otóż, jeśli wartość logiczna takich zdań zostaje z góry przesądzona, owa procedura uściślania pojęć otwartych pociąga za sobą konieczność uznania za fałszywe pewnych zdań, zaliczonych uprzednio do prawdziwych, i na odwrót. Ta właściwość omawianych ostatnio stanowisk wydaje się świadczyć o tym, iż nie oddają one wiernie otwartego charakteru języka *J'*.

Wadliwości tej pozbawione jest ostatnie z uwzględnionych w naszym przeglądzie stanowisk — stanowisko (V). Pojęcie prawdziwości jest tu zdefiniowane tak samo, jak w stosunku do języków całkowicie zinterpretowanych: jako prawdziwość w modelu właściwym $\mathfrak{M}'^*$. Różnica zaś w tym przypadku polega na tym, iż ów model nie jest wyznaczony jednoznacznie. Zakładamy jedynie, że dokładnie jeden taki model istnieje, i że jest nim jeden z modeli należących do rodziny $\mathfrak{K}\mathfrak{M}'$. W rezultacie, wszelkie zdania $Z(Q)$ nie spełniające warunku (OM) są prawdą lub fałszem. Wszelkie prawa klasycznego rachunku logicznego oraz wszelkie klasyczne zależności metalogiczne zachowują w pełni swój walor. Jednocześnie wszelkie zdania $Z(Q)$ nie spełniające warunku (OM), tj. wszelkie zdania „niezdeterminowane", pozostają nierozstrzygalne. Są prawdziwe lub fałszywe, ale niepełna charakterystyka modelu właściwego języka *J'*, a więc tego, o czym się w języku *J'* mówi, uniemożliwia rozstrzygnięcie, która z tych ewentualności zachodzi. W miarę uściślania terminu *Q* przez dołączanie dalszych definicji warunkowych część owych zdań staje się rozstrzygalna. Nie natrafiamy tu jednak nigdy na konieczność uznania za fałszywe (prawdziwe) zdań, zaliczonych uprzednio do prawdziwych (fałszywych). Wydaje się więc, że taka koncepcja prawdziwości odpowiada lepiej otwartemu charakterowi języka *J'*. Z drugiej strony jednakże, uznanie istnienia

20) Pod ogólny schemat stanowiska (IV) podpada m.in. stanowisko, jakie zajmuje T. Kubiński w pracy: „Nazwy nieostre", *Studia Logica* 7 (1958).

21) Por. np. „Pojęcia teoretyczne a doświadczenie", cyt. wyd.

zdań prawdziwych czy fałszywych, lecz zasadniczo nierozstrzygalnych, stanowi niepokojącą pod względem filozoficznym propozycję. Niektóre z rozpatrywanych poprzednio stanowisk sformułowane zostały z myślą uniknięcia tej konsekwencji — stworzenia takiej koncepcji prawdziwości, która by pozbawiła ją charakteru własności niepoznawalnej[22].

6. Semantyczna charakterystyka pojęć otwartych wymaga jeszcze uzupełnienia. Pozostaje nam do rozważenia problem ich denotacji. Problem ten stanowi wyraźny odpowiednik problemu prawdziwości. Toteż dyskusja nad tym zagadnieniem przebiegać będzie w dużym stopniu analogicznie do dyskusji nad zagadnieniem prawdziwości. Stanowiący główny przedmiot naszej analizy termin Q ma charakter predykatu i jako taki pełni dwojaką funkcję semantyczną: desygnowania i denotowania. Pomiędzy tymi dwoma funkcjami zachodzi ścisły związek, który w klasycznej semantyce logicznej polega po prostu na tym, iż zbiór przedmiotów desygnowanych przez dany predykat stanowi jego denotację. W naszych dotychczasowych rozważaniach posługiwaliśmy się wyłącznie „relatywnymi" pojęciami desygnowania i denotowania w modelu. Sens tych pojęć wyjaśniają następujące równoważności, odwołujące się do podstawowego we współczesnej semantyce logicznego pojęcia spełniania w modelu:

> *P* ***desygnuje x w modelu*** $\mathfrak{M}$ *wtedy i tylko wtedy, gdy x spełnia funkcję Px w modelu* $\mathfrak{M}$*;*
>
> *P* ***denotuje X w modelu*** $\mathfrak{M}$ *wtedy i tylko wtedy, gdy X jest zbiorem przedmiotów spełniających funkcję Px w modelu* $\mathfrak{M}$ *(czyli gdy X jest zbiorem przedmiotów desygnowanych przez P w modelu* $\mathfrak{M}$*)*[23].

W poszczególnym przypadku otrzymujemy stąd następujące konsekwencje:

> P_1 *desygnuje x w modelu*
>
> $$\mathfrak{M} = \langle U, a_1, \ldots, a_n, \mathbf{P}_1, \ldots, \mathbf{P}_m \rangle.$$
>
> *wtedy i tylko wtedy, gdy* $x \in \mathbf{P}_1$;
>
> P_1 denotuje $\mathbf{P}_1$ *w modelu*
>
> $$\mathfrak{M} = \langle U, a_1, \ldots, a_n, \mathbf{P}_1, \ldots, \mathbf{P}_m \rangle .$$

Problem nasz dotyczy wszakże nie „relatywnych", lecz „absolutnych" pojęć desygnowania i denotowania. Przejście do owych pojęć „absolutnych" polega — tak samo, jak w przypadku „absolutnego" pojęcia prawdziwości — na wyborze określonego modelu $\mathfrak{M}^*$ języka J jako modelu właściwego, dostarczającego przekładu wyrażeń języka J na metajęzyk MJ, i na zdefiniowaniu desygnowania (denotowania) jako desygnowania (denotowania) w modelu $\mathfrak{M}^*$. „Absolutne" pojęcia desygnowania i denotowania mają więc zastosowanie tylko do języka zinterpretowanego, czyli do takie-

22) Wyraźnie tak stawia sprawę H. Mehlberg w cytowanej książce. Stanowisko (V) stanowi rozwinięcie i precyzację stanowiska, jakie zajmowałem w pracy: „Pojęcia teoretyczne a doświadczenie" (cyt. wyd.) i jakie zakłada milcząco wielu autorów.

23) Zarówno „relatywne", jak i „absolutne" pojęcia desygnowania i denotowania wymagają oczywiście relatywizacji do języka, którą tu dla uproszczenia pomijamy.

go języka J, dla którego dany jest jego model właściwy $\mathfrak{M}^*$. Jeśli dla rozważanego przez nas języka J dany jest jako jego model właściwy $\mathfrak{M}^*$ podany wyżej układ

$$\langle U, a_1, ..., a_n, P_1, ..., P_m \rangle,$$

możemy stwierdzić po prostu, iż

P_1 desygnuje x wtedy i tylko wtedy, gdy $x \in P_1$;

P_1 denotuje P_1.

Zakładamy tutaj tak, jak uczyniliśmy poprzednio, że tak właśnie jest, tzn. że dla języka J dany jest jednoznacznie określony model właściwy $\mathfrak{M}^*$, i że modelem tym jest układ powyższy. Możemy przyjąć wobec tego, że dla dowolnego predykatu P języka J:

P desygnuje x wtedy i tylko wtedy, gdy P desygnuje x w modelu $\mathfrak{M}^*$;

P denotuje X wtedy i tylko wtedy, gdy P denotuje X w modelu $\mathfrak{M}^*$.

W stosunku do języka J' czynimy również te same założenia, co poprzednio. Zakładamy więc, że model właściwy $\mathfrak{M}'^*$ języka J' nie jest wyznaczony jednoznacznie, lecz tylko jako jeden z elementów — zdefiniowanej jak poprzednio, a więc obejmującej więcej niż jeden model — rodziny $\mathrm{R}\mathfrak{M}'$. Jaki więc sens możemy wiązać tutaj ze zwrotami: P desygnuje x, czy też P denotuje X?

Zarysowują się i tym razem możliwości różnych rozwiązań, odpowiadających w pewnym stopniu stanowiskom w sprawie prawdziwości i reprezentowanych na ogół przez przedstawicieli tamtych stanowisk. Wyróżnić możemy przede wszystkim trzy pojęcia desygnowania:

(A) *P desygnuje x wtedy i tylko wtedy, gdy P desygnuje x w każdym modelu $\mathfrak{M}'$ należącym do rodziny $\mathrm{R}\mathfrak{M}'$;*

(B) *P desygnuje x wtedy i tylko wtedy, gdy P desygnuje x w pewnym modelu $\mathfrak{M}'$ należącym do rodziny $\mathrm{R}\mathfrak{M}'$;*

(C) *P desygnuje x wtedy i tylko wtedy, gdy P desygnuje x w modelu $\mathfrak{M}'$,*

gdzie $\mathfrak{M}'^*$ jest modelem właściwym języka J' scharakteryzowanym wyłącznie przez warunek: $\mathfrak{M}'^* \in \mathrm{R}\mathfrak{M}'$.

Odpowiadające powyższym pojęciom desygnowania pojęcia denotacji utworzyć możemy na dwa sposoby. Pierwszy z nich odwołuje się do „relatywnego" pojęcia denotowania i prowadzi do następujących sformułowań:

(a) *P denotuje X wtedy i tylko wtedy, gdy P denotuje X w każdym modelu $\mathfrak{M}'$ należącym do rodziny $\mathrm{R}\mathfrak{M}'$;*

(b) *P denotuje X wtedy i tylko wtedy, gdy P denotuje X w pewnym modelu $\mathfrak{M}'$ należącym do rodziny $\mathrm{R}\mathfrak{M}'$;*

(c) *P denotuje X wtedy i tylko wtedy, gdy P denotuje X w modelu $\mathfrak{M}'^*$,*

gdzie $\mathfrak{M}'^*$ jest określone jak wyżej.

Drugi natomiast z tych sposobów odwołuje się do „absolutnych" pojęć desygnowania i głosi, iż:

P denotuje X wtedy i tylko wtedy, gdy X jest zbiorem przedmiotów desygnowanych (w sensie: (A), (B), (C)) *przez P.*

W przypadku, gdy w grę wchodzi pojęcie desygnowania (C), tak zdefiniowane pojęcie denotowania jest identyczne z pojęciem denotowania (c). Natomiast w przypadku pojęć desygnowania (A) i (B) otrzymujemy pojęcia denotowania (a*) i (b*), różne od pojęć denotowania (a) i (b).

Zilustrujmy konsekwencje przedstawionych stanowisk na przykładzie terminu Q języka J', co do którego poczyniliśmy poprzednio szereg upraszczających założeń. W ich wyniku, jak pamiętamy, rodzina $\mathfrak{RM}'$ obejmuje tylko dwa modele $\mathfrak{M}_1'$ i $\mathfrak{M}_2'$. Łatwo sprawdzić, iż zgodnie z (A) Q desygnuje każdy i tylko taki przedmiot, który należy do zbioru $\boldsymbol{P}_1$; zgodnie z (B) — każdy i tylko taki przedmiot, który należy do zbioru $\boldsymbol{-P}_2$. Zgodnie z (C) wreszcie zachodzi dokładnie jedna z tych ewentualności, ale rozstrzygnięcie tego, która z nich, jest rzeczą niemożliwą. Przedstawione koncepcje denotowania prowadzą z kolei do następujących konkluzji. Wedle (a) Q nie denotuje żadnego zbioru; wedle (b) Q denotuje zarówno zbiór $\boldsymbol{P}_1$, jak i zbiór $\boldsymbol{-P}_2$; wedle (a*) Q denotuje $\boldsymbol{P}_1$, a wedle (b*) — $\boldsymbol{-P}_2$. Wedle (c) wreszcie Q denotuje $\boldsymbol{P}_1$ lub $\boldsymbol{-P}_2$, ale tego, która z tych ewentualności zachodzi, rozstrzygnąć nie jesteśmy w stanie.

Konsekwencje te nasuwają pewne uwagi ogólne na temat każdego z przedstawionych stanowisk. Interesują nas tu głównie stanowiska w sprawie denotacji. Dwa z nich: (a) i (b) odbiegają — w odróżnieniu od pozostałych — od klasycznej koncepcji denotacji. Koncepcja ta zakłada, iż każdy termin denotuje jeden i tylko jeden przedmiot. Tymczasem wedle (a) termin zdefiniowany warunkowo nie denotuje niczego, a wedle (b) denotuje więcej niż jeden przedmiot. Stanowiska (a) i (b) odpowiadają analogicznym stanowiskom w sprawie prawdziwości (I.1) i (II) i są przez przedstawicieli tych ostatnich istotnie reprezentowane[24]. Przy próbach ustosunkowania się do stanowisk (a) i (b) trudno nie przyznać, iż rażą one nasze intuicje dotyczące denotacji, gdyż te związane są wyraźnie z koncepcją klasyczną. Pewną niekonsekwencją poza tym wydaje się łączenie — jak to istotnie ma miejsce — stanowisk (a) i (b) z odpowiadającymi im stanowiskami w sprawie desygnowania (A) i (B). Jeśli zarówno wedle koncepcji desygnowania (A), jak i (B) termin zdefiniowany warunkowo desygnuje każdy i tylko taki przedmiot, który należy do jednoznacznie określonego zbioru, nie widać powodu, dla którego musielibyśmy przyjmować, iż termin taki nie denotuje żadnego zbioru lub że denotuje szereg zbiorów różnych.

Toteż z koncepcjami desygnowania (A) i (B) wydają się harmonizować lepiej koncepcje denotacji (a*) i (b*). Stanowiska te przyporządkowują terminowi zdefiniowanemu warunkowo jako jego denotację jednoznacznie określony zbiór przedmiotów; stanowisko (a*) — zbiór przedmiotów desygnowanych zgodnie z koncepcją desygnowania (A), stanowisko (b*) — zbiór desygnatów pojętych wedle koncepcji (B). Jak uzasadnia się na gruncie tych stanowisk wybór takich właśnie zbiorów desygnatów? Spróbujmy odpowiedzieć na to pytanie w stosunku do pierwszego z tych stanowisk:

24) Por. przypisy 16 i 18.

(A) — (a*), które — w przeciwieństwie do pozostałego — reprezentowane jest faktycznie przez przedstawicieli pewnych spośród uwzględnionych przez nas koncepcji prawdziwości: (I.2) i (IV)[25].

Intuicje, jakie leżą u podstawy koncepcji desygnowania (A), znajdują wyraz w następującej, wyraźnie przez zwolenników tej koncepcji zakładanej zależności:

(D) *P desygnuje x wtedy i tylko wtedy, gdy każdy, kto używa poprawnie terminu P, musi (w określonej sytuacji) orzec ów termin o przedmiocie x.*

Jest rzeczą widoczną, że w przypadku terminów zdefiniowanych warunkowo równoważność powyższa implikuje definicję desygnowania (A). Przy jednoczesnym przyjęciu definicji denotowania (a*) otrzymujemy dla terminu Q opisanego języka J' wspomnianą poprzednio konsekwencję: Q denotuje zbiór P_1.

Stanowisko obecne wiąże się przy tym ze swoistą, odmienną od klasycznej, interpretacją negacji przynazwowej. Załóżmy, że wiemy, jaki zbiór stanowi denotację predykatu P. Co denotuje jego negacja P'? Wszystkie pozostałe koncepcje denotacji odpowiadają na to pytanie w sposób tradycyjny:

P denotuje X wtedy i tylko wtedy, gdy P′ denotuje -X.

Stanowisko obecne daje jednak odpowiedź inną. Zakłada bowiem zależność następującą, chwytającą podobne intuicje, co przytoczona poprzednio zależność (D):

(D′) *P′ desygnuje x wtedy i tylko wtedy, gdy nikt, kto używa poprawnie terminu P, nie może (w określonej sytuacji) orzec owego terminu o przedmiocie x.*

Równoważność ta implikuje następującą definicję desygnowania:

(A′) *P′ desygnuje x wtedy i tylko wtedy, gdy P nie desygnuje x w żadnym modelu $\mathfrak{M}'$ należącym do rodziny $\mathrm{R}\mathfrak{M}'$.*

Przy jednoczesnym przyjęciu dla P' definicji denotowania (a*) otrzymujemy dla terminu Q naszego języka J' następującą konkluzję: Q' denotuje zbiór P_2. Ale zbiór P_2 nie stanowi — wedle założenia — dopełnienia zbioru P_1 (do universum U). Przedmiot a_1 nie należy do żadnego z tych zbiorów. A zatem jest to przedmiot nie należący ani do denotacji predykatu Q, ani do denotacji jego negacji Q'. Czy wobec tego traci walor prawo wyłączonego środka dla negacji przynazwowej: $\bigwedge_x (Qx \vee Q'x)$? Odpowiedź na to pytanie zależy od tego, z jakim stanowiskiem w sprawie prawdziwości — (I.2) czy (IV) — wiążą się omawiane koncepcje desygnowania i denotacji: (A) — (A′) — (a*). Przy rozumieniu prawdziwości zgodnie ze stanowiskiem (I.2) negacja przynazwowa może być rozumiana klasycznie: $\bigwedge_x (Q'x \equiv \sim Qx)$. Koncepcja denotacji terminu Q i jego negacji Q' nie wpływa tu bezpośrednio na wartość logiczną zdań zawierających te terminy. Zachowują więc walor wszystkie prawa klasycznego rachunku nazw, wraz z prawem wyłączonego środka. Mimo iż zdania: Qa_1, $Q'a_1$ są pozbawione wartości logicznej, zdanie $Qa_1 \vee Q'a_1$ pozostaje prawdą. Inaczej jest jednak na gruncie stanowiska (IV). Prawdziwość jest tutaj rozumiana jako prawdziwość w modelu $\mathfrak{M}'_i$. Otóż owym mode-

25) Por. przypisy 17 i 20.

lem $\mathfrak{M}'_i$ ma być taki model należący do rodziny $\mathfrak{K}\mathfrak{M}'$, w którym denotacje terminu Q i jego negacji Q' są zbiorami określonymi zgodnie z definicjami (A), (A'), (a*). Jeśli przyjmiemy sformułowane poprzednio założenia co do języka J', modelem $\mathfrak{M}'_i$ będzie taki model języka J', w którym denotację terminu Q stanowi zbiór P_1, a denotację terminu Q' — zbiór P_2. W tej sytuacji tracą walor pewne prawa klasycznego rachunku nazw, wśród nich — prawo wyłączonego środka. Oba zdania: Qa_1, $Q'a_1$ są fałszywe i fałszem jest ich alternatywa: $Qa_1 \vee Q'a_1$. Negacja przynazwowa nie może być więc rozumiana klasycznie. Odpada w szczególności zależność $\bigwedge_x (\sim Qx \rightarrow Q'x)$. W rezultacie klasyczny rachunek nazw musi być zastąpiony jakimś rachunkiem nieklasycznym[26].

Ta konsekwencja omawianej obecnie koncepcji denotacji przemawia na jej niekorzyść. Ale koncepcja ta wydaje się niezadowalająca i z innych względów. Motywy skłaniające do przyjęcia definicji (A), (A'), (a*), a znajdujące swój wyraz w równoważnościach (D), (D'), świadczą o pewnej zmianie sensu klasycznych pojęć desygnowania i denotowania, jaka ma miejsce na gruncie obecnej koncepcji. To, czy P desygnuje x, uzależnione jest tutaj w rezultacie od tego, czy jest się skłonnym uznać zdanie orzekające P o x, a więc od pewnej postawy osób mówiących danym językiem. Pojęcie desygnowania, a co za tym idzie, i denotacji, nabiera tym samym charakteru pragmatycznego raczej, niż semantycznego. Podobne uwagi mogłyby mieć zastosowanie i do definicji (B), (b*), gdyż i tutaj można by znaleźć motywy skłaniające do ich przyjęcia, o podobnym, pragmatycznym charakterze[27]. Główną jednakże wadą, wspólną dla obu koncepcji, wydaje się co innego. Nie to, że terminowi zdefiniowanemu warunkowo przyporządkowują taką właśnie, a nie inną denotację, lecz to, że w ogóle dokonują takiego jednoznacznego przyporządkowania. Pojęcie odpowiadające takiemu terminowi zatraca przez to charakter pojęcia otwartego. Ustalone zostaje w sposób jednoznaczny, do czego się dany termin odnosi. Dalsze jego uściślanie, polegające na dołączaniu dodatkowych definicji warunkowych, prowadzi do innego, niezgodnego z poprzednim, ustalenia jego denotacji. Za pomocą owego uściślonego terminu zaczynamy mówić o czymś innym, niż mówiliśmy poprzednio.

Tej wady pozbawiona jest ostatnia koncepcja desygnowania i denotacji: (C) — (c). Pojęcie denotacji jest tu rozumiane tak samo, jak w stosunku do języków całkowicie zinterpretowanych: jako denotacja w modelu właściwym $\mathfrak{M}'^*$. Ale ów model nie jest tutaj wyznaczony jednoznacznie. A więc i denotacja terminu Q — jako zdefiniowanego warunkowo terminu języka J' — nie jest ustalona w sposób jednoznaczny. Termin Q denotuje dokładnie jeden zbiór ze ściśle określonej rodziny zbiorów (w przypadku naszego prostego języka J' — zbiór P_1 lub zbiór $-P_2$), ale rozstrzygnięcie, który to zbiór, jest rzeczą niemożliwą. Stanowisko takie łączy w sobie pewne zalety każdego z

26) Taki nieklasyczny rachunek nazw konstruuje w cytowanej pracy T. Kubiński.

27) Takim motywem do przyjęcia definicji (B) mogłoby być następujące zdanie: *P desygnuje x wtedy i tylko wtedy, gdy każdy, kto używa poprawnie terminu P, może (w określonej sytuacji) orzec ów termin o przedmiocie x.*

rozwiązań poprzednich, unikając zarazem ich wad. Pozostajemy tu z jednej strony na gruncie klasycznego rozumienia denotacji i klasycznego rozumienia negacji terminów, z drugiej strony — zachowujemy nieokreśloność funkcji denotowania, charakterystyczną dla pojęć otwartych. Każdy termin, a więc i termin zdefiniowany warunkowo, denotuje jeden i tylko jeden przedmiot. Ale w przypadku terminu zdefiniowanego warunkowo ów przedmiot jest określony tylko jako element pewnej klasy, liczącej więcej niż jeden przedmiot. Przypuśćmy teraz, iż uściślamy dany termin, zaopatrując go w dalsze definicje warunkowe. Rezultatem takiego zabiegu jest zwężenie owej klasy „możliwych" denotacji. Nie pociąga to jednak za sobą nieuchronnie zmiany pierwotnej denotacji danego terminu. Możemy utrzymywać, że za pomocą tak uściślonego terminu mówimy nadal o tym samym, o czym mówiliśmy poprzednio. Wydaje się więc, że jest to koncepcja denotacji zgodna z otwartym charakterem języka J'. Jednocześnie jednak przyznać trzeba, iż twierdzenie głoszące, że pewien termin coś denotuje, choć jest rzeczą niemożliwą rozstrzygnąć co denotuje, brzmi istotnie, jak zarzucają niektórzy, nieco „metafizycznie".

7. Kończymy na tym naszą analizę problemów semantycznych pojęć otwartych. Wyróżniliśmy w jej toku klasę wypowiedzi, które takie problemy nasuwają, oraz sformułowaliśmy zarówno owe problemy, jak i możliwe próby ich rozwiązania. Rozważania nasze opierały się na szeregu upraszczających istniejący stan rzeczy założeń. Niektóre z nich wymagają paru słów komentarza. Najważniejsze bodaj z nich polegało na utożsamieniu pojęć otwartych z terminami zdefiniowanymi warunkowo. Wspominaliśmy już na wstępie o upraszczającym charakterze takiego założenia. Istnieć mogą terminy teoretyczne teorii empirycznych — a więc terminy odpowiadające pojęciom otwartym — nie należące do terminów zdefiniowanych warunkowo. Postulaty znaczeniowe wprowadzające takie terminy do teorii empirycznych przybierać mogą postać nie podpadającą pod schemat definicji warunkowej. Ogólnie — i ogólnikowo zarazem — postulat wprowadzający termin Q do języka J' scharakteryzować można jako takie zdanie języka J' zawierające termin Q, które przy wyznaczonej przez model właściwy $\mathfrak{M}^*$ języka J interpretacji terminów pozostałych dopuszcza co najmniej jedną interpretację terminu Q. Czy rozważania nasze dają się rozciągnąć na terminy wprowadzone za pomocą dowolnych postulatów znaczeniowych, a więc i takich, które nie mają charakteru definicji warunkowych? Próbując odpowiedzieć na to pytanie, stwierdzić możemy w każdym razie, że podstawowe rezultaty naszych rozważań dotyczą wszelkich terminów tak scharakteryzowanych. Każdy taki termin pozwala na zdefiniowanie w sposób przez nas podany klasy zdań „niezdeterminowanych", ów termin zawierających, i na sformułowanie, tak jak to zrobiliśmy poprzednio, uwzględnionych w naszym przeglądzie stanowisk w sprawie prawdziwości takich zdań i denotacji danego terminu. Jedyna wątpliwość powstaje w związku z charakterystyką zdań „niezdeterminowanych" jako zdań, z których omawiany termin ma być eliminowalny. Mówiąc dokładniej, problemem nierozstrzygniętym pozostaje pytanie, czy w rozważanym przypadku ogólnym warunek

„określoności" ((OL), (OT), (OM)) pociąga za sobą logicznie warunek „eliminowalności" ((EL), (ET), (EM))[28]. Czy też istnieje może takie zdanie języka J' zawierające termin Q, które ma określoną wartość logiczną ze względu na Q na podstawie postulatu wprowadzającego ten termin, a które mimo to nie daje się wyrazić w sposób równoważny w języku J na gruncie owego postulatu? Musimy poprzestać tutaj na sformułowaniu owej wątpliwości, a jej rozstrzygnięcie odłożyć do przyszłych rozważań.

Wypada wspomnieć wreszcie o ogólnym założeniu, na jakim oparte są z konieczności wszelkie badania semantyczne stosujące aparat pojęciowy współczesnej semantyki logicznej. Jest to założenie dotyczące rodzaju języków, które stanowią obiekt semantycznej analizy, a których przykładem mogą być rozważane przez nas języki J i J'. Są to, jak widzieliśmy, języki standardowo sformalizowane, stanowiące idealizację i uproszczenie rzeczywistego języka nauki. Rezultaty naszych rozważań stosują się więc, ściśle biorąc, nie do tego języka, którym w rzeczywistości posługują się naukowcy, lecz do języków, które stanowią rekonstrukcję logiczną pewnych jego fragmentów. Jest to cena, jaką płacimy za możność przedstawienia pewnej semantycznej problematyki w sposób ścisły i konsekwentny. Pojęciami umożliwiającymi to zadanie w przypadku omawianych przez nas problemów semantycznych języków otwartych są, jak starałem się to okazać, takie pojęcia współczesnej semantyki logicznej, jak pojęcie języka sformalizowanego, jego modelu, prawdziwości i denotacji w modelu. Nie wystarcza tu w szczególności posługiwanie się potocznym pojęciem języka zinterpretowanego i „absolutnymi" pojęciami prawdy czy denotacji. Języki otwarte bowiem — to języki zinterpretowane tylko częściowo, a zatem dopuszczające możliwość różnych interpretacji. Musimy w związku z tym operować z jednej strony pojęciem języka jako pewnego tworu czysto syntaktycznego, z drugiej strony pojęciem dziedziny rzeczywistości, do której się tak rozumiany język może odnosić. Musimy również dysponować pojęciami prawdziwości zdań danego języka i denotacji jego terminów zrelatywizowanymi do określonej dziedziny rzeczywistości, a więc do określonej interpretacji danego języka. Tych właśnie pojęć dostarcza nam teoria modeli języków sformalizowanych. Zadaniem, jakie sobie stawiałem, było przedstawienie aktualnego w logice współczesnej sporu o prawdziwość i odniesienie przedmiotowe wypowiedzi charakterystycznych dla języków otwartych przy pomocy aparatury pojęciowej tej teorii.

Do zadań pracy obecnej nie należało w zasadzie zajmowanie w tym sporze własnego stanowiska. Jeżeli przytaczałem niejednokrotnie argumentację na rzecz ostatniego z omawianych rozwiązań — koncepcji prawdziwości (V) i denotacji (C)-(c) — to czyniłem to z pewnymi zastrzeżeniami, którymi akceptację takiego rozwiązania zmuszony byłbym opatrzyć. I tak nie sądzę, aby rozważane problemy można było trafnie formułować jako pytania rzeczowe, dotyczące tego, „jak naprawdę jest", a więc: czy

28) Dowód takiej zależności podany w pracy nie daje się zastosować do przypadku ogólnego.

zdania „niezdeterminowane" naprawdę posiadają wartość logiczną, lub czy termin zdefiniowany warunkowo istotnie denotuje jakiś przedmiot. Nie sądzę bowiem, aby „zastane" znaczenie pojęć takich, jak prawda czy denotacja, mogło przesądzać te pytania w sposób definitywny. Mam wrażenie, iż żadne z przedstawionych rozwiązań nie gwałci tego znaczenia w sposób oczywisty. Może się tylko z nim w mniejszym lub większym stopniu zgadzać, a to nie stanowi jeszcze argumentu decydującego. Idzie więc raczej o wybór pewnego konsekwentnego, ścisłego i najbardziej — pod określonymi względami — zadowalającego sposobu mówienia. Otóż w zależności od tego, jakie względy bierzemy przede wszystkim pod uwagę, wypadnie nam opowiedzieć się za tym lub innym rozwiązaniem. Każde bowiem ma pewne wady i pewne zalety. Opowiadając się za koncepcją prawdziwości i denotacji, będącą rozszerzeniem na języki częściowo zinterpretowane tego rozumienia tych pojęć, które ma zastosowanie do języków całkowicie zinterpretowanych, czyniłem to przede wszystkim z pozycji logika. Pod względem logicznym bowiem to stanowisko przedstawia rozwiązanie najbardziej zadowalające. Zajmując wszakże inny punkt widzenia, innego musielibyśmy dokonać wyboru. Rozumienie, które sprawia najmniej kłopotów logicznych, okazuje się rozumieniem ryzykownym pod względem filozoficznym. Toteż wzdraga się przed nim niejeden pozytywistyczny filozof. Ale czy w tym logicznym, bądź co bądź, problemie pierwszeństwo nie należy się logicznym kryteriom oceny? Skłaniając się do wspomnianego rozwiązania, takiemu właśnie przekonaniu dawałem wyraz.

Na koniec — parę uwag o metodologicznym aspekcie naszych rozważań. Zwracałem już uwagę na fakt, że wzorcem języków otwartych są języki empirycznych teorii naukowych. Terminy teoretyczne definiowane są w takich teoriach przez postulaty częściowo tylko ustalające ich denotacje. Pojęcia odpowiadające takim terminom mają więc charakter pojęć otwartych. W tej sytuacji spór o wartość logiczną i odniesienie przedmiotowe wypowiedzi charakterystycznych dla języków otwartych pokrywa się w znacznym stopniu z aktualnym w metodologii nauk sporem o wartość poznawczą teorii naukowych. Dyskusja na tym terenie toczy się bowiem głównie o to, czy twierdzenia teorii mogą być traktowane jako wypowiedzi prawdziwe lub fałszywe. Toteż rezultaty naszych rozważań mają dla tej dyskusji znaczenie bezpośrednie. Okazują one przede wszystkim, iż wątpliwości owe mogą dotyczyć tylko niektórych twierdzeń teoretycznych. Istnieje obszerna klasa twierdzeń zawierających terminy teoretyczne, których wartość logiczna nie podlega dyskusji, gdyż terminy te są z nich eliminowalne. Klasa ta obejmuje nie tylko tautologie, czy ich negacje, ale i twierdzenia, w których terminy teoretyczne występują w sposób istotny. W stosunku do pozostałych twierdzeń teoretycznych zajmować można, jak widzieliśmy, stanowiska różne. W metodologii współczesnej wyróżnia się na ogół dwa zasadnicze typy rozwiązań — „instrumentalistyczne" i „realistyczne"[29]. Pierwsze odmawiają owym

29) Por. np. E. Nagel: *The Structure of Science*, 1961.

problematycznym wypowiedziom wartości logicznej, drugie im taką wartość — rozmaicie zresztą pojmowaną — przyznają. Wśród uwzględnionych przez nas stanowisk znajdujemy przykłady obu typów rozwiązań. Stanowisko, do którego gotów byłbym zgłosić swój akces, reprezentuje typowe rozwiązanie „realistyczne". Twierdzenia zawierające nieeliminowalne terminy teoretyczne uznane są tu za zdania w pełni sensowne, posiadające — pojmowaną klasycznie — wartość logiczną. Niejednoznaczność ich interpretacji pociąga jedynie ich nierozstrzygalność. Co więcej, występujące w nich terminy teoretyczne denotują — w sensie tradycyjnym — pewne przedmioty, a tylko jednoznaczne określenie tych ostatnich pozostaje rzeczą niewykonalną. Zadowalające pod względem logicznym sformułowanie tego „realistycznego" rozwiązania stanowiło jeden z celów niniejszych rozważań.

Filozofia Nauki
Rok I, 1993, Nr 2-3

O identyfikowalności w dziedzinach rozszerzonych

Dla potrzeb przeprowadzanej tu analizy przez definiowalność będę rozumiał to, co zwykle określa się mianem definiowalności bezpośredniej (albo syntaktycznej); przez identyfikowalność — to, co określa się jako definowalność pośrednią (albo semantyczną). Ograniczę się tylko do rozważenia teorii empirycznych, które mogą być sformalizowane w języku logiki pierwszego rzędu. Jak wiadomo, w wypadku takich teorii oba te pojęcia pokrywają się. Termin t jest definiowalny w teorii T zawsze i tylko, gdy jest identyfikowalny w tej teorii. Innymi słowy, definicja terminu t jest twierdzeniem teorii T zawsze i tylko, gdy interpretacja pozostałych terminów w każdym modelu teorii T wyznacza jednoznacznie interpretację t. Z tego właśnie powodu we wszystkich dotąd rozważanych wypadkach nieidentyfikowalność jest w jakiś sposób związana z niedefiniowalnością. Pojęcie jest nieidentyfikowalne tylko wtedy, gdy nie jest definiowalne; np. w sytuacji, gdy jego użyciem rządzą postulaty słabsze niż definicja bezpośrednia (definicja warunkowa, kontekstowa lub cząstkowa — aby wymienić tylko najbardziej typowe wypadki). Chciałbym zająć się pewnymi innymi rodzajami nieidentyfikowalności i ich typowymi źródłami. Nieidentyfikowalność tego rodzaju nie jest związana z niedefiniowalnością. Pojawia się ona w wyniku przyjęcia pewnego słabszego pojęcia interpretacji teorii. W ujęciu standardowym, każde rozszerzenie danego języka ma semantyczny odpowiednik w postaci odpowiedniego przedłużenia struktur tego języka. Wydaje się jednak, że są takie wypadki, w których rozszerzenie języka wiąże się nie tylko z przedłużeniem jego struktur, ale także z rozszerzeniem jego dziedzin. Wydaje się, że tak dzieje się w sytuacji, gdy rozszerza się język obserwacyjny przez wprowadzenie do niego pewnego typu terminów teoretycznych. W takich wypadkach przedstawiona wyżej zależność miedzy definiowalnością i identyfikowalnością już nie zachodzi. Definiowalne w danej teorii pojęcie teoretyczne nie jest identyfikowalne w modelach tej teorii — jeżeli są one strukturami o rozszerzonych dziedzinach.

Aby opisać tę sytuację bardziej precyzyjnie, skorzystamy z języka teorii modeli. Niech L_o będzie językiem pierwszego rzędu, z terminami $o_1,...,o_n$ jako jedynymi stałymi deskryptywnymi, i niech L będzie rozszerzeniem L_o zawierającym t jako dodatkową stałą deskryptywną. Struktury modelowe dla L_o będziemy oznaczać przez $\mathfrak{M}_o$ (albo przez $\mathfrak{N}_o$), natomiast dla L — przez $\mathfrak{M}$ (albo $\mathfrak{N}$). Dziedzina struktury $\mathfrak{M}$ będzie oznaczana przez $|\mathfrak{M}|$, interpretacja w $\mathfrak{M}$ stałej c — przez $c^{\mathfrak{M}}$. Przez $\mathfrak{M}|_o$ będziemy oznaczać obcięcie (redukcję) struktury $\mathfrak{M}$ do języka L_o; $\mathfrak{M}$ jest wtedy nazywane przedłużeniem struktury $\mathfrak{M}|_o$. Jeśli T jest zbiorem zdań języka L, to klasę jego modeli będziemy oznaczać przez Mod (T).

Zgodnie ze standardowym ujęciem, rozszerzenie języka L_o do języka L będziemy związane jest z przedłużeniem struktury $\mathfrak{M}_o$ do struktury $\mathfrak{M}$. W związku z tym, badając związek między (syntaktyczną) definiowalnością i (semantyczną) identyfikowalnością, będziemy brać pod uwagę klasę struktur, $M_{\mathfrak{M}_o}$, zdefiniowaną dla każdego $\mathfrak{M}_o$ w następujący sposób:

(I) $M_{\mathfrak{M}_o} = \{\mathfrak{M}\colon \mathfrak{M}|_o = \mathfrak{M}_o \text{ i } \mathfrak{M} \in \mathrm{Mod}\,(T)\}$.

$M_{\mathfrak{M}_o}$ zawiera te przedłużenia $\mathfrak{M}o$, które są modelami dla T. Stała t jest definiowalna w T (za pomocą terminów $o_1,...,o_n$) zawsze i tylko, gdy dla każdego $\mathfrak{M}_o$, $M_{\mathfrak{M}_o}$ zawiera co najwyżej jedną strukturę. Jeśli T sprowadza się do definicji terminu t (za pomocą terminów $o_1,...,o_n$), to dla każdego $\mathfrak{M}_o$, $M_{\mathfrak{M}_o}$ zawiera dokładnie jedną strukturę: dla każdego $\mathfrak{M}_o$ istnieje dokładnie jedna interpretacja stałej t. Mówimy wtedy, że t jest identyfikowalne w T.

Takie ujęcie można zastosować w tych wszystkich wypadkach, w których wprowadzenie nowych terminów nie wiąże się z rozszerzeniem dotychczasowego uniwersum dyskursu. Jeśli chodzi jednak o teorie empiryczne, to wydaje się, że nie każda procedura wprowadzania nowego terminu podpada pod powyższy schemat. Są wypadki, w których procedura taka jest w sposób istotny związana z rozszerzeniem dziedziny teorii. Typową sytuację tego rodzaju można opisać w następujący sposób. Niech L_o będzie językiem obserwacyjnym i niech $o_1,...,o_n$ będą predykatami obserwacyjnymi. Cokolwiek te określenia miałyby znaczyć, założymy, że każdy model zamierzony $\mathfrak{M}_o$ dla L_o składa się tylko z przedmiotów obserwowalnych: jego dziedzina $|\mathfrak{M}_o|$ jest zbiorem przedmiotów obserwowalnych. Niech t będzie predykatem teoretycznym (powiedzmy predykatem „elektron"), który odnosi się do pewnych nieobserwowalnych, «teoretycznych» bytów. Wprowadzając go do języka L_o musimy rozszerzyć jego modele tak, aby włączyć obiekty nieobserwowalne do ich dziedzin; tylko takie stuktury można zaliczyć do klasy zamierzonych modeli dla L.

W wypadkach takich jak powyższe, interesuje nas przede wszystkim interpretacja t w owych odpowiednio rozszerzonych strukturach. Z jakiego typu rozszerzeniem mamy tu do czynienia? W teorii modeli, $\mathfrak{N}_o$ jest nazywane rozszerzeniem $\mathfrak{M}_o$ zawsze i tylko, gdy $|\mathfrak{M}_o|$ jest zawarte w $|\mathfrak{N}_o|$ i $o_i^{\mathfrak{M}_o}$ jest identyczne z $o_i^{\mathfrak{N}_o}$ ograniczonym do $|\mathfrak{M}_o|$ (dla $i = 1,...,n$). Tak więc, poza $|\mathfrak{M}_o|$, o_i może być interpretowane w $\mathfrak{N}_o$ w dowolny sposób.

Czy predykaty obserwacyjne tak właśnie są zinterpretowane w dziedzinie przedmiotów nieobserwowalnych? Ponieważ pojęcie predykatu obserwacyjnego jest notorycznie wieloznaczne, odpowiedź na to pytanie w oczywisty sposób zależy od tego, jak je będziemy rozumieć. Można wskazać trzy główne stanowiska dotyczące tej kwestii. W dwóch z nich predykat obserwacyjny jest rozumiany jako predykat interpretowany w czysto ostensywny sposób. W konsekwencji, jego interpretacja jest z założenia «ograniczona» jedynie do przedmiotów obserwowalnych. To «ograniczenie» jest jednakże rozumiane na dwa różne sposoby.

(i) Zgodnie z jedną z eksplikacji, interpretacja predykatu obserwacyjnego w dziedzinie wszystkich przedmiotów nieobserwowalnych jest, z założenia, całkowicie niezdeterminowana: poza dziedziną obserwowalną predykat może być interpretowany w dowolny sposób. Zgodnie z tym założeniem, przy definiowaniu modeli zamierzonych dla L będzie brane pod uwagę każde rozszerzenie $\mathfrak{N}_o$ struktury $\mathfrak{M}_o$.

(ii) Zgodnie z inną eksplikacją, zakłada się, że interpretacja predykatu obserwacyjnego jest negatywnie zdeterminowana w stosunku do obiektów nieobserwowalnych: twierdzi się, że predykat można prawdziwie orzec jedynie o przedmiotach obserwowalnych. Przy tym założeniu, przy definiowaniu zamierzonych modeli dla L będą brane pod uwagę tylko te rozszerzenia struktury $\mathfrak{N}_o$ do struktury $\mathfrak{M}_o$, w których $o_i^{\mathfrak{N}_o} = o_i^{\mathfrak{M}_o}$ (dla $i = 1,\dots,n$).

(iii) Trzecia koncepcja jest bardziej liberalną. Nie ogranicza ona predykatów obserwacyjnych jedynie do predykatów definiowalnych ostensywnie. O predykatach obserwacyjnych zakłada się tutaj, że są one interpretowane nie tylko za pomocą pewnych metod bezpośrednich (niewerbalnych), takich jak ostensja, ale również w sposób pośredni (werbalny), mianowicie za pomocą zbioru postulatów. Niebezpośredni sposób interpretacji jest jedynym dopuszczalnym sposobem interpretacji w obrębie dziedziny przedmiotów nieobserwowalnych: predykaty mogą być interpretowane bezpośrednio jedynie w dziedzinie obserwowalnej; poza nią mogą być one interpretowane tylko pośrednio, za pomocą zbioru postulatów. Niech zbiór T_o zdań języka L_o będzie zbiorem postulatów dla predykatów obserwacyjnych $o_1,\dots,o_n$. Przy obecnych założeniach, będziemy brać pod uwagę przy definiowaniu zamierzonych modeli dla L tylko takie rozszerzenia $\mathfrak{N}_o$ struktury $\mathfrak{M}_o$, które są modelami dla postulatów T_o : $\mathfrak{N}_o \in \mathrm{Mod}\,(T_o)$. Aby to zagwarantować, będziemy żądać, by zbiór T zawierał [illegible] T_o: $T_o \subset \mathrm{Cn}\,(T)$.[1]

Tę ostatnią koncepcję predykatów obserwacyjnych proponuję przyjąć jako podstawę naszej analizy. Wynika z niej następująca definicja klasy $N_{\mathfrak{M}_o}$, która ma być odpowiednikiem klasy $M_{\mathfrak{M}_o}$, przystosowanym do rozważanego problemu:

(II) $N_{\mathfrak{M}_o} = \{\mathfrak{M}$: dla pewnego rozszerzenia $\mathfrak{N}_o$ struktury $\mathfrak{M}_o$, $\mathfrak{M}|_o = \mathfrak{N}_o$ i $\mathfrak{M} \in \mathrm{Mod}\,(T)\}$.

1) Predykat identyczności można traktować w ten sam sposób. W dziedzinie obserwowalnej $|\mathfrak{M}_o|$ „=" można traktować jako identyczność, poza $|\mathfrak{M}_o|$ — jako dowolną relację spełniającą aksjomaty identyczności.

$N_{\mathfrak{M}_0}$ zawiera te struktury dla L, które są przedłużeniami rozszerzeń struktury $\mathfrak{M}_0$, oraz modelami dla T.

Definicja klasy $N_{\mathfrak{M}_0}$ odzwierciedla podstawową cechę tej koncepcji interpretacji teorii, która leży u podstaw prezentowanego podejścia. Koncepcja ta jest znana pod nazwą empiryzmu semantycznego. Zgodnie z tą koncepcją jedynie język obserwacyjny może być zinterpretowany bezpośrednio, albo, że tak powiem, «z zewnątrz» — tzn. przez pewne ostensywne lub operacyjne procedury. Język teoretyczny może być interpretowany tylko pośrednio i «od wewnątrz» — tzn. przy pomocy postulatów sformułowanych w języku danej teorii. Zakłada się, że w ten sposób są interpretowane wszystkie terminy teoretyczne. I, co wydaje się nawet ważniejsze, przyjmuje się, że w ten sposób określane jest uniwersum przedmiotów teoretycznych. Zakłada się, że nie istnieje inna droga określenia dziedziny bytów teoretycznych niż postulowanie, aby była ona dziedziną, która spełnia dany zbiór zdań języka L: w naszym schemacie jest to zbiór T. Fakt ten jest kluczowym momentem przy określaniu charakteru klasy $N_{\mathfrak{M}_0}$.

Co obejmuje więc klasa $N_{\mathfrak{M}_0}$? Jak jest interpretowany termin t w jej strukturach? Można odwołać się do pewnych obserwacji, które dają częściową odpowiedź na to pytanie. Niech $\mathfrak{M}_0$ będzie modelem dla T_0 (zgodnie z naszymi założeniami, tylko takie struktury mogą być uważane za zamierzone modele dla L_0). Jeśli tylko zbiór T jest nietwórczy ze względu na T_0, to klasa $N_{\mathfrak{M}_0}$ musi być niepusta. Jest to zagwarantowane przez następujące syntaktyczne kryterium niepustości $N_{\mathfrak{M}_0}$:

> $N_{\mathfrak{M}_0} \neq \emptyset$, dla każdego $\mathfrak{M}_0 \in \mathrm{Mod}(T_0)$, zawsze i tylko, gdy każda czysto uniwersalna L_0-konsekwencja zbioru T jest konsekwencją zbioru T_0.

Jednocześnie łatwo zauważyć, że jeśli $N_{\mathfrak{M}_0}$ jest niepusta, to zawsze zawiera więcej niż jedną strukturę — w rzeczywistości zaś zawiera nieskończenie wiele struktur. Jest to niezależne od tego, jaki jest zbiór T. Tak więc, nawet jeśli t jest definiowalne w T, to ma ono zawsze nieskończenie wiele interpretacji w $N_{\mathfrak{M}_0}$, a w konsekwencji, nie jest identyfikowalne w T — w żadnym dopuszczalnym znaczeniu tego terminu (jedynym wyjątkiem jest wypadek, gdy z T wynika pustość t). Ponadto, nawet jeśli $|\mathfrak{M}_0|$ zawiera przedmioty obserwowalne, do zbioru $|\mathfrak{M}| - |\mathfrak{M}_0|$, dla pewnego $\mathfrak{M} \in N_{\mathfrak{M}_0}$, należeć będą pewne obiekty abstrakcyjne, np. liczby. Takie obiekty będą tworzyć interpretację stałej t w pewnych strukturach z $N_{\mathfrak{M}_0}$.[2]

Rozważmy bardziej szczegółowo interpretację t (dla uproszczenia notacji założę, że t jest predykatem jednoargumentowym). Otóż jakikolwiek byłby zbiór T, termin t, zinterpretowany w klasie $N_{\mathfrak{M}_0}$, okazuje się całkowicie nieostry poza dziedziną obserwowalną $|\mathfrak{M}_0|$. Precyzyjnie tezę tę wyrażają następujące zdania:

---

2) Por. np. Winnie (1967).

(1) $\bigcap_{\mathfrak{M}\in N_{\mathfrak{M}_0}} t^{\mathfrak{M}} \subset |\mathfrak{M}_0|$;

(2) $\bigcap_{\mathfrak{M}\in N_{\mathfrak{M}_0}} (|\mathfrak{M}| - t^{\mathfrak{M}}) \subset |\mathfrak{M}_0|$.

Jedynymi obiektami, które należą do każdej interpretacji t, albo nie należą do żadnej, są przedmioty obserwowalne. Wszystkie przedmioty teoretyczne należą do obszaru nieokreśloności tego predykatu. Załóżmy teraz, że t odnosi się jedynie do bytów teoretycznych. Jest to równoznaczne nałożeniu dodatkowego warunku na jego interpretację:

$$t^{\mathfrak{M}} \cap |\mathfrak{M}_0| = \emptyset, \text{ dla każdego } \mathfrak{M} \in N_{\mathfrak{M}_0}$$

W tym wypadku nieokreśloność stałej t staje się jeszcze bardziej uderzająca. Twierdzenia (1) i (2) redukują się teraz do twierdzeń

(3) $\bigcap_{\mathfrak{M}\in N_{\mathfrak{M}_0}} t^{\mathfrak{M}} = \emptyset$;

(4) $\bigcap_{\mathfrak{M}\in N_{\mathfrak{M}_0}} (|\mathfrak{M}| - t^{\mathfrak{M}}) = |\mathfrak{M}_0|$.

Przypadek definicyjnego rozszerzenia wymaga szczególnej uwagi. T otrzymuje się tutaj z T_0 przez dodanie do tego zbioru definicji terminu t (za pomocą terminów $o_1,\ldots,o_n$), np. w następujący sposób:

(D_t) $\forall x\,(t(x) \leftrightarrow \alpha_0(x))$,

gdzie $\alpha_0(x)$ należy do L_0. Ogólne konsekwencje, które naszkicowaliśmy wyżej są oczywiście prawdziwe również w tym wypadku. Można jednak tutaj poczynić pewne dodatkowe obserwacje. Aby je przedstawić rozróżnimy dwa rodzaje terminów teoretycznych: takie, które odnoszą się do przedmiotów obserwowalnych, i takie, które odnoszą się jedynie do nieobserwowalnych. Rozważmy teraz termin teoretyczny t pierwszego rodzaju i załóżmy, że pewien element w $|\mathfrak{M}_0|$ spełnia w $\mathfrak{M}_0$ definiens definicji D_t. Wtedy dla dowolnego zbioru U zawierającego $|\mathfrak{M}_0|$, istnieje struktura $\mathfrak{M}$ w $N_{\mathfrak{M}_0}$, taka że $|\mathfrak{M}| = U$ i $U - |\mathfrak{M}_0| \subset t^{\mathfrak{M}}$. Analogicznie, jeśli pewien element z $|\mathfrak{M}_0|$ nie spełnia w $\mathfrak{M}_0$ definiensa D_t, wtedy dla dowolnego zbioru U zawierającego $|\mathfrak{M}_0|$, istnieje struktura $\mathfrak{M}$ w $N_{\mathfrak{M}_0}$, taka że $|\mathfrak{M}| = U$ i $t^{\mathfrak{M}} \cap (U - |\mathfrak{M}_0|) = \emptyset$. Tak więc, wśród interpretacji terminu teoretycznego t pierwszego rodzaju, zawsze będą występowały interpretacje, zawierające wszystkie nieobserwowalne elementy danej dziedziny, oraz interpretacje, ktore nie zawierają żadnego z nich — czymkolwiek byłyby te elementy. Jeśli t jest terminem teoretycznym drugiego rodzaju, możemy założyć, że żadne elementy $|\mathfrak{M}_0|$ nie spełniają w $\mathfrak{M}_0$ definiensa D_t. Zatem, dla dowolnego zbioru U zawierającego $|\mathfrak{M}_0|$, zawsze istnieć będzie w $N_{\mathfrak{M}_0}$ struktura $\mathfrak{M}$, taka że $|\mathfrak{M}_0| = U$ i

$t^{\mathfrak{M}} = \emptyset$.Tak więc, wśród interpretacji terminu teoretycznego drugiego rodzaju zawsze znajdzie się interpretacja pusta.[3]

Wszystkie te obserwacje, chociaż fragmentaryczne i niekompletne, pozwalają nam zdać sobie sprawę, jak szeroka musi być klasa $N_{\mathfrak{M}_o}$ i jak niezdeterminowana wskutek tego pozostaje interpretacja terminu t. Czy termin teoretyczny zinterpretowany w ten sposób może spełnić swoje funkcje naukowe? W kwestii tej można argumentować następująco. Mimo bardzo niejednoznacznego zdeterminowania — interpretacja terminu teoretycznego t nie jest arbitralna. Klasa jego interpretacji jest ograniczona do zbiorów, które mają pewne określone własności strukturalne, i które pozostają w pewnych strukturalnych relacjach do interpretacji innych terminów. Dzięki tym ograniczeniom mogą pojawiać się zdania, które (1) zawierają termin t w sposób istotny, i zarazem (2) należą do zdań empirycznie rozstrzygalnych. W powyższym sformułowaniu występują pojęcia, które można wyjaśnić na różne sposoby. Przedstawiony poniżej sposób jest jedną z takich możliwych eksplikacji. Zdanie α języka L zawiera termin t w sposób istotny, jeśli — mówiąc swobodnie — jego wartość logiczna zależy od interpretacji t. Mówiąc ściślej:

α zawiera t w sposób istotny zawsze i tylko, gdy dla pewnego $\mathfrak{M}$ i $\mathfrak{N}$

(i) $\mathfrak{M}|_o = \mathfrak{N}|_o \in$ Mod (T_o);

(ii) $\mathfrak{M} \in$ Mod (α) i $\mathfrak{N} \in$ Mod $(\sim\alpha)$.

Aby wyjaśnić pojęcie empirycznej rozstrzygalności, założę, że zbiór T jest nietwórczy ze względu na T_o; może on być wtedy utożsamiony ze zbiorem postulatów w L. Przy tym założeniu, pojęcie empirycznej rozstrzygalności może być zdefiniowane w następujący sposób. Zdanie α języka L jest empirycznie rozstrzygalne, jeśli — mówiąc swobodnie — jest ono syntetyczne i jego wartość logiczna jest niezmienna w pewnej klasie $N_{\mathfrak{M}_o}$. W zapisie formalnym:

α jest empirycznie rozstrzygalne zawsze i tylko, gdy

(i) $\alpha \notin$ Cn (T) i $\sim\alpha \notin$ Cn (T);

(ii) dla pewnego $\mathfrak{M}_o \in$ Mod (T_o), $N_{\mathfrak{M}_o} \subset$ Mod (α) lub $N_{\mathfrak{M}_o} \subset$ Mod $(\sim\alpha)$.

Jakie dokładnie zdania, zawierające termin t w sposób istotny, będą należeć do klasy zdań rozstrzygalnych empirycznie, zależy teraz oczywiście od tego, jakie są postulaty dla t. Chciałbym dodać jeszcze jedną tylko uwagę na ten temat. Sądzę, że postulaty dla teminu teoretycznego t nie mogą składać się z definicji t zbudowanej za pomocą predykatów obserwacyjnych $o_1,...,o_n$, jeśli klasa empirycznie rozstrzygalnych zdań, które zawierają t w sposób istotny, miałaby zawierać te zdania o t, które są faktycznie uznawane lub wysuwane przez naukowców. Jako przykład bardziej adekwatnych postulatów dla terminu teoretycznego t (typu drugiego), należałoby raczej wymienić postulat następującego rodzaju:

---

3) Fakty te można łatwo wskazać przy założeniu, że struktury dla L nie są ograniczone do struktur normalnych, tzn. struktur ze standardową interpretacją identyczności. Identyczność w takich strukturach może być interpretowana w sposób wspomniany wyżej.

(P_t) $\forall x(\exists y\, (\beta_o(x, y) \wedge t(y)) \leftrightarrow \alpha_o(x))$,

gdzie $\beta_o(x, y)$ i $\alpha_o(x)$ należą do L_o. Jest to typ postulatu proponowany i dyskutowany przez kilku autorów. Dyskusja ta dostarczyła pewnych argumentów na rzecz powyższej tezy.[4]

Omawiając problem identyfikowalności w dziedzinie rozszerzonej, ograniczyliśmy się do pewnego szczególnego przypadku tej ogólnej kwestii i pewnej szczególnej interpretacji rozważanego przypadku. Przypadek ten polegał na rozszerzeniu języka obserwacyjnego przy pomocy pewnych terminów teoretycznych, a jego interpretacja oparta była na założeniach empiryzmu semantycznego. Problem jest jednak ogólniejszy — nie ograniczony do tego szczególnego przypadku i jego interpretacji. Ogólny schemat przedstawiony wyżej może być z powodzeniem zastosowany do innego rodzaju rozszerzeń języka, a ich interpretacja może być oparta na innych koncepcjach semantycznych. W wypadku takich zastosowań, problem nasz wymaga pewnego przeformułowania. Definicja klasy $N_{\mathfrak{M}_o}$ może wymagać pewnych dodatkowych założeń dotyczących charakterystyki rozszerzonego uniwersum i interpretacji dotychczasowych terminów w tym uniwersum. Obie te rzeczy możemy traktować jako wyznaczone nie tylko przez postulaty, lecz także przez pewne metody bespośrednie. W skrajnym wypadku, dziedzina rozszerzonej struktury może być określona jednoznacznie: wszystkie elementy $N_{\mathfrak{M}_o}$ będą miały wspólną dziedzinę:

$|\mathfrak{M}| = U$, dla każdego $\mathfrak{M} \in N_{\mathfrak{M}_o}$.

Jednym z dopuszczalnych sposobów, w jaki określona może być interpretacja dotychczasowych terminów w rozszerzonej dziedzinie, jest sposób przedstawiony wyżej, w związku z analizą interpretacji predykatów obserwacyjnych. Takie dodatkowe ograniczenia muszą wpłynąć na nasze wnioski dotyczące stopnia identyfikowalności (albo raczej nieidentyfikowalności) rozważanych terminów. Możliwości te wydają się warte rozważenia.

Tłumaczyła Anna Lissowska

BIBLIOGRAFIA

Przełęcki, M.: 1976, „Interpretation of theoretical terms: in defence of an empiricist dogma", [w:] M.Przełęcki et al. (wyd.), *Formal methods in the methodology of empirical sciences*, Dordrecht - Wrocław, D.Reidel - Ossolineum.

Winnie, J. A.: 1967, „The implicit definition of theoretical terms", *The British Journal for the Philosophy of Science*, 18, 223-229.

4) Patrz np. Przełęcki (1976).

Filozofia Nauki
Rok I, 1993, Nr 2-3

Ciągłość pojęciowa przy zmianach teorii

I

Niniejszy artykuł, pomyślany jako przyczynek do «teorii racjonalnego rozwoju nauki»[1], dotyczy natury zmian pojęciowych, charakterystycznych dla pewnych typów zmian teorii. Pewien rodzaj ciągłości pojęciowej jest zwykle uważany za warunek konieczny racjonalnego rozwoju nauki. Pogląd ten jest jednak obecnie poważnie kwestionowany. Argumentacja jego przeciwników przebiega, ogólnie rzecz biorąc, w następujący sposób. Niech T_1 i T_2 będą dwiema teoriami, następującymi po sobie w procesie rozwoju nauki. Interpretacja (sens i odniesienie) ich terminów specyficznych zależy od tych właśnie teorii. Ponieważ teorie różnią się tym, co mówią o swoich terminach, interpretacja tych ostatnich jest, jak się wydaje, inna w obu teoriach. W wypadku teorii wzajemnie niezgodnych, ich języki uważa się za wzajemnie nieprzekładalne, a ich terminy specyficzne za «niewspółmierne». To jednak wydaje się wykluczać możliwość logicznego porównania i racjonalnego wyboru między tymi dwiema teoriami. Aby móc traktować teorie jako «konkurencyjne», musimy założyć, że mówią one przynajmniej częściowo o tym samym. A to wydaje się zakładać pewien rodzaj «współmierności» ich struktury pojęciowej. Głównym celem niniejszej analizy jest przedstawienie argumentów na poparcie pewnej wersji tezy o «współmierności». Zmiany w interpretacji terminów jakiejś teorii, pojawiające się w procesie rozwoju tej teorii, mogą pociągać zarówno zmianę ich sensu, jak i odniesienia. Tym, co wydaje się istotne przy porównywaniu dwóch następujących po sobie teorii, szczególnie gdy są one sformułowane w języku ekstensjonalnym, jest pewna identyczność odniesienia. I to

1) Teoria taka jest postulowana jako jeden z tematów w preambule programu badawczego *Foundations of science and ethics*.

właśnie chciałbym pokazać w niniejszym artykule.[2] Przedstawiona argumentacja będzie oparta na pewnej specyficznej koncepcji semantyki dla teorii empirycznych, która z formalnego punktu widzenia może być nazwana teoriomodelową, a z filozoficznego punktu widzenia — empiryczną. Jest ona z grubsza zgodna z Carnapowską semantyką dla języków i teorii empirycznych.[3] Krytyka tego podejścia zatem odnosi się również i do tej koncepcji. Jej uzasadnieniem i obroną nie będziemy się jednak zajmować w tym artykule.[4]

Wedle wspomnianej koncepcji, każdy zinterpretowany język empiryczny może być utożsamiony z systemem semantycznym:

$$L = \langle F, A, M \rangle,$$

składającym się z języka formalnego F i dwóch klas struktur dla F: $M \subseteq A$, gdzie A reprezentuje możliwe interpretacje L («możliwe światy»), zaś M jest faktyczną interpretacją L («światem rzeczywistym»). A może być w związku z tym rozumiane jako czynnik determinujący sens wyrażeń z L, M — jako czynnik determinujący ich odniesienie. Zgodnie z ogólnie przyjętą praktyką, sens terminu t, $s(t)$, będzie utożsamiany z funkcją, która każdej możliwej strukturze $\mathfrak{m}$ dla języka L przyporządkowuje denotację $t^{\mathfrak{m}}$ terminu t w tej strukturze:

$$s(t)\colon\ \mathfrak{m} \in A \rightarrow t^{\mathfrak{m}}.$$

Odniesienie terminu t, $r(t)$, będzie z kolei zdefiniowane jako klasa denotacji terminu t we wszystkich faktycznych strukturach dla L:

$$r(t) = \{t^{\mathfrak{m}}\}_{\mathfrak{m} \in M}.$$

Będziemy się zajmować przede wszystkim tak rozumianymi odniesieniami terminów. Obiektem naszych analiz będą języki L_1 i L_2 dwóch teorii empirycznych T_1 i T_2, które będą reprezentować dwa następujące po sobie etapy rozwoju tego, co zwykle nazywa się daną teorią empiryczną. O interpretacji języka L_i $(i = 1, 2)$ zakłada się, że jest ona zdeterminowana przez następujące założenia, będące głównymi zasadami przyjętego w tym artykule stanowiska empirycznego.

(i) Zakładamy, że pozalogiczne słowniki języków L_i zawierają wspólny podsłownik terminów nazywanych terminami nieteoretycznymi ze względu na obie teorie T_i — w skrócie o-terminami. Ich interpretacja, dana z góry, jest niezależna od żadnej z tych teorii. Zakłada się przy tym, że jest ona taka sama w obu teoriach. (Później osłabimy nieco to założenie.) Tak więc L_1 i L_2 można traktować jako rozszerzenia tego samego systemu semantycznego:

$$L_o = \langle F_o, A_o, M_o \rangle,.$$

Założenie to gwarantuje pewnego rodzaju ciągłość pojęciową między dwiema teoriami T_1 i T_2, taką mianowicie, która jest ograniczona do ich nieteoretycznych struktur.

2) Główna idea tej argumentacji jest przedstawiona w Przełęcki (1978).

3) Por. np. Carnap (1966). Wersja teoriomodelowa tej koncepcji, z której robi się użytek w tym artykule, była rozwijana m.in. w Przełęcki (1969) i Williams (1973).

4) Co do pewnych argumentów na ten temat patrz np. Przełęcki (1969).

Założenie to wydaje się przekonujące pod warunkiem rozumienia takich teorii w dostatecznie szeroki sposób, tak aby do danej teorii weszły wszystkie zakładane przez nią podteorie (w szczególności wszystkie związane z nią teorie pomiaru). Schodząc dostatecznie głęboko w tej hierarchii, możemy mieć nadzieję, że dotrzemy do poziomu, który jest neutralny ze względu na obie teorie, i który dlatego może być uważany za identyczny dla nich obu. Zakwestionowanie tego założenia zwykle wynika z przyjęcia restryktywniejszej koncepcji teorii naukowej, utożsamiającej tę teorię z najwyższą warstwą całej wchodzącej w grę hierarchii teoretycznej.

(ii) Zakłada się, że oprócz *o*-terminów, do pozalogicznego słownika każdego języka L_i należą pewne terminy nazywane terminami teoretycznymi ze względu na teorię T_i — w skrócie: *t*-terminy. Przyjmiemy tutaj, że są one oznaczane przez różne symbole w różnych teoriach (powiedzmy przez t_1 w teorii T_1 i t_2 w teorii T_2). Interpretacja *t*-terminów w teorii T_i zależy od tej teorii. Przy tym ujęciu, oznacza to utożsamienie postulatów znaczeniowych dla *t*-terminów języka L_i z tzw. składnikiem konwencjonalnym teorii T_i. We współczesnej literaturze zaproponowano kilka definicji tego pojęcia. Oryginalna propozycja Carnapa, stosowalna do teorii skończenie aksjomatyzowalnych, utożsamia konwencjonalny składnik teorii T_i ze zdaniem ${}^RT_i \rightarrow T_i$, gdzie ${}^RT_i$ reprezentuje tzw. zdanie Ramseyowskie dla T_i (tzn. egzystencjalne domknięcie formuły otrzymanej z T_i przez jednoczesne podstawienie odpowiednich zmiennych pod wszystkie *t*-terminy). Przy pomocy zwykłej notacji teoriomodelowej (oznaczając, w szczególności, struktury dla L_o i L_i — przez $\mathfrak{m}_o$ i $\mathfrak{m}_i$, fragment $\mathfrak{m}_i$ odpowiadający L_o — przez $\mathfrak{m}_i|_o$, a klasę modeli zdania lub zbioru zdań X przez $\mathrm{Mod}(X)$), zdefiniujemy język L_i teorii T_i jako system semantyczny:

$$L_i = \langle F_i, A_i, M_i \rangle, \text{ gdzie}$$
$$A_i = \{\mathfrak{m}_i : \mathfrak{m}_i|_o \in A_o \text{ i } \mathfrak{m}_i \in \mathrm{Mod}({}^RT_i \rightarrow T_i)\},$$
$$M_i = \{\mathfrak{m}_i : \mathfrak{m}_i|_o \in M_o \text{ i } \mathfrak{m}_i \in \mathrm{Mod}({}^RT_i \rightarrow T_i)\},$$

Tak zdefiniowane A_i zawiera wszystkie przedłużenia struktur w A_o do modeli składnika konwencjonalnego teorii T_i; M_i — wszystkie przedłużenia struktur w M_o do modeli tego składnika. Propozycja ta wydaje się przekonująca pod warunkiem, iż nie ma powodów, aby interpretacja *t*-terminów w teorii T_i zależała nie od całej teorii T_i, lecz od pewnej określonej jej części, wyróżnionej «z zewnątrz», przez pewne czynniki pragmatyczne. W dalszych rozważaniach będziemy zakładać, że takie powody nie istnieją.

II

Można pokazać, że klasa M_i, reperezentująca faktyczną interpretację języka L_i, posiada pewne istotne — ze względu na dalsze rozważania — własności. Mam tu na myśli niejednoznaczność wszelkiej interpretacji empirycznej. M_i nie jest tu co prawda utożsamiana z tym, co czasami nazywa się dziedziną zamierzonych zastosowań teorii T_i.[5] Zgodnie z tradycyjnym, empirystycznym podejściem przyjętym w tym artykule, M_i

5) Patrz Sneed (1971).

odpowiada nie klasie różnych, poszczególnych dziedzin zastosowań, lecz raczej jednej maksymalnej dziedzinie zastosowań, zdefiniowanej np. jako suma wszystkich zastosowań zamierzonych. Mimo to, M_i nie może być utożsamiana z pojedynczą strukturą dla języka L_i — lecz z pewną klasą takich struktur. Wynika to z charakterystycznej własności języka empirycznego: z faktu jego «rozmycia». Zaliczyć do niego można takie zjawiska jak nieostrość w odniesieniu do pojęć jakościowych i aproksymacyjność w odniesieniu do pojęć ilościowych. Każdy jakościowy termin empiryczny, w szczególności każdy predykat empiryczny, jest mniej lub bardziej nieostry. Wynika to ze sposobu, w jaki predykatowi empirycznemu przyporządkowuje się jego interpretację. Każda taka procedura pozostawia pewien obszar niezdeterminowania. Dla pewnych obiektów nie istnieją kryteria podpadania pod dany predykat: nie jest rozstrzygnięte, czy należą one, czy też nie, do jego denotacji. Tak więc, oprócz przypadków pozytywnych i negatywnych zawsze istnieją przypadki graniczne danego predykatu. Jednym ze sposobów zdania sprawy z tego faktu jest utożsamienie interpretacji predykatu nie z jednym zbiorem (albo relacją), ale z pewną klasą zbiorów (albo relacji). Każdy z tych zbiorów (albo relacji) odpowiada możliwej klasyfikacji wszystkich przypadków granicznych danego predykatu, jako jego pozytywnych lub negatywnych przykładów. To z kolei powoduje utożsamienie faktycznej interpretacji języka empirycznego nie z pojedynczą strukturą, lecz z pewną klasą struktur. Każda z tych struktur zawiera, jako denotację danego terminu, jeden element z klasy tworzącej jego interpretację.[6]

Języki teorii empirycznych, będące przedmiotem naszych rozważań, należą do języków ilościowych. Oprócz predykatów jakościowych, zawierają one zawsze pewne terminy ilościowe: symbole wielkości empirycznych. Są one wprowadzane, zgodnie ze współczesnymi teoriami pomiaru bezpośredniego, za pomocą pewnych jakościowych pojęć empirycznych. W konsekwencji, nieostrość tych ostatnich wpływa na charakter tych pierwszych, sprawiając, że są one z istoty aproksymacyjne. Ponieważ zjawisko aproksymacji jest bardzo ważne dla naszych rozważań, przeanalizujmy je dokładniej.[7] Teorie pomiaru właściwego definiują wielkości (cechy ilościowe) jako liczbowe odpowiedniki pewnych struktur empirycznych. W wypadku wielkości ekstensywnych, które stanowią główny typ wielkości w teoriach empirycznych, wspomniane struktury empiryczne należą do tzw. systemów ekstensywnych. Ich istotnym składnikiem jest empiryczna relacja porządkująca, scharakteryzowana przez standardowe aksjomaty dla systemów ekstensywnych jako słaby porządek określony na uniwersum danej struktury. (Często przez system ekstensywny rozumie się system zawierający, oprócz relacji porządkującej, empiryczną operację konkatenacji, ale nie wydaje się, żeby to był warunek konieczny, ponieważ jej rolę może odgrywać operacja sumowania teoriomno-

6) Pełniejsze przedstawienie takiego podejścia do nieostrości (niejednoznaczności) i jej semantyki można znaleźć np. w Przełęcki (1969, 1976).

7) Bardziej szczegółowa analiza problemów pomiaru przybliżonego jest zawarta w Przełęcki (1979). Na temat rozważań dotyczących teorii systemów ekstensywnych patrz np. Suppes (1969).

gościowego.) Relacja porządkująca systemu ekstensywnego jest typowym przykładem pojęcia nieostrego, i właśnie ta jej nieostrość tłumaczy , moim zdaniem, aproksymacyjny charakter danej wielkości. Ten punkt wymaga jednak pewnych komentarzy ponieważ czasami jest źródłem nieporozumień.

Relacja porządkująca zakłada pewną procedurę porównywania dwóch obiektów ze względu na daną wielkość, przy pomocy pewnego przyrządu pomiarowego. Wszelkie takie przyrządy, a w konsekwencji wszelkie takie procedury, są zawsze pod pewnym względem niedokładne. Niedokładność ta, jak zwykle podkreśla się w tym kontekście, polegać ma na ograniczonej czułości przyrządów. Żaden rzeczywisty przyrząd pomiarowy nie jest doskonale czuły. Każdy przyrząd jest niewrażliwy na różnice pod danym względem mniejsze niż pewna skończona wielkość nazywana jego rozdzielczością. W rezultacie, dana relacja porządkująca R_o, zdefiniowana bezpośrednio przy pomocy takiej procedury pomiarowej, nie ma pewnych formalnych własności wymaganych od relacji porządkującej w systemie ekstensywnym: w każdym dostatecznie obszernym zbiorze U istnieją obiekty x, y i z, takie że chociaż R_o nie zachodzi między x i y, i między y i z, to zachodzi między x i z, W efekcie R_o jest tzw. relacją semiporządkującą w zbiorze U. Formuła xR_oy może być rozumiana jako twierdzenie, że obiekt x jest wyraźnie większy pod danym względem niż obiekt y.[8] Ten rodzaj niedokładności procedury pomiarowej może być jednak usunięty przez przejście od czysto obserwacyjnej relacji R_o do bardziej teoretycznej relacji R, zdefiniowanej przy pomocy tej pierwszej w następujący sposób:

xRy zawsze i tylko, gdy dla każdego z: jeśli zR_ox, to zR_oy, i jeśli yR_oz, to xR_oz.

Formuła xRy może być interpretowana jako twierdzenie, że obiekt x jest przynajmniej tak duży pod danym względem jak obiekt y. Można wykazać, że tak zdefiniowana relacja R jest słabym porządkiem w U, i że w związku z tym spełnia (łącznie z operacją konkatenacji) standardowe aksjomaty systemu ekstensywnego. W ten sposób jesteśmy w stanie obejść nieczułość danej procedury pomiarowej i niedokładność odpowiadającego jej pomiaru. Jeśli zbiór U zawiera obiekty dowolnie mało się różniące pod względem danej wielkości, to dokładność z jaką możemy zmierzyć tę wielkość, jest teoretycznie nieograniczona.

Przeczy to jednak, jak się wydaje, faktycznej praktyce naukowej, która przemawia za aproksymacyjną naturą każdej wielkości empirycznej. Aby to wyjaśnić, musimy zdać sobie sprawę, że to, co nazywa się niedokłądnością danego przyrządu pomiarowego i procedury pomiarowej obejmuje dwa różne zjawiska: nieczułość i niejednozna-

8) Jako przykład takiej semiporządkującej relacji R_o można podać relację bycia cięższym, której definicja opiera się na procedurze porównywania ciężaru dwóch obiektów za pomocą wagi szalkowej: obiekt x jest cięższy niż obiekt y zawsze i tylko, gdy jeżeli oba obiekty są umieszczone na różnych szalkach wagi, to szalka z obiektem x znajduje się niżej niż szalka z y. Analiza tego przykładu jest przedstawiona w Przełęcki (1979).

czność. Przyrząd pomiarowy może nie reagować w ogóle, albo reagować niejednoznacznie. Nie rozróżnia on pewnych obiektów lub dokonuje takiego rozróżnienia niejednoznacznie. Na czym taka niejednoznaczność miałaby polegać? Przede wszystkim, poza przypadkami jednoznacznego zachowania przyrządu istnieją zawsze przypadki niezdeterminowane: takie w których nie możemy powiedzieć, czy przyrząd zareagował czy nie.[9] Jeśli procedura porównywania dwóch obiektów nie jest ograniczona do pojedynczego testu, lecz składa się z serii takich testów (jak to się zwykle faktycznie dzieje), to pojawiają się nowe źródła niejednoznaczności: może się zdarzyć, że zastosowany do pewnego obiektu przyrząd nie zachowuje się tak samo podczas wszystkich testów w danej serii — w pewnych sytuacjach reaguje, a w innych nie.[10] Wszystkie wypadki «niezdecydowanego» lub «niekonsekwentnego» zachowania przyrządu, wypadki, w których otrzymujemy niezdeterminowane lub «rozstrzelone» wyniki, tworzą graniczne przypadki naszej wyjściowej relacji R_o. Jaki rodzaj obiektów należy do tych granicznych przypadków? O ile nieczułość danej procedury przejawia się w jej skończonej rozdzielczości, o tyle niejednoznaczność tej procedury jest powodem niedokładności, z jaką ta rozdzielczość jest określona. Chodzi o to, że jest ona równa nie pewnemu q, lecz $q \pm \varepsilon$. Inaczej mówiąc: nasza procedura rozróżnia obiekty, które różnią się pod danym względem więcej niż $q + \varepsilon$, nie rozróżnia tych, które różnią się mniej niż $q - \varepsilon$, różnicuje niejednoznacznie te, których różnica mieści się w przedziale od $q - \varepsilon$ do $q + \varepsilon$. Wszystkie pary obiektów trzeciego rodzaju tworzą graniczne przypadki relacji R_o. W przeciwieństwie do nieczułości danej procedury, jej niejednoznaczność nie może być usunięta przez przejście od relacji R_o do relacji R, zdefiniowanej we wskazany wyżej sposób. Niejednoznaczność R_o przenosi się na R. Nie zostaje ona w ten sposób usunięta, lecz jedynie częściowo zmniejszona. Łatwo zauważyć, że tylko wtedy gdy różnica co do danej wielkości między x i y jest większa niż 2ε, para $\langle x, y\rangle$ jest zdeterminowana, tzn. jest negatywnym albo pozytywnym przypadkiem relacji R. Jeśli ta różnica nie osiąga 2ε, mamy do czynienia z granicznym przypadkiem R.

Przy przedstawionym wyżej ujęciu nieostrości, relacja niejednoznaczna zostaje utożsamiona z pewną klasą «ostrych» relacji, z których każda odpowiada pewnemu możliwemu zaklasyfikowaniu wszystkich przypadków granicznych do przypadków pozytywanych albo negatywnych. Odpowiedniki naszych niejednoznacznych relacji R_o i R są więc pewnymi klasami «ostrych» relacji: R_o^* i R^*. Zakłada się, że klasa R_o^* jest ograniczona tylko do tych relacji, które są semiporządkami w zbiorze U. Klasa R^*, która ma zawierać wszystkie relacje zdefiniowane przez poprzednie relacje w zaproponowany powyżej sposób, będzie w konsekwencji składać się z pewnych słabych porządków w U. Jeżeli teraz potraktujemy wszystkie te relacje jako nasze podstawowe relacje porządkujące, to w konsekwencji otrzymamy całą klasę systemów ekstensyw-

9) Np. czy szalka wagi szalkowej wychyliła się czy nie, w sytuacjach, w których poruszyła się ona tylko nieznacznie.

10) Np. w niektórych testach danej serii szalki wagi będą w równowadze, a w innych nie.

nych S^*, z których każdy zawiera jedną relację z klasy R^*. Na ogół klasa S^* jest nawet obszerniejsza niż klasa R^*, ponieważ pozostałe składniki systemu ekstensywnego — w szczególności empiryczną operację konkatenacji (jeśli oczywiście taka występuje) — również trudno jednoznacznie zdefiniować. Każdemu ekstensywnemu systemowi S w S^* odpowiada jedna «ścisła» wielkość empiryczna, tzn. jedna funkcja F o wartościach rzeczywistych, zdefiniowana jako pewien homomorfizm danego systemu ekstensywnego w strukturę liczbową odpowiedniego typu. Biorąc pod uwagę standardowe systemy ekstensywne z operacją konkatenacji o, możemy przedstawić powyższe rozważania bardziej precyzyjnie w następujący sposób:

Jeśli $S = \langle U, R, o\rangle$ jest systemem ekstensywnym i $u \in U$, to istnieje dokładnie jedna funkcja F o dodatnich wartościach rzeczywistych na U, taka że dla każdego $x,y \in U$:

(i) xRy zawsze i tylko, gdy $F(x) \geq F(y)$;

(ii) $F(xoy) = F(x) + F(y)$;

(iii) $F(u) = 1$.

Wszystkie funkcje zdefiniowane w powyższy sposób tworzą klasę F*, o której można powiedzieć, że reprezentuje przybliżoną wielkość mierzoną przez daną procedurę. Zgodnie z tym ujęciem, przybliżona wielkość nie jest niczym innym, jak tylko klasą pewnych wielkości ścisłych. Skład klasy F^* jest wyznaczony przez skład klasy R^*. Pamiętając, co zakładaliśmy o tej ostatniej, łatwo ustalimy, jaki rodzaj funkcji będzie tworzył klasę F^*. Wartości dwóch takich funkcji dla danego obiektu będą leżały zawsze wewnątrz pewnego ustalonego «przedziału niedokładności»: dla każdej funkcji $F \in F^*$ i obiektu $x \in U$, $F(x) \in [k - \varepsilon, k + \varepsilon]$, dla pewnego k. Taka charakterystyka klasy F^* wydaje się usprawiedliwiać utożsamianie jej z tym, co zwykle nazywa się wielkością przybliżoną.

Interpretując jakościowe i ilościowe terminy empiryczne przez pewne klasy ich standardowych denotacji — symbole relacji przez klasy odpowiednich relacji, a symbole funkcji przez klasy funkcji odpowiedniego typu — utożsamiamy w rezultacie faktyczną interpretację języka L z pewną klasą jego struktur, M. Jest to w szczególności twierdzenie prawdziwe również o naszym empirycznym języku L_i i jego podjęzyku L_o. Zarówno M_i jak i M_o są rozumiane jako klasy struktur, zawierające więcej niż jeden element. Takie podejście do koncepcji interpretacji rodzi pewne problemy związane z pojęciem prawdy dla tak zinterpretowanych języków. Zaproponowano pewne rozwiązania tego problemu, do których możemy jedynie odwołać się w tym miejscu.[11] Ich główna idea jest stosunkowo prosta. Zdanie prawdziwe w L_i jest zdefiniowane jako zdanie prawdziwe w każdej strukturze z M_i. O zdaniu fałszywym w każdej takiej strukturze powiemy, że jest fałszywe w L_i. Zdania, które są prawdziwe w pewnych strukturach a fałszywe w innych, określamy jako zdania ani prawdziwe ani fałszywe w L_i. W stosunku do nich znajduje zastosowanie pojęcie prawdy aproksymacyjnej. Zdanie jest aproksymacynie prawdziwe w języku L_i jeżeli jest prawdziwe przynajmniej w

---

11) Patrz np. Przełęcki (1969, 1976).

jednej strukturze z M_i. Zbiór zdań T_i, jest aproksymacyjnie prawdziwy w L_i jeśli wszystkie zdania należące do T_i są prawdziwe przynajmniej w jednej strukturze z M_i. Zobaczmy, jak stosuje się ta definicja do zdania zawierającego symbol funkcyjny f, interpretowany jako symbol odnoszący się do wielkości przybliżonej w sensie tu wprowadzonym, tzn. do klasy funkcji F^*. Zdanie takie jest prawdziwe, jeśli jest prawdziwe niezależnie od tego, którą funkcję wybierzemy z F^* jako interpretację f; jest fałszywe, jeśli jest fałszywe przy każdej takiej interpretacji; w pozostałych wypadkach jest aproksymacyjnie prawdziwe. Tak więc, przy założeniu, że dla każdego $F \in F^*$, $F(a) \in$ $\in [k - \varepsilon,\ k + \varepsilon]$, zdanie $f(a) \leq k + \varepsilon$ będzie zakwalifikowane jako prawdziwe, jego negacja $f(a) > k + \varepsilon$ jako fałszywa, a zdanie $f(a) = k$ jako aproksymacyjnie prawdziwe. Trzeba zauważyć, że każde ilościowe zdanie tego ostatniego typu może być, ogólnie rzecz biorąc, tylko aproksymacyjnie prawdziwe w L_i (chyba, że okaże się ono konsekwencją postulatów znaczeniowych języka L_i, jak to jest w wypadku zdania $f(u) = 1$).

III

Porównajmy teraz, przy powyższych założeniach, interpretację dwóch terminów teoretycznych t_1 i t_2, odpowiednio w dwóch teoriach T_1 i T_2. Skoncentrujemy się na wypadku, w którym problem ich «współmierności» staje się szczególnie wyraźny. Dzieje się tak w wypadku wzajemnej niezgodności teorii. Co się tutaj przez to rozumie, można wyjaśnić w następujący sposób.

Teorie T_1i T_2 są niezgodne, jeśli istnieją zdania V_o i W_o języka L_o, takie że $W_o \in$ $\in \mathrm{Cn}(T_1 \cup \{V_o\})$, $\neg W_o \in \mathrm{Cn}(T_2 \cup \{V_o\})$ i V_o jest prawdziwe (tzn. prawdziwe we wszystkich strukturach w M_o).

O T_1 i T_2 powiemy więc, że są niezgodne, jeśli przy pewnym dodatkowym założeniu, które jest prawdziwe w L_o (przy jego faktycznej interpretacji), pociągają za sobą sprzeczne o-zdania.

Łatwo zauważyć, że w takich wypadkach nie można utożsamiać sensów terminów t_1 i t_2. Jeśli T_1 i T_2 nie są logicznie równoważne, to ich konwencjonalne składowe ${}^RT_1 \to T_1$ i ${}^RT_2 \to T_2$ również nie są równoważne (z wyjątkiem banalnego wypadku, w którym t-terminy występują w tych teoriach w nieistotny sposób[12]), i, w konsekwencji, $s(t_1) \neq s(t_2)$. Oznacza to niezgodność sensów terminów, jednakże nie pociąga niezgodności ich odniesień. Spróbujemy pokazać, że terminy teoretyczne dwóch niezgodnych teorii mogą w istocie mieć identyczne (lub częściowo identyczne) odniesienia. Ten punkt wymaga jednak pewnego wyjaśnienia ponieważ identyczność ta może mieć dość banalne powody. Jeśli ${}^RT_i$ jest fałszywe w pewnej strukturze w M_o, to wszystkie przedłużenia tej struktury będą należeć do M_i, i klasa odniesień terminu t_i, $r(t_i)$, będzie zawierać wszystkie możliwe interpretacje tego terminu (wewnątrz uniwersum M_o). Będzie to w efekcie gwarantować pewien rodzaj odniesieniowej identyczności między t_1 i t_2: $r(t_1) = r(t_2)$, albo przynajmniej $r(t_1) \cap r(t_2) \neq \emptyset$. Jeśli jednak dana struktura w M_o

12) Patrz Przełęcki (1978).

nie jest modelem ${}^RT_i$, a zatem nie jest przedłużalna do modelu T_i, to wszystkie jej przedłużenia do struktury dla L_i, jak również wszystkie interpretacje t_i, są — ściśle mówiąc — niezamierzone, gdyż nie spełniają waruków wyrażonych przez T_i. Tak więc, kiedy porównujemy $r(t_1)$ i $r(t_2)$, powinniśmy raczej tamte interpretacje pominąć i ograniczyć się do tych interpretacji t_i, które spełniają warunki nałożone przez T_i. Takie interpretacje będziemy nazywać interpretacjami zamierzonymi w ścisłym sensie i definiować w następujący sposób:

$$M_i^* = M_i \cap \mathrm{Mod}(T_i)\,,$$

$$r^*(t_i) = \{t_i^{\mathfrak{m}_i}\}_{\mathfrak{m}_i \in M_i^*}$$

W naszym wypadku więc porównywane winny być klasy odniesień zamierzonych $r^*(t_1)$ i $r^*(t_2)$.

Teorie T_1 i T_2, jako niezgodne, nie mogą być jednocześnie prawdziwe (w rozważanym sensie). Jeśli natomiast jedna z nich, powiedzmy T_1, jest fałszywa, to nie istnieją zamierzone interpretacje jej terminu teoretycznego t_1: $r^*(t_1) = \emptyset$, i problem porównania odniesień staje się banalny. Jest tak tym bardziej, jeśli obie okażą się fałszywe. To, że klasy M_o i M_i są wieloelementowe, otwiera jednak jeszcze inną możliwość: obie teorie mogą być aproksymacyjnie prawdziwe. Ponieważ T_1 jest niezgodna z T_2, nie istnieje struktura w M_o, w której zarówno ${}^RT_1$, jak i ${}^RT_2$ będą prawdziwe. Z powodzeniem jednak mogą istnieć dwie różne struktury w M_o, $\mathfrak{m}_o$ i $\mathfrak{n}_o$, takie że ${}^RT_1$ jest prawdziwe w $\mathfrak{m}_o$, a ${}^RT_2$ w $\mathfrak{n}_o$. Klasy zamierzonych odniesień t_1 i t_2 są wtedy niepuste, i problem ich porównania pozostaje otwarty. Łatwo w tej sytuacji wskazać wypadki, w których $r^*(t_1)$ i $r^*(t_2)$ mają pewne elementy wspólne. Weźmy następujący schemat:

$$T_1 = \bigwedge x\, ((o_1(x) \to t_1(x)) \wedge (o_2(x) \to \neg t_1(x))),$$

$$T_2 = \bigwedge x\, ((o_1(x) \to t_2(x)) \wedge (\neg o_2(x) \to \neg t_2(x))).$$

${}^RT_1$ i ${}^RT_2$ redukują się w tym wypadku do zdań rachunku pierwszego rzędu:

$${}^RT_1 = \bigwedge x\, ((o_1(x) \to \neg o_2(x)),$$

$${}^RT_2 = \bigwedge x\, ((o_1(x) \to o_2(x)).$$

Niech M_o zawiera struktury $\mathfrak{m}_o$ i $\mathfrak{n}_o$, zdefiniowane w następujący sposób:

uniwersum $\mathfrak{m}_o$ = uniwersum $\mathfrak{n}_o = \{a, b, c\}$;

$$o_1^{\mathfrak{m}_o} = o_1^{\mathfrak{n}_o} = \{b\},\quad o_2^{\mathfrak{m}_o} = \{c\},\quad o_2^{\mathfrak{n}_o} = \{b, c\}.$$

Oczywiście $\mathfrak{m}_o \in \mathrm{Mod}({}^RT_1)$, $\mathfrak{n}_o \in \mathrm{Mod}({}^RT_2)$, i wśród ich przedłużeń odpowiednio do modeli T_1 i T_2 istnieją struktury $\mathfrak{m}_1 \in M_1^*$ i $\mathfrak{m}_2 \in M_2^*$, takie że $t_1^{\mathfrak{m}_1} = t_2^{\mathfrak{m}_2} = \{b\}$. Pokazaliśmy więc, że terminy teoretyczne t_1 i t_2 zdefiniowane przez dwie niezgodne teorie T_1 i T_2 mają pewne wspólne odniesienia zamierzone: $r^*(t_1) \cap r^*(t_2) \neq \emptyset$.

Sytuacja ta, najogólniej rzecz biorąc, może być przedstawiona w następujący sposób. Teorie T_1 i T_2 charakteryzują swoje t-terminy t_1 i t_2 na dwa różne sposoby. Ta różnica jest jednakże kompensowana w strukturach, takich jak $\mathfrak{m}_1$ i $\mathfrak{m}_2$, omawiane wyżej, przez odpowiednie różnice w interpretacji pewnych o-terminów ($\mathfrak{m}_1|_o$ i $\mathfrak{m}_2|_o$ nie mogą więc być identyczne). Jest to możliwe, jeśli żądana różnica nie przekracza stopnia niezdeterminowania, charakteryzującego interpretacje o-terminów (zarówno

$\mathfrak{m}_1|_o$ jak i $\mathfrak{m}_2|_o$ muszą należeć do M_o). Jest to schemat w przybliżeniu podobny do schematu tzw. relacji «korespondencji»[13]. Główna różnica, jak się wydaje, dotyczy wspólnej dziedziny, w której dwie rywalizujące ze sobą teorie okazują się aproksymacyjnie prawdziwe. Dziedzina ta była przez nas utożsamiana z całym uniwersum obu teorii, tzn. z uniwersum struktur w M_o i M_i. Zwykle jednak jest to dziedzina węższa. W przeciwieństwie do nowej teorii T_2, dotychczasowa teoria T_1 okazuje się fałszywa w swoim pierwotnym uniwersum. Dziedzina, w której obie teorie okazują się aproksymacyjnie prawdziwe, tworzy tylko pewien podzbiór całego uniwersum; nazwijmy go U. W konsekwencji porównujemy zamierzone interpretacje t_1 i t_2 tylko w tym podzbiorze. Aby to zrobić, musimy zdefiniować pojęcie, które będziemy oznaczać przez $M_o(U)$ ($M_i(U)$). $M_o(U)$ jest klasą tych podstruktur struktur w M_o, które mają uniwersum U. Interpretacje t_1 i t_2, które teraz porównujemy, są więc interpretacjami, które należą do struktur z $M_1(U)^*$ i $M_2(U)^*$. Ponieważ teorie związane relacją «korespondencji» są zwykle teoriami ilościowymi, załóżmy, że o-terminy i t-terminy teorii T_1 i T_2 odnoszą się do pewnych empirycznych wielkości. Aby dać prosty, schematyczny przykład załóżymy, że w zamierzonych strukturach w M_1^* i M_2^*, terminy t_1 i t_2 są zdefiniowane przez następujące równania, które tworzą część odpowiednich teorii:

T_1: $t_1(x) = f_1(o_1(x), ..., o_n(x))$,

T_2: $t_2(x) = f_2(o_1(x), ..., o_n(x))$.

O f_1 i f_2 zakłada się, że są różnymi funkcjami matematycznymi, tzn. że przynajmniej dla niektórych argumentów $k_1, ..., k_n$: $f_1(k_1,..., k_n) \neq f_2(k_1,..., k_n)$. Co więcej, o powyższej nierówności zakłada się, że zachodzi ona dla wszystkich tych argumentów, które reprezentują wartości o-funkcji w faktycznej strukturze dla L_o. Tak więc, dla każdego $\mathfrak{m}_o \in$ $\in M_o$ i każdego x z uniwersum $\mathfrak{m}_o$: $f_1(o_1^{\mathfrak{m}_o}(x), ..., o_n^{\mathfrak{m}_o}(x)) \neq f_2(o_1^{\mathfrak{m}_o}(x), ..., o_n^{\mathfrak{m}_o}(x))$. Otóż sądzimy, że istnieją $\mathfrak{m}_o$ i $\mathfrak{n}_o \in M_o$, takie że dla każdego $x \in U$ (gdzie U jest dziedziną, w której T_1 i T_2 są aproksymacyjnie prawdziwe): $f_1(o_1^{\mathfrak{m}_o}(x), ..., o_n^{\mathfrak{m}_o}(x)) = f_2(o_1^{\mathfrak{n}_o}(x), ..., o_n^{\mathfrak{n}_o}(x))$. Jest to oczywiście możliwe tylko wtedy, gdy przynajmniej dla niektórych o_i ($i = 1,...n$): $o_i^{\mathfrak{m}_o}(x) \neq o_i^{\mathfrak{n}_o}(x)$. Ta różnica w interpretacji o_i kompensuje różnicę między f_1 i f_2. Sądzimy, w konsekwencji, że istnieją $\mathfrak{m}_1 \in M_1(U)^*$ i $\mathfrak{m}_2 \in M_2(U)^*$ (gdzie $\mathfrak{m}_1|_o = \mathfrak{m}_o$, $\mathfrak{m}_2|_o = \mathfrak{n}_o$), takie że $t_1^{\mathfrak{m}_1} = t_2^{\mathfrak{m}_2}$. Pewne zamierzone interpretacje t_1 i t_2 okazują się być identyczne w tej poddziedzinie, w której obie teorie są aproksymacyjnie prawdziwe. Fakt ten, jak się wydaje, jest więc całkowicie zgodny z charakterystyką tych dwóch teorii przedstawioną wyżej.[14]

---

13) Por. np. charakterystykę «reguły korespondencji» w Krajewski (1977).

14) Naczęściej analizowany wypadek klasycznej i relatywistycznej mechaniki wydaje się podpadać pod ten schemat. W dużym uproszczeniu, zasady ruchu charakteryzowane przez te dwie teorie można przedstawić następująco:

(1) $f_1 = ma$; (2) $f_2 = \dfrac{ma}{\sqrt{1-\frac{v^2}{c^2}}}$

Przypadek rozważany przez nas może być uogólniony tak, aby obejmował również inne typy relacji «korespondencji». Jedno z tych uogólnień dotyczy faktycznej interpretacji podjęzyka L_o. Zakłada się, że jest ona identyczna w obu rozważanych teoriach. Wydaje się to jednak zbytnim ograniczeniem. Istnieją takie typy rozwoju teorii, w których interpretacja o-terminów również ulega pewnej zmianie: w nowej teorii T_2 są one interpretowane w bardziej precyzyjny sposób niż w dotychczasowej teorii T_1. Możemy zdać sprawę z tego faktu przez podział podjęzyka L_o na dwa języki L_{o1} i L_{o2} z różnymi klasami faktycznych struktur, takich mianowicie, że

$$M_{o2} \subset M_{o1},$$

i przez definicję klasy faktycznych struktur dla L_1 i L_2, M_1 i M_2, na podstawie odpowiednio M_{o1} i M_{o1}. Porównując zamierzone interpretacje t_1 i t_2, będziemy odwoływać się jedynie do tych zamierzonych struktur dla L_1, które są rozszerzeniami struktur z M_o. Jeśli ograniczy się je w ten sposób, to można pokazać, że pozostają one w tych samych relacjach, co poprzednio.

Inne rozszerzenie rozważanego dotąd przypadku polega na uwzględnianiu takich zmian teorii, w których nowa teoria T_2 ma «wyższy wymiar» niż dotychczasowa teoria T_1, gdyż zawiera terminy z większą ilością argumentów niż odpowiadające im terminy teorii T_1. Sytuację taką można rozwiązać w ten sposób, że porówna się, powiedzmy, k-argumentowy predykat t_1 teorii T_1 nie z odpowiadającym mu k+1-argumentowym predykatem t_2 teorii T_2, lecz z predykatem t_2' zdefiniowanym w następujący sposób:

$$t_2'(x_1, \ldots, x_k) \text{ zawsze i tylko, gdy } \bigvee x_{k+1}\, t_2(x_1, \ldots, x_k, x_{k+1}).$$

Porównanie ich może wtedy przebiegać w sposób zaproponowany wyżej.

Są to tylko przykłady modyfikacji, które mogą przybliżyć naszą analizę do sytuacji typowej dla rzeczywistych zmian teorii. Celem, który przyświecał nam przy analizie pewnych uproszczonych schematów takich sytuacji, było pokazanie możliwości istnienia pewnych wspólnych odniesień terminów teoretycznych, zdefiniowanych przez dwie rywalizujące teorie. Przemawia to, jak się wydaje, na korzyść pewnego rodzaju ciągłości pojęciowej w rozwoju nauki.

Tłumaczyła Anna Lissowska

Prawo klasyczne może być oczywiście zast[illegible]one przez prawo, które zawiera v w sposób nieistotny, i, tak zmodyfikowane, może być porównane z prawem relatywistycznym, zgodnie z przedstawionym wyżej schematem ogólnym. Sądzimy, że chociaż te dwa prawa się różnią, można dowieść, iż pewne zamierzone interpretacje f_1 i f_2 są identyczne w pewnej poddziedzinie U (składającej się z obiektów poruszających się ze względnie małymi prędkościami), ponieważ dla każdego obiektu z U, wartość v może być wyrównana przez różnicę w wartościach m i a, określonych przez pewne zamierzone interpretacje tych terminów.

BIBLIOGRAFIA

Carnap, R.: 1966, *Philosophical Foundations of Physics*, Basic Books, New York.

Krajewski, W.: 1977, *Correspondence Principle and Growth of Science*, Reidel, Dordrecht and Boston.

Przełęcki, M.: 1969, *The Logic of Empirical Theories*, Routledge and Kegan Paul, London.

Przełęcki, M.: 1976, „Fuzziness as multiplicity", *Erkenntnis*, 10, 371-380.

Przełęcki, M.: 1978, „Commensurable referents of incommensurable terms", *Acta Philosophica Fennica*, 30 (Issues 2-4: *The Logic and Epistemology of Scientific Change*), North-Holland, Amsterdam, s. 347-365.

Przełęcki, M.: 1978, „Some approaches to inexact measurement", *Poznań Studies in the Philosophy of the Sciences and the Humanities*, 4, 27-36.

Sneed, J.D.: 1971, *The Logical Structure of Mathematical Physics*, Reidel, Dordrecht.

Suppes, P.: 1969, *Studies in the Methodology and Foundations of Science*, Reidel, Dordrecht.

Williams, P.M.: 1973, „On the logical relations between expressions of different theories", *The British Journal for the Philosophy of Science*, 24, 357-367.

EMPIRYCZNOŚĆ

Filozofia Nauki
Rok I, 1993, Nr 2-3

O pojęciu genotypu

Podejmowane przez Woodgera próby zbudowania precyzyjnego i adekwatnego języka genetyki znalazły w jego książce *Biology and Language* ze wszech miar godną uwagi realizację. W zarysowanej tam strukturze pojęciowej ważne miejsce zajmuje pojęcie genotypu służące do wprowadzenia większości pojęć pozostałych. Podana przez Woodgera definicja genotypu interesująca jest nie tylko dla genetyków. Zasługuje na uwagę wszystkich, którym nie obce są problemy metodologii ogólnej. Genotyp uważany jest zwykle za typowe „pojęcie teoretyczne", pojawiające się na najwyższym szczeblu teorii genetycznej. Pojęcie to denotować ma genetyczną konstytucję organizmu, charakteryzowaną przez zespół genów — przedmiotów niedostępnych bezpośredniej obserwacji. Woodger traktuje genotyp jako „pojęcie elementarne", należące do najniższego szczebla systemu genetyki. Definicję genotypu formułuje za pomocą terminów odnoszących się do bezpośrednio obserwowalnych rzeczy, własności i stosunków. Nie odwołuje się w niej w ogóle do dyskusyjnego pojęcia genu; na odwrót — to ostatnie zdefiniowane zostaje przy pomocy pojęcie genotypu. Woodgerowskie pojęcie genotypu — pojęcie teoretyczne zdefiniowane *explicite* za pomocą pojęć spostrzeżeniowych — zasługuje na dokładną analizę. Analiza ta przyczynić się może do wyjaśnienia pewnych zagadnień dotyczących stosunku teorii do doświadczenia. Nie podejmuję tutaj wszechstronnej analizy podanej przez Woodgera definicji. Chciałbym jedynie zwrócić uwagę na pewną charakterystyczną cechę, która wydaje się przysługiwać wszelkim definicjom tego typu — definicjom pojęć teoretycznych sformułowanym w języku spostrzeżeniowym.

Do najniższego szczebla teorii genetycznej, zwanego w terminologii Woodgera szczeblem zerowym, należą uogólnienia obserwacji dotyczących organizmów rodzicielskich i potomnych oraz środowisk, w jakich się te organizmy rozwijają. Woodger nie wprowadza dla oznaczenia tych pojęć terminów odrębnych. Używa jednego terminu pierwotnego, wyjaśnianego za pomocą następującej reguły semantycznej:

'$F_{X,Y,Z}(W_1, W_2)$' denotuje klasę x–ów takich, iż dla pewnego u i pewnego v : u i v są rodzicami x–a, u należy do W_1, v należy do W_2, u rozwija się w środowisku należącym do X, v — w środowisku należącym do Y, x — w środowisku należącym do Z.

Wyjaśnienie to wyrazić można w postaci symbolicznej:

$$F_{X,Y,Z}(W_1, W_2) = (\hat{x})\,[(\exists u)(\exists v)\,(u \neq v \cdot u \,\varepsilon\, W_1 \cdot u\, Ps\, x \cdot v \,\varepsilon\, W_2 \cdot v\, Ps\, x \cdot en\,(u) \,\varepsilon\, X \cdot \cdot\, en(v) \,\varepsilon\, Y \cdot en\,(x) \,\varepsilon\, Z)],$$

gdzie 'Ps' denotuje relację bycia organizmem rodzicielskim (w procesie rozmnażania płciowego)[1], a '$en(x)$' — środowisko x–a[2], przy czym oba te terminy pojmowane są jako wyrażenia szczebla zerowego. Gdybyśmy wyrażenia te obrali za terminy pierwotne, powyższe wyjaśnienie terminu 'F' mogłoby pełnić rolę jego definicji. Niezależnie jednak od tego, czy termin ten występuje w charakterze terminu pierwotnego czy zdefiniowanego, jest to również wyrażenie szczebla zerowego, odnoszące się do obserwowalnych rzeczy i stosunków. W większości zastosowań owe trzy klasy środowisk: X, Y, Z nie różnią się między sobą. Dlatego też wprowadza się prostszą formę tego wyrażenia:

$$F_X(W_1, W_2) = F_{X,X,X}(W_1, W_2).$$

Termin 'F' denotuje tzw. pierwsze pokolenie filialne. Wprowadzenie notacji dla drugiego pokolenia nie nastręcza trudności:

$$F^2_X(W_1, W_2) = F_X[F_X(W_1, W_2), F_X(W_1, W_2)].$$

Następnym krokiem jest klasyfikacja środowisk oraz organizmów rodzicielskich i potomnych. Klasy środowisk charakteryzuje się pod względem składu gleby, wody czy powietrza, pod względem ilości i składu pożywienia itp. Klasy organizmów wyróżnia się w genetyce na dwa sposoby. Przede wszystkim wyróżniać możemy klasy organizmów na podstawie obserwacji stwierdzających, do jakiego gatunku te organizmy należą i czym się różnią od innych organizmów tego samego gatunku. Oto przykłady takich charakterystyk:

x jest grochem o żółtych liścieniach,

x jest grochem o zielonych liścieniach.

Klasy wyróżnione w ten sposób nazywamy fenotypami. Ale klasyfikacja organizmów rodzicielskich i potomnych na fenotypy nie wystarcza dla zagadnień genetyki. Zachodzi potrzeba dalszej klasyfikacji organizmów należących do poszczególnych fenotypów na klasy zwane genotypami. Klasy te charakteryzuje się zwykle przez odwołanie się do pewnych niespostrzegalnych przedmiotów, postulowanych przez hipotezy najwyższego szczebla teorii genetycznej. Woodger charakteryzuje genotypy bez uciekania się do owych, zagadkowych nieco pojęć. Definicje, które podaje, korzystają wyłącznie z wprowadzonych uprzednio terminów spostrzeżeniowych. Sens terminu „genotyp",

1) Porównaj *Biology and Language*, str. 214.
2) Porównaj *Biology and Language*, str. 208.

jaki te definicje starają się uchwycić, charakteryzuje następujące określenie Haldane'a: „klasa (organizmów), którą można odróżnić od innych za pomocą eksperymentów hodowlanych, zwie się genotypem".

Opiszmy najprostszy tego rodzaju eksperyment — doświadczenia Mendla z odmianami grochu. Potrzebne nam będą w tym celu dwa rodzaje twierdzeń:

(1) $F_X(Y, Z) \subset W$

(2) $F_X(Y, Z)\ \varepsilon\ p_1 W + p_2 W_2 + \ldots + p_n W_n$.

(1) stwierdza, iż każdy element klasy $F_X(Y, Z)$ należy do klasy W, (2) — iż p_i-ta część elementów klasy $F_X(Y, Z)$ należy do klasy W_i (dla każdego i od 1 do n). Przez 'Y' oznaczymy odmianę grochu o żółtych liścieniach, przez 'G' — odmianę grochu o zielonych liścieniach, przez 'A' i 'B' te podzbiory klas Y i G, którymi Mendel posługiwał się w swych doświadczeniach, a przez 'E' — typ środowiska, istniejącego w ogrodzie Mendla. Wyniki doświadczeń Mendla zapisać można w sposób następujący:

I. (1) $A \subset Y$
(2) $F_E(A, A) \subset Y$
(3) $F^2{}_E(A, A) \subset Y$

II. (1) $B \subset G$
(2) $F_E(B, B) \subset G$
(3) $F^2{}_E(B, B) \subset G$

III. (1) $F_E(A, B) \subset Y$
(2) $F^2{}_E(A, B)\ \varepsilon\ \frac{3}{4}Y + \frac{1}{4}G$
(3) $F_E[A, F_E(A, B)] \subset Y$
(4) $F_E[B, F_E(A, B)]\ \varepsilon\ \frac{1}{2}Y + \frac{1}{2}G$.

Te trzy zbiory organizmów: A, B i $F_E(A, B)$ zachowują się różnie pod względem genetycznym i, co za tym idzie, należą do różnych genotypów: do homozygotycznego genotypu odpowiadającego fenotypowi Y — $H(Y)$, do homozygotycznego genotypu odpowiadającego fenotypowi G — $H(G)$, oraz do heterozygotycznego genotypu odpowiadającego fenotypom Y i G — $Ht\ (Y, G)$. $H(Y)$ zostaje zdefiniowany jako klasa organizmów, które zachowują się pod względem genetycznym tak, jak elementy zbioru A, $H\ (G)$ — jako klasa organizmów, które zachowują się pod względem genetycznym tak, jak elementy zbioru B, a $Ht\ (Y,G)$ — jako klasa organizmów, które zachowują się pod względem genetycznym tak, jak elementy zbioru $F_E(A, B)$. Dokładne definicje tych pojęć są formalnie nieco skomplikowane. Ograniczę się wobec tego do przedstawienia tylko jednej z nich — definicji $H\ (Y)$ — i to w sposób trochę uproszczony.

$H\ (Y)$ zawiera wszystkie klasy Z spełniające poniższe warunki:

(1) $Z \subset Y$

(2) $F_{E(Y)}(Z, Z) \subset Y$

(3) $F^2{}_{E(Y)}(Z, Z) \subset Y$

(4) $F^2{}_{E(Y)}(Z, Z) \neq \varnothing$

Warunki (1)–(3) odpowiadają wynikom doświadczeń z elementami zbioru A. Warunek (4), który implikuje:

$$F_{E(Y)}(Z, Z) \neq \emptyset \text{ i } Z \neq \emptyset,$$

stanowić ma gwarancję, że warunki (1)–(3) nie są spełnione w sposób trywialny. $E(Y)$ jest pewną maksymalną klasą środowisk, zawierającą, mówiąc swobodnie, wszelkie klasy środowisk podobne do E, tj. klasy środowisk, w których istnieją organizmy zachowujące się pod względem genetycznym tak, jak elementy zbioru A. Jeśli wprowadzimy '$K(Z)$' jako skrótowy zapis warunków (1)–(4), definicję $H(Y)$ możemy sformułować jak następuje:

$H(Y)$ jest sumą wszystkich klas Z takich, iż $K(Z)$.

Definicja ta równoważna jest formule:

$$H(Y) = (\hat{x})(\exists Z)[K(Z) \cdot x \varepsilon Z]$$

Analogicznie zdefiniować możemy $H(G)$ i $Ht(Y, G)$.

Definicja $H(Y)$ wydaje się intuicyjnie trafna. Organizm jakiś należy do homozygotycznego genotypu odpowiadającego fenotypowi Y wtedy i tylko wtedy, gdy sam należy do fenotypu Y i gdy jego potomstwo w pierwszym i drugim pokoleniu również należy do fenotypu Y. Bliższa analiza pokazuje jednak, iż definicja $H(Y)$ prowadzi do pewnych niepożądanych konsekwencji.

Załóżmy, iż x_1 jest grochem o żółtych liścieniach nie posiadającym potomstwa:

$$x_1 \varepsilon Y \cdot \sim (\exists y)(x_1 Ps\, y).$$

Można łatwo okazać, iż x_1 należeć będzie do $H(Y)$, jeżeli tylko istnieje klasa Z taka, iż $K(Z)$. Niech klasą tą będzie $Z_0 : K(Z_0)$. Ale wówczas klasa $Z_0 \cup \{x_1\}$, powstała przez dołączenie elementu x_1, musi również spełniać warunek $K : K(Z_0 \cup \{x_1\})$. Założenie $K(Z_0)$ sprowadza się do koniunkcji następujących warunków:

(1) $Z_0 \subset Y$

(2) $F_{E(Y)}(Z_0, Z_0) \subset Y$

(3) $F^2_{E(Y)}(Z_0, Z_0) \subset Y$

(4) $F^2_{E(Y)}(Z_0, Z_0) \neq \emptyset$

Okażemy, że warunki te mają walor również dla $Z_0 \cup \{x_1\}$. Warunek:

(1) $Z_0 \cup \{x_1\} \subset Y$

jest spełniony, gdyż zgodnie z naszym założeniem, $x_1 \varepsilon Y$. Po to, aby okazać, iż spełniony jest warunek:

(2) $F_{E(Y)}(Z_0 \cup \{x_1\}, Z_0 \cup \{x_1\}) \subset Y$

wystarczy udowodnić, iż:

$$F_{E(Y)}(Z_0 \cup \{x_1\}, Z_0 \cup \{x_1\}) = F_{E(Y)}(Z_0, Z_0)$$

To zaś jest oczywiste, skoro x_1 nie ma potomstwa. Dołączenie go do klasy Z_0 nie zmienia w niczym klasy złożonej z potomstwa organizmów należących do klasy Z_0. Można to okazać w sposób formalny. Klasa:

$$F_{E(Y)}(Z_0 \cup \{x_1\}, Z_0 \cup \{x_1\})$$

jest, zgodnie z definicją F, identyczna z klasą:

$$(\hat{y})\ [(\exists u)\ (\exists v)\ (u \neq v \cdot u \, \varepsilon \, Z_0 \cup \{x_1\} \cdot u \, Ps \, y \cdot v \, \varepsilon \, Z_0 \cup \{x_1\} \cdot v \, Ps \, y \cdot en\,(u) \, \varepsilon \, E(Y) \cdot$$
$$\cdot \, en\,(v) \, \varepsilon \, E\,(Y) \cdot en\,(y) \, \varepsilon \, E\,(Y))]$$

Ponieważ $u \, \varepsilon \, Z_0 \cup \{x_1\} \equiv u \, \varepsilon \, Z_0 \vee u = x_1$, sformułowanie powyższe przedstawić można w następującej rozwiniętej postaci:

$$(\hat{y})\ [(\exists u)\ (\exists v)\ (u \neq v \cdot u \, \varepsilon \, Z_0 \cdot u \, Ps \, y \cdot v \, \varepsilon \, Z_0 \cdot v \, Ps \, y \cdot en\,(u) \, \varepsilon \, E\,(Y) \cdot en\,(v) \, \varepsilon \, E\,(Y) \cdot$$
$$\cdot \, en\,(y) \, \varepsilon \, E\,(Y) \vee u \neq v \cdot u \, \varepsilon \, Z_0 \cdot u \, Ps \, y \cdot v = x_1 \cdot v \, Ps \, y \cdot en\,(u) \, \varepsilon \, E\,(Y) \cdot en\,(v) \, \varepsilon \, E\,(Y) \cdot$$
$$\cdot \, en\,(y) \, \varepsilon \, E\,(Y) \vee u \neq v \cdot u = x_1 \cdot u \, Ps \ y \cdot v \, \varepsilon \, Z_0 \cdot v \, Ps \, y \cdot en\,(u) \, \varepsilon \, E\,(Y) \cdot en\,(v) \, \varepsilon \, E\,(Y) \cdot$$
$$\cdot \, en\,(y) \, \varepsilon \, E\,(Y) \vee u \neq v \cdot u = x_1 \cdot u \, Ps \, y \cdot v = x_1 \cdot v \, Ps \, y \cdot en\,(u) \, \varepsilon \, E\,(Y) \cdot en\,(v) \, \varepsilon \, E\,(Y) \cdot$$
$$\cdot \, en\,(y) \, \varepsilon \, E\,(Y))]$$

Otóż składnik ostatni powyższej alternatywy jest fałszywy dla każdego y ze względów czysto formalnych:

$$\sim (\exists u)\ (\exists v)\ (u \neq v \cdot u = x_1 \cdot v = x_1)\,.$$

Natomiast składniki: drugi i trzeci są fałszywe dla każdego y dlatego, że zgodnie z naszymi założeniami dotyczącymi x_1:

$$\sim (\exists \mathrm{y})\ (\exists \mathrm{u})\ (u = x_1 \cdot u \, Ps \, y)\,.$$

Tak więc cała alternatywa jest równoważna składnikowi pierwszemu, a rozważana klasa jest identyczna z klasą:

$$(\hat{y})\ [(\exists u)\ (\exists v)\ (u \neq v \cdot u \, \varepsilon \, Z_0 \cdot u \, Ps \, y \cdot v \, \varepsilon \, Z_0 \cdot v \, Ps \, y \cdot en\,(u) \, \varepsilon \, E\,(Y) \cdot en\,(v) \, \varepsilon \, E\,(Y) \cdot$$
$$\cdot \, en\,(y) \, \varepsilon \, E\,(Y))],$$

czyli z klasą: $F_{E(Y)}\,(Z_0, Z_0)$. Identyczność pokoleń pierwszych:

$$F_{E(Y)}\,(Z_0 \cup \{x_1\}, Z_0 \cup \{x_1\}) = F_{E(Y)}\,(Z_0, Z_0)$$

pociąga za sobą identyczność pokoleń drugich:

$$F^2_{E(Y)}\,(Z_0 \cup \{x_1\}, Z_0 \cup \{x_1\}) = F^2_{E(Y)}\,(Z_0, Z_0)$$

ponieważ:

$$F^2_{E(Y)}\,(Z_0 \cup \{x_1\}, Z_0 \cup \{x_1\}) =$$
$$F_{E(Y)}\,(F_{E(Y)}\,(Z_0 \cup \{x_1\}, Z_0 \cup \{x_1\}), F_{E(Y)}\,(Z_0 \cup \{x_1\}, Z_0 \cup \{x_1\}))\,.$$

A zatem warunki:

(3) $F^2_{E(Y)}\,(Z_0 \cup \{x_1\}, Z_0 \cup \{x_1\}) \subset Y$

(4) $F^2_{E(Y)}\,(Z_0 \cup \{x_1\}, Z_0 \cup \{x_1\}) \neq \emptyset$

spełnione muszą być również. Wynika stąd, iż $K\,(Z_0 \cup \{x_1\})$. Ale jeśli $K\,(Z_0 \cup \{x_1\})$, to $(\exists Z)\ (K(Z) \cdot x_1 \, \varepsilon \, Z)$, a stąd, zgodnie z definicją $H(Y)$, $x_1 \, \varepsilon \, H(Y)$.

Ta paradoksalna konsekwencja dotyczy nie tylko tych elementów Y, które w ogóle nie posiadają potomstwa. Jeśli pewien organizm o fenotypie Y posiada potomstwo, ale potomstwo nie pochodzące od żadnego z elementów klasy Z_0, należy on do $H(Y)$ z analogicznych względów. To samo dotyczy tych elementów Y, które rozwinęły się w środowisku nie należącym do $E\,(Y)$ lub których potomstwo rozwinęło się w takim środowisku. Mówiąc ogólnie, jeśli $x_1 \, \varepsilon \, Y$ i $\sim (\exists y)\ (\exists v)\ (x_1 \neq v \cdot x_1 \, Ps \, y \cdot v \, \varepsilon \, Z_0 \cdot v \, Ps \, y \cdot$ $\cdot \, en\,(x_1) \, \varepsilon \, E\,(Y) \cdot en\,(v) \, \varepsilon \, E\,(Y) \cdot en\,(y) \, \varepsilon \, E\,(Y))$, to $x_1 \, \varepsilon \, H\,(Y)$.

Innymi słowy, jeśli $x_1 \, \varepsilon \, Y$ i $F_{E(Y)}\,(\{x_1\}, Z_0) = \emptyset$, to $x_1 \, \varepsilon \, H\,(Y)$. Rzecz w tym, że dla x_1 warunki (2)–(4) naszej definicji $H\,(Y)$ spełnione są w sposób trywialny. Jeśli $F^2_{E(Y)}\,(\{x_1\}, Z_0) = \emptyset$, w takiej samej sytuacji są w stosunku do x_1 warunki (3)–(4).

Jest to bez wątpienia konsekwencja niepożądana. Tak zdefiniowana klasa $H(Y)$ nie jest najwyraźniej tym, czym być miała. Jest stanowczo za obszerna, gdyż — prócz «właściwych» elementów homozygotycznego genotypu odpowiadającego fenotypowi Y — obejmuje te wszystkie elementy Y, które nie posiadają potomstwa z żadnym elementem jakiejś klasy spełniającej warunek K. Podobne konsekwencje mają miejsce w przypadku skonstruowanej wedle tego samego wzoru definicji heterozygotycznego genotypu odpowiadającego fenotypom Y i G. W rezultacie klasy $H(Y)$ i $Ht(Y, G)$ okazują się nierozłączne. Pewne elementy mają wspólne. Tak np. wszystkie bezpotomne organizmy o fenotypie Y należą zarazem do $H(Y)$ i do $Ht(Y, G)$[3].

Jeśli chcemy uniknąć paradoksalnych konsekwencji naszej definicji, musimy wzmocnić warunek K tak, aby wykluczyć z klasy $H(Y)$ te wszystkie organizmy o fenotypie Y, które należą do niej wyłącznie wskutek braku potomstwa. Możemy to uczynić przez dołączenie do warunków (1)–(4) następującego postulatu:

(5) $(x)(x \varepsilon Z \supset F^2_{E(Y)}(\{x\}, Z) \neq \emptyset)$.

Gwarantuje on, iż klasa Z obejmuje tylko takie organizmy, które posiadają potomstwo (w pierwszym i drugim pokoleniu) z pewnym elementem klasy Z. Wprowadźmy '$K'(Z)$' jako skrótowy zapis warunków (1)–(5). Twierdzenie: $(\exists Z)(K'(Z) \cdot x_1 \varepsilon Z)$ implikuje teraz, że x_1 należy do Y, że posiada potomstwo (w pierwszym i drugim pokoleniu), i że potomstwo to również należy do Y. Jeśli któryś z tych faktów nie ma miejsca, twierdzenie to nie może być prawdziwe. A więc jeśli x_1 nie ma potomstwa, nie może należeć do żadnej klasy spełniającej warunek K'. Definicja:

$$H(Y) = (x)(\exists Z)(K'(Z) \cdot x \varepsilon Z)$$

wyklucza w ten sposób z klasy $H(Y)$ te wszystkie «niewłaściwe» elementy, które należały do niej zgodnie z definicją poprzednią. Ale i ta definicja nie jest zadowalająca. Wszystkie bezpotomne organizmy o fenotypie Y zmuszeni jesteśmy teraz zaliczyć do $\overline{H}(Y)$. Mówiąc ściślej, wszystkie organizmy o fenotypie Y, które nie posiadają potomstwa z żadnym elementem jakiejkolwiek klasy spełniającej warunek K', należeć będą do $\overline{H}(Y)$. Jest to konsekwencja równie nie do przyjęcia, jak poprzednia.

Wskazane wadliwości wydają się nieuleczalne. Każda zupełna definicja $H(Y)$ korzystająca wyłącznie z naszego ograniczonego zasobu terminów pierwotnych musi prowadzić do konkluzji zaliczającej wszystkie bezpotomne organizmy o fenotypie Y bądź do klasy $H(Y)$ bądź do klasy $\overline{H}(Y)$. Nie mamy żadnych podstaw do czynienia pomiędzy nimi jakichkolwiek różnic, ponieważ wszelkie nasze decyzje oparte są na obserwacjach dotyczących fenotypu danego organizmu i jego potomstwa. Wobec faktu, że wszystkie omawiane organizmy należą do tego samego fenotypu i żadnego potomstwa nie posiadają, wszystkie znajdą się w tej samej klasie: $H(Y)$ lub $\overline{H}(Y)$. Obie

3) W pracy „Zastosowanie pojęć logiki matematycznej do wyjaśnienia niektórych pojęć przyrodoznawstwa" (*Studia Logica* 4, (1956)) M. Kokoszyńska, T Kubiński i J. Słupecki rozważają definicję genotypu zbliżoną do definicji Woodgerowskiej i okazują, że definicja ta pociąga konsekwencje zbliżone do konsekwencji omawianych.

decyzje są równie nieuzasadnione i arbitralne. Co więcej, decyzje te mogą popaść w konflikt z wynikami przyszłych doświadczeń. Genetyka rozszerza stopniowo swoją doświadczalną bazę, włączając w nią np. pewne obserwacje cytologiczne. Prowadzi to do sformułowania nowych kryteriów stosowalności dla $H(Y)$, odwołujących się do owych obserwacji. Kryteria te pozwalają na zaliczenie pewnych organizmów bezpotomnych do $H(Y)$, innych — do $\overline{H}(Y)$. Ale decyzje takie kolidują z definicjami zaliczającymi arbitralnie wszystkie te organizmy do $H(Y)$ lub $\overline{H}(Y)$.

Genetyka w swojej «elementarnej», Mendlowskiej postaci nie pozwala nam powiedzieć niczego o genotypie organizmów bezpotomnych. Dlatego też i definicja genotypu formułowana na gruncie tej teorii musi kwestię tę pozostawiać otwartą. Nie może to być zatem definicja zupełna. Pojęcie genotypu może zostać zdefiniowane adekwatnie tylko przy pomocy jakiejś bardziej «liberalnej» procedury. Procedurę taką reprezentuje tzw. definicja cząstkowa lub warunkowa (inaczej redukcja). Ten rodzaj definicji został wprowadzony przez Carnapa[4], a potem uogólniony przez innych autorów[5].

Najprostszą postać definicji cząstkowej terminu Q przedstawiają wypowiedzi następujące:

(1) $Px \supset Qx$

(2) $Rx \supset {\sim} Qx$

Główna różnica pomiędzy definicją cząstkową a definicją zupełną polega na tym, iż warunki definicyjne: P i R ani się logicznie nie dopełniają, ani nie wykluczają, gdy tymczasem w przypadku definicji zupełnej warunki te sprowadzają się do: P i $\sim P$. Ponieważ warunki P i R nie wyczerpują wszystkich możliwości, mogą istnieć przedmioty, które nie są ani P, ani R. W stosunku do takich przedmiotów znaczenie terminu Q pozostaje nieustalone. Nie rozporządzamy w tych przypadkach żadnym kryterium stosowalności tego terminu. Q jest pojęciem „otwartym". Dalsze ustalanie jego znaczenia wymaga kryteriów dodatkowych.

Przeprowadzona analiza świadczy o tym, że genotyp jest pojęciem „otwartym", które w pewnych przypadkach pozostać musi niezdeterminowane. Toteż — jako pojęcie szczebla zerowego — genotyp może być zdefiniowany jedynie cząstkowo. Rolę jego definicji cząstkowej opartej na przyjętym słowniku terminów pierwotnych pełnić mogą następujące wypowiedzi:

(1) $(\exists Z)\,[K'(Z) \cdot x \varepsilon Z] \supset x \varepsilon H(Y)$

(2) $(\exists Z)\,[K'(Z) \cdot ({\sim} x \varepsilon Y \vee {\sim} F_{E(Y)}(\{x\}, Z) \subset Y \vee {\sim} F^2{}_{E(Y)}(\{x\}, Z) \subset Y)] \supset {\sim} x \varepsilon$
$\varepsilon H(Y)$.

Wypowiedź (1) formułuje warunek wystarczający należenia do $H(Y)$, wypowiedź (2) — warunek konieczny. Starałem się poprzednio okazać, iż (1) jest sformułowaniem adekwatnym. (2) wyłącza z klasy $H(Y)$ te organizmy, które nie należą do Y, lub które

4) „Testability and Meaning", *Philosophy of Science* 3–4, (1936–1937).

5) Na przykład H. Mehlberg, „Positivisme et Science", *Studia Philosophica* 3, (1948); H. V. Stopes-Roe, „Some Considerations Concerning «Interpretative Systems»", *Philosophy of Science* 25, (1958).

posiadają z elementami jakiejś klasy spełniającej warunek K' potomstwo nie należące do Y. I to zatem wydaje się sformułowaniem trafnym. Warunki wyrażone przez wypowiedzi (1) i (2) logicznie się nie dopełniają. Mogą istnieć organizmy, które nie spełniają żadnego z nich. Właśnie rozważane przez nas uprzednio organizmy do takich należą. Jeśli jakiś organizm nie posiada potomstwa z elementami żadnej klasy spełniającej warunek K', organizm ten nie spełnia poprzednika wypowiedzi (1). A jeśli przy tym jest to organizm o fenotypie Y, nie spełnia on również poprzednika wypowiedzi (2). (Należy zwrócić uwagę na fakt, iż twierdzenie: $\sim F_{E(Y)}(\{x\}, Z) \subset Y$ implikuje: $F_{E(Y)}(\{x\}, Z) \neq \emptyset$). W stosunku do takiego organizmu nie dysponujemy żadnym kryterium stosowalności terminu $H(Y)$. Nie mamy prawa orzec ani tego terminu, ani jego negacji. A o to właśnie chodziło.

Gdyby żadna klasa nie spełniała warunku K', znaczenie $H(Y)$ byłoby całkowicie nieustalone. Możemy jednak przyjąć (jak to uczynił Woodger), iż w tej sytuacji $H(Y) = \emptyset$, dodając dodatkowe zastrzeżenie do warunku (2):

(2′) $\sim (\exists Z)\, K'(Z) \vee (\exists Z)\, [K'(Z) \cdot (\sim x \varepsilon Y \vee \sim F_{E(Y)}(\{x\}, Z) \subset Y \vee \sim F^2_{E(Y)}(\{x\}, Z) \subset Y)] \supset \sim x \varepsilon H(Y)$.

Warunki, o których mowa w wypowiedziach (1) i (2) logicznie się nie wykluczają. Ale wypowiedzi te implikują konsekwencję stwierdzającą ich wzajemne wykluczanie. Konsekwencja ta musi mieć zatem charakter twierdzenia rzeczowego. Doświadczenie decyduje, czy przedmiot, który spełnia poprzednik wypowiedzi (1), spełniać może również poprzednik wypowiedzi (2). Doświadczenie zdaje się wykluczać taką możliwość i tym samym potwierdzać empiryczną konsekwencję naszej definicji. Ale świadectwo doświadczenia nigdy nie jest ostateczne. Chcąc zapewnić logiczną rozłączność warunków (1) i (2), musimy jeden z nich, np. warunek (2), poddać pewnej modyfikacji:

(2″) $\sim (\exists Z)\, [K'(Z) \cdot x \varepsilon Z] \cdot (\exists Z)\, [K'(Z) \cdot (\sim x \varepsilon Y \vee \sim F_{E(Y)}(\{x\}, Z) \subset Y \vee \sim F^2_{E(Y)}(\{x\}, Z) \subset Y)] \supset \sim x \varepsilon H(Y)$.

Podana przez nas cząstkowa definicja $H(Y)$ wydaje się wyposażać ten termin w zamierzony sens. Unika ona wadliwości przytaczanej uprzednio definicji zupełnej, a jednocześnie opiera się na tym samym, „elementarnym" słowniku. Jej adekwatność okupiona została jednak niewygodną, skomplikowaną formą. Wydaje się, iż można wprowadzić pojęcie genotypu za pośrednictwem prostszej definicji cząstkowej, sformułowanej w tym samym, Woodgerowskim języku. Ponieważ pojęcie to różni się pod pewnym względem od $H(Y)$, oznaczymy je przez '$H^*(Y)$'. Jego definicja ma postać następującą:

(1) $\{x, y\} \subset Y \cdot F_{E(Y)}\{x, y\} \subset Y \cdot F^2_{E(Y)}\{x, y\} \subset Y \cdot F^2_{E(Y)}\{x, y\} \neq \emptyset \supset \{x, y\} \varepsilon H^*(Y)$

(2) $\sim \{x, y\} \subset Y \vee \sim F_{E(Y)}\{x, y\} \subset Y \vee \sim F^2_{E(Y)}\{x, y\} \subset Y \supset \sim \{x, y\} \varepsilon H^*(Y)$,

gdzie $F_{E(Y)}\{x, y\}$ jest po prostu skrótem: $F_{E(Y)}(\{x\}, \{y\})$. Różnica między $H(Y)$ a $H^*(Y)$ leży w tym, iż $H^*(Y)$ nie jest klasą indywiduów, lecz klasą (nieuporządkowanych) par indywiduów. Definicja $H^*(Y)$ pozwala nam rozstrzygać, czy jakaś para organizmów należy do danego genotypu, nie pozwala natomiast na przypisywanie genotypu poszczególnemu organizmowi. Jeśli pominąć tę ważną skądinąd różnicę, znaczenie $H^*(Y)$ uważać można za takie samo, jak znaczenie $H(Y)$. Warunki wyszczególnione w definicji $H^*(Y)$ są analogiczne do warunków z definicji $H(Y)$ i analogiczne pociągają konsekwencje. $H^*(Y)$ jest również pojęciem «otwartym». Para bezpotomnych organizmów o fenotypie Y nie spełnia ani poprzednika wypowiedzi (1), ani poprzednika wypowiedzi (2). A zatem nie mamy prawa orzec o niej ani $\overline{H}^*(Y)$ ani $H^*(Y)$. Fakt ten wydaje się być zgodny z tym, co zostało powiedziane o sytuacji w genetyce klasycznej. Genetyka oparta wyłącznie na klasycznych eksperymentach genetycznych, tj. na eksperymentach hodowlanych, nie może powiedzieć nam niczego o genotypie takich organizmów. Dlatego też pojęcie genotypu definiowalne jest jedynie cząstkowo.

Tutaj jednak wypada zrobić pewne zastrzeżenie. Są sytuacje, w których określenie genotypu organizmu bezpotomnego możliwe jest również na szczeblu elementarnym. Tak jest wtedy, gdy odpowiednie twierdzenie wynika z uznanych twierdzeń dotyczących genotypu innych organizmów. Przypuśćmy, iż x_1 jest organizmem o fenotypie Y, który sam potomstwa nie posiada, ale którego rodzice: y_1 i y_2 mają potomstwo w pierwszym i drugim pokoleniu. Jesteśmy wówczas w stanie określić genotyp obojga rodziców, a — w pewnych przypadkach — również genotyp x_1. Jeśli y_1 i y_2 należą do $H(Y)$, taki sam musi być genotyp x_1. Jeśli jedno z rodziców należy do $H(Y)$ a drugie do $H(G)$, x_1 musi należeć do $Ht(Y, G)$. Tak samo będzie wtedy, gdy genotypami rodziców są odpowiednio $Ht(Y, G)$ i $H(G)$. Ale gdy oboje rodzice należą do $Ht(Y, G)$, lub jedno do $Ht(Y, G)$, a drugie do $H(Y)$, genotypem x_1 może być równie dobrze $H(Y)$, jak i $Ht(Y, G)$. Wyłącznie zbadanie potomstwa x_1 mogłoby rozstrzygnąć tę kwestię. Opisane przez nas sytuacje są całkowicie możliwe. Proponowane definicje $H(Y)$ czy $H^*(Y)$ nie implikują niczego, co by taką możliwość miało wyłączać.

Przejście od definicji zupełnej do cząstkowej pozwala, jak widzieliśmy, uniknąć pewnych niepożądanych konsekwencji pierwotnej definicji $H(Y)$. Wobec tej definicji wysunąć można jednak zastrzeżenia inne, które nie dadzą się usunąć w ten sposób. Istnieją wadliwości, którymi obarczona jest zarówno zupełna definicja $H(Y)$ jak i cząstkowe definicje tego pojęcia. Załóżmy, że x_1 należący do genotypu $Ht(Y, G)$ posiada potomstwo składające się w pierwszym i drugim pokoleniu wyłącznie z organizmów o fenotypie Y. Jest to całkowicie prawdopodobny zbieg okoliczności, zwłaszcza gdy liczba tego potomstwa jest niewielka. Ale wówczas wszystkie dotychczasowe definicje $H(Y)$ — zarówno zupełne, jak i cząstkowe — prowadzić muszą do konkluzji, iż x_1 należy do $H(Y)$! Jest to jaskrawo niezgodne z tym znaczeniem, jakie terminowi $H(Y)$ chcieliśmy nadać. Rzecz w tym, że zależność pomiędzy genotypem danego organizmu a fenotypem jego potomstwa jest zależnością probabilistyczną. Toteż adekwatna definicja $H(Y)$ sformułowana w języku spostrzeżeniowym musi być również definicją «probabilistyczną». Takie probabilistyczne definicje cząstkowe wprowadzone zostały przez pewnych autorów[6], jednakże wymagają one jeszcze bliższego opracowania. Sprawą tą nie będę się zajmował w pracy obecnej, poświęconej jednemu tylko aspektowi analizowanego pojęcia: jego cząstkowej definiowalności.

Charakter ten przysługuje nie tylko pojęciu genotypu, lecz i szeregowi innych pojęć należących do skonstruowanego przez Woodgera aparatu pojęciowego genetyki. Przede wszystkim terminy zdefiniowane przy pomocy $H(Y)$ dziedziczą «otwarty» charakter tego ostatniego. Ale i niezależnie od tego pewne terminy uzyskują tę właściwość. Ich definicje można poddać analizie takiej samej, jak definicję $H(Y)$. I tak samo, jak tam, można usunąć ich wadliwości przechodząc do definicji cząstkowych. Jedno z tych pojęć zasługuje jednakże na szczególną uwagę. Definicja terminu $Gm(P)$ prowadzi do pewnych paradoksalnych konsekwencji z powodów niezmiernie podobnych do tych, któreśmy rozważali dotychczas. Woodger sugeruje jednak w tym przypadku pewne rozwiązanie, i to rozwiązanie polegające nie na rezygnacji z definicju zupełnej, lecz na jej nieznacznej modyfikacji. Ta interesująca sugestia wymaga bliższego rozpatrzenia.

'$Gm(P)$' ma denotować klasę gamet takich, iż powstałe z ich połączenia zygoty rozwijają się w organizmy o genotypie $H(P)$. Dla sformułowania definicji $Gm(P)$ trzeba dwóch dodatkowych wyrażeń:

(1) '$dlz(x, y, z)$', które głosi, iż zygota x rozwija się w środowisku y w organizm z, oraz

(2) '$U(\alpha, \beta)$', które denotuje klasę wszystkich zygot powstałych z połączenia gamet należących do zbioru α z gametami należącymi do zbioru β.

Przyjmijmy '$L(\alpha)$' jako skrótowy zapis następujących warunków:

(1) $\alpha \neq \emptyset$

(2) $(x)(y)(z)[dlz(x, y, z) \cdot x \varepsilon U(\alpha, \alpha) \cdot y \varepsilon E(P) \supset z \varepsilon H(P)]$.

$Gm(P)$ zdefiniowane zostaje następująco:

$$Gm(P) = (\hat{u})(\exists\alpha)(L(\alpha) \cdot u \varepsilon \alpha).$$

W stosunku do tej definicji Woodger sam wysuwa pewne obiekcje. „Można zawsze dołączyć do zbioru α z tej definicji przedmioty, dla których poprzednik zawartego *implicite* w tej definicji okresu warunkowego byłby fałszywy, i które z tej racji spełniałyby ową definicję. Zbiór α stałby się w ten sposób za obszerny i obejmowałby różne rodzaje przedmiotów (np. przedmioty nie będące w ogóle gametami), których nie chcielibyśmy zaliczyć do klasy $Gm(P)$". Tak jest istotnie i łatwo można to okazać w sposób analogiczny do tego, jakim posłużyliśmy się w przypadku $H(Y)$.

Przypuśćmy, iż klasa α_0 czyni zadość warunkowi L i że u_1 jest przedmiotem, który nie utworzył zygoty z żadnym elementem klasy α_0: $\sim(\exists x)(x \varepsilon U(\{u_1\}, \alpha_0))$. Klasa $U(\alpha_0 \cup \{u_1\}, \alpha_0 \cup \{u_1\})$ jest wówczas identyczna z klasą: $U(\alpha_0, \alpha_0)$, ponieważ nie zmienimy w niczym klasy wszystkich zygot powstałych z połączenia elementów klasy α_0 przez dodanie do klasy α_0 przedmiotu, który nie tworzy zygoty z żadnym elementem

6) Na przykład A. Kaplan, „Definition and Specification of Meaning", *Journal of Philosophy* 43, (1946); H. Mehlberg, op. cit.; H. Mehlberg, *The Reach of Science* (1958).

klasy α_0. Identyczność ta implikuje: $L(\alpha_0 \cup \{u_1\})$. Ale skoro $L(\alpha_0 \cup \{u_1\})$, to $(\exists\alpha)(L(\alpha) \cdot u_1 \,\varepsilon\, \alpha)$, a stąd zgodnie z naszą definicją $Gm(P)$, $u_1 \,\varepsilon\, Gm(P)$. A zatem $Gm(P)$ obejmuje wszystkie przedmioty, które nie utworzyły zygot z elementami jakiejś klasy spełniającej warunek L. Co więcej, klasa ta obejmuje również wszystkie gamety, które odznaczają się tym, iż powstałe z ich połączenia zygoty rozwijają się w środowisku nie należącym do $E(P)$. Ściśle mówiąc, jeśli α_0 jest klasą spełniającą warunek L, wszystkie przedmioty u takie, iż:

$$\sim (\exists x)(\exists y)(\exists z)\,[dlz(x, y, z) \cdot x \,\varepsilon\, U(\{u\}, \alpha_0) \cdot y \,\varepsilon\, E(P)]$$

należą do $Gm(P)$.

Konsekwencji tej można uniknąć stosując procedurę analogiczną do procedury zastosowanej w przypadku $H(Y)$, tj. wzmacniając odpowiednio warunek L i konstruując za jego pomocą definicję cząstkową. Woodger zarysowuje inne wyjście. Pierwotną definicję $Gm(P)$ proponuje zastąpić definicją następującą:

$$Gm(P) = (\hat{u})(\exists v)(\exists X)\,[u(u, v) \,\varepsilon\, X \cdot (x)(y)(z)(dlz(x, y, z) \cdot x \,\varepsilon\, X \cdot y \,\varepsilon\, E(P) \supset z \,\varepsilon\, H(P) \cdot H(P) \neq \emptyset],$$

gdzie $u(u, v)$ ma być funktorem denotującym zygotę powstałą z połączenia gamety u z gametą v. Wydaje się jednak, że propozycja ta nie usuwa wszystkich wadliwości definicji pierwotnej. Niejasny jest przede wszystkim charakter logiczny wyrażenia $u(u, v)$. Gdyby to miał być funktor w sensie właściwym[7], spełniony by być musiał warunek następujący: Dla każdego u i v: istnieje jedna i tylko jedna zygota powstała z połączenia u i v. Ale jest to notorycznym fałszem i pozostałoby fałszem nawet, gdyby zakres zmiennych u i v ograniczyć do klasy gamet, które istotnie połączyły się tworząc zygoty. $u(u, v)$ nie może być zatem uważany za funktor właściwy. Może być natomiast pojmowany jako Russelowska deskrypcja. Ale to w dalszym ciągu prowadzi do pewnych niepożądanych konsekwencji. Przypuśćmy, że u_1 nie utworzył jakiejkolwiek zygoty. Wówczas: $\sim (\exists v)(\exists X)(u(u_1, v) \,\varepsilon\, X)$ i, zgodnie z naszą definicją, $\sim u_1 \,\varepsilon\, Gm(P)$. To konsekwencja nie do przyjęcia. A jeśli u_1 utworzył zygotę, która rozwija się w środowisku nie należącym do $E(P)$, u_1 tym samym należy do $Gm(P)$. W tym przypadku nowa definicja $Gm(P)$ prowadzi do tej samej paradoksalnej konkluzji, co poprzednia. Jedynym rozwiązaniem wydaje się i tu cząstkowa definicja $Gm(P)$.

Przypadek rozważany w pracy niniejszej nie jest czymś wyjątkowym. Przeciwnie, jest to sytuacja typowa dla «pojęć teoretycznych». Wyraźnie stwierdził to Carnap, którego teoria zdań redukcyjnych zdawać miała z tej sytuacji sprawę. Przeprowadzona analiza dostarcza przykładu ilustrującego tę teorię. Streśćmy jej wyniki w metaforyczny nieco sposób. Terminy teoretyczne traktować możemy jako terminy odnoszące się do pewnych niespostrzegalnych struktur. Struktury te przejawiają się w pewnych spostrzegalnych warunkach w postaci spostrzegalnych zjawisk. Tam, gdzie warunki takie nie istnieją, struktury owe nie zdradzają niczym swojej obecności. Tak więc na szczeblu spostrzeżeniowym przybierają one charakter dyspozycji. Tak samo traktować można

7) Porównaj np. R. Carnap, *Introduction to Symbolic Logic and Its Applications*, (1958), str. 71 i nn.

pojęcie genotypu. Odnosi się ono do pewnej niespostrzegalnej struktury — genetycznej konstytucji organizmu — która przejawia się w postaci pewnego spostrzegalnego zachowania — «genetycznego» zachowania się organizmu. Jeśli organizm jakiś w ogóle potomstwa nie posiada, nie można orzec, jak zachowuje się pod względem genetycznym i, co za tym idzie, jaki jest jego genotyp. Dlatego też sens tego pojęcia, podobnie jak innych pojęć dyspozycyjnych, może być — na szczeblu spostrzeżeniowym — ustalony jedynie częściowo. To właśnie miała okazać nasza analiza w bardziej formalny i szczegółowy sposób. Zadanie takie mogło zostać podjęte tylko dzięki uprzedniej konstrukcji precyzyjnego języka genetycznego, który potrafi „ujawnić rzeczywistą złożoność przedmiotu maskowaną przez język naturalny". Wydaje się to niezbędnym założeniem wszelkiej metodologicznej analizy.

Filozofia Nauki
Rok I, 1993, Nr 2-3

Pojęcia teoretyczne a doświadczenie

I

1. Problem stosunku teorii do doświadczenia ma dwa różne, choć ściśle ze sobą związane, aspekty: jeden dotyczący twierdzeń, drugi — pojęć teorii. Możemy pytać bądź o stosunki logiczne łączące twierdzenia teoretyczne z twierdzeniami elementarnymi, bądź o związki definicyjne pomiędzy pojęciami teoretycznymi a pojęciami elementarnymi. Praca niniejsza poświęcona jest rozważeniu pewnych zagadnień związanych z tym drugim aspektem naczelnego problemu. Idzie w niej głównie o to, jakiego rodzaju związki zachodzić mogą pomiędzy tymi dwoma typami pojęć na gruncie teorii empirycznych. Rozważania nasze będą miały charakter dość abstrakcyjny. Chodzić będzie nie tyle o zdanie sprawy z faktycznego stanu rzeczy, ile raczej o przegląd możliwości, jakie się tu zarysowują. Rozważania takie dostarczyć mogą jednak pewnej aparatury pojęciowej, którą wykorzystać można dla analizy istniejących teorii naukowych. Próbę takich zastosowań zawiera m.in. praca „O pojęciu genotypu" [21] Analiza tego pojęcia stanowić może egzemplifikację szeregu pojęć i zależności omawianych w pracy niniejszej. Szereg innych prac cytowanych w toku dalszych rozważań dostarcza również pewnego materiału przykładowego.

Sformułowanie zagadnień będących przedmiotem obecnych dociekań wymaga paru wstępnych wyjaśnień i założeń. Przede wszystkim — wyjaśnień dotyczących terminów: „pojęcie elementarne" i „pojęcie teoretyczne". Problem pojęć teoretycznych jest od szeregu lat obiektem żywych zainteresowań metodologów nauk empirycznych. Spośród wielu prac poświęconych tej problematyce chciałbym zwrócić uwagę na ostatnie prace Carnapa [6], Hempla [9] i Braithwaite'a [2]. We wszystkich tych pracach występuje przeciwstawienie pojęć elementarnych i teoretycznych, rozumianych w podobny, w zasadzie, sposób. W ten sam też sposób używać będę tych terminów w rozważaniach niniejszych. Nie poddając ich zatem bliższej analizie, przypomnę tylko pokrótce, o co w owym przeciwstawieniu idzie. Zamiast o pojęciach mówmy raczej o terminach, do których zaliczymy, w każdym razie, wszelkie pozalogiczne predykaty jedno- lub wieloargumentowe odnoszące się do własności rzeczy lub do stosunków zachodzących pomiędzy rzeczami. Za terminy elementarne uważać będziemy te spośród predykatów, które odnoszą się do spostrzegalnych własności lub stosunków, za terminy teoretyczne — predykaty pozostałe. Nieostre to niewątpliwie rozróżnienie, bo też nieostre jest pojęcie spostrzegalnej własności. Nie będę usiłował go tutaj precyzować, bo dla dalszych rozważań nie jest rzeczą ważną, gdzie dokładnie przeprowadzi się linię graniczną pomiędzy terminami elementarnymi i teoretycznymi.[1] Dodać tylko wypada, że idzie tu w każdym razie o pewne własności rzeczy materialnych, a nie o treści naszych wrażeń. Własności takie uważa się za spostrzegalne, jeśli w odniesieniu do pewnych przedmiotów można bezpośrednio na podstawie spostrzeżenia rozstrzygnąć, czy własności te im przysługują. Jeśli to w stosunku do żadnego przedmiotu nie jest wykonalne, mamy do czynienia z własnością niespostrzegalną. Predykaty spostrzeżeniowe odnoszące się do spostrzegalnych w sensie powyższym własności lub stosunków stanowią najbliżej z bezpośrednim doświadczeniem związane terminy nauk szczegółowych. Mamy więc prawo traktować je jako terminy elementarne na gruncie teorii nauki. Ich ewentualna analiza odwołująca się do treści naszych wrażeń uważana być może za zadanie należące do teorii poznania, a więc do filozofii raczej niż metodologii.

Rola terminów elementarnych w teorii empirycznej nie budzi poważniejszych wątpliwości. Terminy te odnoszą się do bezpośrednio obserwowalnych własności. Dopuszczają zatem bezpośrednią interpretację przyporządkowującą im takie własności. Inaczej jest z terminami teoretycznymi. Nie odnoszą się one do niczego, co byłoby dostępne bezpośredniej obserwacji. Bezpośrednia interpretacja takich terminów nie jest zatem możliwa. Na czym wobec tego polega ich stosowalność do badanej dziedziny przedmiotów? W jaki sposób rozstrzygamy, czy dany termin teoretyczny stosuje się do któregoś z tych przedmiotów, czy nie? Wydaje się, iż stosowalność terminu teoretycznego zagwarantować mogą tylko jego związki z terminami elementarnymi. Dzięki nim uzyskuje on pewną interpretację pośrednią. O jakie związki tutaj idzie? Związki te charakteryzuje się ogólnie, i ogólnikowo zarazem, jako związki definicyjne. Pojmuje je się jednak w sposób bardzo różnorodny. We współczesnych badaniach dotyczących terminów teoretycznych wyróżnić można szereg etapów, charakteryzujących się coraz to bardziej liberalnym pojmowaniem owych związków definicyjnych. I tak, we wczesnych pracach Carnapa dominował pogląd traktujący pojęcia teoretyczne jako „logiczne konstrukcje" z pojęć elementarnych. Terminy teoretyczne przybierały przy tej interpretacji charakter terminów definiowanych *explicite* za pomocą terminów elementarnych. Z chwilą ukazania się *Testability and Meaning* Carnapa zapanowało przekonanie, iż pewne terminy teoretyczne nie mogą zostać zdefiniowane *explicite* za pomocą termi-

1) Dlatego też zamiast o „terminach spostrzeżeniowych" będę mówił raczej o „terminach elementarnych".

nów elementarnych. Właściwym sposobem wprowadzania tych terminów do języka nauki okazały się tzw. definicje cząstkowe, stwierdzające luźniejszy związek z terminami elementarnymi niż definicje zupełne. Ale i takie postawienie sprawy okazało się wkrótce zbyt rygorystyczne. I tak, w pracy *Foundations of Logic and Mathematics* Carnap zwrócił uwagę na to, że w większości teorii przyrodniczych terminy teoretyczne występujące w naczelnych postulatach teorii mają charakter terminów pierwotnych i nie są definiowane za pomocą terminów elementarnych ani na drodze definicji zupełnych, ani cząstkowych. Przeciwnie, to terminy elementarne definiowane są za pomocą teoretycznych, często za pośrednictwem długich łańcuchów definicyjnych. I w ten sposób, nie wprost niejako, zapewniony jest związek terminów teoretycznych z elementarnymi.

Pogląd ten wydaje się trafny. Sugeruje on zarazem inne nieco podejście do problemu terminów teoretycznych, które m. in. znalazło wyraz w cytowanych na wstępie pracach Carnapa, Hempla i Braithwaite'a. Rozpatruje się tu terminy teoretyczne jako wyrażenia systemów aksjomatycznych reprezentujących teorie empiryczne. W języku tych systemów wyróżnia się obok terminów logicznych dwa rodzaje terminów pozalogicznych: terminy teoretyczne i elementarne. Szukając związków łączących terminy teoretyczne z terminami elementarnymi bierzemy pod uwagę nie tylko twierdzenia pełniące w danej teorii rolę definicji, ale i pozostałe twierdzenia teorii. Badamy, czy z postulatów teorii wynikają logicznie twierdzenia formułujące kryteria stosowalności terminów teoretycznych za pomocą terminów elementarnych. Twierdzenia takie mogą mieć postać definicji zupełnych lub definicji cząstkowych; mogą również przybierać postać wyrażającą luźniejsze jeszcze związki pomiędzy tymi dwoma rodzajami terminów. Nie jest tutaj ważne to, czy tezy takie są istotnie twierdzeniami definicyjnymi, czy rozważane terminy teoretyczne rzeczywiście przy pomocy tych twierdzeń do języka danej teorii zostały wprowadzone. To zresztą, wobec nieokreśloności istniejącego języka naukowego, jest często nierozstrzygalne. Trudno jest na ogół odpowiedzieć na pytanie, które z twierdzeń zawierających jakiś termin teoretyczny jest twierdzeniem definicyjnym, a które — rzeczowym. Wyróżnia się niekiedy wśród postulatów teorii hipotezy i postulaty znaczeniowe w ten sposób, iż do hipotez zalicza się twierdzenia sformułowane wyłącznie w terminach teoretycznych, a do postulatów znaczeniowych — twierdzenia, w których figurują zarówno terminy teoretyczne jak i elementarne.[2] Odróżnienie takie ma jednak charakter całkowicie arbitralny. Co więcej, w odniesieniu do obszernej klasy systemów wykazano, że każdy taki system zawierający postulaty pierwszego i drugiego typu może zostać przekształcony na równoważny mu, w pewnym sensie,[3] system zawierający wyłącznie postulaty drugiego rodzaju [3]. W dalszym

2) Tym dwom rodzajom postulatów odpowiadają np. „postulates" i „correspondence rules" u Carnapa [6], „postulates" i „interpretative systems" u Hempla [9], „Campbellian axioms" i „identificatory axioms" u Braithwaite'a [2], [3].

3) Idzie tu o tzw. empiryczną równoważność, omawianą w drugiej części niniejszej pracy.

ciągu rozważań uwzględniać przeto będziemy wszelkie kryteria stosowalności terminów teoretycznych, jakie wynikają z ogółu postulatów teorii. Jeśli kryteria takie mają postać definicji, terminy teoretyczne uważać będziemy za definiowalne przez terminy elementarne (w przypadku definicji cząstkowych — za definiowalne cząstkowo), niezależnie od tego, czy w rozważanej teorii terminy te faktycznie za pomocą takich definicji (ew. definicji cząstkowych) zostały zdefiniowane. Jeżeli zatem z postulatów teorii T wynika twierdzenie o postaci:

$$(x)\,(Qx \equiv \Phi x),$$

gdzie Q jest terminem teoretycznym, a Φ zawiera wyłącznie terminy elementarne, termin Q uważać będziemy za definiowalny w teorii T przez terminy elementarne.[4]

Takie pojęcie związku definicyjnego jest najwyraźniej pojęciem względnym, zrelatywizowanym do określonej teorii. Nabiera ono przy tym precyzyjnego sensu tylko na gruncie teorii stanowiących systemy aksjomatyczne. Takie też tylko teorie będą przedmiotem dalszych rozważań. Ograniczymy się w zasadzie do teorii możliwie prostych pod względem formalnym, a więc do systemów, których środki logiczne nie przekraczają węższego rachunku funkcyjnego. Terminy pozalogiczne takich systemów — to głównie jedno- i wieloargumentowe predykaty. Rozpatrywać będziemy oczywiście tylko takie teorie, których terminy pozalogiczne zawierają zarówno terminy elementarne jak i teoretyczne. Powstaje pytanie, czy istnieją teorie, które nie spełniają tego warunku. Jeśli teoria jakaś ma mieć charakter teorii empirycznej, jej terminy muszą zawierać pewne terminy elementarne. Interpretacja jakiegoś systemu formalnego jako teorii empirycznej polega z reguły na interpretacji pewnych jego wyrażeń jako terminów spostrzeżeniowych. Umożliwia to interpretację pewnych jego tez jako zdań spostrzeżeniowych i, w konsekwencji, empiryczną sprawdzalność systemu. Przytacza się co prawda czasami przykłady teorii, sformułowanych wyłącznie w terminach teoretycznych, np. z dziedziny fizyki teoretycznej. Jeśli jednak teorie takie uważać chcemy za teorie empiryczne, musimy je traktować jako pewne fragmenty jedynie systemów obszerniejszych, w których prócz terminów teoretycznych występują również terminy elementarne. Przyjąć możemy zatem, iż terminy elementarne stanowią niezbędny składnik każdej teorii empirycznej. Czy również terminy teoretyczne występować muszą w każdej takiej teorii? Mówi się niekiedy o teoriach typu „fenomenalistycznego", które sformułowane być mają wyłącznie w języku elementarnym. Nie przesądzając tej sprawy, pominiemy w naszej analizie ten typ teorii, ograniczając się do teorii, w których reprezentowany jest zarówno jeden jak i drugi rodzaj terminów. Wydaje się zresztą, że takie tylko systemy twierdzeń zasługują w pełni na miano „teorii". One też tylko przedstawiają istotne problemy metodologiczne dotyczące stosunku teorii do doświadczenia.

4) Pojęcie to odpowiada, z grubsza biorąc, pojęciu definiowalności terminów wprowadzonemu przez Tarskiego [24].

Jako prosty przykład teorii spełniającej wszystkie wymienione wyżej warunki służyć może teoria skonstruowana przez Braithwaite'a [2], której zresztą w toku późniejszych rozważań poświęcimy szczegółową uwagę. Teoria ta stanowi system aksjomatyczny oparty na węższym rachunku predykatów. System ten składa się z następujących postulatów:

(P1) $(x)\,(Ax \equiv Lx \cdot Mx)$

(P2) $(x)\,(Bx \equiv Mx \cdot Nx)$

(P3) $(x)\,(Cx \equiv Nx \cdot Lx)$.

Wśród terminów pozalogicznych, które występują w tych postulatach, wyróżnić możemy terminy elementarne: A, B, C, odnoszące się do pewnych spostrzegalnych własności, oraz terminy teoretyczne: L, M, N, odnoszące się do pewnych niespostrzegalnych «czynników». Ta fikcyjna, schematyczna teoria stanowi, zdaniem autora, uproszczony wzorzec pewnego rozpowszechnionego typu teorii przyrodniczych.

2. Po tych wstępnych wyjaśnieniach i założeniach mających na celu ustalenie i sprecyzowanie naszego problemu przejdźmy do prób bliższego scharakteryzowania związków definicyjnych łączących terminy teoretyczne z terminami elementarnymi na gruncie teorii empirycznych. W przypadku skrajnym związki te polegać mogą na całkowitej definiowalności terminów teoretycznych przez terminy elementarne. Sytuacji takiej nie będziemy jednak poddawali obecnie szczegółowej analizie. Istnieją być może teorie, które zapewniają terminom teoretycznym kryteria stosowalności o postaci definicji zupełnych. Teorie takie nie stanowią jednakże poważniejszego problemu metodologicznego. Terminy teoretyczne, równoważne na ich gruncie wyrażeniom złożonym wyłącznie z terminów elementarnych, mogą być przez te ostatnie całkowicie wyrugowane. W konsekwencji, teorie tego typu mogą być zawsze zastąpione przez równoważne im teorie sformułowane wyłącznie w języku elementarnym. Warto jednak dodać, że wśród terminów teoretycznych definiowalnych *explicite* za pomocą terminów elementarnych zachodzić mogą istotne różnice pod względem ich empirycznego charakteru. Różnice te zależne są od struktury logicznej wchodzących w grę definicji. Jeśli ich człon definiujący nie zawiera kwantyfikatorów, termin teoretyczny jest stosowalny na podstawie skończonej ilości spostrzeżeń. Obecność kwantyfikatorów sprawia, iż tylko nieskończona liczba obserwacji mogłaby w sposób ostateczny uzasadnić jego zastosowanie w jakimś konkretnym przypadku. Jako przykład terminu pierwszego rodzaju służyć może termin Q_1, definiowalny w sposób następujący:

$$(x)\,(Q_1x \equiv \Phi_1x),$$

terminy drugiego rodzaju ilustrować może termin Q_2:

$$(x)\,(Q_2x \equiv (\exists y)\,(z)\,\Phi_2(x, y, z)).$$

Sprawy te omawiałem bliżej na innym miejscu [19]. Nie zatrzymując się więc nad nimi obecnie, przejdę do omówienia takich związków łączących terminy teoretyczne z terminami elementarnymi, które nie pociągają za sobą całkowitej definiowalności terminów teoretycznych przez terminy elementarne.

Istnienie terminów teoretycznych nie definiowalnych *explicite* przez terminy elementarne wydaje się od czasu *Testability and Meaning* nie ulegać wątpliwości. Należą do nich w pierwszym rzędzie terminy dyspozycyjne, w rodzaju klasycznego już dziś terminu „rozpuszczalny" stanowiącego punkt wyjścia dla Carnapowskich rozważań. Ich rezultatem było stwierdzenie, iż terminy dyspozycyjne są definiowalne przez terminy spostrzeżeniowe tylko częściowo. Właściwym sposobem ich określania jest nie definicja zupełna, lecz tzw. redukcja, stanowiąca poszczególny przypadek definicji cząstkowej (inaczej warunkowej).[5] Do tej samej klasy terminów należy też szereg terminów teoretycznych, których charakter dyspozycyjny nie jest widoczny na pierwszy rzut oka, np. szereg wielkości fizykalnych. W pracy „O pojęciu genotypu" starałem się okazać, iż taki sam charakter przysługuje pewnym biologicznym terminom teoretycznym, w szczególności terminowi „genotyp". Termin ten daje się w sposób adekwatny zdefiniować w terminach elementarnych jedynie na drodze definicji cząstkowej. Powstaje pytanie, czy klasa terminów cząstkowo definiowalnych nie obejmuje wszelkich nie definiowalnych *explicite* terminów teoretycznych. Wrócimy jeszcze do tego zagadnienia w dalszym ciągu pracy. Niezależnie jednak od tego, czy tak jest, czy nie, klasa terminów cząstkowo definiowalnych, stanowiąca obszerną i doniosłą klasę terminów teoretycznych, zasługuje na bliższą analizę.

Teoria definicji cząstkowych jest zbyt znana, aby trzeba ją było tutaj szczegółowo referować.[6] Przypomnę zatem tylko krótko, na czym ów rodzaj definicji — czy raczej pseudodefinicji — polega, oraz omówię pewne logiczne problemy, jakie ta procedura definicyjna nasuwa. Najogólniejszy i najprostszy zarazem schemat definicji cząstkowej terminu Q przedstawić można w postaci następujących dwóch wypowiedzi:

(1) $$(x)\,(\Phi x \supset Qx)$$

(2) $$(x)\,(\Psi x \supset \sim Qx),$$

w których Φ i Ψ reprezentują wyrażenia o uprzednio już ustalonych znaczeniach. W sytuacjach stanowiących przedmiot naszych rozważań Q jest pewnym terminem teoretycznym, a Φ i Ψ — wyrażeniami, w których jako terminy pozalogiczne występują wyłącznie terminy elementarne. Schemat powyższy można z łatwością rozszerzyć na terminy należące do innych kategorii syntaktycznych, m. in. na predykaty wieloargumentowe. Poszczególnym przypadkiem tak rozumianej definicji cząstkowej są wprowadzone przez Carnapa zdania redukcyjne:

(1′) $$(x)\,(\Phi_1 x \supset (\Phi_2 x \supset Qx))$$

(2′) $$(x)\,(\Psi_1 x \supset (\Psi_2 x \supset \sim Qx)),$$

dostosowane do określania terminów dyspozycyjnych. Zwłaszcza pewne ich uszczegółowienie zwane obustronnym zdaniem redukcyjnym:

(1″) $$(x)\,(\Phi_1 x \supset (Qx \equiv \Phi_2 x))$$

5) Tak jest w każdym razie na gruncie ekstensjonalnych systemów logicznych. Ale też do takich tylko systemów ograniczamy nasze rozważania.

6) Por. np. Carnap [4], Mehlberg [13], [14], Stopes-Roe [23], Kamińska [11], Przełęcki [17], [18], [19].

znajduje częste zastosowanie w praktyce definiowania terminów dyspozycyjnych. Φ_1 stanowi tu opis pewnej sytuacji doświadczalnej, a Φ_2 — opis pewnego zachowania się w takiej sytuacji. Jako przykład schematu (1″) — a tym samym ogólniejszych od niego schematów (1′)-(2′) oraz (1)-(2) — służyć może następująca, uproszczona oczywiście z fizykalnego punktu widzenia, cząstkowa definicja „magnesu":

Jeśli w pobliżu *x* znajduje się niewielki kawałek żelaza, to *x* jest magnesem wtedy i tylko wtedy, gdy *x* przyciąga ten kawałek.

Co odróżnia definicję cząstkową od definicji zupełnej? Wypowiedzi (1)-(2) pozwalają na zastosowanie terminu Q do przedmiotów spełniających warunek Φ, a negacji terminu Q do przedmiotów spełniających warunek Ψ. Gdyby warunek Ψ był równoważny logicznie negacji warunku Φ, każdy przedmiot musiałby spełniać któryś z tych warunków i o każdym z nich zatem można by na podstawie takiej definicji orzec bądź termin Q, bądź jego negację. Byłaby to sytuacja charakterystyczna dla definicji zupełnej. Istotnie, gdy:

$$(x)\,(\Psi x \equiv \sim \Phi x),$$

wypowiedzi (1)-(2) przechodzą w definicję zupełną terminu Q:

$$(x)\,(Qx \equiv \Phi x).$$

Definicję zupełną traktować można zatem jako poszczególny przypadek definicji cząstkowej. Ta ostatnia przedstawia pewną ogólniejszą procedurę definicyjną, obejmującą również takie przypadki, w których powyższa równoważność nie zachodzi. Mogą wówczas istnieć przedmioty, które nie spełniają ani warunku Φ ani Ψ, i o których zatem ani terminu Q, ani jego negacji na podstawie takiej definicji orzec nie można. Na tym polega cząstkowy charakter omawianej procedury. Formułuje ona kryteria stosowalności terminu Q dla pewnej tylko klasy przedmiotów i tym samym częściowo tylko determinuje jego znaczenie.

2.1. Ta charakterystyczna cecha definicji cząstkowej odróżniająca ją od definicji w ścisłym tego słowa znaczeniu nasuwa pewien logiczny problem, któremu chciałbym obecnie poświęcić nieco uwagi. Mam na myśli wysuwany niekiedy problem sensowności terminów zdefiniowanych cząstkowo. Sensowność takich terminów podawano w wątpliwość powołując się głównie na to, iż w stosunku do szeregu przedmiotów nie dysponujemy żadnym kryterium ich stosowalności.[7] Czy mamy wobec tego prawo traktować je jako terminy w pełni sensowne, wyposażone w ściśle określone znaczenie? Te wątpliwości dotyczące sensowności terminów zdefiniowanych cząstkowo mogą przybierać różny zasięg. Nie zawsze wyraźnie zdawano sobie z tego sprawę. Dobrze więc będzie rozejrzeć się bliżej w różnych wersjach takiego stanowiska.

Najradykalniejsza z nich traktuje terminy cząstkowo definiowalne po prostu jako wyrażenia bezsensowne. Taki sam, w konsekwencji, charakter przypisuje wszelkim

---

7) Podkreśla się czasami i to, że zakres takich terminów nie jest wyznaczony w sposób jednoznaczny. Definicja cząstkowa typu (1)-(2) charakteryzuje zakres terminu Q jako klasę zawierającą klasę Φ i zawartą w klasie $\sim \Psi$. Nie jest to oczywiście charakterystyka jednoznaczna, gdyż warunek taki spełniony być może przez szereg różnych klas.

wypowiedziom, w których figuruje jakiś termin cząstkowo definiowalny. Prowadzi to do poważnych trudności przy próbach ustosunkowania się do istniejących teorii naukowych. Przeważająca ich większość zawiera terminy teoretyczne, które definiowalne są jedynie cząstkowo za pomocą terminów elementarnych. A te ostatnie tylko uważa się za wyrażenia wyposażone z góry w określony empiryczny sens. Naczelne hipotezy teorii przyrodniczych z reguły sformułowane bywają za pomocą terminów teoretycznych tego właśnie rodzaju. Czyż możemy hipotezy te uważać za wypowiedzi pozbawione jakiegokolwiek znaczenia? A jednak do takiego właśnie wniosku prowadzi omawiane stanowisko. Jego zwolennicy traktują jako właściwe twierdzenia teorii tylko te jej wypowiedzi, które sformułowane są wyłącznie w terminach elementarnych, ew. w terminach *explicite* za pomocą tych ostatnich zdefiniowanych. Nazwijmy je twierdzeniami elementarnymi. Wszystkie wypowiedzi pozostałe, a wśród nich i naczelne hipotezy teorii, uważane są nie za właściwe twierdzenia, lecz wyłącznie za środki umożliwiające systematyzację twierdzeń elementarnych polegającą na włączeniu ich do pewnego systemu aksjomatycznego. Usiłuje się niekiedy dokonać tej systematyzacji bez odwoływania się do owych problematycznych wypowiedzi teoretycznych, wyłącznie w obrębie twierdzeń elementarnych. Osiągnięto w tej dziedzinie pewne interesujące wyniki, które wskazują na to, iż zawsze, w zasadzie, istnieje procedura umożliwiająca przedstawienie ogółu elementarnych twierdzeń teorii jako systemu aksjomatycznego — o nieskończonej, co prawda, liczbie aksjomatów.[8] Na ogół jednakże nie rezygnuje się całkowicie z wypowiedzi teoretycznych jako części składowych teorii, tylko poddaje się je pewnym interpretacjom czy modyfikacjom. Takie stanowisko reprezentuje np. Ramsey.[9] Postulaty teorii zawierające terminy nie definiowalne *explicite* przez terminy elementarne nie są, wedle tego poglądu, zdaniami posiadającymi określoną wartość logiczną, lecz raczej czymś w rodzaju funkcji zdaniowych, w których terminy teoretyczne pełnią rolę zmiennych. Zastąpmy je przeto we wszystkich postulatach teorii przez odpowiednie zmienne, a przed koniunkcją tak zmodyfikowanych postulatów postawmy kwantyfikatory szczegółowe wiążące owe zmienne. Tego rodzaju wyrażenie jest w przeciwieństwie do koniunkcji oryginalnych postulatów zdaniem o określonej wartości logicznej. Pociąga przy tym za sobą dokładnie te same twierdzenia elementarne, co oryginalne postulaty teorii.

Zilustrujmy opisaną procedurę na prostym przykładzie. Sięgnijmy w tym celu do cytowanej we wstępie teorii Braithwaite'a. Występujące w postulatach tej teorii terminy teoretyczne: L, M, N, okazują się nie definiowalne *explicite* na gruncie tej teorii przez terminy elementarne A, B, C. Jeżeli terminom tym odmawiamy z tej racji charakteru wyrażeń sensownych, możemy je wyrugować zastępując postulaty (P1)-(P3) następującym twierdzeniem sformułowanym wyłącznie za pomocą terminów elementarnych:

8) Por. Craig [7], [8]. Wyniki te omawia m. in. Hempel [9].
9) Praca pt. „Theories" w zbiorze [22]. Poglądy Ramsey'a przytacza Braithwaite [2], Hempel [9].

(P') $(\exists X)(\exists Y)(\exists Z)(x)((Ax \equiv Xx \cdot Yx) \cdot (Bx \equiv Yx \cdot Zx) \cdot (Cx \equiv Zx \cdot Xx))$.
Stwierdza ono, iż istnieją pewne własności czyniące zadość postulatom (P1)-(P3), nie mówiąc, o jakie to konkretne własności chodzi. Z postulatu (P') wynikają przy tym te same dokładnie twierdzenia zawierające wyłącznie terminy elementarne: A, B, C, co z postulatów (P1)-(P3).

Postępowanie powyższe ma jednak poważną niedogodność. Wyprowadza nas ono poza zakres systemów elementarnych, angażując środki logiczne szerszego rachunku funkcyjnego. Wszak w postulacie (P') występują oprócz zmiennych indywidualnych, również zmienne wyższego rzędu jako zmienne związane. Dla nikogo więc, kto nie chce wykraczać poza systemy elementarne, procedura taka nie jest do przyjęcia.

Obok najbardziej rygorystycznej wersji kryterium sensowności, odmawiającej sensownego charakteru wszelkim wyrażeniom zawierającym terminy cząstkowo definiowalne, spotkać możemy wersje bardziej liberalne, kwestionujące sensowność pewnych tylko wypowiedzi posługujących się takimi terminami. Jakie wypowiedzi ma się tu na myśli? Załóżmy, iż jedynym kryterium stosowalności terminu Q jest definicja cząstkowa cytowana poprzednio:

(1) $(x)(\Phi x \supset Qx)$

(2) $(x)(\Psi x \supset \sim Qx)$.

Definicja ta formułuje kryteria stosowalności terminu Q dla przedmiotów, które są Φ lub Ψ, pozwalając tym samym na rozstrzyganie twierdzeń przypisujących termin Q takim przedmiotom. Nie widać zatem powodu, dla którego sensowność takich powiedzeń trzeba by podawać w wątpliwość. Inaczej jednak przedstawia się sprawa twierdzeń orzekających termin Q o przedmiotach, które nie są ani Φ, ani Ψ. Twierdzenia takie są zasadniczo nierozstrzygalne, gdyż definicja terminu Q żadnych kryteriów stosowalności dla tych przypadków nie przewiduje. Formułując takie twierdzenia wykraczamy poza ustalone dla terminu Q kryteria stosowalności. Toteż wypowiedzi takie uważa się za wyrażenia bezsensowne.[10]

Stanowisko to wydaje się dość przekonujące. Ale i ono prowadzi do pewnych niewygodnych konsekwencji. Sensowność pewnych wyrażeń uzależnia się tu od faktów pozajęzykowych. Sama poprawność budowy nie gwarantuje sensowności wyrażeniom zawierającym terminy cząstkowo definiowalne. To, czy przedmiot a spełnia warunek Φ lub Ψ, jest pewnym faktem empirycznym, o którym w zasadzie może nas przekonać tylko doświadczenie. A od tego przecież zależy, czy wyrażenie: Qa jest wypowiedzią sensowną. Trzeba jednak zauważyć, iż nie jest to jedyna sytuacja, w której mamy do czynienia z pozaformalnymi kryteriami sensowności wyrażeń. Analogiczny charakter ma, wedle pewnych systemów logicznych,[11] tzw. operator deskrypcyjny. To, czy wyrażenie: $\iota x\, \Phi x$ jest wyrażeniem sensownym, zależy na gruncie tych systemów od tego, czy warunek Φ spełniony jest dokładnie przez jeden przedmiot.

10) Tak stawiałem sprawę w pracy [20].

11) Np. w systemie Hilbert-Bernays [10].

O tym zaś, w przypadku terminów pozalogicznych, przekonać się możemy w zasadzie tylko na drodze doświadczalnej. Jest to jednak niewątpliwie konsekwencja wysoce kłopotliwa. Starano się jej też uniknąć, wprowadzając w pewnych systemach logicznych operator deskrypcyjny na innej drodze, mimo iż interpretacja poprzednia oddaje chyba najlepiej sens zwrotów potocznych odpowiadających temu operatorowi.

Spotkać się można również z argumentacją następującą, prowadzącą do jeszcze bardziej tolerancyjnego stanowiska. Twierdzenie przypisujące termin Q przedmiotowi a nie spełniającemu żadnego z warunków: Φ lub Ψ jest co prawda nierozstrzygalne, lecz niemożliwość jego rozstrzygnięcia nie ma charakteru logicznego. To że: $\sim \Phi a \cdot \sim \Psi a$ jest pewnym faktem «przypadkowym». Nie ma sprzeczności w przypuszczeniu, iż jest przeciwnie. Gdyby zaś okazało się, iż: $\Phi a \vee \Psi a$, twierdzenie Qa byłoby twierdzeniem rozstrzygalnym. Biorąc to pod uwagę, niektórzy nie są skłonni odmawiać sensownego charakteru powyższym twierdzeniom, które tylko wskutek «przypadkowego» stanu rzeczy okazują się nierozstrzygalne. Jedynie logiczna niemożliwość rozstrzygnięcia danego twierdzenia dyskwalifikuje je pod względem sensowności. A taki właśnie charakter mają mieć np. twierdzenia następujące: $\sim \Phi a \cdot \sim \Phi a \cdot Qa$ lub $\sim \Phi a \cdot \sim \Phi a \cdot \sim Qa$ orzekające termin Q lub jego negację o przedmiotach, którym jednocześnie odmawia się własności Φ i Ψ. Jeśli przytoczona definicja cząstkowa terminu Q stanowi jedyne kryterium jego stosowalności, okazanie prawdziwości takich wypowiedzi jest istotnie logiczną niemożliwością, gdyż trzeba by w tym celu stwierdzić, iż: $\sim \Phi a \cdot \sim \Psi a \cdot \Phi a$ lub $\sim \Phi a \cdot \sim \Psi a \cdot \Psi a$.[12] Uważając tego rodzaju konstrukcje za wypowiedzi bezsensowne, nie wykraczamy poza formalne kryteria sensowności wyrażeń. Kryteria te komplikujemy jednak znacznie. Terminami cząstkowo definiowalnymi operować musimy ostrożniej niż pozostałymi, gdyż w pewnych kontekstach prowadzą one do pozbawionych sensu wyrażeń.

Najbardziej liberalne z rozważanych stanowisk nie kwestionuje sensownego charakteru terminów cząstkowo definiowalnych, i to niezależnie od sytuacji czy kontekstu, w których terminów tych używamy. Istnienie jakichś, choćby częściowych tylko, kryteriów stosowalności danego terminu wystarcza dla uznania go za wyrażenie w pełni sensowne. W stosunku do terminu Q obowiązują takie same kryteria sensowności co w stosunku do wszelkich jednoargumentowych predykatów. Unika się w ten sposób owych komplikacji «technicznych», o których wspominałem poprzednio. Cel ten osiąga się jednak za cenę pewnych komplikacji «filozoficznych». Stosowanie terminów Q do przedmiotów nie spełniających warunków Φ i Ψ nie prowadzi tu do wypowiedzi bezsensownych. Nie jest też wypowiedzią bezsensowną stwierdzenie, iż dany przedmiot nie spełnia żadnego z tych warunków a jednocześnie podpada pod termin Q.

12) Można jednak okazać fałszywość takich wypowiedzi. Odwrotnie oczywiście przedstawia się sprawa ich negacji: $\Phi a \vee \Psi a \vee \sim Qa$ lub $\Phi a \vee \Psi a \vee Qa$. Można okazać ich prawdziwość, natomiast okazanie ich fałszywości staje się logicznie niemożliwe. Zarówno jedne jak i drugie można by uważać z tej racji za konstrukcje niedopuszczalne.

Wypowiedzi te, będąc rzetelnymi zdaniami, są jednakże twierdzeniami — «faktycznie» lub nawet «logicznie» — nierozstrzygalnymi. Wprowadza się tu zatem do języka nauki zdania, które mają określone znaczenie, a których mimo to nigdy nie potrafimy rozstrzygnąć. Nie jest to rozwiązanie zadowalające, ale nie są nimi również, jak widzieliśmy, rozwiązania poprzednie. To zaś ma nad tamtymi tę wyższość, iż wydaje się bardziej zgodne z faktyczną praktyką badawczą. Terminy teoretyczne funkcjonujące w języku naukowym traktowane są z reguły jako wyrażenia w pełni sensowne niezależnie od tego, czy ich empiryczne kryteria stosowalności mają charakter zupełny czy częściowy. Język ten wydaje się zatem tolerować istnienie sensownych choć nierozstrzygalnych twierdzeń, będących rezultatem nieskrępowanego używania takich terminów.

2.2. Definicja cząstkowa:

(1) $$(x)\,(\Phi x \supset Qx)$$

(2) $$(x)\,(\Psi x \supset \sim Qx)$$

różni się, jak widzieliśmy, od definicji zupełnej:

$$(x)\,(Qx \equiv \Phi x)$$

tym, iż warunki Φ i Ψ, w przeciwieństwie do warunków Φ i $\sim\Phi$, logicznie się nie dopełniają. Mogą więc istnieć przedmioty, które nie spełniają żadnego z tych warunków. Ale warunki Φ i Ψ różnią się od warunków Φ i $\sim\Phi$ również i tym, że logicznie się nie wykluczają. Mogą istnieć przedmioty, które spełniają oba te warunki zarazem. Nie ma logicznej sprzeczności w przypuszczeniu, iż:

$$(\exists x)\,(\Phi x \cdot \Psi x).$$

Koniunkcja wypowiedzi (1) i (2) pociąga jednak za sobą twierdzenie wykluczające taką ewentualność:

$$(x) \sim (\Phi x \cdot \Psi x).$$

W twierdzeniu tym nie występuje termin definiowany Q, lecz wyłącznie terminy o ustalonych już uprzednio znaczeniach. Skoro nie ma ono charakteru tezy logicznej, jaki ma analogiczne twierdzenie implikowane przez definicję zupełną:

$$(x) \sim (\Phi x \cdot \sim \Phi x),$$

mamy tu do czynienia ze zdaniem syntetycznym, wyrażającym pewną doświadczalną zależność. A więc i definicja cząstkowa, której jest ono konsekwencją, przybiera — w przeciwieństwie do definicji zupełnej będącej zawsze zdaniem analitycznym — charakter twierdzenia doświadczalnego. Zdania redukcyjne:

(1′) $$(x)\,(\Phi_1 x \supset (\Phi_2 x \supset Qx))$$

(2′) $$(x)\,(\Psi_1 x \supset (\Psi_2 x \supset \sim Qx)),$$

stanowiące pewną szczegółową odmianę definicji cząstkowej, implikują podobną konsekwencję:

$$(x) \sim (\Phi_1 x \cdot \Phi_2 x \cdot \Psi_1 x \cdot \Psi_2 x).$$

I one zatem mają charakter twierdzeń doświadczalnych. Jedynie obustronne zdanie redukcyjne:

(1″) $$(x)\,(\Phi_1 x \supset (Qx \equiv \Phi_2 x))$$

reprezentuje takie uszczegółowienie definicji cząstkowej, które przybiera charakter zdania analitycznego, gdyż konsekwencja:

$$(x) \sim (\Phi_1 x \cdot \Phi_2 x \cdot \Phi_1 x \cdot \sim \Phi_2 x)$$

jest twierdzeniem logicznie prawdziwym. Ale i w przypadku obustronnych zdań redukcyjnych napotykamy problem analogiczny. Zdanie takie, podobnie jak wszelka definicja cząstkowa, determinuje znaczenie terminu Q tylko częściowo. Nie przewiduje żadnych kryteriów stosowalności dla przedmiotów nie spełniających warunku Φ_1. Dlatego też chcąc rozszerzyć zakres stosowalności terminu Q musimy wprowadzić dalsze zdania redukcyjne odwołujące się do innych kryteriów stosowalności:

(2″) $$(x)(\Phi_3 x \supset (Qx \equiv \Phi_4 x)).$$ [13]

Kryteria te z reguły zachodzą na siebie. Koniunkcja obustronnych zdań redukcyjnych (1″) i (2″) przybiera wskutek tego charakter twierdzenia doświadczalnego, implikuje bowiem następującą syntetyczną konsekwencję:

$$(x)(\Phi_1 x \cdot \Phi_3 x \supset (\Phi_2 x \equiv \Phi_4 x)).$$

Ta charakterystyczna cecha definicji cząstkowych stała się, podobnie jak i poprzednia, źródłem pewnej logicznej trudności. Istnienie tego typu twierdzeń wysuwano mianowicie jako argument przeciwko tradycyjnemu podziałowi zdań na zdania analityczne i syntetyczne.[14] Przyjrzyjmy się bowiem sytuacji, w jakiej znajdują się definicje cząstkowe traktowane jako pewne procedury definicyjne. Z jednej strony wydają się mieć one charakter zdań analitycznych. Termin Q występujący w wypowiedziach (1)-(2) nie jest skądinąd wyposażony w jakiekolwiek znaczenie. Właśnie wypowiedzi (1)-(2) określają jego znaczenie jako takie, przy którym wypowiedzi te stają się prawdziwe. Tego rodzaju wypowiedzi zalicza się zwykle do zdań analitycznych. Z drugiej strony jednak, definicje cząstkowe pociągają za sobą pewne twierdzenia, o których prawdziwości decyduje w zasadzie doświadczenie. Same więc podlegają kontroli doświadczenia, a to uważa się na ogół za charakterystyczną cechę zdań syntetycznych.[15] Definicje te wydają się pełnić jednocześnie dwie role: twierdzeń definicyjnych i twierdzeń rzeczowych. Z tego powodu dzielą one zarówno pewne charakterystyczne własności zdań analitycznych jak i syntetycznych.

Wyciąganie z powyższego faktu wniosku o niemożliwości utrzymania tradycyjnego rozróżnienia pomiędzy zdaniami analitycznymi a syntetycznymi wydaje się jednak zbyt pochopne. Można bronić poglądu, iż definicja cząstkowa (koniunkcja wypowiedzi (1) i (2) lub zdań redukcyjnych (1′) i (2′) lub obustronnych zdań redukcyjnych (1″) i (2″)) jest po prostu zdaniem syntetycznym, a jej charakterystyczne właściwości płyną stąd, iż jest ona logicznie równoważna koniunkcji zdania analitycznego i zdania syntetycznego.

---

13) Sytuacje takie rozważałem w pracy [19].

14) Stanowisko takie zajmuje np. Pap [16].

15) Zwracano uwagę na to, że i zdania analityczne zależne są w pewien sposób od doświadczenia (por. Ajdukiewicz [1]). Ale nawet jeśli zgodzimy się z tym poglądem, stwierdzić musimy, że w przypadku definicji cząstkowej mamy do czynienia z pewną zależnością swoistą, nie występującą w przypadku definicji zupełnej.

Pierwsze z nich ma charakter twierdzenia definicyjnego ustalającego znaczenie terminu Q, drugie — twierdzenia rzeczowego, w którym termin Q w ogóle nie występuje. Sposób rozbicia definicji cząstkowej na te dwa różne co do swego logicznego charakteru czynniki nie jest jednak czymś narzucającym się w sposób oczywisty. Nie polega ona w każdym razie na tym, aby jedną z wchodzących tu w grę wypowiedzi ((np. (1) lub (1′) lub (1″)) traktować jako twierdzenie definicyjne, a drugą ((2) lub (2′) lub (2″)) — jako twierdzenie rzeczowe, w którym termin Q wzięty jest w znaczeniu nadanym mu przez wypowiedź poprzednią.[16] Rozwiązanie takie, całkowicie arbitralne, nie zdaje w sposób adekwatny sprawy z tego, co w definicji cząstkowej ma charakter postulatu znaczeniowego, a co — przyrodzonego prawa. Elementy te wyodrębnić musimy w inny, bardziej skomplikowany sposób. Wydaje się zresztą, iż definicja cząstkowa dopuszcza pod tym względem różne interpretacje. Rozpatrzmy je na najczęściej dyskutowanym przykładzie obustronnych zdań redukcyjnych.

Dwa obustronne zdania redukcyjne formułujące dwa różne kryteria stosowalności terminu Q:

(1) $$(x)\,(\Phi_1 x \supset (Qx \equiv \Phi_2 x))$$

(2) $$(x)\,(\Phi_3 x \supset (Qx \equiv \Phi_4 x))$$

implikują, jak wiemy, następującą konsekwencję o charakterze przyrodzonego prawa:

(3) $$(x)\,(\Phi_1 x \cdot \Phi_3 x \supset (\Phi_2 x \equiv \Phi_4 x)).$$

Ona to właśnie reprezentuje syntetyczny składnik zdań redukcyjnych (1)-(2). A co stanowi ich składnik analityczny? Aby go sformułować, winniśmy wypowiedzi (1)-(2) zmodyfikować tak, aby nie pociągały żadnej konsekwencji doświadczalnej a jednocześnie zaopatrywały termin Q w te same, co poprzednio, kryteria stosowalności. W pracy [19] proponowałem w tym celu następującą procedurę. Zdanie redukcyjne (1) określa znaczenie terminu Q dla przedmiotów spełniających warunek Φ_1. Zdanie redukcyjne (2) należy wobec tego sformułować tak, aby określało znaczenie terminu Q tylko dla tych przedmiotów, dla których zdanie redukcyjne (1) żadnych kryteriów stosowalności terminu Q nie przewiduje, tj. dla przedmiotów nie spełniających warunku Φ_1. Przybierze ono teraz postać następującą:

(2′) $$(x)\,(\Phi_3 x \cdot {\sim}\,\Phi_1 x \supset (Qx \equiv \Phi_4 x)).$$

Koniunkcja wypowiedzi (1) i (2′) nie pociąga już żadnych konsekwencji doświadczalnych i uważana być może za analityczny składnik zdań redukcyjnych (1)-(2). Tak więc koniunkcja obustronnych zdań redukcyjnych (1) · (2) przedstawiona być może w sposób równoważny jako koniunkcja zdania analitycznego: (1) · (2′) oraz zdania syntetycznego: (3). Czynnik pierwszy jest postulatem znaczeniowym wprowadzającym termin Q, czynnik drugi — prawem przyrodzonym sformułowanym bez użycia terminu Q.

16) Taką koncepcję wydają się sugerować wywody Nagla [15]. Słusznej krytyce poddają ją Braithwaite [3] i Pap [16], którzy jednak nie dostrzegają innych możliwości wyodrębnienia omawianych komponentów definicji cząstkowych.

Całość jako koniunkcja zdania analitycznego i syntetycznego jest oczywiście zdaniem syntetycznym i mieści się doskonale w przyjętej powszechnie klasyfikacji twierdzeń.

Zarysowana procedura ma jednak tę wadę, że nie traktuje owych wyjściowych zdań redukcyjnych (1)-(2) w sposób symetryczny. Jedno z nich: (1) uważane jest za podstawowe, drugie: (2) — za dodatkowe, formułujące jedynie uzupełniające w stosunku do pierwszego kryteria stosowalności. W przypadku większej liczby takich zdań redukcyjnych musielibyśmy ustalić, z tego punktu widzenia, określoną ich kolejność. Bywają chyba sytuacje, kiedy taka interpretacja zdań redukcyjnych odpowiada praktyce naukowej. Pewne kryterium stosowalności danego terminu teoretycznego wyróżnione jest jako podstawowe, a dalsze wprowadzane są tylko jako kolejne uzupełnienia kryterium naczelnego. Ale na pewno nie wszelkie przypadki podpadają pod taki schemat. Wydaje się, iż na ogół różne kryteria stosowalności danego terminu teoretycznego traktowane są równorzędnie. Żadne z nich nie jest wyróżnione jako kryterium podstawowe, a tym bardziej, nie ma mowy o istnieniu wśród nich jakiejś określonej hierarchii. Jeszcze w większym stopniu dotyczy to pozostałych sytuacji, o których mowa była poprzednio. Definicja cząstkowa w postaci ogólnej lub w formie pary zdań redukcyjnych składa się z dwóch wypowiedzi, z których jedna formułuje kryterium stosowalności terminu Q, druga — jego negacji. Nie widać żadnego powodu do tego, aby którąś z tych wypowiedzi wyróżniać jako podstawową i odpowiednio do tego modyfikować pozostałą. A to właśnie zaleca omawiana procedura w celu nadania tym wypowiedziom charakteru zdań analitycznych.

Inne nieco rozwiązanie, unikające tej konsekwencji, proponuje Mehlberg [14]. W rozważanym przez nas przypadku obustronnych zdań redukcyjnych (1)-(2) rozwiązanie to sprowadza się do następującej ich modyfikacji:

(1′) $(x)\,(\Phi_1 x \cdot \sim \Phi_3 x \supset (Qx \equiv \Phi_2 x))$

(2′) $(x)\,(\Phi_3 x \cdot \sim \Phi_1 x \supset (Qx \equiv \Phi_4 x))$.

Wypowiedzi (1′)-(2′), podając wyłączające się wzajemnie kryteria, nie pociągają żadnych doświadczalnych konsekwencji i mogą być uważane za analityczny komponent zdań redukcyjnych (1)-(2). W porównaniu z modyfikacją proponowaną poprzednio rozwiązanie obecne ma tę wyższość, iż oba zdania redukcyjne traktuje równorzędnie. Każde z nich zostaje ograniczone tak, aby podawało kryteria stosowalności terminu Q dla tych tylko przedmiotów, dla których pozostałe nie przewidywało żadnych. Nie narażone też jest z tego powodu na zarzuty wysuwane uprzednio. Obciążone jest natomiast wadliwością inną. Nie zaopatruje mianowicie terminu Q w te same kryteria stosowalności co pierwotne zdania redukcyjne (1)-(2). Termin ten ma obecnie węższy zakres stosowalności niż poprzednio. Zakres ten obejmuje z jednej strony przedmioty, które są Φ_1 i $\sim \Phi_3$, z drugiej — przedmioty, które są Φ_3 i $\sim \Phi_1$. Nie należą do niego natomiast żadne spośród przedmiotów spełniających zarówno warunek Φ_1 jak i Φ_3, mimo iż pierwotne zdania redukcyjne (1)-(2) przewidywały również dla takich przedmiotów kryteria stosowalności terminu Q. Wady tej nie posiada rozwiązanie poprzednie. Zgodnie z nim, w stosunku do przedmiotów spełniających warunki Φ_1 i Φ_3

obowiązują kryteria stosowalności terminu Q z pierwszego zdania redukcyjnego. Proponowane obecnie wypowiedzi (1')-(2') nie formułują zatem w sposób wyczerpujący postulatu znaczeniowego zawartego w zdaniach redukcyjnych (1)-(2) i wymagają istotnego rozszerzenia.

Propozycję uwzględniającą te zastrzeżenia przedstawić chciałbym w sposób ogólny, dotyczący definicji cząstkowej o postaci dowolnej. Jako konsekwencję tego rozwiązania ogólnego otrzymamy pewną modyfikację omawianych obecnie obustronnych zdań redukcyjnych, zapewniającą im charakter analityczny a zarazem dostatecznie szeroko formułującą kryteria definicyjne w nich zawarte. Definicja cząstkowa:

(1) $$(x)\,(\Phi x \supset Qx)$$

(2) $$(x)\,(\Psi x \supset {\sim} Qx)$$

wyposaża termin Q w kryteria stosowalności w stosunku do przedmiotów: $\Phi \vee \Psi$. Implikując jednak konsekwencję:

(3) $$(x) \sim (\Phi x \cdot \Psi x)$$

ogranicza faktycznie tę dziedzinę stosowalności do przedmiotów: $\Phi \cdot {\sim} \Psi$ oraz $\Psi \cdot {\sim} \Phi$. Pierwsze z nich kwalifikuje jako podpadające pod termin Q, drugie — pod jego negację. Formułując zatem zdania (1)-(2) jako:

(1') $$(x)\,(\Phi x \cdot {\sim} \Psi x \supset Qx)$$

(2') $$(x)\,(\Psi x \cdot {\sim} \Phi x \supset {\sim} Qx)$$

formułujemy je tak, iż po pierwsze, żadnej syntetycznej konsekwencji za sobą nie pociągają, po drugie — zaopatrują termin Q w dokładnie te same kryteria stosowalności co zdania (1)-(2). Zarazem obie wypowiedzi wyjściowe (1)-(2) traktowane są w sposób ściśle symetryczny. Koniunkcja zdań (1') · (2') · (3) jest równoważna logicznie koniunkcji wypowiedzi (1) · (2) a charakteryzuje się tym, iż jeden jej czynnik: (1') · (2') ma charakter zdania analitycznego, drugi: (3) — syntetycznego. W ten sposób rozwiązanie to czyni zadość wszystkim wysuniętym wyżej warunkom.

Z łatwością daje się ono zastosować do tego uszczegółowienia definicji cząstkowej, jakim jest para zdań redukcyjnych:

(1) $$(x)\,(\Phi_1 x \supset (\Phi_2 x \supset Qx))$$

(2) $$(x)\,(\Psi_1 x \supset (\Psi_2 x \supset {\sim} Qx)),$$

Zgodnie z proponowanym rozwiązaniem, koniunkcja tych zdań przedstawiona być może w sposób logicznie równoważny jako koniunkcja zdań analitycznych:

(1') $$(x)\,(\Phi_1 x \cdot \Phi_2 x \cdot {\sim} (\Psi_1 x \cdot \Psi_2 x) \supset Qx)$$

(2') $$(x)\,(\Psi_1 x \cdot \Psi_2 x \cdot {\sim} (\Phi_1 x \cdot \Phi_2 x) \supset {\sim} Qx)$$

oraz zdania syntetycznego:

(3) $$(x) \sim (\Phi_1 x \cdot \Phi_2 x \cdot \Psi_1 x \cdot \Psi_2 x)\,.$$

Czynnik pierwszy reprezentuje postulat znaczeniowy zawarty w zdaniach redukcyjnych (1)-(2), czynnik drugi — implikowane przez nie prawo przyrodzone.

A jakie modyfikacje przewiduje proponowane rozwiązanie w przypadku obustronnych zdań redukcyjnych? Załóżmy, jak poprzednio, iż dla terminu Q istnieją dwa takie zdania odwołujące się do dwóch różnych kryteriów stosowalności:

(1) $(x)\,(\Phi_1 x \supset (Qx \equiv \Phi_2 x))$

(2) $(x)\,(\Phi_3 x \supset (Qx \equiv \Phi_4 x))$.

Zapiszmy je w postaci następujących wypowiedzi:

(1a) $(x)\,(\Phi_1 x \cdot \Phi_2 x \supset Qx)$

(1b) $(x)\,(\Phi_1 x \cdot \sim \Phi_2 x \supset \sim Qx)$

(2a) $(x)\,(\Phi_3 x \cdot \Phi_4 x \supset Qx)$

(2b) $(x)\,(\Phi_3 x \cdot \sim \Phi_4 x \supset \sim Qx)$.

Zastosowanie do tych ostatnich przyjętej przez nas procedury nie nastręcza żadnych trudności. Trzeba po prostu, tak jak poprzednio, każdą z tych wypowiedzi przekształcić w ten sposób, aby przypisywała termin Q (ew. negację terminu Q) tylko tym przedmiotom, którym żadna z wypowiedzi pozostałych nie przypisuje negacji terminu Q (ew. terminu Q). Modyfikacja taka prowadzi do następujących sformułowań:

(1a′) $(x)\,(\Phi_1 x \cdot \Phi_2 x \cdot \sim (\Phi_3 x \cdot \sim \Phi_4 x) \supset Qx)$

(1b′) $(x)\,(\Phi_1 x \cdot \sim \Phi_2 x \cdot \sim (\Phi_3 x \cdot \Phi_4 x) \supset \sim Qx)$

(2a′) $(x)\,(\Phi_3 x \cdot \Phi_4 x \cdot \sim (\Phi_1 x \cdot \sim \Phi_2 x) \supset Qx)$

(2b′) $(x)\,(\Phi_3 x \cdot \sim \Phi_4 x \cdot \sim (\Phi_1 x \cdot \Phi_2 x) \supset \sim Qx)$.

Wypowiedzi te nie pociągają, w przeciwieństwie do zdań (1)-(2), żadnych syntetycznych konsekwencji i zaopatrują termin Q w te same, co zdania (1)-(2), kryteria stosowalności. W szczególności, w odróżnieniu od propozycji Mehlberga, wypowiedzi (1a′)-(2b′) pozwalają na stosowanie terminu Q do przedmiotów spełniających zarazem warunek Φ_1 i Φ_3, przy czym w stosunku do tych przedmiotów mają walor zarówno kryteria definicyjne pierwszego jak i drugiego zdania redukcyjnego, co z kolei odróżnia rozwiązanie obecne od rozwiązania proponowanego przeze mnie poprzednio. Jedynie te spośród przedmiotów spełniających warunki Φ_1 i Φ_3, które są $\Phi_2 \cdot \sim \Phi_4$ lub $\sim \Phi_2 \cdot \Phi_4$, pominięte zostały w wypowiedziach (1a′)-(2b′). Ale istnienie takich przedmiotów wyłączone zostało przez zdania redukcyjne (1)-(2). Wszak implikują one twierdzenie:

(3) $(x)\,(\Phi_1 x \cdot \Phi_3 x \supset (\Phi_2 x \equiv \Phi_4 x))$,

lub, co na jedno wychodzi, twierdzenia:

(3a) $(x) \sim (\Phi_1 x \cdot \Phi_2 x \cdot \Phi_3 x \cdot \sim \Phi_4 x)$

(3b) $(x) \sim (\Phi_1 x \cdot \sim \Phi_2 x \cdot \Phi_3 x \cdot \Phi_4 x)$.

Twierdzenia te dołączone do zdań (1a′)-(2b′) dają w rezultacie całość równoważną logicznie początkowym zdaniom redukcyjnym (1)-(2). Całość ta, podobnie jak w przypadkach poprzednich, jest koniunkcją dwóch różnych co do swego logicznego charakteru czynników: twierdzenia definicyjnego o charakterze analitycznym (1a′) · (1b′) · (2a′) · (2b′) oraz twierdzenia rzeczowego o charakterze syntetycznym (3a) · (3b).

Tak więc, we wszystkich omawianych przypadkach analiza nasza prowadzi do stwierdzenia niejednorodnego charakteru definicji cząstkowych i pozwala na adekwatne, jak się wydaje, wyodrębnienie ich różnych pod względem logicznym czynników. W świetle jej wyników, istnienie definicji cząstkowych wydaje się nie naruszać w niczym tradycyjnego i doniosłego z wielu względów podziału ogółu zdań na dwie rozłączne i wyczerpujące klasy: zdań analitycznych i syntetycznych. Wynik ten nie ma decydującego znaczenia dla dalszych rozważań, w których omawiając kryteria stosowalności terminów teoretycznych abstrahujemy od tego, czy kryteria te istotnie pełnią rolę twierdzeń definicyjnych, a więc twierdzeń uważanych za zdania analityczne. Świadczy on jednak o tym, że kryteria te, w odpowiednim sformułowaniu, do roli takiej się nadają. Sam ich cząstkowy charakter sprawy tej zatem nie przesądza.

3. Rozważania dotychczasowe poświęcone były jednej tylko klasie terminów teoretycznych: terminom cząstkowo definiowalnym. Rozpatrywaliśmy sytuacje takie, w których teoria jakaś wyposaża terminy teoretyczne w kryteria stosowalności o postaci definicji cząstkowych. Poddaliśmy szczegółowej analizie problem sensowności takich terminów oraz problem logicznego charakteru takich definicji. Nie rozpatrywaliśmy bliżej terminów teoretycznych definiowalnych *explicite* za pomocą terminów elementarnych, gdyż terminy takie nie nasuwają specjalnych problemów metodologicznych. Ale nie rozpatrywaliśmy również do tej pory takich terminów teoretycznych, które nie są definiowalne przez terminy elementarne ani całkowicie, ani częściowo. A te terminy przedstawiają interesujący problem metodologiczny. Postulaty teorii nie implikują w stosunku do terminu teoretycznego tego rodzaju żadnych kryteriów stosowalności sformułowanych w języku terminów elementarnych i mających postać definicji zupełnej lub cząstkowej. Jeżeli jednak termin taki posiadać ma jakiś sens empiryczny umożliwiający jego stosowanie do badanych przedmiotów, postulaty teorii jakieś kryteria jego stosowalności implikować muszą. Będą to jednak kryteria ustanawiające luźniejsze jeszcze związki z terminami elementarnymi, niż kryteria o postaci cząstkowej. Powstaje pytanie, o jakie związki tu chodzi. Jaką postać przybierają kryteria definicyjne, które te związki wyrażają? W szczególności, czym różnią się one od omówionych przez nas definicji cząstkowych? Tym zagadnieniom poświęcić chciałbym z kolei parę uwag.

Zwracano przede wszystkim uwagę na to[17], iż związek pomiędzy terminami teoretycznymi a terminami spostrzeżeniowymi ma często charakter probabilistyczny. Wyniki obserwacji pozwalają z pewnym prawdopodobieństwem tylko wnosić o tym, iż termin teoretyczny stosuje się do badanego przedmiotu. Tak ma być np. w przypadku pewnych wielkości fizykalnych, gdzie wynik pomiaru nie stanowi nigdy absolutnie pewnego kryterium przysługiwania danej wielkości w dokładnie określonym stopniu przedmiotowi mierzonemu. Tak ma być również w wypadku pojęć psychologicznych, których związek z obserwowalnymi objawami nie ma nigdy charakteru bezwyjątkowego. W cytowanej wyżej pracy [21] wskazywałem na to, iż analizowane tam pojęcie genotypu również należy zaliczyć do pojęć tego rodzaju. Jego związek ze spostrzegalnymi cechami organizmów jest związkiem natury statystycznej. We wszystkich tych przypadkach kryteria stosowalności terminu teoretycznego sformułowane w języku elementarnym stwierdzać muszą również związki probabilistyczne. Nie mogą to być

17) Por. np. Kaplan [12], Mehlberg [13] i [14], Carnap [6], Hempel [9].

zatem definicje cząstkowe, lecz co najwyżej pewne ich modyfikacje, zwane probabilistycznymi definicjami cząstkowymi. Definicje takie opisane zostały w sposób systematyczny przez Mehlberga [13]. W pracy [19] przedstawiłem ten typ definicji oraz pewne wiążące się z nimi zagadnienia. Tutaj więc poprzestanę na krótkiej informacji.

Probabilistyczną definicję cząstkową terminu Q przedstawić można w postaci następujących dwóch wypowiedzi:

(1) $$P(\Phi, Q) = p$$

(2) $$P(\Psi, \sim Q) = q,$$

z których pierwsza stwierdza, iż prawdopodobieństwo Q ze względu na Φ równa się p, druga — iż prawdopodobieństwo $\sim Q$ ze względu na Ψ równa się q, gdzie p i q są liczbami rzeczywistymi większymi od zera i niewiększymi od jedności. Prawdopodobieństwo, o którym mowa, Mehlberg interpretuje w sposób częstościowy. Przy tej interpretacji zdanie (1) głosi, iż częstość względna klasy Q w stosunku do klasy Φ równa jest p, zdanie (2) — iż częstość względna klasy $\sim Q$ w stosunku do klasy Ψ równa jest q. W przypadku granicznym, gdy $p = q = 1$, probabilistyczna definicja cząstkowa przechodzi w omówioną przez nas definicję cząstkową. Przykłady podawane jako ilustracja tej procedury wydają się wskazywać na to, iż probabilistyczne definicje cząstkowe spotykane w aktualnej praktyce naukowej stwierdzają z reguły częstości bliskie jedności. Formułują one zatem kryteria stosowalności, które o prawie każdym przedmiocie spełniającym warunek Φ pozwalają orzec termin Q, a o prawie każdym przedmiocie spełniającym warunek Ψ — negację terminu Q.

Probabilistyczne definicje cząstkowe nasuwają podobne problemy logiczne, co zwykłe definicje cząstkowe. Podobne też nasuwają się tutaj rozstrzygnięcia. W szczególności — analogiczne modyfikacje zmierzające do zapewnienia im charakteru zdań analitycznych. W związku z tą ostatnią sprawą warto może wspomnieć o pewnym zarzucie skierowanym pod adresem tych definicji traktowanych jako pewne probabilistyczne postulaty znaczeniowe.[18] Chodzi mianowicie o to, jaki jest sens występującego w nich pojęcia prawdopodobieństwa, żadne bowiem z obiegowych znaczeń tego terminu nie wydaje się w tym przypadku odpowiednie. Dotyczyć to ma również prawdopodobieństwa częstościowego. Twierdzenie:

$$P(\Phi, Q) = p$$

ma mieć bowiem określony sens tylko wtedy, gdy potrafimy niezależnie od stwierdzenia, czy przedmiot a spełnia warunek Φ, rozstrzygnąć, czy należy on do klasy Q. A to właśnie w przypadku definicji probabilistycznej jest niemożliwe. Zarzut ten nie wydaje się jednak słuszny. Pojęcie prawdopodobieństwa — również prawdopodobieństwa częstościowego — budzi szereg znanych wątpliwości, których nie można zlekceważyć i w przypadku definicji probabilistycznych. Nie wydaje się jednak, aby powstawały tu jakieś trudności dodatkowe. Fakt, na który się wspomniany zarzut powołuje, wydaje się świadczyć tylko o tym, że twierdzenie:

18) Por. Pap [16].

$$P(\Phi, Q) = p$$

nie ma charakteru empirycznego. Jest to twierdzenie analityczne, oparte na konwencji terminologicznej nakazującej używanie terminu Q w sensie takim, aby twierdzenie to było prawdziwe. Używamy go zaś tak wtedy, gdy — mówiąc w uproszczeniu — zaliczamy do klasy Q p-tą część przedmiotów spełniających warunek Φ. Sytuacja wydaje się tu analogiczna do sytuacji zwykłych definicji cząstkowych:

$$(x)\,(\Phi x \supset Qx).$$

I tu można by powoływać się na fakt, iż nie potrafimy o żadnym przedmiocie rozstrzygnąć, czy należy do klasy Q niezależnie od stwierdzenia, czy spełnia on warunek Φ. I tu fakt ten wydaje się świadczyć jedynie o analitycznym charakterze powyższej wypowiedzi. Tyle tylko, że konwencja, na której opiera się ta ostatnia, nakazuje zaliczanie do klasy Q wszystkich przedmiotów spełniających warunek Φ.

Inna rzecz, czy probabilistyczne postulaty znaczeniowe znajdują istotnie tak szerokie zastosowanie, jak to sugerują niektórzy. Wydaje się, iż pewne przynajmniej spośród podawanych przykładów dopuszczają również interpretację inną. Idzie tu o sytuacje następujące. Przypuśćmy, iż termin Q wprowadzony został przez zwykłą definicję cząstkową zawierającą wypowiedź:

(1) $$(x)\,(\Phi x \supset Qx).$$

Posługując się tym częściowo zdefiniowanym terminem odkryto szereg praw, które — rozszerzone konwencjonalnie na dziedziny, w których termin Q pozostawał do tej pory nieokreślony — zaopatrzyły ten termin w nowe kryteria stosowalności i nowe, oparte na nich, definicje cząstkowe. Okazało się jednak następnie, iż nowe kryteria kolidują w pewnych przypadkach z kryteriami pierwotnymi, prowadząc np. do wniosku, iż:

$$(\exists x)\,(\Phi x \cdot \sim Qx).$$

Któreś z nich zatem wymagają rewizji. W przypadku, gdy kryteria późniejsze mają teoretyczną wyższość nad kryteriami pierwotnymi (np. szerszy zakres stosowalności), modyfikacji poddane zostają raczej te ostatnie. W pewnych wypadkach powstałą niezgodność usunąć można zastępując bezwyjątkowe kryteria pierwotne kryteriami probabilistycznymi, a więc np. wypowiedź (1) sformułowaniem następującym:

(1′) $$P(\Phi, Q) \approx 1.$$

Ma to być jedna z dróg, jakimi do języka naukowego dostają się definicje probabilistyczne.

Czy nie nasuwa się jednak odmienna nieco interpretacja takich sytuacji? Wydaje się, że wobec powstałej niezgodności, z pierwotnych kryteriów definicyjnych terminu Q po prostu rezygnujemy na korzyść kryteriów późniejszych. Nie wprowadza się tym samym żadnych postulatów znaczeniowych o charakterze probabilistycznym. Zamiast nich stwierdza się jedynie pewne zależności statystyczne ograniczone do dziedziny, w której termin Q ma ustalony przez owe późniejsze kryteria sens. Zależności te mają wskutek tego charakter twierdzeń empirycznych. Tak przynajmniej można by potraktować przykład pojęcia temperatury, stanowiący klasyczny wzorzec opisanej przed chwilą sytuacji. Jego pierwotna definicja cząstkowa głosi: jeśli termometr stykający się z ciałem x wskazuje cyfrę t, to ciało x ma temperaturę t. Dalsze jednak kryteria definicyjne prowadzą na gruncie praw termodynamiki do wniosku, iż w pewnym, niewielkim zresztą, procencie przypadków zależność powyższa nie zachodzi. Możemy wówczas, obstając przy kryterium pierwotnym jako kryterium definicyjnym, sformułować je jako definicję probabilistyczną. Możemy jednak wybrać drogę inną: zrezygnować z kryterium pierwotnego jako twierdzenia ustalającego sens terminu „temperatura". Sens ten byłby wyznaczony tylko przez owe kryteria późniejsze. Biorąc termin „temperatura" w tym właśnie znaczeniu, dochodzilibyśmy na drodze doświadczalnej do stwierdzenia pewnej empirycznej zależności statystycznej pomiędzy wskazaniami termometru a temperaturą ciała. Mam wrażenie, iż w faktycznym postępowaniu naukowym taki właśnie, empiryczny czysto, charakter przypisujemy na obecnym etapie rozwoju fizyki tej zależności. W ten sposób postulaty znaczeniowe o charakterze probabilistycznym mogłyby być ograniczone do tych tylko terminów, dla których w ogóle nie istnieją kryteria stosowalności o charakterze bezwyjątkowym. Stanowiłoby to znaczne zredukowanie zakresu zastosowań probabilistycznych definicji cząstkowych, choć nie eliminowałoby ich całkowicie. Wydaje się bowiem, iż istotnie istnieją terminy teoretyczne, dla których jedynymi adekwatnymi kryteriami stosowalności sformułowanymi w języku terminów spostrzeżeniowych są kryteria probabilistyczne.

Definicje probabilistyczne reprezentują jeden z kierunków rozszerzenia pojęcia definicji cząstkowej. Pojęcie to jednak dopuszcza i inne modyfikacje, prowadzące, podobnie jak i poprzednia, do luźniejszych, niż definicja cząstkowa, kryteriów stosowalności. Dokonując — niewyczerpującego zresztą i ogólnikowego — przeglądu takich kryteriów, abstrahować będziemy od tego, czy wszystkie one znajdują faktyczne zastosowania w praktyce naukowej. Kwestię tę pozostawiamy otwartą. Jej rozstrzygnięcie wymaga szczegółowych badań empirycznych wykraczających poza zadania obecnej pracy.

Definicja cząstkowa terminu Q formułuje kryteria stosowalności zarówno dla terminu Q jak i dla jego negacji. Ogólniejszą procedurę możemy otrzymać podając tylko jedno z tych kryteriów: bądź dla terminu Q, bądź dla jego negacji. A zatem jedynym kryterium stosowalności terminu Q na gruncie pewnej teorii może być np. wypowiedź następująca:

$$(x)\,(\Phi x \supset Qx).$$

Określa ona znaczenie terminu Q w sposób jeszcze bardziej niezupełny niż definicja cząstkowa. I tutaj jednak znaczenie to nie pozostaje całkowicie niezdeterminowane. O żadnym, co prawda, przedmiocie nie jesteśmy w stanie stwierdzić na podstawie doświadczenia, iż nie podpada pod termin Q. Pewne jednak przedmioty, spełniające sformułowany w terminach spostrzeżeniowych warunek Φ, możemy do zakresu terminu Q zaliczyć. W przypadku wypowiedzi formułującej negatywne kryterium stosowalności terminu Q:

$$(x)\,(\Psi x \supset \sim Qx)$$

mamy do czynienia z sytuacją odwrotną.

Wszystkie rozważane do tej pory wypowiedzi definicyjne, a więc zarówno definicja zupełna, jak i definicja cząstkowa wraz z jej luźniejszymi odmianami: probabilistyczną i, wspomnianą ostatnio, «jednostronną» definicją cząstkową — formułowały takie kryteria stosowalności terminu teoretycznego Q, które pozwalały nam o pewnych przynajmniej przedmiotach orzekać na podstawie doświadczenia — stanowczo lub z określonym prawdopodobieństwem — bądź sam termin Q, bądź jego negację. Mogą jednakże istnieć na gruncie teorii empirycznych takie kryteria stosowalności terminów teoretycznych, które powyższego warunku nie spełniają. Wypowiedzi formułujące te kryteria przybierać muszą zatem postać różną od wypowiedzi definicyjnych uwzględnionych przez nas do tej pory. Pewien typ takich wypowiedzi, traktowany jako uogólnienie definicji cząstkowych, scharakteryzowany został przez Stopes-Roe [23]. Uogólniona definicja cząstkowa różni się od zwykłej definicji cząstkowej tym, iż podaje kryteria stosowalności nie dla jednego terminu teoretycznego, lecz dla szeregu takich terminów jednocześnie. Zilustrujemy to na przykładzie uogólnionej definicji cząstkowej wprowadzającej dwa terminy teoretyczne: Q_1 i Q_2. Definicja taka przybierać może postać jednej z czterech następujących wypowiedzi, ewentualnie pewnej ich koniunkcji:

(1) $$(x)\,(\Phi_1 x \supset Q_1 x \vee Q_2 x)$$

(2) $$(x)\,(\Phi_2 x \supset Q_1 x \vee {\sim}\, Q_2 x)$$

(3) $$(x)\,(\Phi_3 x \supset {\sim}\, Q_1 x \vee Q_2 x)$$

(4) $$(x)\,(\Phi_4 x \supset {\sim}\, Q_1 x \vee {\sim}\, Q_2 x).$$

Wypowiedzi te podają kryteria stosowalności nie dla poszczególnych terminów teoretycznych, lecz dla pewnych ich funkcji. Ich następniki nie są zdaniami «atomowymi» (ew. ich negacjami), lecz «molekularnymi». Ograniczenie do alternatywy zdań atomowych i ich negacji jest tutaj nieistotne. Każde zdanie molekularne można, jak wiadomo, przedstawić w normalnej postaci koniunkcyjnej jako koniunkcję takich alternatyw. Każdą więc wypowiedź podającą kryteria stosowalności dla jakiegoś zdania molekularnego przedstawić można jako koniunkcję wypowiedzi definicyjnych postaci (1)-(4). Tak np. wypowiedź formułującą warunki prawdziwości dla pewnej równoważności:

$$(x)\,(\Phi x \supset Q_1 x \equiv Q_2 x)$$

sformułować można jako koniunkcję wypowiedzi:

$$(x)\,(\Phi x \supset {\sim}\, Q_1 x \vee Q_2 x)$$
$$(x)\,(\Phi x \supset Q_1 x \vee {\sim}\, Q_2 x),$$

podpadających pod opisane schematy. Łatwo przewidzieć, jak wyglądać będą wypowiedzi formułujące kryteria stosowalności dla większej liczby terminów. W przypadku trzech terminów wypowiedzi te przybierać mogą jedną z ośmiu form, wśród których znajdzie się np. taka:

$$(x)\,(\Phi x \supset Q_1 x \vee Q_2 x \vee {\sim}\, Q_3 x).$$

Powróćmy jednak do uogólnionej definicji cząstkowej dwóch terminów teoretycznych. Wypowiedzi typu (1)-(4) nie pozwalają na ogół rozstrzygnąć o żadnym przedmiocie, czy podpada on pod termin Q_1 (ew. Q_2). Mimo to ustanawiają pewne związki pomiędzy tymi terminami a terminami elementarnymi, determinując w jakimś stopniu sens empiryczny terminów Q_1 i Q_2. Jeśli wiemy np., iż przedmiot a spełnia warunek Φ_1, możemy na podstawie wypowiedzi (1) wnosić o tym, iż jest on Q_1 lub Q_2. W pewnych przypadkach szczególnych uogólniona definicja cząstkowa terminów Q_1 i Q_2 pozwala na zastosowanie wobec pewnych przedmiotów jednego z tych terminów. I tak, jeśli oprócz wypowiedzi (1) przyjęte zostały również wypowiedzi (2) i (4), możemy na ich podstawie o przedmiocie spełniającym jednocześnie warunki Φ_1, Φ_2 i Φ_4 orzec termin Q_1 — chyba, żeby warunki te okazały się wzajem sprzeczne. Podobna sytuacja zaistnieć może wtedy, gdy dla któregoś z terminów Q_1 i Q_2 istnieje oprócz uogólnionej definicji cząstkowej definicja cząstkowa zwykła. Przypuśćmy, iż dla terminu Q_1 ma walor następujące kryterium:

$$(x)\,(\Phi x \supset Q_1 x).$$

Jeśli dla terminów Q_1 i Q_2 przyjęte zostało ponadto kryterium w postaci (3), mamy prawo przedmiotowi, który spełnia warunki Φ i Φ_3 przypisać cechę Q_2 — o ile oczywiście warunki Φ i Φ_3 są wzajem niesprzeczne. Są to jednak sytuacje szczególne, nie zmieniające faktu istotnego rozszerzenia zakresu kryteriów stosowalności przez dopuszczenie w tym charakterze uogólnionych definicji cząstkowych.

Czy mamy tu do czynienia z koncepcją dostatecznie szeroką, aby podpadały pod nią wszelkie rodzaje twierdzeń wyrażających związki zachodzące pomiędzy terminami teoretycznymi a terminami elementarnymi na gruncie teorii empirycznych? Na pewno nie. Nawet jeśli ograniczymy się do najprostszych formalnie systemów opartych na rachunku predykatów pierwszego rzędu, znajdziemy bez trudu takie twierdzenia formułujące kryteria stosowalności terminów teoretycznych, które nie mieszczą się w powyższej koncepcji. Należą do nich twierdzenia zawierające terminy teoretyczne «kontrolowane» przez kwantyfikatory szczegółowe. Nie wdając się w ogólną charakterystykę tego typu twierdzeń, podaną przez Stopes-Roe [23], przytoczę najprostsze przykłady takich wypowiedzi. Mogą być nimi np. twierdzenia następujące:

$$(x)\,(\Phi x \supset (\exists y)\, Qy)$$
$$(x)\,(\Psi x \supset (\exists y)\, R(x, y)).$$

Nie dają się one wtłoczyć w schemat zwykłych czy uogólnionych definicji cząstkowych. Mimo iż użyteczność tego rodzaju kryteriów wydaje się dość problematyczna, nie można wyłączyć ich z góry z grona możliwych kryteriów stosowalności terminów teoretycznych.

Dotychczasowe rozważania obracały się wyłącznie w kręgu teorii, których aparat formalny mieścił się w obrębie węższego rachunku funkcyjnego. Nie rozważyliśmy nawet tych komplikacji, które pociąga za sobą wprowadzenie wyrażeń funkcyjnych czy operatora deskrypcyjnego. Jest to aparat formalny zbyt ubogi dla sformułowania większości teorii przyrodniczych. Środki logiczne, które teorie te angażują, wykraczają z reguły poza węższy rachunek funkcyjny. Wzbogacają się też tym samym rodzaje związków, jakie na gruncie tych teorii zachodzić mogą pomiędzy terminami teoretycznymi a terminami elementarnymi. Jest rzeczą oczywistą, iż nie dają się one wyrazić za pomocą rozważanych przez nas do tej pory twierdzeń. Wobec wielkiej rozmaitości, jaką odznaczać się mogą takie sformułowania, beznadziejna wydaje się próba uchwycenia ich w jakieś ogólne schematy. Nasuwa się jednak inny sposób ujęcia istoty tych powiązań.

Jeśli dany system ma mieć charakter teorii empirycznej, jego terminy teoretyczne winny posiadać empiryczny sens. Otóż związki zachodzące pomiędzy nimi a terminami elementarnymi można scharakteryzować ogólnie jako związki zapewniające terminom teoretycznym empiryczny sens. Aby charakterystyka taka posiadała jakąś wartość informacyjną, musimy rozporządzać dostatecznie precyzyjnym pojęciem empirycznego sensu. Termin ten jest jednym z najczęściej używanych i zarazem najróżniej interpretowanych terminów metodologii współczesnej. Jego najbardziej, moim zdaniem, adekwatną a jednocześnie precyzyjną charakterystykę podaje, w swej ostatniej pracy [6], Carnap. Zasadniczą myśl jego propozycji wyrazić można jak następuje:

Termin jakiś ma sens empiryczny, jeżeli pewne jego zastosowanie pociąga za sobą obserwowalne konsekwencje.

Nie będę przytaczał tutaj dokładnego sformułowania Carnapowskiej definicji. Podam tylko — i to w pewnym uproszczeniu — konsekwencje tej definicji, charakteryzujące klasę terminów empirycznie sensownych.

Pojęcie terminu empirycznie sensownego jest tutaj zrelatywizowane do określonej teorii T. Przyjmijmy następujące oznaczenia. Niech Q będzie, jak dotąd, terminem teoretycznym, Z_Q — zdaniem zawierającym Q jako jedyny termin pozalogiczny, a Z_S — zdaniem elementarnym, zawierającym jako terminy pozalogiczne wyłącznie terminy elementarne (zdaniem języka «spostrzeżeniowego»). Termin Q uważać będziemy za empirycznie sensowny na gruncie teorii T, jeżeli istnieją zdania Z_Q i Z_S takie, iż ze zdania Z_Q i postulatów teorii T wynika logicznie zdanie Z_S , natomiast z samych postulatów teorii T zdanie Z_S nie wynika. Definicja Carnapa nie ogranicza jednak klasy terminów empirycznie sensownych do terminów spełniających powyższy warunek. Załóżmy, iż K jest pewną klasą terminów teoretycznych empirycznie sensownych na gruncie teorii T Niech Z_K będzie zdaniem zawierającym jako terminy pozalogiczne wyłącznie terminy klasy K. W tej sytuacji termin Q uważać będziemy za empirycznie sensowny na gruncie teorii T, jeżeli istnieją zdania Z_Q, Z_K i Z_S takie, iż ze zdań Z_Q, Z_K i postulatów teorii T wynika logicznie zdanie Z_S, natomiast ze zdania Z_K i postulatów teorii T zdanie Z_S nie wynika.[19] To szerokie kryterium empirycznej sensowności terminów teoretycznych rozszerzyć można jeszcze bardziej dopuszczając oprócz wynikania logicznego wynikanie „probabilistyczne". Warunki powyższe stwierdzałyby wówczas, iż prawdopodobieństwo zdania Z_S ze względu na zdanie Z_Q i postulaty T (ew. zdania Z_Q, Z_K i postulaty T) jest różne od prawdopodobieństwa zdania Z_S ze względu na same postulaty T (ew. zdanie Z_K i postulaty T).

19) Jest to uproszczenie Carnapowskiego kryterium, pomijające pewne niezbędne zastrzeżenia.

Definicję Carnapa uważać można za końcowy etap ewolucji pojęcia empirycznej sensowności w kierunku coraz to większego liberalizmu. Kryterium proponowane przez Carnapa jest niezmiernie tolerancyjne. Autor pokazuje jednak w sposób przekonujący, iż tolerancja ta nie idzie zbyt daleko. Kryterium to nie pozwala nam uznać za empirycznie sensowne wyrażeń, które uważa się zgodnie za wyrażenia metafizyczne i którym odmawia się z tej racji prawa obywatelstwa w nauce. Z drugiej strony kryterium to nie jest zbyt rygorystyczne. Przyznaje empiryczną sensowność m. in. takim terminom, którym kryteria wcześniejsze tej cechy odmawiały, a które należą niewątpliwie do współczesnego języka naukowego. W szczególności okazuje się, iż wszystkie omawiane przez nas rodzaje kryteriów stosowalności terminów teoretycznych z reguły zapewniają tym terminom tak rozumiany sens empiryczny.

Jest to oczywiste w przypadku zwykłej definicji cząstkowej — a przy wspomnianym rozszerzeniu Carnapowskiego kryterium — również w przypadku probabilistycznej definicji cząstkowej. Podobnie rzecz się ma, gdy jedynym kryterium stosowalności terminu teoretycznego jest «jednostronna» definicja cząstkowa, np. postaci następującej:

(1) $(x)\,(\Phi x \supset Q_1 x).$

Wszak istnieje wówczas zdanie typu Z_{Q_1}: $\sim Q_1 a$, które pociąga za sobą na gruncie postulatów obejmujących definicję (1) zdanie typu Z_S: $\sim \Phi a$, nie wynikające z samych postulatów.[20]

A jak przedstawia się sprawa empirycznej sensowności terminu teoretycznego, dla którego jedyne kryterium stosowalności stanowi uogólniona definicja cząstkowa? Załóżmy, iż dla terminu Q_2 definicja taka ma postać następującą:

(2) $(x)\,(\Phi_1 x \supset Q_1 x \vee Q_2 x).$

Nie istnieje w tym przypadku zdanie typu Z_{Q_2}, które na gruncie powyższej definicji pociągałoby za sobą zdanie typu Z_S. Dopiero w połączeniu ze zdaniem typu Z_{Q_1}: $\sim Q_1 a$ zdanie typu Z_{Q_2}: $\sim Q_2 a$ implikuje zdanie elementarne $\sim \Phi_1 a$. A zatem, zgodnie z definicją Carnapa, termin Q_2 może mieć sens empiryczny tylko wtedy, gdy termin Q_1 już taki sens posiada. Przypuśćmy, że tak jest i że dla terminu Q_1 ma walor przytoczona wyżej definicja (1). Wówczas uogólniona definicja cząstkowa (2) zapewnia sens empiryczny również terminowi Q_2. Zdanie elementarne: $\sim \Phi_1 a$ wynikające z koniunkcji: $\sim Q_1 a \cdot \sim Q_2 a$ nie wynika bowiem z samego zdania: $\sim Q_1 a$ na gruncie postulatów obejmujących definicje (1) i (2). Analogiczna sytuacja zachodzi niezależnie od tego, jaką postać ma uogólniona definicja cząstkowa formułująca kryteria stosowalności dla ter-

20) Zdanie: $\sim \Phi a$ nie wynika oczywiście z samej definicji (1). Możemy przyjąć, iż nie wynika ono również z pozostałych postulatów danej teorii, gdyż postulaty takie są z reguły twierdzeniami uniwersalnymi, nie zawierającymi nazw indywidualnych i nie pociągającymi żadnych jednostkowych konsekwencji. Jednocześnie trzeba dodać, iż zdanie: $\sim Q_1 a$ możemy traktować jako zdanie typu Z_{Q_1} ze względu na to, iż istnieją języki (w rodzaju opisanego przez Carnapa języka «koordynatów»), w których nazwy indywidualne nie należą do terminów deskryptywnych.

minów Q_1 i Q_2, jak również niezależnie od tego, jakie kryteria stosowalności mają walor dla jednego z tych terminów. Jeżeli tylko zapewniają mu sens empiryczny, termin pozostały zostaje z reguły również w taki sens wyposażony. Może być przy tym, jak widzieliśmy, tak, iż wszystkie owe kryteria łącznie implikują dla terminu Q_2 kryterium o postaci definicji cząstkowej (zwykłej lub «jednostronnej»). Wówczas przypadek taki podpada pod schemat omawiany poprzednio. Na ogół jednakże taki stan rzeczy nie zachodzi. Mamy wtedy do czynienia z terminem empirycznie sensownym nie definiowalnym nawet cząstkowo. Tak właśnie jest w przytaczanym przykładzie. Termin Q_2 posiada sens empiryczny, mimo iż o żadnym przedmiocie nie potrafimy na podstawie doświadczenia orzec ani terminu Q_2 ani jego negacji.[21]

Carnapowskie kryterium empirycznej sensowności spełniają również terminy, których kryteria stosowalności nie dają się wtłoczyć w żaden z omawianych schematów. Tak np. wspomniane poprzednio twierdzenia o «kontrolowanych» przez kwantyfikator szczegółowy terminach teoretycznych również mogą terminy te zaopatrywać w empiryczny sens. Weźmy dla przykładu jedno z cytowanych wyżej twierdzeń:

$$(x)\,(\Psi x \supset (\exists y)\, R(x, y)).$$

Ze zdania zawierającego termin teoretyczny R jako jedyny termin pozalogiczny: $\sim (\exists y)\, R(a, y))$ wynika na gruncie tego postulatu zdanie elementarne: $\sim \Psi a$. Ponieważ nie wynika ono z samego tego postulatu, termin R jest terminem empirycznie sensownym. Rozmaitość twierdzeń ustanawiających takie związki pomiędzy terminami teoretycznymi a terminami elementarnymi, które gwarantują terminom teoretycznym posiadanie empirycznego sensu, wymyka się próbom schematyzacji. Zwłaszcza iż mogą to być twierdzenia angażujące środki logiczne wykraczające daleko poza węższy rachunek funkcyjny. Jedno można stwierdzić na pewno: muszą to być twierdzenia, w których przynajmniej jeden termin teoretyczny i jeden termin elementarny występują w sposób istotny, a więc twierdzenia, które nie są logicznie równoważne wyrażeniom pozbawionym któregoś z tych terminów. Jest to warunek niezbędny do tego, aby ustanawiały one jakiekolwiek związki pomiędzy tymi dwoma rodzajami terminów. Oczywiście nie wszystkie twierdzenia zapewniające empiryczny sens terminom teoretycznym mogą być traktowane na równi. Twierdzenia takie determinują sens terminów teoretycznych w stopniu różnym. Prowadzą zatem do terminów różniących się znacznie swoją rolą w procesie wyjaśniania i przewidywania danych nam w doświadczeniu zjawisk. Wydaje się, iż omówione przez nas rodzaje twierdzeń zasługują z tego punktu widzenia na specjalną uwagę. Definicja cząstkowa w jej różnych odmianach reprezen-

21) Przykład ten świadczy, iż interpretacja kryterium Carnapa, jaką w swej książce [14] podaje Mehlberg, jest nietrafna. Nie jest prawdą, iż klasa terminów empirycznie sensownych pokrywa się, jak twierdzi Mehlberg, z klasą terminów, dla których istnieją kryteria w postaci definicji zupełnych lub cząstkowych. Termin Q_2 należy do pierwszej, a nie należy do drugiej. Jest to termin empirycznie sensowny, a jednocześnie całkowicie nieostry, wg terminologii Mehlberga. Autor przyznaje co prawda, iż jego interpretacja ma charakter przybliżony. Wydaje się jednak, iż różnice, jakie tu zachodzą, są zbyt duże, aby interpretacja ta mogła być uważana za trafne przybliżenie.

tuje najczęstszy chyba i najdonioślejszy typ twierdzeń formułujących empiryczne kryteria stosowalności terminów teoretycznych.

II

4. Pytanie, czy termin teoretyczny Q jest definiowalny na gruncie teorii T przez terminy elementarne, jest niewątpliwie pytaniem wieloznacznym. Na wstępie naszych rozważań wyróżniliśmy dwie interpretacje takiego sformułowania. Może w nim przede wszystkim chodzić o to, czy wśród postulatów teorii T istnieje postulat będący definicją terminu Q w języku terminów elementarnych. W pytaniu tym jednak może chodzić również o coś innego, mianowicie o to, czy z postulatów teorii T wynika twierdzenie mające postać definicji terminu Q w języku terminów elementarnych. To właśnie pojęcie definiowalności było przedmiotem naszych dotychczasowych rozważań. Jest to pojęcie szersze od poprzedniego. Jeśli termin jakiś jest definiowalny w sensie poprzednim, jest on tym samym definiowalny w sensie obecnym, lecz nie na odwrót. Definiowalność jednak pojmowana być może jeszcze inaczej. W *Scientific Explanation* Braithwaite wprowadza pojęcie definiowalności terminów teoretycznych przez terminy elementarne różne od obu pojęć poprzednich i od obu z nich szersze. Tak rozumianej definiowalności terminów teoretycznych poświęcimy dalsze rozważania. Pojęcie to wyjaśnimy na przykładzie prostej, omawianej już przez nas na wstępie, teorii.

Teoria ta — oznaczmy ją przez T_1 — składa się z trzech postulatów o postaci następującej:

(P1) $(x)\,(Ax \equiv Lx \cdot Mx)$

(P2) $(x)\,(Bx \equiv Mx \cdot Nx)$

(P3) $(x)\,(Cx \equiv Nx \cdot Lx)$.

Występujące w nich terminy pozalogiczne podzielić można na dwie klasy. Terminy: A, B, C — to terminy elementarne. Każdy z nich, na mocy przyjętej interpretacji, odnosi się do określonej obserwowalnej własności. Terminy: L, M, N — to terminy teoretyczne, dla których żadna bezpośrednia interpretacja nie istnieje. Z postulatów P wynikają logicznie następujące twierdzenia sformułowane wyłącznie w języku terminów elementarnych:

(T1) $(x)\,(Ax \cdot Bx \supset Cx)$

(T2) $(x)\,(Bx \cdot Cx \supset Ax)$

(T3) $(x)\,(Cx \cdot Ax \supset Bx)$.

Teorię T_1 traktować można jako teorię wyjaśniającą owe doświadczalne uogólnienia. Mimo swej niezmiernej prostoty teoria T_1 służyć może jako przykład pewnego charakterystycznego dla nauk przyrodniczych typu teorii. Teorie takie wyjaśniają pewne obserwowalne zjawiska postulując istnienie ograniczonej liczby nieobserwowalnych «czynników» (genów, atomów, cząstek elementarnych), których określone kombinacje odpowiadać mają owym obserwowalnym fenomenom.

Terminy teoretyczne: L, M, N nie są na gruncie teorii T_1 definiowalne przez terminy elementarne: A, B, C ani w pierwszym, ani w drugim z wyróżnionych przez nas znaczeń. Żaden z postulatów P nie jest definicją któregoś z terminów *L*, *M*, *N* za pomocą terminów: *A*, *B*, *C*. Co więcej, z postulatów P nie wynika, jak się łatwo przekonać, żadne twierdzenie, które by miało postać takiej definicji. Czy nie można jednak mówić o definiowalności tych terminów w innym, słabszym nieco sensie? Skoro teoria T_1 definicji takich nie implikuje, spróbujmy je skonstruować niezależnie od niej. Sformułujmy pewne twierdzenia mające postać definicji terminów: *L*, *M*, *N* za pomocą terminów *A*, *B*, *C* i dołączmy je do postulatów P. Przejdziemy w ten sposób od teorii T_1 do bogatszej teorii T_1^*. Może się jednak okazać, iż to wzbogacenie teorii T_1 nie pociąga za sobą wzbogacenia zasobu tych twierdzeń, które sformułowane są wyłącznie za pomocą terminów elementarnych. Teorie T_1 i T_1^* nazwiemy wówczas teoriami „empirycznie równoważnymi", a dołączone twierdzenia — formułami „jałowymi". Jeśli fakt taki ma miejsce, terminy teoretyczne uważać będziemy za definiowalne w sensie obecnym przez terminy elementarne w teorii T_1. A zatem, termin teoretyczny *Q* jest terminem definiowalnym przez terminy elementarne w teorii *T*, jeżeli istnieje takie twierdzenie, które ma postać definicji terminu *Q* w języku terminów elementarnych i które w teorii *T* jest twierdzeniem jałowym. Możemy wówczas powiedzieć, iż teoria *T* co prawda definicji takiej nie narzuca, ale na nią pozwala. Dołączenie jej do postulatów teorii *T* nie zmienia w niczym ogółu elementarnych konsekwencji tych postulatów; a tylko takie konsekwencje uważa się na ogół za doświadczalne, bezpośrednio sprawdzalne twierdzenia teorii.

Jak zatem przedstawia się sytuacja terminów teoretycznych w teorii T_1? Czy są one definiowalne w sensie powyższym przez terminy elementarne? Braithwaite uzasadnia twierdzącą odpowiedź na to pytanie podając twierdzenia, które mają postać definicji terminów: *L*, *M*, *N* w języku terminów: *A*, *B*, *C* i które w teorii T_1 są formułami jałowymi. Definicje te przedstawiają się następująco:

(D1) $(x)\ (Lx \equiv Cx \vee Ax)$

(D2) $(x)\ (Mx \equiv Ax \vee Bx)$

(D3) $(x)\ (Nx \equiv Bx \vee Cx)$.

Są to, jak pokazuje Braithwaite, w zasadzie jedyne twierdzenia, które nadają się do tego celu. Doszliśmy w ten sposób do dość nieoczekiwanej konkluzji. Terminy teoretyczne omawianej teorii okazują się, w pewnym przynajmniej sensie, definiowalne *explicite* przez terminy elementarne. Owe niespostrzegalne «czynniki» potraktowane być mogą jako proste konstrukcje z własności spostrzegalnych. Wydaje się to niezgodne z faktyczną praktyką naukową. W rzeczywistych teoriach, których wzorcem ma być teoria omawiana, nie utożsamia się na ogół owych hipotetycznych «czynników» z pewnymi dostępnymi obserwacji przedmiotami. Nie definiuje się, co za tym idzie, terminów teoretycznych przez terminy elementarne, poprzestając, jak widzieliśmy wyżej, na stwierdzaniu luźniejszych znacznie związków pomiędzy tymi dwoma rodzajami terminów. Braithwaite stara się wyjaśnić ten fakt pokazując na przykładzie omawianej teorii, iż wprowadzenie definicji terminów teoretycznych — i to takich, na które dana teoria «pozwala» — pociąga w rezultacie pewne niepożądane konsekwencje. Jeśli ich chcemy uniknąć, musimy zrezygnować z definiowania terminów teoretycznych przez terminy elementarne. Wywody swe ogranicza autor do skonstruowanej przez siebie teorii T_1. Pokazuje, iż dołączenie do postulatów P definicji D prowadzi do wspomnianych konsekwencji. Niewątpliwie jednak intencje autora idą dalej. Teoria T_1 ma być tylko przykładem obszernej klasy teorii empirycznych i to, co zostało pokazane na tym przykładzie, ma mieć walor w stosunku do wszystkich teorii owej klasy. W pewnych punktach sugeruje nawet autor sposób uogólnienia podawanej argumentacji. Celem dalszych rozważań będzie próba uogólnienia wyników autora. Wywody przez niego podawane chciałbym uniezależnić od specyficznych własności obranej za przykład teorii T_1, formułując je w sposób dostatecznie ogólny na to, aby obejmowały i inne pokrewne teorie.

4.1. Niepożądane konsekwencje wprowadzenia do teorii T_1 definicji D dotyczą zarówno samej teorii T_1 jak i jej dalszych etapów rozwojowych. Rozpatrzmy kolejno te dwa rodzaje konsekwencji. Na czym polegają konsekwencje dotyczące teorii T_1? Z postulatów P teorii T_1 wynikają logicznie twierdzenia doświadczalne T. Wynikanie odwrotne jednak nie zachodzi. W tej sytuacji teorię T_1 uważać możemy za teorię wyjaśniającą owe doświadczalne uogólnienia. Wprowadzenie do teorii T_1 definicji terminów teoretycznych D sytuację tę zasadniczo zmienia. Obecnie nie tylko z postulatów teorii wynikają jej doświadczalne konsekwencje, ale i na odwrót: z ogółu twierdzeń doświadczalnych T wynikają na gruncie definicji D postulaty P. Postulaty te stają się więc po przyjęciu definicji terminów teoretycznych równoważne logicznie ogółowi doświadczalnych konsekwencji teorii. Stanowią zatem nie tyle wyjaśnienie owych doświadczalnych uogólnień, ile po prostu sformułowanie ich w nieco inny, bardziej skomplikowany sposób. Tak zmodyfikowana teoria nie stwierdza nic ponad to, co stwierdzają te doświadczalne uogólnienia, dla których wytłumaczenia została ona przyjęta.

To, że z twierdzeń T wynikają po przyjęciu definicji D postulaty P, udowadnia Braithwaite w sposób mający walor tylko dla teorii T_1. Można to jednak uczynić w sposób, który daje się zastosować do teorii dowolnego rodzaju. Definicje D pozwalają na zastąpienie w postulatach P terminów teoretycznych: L, M, N równoznacznymi wyrażeniami zawierającymi wyłącznie terminy elementarne: A, B, C. Tak więc, z postulatów P i definicji D wynikają logicznie twierdzenia:

(P1*) $(x)\,(Ax \equiv (Cx \vee Ax) \cdot (Ax \vee Bx))$

(P2*) $(x)\,(Bx \equiv (Ax \vee Bx) \cdot (Bx \vee Cx))$

(P3*) $(x)\,(Cx \equiv (Bx \vee Cx) \cdot (Cx \vee Ax))$.

Są to twierdzenia doświadczalne, sformułowane wyłącznie za pomocą terminów elementarnych. Czy wynikają one z twierdzeń doświadczalnych T? Twierdzenia T reprezentują, jak wiemy, ogół doświadczalnych konsekwencji teorii T_1. Gdyby twierdzenia P* z nich nie wynikały, gdyby zatem — swobodnie mówiąc — nie były w nich zawarte, dołączenie do teorii T_1 definicji D pozwalałoby na wyprowadzenie takich twierdzeń doświadczalnych, których przed tym udowodnić się nie dało. A zatem wzbogacona o definicje D teoria T_1^*nie byłaby empirycznie równoważna teorii T_1, a definicje D nie byłyby formułami jałowymi. To jednak sprzeczne jest z naszym założeniem. A więc twierdzenia P* muszą wynikać z twierdzeń T. Z twierdzeń P* wynikają jednak na mocy definicji D postulaty P. Skoro tedy twierdzenia T pociągają za sobą twierdzenia P*, a te ostatnie łącznie z definicjami D — postulaty P, z twierdzeń T wynikają przy założeniu definicji D postulaty P. Istotnie więc wprowadzenie do teorii T_1 definicji terminów teoretycznych sprawia, iż postulaty tej teorii stają się równoważne ogółowi jej doświadczalnych konsekwencji.

Rozumowanie to przedstawić można w sposób ogólny, nie odwołujący się do swoistych własności teorii T_1. Załóżmy, iż teoria T zawiera n terminów elementarnych: A_1, A_2, ..., A_n i m terminów teoretycznych: $L_1, L_2, ..., L_m$. Niech

(1) $$\Phi(A_1, A_2, ..., A_n, L_1, L_2, ..., L_m)$$

będzie koniunkcją postulatów teorii T, a

(2) $$\mathrm{E}_{A_1, A_2 ... A_n}$$

zbiorem wszystkich elementarnych konsekwencji. Załóżmy następnie, iż wszystkie terminy teoretyczne zdefiniowane zostały za pomocą terminów elementarnych przez dołączenie do teorii T m definicji o postaci

(3) $$L_i = \chi_i$$

gdzie χ_i jest pewnym wyrażeniem elementarnym. Na definicje te nakładamy przy tym warunek żądający, aby były to twierdzenia jałowe w teorii T. Z postulatów teorii T wynika na mocy dołączonych definicji twierdzenie:

(4) $$\Phi^*(A_1, A_2, ..., A_n)$$

powstałe z postulatów (1) przez zastąpienie wszystkich terminów teoretycznych równoznacznymi wyrażeniami elementarnymi. Jest to zatem pewne twierdzenie elementarne i jako takie należeć musi do zbioru twierdzeń elementarnych (2). W przeciwnym wypadku definicje (3) nie byłyby twierdzeniami jałowymi, gdyż prowadziłyby do uzyskania pewnego twierdzenia elementarnego, które poprzednio w teorii T uzyskać się nie dawało. Wobec tego iż ze zbioru (2) wynika twierdzenie (4), a z twierdzenia (4) i definicji (3) — postulaty (1), postulaty te dają się wyprowadzić ze zbioru twierdzeń elementarnych (2) i definicji terminów teoretycznych (3). Są one zatem na mocy tych definicji równoważne logicznie zbiorowi elementarnych konsekwencji teorii T.

4.2. Ujemne skutki definiowania terminów teoretycznych przez terminy elementarne pojawiają się, jak wspomniałem, nie tylko w danej teorii, ale również w teoriach, które uważane być mogą za jej dalsze fazy rozwojowe. Jako przykład teorii reprezentującej rozwinięcie teorii T_1 podaje Braithwaite następującą teorię T_2. Załóżmy, że przyjęta jest opisana wyżej teoria T_1 wyjaśniająca pewne zależności pomiędzy obserwowalnymi własnościami: A, B, C przy pomocy hipotetycznych «czynników»: L, M, N. Przypuśćmy następnie, iż odkryto pewne zależności dalsze zachodzące pomiędzy własnościami: A, B, C a pewnymi obserwowalnymi własnościami: D, E, F. Oto niektóre z nich:

(T4) $(x)\,(Ax \cdot Dx \supset Ex)$

(T5) $(x)\,(Bx \cdot Ex \supset Fx)$ itp.

Zależności te możemy wyjaśnić nie budując nowej teorii, lecz jedynie rozbudowując istniejącą teorię T_1. Postulujemy w tym celu istnienie jeszcze jednego hipotetycznego «czynnika»: R, który wraz z «czynnikami»: L, M, N ma zdawać sprawę z zaobserwowanych zjawisk. Powstała w ten sposób teoria T_2 zawiera jako swoją część teorię T_1. Postulaty teorii T_2 składają się z postulatów teorii T_1 oraz z trzech postulatów dodatkowych:

(P4) $(x)\,(Dx \equiv Lx \cdot Rx)$

(P5) $(x)\,(Ex \equiv Mx \cdot Rx)$

(P6) $(x)\,(Fx \equiv Nx \cdot Rx)$.

Teoria T_2 wyjaśnia zarówno uogólnienia dawne jak i nowe. Wśród jej konsekwencji znajdują się wszystkie wymienione wyżej twierdzenia (T1)-(T5) oraz szereg twierdzeń innych.

Taki sposób wyjaśniania nowych zależności polegający na podciąganiu ich pod odpowiednio wzbogaconą teorię stojącą już do naszej dyspozycji — odznacza się znaczną ekonomią i odpowiada, zdaniem Braithwaite'a, rzeczywistej praktyce naukowej. Sposób taki byłby jednak niemożliwy, gdybyśmy we wchodzących tu w grę teoriach definiowali terminy teoretyczne przez terminy elementarne. Okazuje się bowiem, iż definicje terminów teoretycznych, które w teorii T_1 są formułami jałowymi, w teorii T_2 bynajmniej nimi nie są. Wzbogacona przez nie teoria T_2^* nie jest empirycznie równoważna teorii T_2. Dają się w niej udowodnić takie twierdzenia sformułowane wyłącznie w języku terminów elementarnych, które nie należą do konsekwencji teorii T_2. Twierdzenia te oczywiście mogą okazać się prawdziwe. Wykraczają one jednak poza to, co stwierdza teoria T_2. Jeżeli zatem chcemy mieć możność przejścia od teorii T_1 do opisanej wyżej teorii T_2, nie możemy dołączać do teorii T_1 definicji terminów teoretycznych przez terminy elementarne. Fakt ten ilustruje Braithwaite na konkretnym przykładzie teorii T_1 i T_2. Pokazuje, iż definicje terminów teoretycznych (D1)-(D3), które w teorii T_1 są formułami jałowymi, dołączone do postulatów (P1)-(P6) teorii T_2 prowadzą do następujących twierdzeń zawierających wyłącznie terminy elementarne i nie dających się wyprowadzić z samych postulatów (P1)-(P6):

(T1*) $(x)\,(Dx \supset Ax \vee Bx \vee Cx)$

(T2*) $(x)\,(Ex \supset Ax \vee Bx \vee Cx)$

(T3*) $(x)\,(Fx \supset Ax \vee Bx \vee Cx)$.

Definicje te nie są, co za tym idzie, formułami jałowymi w teorii T_2.

Jak uogólnić te wyniki? Wywód autora ograniczony jest do skonstruowanych przez niego teorii T_1 i T_2. Postaramy się obecnie przedstawić jego argumentację tak, aby mogła znaleźć zastosowanie i w innych przypadkach. Autor sam sugeruje pewien kierunek uogólnienia swych wywodów. Sugestia ta nie prowadzi nas jednak daleko.

Istotę ograniczenia nałożonego na terminy: L, M, N przez definicje (D1)-(D3) upatruje Braithwaite w ograniczeniu zakresu tych terminów do klasy przedmiotów, którym przysługuje co najmniej jedna z własności: A, B, C. Istotnie definicje (D1)-(D3) pociągają takie ograniczenie:

(T4*) $(x)\,(Lx \supset Ax \vee Bx \vee Cx)$

(T5*) $(x)\,(Mx \supset Ax \vee Bx \vee Cx)$

(T6*) $(x)\,(Nx \supset Ax \vee Bx \vee Cx)$.

To właśnie ograniczenie prowadzi w teorii T_2 do analogicznego ograniczenia zakresu terminów elementarnych: D, E, F, z którego zdają sprawę przytaczane poprzednio twierdzenia (T1*)-(T3*). Ograniczenie to nie ma jednak charakteru istotnego. Jest ono najwyraźniej zależne od specyficznych własności teorii T_1 i nie występuje w innych, nieznacznie tylko od T_1 różniących się teoriach. W teoriach tych definicje terminów teoretycznych prowadzą do podobnych konsekwencji nie ograniczając wcale ich zakresu do przedmiotów posiadających którąś z obserwowalnych cech. Taką ściśle analogiczną do teorii T_1 teorię T_1' otrzymać możemy zastępując w teorii T_1 wszystkie terminy elementarne ich negacjami. Otrzymaną w ten sposób teorię T_1' przedstawić można zresztą w postaci takiej, iż zarówno jej postulaty jak i jej doświadczalne konsekwencje sformułowane będą za pomocą terminów: A, B, C a nie ich negacji. Postulatami tej teorii będą twierdzenia następujące:

(P1′) $(x)\,(Ax \equiv {\sim} Lx \vee {\sim} Mx)$

(P2′) $(x)\,(Bx \equiv {\sim} Mx \vee {\sim} Nx)$

(P3′) $(x)\,(Cx \equiv {\sim} Nx \vee {\sim} Lx)$

a definicjami, odpowiadającymi definicjom (D1)-(D3), twierdzenia:

(D1′) $(x)\,(Lx \equiv {\sim} Cx \vee {\sim} Ax)$

(D2′) $(x)\,(Mx \equiv {\sim} Ax \vee {\sim} Bx)$

(D3′) $(x)\,(Nx \equiv {\sim} Bx \vee {\sim} Cx)$.

Definicje te pociągają takie same niepożądane konsekwencje co poprzednie, a jednak nie ograniczają zakresu terminów: L, M, N do klasy przedmiotów, którym przysługuje co najmniej jedna z własności: A, B, C. Można co prawda stwierdzić, iż pociągają one «ograniczenie» inne: do klasy przedmiotów, którym co najmniej jedna z własności: A, B, C nie przysługuje, ale to ograniczenie również będzie miało walor tylko dla pewnych teorii.

Na czym zatem polega istota ograniczeń, jakie na terminy teoretyczne nakładają przyjmowane w danej teorii definicje? Spróbujemy odpowiedzieć na to pytanie posługując się wprowadzonym w pierwszej części pracy pojęciem definicji cząstkowej. Wynik dalszych rozważań wyrazić można ogólnikowo, stwierdzając, iż ograniczenie, o którym mowa, polega po prostu na ograniczeniu zakresu terminów teoretycznych do jakiejś określonej — obojętnie zresztą jakiej — klasy przedmiotów. Ograniczenie takie pociąga każda definicja zupełna i ono to właśnie z reguły uniemożliwia swobodny rozwój teorii.

Przyjrzyjmy się raz jeszcze pierwotnej teorii T_1. Wspomniałem na wstępie, iż terminy teoretyczne tej teorii: *L, M, N* nie są definiowalne *explicite* — w drugim z wyróżnionych przez nas znaczeń — przez terminy elementarne: *A, B, C.* Z postulatów tej teorii (P1)-(P3) nie wynikają żadne kryteria stosowalności terminów: *L, M, N* sformułowane wyłącznie w terminach: *A, B, C* i mające postać definicji zupełnych. Jedyne kryteria, w jakie terminy: *L, M, N* zaopatruje teoria T_1 — to kryteria cząstkowe. Weźmy pod uwagę, aby uniknąć zbędnych powtórzeń, jeden z tych terminów: *L*. Sytuacja pozostałych terminów *M* i *N* jest ściśle analogiczna. Z postulatów (P1)-(P3) wynikają następujące kryteria dla terminu *L*:

(1) $$(x)\,(Ax \vee Cx \supset Lx)$$

(2) $$(x)\,((\sim Ax \vee \sim Cx) \cdot Bx \supset \sim Lx).$$

Kryteria (1)-(2) mają postać zwykłej definicji cząstkowej. Posiadają też, w konsekwencji, wszystkie charakterystyczne cechy takiej definicji. A więc, po pierwsze, formułują warunki, które się logicznie nie dopełniają. Mogą istnieć przedmioty, które nie spełniają żadnego z nich. Będą to przedmioty, które nie są ani *A*, ani *B*, ani *C*. O przedmiotach tych kryteria (1)-(2) nie pozwalają orzec ani terminu *L*, ani jego negacji. Po drugie, kryteria (1)-(2) formułują warunki, które się logicznie nie wykluczają. Mogą więc istnieć przedmioty, które spełniają oba z nich zarazem, mianowicie przedmioty, które są *A, B* i nie *C* — lub *B, C* i nie *A*. Takie przedmioty byłyby jednak na mocy kryteriów (1)-(2) *L* i nie *L* zarazem. A zatem kryteria te pociągać muszą twierdzenia, które istnienie takich przedmiotów wykluczają. Twierdzenia te — to (T1) i (T2), przytaczane przez nas poprzednio. Należą one do doświadczalnych konsekwencji teorii T_1. Kryteria (1)-(2) mogą zostać zatem, zgodnie z proponowaną w części pierwszej procedurą, rozbite na dwa składniki, z których pierwszy nie pociągałby już żadnych konsekwencji doświadczalnych, a drugi byłby identyczny z twierdzeniami: (T1) i (T2). Ograniczone w ten sposób kryteria (1)-(2) przybierają postać następującą:

(1′) $$(x)\,(Ax \vee Cx \supset Lx)$$

(2′) $$(x)\,((\sim Ax \cdot Bx \cdot \sim Cx \supset \sim Lx)^{22}.$$

Koniunkcja wypowiedzi (1′)-(2′) i twierdzeń (T1)-(T2) równoważna jest logicznie pierwotnym kryteriom (1)-(2). Kryteria stosowalności terminu *L* w formie obecnej zaopatrują ten termin w określony sens w tej samej *de facto* dziedzinie przedmiotów, co kryteria pierwotne. W szczególności, pozostaje on nieokreślony w stosunku do wszelkich przedmiotów, którym nie przysługuje ani jedna z cech: *A, B, C.* Podobnie przedstawia się sytuacja terminów: *M* i *N.* Kryteria stosowalności tych terminów odznaczają się podobnymi własnościami, dopuszczają analogiczne modyfikacje, i tak samo, jak kryte-

22) Właściwie, zgodnie z proponowaną w części pierwszej procedurą, kryterium (1) również powinno zostać ograniczone w sposób następujący:
(1″) $(x)\,((Ax \vee Cx) \cdot (Ax \cdot Cx \vee \sim Bx) \supset Lx)$
.Kryterium (1″) okazuje się jednak na mocy twierdzeń (T1) i (T2) równoważne kryterium (1).

rium omawiane, pozostawiają te terminy nieokreślonymi w stosunku do wszelkich przedmiotów, które nie są ani A, ani B, ani C.

Tutaj trzeba dodać jedno uzupełnienie. Układ kryteriów (1)-(2) i analogicznych kryteriów dla terminów pozostałych, ewentualnie układ kryteriów (1′)-(2′), analogicznych kryteriów dla terminów pozostałych oraz twierdzeń doświadczalnych (T1)-(T3) — nie wyczerpuje zawartości postulatów (P1)-(P3). Postulaty te implikują ponadto pewne kryteria stosowalności terminów teoretycznych o postaci uogólnionych definicji cząstkowych; twierdzenie formułujące kryterium stosowalności dla pary L, M:

$$(x)\,(\sim Ax \cdot \sim Bx \cdot \sim Cx \supset \sim Lx \vee \sim Mx)$$

oraz analogiczne twierdzenia dla par: M, N i N, L.[23] Twierdzenia te nie prowadzą jednakże do żadnych dodatkowych kryteriów dla poszczególnych terminów teoretycznych, gdyż podają warunki, które wykluczają się z warunkami sformułowanymi w definicjach cząstkowych poszczególnych terminów teoretycznych. Nadal więc o żadnym przedmiocie nie będącym ani A, ani B, ani C nie potrafimy rozstrzygnąć, czy jest, czy też nie jest L (ew. M czy N).

Terminy teoretyczne: L, M, N reprezentują więc tzw. pojęcia otwarte. Ich zakres nie jest ograniczony do żadnej określonej klasy przedmiotów. Istnieje sfera nieokreśloności, w której terminy te pozostają niezdeterminowane. Wprowadzenie do teorii T_1 definicji (D1)-(D3) powoduje «domknięcie» tych terminów. Ogranicza ono ich zakres do określonej klasy przedmiotów, likwidując tym samym ową sferę nieokreśloności. Definicja (D1) terminu L wyposaża ten termin w uniwersalne kryteria stosowalności, pozwalające o każdym przedmiocie rozstrzygnąć, czy jest, czy też nie jest L. A więc i o przedmiotach, którym żadna z własności: A, B, C nie przysługuje i w stosunku do których rozstrzygnięcie takie było do tej pory niemożliwe. Definicja (D1) polega właśnie na dołączeniu do teorii T_1 twierdzenia, które wszystkie te przedmioty kwalifikuje jako nie L:

(T7*) $$(x)\,(\sim Ax \cdot \sim Bx \cdot \sim Cx \supset \sim Lx).$$

Analogiczne twierdzenia implikują definicje pozostałych terminów teoretycznych.

Twierdzenie (T7*) nie prowadzi w teorii T_1 do żadnych konsekwencji doświadczalnych, gdyż właśnie w stosunku do tych przedmiotów, o których mowa w tym twierdzeniu, termin L pozbawiony był jakichkolwiek kryteriów stosowalności. Mamy więc całkowitą swobodę w decydowaniu, czy przedmiotom tym przypisać termin L, czy jego negację. Z tego też powodu definicja (D1), która się do takiej decyzji sprowadza, jest na gruncie teorii T_1 formułą jałową.

Inaczej jest jednak na gruncie teorii T_2. Teoria ta zawiera trzy dalsze postulaty (P4)-(P6). Postulaty te dostarczają dodatkowe kryteria stosowalności dla terminów: L, M, N sformułowane za pomocą terminów elementarnych: D, E, F. Kryteria te również mają postać definicji cząstkowych. Oto definicja taka dla terminu L:

---

23) Przedstawienie teorii T_1 jako układu zwykłych i uogólnionych definicji cząstkowych podaje również w sposób zbliżony do omawianego, Stopes-Roe [23].

(3) $$(x)\,(Dx \supset Lx)$$

(4) $$(x)\,(\sim Dx \cdot (Ex \vee Fx) \supset \sim Lx).$$

Analogicznie — dla M i N. Twierdzenia (3)-(4) podają kryteria stosowalności terminu L m.in. dla przedmiotów, dla których twierdzenia (1′)-(2′) również pewne kryteria przewidywały. Nowe kryteria zachodzą zatem na dawne, prowadząc w ten sposób do pewnych doświadczalnych konsekwencji, których przykładem mogą być twierdzenia następujące:

(T6) $$(x)\,(Ax \vee Cx \supset Dx \vee \sim Ex \cdot \sim Fx)$$

(T7) $$(x)\,(\sim Ax \cdot Bx \cdot \sim Cx \supset \sim Dx),$$

należące do zasobu elementarnych konsekwencji teorii T_2. Dla naszych rozważań ważne jest to, iż twierdzenia (3)-(4) podają kryteria stosowalności terminu L również dla takich przedmiotów, dla których twierdzenia (1′)-(2′) żadnych kryteriów nie przewidywały, tj. dla przedmiotów będących nie A, nie B i nie C. Pewne z nich przy tym zostają na mocy kryteriów (3)-(4) zakwalifikowane jako L, inne — jako nie L:

$$(x)\,(\sim Ax \cdot \sim Bx \cdot \sim Cx \cdot Dx \supset Lx)$$

$$(x)\,(\sim Ax \cdot \sim Bx \cdot \sim Cx \cdot \sim Dx \cdot (Ex \vee Fx) \supset \sim Lx).$$

Zwężona zostaje w ten sposób sfera nieokreśloności charakteryzująca termin L w teorii T_1. Przypuśćmy teraz, iż do teorii T_2 dołączona zostaje definicja (D1). Definicja ta pociąga, jak widzieliśmy, twierdzenie (T7*), przypisujące wszystkim przedmiotom będącym nie A, nie B i nie C cechę nie L. Twierdzenie to pozostaje w sprzeczności z kryteriami (3)-(4), które pewnym przedmiotom będącym nie A, nie B i nie C — tym mianowicie, które są D — przypisują cechę L. Sprzeczności tej można uniknąć tylko przez przyjęcie, iż przedmioty takie nie istnieją. Twierdzenie (T7*) prowadzi przeto na gruncie teorii T_2 do twierdzenia:

(T8*) $$(x)\,(\sim Ax \cdot \sim Bx \cdot \sim Cx \supset \sim Dx).$$

Jest to twierdzenie równoważne logicznie twierdzeniu (T1*), przytaczanemu poprzednio jako przykład twierdzenia elementarnego nie dającego się udowodnić w teorii T_2. Tak więc dołączenie do teorii T_2 definicji terminu teoretycznego L pociąga za sobą pewne twierdzenie doświadczalne, które bez tej definicji uzyskać się nie daje. Analogicznie jest w przypadku terminów: M i N, których definicje pociągają twierdzenia równoważne logicznie twierdzeniom (T2*) i (T3*). Definicje terminów: L, M, N nie są wobec tego w teorii T_2 twierdzeniami jałowymi.

Chcąc uogólnić osiągnięte wyniki, musimy zdać sprawę z tego, jakie własności omawianych teorii T_1 i T_2 odgrywają w przedstawionej argumentacji rolę istotną. Naszkicujmy ją zatem raz jeszcze w sposób ogólny, podkreślając założenia, które wydają się niezbędne. Weźmy pod uwagę dowolną teorię T zawierającą termin teoretyczny Q i załóżmy, iż postulaty tej teorii zaopatrują termin Q w cząstkowe jedynie kryteria stosowalności. Istnieje wówczas klasa przedmiotów K, w stosunku do których termin Q żadnych kryteriów stosowalności nie posiada. Załóżmy następnie, iż przyjmujemy definicję terminu Q, która w teorii T jest formułą jałową. Definicja taka sprowadza się do twierdzenia orzekającego o przedmiotach klasy K bądź termin Q, bądź jego negację.

Załóżmy wreszcie, iż w procesie rozwoju teorii T dołączamy do niej dalsze postulaty, które zaopatrują termin Q w nowe kryteria stosowalności. Jeśli teraz owe kryteria okazują się niezgodne z kryteriami ustalonymi przez definicję terminu Q, definicja ta przestaje być formułą jałową. Prowadzi bowiem do twierdzenia, które głosi, iż nie istnieją przedmioty, którym definicja terminu Q przypisywałaby termin Q, a owe kryteria dodatkowe — jego negację, lub na odwrót. Tak jest właśnie w teoriach T_1 i T_2. Czy tak bywa zawsze? Na pewno nie. Sytuacja taka — nie będąc powszechną — wydaje się jednak typowa. Warunki, o których mowa, wydają się być spełnione przez obszerną klasę teorii empirycznych. Zastanówmy się bowiem, do czego się one sprowadzają.

Definicja terminu Q nie prowadzi do żadnych konsekwencji doświadczalnych tylko wtedy, gdy postulaty teorii T wykluczają istnienie takich przedmiotów klasy K, o których definicja terminu Q orzekałaby termin Q, a kryteria stosowalności terminu Q implikowane przez nowe postulaty teorii T — jego negację, lub na odwrót. Sytuacja taka miałaby miejsce np. wtedy, gdyby postulaty teorii T wykluczały w ogóle istnienie przedmiotów klasy K podpadających pod dodatkowe kryteria terminu Q. Kryteria takie nie zmniejszałyby wówczas sfery nieokreśloności, jaką terminowi Q pozostawiały kryteria pierwotne. Takie rozszerzenie teorii T uważać jednak można za rozszerzenie pod względem pojęciowym nieistotne. Każde istotne rozszerzenie teorii dostarcza takich kryteriów stosowalności terminów teoretycznych, które zwężają ich początkową sferę nieokreśloności. Może być jednak i tak, że postulaty teorii T, nie wykluczając istnienia wszelkich przedmiotów klasy K podpadających pod owe kryteria dodatkowe, wykluczają istnienie tylko takich spośród nich, co do których kryteria te prowadziłyby do orzeczeń sprzecznych z orzeczeniami implikowanymi przez definicję terminu Q. Taki zbieg okoliczności wydaje się mało prawdopodobny. Definicja terminu Q ma z konieczności charakter arbitralny. Przyjmując ją dysponujemy jedynie pierwotnymi kryteriami stosowalności terminu Q o charakterze cząstkowym. Kryteria te nie narzucają, ani nie sugerują żadnej decyzji odnośnie przedmiotów klasy K. Decyzja ta ma charakter czysto konwencjonalny. Polega też zwykle na zaliczeniu wszystkich tych przedmiotów bądź do zakresu terminu Q, bądź do zakresu jego negacji. Dołączone następnie do teorii T postulaty nie mają charakteru arbitralnego. Przyjęte zostają po to, aby wyjaśnić pewne stwierdzone doświadczalnie zależności. Implikowane przez nie kryteria stosowalności terminu Q są więc również zależne od doświadczenia. Odwołują się przy tym z reguły do innych niż kryteria pierwotne obserwowalnych własności i pozwalają na zróżnicowanie przedmiotów klasy K przez zaliczenie jednych z nich do zakresu terminu Q, innych — do zakresu jego negacji. Jest rzeczą mało prawdopodobną, aby kryteria takie okazały się zgodne z przyjętą poprzednio w sposób arbitralny definicją. Byłby to wyjątkowy raczej zbieg okoliczności, gdyby okazało się, że nie istnieją przedmioty, w stosunku do których owe kryteria i definicja prowadziłyby do wniosków sprzecznych. Jeszcze mniej prawdopodobne jest to, aby twierdzenia wykluczające istnienie takich przedmiotów wynikały z postulatów teorii T. A tylko w takim wypadku definicja terminu Q nie prowadziłaby do konsekwencji doświadczalnych w rozszerzonej teorii.

Na ogół przeto bywa inaczej. A wówczas pojawiają się owe niepożądane skutki definiowania terminów teoretycznych przez terminy elementarne, na które wskazuje Braithwaite. Definicje takie ograniczają rozwój teorii. Prowadzą w dalszych fazach rozwojowych do pewnych niezamierzonych konsekwencji, które z reguły okazują się sprzeczne z doświadczeniem i zmuszają do rewizji przyjętej poprzednio teorii.

* * *

Odtwórzmy na zakończenie w paru słowach przebieg naszych rozważań. Pierwsza ich część poświęcona była kryteriom stosowalności, w jakie teorie empiryczne zaopatrują terminy teoretyczne. Kryteria takie wyrażać mogą różnego rodzaju związki pomiędzy terminami teoretycznymi a terminami elementarnymi. Mogą więc przybierać rozmaite formy: od definicji zupełnej, poprzez różne odmiany definicji cząstkowej: zwykłą, probabilistyczną, jednostronną, uogólnioną — aż do twierdzeń, które scharakteryzować się dają jedynie przez fakt wyposażania terminów teoretycznych w empiryczny sens. Dokonaliśmy przeglądu takich kryteriów, główną uwagę poświęcając definicji cząstkowej, w jej różnych wersjach, jako doniosłej a mało stosunkowo opracowanej procedurze definicyjnej. Rozpatrzyliśmy zwłaszcza bliżej pewne problemy logiczne, jakie ten rodzaj definicji nasuwa. Pozostała część rozważań poświęcona była problemowi definiowalności terminów teoretycznych, rozumianej jako możliwość dołączania definicji zupełnych terminów teoretycznych do teorii, która takich definicji nie zakłada. Negatywne rozstrzygnięcie tego problemu przez Braithwaite'a polega na okazaniu, iż definicje takie prowadzą do pewnych niepożądanych konsekwencji dotyczących zarówno aktualnej teorii jak i jej dalszych możliwości rozwojowych. Rezultat ten, udowodniony przez Braithwaite'a dla pewnej przykładowo wybranej teorii, został uogólniony w pracy niniejszej.

Bibliografia

[1] K. Ajdukiewicz: „Le problème du fondement des propositions analytiques". *Studia Logica*, 8 (1959).

[2] R. B. Braithwaite: *Scientific Explanation*. Cambridge, 1953.

[3] R. B. Braithwaite: „Axiomatizing a Scientific System by Axioms in the Form of Identifications". *The Axiomatic Method...* Amsterdam, 1959.

[4] R. Carnap: „Testability and Meaning". *Philosophy of Science* 3 (1936), 4 (1937).

[5] R. Carnap: *Foundations of Logic and Mathematics*. Chicago, 1939.

[6] R. Carnap: „The Methodological Character of Theoretical Concepts". *Minnesota Studies in the Philosophy of Science* I, 1956.

[7] W. Craig: „On Axiomatizability within a System". *Journal of Symbolic Logic*, 18 (1953).

[8] W. Craig: „Replacement of Auxiliary Expressions". *Philosophical Review*, 65 (1956).

[9] C. G. Hempel: „The Theoretician's Dilemma". *Minnesota Studies in the Philosophy of Science* II, 1958.

[10] D. Hilbert, P. Bernays: *Grundlagen der Mathematik I.* Berlin, 1934.

[11] J. Kamińska: „Ewolucja Koła Wiedeńskiego". *Myśl Współczesna*, 2 (1947).

[12] A. Kaplan: „Definition and Specification of Meaning". *Journal of Philosophy*, 43 (1946).

[13] H. Mehlberg: „Positivisme et Science". *Studia Philosophica* 3 (1948).

[14] H. Mehlberg: *The Reach of Science.* Toronto, 1958.

[15] E. Nagel: „A Budget of Problems in the Philosophy of Science". *Philosophical Review*, 66 (1957).

[16] A. Pap: „Reduction Sentences and Open Concepts". *Methodos*, 5 (1953).

[17] M. Przełęcki: „O tzw. definicjach operacyjnych". *Studia Logica*, 3 (1955).

[18] M. Przełęcki: „Operacjonizm". *Archiwum Historii Filozofii i Myśli Społecznej*, 5 (1959).

[19] M. Przełęcki: „Postulat empiryczności terminów przyrodniczych". *Fragmenty Filozoficzne — Seria Druga*, Warszawa. 1959.

[20] M. Przełęcki: „W sprawie terminów nieostrych". *Studia Logica*, 8 (1958).

[21] M. Przełęcki: „On the Concept of Genotype". W druku.

[22] F. P. Ramsey: *The Foundation of Mathematics and other Logical Essays.* London, 1931.

[23] H. V. Stopes-Roe: „Some Considerations Concerning «Interpretative Systems»". *Philosophy of Science*, 25 (1958).

[24] A. Tarski: „Z badań metodologicznych nad definiowalnością terminów". *Przegląd Filozoficzny*, 37 (1934).

Filozofia Nauki
Rok I, 1993, Nr 2-3

W sprawie uzasadniania zdań spostrzeżeniowych

Obok pośrednich sposobów uzasadniania twierdzeń, odwołujących się do innych, uprzednio uzasadnionych, twierdzeń jako do przesłanek, wyróżnia się zwykle bezpośrednie sposoby uzasadniania twierdzeń, które do żadnych innych twierdzeń się nie odwołują. Do sposobów bezpośrednich zalicza się przede wszystkim uzasadnianie na podstawie «bezpośredniego świadectwa doświadczenia». Owo «bezpośrednie świadectwo doświadczenia» utożsamia się z rezultatami naszych spostrzeżeń, a zdania w ten sposób uzasadnialne — ze zdaniami spostrzeżeniowymi. Zdanie *Z* jest więc uważane za zdanie spostrzeżeniowe, jeśli można rozstrzygnąć, czy jest tak, jak to zdanie głosi (innymi słowy: uzasadnić zdanie *Z* lub jego negację), bezpośrednio na podstawie spostrzeżenia, bez uzasadniania jakichkolwiek zdań innych. Określenie to wymaga oczywiście dalszej precyzacji, w szczególności — relatywizacji do osoby *O* lub języka *J*.

Powstaje pytanie, czy istnieją zdania spełniające powyższy warunek. Na pytanie to odpowiada się z reguły twierdząco, a zdania takie usiłuje się charakteryzować na różne sposoby. Przede wszystkim — w sposób, który nazwać można «ontologicznym». Zdaniami spostrzeżeniowymi mają być zdania, które stwierdzają spostrzegalne fakty: przysługiwanie spostrzegalnej własności spostrzegalnemu przedmiotowi lub zachodzenie spostrzegalnego stosunku pomiędzy spostrzegalnymi przedmiotami. Innymi słowy, zdaniem spostrzeżeniowym jest zdanie atomowe, w którym zarówno nazwy indywiduowe, jak i predykaty, denotują przedmioty spostrzegalne; w przypadku najprostszym — zdanie „a ε P”, gdzie „a” denotuje spostrzegalne indywiduum, a „P” — spostrzegalną własność.

Charakterystykę tę trudno uznać za zadowalającą ze względu na niejasność pojęcia przedmiotu spostrzegalnego. Nawet jeśli przyjmiemy, iż pojęcie spostrzegalnego indywiduum jest dostatecznie zrozumiałe, sens pojęcia spostrzegalnej własności (czy klasy) pozostaje wysoce zagadkowy. W sensie dosłownym — spostrzegać możemy tylko konkretne indywidua, gdyż tylko one mogą być bodźcami dla naszych narządów zmysłowych. O spostrzegalności przedmiotów abstrakcyjnych, takich jak własność (czy klasa), w tym samym literalnym znaczeniu mówić nie można. Jakiż więc sens może mieć pojęcie spostrzegalnej własności? Jest to w każdym razie pojęcie niezmiernie nieostre. Jego zakres pozostaje w znacznym stopniu nieustalony. W tych zaś przypadkach, gdzie rozróżnianie spostrzegalnych i niespostrzegalnych własności istotnie ma miejsce, dotyczy ono własności pojmowanych jako coś różnego od odpowiadających im klas. Barwę zieloną zalicza się do własności spostrzegalnych, a własność wysyłania fal elektromagnetycznych o odpowiedniej długości — do niespostrzegalnych, choć klasa przedmiotów zielonych jest identyczna z klasą przedmiotów wysyłających fale elektromagnetyczne o danej długości. Rozróżnienie owo dotyczy zatem nie denotacji, ale raczej treści (czy znaczenia) terminów występujących w zdaniu spostrzeżeniowym, czyli tzw. terminów spostrzeżeniowych. Ma więc w gruncie rzeczy charakter nie tyle «ontologiczny», co «językowy».

Charakter wyraźnie «językowy» nosi inna próba określenia zdań spostrzeżeniowych. Bierzemy w niej pod uwagę nie to, jakimi przedmiotami są denotacje terminów spostrzeżeniowych, lecz to, w jaki sposób denotacje zostały tym terminom przyporządkowane. Wyróżnia się zwykle dwa rodzaje interpretacji terminów deskryptywnych: interpretację pośrednią i bezpośrednią. Interpretując dany termin deskryptywny pośrednio, korzystamy z innych terminów deskryptywnych zinterpretowanych uprzednio. Natomiast interpretacja bezpośrednia nie odwołuje się do żadnych innych terminów deskryptywnych o ustalonych denotacjach. Pośrednia interpretacja danego terminu przybierać może postać zwykłej, werbalnej definicji tego terminu. Przedmiot, który stanowić ma jego denotację, zostaje w niej scharakteryzowany za pomocą słownego opisu. Bezpośrednią interpretację danego terminu utożsamia się z jego definicją ostensywną. Przyporządkowuje ona temu terminowi przedmiot, który stanowić ma jego denotację, nie przez słowny opis tego przedmiotu, lecz po prostu przez jego wskazanie. Wydaje się wobec powyższego, iż podział na terminy spostrzeżeniowe i niespostrzeżeniowe pokrywa się z podziałem na terminy zinterpretowane bezpośrednio i terminy zinterpretowane pośrednio. Termin, któremu nadano interpretację przez proste wskazanie pewnych przedmiotów, może być stosowany bezpośrednio na podstawie spostrzeżenia. Natomiast zastosowanie terminu zinterpretowanego przy pomocy innych terminów sprowadza się bezpośrednio do zastosowania owych terminów, a tylko pośrednio — do spostrzeżeń. A zatem zdanie „$a \varepsilon P$" jest zdaniem spostrzeżeniowym, jeśli zarówno „a" jak i „P" są terminami zinterpretowanymi bezpośrednio, czyli terminami wprowadzonymi do języka przez definicje ostensywne.

Problem definicji ostensywnych rozważany był szczegółowo na innym miejscu[1]. Tutaj streścimy pokrótce pewne rezultaty tych rozważań, pozostające w związku z problemem uzasadniania zdań spostrzeżeniowych. Definicja ostensywna nazwy indy-

1) M. Przełęcki: „O definiowaniu terminów spostrzeżeniowych".

widuowej okazuje się zabiegiem nie nastręczającym jakichś trudności zasadniczych. Wskazanie pewnego przedmiotu i stwierdzenie, iż przedmiot ten jest identyczny z *a*, przyporządkowuje wskazany przedmiot jako denotację nazwie indywiduowej „*a*". Trudności powstają przy ostensywnym definiowaniu predykatów. Przedmiotu, który stanowić ma denotację predykatu, wskazać nie można, gdyż jest to przedmiot abstrakcyjny: własność (klasa). Trzeba się ograniczyć do wskazania pewnych przedmiotów konkretnych (takich, które do danej klasy należą, i takich, które do niej nie należą) w charakterze przedmiotów wzorcowych (pozytywnych i negatywnych). Tak więc, predykat „zielony" definiujemy ostensywnie, wskazując poszczególne przedmioty zielone (*resp.* niezielone) i orzekając o nich ten predykat (*resp.* jego negację). Ale układ tego rodzaju postulatów stwierdzających, iż ten przedmiot jest zielony, a tamten nie jest, wyznacza denotację predykatu „zielony" w sposób niezmiernie wieloznaczny. Postulaty te spełnia dowolna klasa zawierająca wszystkie wzorce pozytywne i nie zawierająca żadnego z wzorców negatywnych. Wszelka taka klasa może być zatem denotacją predykatu „zielony". Usiłuje się niekiedy definicję ostensywną predykatu ująć w postaci definicji równoważnościowej stwierdzającej, iż przedmiot zielony — to taki jak przedmioty wskazane. Ale wszelkie próby precyzacji takiej definicji w sposób nie odwołujący się do innych predykatów deskryptywnych dają w wyniku sformułowania nieadekwatne. Klasy przyporządkowywane przez tego rodzaju definicje predykatowi „zielony" nie pokrywają się bynajmniej z klasą przedmiotów zielonych; są jawnie za szerokie, lub za wąskie. Nasuwa się wniosek, iż jednoznaczna i adekwatna może być tylko taka definicja ostensywna predykatu „zielony", która odwołuje się do innych, już zinterpretowanych predykatów deskryptywnych (takich jak „barwa" lub „równobarwny"). Omawiany układ postulatów tylko wtedy wyznacza denotację predykatu „zielony" w sposób jednoznaczny, gdy zawiera założenie ograniczające wchodzące w grę klasy do barw, czyli klas przedmiotów równobarwnych. Tego rodzaju definicja ostensywna przestaje być jednak bezpośrednim sposobem interpretacji predykatu „zielony". Predykat ten przestaje być tym samym — zgodnie z przyjętym określeniem — terminem spostrzeżeniowym. Nie jest przy tym nim również predykat „równobarwny", gdyż jego bezpośrednia interpretacja nasuwa te same problemy, co interpretacja predykatu „zielony".

Faktem pozostaje jednak, iż pewne predykaty, w rodzaju predykatu „zielony", interpretowane bywają bezpośrednio i że uzyskują na tej drodze interpretację w przybliżeniu jednoznaczną. Definicja ostensywna predykatu „zielony", traktowana jako sposób bezpośredni, jest więc procedurą faktycznie skuteczną. Musimy zatem przyjąć, iż zabieg polegający na wskazaniu odpowiednich przedmiotów i wygłoszeniu wspomnianych postulatów może wywołać u osoby, do której jest skierowany, skutek polegający na rozumieniu predykatu „zielony" jako predykatu denotującego klasę przedmiotów zielonych. Może więc czynić tę osobę zdolną do właściwego posługiwania się owym predykatem. Rozważania poprzednie pokazują jednak, iż omawiana procedura nie dostarcza uzasadnienia dla takiego wniosku. Ów układ postulatów nie pociąga logicznie twierdzenia, iż denotacją predykatu „zielony" jest klasa przedmiotów zielonych. Denotację tę determinuje, jak widzieliśmy, w stopniu jedynie nieznacznym. Wyłącznie o przedmiotach tożsamych z którymś z przedmiotów wzorcowych możemy na gruncie tak pojętej definicji ostensywnej rozstrzygać zasadnie, czy należą one do denotacji predykatu „zielony". Nie mamy żadnych racji logicznych do rozstrzygnięcia tego pytania w stosunku do przedmiotów pozostałych. A zatem zdanie orzekające tak zdefiniowany predykat „zielony" o którymkolwiek z tych przedmiotów, nie daje się rozstrzygnąć w sposób zasadny. Jeśli faktycznie dokonujemy pewnego rozstrzygnięcia, nie stanowi ono uzasadnienia otrzymanej odpowiedzi. Można oczywiście nadać predykatowi „zielony" interpretację taką, przy której zdanie przypisujące ten predykat dowolnemu przedmiotowi spostrzegalnemu będzie zdaniem uzasadnialnym. Interpretacja taka nie może być jednak, jak widzieliśmy, interpretacją bezpośrednią. Zdanie takie nie będzie, co za tym idzie, zdaniem spostrzeżeniowym.

Filozofia Nauki
Rok I, 1993, Nr 2-3

Interpretacja terminów teoretycznych: w obronie dogmatu empiryzmu.

Na wstępie chciałbym zaznaczyć, że niniejszy artykuł nie zawiera żadnych nowych rezultatów z formalnej metodologii nauk. Nie ma on charakteru formalnego, lecz raczej filozoficzny. Zamierzam przeanalizować w nim problem interpretacji terminów teoretycznych znany pod nazwą cząstkowej interpretacji znaczenia terminów teoretycznych — skupiając się przede wszystkim na jego głównych założeniach i konsekwencjach. Interpretacja ta jest oparta zarówno na pewnych założeniach formalnych, jak i filozoficznych. Założenia formalne są założeniami charakterystycznymi dla podejścia teoriomodelowego. Założenia filozoficzne są tożsame z tezami tzw. semantycznego empiryzmu. Założenia te pociągają za sobą pewne paradoksalnie brzmiące konsekwencje, dotyczące interpretacji ważnej klasy terminów teoretycznych. Konsekwencje te zostały najjaśniej przedstawione przez Johna A. Winniego w jego znanym artykule „The implicit definition of theoretical terms" (*The British Journal for the Philosophy of Science*, 1967). Ze względu na swój nieco instrumentalistyczny charakter, są one uważane za jawnie fałszywe przez niektórych filozofów nauki, którzy w następstwie tego uważają, że owe konsekwencje obalają założenia, z których wynikają. Spróbuję podważyć tę opinię. Krytyka pod adresem interpretacji cząstkowej wydaje mi się nieuzasadniona. O ile potrafię to ocenić, interpretacja ta jest zgodna z faktyczną interpretacją terminów teoretycznych w teoriach naukowych, a alternatywne — naiwnie realistyczne — podejście do ich interpretacji jest trudne do utrzymania. Jest to główna teza, którą uzasadniam w tym artykule.

Przedstawmy w skrócie pogląd, który będziemy omawiać, i najbardziej kontrowersyjne implikacje tego poglądu. Niech L będzie językiem pewnej teorii empirycznej. Jako formalną strukturą dla semantyki języka L będę posługiwał się tutaj aparatem pojęciowym teorii modeli — na tyle znanym, by pominąć jego szczegółowe objaśnienie. Zgodnie z tym podejściem, interpretacja języka *L* jest tożsama z odpowiednim tworem teoriomnogościowym, zwanym *modelem dla L*. Dla języka *L* najprostszego typu będzie on składał się z uniwersum dla zmiennych języka *L* i z denotacji dla predykatów pozalogicznych z *L*. Przez uniwersum będziemy rozumieć pewien niepusty zbiór, a przez denotacje — zbiory obiektów należących do tego uniwersum (albo zbiory *n*-tek takich obiektów — w wypadku predykatów *n*-argumentowych). Obiekty takie będziemy nazywać *obiektami desygnowanymi przez dany predykat*. Zakładamy, że znana jest definicja prawdziwości w modelu $\mathfrak{M}$ zdania α należącego do języka *L*.

Jak zauważyliśmy wyżej, filozoficzne przesłanki leżące u podstaw przedstawianego poglądu są wyrazem pewnego rodzaju empirystycznej epistemologii i należą do doktryny nazywanej czasami *semantycznym empiryzmem*. Naszkicujmy główne założenia tej doktryny. Zgodnie z pewną jej wersją , terminy empiryczne języka *L* mogą być podzielone na trzy klasy: terminy obserwacyjne, albo inaczej *O*-terminy, terminy «mieszane», albo inaczej *M*-terminy i terminy teoretyczne, albo inaczej *T*-terminy. (Stosuję tu terminologię Winniego; zwykle zarówno *T*-terminy, jaki i *M*-terminy, są nazywane terminami teoretycznymi w szerszym tego słowa znaczeniu, a *T*-terminy są odróżniane od innych jako terminy teoretyczne *sensu stricto*.) *O*-terminy odnoszą się do własności (i relacji) obserwacyjnych i przypisują je tylko przedmiotom obserwowalnym. Wszystkie pozostałe terminy odnoszą się do własności (i relacji) teoretycznych; o ile jednak *M*-terminy stosują się zarówno do obiektów obserwowalnych, jak i nieobserwowalnych, o tyle *T*-terminy stosują się wyłącznie do przedmiotów nieobserwowalnych. Terminy takie, jak „zielony", „masa", „elektron", są odpowiednio przykładami tych rodzajów terminów.

Powyższe rozróżnienia są dość nieokreślone i niejasne. Jakiekolwiek próby zaostrzenia ich znaczenia są zadaniem wykraczającym poza ramy tego artykułu. Chciałbym ograniczyć się tylko do pewnego komentarza w celu rozwiania przynajmniej niektórych niejasności. Rozróżnienie między przedmiotami obserwowalnymi i nieobserwowalnymi sprowadza się z grubsza rzecz biorąc do rozróżnienia między makro- i mikro-obiektami. Pojęcie cechy (albo relacji) obserwacyjnej wydaje się jeszcze bardziej zagadkowe. Jeden ze sposobów wyjaśnienia jego znaczenia odwołuje się do pojęcia ostensji. Przy takim jego rozumieniu, różnica między cechami obserwacyjnymi i teoretycznymi sprowadza się do różnicy między cechami ostensywnymi i nieostensywnymi (cokolwiek miałoby to znaczyć). Dla naszych późniejszych rozważań nie jest ważne jak dokładnie przeprowadzamy ten podział. Istotne jest jedynie założenie, że istnieją nieobserwowalne rzeczy i nieobserwacyjne cechy. A to jest, jak się zdaje, trudne do podważenia.

Podstawowym problemem, powstającym w związku z tak scharakteryzowanym językiem *L*, jest problem interpretacji jego terminów teoretycznych: w jaki sposób terminom teoretycznym przyporządkowywane są ich denotacje? Albo mówiąc słowami Winniego: jak to się dzieje, że terminy te desygnują przedmioty nieobserwowalne? Trzeba zaznaczyć, że pytanie to traktuje się tutaj jako pytanie o charakterze logicznym (albo metodologicznym), a nie jako pytanie faktualne (np. historyczne albo psychologiczne): jako pytanie *quid iuris*, a nie *quid facti*. Odpowiedź na nie stanowi podstawowy dogmat semantycznego empiryzmu. Sprowadza się on do następujących stwierdzeń. Istnieją tylko dwa sposoby interpretacji terminów: poprzez to, co robimy, albo poprzez to, co mówimy. To, co dany termin desygnuje może być mu przypisane bezpośrednio — przez wskazanie odpowiednich obiektów, albo pośrednio — przez opisanie tych obiektów. Oczywiście tylko obiekty obserwowalne mogą być przyporządkowywane terminom bezpośrednio; obiekty nieobserwowalne mogą być przyporządkowywane tylko pośrednio. W konsekwencji jedynymi terminami, które mogą być interpretowane bezpośrednio, są *O*-terminy. *M*- i *T*-terminy muszą być interpretowane pośrednio. Każdy sposób bezpośredniej interpretacji sprowadza się do pewnego rodzaju procedury ostensywnej. Zakłada się, że w ten sposób *O*-terminom nadaje się ich właściwą interpretację. Pośredni sposób interpretacji polega na przyjęciu pewnych zdań, nazywanych postulatami, które łączą interpretowane terminy z pewnymi terminami już zinterpretowanymi. Właściwą interpretację *M*-terminów i *T*-terminów stanowi więc taka interpretacja, która zapewnia prawdziwość odpowiednim postulatom, pozostawiając przy tym niezmienioną właściwą interpretację *O*-terminów.

Formalizację tych idei można naszkicować w sposób następujący. Niech L_O będzie obserwacyjnym podjęzykiem języka *L*, zawierającym *O*-terminy jako jedyne terminy pozalogiczne. Zakładamy, że *model właściwy* dla L_O, $\mathfrak{M}^*_O$, jest zdefiniowany w sposób jednoznaczny poprzez pewne procedury ostensywne. Jego uniwersum, $U_{\mathfrak{M}^*_O}$, i denotacje przyporządkowane *O*-terminom, składają się wyłącznie z obiektów obserwowalnych. *Model właściwy całego języka L*, $\mathfrak{M}^*$, jest scharakteryzowany przez następujące warunki:

(i) postulaty dla *M*- i *T*-terminów są prawdziwe w $\mathfrak{M}^*$;

(ii) uniwersum modelu $\mathfrak{M}^*$, $U_{\mathfrak{M}^*}$ zawiera uniwersum modelu $\mathfrak{M}^*_O$, $U_{\mathfrak{M}^*_O}$;

(iii) denotacje *O*-terminów w $\mathfrak{M}^*$ są identyczne z denotacjami tych terminów w $\mathfrak{M}^*_O$;

(iv) denotacje *T*-terminów w $\mathfrak{M}^*$ są ograniczone do zbioru $U_{\mathfrak{M}^*} - U_{\mathfrak{M}^*_O}$, tzn. do zbioru nieobserwowalnych obiektów uniwersum $\mathfrak{M}^*$.

Jest oczywiste, że powyższe warunki nie charakteryzują modelu $\mathfrak{M}^*$ jednoznacznie; nie są one tożsame z definicją modelu $\mathfrak{M}^*$. Definiują one jedynie klasę modeli, M^*, którą będziemy nazywać klasą modeli właściwych języka *L*.

Spostrzeżenie Winniego pozwala nam uświadomić sobie, jak słabe są powyższe warunki, i jak w efekcie obszerna jest klasa M^*. Jeśli do M^* należy model $\mathfrak{M}$, którego uniwersum zawiera pewne obiekty nieobserwowalne, to do M^* będzie również należał model $\mathfrak{M}'$, którego uniwersum zawiera w charakterze tych obiektów pewne byty abstrakcyjne, np. liczby. Innymi słowy, jeśli dla pewnego modelu $\mathfrak{M} \in M^*$, zbiór $U_{\mathfrak{M}} - U_{\mathfrak{M}^*_O}$ jest niepusty, to istnieje model $\mathfrak{M}' \in M^*$, taki że zbiór $U_{\mathfrak{M}'} - U_{\mathfrak{M}^*_O}$ jest identyczny ze zbiorem liczb. Zatem obiekty desygnowane przez *T*-terminy mogą również być liczbami. Wśród zamierzonych interpretacji terminów teoretycznych jest zawsze interpretacja liczbowa. Podobny wniosek stosuje się do *M*-terminów, chociaż w bardziej ograniczonym zakresie. Pewne obiekty desygnowane przez te terminy, mianowicie obiekty nieobserwowalne, także mogą być utożsamione z liczbami.

Konsekwencje te uznane zostały za niedopuszczalne — zwłaszcza przez tych, którzy deklarowali się jako realiści w sporze między realizmem i instrumentalizmem. Ich argumentacja przebiegała następująco. Obiekty desygnowane przez terminy teoretyczne w teoriach naukowych, chociaż nieobserwowalne, powinny być bytami fizycznymi. Każda interpretacja, która identyfikuje je z liczbami, albo z jakimiś innymi bytami abstrakcyjnymi, jest w sposób oczywisty interpretacją niezamierzoną. Takie niezamierzone interpretacje nie zostały jednak wyłączone ze zdefiniowanej wyżej klasy interpretacji właściwych . Tak więc, „coś tu musi być nie w porządku" — jak konkluduje Winnie. Chciałbym podważyć ten wniosek. Według mnie, interpretacja terminów teoretycznych przedstawiona powyżej jest zgodna z ich faktyczną interpretacją w teoriach naukowych. Pozornie paradoksalne konsekwencje wskazane przez Winniego są wbrew pozorom zgodne z tą interpretacją.

Rozważmy, dla ustalenia uwagi, termin „elektron" — klasyczny przykład terminu teoretycznego w fizyce. Zgodnie z powyższym ujęciem interpretacja tego terminu jest, jak się okazuje, wyznaczona w niezwykle słaby i niejasny sposób. Nie jest określona jednoznacznie; nie jest nawet ograniczona do zbioru przedmiotów fizycznych: dopuszcza również pewne zbiory przedmiotów abstrakcyjnych, np. liczby. W konsekwencji okazuje się, że predykat „elektron" jest całkowicie nieostry w swojej dziedzinie właściwej, tzn. w dziedzinie przedmiotów nieobserwowalnych: dla każdego takiego obiektu nieobserwowalnego istnieje zawsze pewien model właściwy danego języka, który przypisuje ten obiekt denotacji danego predykatu, i taki, który tego nie czyni. Czy tak się rzecz ma z interpretacjami predykatów typu „elektron" we współczesnych teoriach fizycznych? Nie jestem w stanie przedstawić kompetentnej i szczegółowej analizy interpretacji takich terminów. Pewne ogólne rozważania skłaniają jednak, jak się wydaje, do twierdzącej odpowiedzi na to pytanie.

Interpretację terminu „elektron" w teoriach fizycznych charakteryzuje się zwykle przez odwołanie się do takich założeń, jak definiowalność „elektronu" za pomocą wielkości fizycznych takich jak masa, ładunek elektryczny itp., czy też rzekoma «obserwowalność» elektronów. Czy jest to zgodne z poglądem o cząstkowej interpretowalności terminu „elektron"? Czy możemy to wytłumaczyć w ramach naszej aparatury pojęciowej? Załóżmy, że elektron może być scharakteryzowany jako obiekt o pewnej masie, pewnym ładunku elektrycznym itp.; innymi słowy, że termin „elektron" może być zdefiniowany za pomocą terminów: „masa", „ładunek elektryczny" itp. Jakie pociąga to konsekwencje co do interpretacji tego terminu? Warto podkreślić, że terminy użyte do jego zdefiniowania nie należą do terminów obserwacyjnych: są albo terminami teoretycznymi, albo w najlepszym razie terminami mieszanymi. Rozważmy jako przykład możliwą konsekwencję takiej definicji, zgodnie z którą elektron jest obiektem o masie równej *k* (powiedzmy $9 \cdot 10^{-28}$g):

(I) $E(x) \rightarrow m(x) = k$.

Termin m jest typowym terminem mieszanym. Zgodnie z ogólnie przyjętym założeniem jego interpretacja jest scharakteryzowana przez zbiór postulatów następującego rodzaju:

(1) $R(x, y) \leftrightarrow m(x) \leq m(y)$,

(2) $m(x \, o \, y) = m(x) + m(y)$,

(3) $m(a) = 1$,

które wiążą termin ilościowy m z terminami jakościowymi R i o. Jednakże, jak już próbowałem pokazać gdzie indziej („Empirical meaningfulness of quantitative statements", *Synthese* 1974), dwa ostatnie terminy nie są terminami obserwacyjnymi. Są one również terminami mieszanymi i ich związek z terminami obserwacyjnymi przyjmuje postać pewnych zdań redukcyjnych:

(i) $O_1(x, y) \rightarrow (R(x, y) \leftrightarrow O_2(x, y))$,

(ii) $O_3(x, y, z) \rightarrow (x \, o \, y = z \leftrightarrow O_4(x, y, z))$.

Problem polega na tym, że postulaty tego rodzaju nie mogą ustalić dla terminów R i o żadnej określonej interpretacji poza dziedziną przedmiotów obserwowalnych, tzn. poza zbiorem $U_{\mathfrak{M}^*_o}$. W szczególności nie mogą one wykluczyć liczbowej interpretacji relacji R poza tą dziedziną. Charakterystyka ta przenosi się też na termin m jako termin zdefiniowany przez postulaty (1) - (3). Argumentami funkcji m mogą być między innymi liczby. W konsekwencji ani zdanie (I), ani żadna adekwatna definicja terminu E, z której ono wynika, nie może ustalić jednej, określonej fizycznej interpretacji terminu E i wykluczyć wszystkich innych niezamierzonych interpretacji, a w szczególności nie może wykluczyć interpretacji liczbowej.

Co się tyczy rzekomej «obserwowalności» elektronu, to jest oczywiste, że to określenie nie może być traktowane literalnie. Elektron nie jest obserwowalny w żadnym dosłownym tego słowa znaczeniu. To, co jest obserwowalne, to pewne makroobiekty, charakteryzowane za pomocą tego terminu, np. obiekty «zawierające wolne elektrony» lub tp. Tylko do takich obiektów odnoszą się bezpośrednio pewne dobrze znane kryteria obserwacyjne. W bardzo uproszczonej formie, kryterium tego typu może być przedstawione w następujący sposób:

(II) $O_5(x) \rightarrow \exists y \, (P(y, x) \wedge E(y))$.

Predykat obserwacyjny O_5 można traktować tutaj jako predykat opisujący pewien obserwowalny stan komory Wilsona, a predykat P — jako relację bycia częścią. Mimo obserwacyjnego charakteru, żadne kryterium typu (II) nie nadaje terminowi E ustalonej interpretacji fizycznej. Łatwo to zauważyć, jeżeli uświadomimy sobie, że termin P nie jest terminem obserwacyjnym lecz mieszanym, a więc tylko pośrednio i luźno związanym z terminami obserwacyjnymi. Postulaty dla P zwykle będą zawierały dwa rodzaje zdań: postulaty teoretyczne i reguły korespondencji. Pierwsze można traktować jako formuły składające się z aksjomatów mereologii (zawierających P jako jedyny termin pozalogiczny), drugie — jako pewne zdania redukcyjne typu:

$O_6(x, y) \rightarrow (P(x, y) \leftrightarrow O_7(x, y))$.

Nieobserwowalnymi desygnatami tak zinterpretowanego predykatu P mogą być również liczby. I takie właśnie obiekty będą w konsekwencji tworzyć pewne denotacje terminu E — mimo jego obserwacyjnych kryteriów.

Okazuje się więc, że przy naszych założeniach ani definiowalność, ani «obserwowalność» elektronów nie gwarantują temu terminowi interpretacji zwykle mu przypisywanej: przyporządkowania mu pewnego ustalonego zbioru obiektów fizycznych jako denotacji. Czy ten fakt podważa nasze założenia? Czy też możemy mimo to traktować takie założenia jako właściwe? Uważam, że możemy. Przemawiają za tym, jak się wydaje następujące argumenty. Chociaż termin „elektron" jest całkowicie nieostry (w dziedzinie przedmiotów nieobserwowalnych), jego interpretacja nie jest zupełnie arbitralna. Nie jest tak, że dowolny zbiór obiektów nieobserwowalnych może być mu przypisany jako denotacja. Klasa takich denotacji jest ograniczona do zbiorów, które mają pewne własności strukturalne (takie, jak np. niepustość), i które pozostają w pewnych określonych relacjach strukturalnych do właściwych denotacji innych terminów. Dzięki tym ograniczeniom, niektóre zdania zawierające w sposób istotny termin „elektron", stają się zdaniami rozstrzygalnymi empirycznie. To stwierdzenie wymaga pewnego komentarza. Mówimy, że zdanie α zawiera w istotny sposób termin E, jeśli jego wartość logiczna zależy od interpretacji tego terminu; tzn. jeśli istnieje taka właściwa interpretacja pozostałych terminów, że przy pewnej interpretacji terminu E zdanie α będzie prawdziwe, a przy innej — fałszywe. Aby wyjaśnić pojęcie empirycznej rozstrzygalności musimy najpierw zdefiniować pojęcie rozstrzygalności. Będziemy mówić, że zdanie α jest *rozstrzygalne* zawsze i tylko, gdy jest prawdziwe przy wszystkich jego właściwych interpretacjach, albo jest fałszywe przy wszystkich takich interpretacjach. Będziemy mówić, że α jest *empirycznie rozstrzygalne* zawsze i tylko, gdy jest rozstrzygalne i ani zdanie α, ani jego negacja nie są logicznymi konsekwencjami postulatów języka. To, które dokładnie zdania zawierające termin E w sposób istotny będą należeć do klasy zdań empirycznie rozstrzygalnych, zależy oczywiście od tego, jakie są postulaty dla terminu E. Załóżmy, że zawierają one postulat (II). Wtedy, jeśli tylko $O_5(a)$ jest prawdziwe dla pewnego a, zdania: $\exists y\,(P(y, a) \wedge E(y))$ i $\exists y\,E(y)$ okażą się empirycznie rozstrzygalne. Zgodnie z naszą interpretacją zamierzoną hipoteza ta jest oczywiście prawdziwa, a zdania powyższe głoszą co następuje: „ta komora Wilsona zawiera (wolne) elektrony" i „elektrony istnieją". Z drugiej strony zdanie postaci $E(a)$, gdzie a denotuje obiekt nieobserwowalny, może stanowić przykład zdania empirycznie nierozstrzygalnego — przy wszystkich dopuszczalnych interpretacjach terminu E. Nie ma w tym nic dziwnego. Zdania „a jest elektronem" nie używa się nigdy w praktyce naukowej. Sądzę, że klasa zdań rozstrzygalnych, zdefiniowanych jak wyżej, zawiera wszystkie zdania o elektronach faktycznie uznawane przez naukowców. Nie mogę uzasadnić tego poglądu w żaden przekonujący sposób. Wysuwam go tylko jako hipotezę.

Niezależnie od wszystkich przedstawionych tu argumentów trzeba stwierdzić, że wyjaśnienie sposobu interpretacji terminów teoretycznych, zaproponowane w niniejszym artykule, jest wyraźnie niezgodne z potocznym poglądem na tę sprawę. Ten ostatni jednak jest, jak się wydaje, nie do utrzymania. Wiara w to, że wszystkie terminy teoretyczne mają ustaloną empiryczną interpretację jest, moim zdaniem, iluzją. Nie widzę sposobu, w jaki wiara taka mogłaby zostać uzasadniona. Nie istnieje procedura «zmuszająca» zbiór elektronów do tego, aby stał się denotacją tego terminu. Co więcej, wydaje się, że nie istnieje procedura, która by zagwarantowała, że denotacja ta będzie ograniczona jedynie do obiektów fizycznych. Możemy oczywiście żądać (i faktycznie to robimy), aby elektrony były obiektami czasoprzestrzennymi, obdarzonymi masą itp. Ale wszystko to są słowa tylko. Charakteryzują one interpretację terminu „elektron" tylko w takim zakresie, w którym same są zinterpretowane. Kiedy postuluję, aby x był (nieobserwowalnym) obiektem obdarzonym pewną masą, to to, co robię faktycznie, sprowadza się do żądania, aby x był argumentem funkcji, która spełnia określone aksjomaty, i która w pewnej poddziedzinie (mianowicie w poddziedzinie rzeczy obserwowalnych) przyjmuje takie to a takie obiekty jako argumenty. Jak widać z powyższych przykładów, nie wyklucza się w ten sposób sytuacji, w której x jest obiektem abstrakcyjnym, np. liczbą. Jest to konsekwencja empirystycznych założeń. Nie widzę, jak można by je ominąć. Nasz jedyny kontakt z rzeczywistością przebiega na makropoziomie. Bezpośredni dostęp mamy wyłącznie do makroobiektów. Dostęp do mikroobiektów mamy jedynie za pośrednictwem słów. O obiektach nieobserwowalnych możemy tylko mówić. Nie można ich «uchwycić» w sposób niewerbalny. Nie ma też takiej potrzeby — jeżeli stawiamy przed sobą cele naukowe. Ta «nieuchwytnośc» nie przeszkadza obiektom nieobserwowalnym odgrywać istotnej roli w teoriach naukowych.

Omawiane przez nas stanowisko oparte jest nie tylko na pewnych założeniach filozoficznych, ale również na założeniach formalnych, dokładniej — teoriomodelowych. Założenia te są niekiedy westionowane przez tych, którzy przedstawione ujęcie uważają za nie do przyjęcia. Opis teoriomodelowy zwykle uważany jest za opis ekstensjonalny, i właśnie tę jego ekstensjonalność obwinia się za paradoksalne konsekwencje niniejszego stanowiska, a wyjście z trudności upatruje się w opisie intensjonalnym. Ten problem jest zbyt skomplikowany, aby analizować go w tym miejscu. A więc tylko krótki komentarz. Pojęcie intensji jest nieokreślone i niejasne, a te jego eksplikacje, które są dostatecznie precyzyjne, nie są, jak się wydaje, pomocne w wyjaśnieniu problemu interpretacji terminów teoretycznych. W szczególności dotyczy to eksplikacji dokonywanej na gruncie teorii modeli. Zgodnie z nią, intensja terminu może być zdefiniowana przez odwołanie się do klasy możliwych modeli danego języka («możliwych światów») — np. jako funkcja, która każdemu możliwemu modelowi przypisuje denotację danego terminu w tym modelu. Tego typu definicję pojęcia intensji można zastosować w naszym podejściu teoriomodelowym: wystarczy utożsamić możliwy model danego języka z modelem jego postulatów. Nie sądzę jednak, aby odwołanie się do takiego pojęcia mogło przyczynić się do powstania istotnie nowego i zadowalającego rozwiązania problemu interpretacji terminów teoretycznych.

Na koniec rozważmy zaproponowane w tym artykule ujęcie z punktu widzenia sporu między realizmem a instrumentalizmem. Czy można je nazwać realistycznym? A jeżeli tak, to w jakim sensie? Nie ma prostej odpowiedzi na to pytanie. Stanowisko to wydaje się realistyczne, ponieważ nadaje interpretację wszystkim terminom teoretycznym, a co się z tym wiąże, pozwala zdefiniować pojęcie prawdy w odniesieniu do wszystkich zdań teoretycznych. Wydaje się jednak instrumentalistyczne, jeśli chodzi o sposób tej interpretacji. Interpretacja terminu teoretycznego jest wyznaczona nie przez pojedynczą denotację zamierzoną, ale przez całą klasę takich denotacji, wśród których znajdują się denotacje ewidentnie niezamierzone. Zakłada się, że nie istnieje sposób identyfikacji denotacji zamierzonej, np. fizycznej, i odróżnienia jej od denotacji niezamierzonych, np. liczbowych. Każdemu terminowi teoretycznemu przyporządkowanych jest wiele różnych denotacji. Wynikiem ich różnorodności jest całkowita nieostrość danego terminu (w jego właściwym obszarze zastosowań). Mimo to skłonny jestem traktować to ujęcie jako w istocie realistyczne. Ale jeśli ktokolwiek ma poczucie, że powinno się je nazwać raczej instrumentalistycznym, nie będę protestował. Nie traktuję tej nazwy jako obraźliwej. Gdybym miał wskazać filozoficzne stanowisko leżące u podstaw tego ujęcia, wymieniłbym pogląd przedstawiony przez Quine'a w jego głośnym eseju „Ontological relativity". Nieunikniona wielość i różnorodność interpretacji wszystkich terminów teoretycznych odzwierciedla ontologiczny relatywizm naszej siatki pojęciowej, o czym tak przekonywająco pisał Quine.

Pozwolę sobie zakończyć te uwagi kilkoma słowami poświęconymi współczesnej krytyce jednego z głównych założeń leżących u podstaw prezentowanego w tym artykule stanowiska. Jest to założenie związane z rozróżnieniem pomiędzy pojęciami obserwacyjnymi i teoretycznymi, głoszące, iż pojęcia obserwacyjne występują w języku teorii empirycznych. Są to pojęcia obserwacyjne w dość ścisłym i absolutnym sensie — rozumiane jako terminy interpretowane przez pewne procedury ostensywne. Można postawić zarzut, że faktycznie takie terminy nie występują w języku teorii naukowych; ich język składa się wyłącznie z terminów teoretycznych, jeżeli te ostatnie będziemy rozumieć zgodnie z powyższym rozróżnieniem. W odpowiedzi na ten zarzut chciałbym uczynić dwie uwagi.

Zauważmy przede wszystkim, że kwestionowane założenie nie odgrywa istotnej roli w przedstawionej argumentacji. Można ją łatwo tak uogólnić, żeby całkowicie zrezygnować z pojęcia terminu obserwacyjnego. W tej uogólnionej formie argumentacja opiera się na następującym, mniej kontrowersyjnym stwierdzeniu:

Istnieją terminy empiryczne interpretowane jedynie w sposób pośredni, które desygnują obiekty nie desygnowane przez żaden termin interpretowany w sposób bezpośredni.

Założenie to wystarczy, aby uzasadnić główny wniosek.

Identyfikacja bezpośredniego sposobu interpretacji z procedurą ostensywną i, w konsekwencji, terminów interpretowanych w ten sposób z terminami obserwacyjnymi, stanowi pewien sposób wyjaśnienia powyższego założenia. Ale czy tak zinterpretowane założenie jest jeszcze do przyjęcia? Czy jakiekolwiek terminy obserwacyjne należą do języka rzeczywistych teorii naukowych? Wydaje się, że odpowiedź na to pytanie zależy od tego, z czym utożsamimy teorie naukowe. Przy pewnym wąskim rozumieniu teorii naukowej, większość teorii empirycznych (w szczególności wszystkie teorie fizyczne) nie zawiera niczego, co można by nazwać terminem obserwacyjnym w dosłownym znaczeniu. Ale przy szerszym rozumieniu, każda teoria empiryczna musi zawierać jakieś terminy obserwacyjne. Weźmy typowy przykład — mechanikę klasyczną. Co wchodzi w skład jej słownika? Zgodnie z obiegowym poglądem, jest to teoria zawierająca prawa Newtona jako jedyne aksjomaty fizyczne i symbole funkcyjne: „położenie", „masa" i „siła" — jako jedyne terminy fizyczne (tzn. pozalogiczne i pozamatematyczne). Żaden z tych terminów nie jest oczywiście terminem obserwacyjnym w przyjętym sensie. Teoria ta może być jednak rozumiana w inny, szerszy sposób. Zgodnie z rozważanym przez nas stanowiskiem utrzymuje się, że teoria ta nie zawiera żadnych podteorii, które leżą u jej podstaw — w szczególności żadnej teorii pomiaru jej podstawowych wielkości fizycznych. Te podteorie nie są częścią teorii właściwej. Teoria fizyczna, taka jak np. mechanika klasyczna, jest tu utożsamiana z najwyższą warstwą odpowiedniej struktury pojęciowej, jej najbardziej teoretycznym poziomem. Istnieje niewątpliwie wiele takich problemów, że do ich rozstrzygania takie pojęcie teorii fizycznej wydaje się odpowiednie. Jest jednak równie pewne, że istnieją inne problemy, dla których jest ono zbyt ograniczone. Jeżeli np. interesuje nas problem zawartości empirycznej danej teorii fizycznej, to nie możemy przy jego analizie abstrahować od teorii pomiaru, leżących u jej podstawy. Tylko rekonstrukcja procedur pomiarowych, charakterystycznych dla odpowiednich wielkości fizycznych, może wydobyć empiryczną treść podstawowych pojęć takiej teorii. Przy badaniu problemów tego rodzaju musimy traktować daną teorię w szerszy sposób, dołączając do jej teoretycznych praw tzw. reguły korespondencji. W teoriach fizycznych reguły korespondencji są zdaniami zawierającymi opis sposobu pomiaru odpowiednich wielkości fizycznych. Jeżeli rozumiemy teorię empiryczną w tak szeroki sposób (co istotnie robimy w tym artykule), mamy prawo założyć, że jej język zawiera pewne terminy obserwacyjne. Jeżeli przeprowadzimy analizę procedur pomiarowych dostatecznie głęboko, będziemy zmuszeni stwierdzić występowanie pewnych terminów obserwacyjnych w ścisłym, rozpatrywanym wyżej sensie tego słowa.

Trzeba jednak uczciwie powiedzieć, jaką cenę płacimy za możliwość prowadzenia tego rodzaju rozważań. Włączając do danej teorii, jako jej istotne części, wszystkie zakładane przez nią teorie pomiaru, zastępujemy dobrze zdefiniowaną, czystą strukturę przez niejasną i nieporęczną całość, której analiza stanowi niezwykle trudne i niewdzięczne zadanie. Tego zadania nie możemy uniknąć, jeśli zajmujemy się treścią, czy interpretacją, podstawowych pojęć tej teorii. Nie możemy wtedy traktować tej interpretacji jako czegoś danego; musimy przeanalizować sposób, w jaki powstaje. Trzeba przyznać, że wszystkie proponowane rozwiązania tego problemu są dalekie od ideału: są albo nieprecyzyjne, wyrażone przy pomocy luźnych, metaforycznych pojęć, albo nierealistyczne, oparte na zbytnich uproszczeniach i idealizacjach — albo mają obie te cechy. Moim zdaniem jednak, w odniesieniu do istotnych pytań teoretycznych lepsza jest zła odpowiedź niż żadna, ponieważ stanowi ona pobudkę do dalszych dociekań i ulepszeń.

Tłumaczyła Anna Lissowska

IDEALIZACJA

Filozofia Nauki
Rok I, 1993, Nr 2-3

Interpretacja systemów aksjomatycznych

1. Przedmiotem naszych rozważań będą systemy aksjomatyczne reprezentujące teorie empiryczne. Struktura takich systemów jest rzeczą znaną. System aksjomatyczny traktujemy jako zbiór konsekwencji logicznych pewnego — na ogół skończonego — układu aksjomatów. System taki opiera się na określonym rachunku logicznym. W aksjomatach systemu aksjomatycznego reprezentującego teorię empiryczną wyróżnić można zatem dwa rodzaje terminów: stałe logiczne należące do rachunku logicznego, na którym oparta jest dana teoria, oraz stałe specyficzne owej teorii, a więc pewne terminy fizykalne, biologiczne czy psychologiczne. Terminy logiczne traktować będziemy jako wyrażenia wyposażone w swój zwykły sens. Problem, który stoi przed nami, dotyczy sensu terminów specyficznych. System aksjomatyczny scharakteryzowany jest w sposób czysto formalny, nie odwołujący się do sensu terminów specyficznych. W jaki sposób zatem terminom tym nadane zostaje określone znaczenie? I to znaczenie empiryczne, pozwalające na stosowanie owych terminów na podstawie doświadczenia?[1].

Wśród terminów specyficznych teorii empirycznych wyróżnia się na ogół terminy pierwotne i pochodne. Pierwsze z nich — to terminy niezdefiniowane, występujące w aksjomatach teorii, drugie — to terminy wprowadzone na drodze definicji za pomocą terminów pierwotnych. Pytanie o sens terminów specyficznych danej teorii sprowadza się więc zwykle do pytania o sens terminów pierwotnych. Gdy bowiem znaczenie tych ostatnich jest ustalone, definicje określają znaczenie terminów pozostałych. W dalszych rozważaniach nie będziemy jednak czynili tego rozróżnienia. Definicje występujące w danej teorii traktować będziemy jako poszczególny przypadek aksjomatów. Mówiąc

1) Pojęcie empirycznego sensu terminów omawiam m.in. w pracy „Pojęcia teoretyczne a doświadczenie", *Studia Logica* 11.

zatem o układzie aksjomatów, nazwą tą obejmować będziemy zarówno aksjomaty właściwe, jak i definicje. Problem sensu empirycznego terminów danej teorii dotyczyć więc będzie ogółu jej terminów specyficznych. Odróżnienie aksjomatów od definicji nie zawsze zresztą jest rzeczą łatwą. Różnica między tymi dwoma rodzajami twierdzeń ma niewątpliwie charakter pragmatyczny. Sprowadza się chyba głównie do tego, iż twierdzeniami, które w procesie rozwoju teorii ulegają z reguły zmianie w przypadku sprzeczności z doświadczeniem, są aksjomaty, a nie definicje. Dla dalszych rozważań różnica ta nie jest ważna. Ważne okazuje się natomiast pojęcie definiowalności terminów wprowadzone przez Tarskiego. Jest to pojęcie czysto syntaktyczne. Zilustrujemy je na prostym przykładzie predykatu jednoargumentowego Q. Termin Q jest definiowalny w teorii T przez terminy $P_1, ..., P_n$ wtedy i tylko wtedy, gdy wyrażenie postaci:

$$(x)(Qx \equiv \Phi x),$$

gdzie Φ jest wyrażeniem zawierającym jako terminy pozalogiczne wyłącznie terminy $P_1, ..., P_n$ — jest twierdzeniem teorii T. Jak widać, jeśli dany termin jest terminem wprowadzonym do teorii na drodze definicji, jest on też terminem definiowalnym w tej teorii, lecz nie na odwrót.

Rozważając zagadnienie sensu terminów specyficznych teorii empirycznych ograniczymy się — zwłaszcza jeśli idzie o ilustrację wywodów ogólnych — do teorii możliwie prostych pod względem formalnym. Weźmiemy pod uwagę głównie teorie elementarne. Rachunek logiczny, na którym się teorie takie opierają, nie przekracza węższego rachunku funkcyjnego z identycznością. Ich język zatem obejmuje tylko jeden typ zmiennych — zmienne indywidualne, a spośród stałych logicznych — spójniki rachunku zdań, kwantyfikatory i znak identyczności. Ograniczenie do teorii elementarnych nie stanowi zresztą jakiejś drastycznej restrykcji, gdyż, jak wiadomo, wszelkie w zasadzie teorie dadzą się przedstawić w postaci elementarnej. Główny rodzaj terminów specyficznych występujących w tego typu teoriach — to jedno- i wieloargumentowe predykaty, denotujące klasy indywiduów i relacje zachodzące pomiędzy indywiduami. Na takich też terminach skoncentrujemy się w naszych rozważaniach, pomijając dla uproszczenia m.in. terminy funkcyjne. Wśród terminów specyficznych teorii empirycznych zwykło się wyróżniać dwie klasy wyrażeń: terminy elementarne i teoretyczne[2]. Rozróżnienie to, odwołujące się do sensu empirycznego terminów specyficznych, będzie w dalszych wywodach odgrywało istotną rolę. Jako niezmiernie prosty przykład teorii omawianego typu służyć może fikcyjna teoria skonstruowana przez Braithwaite'a[3]. Jest to teoria elementarna. Jej terminami specyficznymi są predykaty jednoargumentowe w liczbie sześciu: A, B, C, L, M, N; trzy pierwsze zaliczone są do terminów elementarnych, pozostałe — do teoretycznych. A oto układ aksjomatów tej teorii:

2) Por. „Pojęcia teoretyczne...", cyt. wyd.

3) Por. R. B. Braithwaite, *Scientific Explanation*, Cambridge 1953. Teorię tę omawiałem szerzej w pracy „Pojęcia empiryczne...", cyt. wyd.

$$(x)\,(Ax \equiv Lx \cdot Mx),$$
$$(x)\,(Bx \equiv Mx \cdot Nx),$$
$$(x)\,(Cx \equiv Nx \cdot Lx).$$

Oczywiście, przedstawiony system aksjomatyczny jest zbyt prosty, aby mógł reprezentować jakąś rzeczywistą teorię empiryczną. Ma on jednak, zdaniem autora, pewne ważne pod względem logicznym własności, charakterystyczne dla szeregu teorii przyrodniczych.

Na jakiej więc drodze terminy specyficzne podobnych teorii empirycznych wyposażone zostają w sens empiryczny? W jaki sposób terminom tym przyporządkowane zostają jako ich denotacje określone klasy czy relacje? Do tego ostatniego zresztą pytania sprowadzić chciałbym nasz problem. Chodzić w nim więc będzie nie o ustalenie znaczenia, lecz jedynie o ustalenie denotacji rozważanych terminów, o ich, jak to się często mówi, interpretację[4]. Interpretacja terminów specyficznych systemu aksjomatycznego pokrywa się z tym, co w języku współczesnej teorii systemów aksjomatycznych nosi nazwę konstrukcji modelu takiego systemu. Teoria modeli systemów aksjomatycznych stanowi więc dział logiki matematycznej mający bezpośrednie znaczenie dla wysuniętego przez nas problemu interpretacji terminów specyficznych teorii empirycznych. Nie mogę przytaczać na tym miejscu definicji podstawowych pojęć tej teorii, w szczególności ogólnego pojęcia modelu[5]. Przypomnę tylko w skrócie i uproszczeniu, o co chodzi w przypadku najprostszych teorii aksjomatycznych omawianych poprzednio. Niech $\Phi\,(P_1, ..., P_n)$ będzie koniunkcją aksjomatów teorii T o terminach specyficznych: $P_1, ..., P_n$. Niech $\mathfrak{M} = \langle U, R_1, ..., R_n\rangle$ będzie dziedziną składającą się z niepustego zbioru indywiduów U oraz z relacji $R_1, ..., R_n$ zachodzących pomiędzy indywiduami należącymi do zbioru U, przy czym jeśli P_i jest predykatem k-argumentowym, R_i jest relacją k-argumentową[6]. Dziedzinę $\mathfrak{M} = \langle U, R_1, ..., R_n\rangle$ nazywamy modelem zdania $\Phi\,(P_1, ..., P_n)$ wtedy i tylko wtedy, gdy zdanie to jest prawdziwe w dziedzinie $\mathfrak{M}$ przy rozumieniu predykatów $P_1, ..., P_n$ jako nazw relacji $R_1, ..., R_n$. Model zdania $\Phi\,(P_1, ..., P_n)$ uważać możemy za model teorii T, gdyż z określenia modelu i konsekwencji logicznej wynika, iż dziedzina będąca modelem koniunkcji aksjomatów danej teorii jest zarazem modelem ogółu jej twierdzeń. Tak pojęty model teorii T stanowi interpretację jej terminów specyficznych. Zbiór U odpowiada universum teorii, czyli zbiorowi przedmiotów przebieganemu przez jej zmienne, a relacje $R_1, ..., R_n$ zachodzą-

4) Terminem „interpretacja" nazywać będę w dalszym ciągu zarówno czynność interpretowania, jak i rezultat tej czynności.

5) W sposób przystępny przedstawia pewne podstawowe pojęcia teorii modeli R. Suszko w pracy „Logika formalna a niektóre zagadnienia teorii poznania", *Myśl Filozoficzna* 1957, 2(28)-3(29).

6) W przypadku predykatu jednoargumentowego odpowiadająca mu relacja jednoargumentowa utożsamiona może być z klasą.

ce pomiędzy elementami zbioru U — denotacjom terminów specyficznych teorii $P_1, \ldots, P_n$. Podając pewien model teorii empirycznej podajemy tym samym pewną interpretację jej stałych specyficznych.

Zanim przejdziemy do analizy sposobów nadawania interpretacji terminom specyficznym systemu aksjomatycznego, zwróćmy uwagę na fakt, iż istnieje taki sposób ujęcia systemów aksjomatycznych, przy którym problem ów w ogóle nie powstaje[7]. Mam na myśli ujęcie zwane formalnym. Ujęcie to charakteryzuje się właśnie tym, iż terminy specyficzne pozostają niezinterpretowane. Zastępujemy je zmiennymi wolnymi, przez co zarówno aksjomaty, jak i twierdzenia przybierają charakter formuł zdaniowych, a nie zdań. Jeśli $\Phi (P_1, \ldots, P_n)$ jest koniunkcją aksjomatów teorii T, a $\Psi (P_1, \ldots, P_n)$ — jednym z jej twierdzeń, to przy ujęciu formalnym zdaniom tym odpowiadają formuły ze zmiennymi wolnymi: $\Phi (X_1, \ldots, X_n)$ i $\Psi (X_1, \ldots, X_n)$. Oczywiście, ani jedne, ani drugie nie mogą mieć charakteru tez. Jako tezy mogą być traktowane co najwyżej zdania stwierdzające wynikanie jednych formuł z innych. Zdania takie interpretuje się przy tym rozmaicie. Bądź jako twierdzenia sformułowane w języku danej teorii o postaci:

$$(X_1) \ldots (X_n) (\Phi (X_1, \ldots, X_n) \supset \Psi (X_1, \ldots, X_n));$$

bądź jako twierdzenia należące do metateorii o postaci:

Formuła „$\Psi (X_1, \ldots, X_n)$" wynika logicznie z formuły „$\Phi (X_1, \ldots, X_n)$".

Dla dalszych rozważań różnice te są nieistotne. Decydujące jest to, że przy formalnym ujęciu systemu aksjomatycznego, przy którym jego terminy specyficzne pozostają niezinterpretowane, system taki, pozbawiony całkowicie terminów o sensie empirycznym, nie może stanowić teorii o charakterze empirycznym.

Przejdźmy zatem do takich ujęć systemów aksjomatycznych, które zapewniają pewną interpretację ich terminów specyficznych. Przy ujęciach tych aksjomaty i twierdzenia systemu przybierają charakter zdań, a nie formuł zdaniowych. Mogą, co za tym idzie, być traktowane jako tezy. Tutaj dopiero nasz problem staje się aktualny. Przyglądając się próbom jego rozwiązania wyróżnić możemy dwa główne sposoby, w jakie usiłowano zdać sprawę z interpretacji terminów specyficznych. Nazwijmy je interpretacją intra— i ekstrasystemową. W odróżnieniu tym idzie, najogólniej mówiąc, o to, czy interpretacja owa wyznaczona jest przez twierdzenia samego systemu, czy też dana jest «z zewnątrz». Interpretacja intrasystemowa terminów specyficznych wyznaczona ma być przez aksjomaty systemu, w których owe terminy figurują, interpretacja ekstrasystemowa — przez „reguły semantyczne", a więc przez pewne twierdzenia metasystemu. Rozpatrzmy kolejno obie ewentualności.

2. Interpretacja intrasystemowa odpowiada temu, co zwykło się nazywać definicją (lub pseudodefinicją) przez postulaty. Pogląd, iż to na tej właśnie drodze terminy specyficzne uzyskują znaczenie, formułuje się często mówiąc, iż „aksjomaty konstytu-

7) Podkreśla to m.in. M. Kokoszyńska w odczycie pt. *Wprowadzanie pojęć w systemach aksjomatycznych drogą definicji* wygłoszonym na zebraniu naukowym Zakładu Logiki PAN, 1958. W sprawie różnych ujęć systemów aksjomatycznych por. również K. Ajdukiewicz, „The Axiomatic Systems from the Methodological Points of View", *Studia Logica* 9, 1960.

ują znaczenie terminów pierwotnych". Owo „konstytuowanie znaczenia" sprowadzać się ma do przyjęcia umowy terminologicznej nakazującej rozumienie terminów specyficznych jako nazw o takim znaczeniu, przy którym aksjomaty teorii stają się zdaniami prawdziwymi. Ponieważ interpretację terminów specyficznych pojęliśmy jako ustalenie ich denotacji jedynie, a nie znaczenia, interpretację intrasystemową sprowadzić możemy do przyjęcia umowy terminologicznej nakazującej rozumienie terminów specyficznych jako nazw takich przedmiotów, które spełniają aksjomaty teorii. Interpretacja ta polega zatem na przyporządkowaniu terminom specyficznym jako denotacji przedmiotów spełniających układ aksjomatów. Zabieg taki sprowadza się, jak widzieliśmy, do konstrukcji modelu danej teorii. Wobec tego pytanie, czy terminy specyficzne danej teorii empirycznej mogą uzyskać na tej drodze określoną — a więc jednoznaczną — interpretację, sprowadza się do problemu, czy dana teoria ma jeden i tylko jeden model. Innymi słowy: czy istnieje dziedzina, w której aksjomaty teorii są prawdziwe, i czy tylko jedna dziedzina czyni zadość temu warunkowi? Jeśli tak, to ów jedyny model stanowi interpretację terminów specyficznych teorii wyznaczoną przez układ aksjomatów, w których te terminy figurują. Jeśli nie, to wyznaczenie określonego modelu, a z nim i określonej interpretacji terminów specyficznych, nastąpić musi «z zewnątrz». Rozstrzygnięcie tego problemu umożliwiają pewne wyniki osiągnięte w teorii modeli. Omawiając je ograniczymy się do najprostszego typu teorii, opisanego poprzednio.

Sprawa istnienia co najmniej jednego modelu danej teorii rozstrzygnięta jest pozytywnie. Każda niesprzeczna teoria ma model. Walor ma przy tym również pewne twierdzenie mocniejsze. Spośród ogółu modeli teorii zawierającej znak identyczności wyróżnić można modele charakteryzujące się tym, że znak ów interpretowany jest istotnie jako relacja identyczności pomiędzy elementami universum, a nie jako dowolna relacja spełniająca aksjomaty identyczności. Modele takie nazywamy modelami z absolutnym pojęciem identyczności. Ten też typ modeli będziemy mieli przede wszystkim na uwadze w dalszych rozważaniach. Otóż prawdą jest, iż każda niesprzeczna teoria (z identycznością) ma model z absolutnym pojęciem identyczności.

W przeciwieństwie do sprawy istnienia co najmniej jednego modelu danej teorii, problem istnienia tylko jednego modelu rozstrzygnięty jest na gruncie teorii modeli w sposób negatywny. Decydujące jest tutaj znane twierdzenie o izomorfizmie. Żadna teoria nie może mieć tylko jednego modelu. Jeżeli dziedzina $\mathfrak{M}'$ jest modelem teorii T, to jest nim również każda dziedzina $\mathfrak{M}''$ izomorficzna z $\mathfrak{M}'$. A zatem maksimum tego, czego oczekiwać możemy od teorii, to — żeby wszystkie jej modele były izomorficzne. Teorię spełniającą ten warunek nazywamy teorią kategoryczną. Z faktu tego płyną doniosłe konsekwencje dla problemu interpretacji intrasystemowej terminów specyficznych teorii empirycznej. Spróbujmy uprzytomnić je sobie obecnie. Przypomnimy w tym celu przede wszystkim pojęcie izomorfizmu dwóch dziedzin. Niechaj będą nimi dziedziny: $\mathfrak{M}' = \langle U', R_1', ..., R_n' \rangle$ i $\mathfrak{M}'' = \langle U'', R_1'', ..., R_n'' \rangle$. Dziedziny te są izomorficzne, gdy istnieje wzajemnie jednoznaczna relacja R przyporządkowująca elementom universum U' elementy universum U'' w sposób taki, iż relacja R_i' zachodzi między elementami universum U' wtedy i tylko wtedy, gdy relacja R_i'' zachodzi między przyporządkowanymi tamtym elementom elementami universum U''. Niech Φ $(P_1, ..., P_n)$ będzie, jak poprzednio, koniunkcją aksjomatów teorii T. Twierdzenie o izomorfizmie głosi, iż jeśli dziedzina $\mathfrak{M}'$ jest modelem teorii T, dziedzina $\mathfrak{M}''$ jest nim również. Układ aksjomatów teorii T wyznacza więc nie jeden model tej teorii, lecz całą ich rodzinę. W najlepszym razie — w przypadku kategoryczności teorii T — wszystkie modele należące do tej rodziny okazują się wzajemnie izomorficzne.[8] Układ aksjomatów ogranicza wtedy co prawda możliwości interpretacyjne terminów specyficznych teorii, ale — nie mówiąc już o tym, iż nie ogranicza tych możliwości do jedynej — czyni to pod pewnym tylko względem. Nie przyporządkowuje danemu terminowi specyficznemu określonej denotacji, lecz dowolną denotację spośród niezmiernie obszernej klasy denotacji, których jedyną wspólność stanowi fakt ich wzajemnej izomorficzności.[9]

Weźmy dla przykładu sytuację najprostszą: teorię aksjomatyczną T o predykacie jednoargumentowym P jako jedynym terminie specyficznym. Niech $\Phi(P)$ będzie koniunkcją aksjomatów tej teorii, a $\mathfrak{M}' = \langle U', K' \rangle$, gdzie K' jest pewną klasą elementów universum U' — jednym z jej modeli. Modelem teorii T będzie również każda dziedzina $\mathfrak{M}'' = \langle U'', K'' \rangle$ izomorficzna z $\mathfrak{M}'$, a przy założeniu, iż T jest teorią kategoryczną — tylko taka dziedzina. Układ aksjomatów teorii T przyporządkowuje więc predykatowi P jako denotację dowolną klasę izomorficzną z klasą K'. Ale izomorfizm klas sprowadza się, jak wiadomo, do ich równoliczności. Denotacją predykatu P może być więc jakakolwiek klasa równoliczna z klasą K'. Aksjomaty teorii T determinują klasę denotowaną przez P tylko co do liczby jej elementów. Nie wiemy, które przedmioty podpadają pod termin P. Wiemy tylko, iż przedmiotów tych jest pewna określona liczba. Aksjomaty te zatem nie tylko nie wyznaczają jednoznacznie denotacji terminu P, ale denotację tę charakteryzują w sposób bardzo jednostronny. Trudno w tej sytuacji twierdzić, iż aksjomaty teorii T wyznaczają interpretację jej terminu specyficznego, a tym bardziej — iż „konstytuują jego znaczenie".

W przypadku teorii aksjomatycznych o innych rodzajach terminów specyficznych sytuacja jest analogiczna. Aksjomaty tych teorii determinują denotacje terminów specyficznych pod pewnym tylko względem. Załóżmy, tak jak dotychczas, iż terminami tymi są k-argumentowe predykaty, denotujące k-argumentowe relacje. Mówimy ogólnie, iż dwie relacje mają tę samą strukturę, wtedy i tylko wtedy, gdy są wzajemnie izomorficzne. Aksjomaty teorii kategorycznych determinują zatem strukturę denotacji terminów

8) A. Tarski wprowadził w pracy „Z badań metodologicznych nad definiowalnością terminów", *Przegląd Filozoficzny* 37, 1934, pojęcie jednoznacznej kategoryczności, węższe od pojęcia kategoryczności. Teoria T jest kategoryczna, jeśli dla dwóch dowolnych modeli teorii T istnieje co najmniej jedna relacja ustalająca izomorfizm tych modeli; teoria T jest natomiast jednoznacznie kategoryczna, jeśli dla dwóch dowolnych modeli teorii T istnieje jedna i tylko jedna taka relacja.

9) Podkreślali to m. in. R. Suszko w pracy „Logika formalna..." cyt. wyd., M. Kokoszyńska we wspomnianym odczycie, H. Mehlberg w książce *The Reach of Science*, Toronto, 1958.

specyficznych. Taką strukturę relacji jednoargumentowych, czyli klas, stanowi właśnie ich liczba kardynalna. Przykładem struktury relacji dwuargumentowych może być np. progresja. Tak więc tym, co w najlepszym razie — w przypadku kategoryczności teorii — zostaje wyznaczone przez układ aksjomatów, jest jedynie struktura denotacji terminów specyficznych. Widzieliśmy na przykładzie predykatu jednoargumentowego, iż nie można tu mówić o wyznaczeniu interpretacji.

Wzajemny izomorfizm wszystkich modeli danej teorii aksjomatycznej istnieje tylko w przypadku kategoryczności teorii. Powstaje wobec tego pytanie, czy rozpatrywane przez nas aksjomatyczne teorie empiryczne należą do teorii kategorycznych. Pewne rezultaty osiągnięte w teorii modeli implikują częściowe odpowiedzi na to pytanie, przesądzając je w zasadzie w sposób negatywny. I tak, okazuje się, iż każda teoria elementarna na modele nieizomorficzne. Żadna teoria elementarna nie jest więc teorią kategoryczną. Płynie to z faktu, iż każda teoria elementarna ma modele dowolnie wysokich mocy,[10] a modele o różnej mocy nie są modelami izomorficznymi. Nawet przy ograniczeniu się do modeli z absolutnym pojęciem identyczności okazać można, iż każda teoria elementarna mająca model nieskończony ma modele nieizomorficzne z absolutnym pojęciem identyczności. Teorią kategoryczną w tym rozumieniu może być więc tylko taka teoria elementarna, która ma wyłącznie modele skończone. Teoria taka zakładać musi istnienie skończonej, np. nie większej od k, liczby przedmiotów składających się na jej uniwersum. Konsekwencją jej aksjomatów musi być twierdzenie ograniczające do k liczbę indywiduów, o których w danej teorii mowa. Wydaje się, iż teorie empiryczne — fizykalne czy biologiczne — z reguły tego warunku nie spełniają. Nie mogą to być więc teorie kategoryczne, o ile oczywiście należą do teorii elementarnych. A wspomniałem już na wstępie, iż każda w zasadzie teoria dopuszcza możliwość formalizacji w postaci elementarnej. Teorie empiryczne z reguły zatem są teoriami niekategorycznymi.[11] Oczywiście i przytaczana wyżej fikcyjna teoria Braithwaite'a należy do teorii niekategorycznych. Łatwo podać dwa nieizomorficzne modele tej teorii. Niechaj będą nimi dziedziny: $\mathfrak{D}' = \langle N^*, A', B', C', L', M', N' \rangle$ oraz $\mathfrak{D}'' = \langle N^*, A', B', C', L'', M'', N'' \rangle$, gdzie N^* jest zbiorem liczb naturalnych, $A' = \{2, 3\}$, $B' = \{2, 4\}$, $C' = \{1, 2\}$, $L' = \{1, 2, 3\}$, $M' = \{2, 3, 4\}$, $N' = \{1, 2, 4\}$, $L'' = \{1, 2, 3, 5\}$, $M'' = \{2, 3, 4, 6\}$, $N'' = \{1, 2, 4, 7\}$. Dziedziny $\mathfrak{D}'$ i $\mathfrak{D}''$ stanowią, jak się łatwo przekonać, modele układu aksjomatów owej teorii. Nie są to przy tym modele izomorficzne (ze względu na nierównoliczność klas: L' i L'', M' i M'' oraz N' i N'', stanowiących denotacje terminów specyficznych L, M, N).

10) Przez moc modelu rozumie się moc (liczbę kardynalną) jego universum.

11) J. Łoś wprowadził w pracy „On the Categoricity in Power of Deductive Theories and Some Related Problems", *Collequim Mathematicum* 3 1954, pojęcie kategoryczności w mocy, szersze od pojęcia kategoryczności. Teoria T jest kategoryczna w mocy m, jeśli wszystkie modele teorii T o mocy m z absolutnym pojęciem identyczności są wzajemnie izomorficzne. Pojęcie to nie jest puste. Istnieją teorie elementarne kategoryczne w pewnej mocy. Wydaje się jednak, iż w przypadku rozważanych przez nas teorii empirycznych pojęcie to nie znajduje większego zastosowania.

Tak więc układ aksjomatów teorii empirycznej nie tylko nie wyznacza jednego modelu tej teorii, lecz z reguły wyznacza taką rodzinę modeli, w której skład wchodzą modele wzajemnie nieizomorficzne. Do rodziny tej należą przy tym zawsze — również w przypadku teorii kategorycznej — modele odznaczające się tym, iż ich uniwersum składa się z wyrażeń należących do języka danej teorii, oraz modele takie, których universum stanowi zbiór liczb naturalnych! Trudno wobec tego stanu rzeczy utrzymywać, iż układ aksjomatów wyznacza to, o czym mowa w danej teorii, w szczególności — denotacje jej terminów specyficznych. Denotacje te scharakteryzowane zostają pod pewnymi tylko względami. Aksjomaty teorii wyznaczają niektóre ze strukturalnych własności tych denotacji, składających się na ich strukturę. Taką strukturalną własnością relacji jest np. symetryczność czy przechodniość. Za pomocą żadnego układu aksjomatów niepodobna jednak przyporządkować danemu terminowi jego denotacji w sposób taki, który by pozwalał o jakimś przedmiocie stwierdzić, czy pod ten termin podpada. Trzeba tę sytuację odróżnić od sytuacji nieostrych terminów empirycznych, których denotacje również nie są wyznaczone w sposób jednoznaczny. Przejawem tego jest fakt, iż o pewnych przedmiotach nie sposób rozstrzygnąć, czy pod taki termin podpadają. Tam jednak istnieją przedmioty, w stosunku do których stwierdzenie takie jest możliwe. Tutaj przedmiotów takich nie ma. Można powiedzieć, iż mamy tu do czynienia z terminami całkowicie nieostrymi. Termin specyficzny wyposażony jedynie w interpretację intrasystemową uznać musimy za pozbawiony sensu empirycznego — jakkolwiek byśmy ten ostatni określili. Na drodze interpretacji intrasystemowej niepodobna zapewnić terminom specyficznym teorii — a więc i samej teorii — charakteru empirycznego.[12] Wniosek ten wydaje się oczywisty, jeśli zważymy, iż jedynymi terminami figurującymi w aksjomatach teorii i mającymi z góry określony sens — są stałe logiczne. A za pomocą samych terminów logicznych nie można określić żadnego terminu o sensie empirycznym.[13]

3. Interpretację zapewniającą terminom specyficznym teorii charakter empiryczny nadać zatem można jedynie «z zewnątrz». Spośród całej rodziny modeli danej teorii aksjomatycznej wskazać trzeba w jakiś sposób ten, który ma stanowić jej model właściwy.[14] Realizuje się to za pomocą tzw. reguł semantycznych, które poszczególnym terminom specyficznym teorii przyporządkowują określone denotacje. Reguły te oczy-

12) Do podobnej konkluzji dochodzi H. Mehlberg w cytowanej książce.

13) Powyższą sytuację zilustrować można również za pomocą tzw. pojęcia wyraźnego (*explicit concept*) teorii, omawianego np. przez R. Carnapa w *Introduction to Symbolic Logic and its Applications*, New York 1958. Jeśli koniunkcją aksjomatów teorii T jest zdanie Φ (P_1, ..., P_n), definicję pojęcia wyraźnego R teorii T sformułować można następująco:

$$R(X_1, ..., X_n) \equiv \Phi(X_1, ..., X_n).$$

Ponieważ wyrażenie $\Phi(X_1, ..., X_n)$ zawiera jedynie stałe logiczne, R jest również stałą logiczną. W przypadku omawianej teorii Braithwaite'a owym pojęciem wyraźnym jest relacja zachodząca pomiędzy dowolnymi sześcioma klasami takimi, iż pierwsze trzy są iloczynami logicznymi trzech różnych par klas pozostałych. Mówiąc swobodnie, teoria Braithwaite'a pozbawiona jakiejś interpretacji ekstrasystemowej jest teorią dowolnych klas spełniających powyższy warunek.

14) Sposób ten charakteryzuje z pewnego punktu widzenia R. Suszko w pracy „Logika formalna..." cyt. wyd.

wiście nie są twierdzeniami danej teorii, lecz należą do twierdzeń metateorii. Niech terminami specyficznymi teorii *T* będą, jak poprzednio, predykaty $P_1, ..., P_n$. Reguły semantyczne wyznaczające interpretację teorii *T* przybierać mogą np. postać następującą:

Predykat „P_i” denotuje relację R_i.

Reguła taka przyporządkowuje predykatowi P_i określoną denotację pod warunkiem, iż termin metateorii R_i już taką denotację ma. A zatem reguły semantyczne zapewniają terminom specyficznym teorii określoną interpretację tylko wtedy, gdy w metateorii istnieją terminy (proste lub złożone), które już taką interpretację mają. Takie postawienie sprawy stanowi, jak widać, nie tyle rozwiązanie, ile raczej przesunięcie naszego problemu. Zapytać bowiem musimy z kolei, w jaki sposób terminy owej metateorii uzyskują określoną interpretację. A tutaj podobne zarysowują się możliwości, jak w przypadku teorii pierwotnej.

W jaki sposób zatem można zdać sprawę z operacji nadawania określonej interpretacji terminom specyficznym teorii empirycznej? Przede wszystkim chciałbym zwrócić uwagę na fakt, iż terminy te są przez układ aksjomatów, w których figurują, wzajemnie ze sobą powiązane. Owe związki logiczne przybierać mogą postać różnorodną. W przypadku skrajnym sprowadzać się mogą do tego, iż pewne spośród terminów specyficznych są definiowalne przez terminy pozostałe. Aby nadać ogółowi terminów specyficznych określoną interpretację, wystarczy zatem na ogół zinterpretować za pomocą reguł semantycznych tylko niektóre spośród nich, a pozostałe uzyskają wówczas interpretację pośrednią dzięki związkom logicznym, jakie między nimi na gruncie danego układu aksjomatów zachodzą. Jak jednak nadajemy interpretację owym terminom wybranym? Tutaj przypomnieć chciałbym wspomniany już na wstępie fakt istnienia wśród ogółu terminów specyficznych teorii empirycznych dwóch klas wyrażeń: terminów elementarnych i teoretycznych. Ograniczając się do terminów o charakterze predykatów, terminy elementarne określić możemy jako predykaty denotujące spostrzegalne własności lub stosunki, terminy teoretyczne — jako predykaty denotujące własności lub stosunki niedostępne bezpośredniej obserwacji. Przykładem pierwszych może być termin „żółty”, przykładem drugich — termin „gen”. Wydaje się, iż każda teoria empiryczna zawiera terminy specyficzne obu rodzajów.[15] W szczególności każda taka teoria zawiera pewne terminy spostrzeżeniowe. Terminy te najczęściej nie należą do terminów pierwotnych danej teorii, lecz wprowadzone zostają za pomocą długiego nieraz łańcucha definicji. Ponieważ jednak definicje traktujemy tutaj jako poszczególny rodzaj aksjomatów i, w związku z tym, pod uwagę bierzemy ogół terminów specyficznych teorii, a nie tylko jej terminy pierwotne, mamy prawo przyjąć, iż wśród tych terminów specyficznych znajdują się terminy spostrzeżeniowe. Terminy te, odnoszące się do przedmiotów spostrzegalnych, dopuszczają, jak się na ogół przyjmuje, interpretację bezpośrednią. Ów spostrzegalny przedmiot mający stanowić denotację

15) Por. „Pojęcia teoretyczne...” cyt. wyd.

terminu spostrzeżeniowego można przyporządkować temu terminowi nie za pośrednictwem innego terminu denotującego ów przedmiot, lecz przez bezpośrednie wskazanie samego przedmiotu. Reguła semantyczna wyznaczająca to przyporządkowanie utożsamiona być może z tzw. definicją dejktyczną (lub ostensywną) danego terminu spostrzeżeniowego. Nie prowadzi ona do konsekwencji wskazanych poprzednio, gdyż nie wymaga dla swego sformułowania istnienia terminu o takiej denotacji, jaką chcemy przyporządkować terminowi interpretowanemu. Owa bezpośrednia interpretacja terminów spostrzeżeniowych nasuwa szereg wątpliwości, które rozważam na innym miejscu.[16] Wątpliwości te powstają głównie w przypadku takich terminów spostrzeżeniowych, jak terminy rozważane przez nas obecnie. Nie są to nazwy indywiduowe, lecz predykaty. Ich denotacjami nie są więc jednostkowe rzeczy, lecz własności takich rzeczy lub zachodzące pomiędzy nimi stosunki. Tymczasem, ściśle biorąc, spostrzegać czy wskazywać możemy tylko jednostkowe rzeczy. Owa bezpośrednia interpretacja predykatów spostrzeżeniowych przedstawia się więc o wiele bardziej skomplikowanie, niż się to na ogół sugeruje.[17] Zakładamy jednak tutaj, iż możliwość takiej interpretacji istnieje i że rozpatrywane przez nas teorie empiryczne zawierają terminy specyficzne, które taką interpretację dopuszczają. Za takie terminy uważać będziemy wszystkie elementarne terminy teorii. Natomiast terminy teoretyczne, odnoszące się do przedmiotów niespostrzegalnych, traktować będziemy jako terminy, które bezpośredniej interpretacji nie dopuszczają.

Operację nadawania określonej interpretacji terminom specyficznym teorii empirycznej przedstawić możemy teraz jak następuje. Załóżmy, iż terminami specyficznymi teorii T są predykaty elementarne $P_1, ..., P_n$ oraz predykaty teoretyczne $Q_1, ..., Q_m$. Niech $\Phi (P_1, ..., P_n, Q_1, ..., Q_m)$ stanowi koniunkcję aksjomatów teorii T. Wszystkie predykaty elementarne mają interpretację ekstrasystemową. Reguły semantyczne w postaci definicji dejktycznych przyporządkowują im jako denotacje określone obserwowalne relacje $R_1, ..., R_n$. Żaden z predykatów teoretycznych nie ma interpretacji ekstrasystemowej. Interpretację predykatów teoretycznych zapewniają im wyłącznie związki logiczne z predykatami elementarnymi implikowane przez aksjomaty teorii. Interpretacja predykatów teoretycznych ma zatem charakter intrasystemowy.[18]

---

16) Por. „O terminach spostrzeżeniowych" (przygotowane do druku).

17) Wnikliwą analizę tego zabiegu zawiera praca J. Kotarbińskiej pt. „Tak zwana definicja dejktyczna", *Fragmenty filozoficzne*, seria 2, Warszawa 1959. Jednym z jej rezultatów jest stwierdzenie, iż owa bezpośrednia interpretacja terminów spostrzeżeniowych z reguły nie przyporządkowuje tym terminom denotacji w sposób jednoznaczny. Przejawem tego jest notoryczna nieostrość terminów spostrzeżeniowych.

18) Podobnie stawia sprawę R. Carnap w pracy *Foundations of Logic and Mathematics*, Chicago 1939. Stanowisko Carnapa precyzuje R. Montague przez wprowadzenie w pracy *Deterministic Theories* (egzemplarz powielony) pojęcia modelu częściowego (*partial model*). Niechaj $\mathfrak{M} = \langle U, R_1, ..., R_n, S_1, ..., S_m \rangle$ będzie modelem powyższej teorii T. Modelem częściowym odpowiadającym modelowi $\mathfrak{M}$ jest ciąg denotacji terminów elementarnych $\langle R_1, ..., R_n \rangle$.

Powstaje natychmiast kwestia, czy terminy teoretyczne mogą na tej drodze otrzymać interpretację jednoznaczną. Czy aksjomaty teorii, w których terminy elementarne mają z góry określone denotacje, wyznaczają w sposób jednoznaczny denotację terminów teoretycznych? Odpowiedź na to pytanie zależy najwyraźniej od charakteru owych aksjomatów, w szczególności — od rodzaju związków logicznych, jakie owe aksjomaty ustanawiają pomiędzy terminami elementarnymi a terminami teoretycznymi. Istnieje w teorii modeli twierdzenie, które ma dla tej sprawy znaczenie decydujące. Przyjmijmy dla uproszczenia, iż w teorii T występuje tylko jeden termin teoretyczny Q oraz n terminów elementarnych $P_1, ..., P_n$. Niech $\Phi(P_1, ..., P_n, Q)$ będzie koniunkcją jej aksjomatów. Twierdzenie, o którym mowa głosi, iż termin Q jest definiowalny w teorii T przez terminy $P_1, ..., P_n$ wtedy i tylko wtedy, gdy dla każdej dziedziny $\mathfrak{M} = \langle U, R_1, ..., R_n, S', S''\rangle$ prawdą jest, że jeżeli $\mathfrak{M}' = \langle U, R_1, ..., R_n, S'\rangle$ jest modelem teorii T i $\mathfrak{M}'' = \langle U, R_1, ..., R_n, S''\rangle$ jest modelem teorii T, to relacja S' jest identyczna z relacją S''. A zatem wtedy i tylko wtedy, gdy termin teoretyczny Q jest definiowalny w teorii T przez terminy elementarne $P_1, ..., P_n$ przyporządkowanie denotacji terminom elementarnym $P_1, ..., P_n$ wyznacza w sposób jednoznaczny denotację terminu teoretycznego Q. Mówiąc ogólnie, wtedy i tylko wtedy, gdy terminy teoretyczne są definiowalne w danej teorii empirycznej przez terminy elementarne, interpretacja tych ostatnich gwarantuje jednoznaczną interpretację terminów teoretycznych.

Jakże więc przedstawia się pod tym względem sytuacja w rozważanych obecnie teoriach empirycznych? Czy teorie te implikują definiowalność terminów teoretycznych przez terminy elementarne? Zagadnieniem tym zajmowałem się w pracy „Pojęcia teoretyczne a doświadczenie", poświęconej szczegółowej analizie związków łączących terminy teoretyczne z terminami elementarnymi. Wyniki tych rozważań wydają się przesądzać sprawę w sposób negatywny. Terminy teoretyczne z reguły nie są definiowalne na gruncie teorii empirycznych przez terminy elementarne. Związki, jakie aksjomaty teorii empirycznych ustanawiają pomiędzy tymi dwoma rodzajami terminów, mają charakter luźniejszy. Nie sprowadzają się do definicji terminów teoretycznych za pomocą terminów elementarnych. To raczej te ostatnie bywają definiowalne przez terminy teoretyczne. A zatem fakt bezpośredniej interpretacji terminów elementarnych nie pociąga za sobą jednoznacznej interpretacji terminów teoretycznych. Aksjomaty, w których terminy elementarne mają z góry wyznaczone denotacje, nie determinują w sposób jednoznaczny denotacji terminów teoretycznych. Weźmy dla przykładu omawianą teorię Braithwaite'a. W teorii tej terminy teoretyczne L, M, N nie są definiowalne przez terminy elementarne A, B, C. Aksjomaty tej teorii nie implikują żadnych twierdzeń, które by miały postać definicji zupełnych terminów teoretycznych za pomocą terminów elementarnych.[19] Toteż można z łatwością znaleźć dla tej teorii interpretacje takie, w których przy identycznej interpretacji terminów elementarnych terminy teoretyczne interpretowane są różnie. Takie właśnie interpretacje stanowią dwa

19) Por. „Pojęcia teoretyczne..." cyt. wyd.

opisane wyżej nieizomorficzne modele tej teorii: $\mathcal{M}'$ i $\mathcal{M}''$. W modelach tych terminom elementarnym A, B, C przyporządkowane zostały jako denotacje klasy identyczne (zarówno w modelu $\mathcal{M}'$, jak i w modelu $\mathcal{M}''$ — klasy A', B', C'), a terminom teoretycznym L, M, N — różne (w modelu $\mathcal{M}'$ — klasy L', M', N', w modelu $\mathcal{M}''$ — różne od nich klasy L'', M'', N'').

W jakim stopniu zatem interpretacja terminów elementarnych determinuje interpretację terminów teoretycznych, skoro nie czyni tego w sposób jednoznaczny? Stopień ten zależy oczywiście od stopnia ścisłości związków logicznych, jakie na gruncie danej teorii zachodzą pomiędzy tymi dwoma klasami terminów. Związki te, jak się okazuje,[20] mogą przybierać formy różnorodne. Najczęstszą jednak formą, luźniejszą od definicji zupełnej, jest tzw. definicja cząstkowa. Toteż uwagę naszą poświęcimy obecnie teoriom empirycznym, w których terminy teoretyczne są definiowalne cząstkowo przez terminy elementarne. Taki właśnie charakter ma m. in. fikcyjna teoria Braithwaite'a.[21] Najprostsza postać definicji cząstkowej terminu teoretycznego Q za pomocą terminów elementarnych P_1 i P_2 składa się z następujących wypowiedzi:

(1) $$(x)\,(P_1x \supset Qx),$$

(2) $$(x)\,(P_2x \supset {\sim}Qx).$$

Załóżmy dla uproszczenia, iż powyższe wypowiedzi stanowią jedyne aksjomaty teorii T o terminach specyficznych P_1, P_2, Q i oznaczmy ich koniunkcję przez $\Phi\,(P_1, P_2, Q)$. Interpretacja terminów elementarnych P_1, P_2 nie gwarantuje, jak widzieliśmy, jednoznacznej interpretacji terminu teoretycznego Q. Z łatwością można dobrać, podobnie jak w przypadku teorii Braithwaite'a, takie dwa modele teorii T, w których P_1 i P_2 będą mieć denotacje identyczne, a Q — różne. Nasuwa się jednak pytanie, czy wśród możliwych interpretacji terminów elementarnych P_1, P_2 nie ma takiej określonej interpretacji, która by wyznaczała interpretację terminu teoretycznego Q w sposób jednoznaczny. Zwróćmy uwagę, iż cytowane poprzednio twierdzenie możliwości takiej nie wyłącza. Głosi ono jedynie, iż nie wszystkie możliwe interpretacje terminów P_1, P_2 ten warunek spełniają; iż zawsze znaleźć można taką określoną interpretację tych terminów, która dopuszcza różne interpretacje terminu Q. Nie przeczy ono jednak temu, iż może istnieć taka określona interpretacja terminów P_1, P_2, która dopuszcza tylko jedną interpretację terminu Q. Nie jest więc wykluczone, że rzeczywista interpretacja terminów specyficznych rozważanych teorii empirycznych temu właśnie warunkowi czyni zadość.

Zastanówmy się zatem, jaka by to musiała być interpretacja. Klasy stanowiące denotacje terminów elementarnych P_1, P_2 musiałyby spełniać nie tylko warunki nałożone na nie przez aksjomaty $\Phi\,(P_1, P_2, Q)$, ale i pewne warunki dodatkowe. Warunki te musiałyby gwarantować, iż jedna tylko klasa spełnia aksjomaty, w których terminy P_1, P_2 mają ową interpretację:

---

20) Por. „Pojęcia teoretyczne.." cyt. wyd.

21) Por. „Pojęcia teoretyczne..." cyt. wyd.

$$(X)\,(Y)\,(\Phi\,(P_1, P_2, X) \cdot \Phi\,(P_1, P_2, Y) \supset X = Y)$$

Warunek powyższy równoważny jest logicznie stwierdzeniu:

(3) $$(x)\,(P_1x \vee P_2x).$$

A zatem taka i tylko taka interpretacja terminów elementarnych P_1, P_2, która przyporządkowuje im jako denotacje klasy spełniające warunek (3), wyznacza jednoznacznie klasę stanowiącą denotację terminu teoretycznego Q. Warunek (3) orzeka, iż denotacje terminów P_1 i P_2 wyczerpują uniwersum teorii; każdy przedmiot jest bądź P_1 bądź P_2. Czy warunek taki bywa spełniony w przypadku rozważanych przez nas teorii empirycznych? Na pewno nie. Gdyby było tak, jak on głosi, termin Q byłby faktycznie definiowalny przez terminy elementarne P_1, P_2. Aksjomaty $\Phi\,(P_1, P_2, X)$ stanowią przy założeniu prawdziwości twierdzenia (3) definicję zupełną terminu Q za pomocą terminów P_1, P_2. Implikują one łącznie z twierdzeniem (3) twierdzenie o postaci:

$$(x)\,(Qx \equiv P_1x).$$

Tymczasem w rzeczywistości termin teoretyczny Q nie daje się zdefiniować przez terminy elementarne P_1, P_2. Aksjomaty $\Phi\,(P_1, P_2, X)$ stanowią cząstkową jedynie definicję terminu Q za pomocą terminów P_1, P_2. Nie ma więc waloru żadne twierdzenie równoważne twierdzeniu (3). Denotacje terminów P_1, P_2 nie wyczerpują uniwersum teorii; istnieją przedmioty, które nie są ani P_1, ani P_2. A zatem faktyczna interpretacja terminów elementarnych P_1, P_2 nie wyznacza interpretacji terminu teoretycznego Q w sposób jednoznaczny, lecz dopuszcza możliwości interpretacji różnych. Zakres tych możliwości zostaje w procesie rozwoju teorii zwężany nie przez dołączanie twierdzeń w rodzaju warunku (3), które okazują się z reguły fałszywe, lecz przez dołączanie dodatkowych definicji cząstkowych dla terminu Q odwołujących się do innych terminów elementarnych P_3, P_4:

(1′) $$(x)\,(P_3x \supset Qx),$$

(2′) $$(x)\,(P_4x \supset {\sim}Qx).$$

Mimo iż takie wzbogacenie pierwotnych aksjomatów teorii stanowi dalsze ograniczenie możliwości interpretacyjnych terminu Q, jednakże z reguły nie ogranicza tych możliwości do jedynej. Miałoby to miejsce tylko wtedy, gdyby prawdą okazywało się twierdzenie (3′), analogiczne do (3):

(3′) $$(x)\,(P_1x \vee P_2x \vee P_3x \vee P_4x).$$

Tak jednak na ogół nie jest. W dalszym więc ciągu termin Q dopuszcza różne możliwości interpretacji.

W pierwszej części naszych rozważań doszliśmy do wniosku, iż na drodze interpretacji intrasystemowej nie sposób terminom specyficznym przyporządkować denotacji w sposób jednoznaczny. Co więcej, na drodze takiej niepodobna w ogóle scharakteryzować owych denotacji tak, aby terminom tym nadać sens empiryczny. Obecnie doszliśmy do konkluzji, iż również przy założeniu ekstrasystemowej (bezpośredniej) interpretacji terminów elementarnych nie można na drodze interpretacji intrasystemowej terminom pozostałym, tj. teoretycznym, przyporządkować denotacji w sposób jednoznaczny. Pod pewnym istotnym względem sytuacja jednak przedstawia się tu inaczej. Terminy teoretyczne, mimo iż nie mają interpretacji jednoznacznej, posiadają interpretację taką, która zapewnia im sens empiryczny. Rozpatrzmy to na przykładzie terminu teoretycznego Q zdefiniowanego cząstkowo przez terminy elementarne P_1, P_2 za pomocą wypowiedzi (1) i (2). W pracy „Pojęcia teoretyczne a doświadczenie" starałem się okazać, iż termin taki ma sens empiryczny, rozumiany w pewien przyjęty sposób. Wystarczy w tym celu zdać sobie sprawę z tego, iż terminy elementarne, którym jako denotacje zostały przyporządkowane obserwowalne klasy przedmiotów, już taki sens mają. Denotacja terminu Q nie jest przez nie co prawda wyznaczona jednoznacznie. Może nią być dowolna klasa spełniająca aksjomaty Φ (P_1, P_2, Q). Ale każda taka klasa charakteryzować się musi tym, iż zawiera w sobie klasę P_1, a zawarta jest w klasie $\sim P_2$. A zatem o każdym przedmiocie, który jest P_1, wiemy, iż podpada pod termin Q, o każdym, który jest P_2 wiemy, iż pod termin Q nie podpada. A to, czy jakiś przedmiot jest P_1 (ew. P_2), rozstrzygamy bezpośrednio na podstawie doświadczenia. Istnieją oczywiście przedmioty — mianowicie przedmioty nie będące ani P_1 ani P_2 — w stosunku do których nie rozporządzamy żadnym kryterium stosowalności terminu Q. Fakt ten jednak nie pozbawia terminu Q sensu empirycznego. Świadczy tylko o tym, iż jest to termin nieostry. Ale taki charakter mają wszelkie terminy empiryczne.[22] Przeprowadzona analiza sposobu interpretacji specyficznych terminów teorii empirycznych zdaje w ten sposób sprawę z owej charakterystycznej cechy języka empirycznego.

22) Por. odnośnik 17.

Filozofia Nauki
Rok I, 1993, Nr 2-3

Problem interpretacji języka empirycznego w ujęciu teorio–modelowym

Jednym z podstawowych problemów logicznej metodologii nauk empirycznych jest problem interpretacji języka tych nauk — problem sposobu, w jaki wyrażeniom języka empirycznego nadane zostaje znaczenie i odniesienie przedmiotowe. Zakładamy tu, iż język każdej teorii empirycznej jest językiem sensownym, językiem odnoszącym się do określonej dziedziny rzeczywistości (lub do określonej klasy takich dziedzin), i pytamy, na czym polega procedura, która mu taki charakter zapewnia. Idzie tu — tak jak i w innych zagadnieniach metodologii logicznej — nie o opis sprawozdawczy faktycznie stosowanych zabiegów interpretacyjnych, lecz o ich rekonstrukcję logiczną — o taką charakterystykę procedury interpretacyjnej, która by dostarczała racji dla przypisania danemu językowi interpretacji faktycznie mu przysługującej. Sprawa ta nabiera szczególnego znaczenia wtedy, gdy od nauk empirycznych posługujących się językiem potocznym, lub niewiele od potocznego odbiegającym, przechodzimy do teorii empirycznych, których język zawiera wyrażenia specyficzne i językowi potocznemu obce. Zarówno ich sens, jak i sposób jego określania dalekie są od oczywistości. Terminy tych teorii odnosić się mają z założenia do przedmiotów odległych od przedmiotów naszego codziennego doświadczenia, toteż i sposób przyporządkowania im tych przedmiotów przybierać tu musi charakter szczególny.

Aby zdać z niego sprawę, trzeba w całości, jaką tworzy zinterpretowany język empiryczny, wyróżnić dwa różne składniki: język niezinterpretowany, stanowiący pewien twór czysto formalny, i jego interpretację, pojętą jako pewna dziedzina rzeczywistości, do której się tak rozumiany język odnosi. Oba te pojęcia znajdują swoje ścisłe odpowiedniki w aparacie formalnym współczesnej semantyki logicznej: są to pojęcia języka sformalizowanego i jego modelu. Teoria modeli jest tym działem logiki współczesnej, który bada związki semantyczne między językami sformalizowanymi a ich modelami. Toteż środki logiczne tej właśnie teorii posłużą nam do analizy problemu interpretacji języka empirycznego. Jej głównym celem jest dostarczenie przykładu stosowalności pojęć teorio–modelowych do semantyki języków empirycznych. Stanowi ona tym samym kontynuację podobnych analiz zawartych zarówno w moich pracach wcześniejszych (w szczególności w monografii *The Logic of Empirical Theories*, 1969), jak i w pracach innych autorów (Suszko, Wójcicki, Montague). Aparat formalny, z którego się tu korzysta, nie wykraczający poza elementarne pojęcia teorii modeli, znany jest dziś nie tylko logikom, lecz i metodologom; co więcej, należy do zasobu standardowych narzędzi badawczych metodologii współczesnej. Nie ma zatem potrzeby wyjaśniać go tu raz jeszcze w sposób szczegółowy i systematyczny. Ograniczę się więc jedynie do paru słów przypomnienia[1].

I

1. Problem interpretacji języka empirycznego rozważymy na przykładzie języków możliwie najprostszych. Będą to języki o tzw. standardowej formalizacji zawierające predykaty jako jedyne terminy logiczne. (To upraszczające założenie nie zmniejszy, jak zobaczymy, w sposób niepożądany ogólności naszych rozważań.) Słownik takiego języka L zawiera, prócz zmiennych indywiduowych, spójniki zdaniowe, kwantyfikatory i predykat identyczności — jako stałe logiczne, oraz predykaty $r_1, \ldots, r_n$ — jako stałe pozalogiczne. W znany, czysto syntaktyczny sposób definiujemy zbiór formuł zdaniowych języka L, oraz zbiór zdań (formuł zdaniowych nie zawierających zmiennych wolnych) języka L. W czysto syntaktyczny sposób charakteryzujemy również operację konsekwencji logicznej w języku L, Cn, a stąd i zbiór twierdzeń logicznych języka L jako ogół konsekwencji logicznych zbioru pustego, $Cn(\emptyset)$ (obejmujący twierdzenia węższego rachunku predykatów z identycznością).

Tak scharakteryzowany język L pozostaje oczywiście językiem niezinterpretowanym. Nie mówiąc o żadnej określonej dziedzinie rzeczywistości, nadaje się jednak do mówienia o różnych takich dziedzinach. Każda dziedzina, o której można mówić w języku L, stanowi jeden z modeli tego języka. Modelem języka L jest więc każdy układ

$$\mathfrak{M} = \langle U; R_1, \ldots, R_n \rangle$$

który składa się z niepustego zbioru U, zwanego uniwersum modelu $\mathfrak{M}$, i z relacji $R_1, \ldots, R_n$, zachodzących między jego elementami, z których każda ma tyle członów, ile argumentów ma odpowiadający jej predykat języka L. Każdy model $\mathfrak{M}$ języka L wyznacza jedną z możliwych interpretacji tego języka. Jego zmiennym przyporządkowuje jako zbiór wartości uniwersum U, a predykatom $r_1, \ldots, r_n$ jako ich denotacje relacje $R_1, \ldots, R_n$.

1) Wszystkie pojęcia formalne używane w tej pracy wyjaśnione są w każdym obszerniejszym podręczniku logiki. Ważniejsze z nich wprowadza w sposób przystępny R. Wójcicki w pracy „Analityczność, syntetyczność, empiryczna sensowność zdań", *Studia Filozoficzne* 3 (46), 1966.

Pojęcie modelu $\mathfrak{M}$ języka L pozwala na wprowadzenie, w znany dobrze sposób, podstawowych pojęć semantycznych zrelatywizowanych do modelu $\mathfrak{M}$. Pojęciem wyjściowym jest w tej konstrukcji pojęcie spełniania. Nie sposób przedstawić tu jego, skomplikowanej nieco, definicji, ale i nie ma potrzeby, bo jest to pojęcie znane i intuicyjnie jasne. Poprzestanę zatem na krótkim, nieformalnym wyjaśnieniu. Niech $\alpha(x_1, ..., x_k)$ będzie formułą języka L o zmiennych wolnych $x_1, ..., x_k$, $\mathfrak{M} = \langle U; R_1, ..., R_n\rangle$ — modelem tego języka, a $\langle a_1, ..., a_k\rangle$ — ciągiem elementów uniwersum U.

Formuła $\alpha(x_1, ..., x_k)$ jest spełniona w modelu $\mathfrak{M}$ przez ciąg $\langle a_1, ..., a_k\rangle$ wtedy, gdy jest tak jak głosi formuła $\alpha(x_1, ..., x_k)$ zinterpretowana w sposób następujący: jej zmienne związane przebiegają zbiór U, zmienne wolne $x_1, ..., x_k$ pełnią funkcje nazw elementów $a_1, ..., a_k$, predykaty $r_1, ..., r_n$ denotują relacje $R_1, ..., R_n$, a stałe logiczne przybierają swą zwykłą, klasyczną interpretację.

Pojęcie spełniania pozwala z kolei na zdefiniowanie pojęcia prawdziwości w modelu $\mathfrak{M}$.

Formuła $\alpha(x_1, ..., x_k)$ jest prawdziwa w modelu $\mathfrak{M}$ wtedy, gdy formuła $\alpha(x_1, ..., x_k)$ jest spełniona w modelu $\mathfrak{M}$ przez każdy ciąg $\langle a_1, ..., a_k\rangle$ elementów zbioru U.

Dla formuł nie zawierających zmiennych wolnych otrzymujemy w ten sposób pojęcie zdania prawdziwego w modelu $\mathfrak{M}$. Zbiór zdań języka L prawdziwych w modelu $\mathfrak{M}$ symbolizować będziemy przez *Ver*$(\mathfrak{M})$, a zbiór ich negacji — zdań języka L fałszywych w modelu $\mathfrak{M}$ — przez *Fls*$(\mathfrak{M})$. Gdy wszystkie zdania zbioru X są prawdziwe w $\mathfrak{M}$, mówimy, iż $\mathfrak{M}$ jest modelem zbioru zdań X, w skrócie $\mathfrak{M} \in M(X)$. Przy pomocy tego pojęcia podać możemy semantyczną charakterystykę wprowadzonych wyżej w sposób syntaktyczny pojęć konsekwencji i twierdzenia logicznego:

$\alpha \in Cn(X)$ wtedy, gdy każdy model zbioru zdań X jest modelem zdania α;

$\alpha \in Cn(\emptyset)$ wtedy, gdy każdy model języka L jest modelem zdania α.

Przyjmuje się na ogół, iż język L staje się językiem zinterpretowanym wtedy, gdy spośród wszystkich dziedzin, o których można mówić w języku L, wyróżniona zostaje jedna jako ta, o której język ten faktycznie mowi; innymi słowy, gdy spośród wszystkich modeli języka L wyróżniony zostaje jeden jako tzw. model właściwy lub zamierzony. Język zinterpretowany utożsamia się wobec tego najczęściej z parą $\langle L, \mathfrak{M}^*\rangle$ złożoną z języka sformalizowanego L i jego modelu właściwego $\mathfrak{M}^*$. W zastosowaniu do języka zinterpretowanego wprowadzić możemy pojęcia semantyczne «absolutne» (a nie «relatywne», jak poprzednie), w szczególności pojęcie zdania prawdziwego *tout court*, utożsamiając prawdziwość z prawdziwością w modelu właściwym. Zdanie α jest prawdziwe, gdy jest prawdziwe w $\mathfrak{M}^*$; fałszywe, gdy jest fałszywe w $\mathfrak{M}^*$. Oznaczając zbiór zdań prawdziwych języka L przez *Ver*, a fałszywych przez *Fls*, notujemy, iż:

$$Ver = Ver(\mathfrak{M}^*),\ Fls = Fls(\mathfrak{M}^*).$$

Taka koncepcja języka zinterpretowanego wydaje się jednak zbyt rygorystyczna, zwłaszcza w zastosowaniu do języków empirycznych. To, o czym mówi język empiryczny, nie jest prawie nigdy zdeterminowane w sposób jednoznaczny. Czynniki pragmatyczne decydujące o tym, do czego się odnoszą wyrażenia takiego języka, określają jego model właściwy w sposób wieloznaczny. Nieuchronna nieostrość wszelkich terminów empirycznych jest tego wyraźnym dowodem. W istocie więc czynniki te wyznaczają nie jeden model właściwy $\mathfrak{M}^*$, lecz klasę takich modeli M^*, zawierającą z reguły więcej niż jeden element. Dalsza analiza interpretacji języka empirycznego przynosi wyraźne potwierdzenie tego założenia. Przyjmiemy tu zatem bardziej liberalną koncepcję języka zinterpretowanego, zgodnie z którą o języku takim możemy mówić już wtedy, gdy mamy wyróżnioną pewną klasę jego modeli właściwych M^* — niepustą i nie pokrywającą się z ogółem jego modeli. Utożsamiamy więc każdy język zinterpretowany z parą $\langle L, M^* \rangle$, gdzie L — to język sformalizowany, a M^* — to klasa jego modeli spełniająca wymienione warunki. Pojęcie to obejmuje pojęcie poprzednie jako warunek graniczny. Gdy M^* jest klasą jednostkową, otrzymujemy język zinterpretowany w sposób jednoznaczny. Problem zdefiniowania dla języków o interpretacji wieloznacznej «absolutnych» pojęć semantycznych, w szczególności «absolutnego» pojęcia prawdy, natrafia na pewne trudności i do dziś pozostaje problemem otwartym. Wysuwano i dyskutowano różne propozycje jego rozwiązania[2]. Jakkolwiek różnią się one między sobą poza tym, wspólne jest im wszystkim następujące założenie:

Jeśli zdanie α jest prawdziwe w każdym modelu klasy M^*, zdanie to jest prawdziwe; jeśli zdanie α jest fałszywe w każdym modelu klasy M^*, zdanie to jest fałszywe.

Różnice pomiędzy poszczególnymi rozwiązaniami dotyczą wyłącznie kwalifikacji tzw. zdań niezdeterminowanych, tj. zdań, które w pewnych modelach klasy M^* są prawdziwe, a w innych — fałszywe. W rozważaniach naszych nie ma potrzeby przesądzania tej spornej kwestii.

2. Spróbujmy obecnie scharakteryzować w sposób najogólniejszy interpretację dowolnego języka empirycznego. Co możemy założyć o własnościach klasy M^* i o sposobach jej wyznaczania? Wyróżnimy na wstępie dwa główne sposoby interpretowania języka sformalizowanego: werbalny i niewerbalny. Werbalna interpretacja języka L polega na scharakteryzowaniu klasy jego modeli właściwych M^* jako klasy modeli określonego zbioru zdań języka L, tj. jako klasy modeli, w których wszystkie zdania tego zbioru są prawdziwe. Ów zbiór zdań, oznaczany odtąd przez P, nazywać będziemy, zgodnie z przyjętą terminologią, zbiorem postulatów dla języka L (lub dla jego terminów). Werbalna interpretacja języka L sprowadza się do definicji:

$$M^* = M(P)$$

utożsamiającej klasę M^* z pewną klasą definiowalną w języku L. Wszelki inny sposób wyznaczania klasy M^* zaliczymy do sposobów niewerbalnych. Otóż stwierdźmy przede wszystkim, co następuje: jeśli język L ma być językiem empirycznym, jego interpretacja nie może być interpretacją czysto werbalną. Jest to fakt dość oczywisty i niejednokrotnie stwierdzany. Toteż nie uzasadniając go już tutaj szerzej (uczyniłem to m.in. w cytowanej wyżej monografii), ograniczę się do paru luźnych uwag. Jakikolwiek niesprzeczny zbiór języka L obierzemy jako zbiór postulatów P, klasa jego modeli

---

2) Omawiam je m.in. w pracy „Z semantyki pojęć otwartych", *Studia Logica* 15, 1964.

$M(P)$ będzie klasą tak obszerną i zawierać będzie modele tak różnorodne, iż nie może być uznana za interpretację języka empirycznego. Wśród jej modeli znajdą się zawsze modele takie, których uniwersum składa się z liczb naturalnych, i takie, których uniwersum składa się z wyrażeń językowych! Toteż język, którego klasa modeli właściwych obejmuje wszystkie modele klasy $M(P)$, nie może mieć charakteru empirycznego, jakkolwiek szeroko pojmowalibyśmy to określenie. Świadczy o tym choćby to, że w języku tak zinterpretowanym wszelki termin pozalogiczny pozostaje terminem całkowicie nieostrym.

Przyjmujemy zatem, iż interpretacja naszego języka L nie jest interpretacją czysto werbalną. Klasa jego modeli właściwych M^* nie jest tożsama z żadną klasą definiowalną w języku L. W jaki więc sposób zostaje wyznaczona? Jakie niewerbalne metody wchodzą tu w grę? Założenie, jakie w tej sprawie przyjmiemy, uważane być może za wyraz semantycznego empiryzmu. Zgodnie z nim, dwa istnieją tylko sposoby przyporządkowania terminom empirycznym ich denotacji: jeden — bezpośredni, sprowadzający się w rezultacie do wskazania, w ten czy inny sposób, przedmiotów, które do denotacji danego terminu mają należeć (zabieg taki nazywany bywa definicją ostensywną terminu interpretowanego); drugi — pośredni, polegający na scharakteryzowaniu denotacji danego terminu za pomocą postulatów językowych odwołujących się do terminów empirycznych już zinterpretowanych (w pewnych przypadkach postulaty te przybierają postać definicji równoważnościowej). Bezpośredni sposób interpretacji ma zastosowanie tylko w stosunku do tzw. terminów obserwacyjnych, czyli, mówiąc swobodnie, terminów denotujących obserwowalne własności przedmiotów (lub obserwowalne stosunki między takimi przedmiotami). Wszelkie inne terminy empiryczne, zaliczane do tzw. terminów teoretycznych, interpretowane być mogą jedynie w sposób pośredni. Tak więc, zgodnie z powyższymi założeniami, klasę modeli właściwych M^* języka empirycznego L pojmować będziemy zawsze jako pewną niepustą podklasę właściwą klasy wszystkich modeli zbioru postulatów P, $M(P)$:

$$\emptyset \neq M^* \subset M(P)$$

wyodrębnioną z tej ostatniej za pomocą bezpośrednich procedur interpretacyjnych o charakterze definicji ostensywnych pewnych terminów języka L. Zbiór P reprezentuje tu zbiór wszystkich postulatów dla terminów języka L. Prócz nich — i ich logicznych konsekwencji $Cn(P)$ — nie ma innych zdań języka L, których prawdziwość w modelach klasy M^* byłaby zagwarantowana przez sam sposób wyznaczenia tej klasy.

Bezpośredni sposób interpretacji terminów empirycznych jest procedurą, której analiza wymyka się metodom formalno–logicznym[3]. Dalsze rozważania poświęcimy analizie interpretacji pośredniej. Przedmiotem tych rozważań będzie procedura wzbogacania języka empirycznego o nowe terminy teoretyczne, a więc terminy dopuszczające wyłącznie interpretację pośrednią. A idzie nam o to, jak zdać sprawę — przy

3) Problem ten rozważałem m.in. w pracy „O definiowaniu terminów spostrzeżeniowych", *Rozprawy Logiczne*, 1964.

założeniu określonej interpretacji języka wyjściowego — z interpretacji języka wzbogaconego. Na to pytanie, jak zobaczymy, nasuwają się odpowiedzi różne — zależne od typu sytuacji, w której owo wzbogacenie ma miejsce. Aby odpowiedzi te sformułować w sposób precyzyjny, musimy wprowadzić pewne formalne pojęcia i symboliczne oznaczenia.

3. Niech naszym wyjściowym językiem empirycznym będzie język L_1 zawierający jako jedyne terminy pozalogiczne predykaty $r_1, ..., r_n$. Język ten wzbogacamy o nowe predykaty $q_1, ..., q_m$ przechodząc tym sposobem do języka L_2, zawierającego L_1 jako swoją część. Przyjmiemy tu konwencję, zgodnie z którą wszelkie synktatyczne i semantyczne pojęcia dotyczące języka L_i (i=1, 2) zaopatrzone będą we wskaźnik i. W szczególności, modele języka L_1, czyli układy $\langle U; R_1, ..., R_n\rangle$, symbolizować będziemy przez $\mathfrak{M}_1$, a modele języka L_2, czyli układy $\langle U; R_1, ..., R_n, Q_1, ..., Q_m\rangle$ — przez $\mathfrak{M}_2$. Symbol $\mathfrak{M}_2|_1$ oznaczać będzie fragment modelu $\mathfrak{M}_2$ odpowiadający językowi L_1, tj. model języka L_1 otrzymany z $\mathfrak{M}_2$ przez eliminację denotacji predykatów $q_1, ..., q_m$. Zakładamy, iż P_1 jest zbiorem postulatów dla predykatów $r_1, ..., r_n$ języka L_1, a P_2 — zbiorem postulatów dla nowo wprowadzonych predykatów $q_1, ..., q_m$ języka L_2. Zgodnie z tym, co ustaliliśmy wyżej, przyjmujemy, że język L_1 jest językiem zinterpretowanym, a klasa modeli M_1^* wyznaczająca tę interpretację, stanowi niepustą podklasę właściwą klasy modeli $M(P_1)$, wyodrębnioną z niej przez bezpośrednie zabiegi interpretacyjne:

$$\emptyset \neq M_1^* \subset M(P_1).$$

A jak zdać sprawę z interpretacji języka L_2? Jak scharakteryzować klasę M_2^* jego modeli właściwych? Odpowiedź na to pytanie zależy od pewnych czynników pragmatycznych, w szczególności od intencji użytkowników języka L_2 decydujących o charakterze jego interpretacji. W rozważaniach obecnych weźmiemy pod uwagę takie tylko — niewątpliwie najczęstsze — sytuacje, w których owo wzbogacenie języka L_1 ma mieć, zgodnie z intencjami jego użytkowników, charakter «konserwatywny», zachowujący jego interpretację dotychczasową. W sytuacjach tych język L_2 ma być, z założenia, interpretowany tak, aby, po pierwsze, wszystkie terminy należące do języka L_1, a więc predykaty $r_1, ..., r_n$, zachowały swą interpretację dotychczasową, wyznaczoną przez klasę M_1^*; po drugie, wszystkie terminy wprowadzane, a więc predykaty q_1, ..., q_m, były interpretowane zgodnie z postulatami P_2, tj. w taki sposób, aby postulaty te były prawdziwe. Wszelka definicja klasy M_2^*, wyznaczającej interpretację języka L_2, musi respektować te żądania. Warunek drugi dopuszcza, jak się wydaje, jeden tylko sposób precyzacji. Żąda po prostu tego, aby każdy model $\mathfrak{M}_2$, należący do klasy M_2^* był modelem zbioru postulatów P_2 : $\mathfrak{M}_2 \in M(P_2)$. Warunek pierwszy natomiast dopuszcza eksplikacje różne. Owo zachowanie przez terminy języka L_1 ich interpretacji dotychczasowej może być rozumiane w sposób bardziej lub mniej rygorystyczny, i wydaje się, że w różnych sytuacjach faktycznych różnie też bywa rozumiane. Aby uchwycić te różnice, musimy odwołać się do pewnych pojęć teorii modeli — pojęć, które w zastosowaniach tej teorii do semantyki języków empirycznych zdają się odgrywać szczególnie ważną rolę. Mam na myśli takie pojęcia teorio–modelowe, jak pojęcie rozszerzenia i elementarnego rozszerzenia modeli. Wyjaśnimy ich sens w zastosowaniu do omawianego przez nas języka L.

Niech $\mathfrak{M} = \langle U; R_1, ..., R_n \rangle$ i $\mathfrak{M}' = \langle U'; R_1', ..., R_n' \rangle$ będą dwoma modelami języka L.

$\mathfrak{M}'$ jest rozszerzeniem $\mathfrak{M}$, w skrócie $\mathfrak{M} \subset \mathfrak{M}'$, wtedy gdy (i) $U \subset U'$, oraz (ii) $R_i' |_U = R_i$ $(i = 1, ..., n)$; a więc gdy uniwersum modelu $\mathfrak{M}'$ zawiera w sobie uniwersum modelu $\mathfrak{M}$, a relacje modelu $\mathfrak{M}'$ ograniczone do uniwersum modelu $\mathfrak{M}$ pokrywają się z odpowiadającymi im relacjami modelu $\mathfrak{M}$.

$\mathfrak{M}'$ jest elementarnym rozszerzeniem $\mathfrak{M}$, w skrócie $\mathfrak{M} < \mathfrak{M}'$, wtedy, gdy (i) $\mathfrak{M}'$ jest rozszerzeniem $\mathfrak{M}$ oraz (ii) dla dowolnej formuły $\alpha(x_1, ..., x_k)$ języka L i dla dowolnego ciągu $\langle a_1, ..., a_k \rangle$ elementów uniwersum modelu $\mathfrak{M}$: formuła $\alpha(x_1, ..., x_k)$ jest spełniona przez ciąg $\langle a_1, ..., a_k \rangle$ w modelu $\mathfrak{M}'$ wtedy i tylko wtedy, gdy formuła $\alpha(x_1, ..., x_k)$ jest spełniona przez ciąg $\langle a_1, ..., a_k \rangle$ w modelu $\mathfrak{M}$.

Zwróćmy uwagę na pewne konsekwencje tych definicji. Jeśli $\mathfrak{M}'$ jest zwykłym lub elementarnym rozszerzeniem $\mathfrak{M}$, uniwersum modelu $\mathfrak{M}'$ może zawierać elementy nowe, nie należące do uniwersum modelu $\mathfrak{M}$. W obu przypadkach interpretacja predykatów $r_1, ..., r_n$ w dawnej części uniwersum modelu $\mathfrak{M}'$, tj. w zbiorze U, musi być identyczna z ich interpretacją w modelu $\mathfrak{M}$. A jak się przedstawia ich interpretacja w nowej części uniwersum modelu $\mathfrak{M}'$, tj. w zbiorze U'–U? Gdy $\mathfrak{M}'$ jest zwykłym rozszerzeniem $\mathfrak{M}$, interpretacja ta może być najzupełniej dowolna. Gdy natomiast $\mathfrak{M}'$ jest elementarnym rozszerzeniem $\mathfrak{M}$, interpretacja ta spełniać musi warunek (ii). Jaki jest jego sens intuicyjny? Zauważmy, iż pociąga in warunek tzw. elementarnej równoważności modeli $\mathfrak{M}'$ i $\mathfrak{M}$:

$$Ver(\mathfrak{M}') = Ver(\mathfrak{M}).$$

Każde zdanie języka L prawdziwe w modelu $\mathfrak{M}'$ musi być prawdziwe w modelu $\mathfrak{M}$, i na odwrót. A zatem predykaty $r_1, ..., r_n$ muszą być w modelu $\mathfrak{M}'$ zinterpretowane w taki sposób, który by nie zmieniał wartości logicznej żadnego zdania języka L zawierającego owe predykaty. Cokolwiek (wyrażalnego w języku L) było o nich prawdą w modelu $\mathfrak{M}$, musi pozostać prawdę w modelu $\mathfrak{M}'$. W rzeczywistości ów związek między interpretacją predykatów $r_1, ..., r_n$ w modelu $\mathfrak{M}$ i $\mathfrak{M}'$ jest jeszcze ściślejszy. Warto może zwrócić uwagę na fakt, iż w przypadku modelu $\mathfrak{M}$ o uniwersum skończonym, każde elementarne rozszerzenie modelu $\mathfrak{M}$ musi być z nim identyczne.

Wprowadzone obecnie pojęcia pozwalają na eksplikację warunku postulującego zachowanie dotychczasowej interpretacji języka L_1 przy przejściu do języka L_2. Przy ich pomocy wyróżnić możemy co najmniej trzy wersje tego warunku, formułujące żądania coraz to mniej rygorystyczne. Interpretacja języka L_1 wyznaczona jest, jak pamiętamy, przez klasę modeli M_1^*, interpretacja języka L_2 — przez klasę modeli M_2^*. Otóż żądać możemy, aby fragment dowolnego modelu $\mathfrak{M}_2 \in M_2^*$ odpowiadający językowi L_1 był identyczny z pewnym modelem $\mathfrak{M}_1 \in M_1^*$: $\mathfrak{M}_1 = \mathfrak{M}_2 |_1$; był jego elementarnym rozszerzeniem: $\mathfrak{M}_1 < \mathfrak{M}_2|_1$; a wreszcie, był jego rozszerzeniem zwykłym: $\mathfrak{M}_1 \subset \mathfrak{M}_2|_1$. Każdej z tych możliwości odpowiada inna definicja klasy M_2^*. Podamy naprzód ich zapis formalny[4], a następnie rozważymy sens intuicyjny.

D1. $M_2^* = \{\mathfrak{M}_2 : \mathfrak{M}_2 \in M(P_2)$ i $\mathfrak{M}_1 = \mathfrak{M}_2|_1$, dla pewnego $\mathfrak{M}_1 \in M_1^*\}$.

D2. $M_2^* = \{\mathfrak{M}_2 : \mathfrak{M}_2 \in M(P_2)$ i $\mathfrak{M}_1 < \mathfrak{M}_2|_1$, dla pewnego $\mathfrak{M}_1 \in M_1^*\}$.

D3. $M_2^* = \{\mathfrak{M}_2 : \mathfrak{M}_2 \in M(P_2)$ i $\mathfrak{M}_1 \subset \mathfrak{M}_2|_1$, dla pewnego $\mathfrak{M}_1 \in M_1^*\}$.

Każda z tych definicji zalicza, jak widać, do klasy M_2^*, determinującej interpretację języka L_2, tylko takie modele języka L_2, które są modelami zbioru postulatów P_2. Tym samym każda z tych definicji realizuje żądanie interpretowania nowo wprowadzonych predykatów $q_1, ..., q_m$ w sposób zgodny z postulatami P_2. Czy można również przyjąć, że każda z podanych definicji realizuje żądanie interpretowania dawnych predykatów $r_1, ..., r_n$ w sposób zgodny z ich interpretacją dotychczasową, daną przez klasę M_1^*? Niewątpliwie, każda z nich w jakimś stopniu czyni temu żądaniu zadość. Zgodnie z każdą z tych definicji, uniwersum dowolnego modelu $\mathfrak{M}_2$ klasy M_2^* obejmuje uniwersum pewnego modelu $\mathfrak{M}_1$ klasy M_1^*. Zgodnie z każdą z nich, dawne predykaty $r_1, ..., r_n$ interpretowane są w modelu $\mathfrak{M}_2$ w obrębie dawnej części jego uniwersum (tj. w uniwersum modelu $\mathfrak{M}_1$) dokładnie tak samo, jak w modelu $\mathfrak{M}_1$, W tym też sensie każda z tych definicji zachowuje dotychczasową interpretację terminów języka L_1. A jakie zachodzą między nimi różnice? Definicja D1 zachowuje tę interpretację w sensie możliwie najściślejszym. Ponieważ wedle niej uniwersum każdego modelu $\mathfrak{M}_2$ klasy M_2^* pokrywa się dokładnie z uniwersum pewnego modelu $\mathfrak{M}_1$ klasy M_1^*, interpretacja predykatów $r_1, ..., r_n$ w modelu $\mathfrak{M}_2$ jest po prostu identyczna z ich interpretacją w modelu $\mathfrak{M}_1$; predykaty te w obu modelach denotują te same relacje. Wedle definicji pozostałych, uniwersa modeli klasy M_2^* mogą zawierać przedmioty nie wchodzące w skład żadnego z uniwersów modeli klasy M_1^*. Interpretacje predykatów języka L_1 w modelach klasy M_2^* mogą więc różnić się od interpretacji w modelach klasy M_1^*; do denotacji tych predykatów w jakimś modelu klasy M_2^* należeć mogą przedmioty z nowej części jego uniwersum, a więc przedmioty, które nie należały do ich denotacji w żadnym modelu klasy M_1^*. Definicja D2 nakłada jednak na takie zmiany interpretacji daleko idące ograniczenia. Żądając, aby fragment każdego modelu $\mathfrak{M}_2$ klasy M_2^* odpowiadający językowi L_1 był elementarnym rozszerzeniem pewnego modelu $\mathfrak{M}_1$ klasy M_1^*, żąda tym samym, jak widzieliśmy, aby były to zmiany niewyrażalne w języku L_1 (a nawet w pewnych wzbogaceniach tego języka), aby więc owo rozszerzenie interpretacji języka L_1 było dla użytkowników tego języka w pewnym sensie niezauważalne. Definicja D3 tego rodzaju ograniczeń nie nakłada. Interpretacja języka L_1 może tu ulegać dowolnym rozszerzeniom (byle by zgodnym z postulatami P_2).

4) Skrót $\{x : W(x)\}$ symbolizuje zbiór x-ów spełniających warunek W.

Pierwotna interpretacja języka L_1 jest jednak charakteryzowana i przez to, że jest to interpretacja zgodna z postulatami P_1: każdy model klasy M_1^* — to model zbioru postulatów P_1. Czy ta jej własność zostaje zachowana przy przejściu do języka L_2? Zachodzi pod tym względem wyraźna różnica między definicją D3 a definicjami pozostałymi. Jeśli klasa M_2^* określona jest zgodnie z definicją D1 lub D2, każdy należący do niej model musi być modelem zbioru postulatów P_1, skoro jego fragment odpowiadający językowi L_1 jest identyczny z jakimś modelem klasy M_1^*, lub jest jego elementarnym rozszerzeniem. Inaczej jest jednak w przypadku definicji D3. Fakt, iż fragment modelu $\mathfrak{M}_2$ klasy M_2^* odpowiadający językowi L_1 jest rozszerzeniem jakiegoś modelu $\mathfrak{M}_1$ klasy M_1^*, gwarantuje prawdziwość w modelu $\mathfrak{M}_2$ tylko tym zdaniom prawdziwym w modelu $\mathfrak{M}_1$, które są zdaniami czysto egzystencjalnymi (tj. zdaniami równoważnymi logicznie takim zdaniom o postaci normalnej, które nie zawierają kwantyfikatorów ogólnych). Mogą więc istnieć wśród postulatów P_1 zdania takie, które okażą się zdaniami fałszywymi w modelach klasy M_2^*. I tutaj jednak zachodzi fakt zachowania postulatów P_1 w pewnej ograniczonej postaci. Załóżmy, iż w języku L_2 istnieje jednoargumentowy predyktat q_1, który w każdym modelu $\mathfrak{M}_2$ klasy M_2^* denotuje dawne uniwersum, tj. uniwersum tego modelu $\mathfrak{M}_1$ klasy M_1^*, którego rozszerzeniem jest model $\mathfrak{M}_2$. Dokonajmy w każdym zdaniu zbioru P_1 relatywizacji wszystkich zmiennych związanych do predykatu q_1. (Operacja ta polega na zastąpieniu kwantyfikatorów zwykłych przez odpowiednie kwantyfikatory ograniczone do predykatu q_1). Zbiór tak przekształconych postulatów oznaczmy przez P_1 (q_1). Otóż definicja D3 klasy M_2^* gwarantuje nam, że każdy model tej klasy jest modelem zbioru zdań P_1 (q_1). W każdym z nich zachowują więc prawdziwość postulaty P_1 ograniczone do dawnego uniwersum.

4. Rozważyliśmy trzy możliwe typy interpretacji języka L_2. Powstaje oczywiście pytanie, czy możliwości te odpowiadają pewnym sytuacjom faktycznym, w szczególności sytuacjom typowym dla rozwoju nauk empirycznych. Jest rzeczą jasną, że zarysowane tu schematy stanowić mogą co najwyżej daleko idące uproszczenie i idealizację jakiegokolwiek faktycznego stanu rzeczy. Z tym jednakże zastrzeżeniem można, jak sądzę, na pytanie to odpowiedzieć twierdząco. Chciałbym wysunąć tu przypuszczenie, iż każdy z opisanych wyżej rodzajów interpretacji odpowiada pewnym charakterystycznym dla nauk empirycznych procedurom wzbogacania języka empirycznego o nowe terminy teoretyczne. Uzasadnienie tego przypuszczenia jest zadaniem odrębnym. Tutaj chcę w paru tylko słowach zaznaczyć, o jakie to sytuacje może chodzić.

Sytuacja odpowiadająca definicji D1 — sytuacja, w której wzbogacając dany język empiryczny zachowujemy jego interpretację, a w szczególności jego uniwersum, niezmienione — to sytuacja bodaj najczęstsza. Taki w szczególności charakter wydają się mieć wszelkie rozszerzenia definicyjne. Wprowadzając do języka L_1 nowy termin za pomocą definicji równoważnościowej, nie mamy potrzeby zmieniania jego dotychczasowej interpretacji, bo mamy w tym przypadku gwarancję, że przy interpretacji istniejącej znajdziemy zawsze żądaną (tj. zgodną z ową definicją) interpretację dla terminu wprowadzanego.

Gwarancji takiej możemy nie mieć jednak wtedy, gdy postulaty dla terminu wprowadzanego przybierają postać różną od definicji równoważnościowej. Wówczas okazać się może, że do tego, aby znaleźć zgodną z tymi postulatami interpretację dla terminu wprowadzanego, musimy rozszerzyć nasze dotychczasowe uniwersum. W pewnych przypadkach — odpowiadających definicji D2 — może to być rozszerzenie elementarne. Jest to możliwe wtedy, gdy uniwersum nasze rozszerzamy o przedmioty tego samego typu, co przedmioty już do niego należące, tj. o przedmioty, które mają te same własności (wyrażalne w dawnym języku L_1), co przedmioty dotychczasowe. Takie rozszerzenie nie zmienia, jak widzieliśmy, istniejącej interpretacji języka L_1 w sposób widoczny. Dla kogoś, kto nie wykracza poza ów język, zmiana taka pozostaje niedostrzegalna.

Istnieją jednak i sytuacje takie, kiedy wzbogacając nasz język świadomie zmieniamy jego dotychczasową interpretację, a zwłaszcza jego dotychczasowe uniwersum, w sposób widoczny. Ma to miejsce wtedy, gdy wprowadzamy terminy odnoszące się do przedmiotów zasadniczo różnych od tych, które składały się na uniwersum dawne. Tak jest na przykład wtedy, gdy do języka obserwacyjnego, którego uniwersum składa się wyłącznie z przedmiotów obserwowalnych, wprowadzamy terminy teoretyczne odnoszące się do przedmiotów zasadniczo nieobserwowalnych. Jest rzeczą jasną, że takie rozszerzenie dotychczasowego uniwersum nie może zostać niezauważone. Nowe przedmioty mają z reguły inne własności, niż przedmioty dawne — i to własności wyrażalne w dawnym języku L_1. Nie wszystko więc, co było prawdą przy poprzedniej interpretacji języka L_1, pozostanie prawdą przy przy obecnej. Dotyczy to również postulatów P_1; pozostaną one prawdziwe tylko wtedy, gdy ograniczymy je do dawnego uniwersum. Zilustrujmy to na najprostszym przykładzie. Przypuśćmy, iż jeden z postulatów języka obserwacyjnego głosi, że „każdy przedmiot jest barwny". Postulat ten, prawdziwy w modelach o uniwersum złożonym z przedmiotów spostrzegalnych, przestaje być prawdziwy, gdy uniwersum to rozszerzymy o przedmioty zasadniczo niespostrzegalne (takie, jak atomy czy elektrony). Prawdą w takich modelach pozostaje natomiast nadal ów postulat ograniczony do uniwersum dawnego, a więc twierdzenie głoszące, iż „każdy przedmiot spostrzegalny jest barwny". Rozszerzenie, które w takich przypadkach wchodzi w grę, nie może być więc rozszerzeniem elementarnym. Są to sytuacje, które zdają się podpadać pod schemat odpowiadający definicji D3[5].

5) Ten rodzaj rozszerzenia języka i jego interpretacji wyróżniony został przez R. Suszkę w pracy „Logika formalna a rozwój poznania", *Studia Filozoficzne* 1 (44), 1966. W pracy tej sformułowana też została zasada zachowania postulatów w postaci ograniczonej.

II

1. To, czy podane przez nas definicje klasy modeli właściwych języka L_2 — D1, D2, D3 — charakteryzują interpretację języka empirycznego w sposób trafny, zgodny z rodzajami interpretacji, które faktycznie takiemu językowi przysługują — zależy m.in. od tego, jakim zbiorem jest zbiór P_2, reprezentujący ogół postulatów dla terminów nowo wprowadzonych. Jeśli definicje te mają charakteryzować interpretację języka L_2 w sposób właściwy, nie może to być zbiór dowolny. Spróbujmy sformułować warunki, jakie każda z tych definicji nakłada na zbiór P_2. Problem ten — w zastosowaniu do pewnych spośród owych definicji — rozważany był już niejednokrotnie. Ograniczę się więc tym razem do uwag ogólnych i skrótowych, prezentujących niektóre uzyskane uprzednio rezultaty.

Taka, czy inna odpowiedź na powyższe pytanie zależy od pewnych ogólnych założeń dotyczących semantycznych własności języka empirycznego. I my przyjmiemy tu takie założenia. Zakładamy mianowicie, że język empiryczny — język pewnej nauki czy teorii empirycznej — jest zawsze językiem zinterpretowanym, i że fakt ten jest niezależny od doświadczenia. Odpowiada to tendencji — dominującej w semantyce współczesnej — do uniezależniania sprawy sensowności wyrażeń językowych od rezultatów doświadczenia. Doświadczenie decydować ma o prawdziwości, czy fałszywości twierdzenia empirycznego, a nie o jego sensowności; ta ma być zagwarantowana z góry, przez samą konstrukcję danego języka empirycznego.

Język zinterpretowany utożsamiliśmy uprzednio z parą $\langle L, M^* \rangle$, w której L — to język sformalizowany, a M^* — to niepusta podklasa właściwa klasy jego modeli. Dany język uznać więc możemy za zinterpretowany tylko pod tym warunkiem, iż klasa jego modeli właściwych jest niepusta. Co więcej, zgodnie z przyjętym przed chwilą założeniem, jej niepustość ma być faktem zagwarantowanym z góry. Tak też potraktować musimy analizowany przez nas język L_2. Stwierdzenie niepustości klasy jego modeli właściwych:

$$M_2^* \neq \emptyset$$

ma być prawdą, i to prawdą niezależną od doświadczenia. Kiedy warunek taki uważać możemy za spełniony? Rozważmy go w zastosowaniu do którejś z podanych przez nas definicji klasy M_2^*, np. definicji D1. Twierdzenie, iż klasa M_2^* jest niepusta, równoważne jest na gruncie tej definicji twierdzeniu głoszącemu, co następuje:

(*) Dla pewnego modelu $\mathfrak{M}_1 \in M_1^*$ istnieje model $\mathfrak{M}_2 \in M(P_2)$ taki, iż $\mathfrak{M}_1 = \mathfrak{M}_2 |_1$.

Kiedy twierdzeniu temu przypisać możemy prawdziwość niezależną od doświadczenia? Sądzę, że wtedy tylko, gdy prawdą jest następujące twierdzenie ogólne:

C1. Dla każdego modelu $\mathfrak{M}_1 \in M(P_1)$ istnieje model $\mathfrak{M}_2 \in M(P_2)$ taki, iż $\mathfrak{M}_1 = \mathfrak{M}_2 |_1$.

Zważmy, iż o klasie modeli właściwych języka L_1 założyliśmy, co następuje:

$$\emptyset \neq M_1^* \subset M(P_1).$$

W świetle tego założenia prawdziwość twierdzenia C1 gwarantuje prawdziwość twierdzenia (*), sama zaś jest najwyraźniej od doświadczenia niezależna. Twierdzenie C1 formułuje pewną zależność teorio–modelową i — dla danych zbiorów P_1 i P_2 — na gruncie samej teorii modeli może być rozstrzygnięte. Z drugiej strony, tylko wtedy, gdy prawdą jest twierdzenie C1, prawdziwość twierdzenia (*) może być zagwarantowana z góry. Klasa M_1^* została, jak wiemy, wyodrębniona z klasy $M(P_1)$ za pomocą bezpośrednich procedur interpretacyjnych typu definicji ostensywnej. Jedyną, z góry założoną własnością jej modeli jest to, iż są to modele postulatów P1. Jeżeli więc chcemy z góry mieć pewność, że wśród modeli klasy M_1^* istnieje model taki, który jest fragmentem pewnego modelu postulatów P_2, musimy z góry wiedzieć, że każdy model klasy $M(P_1)$ spełnia taki warunek. A to właśnie głosi twierdzenie C1[6].

Twierdzenie to traktować możemy jako pewien warunek nałożony na zbiór postulatów P_2. Taki tylko zbiór P_2 może być uznany za zbiór postulatów dla predyktatów $q_1, \ldots, q_m$ języka L_2, który spełnia — dla danego P_1 — warunek sformułowany w tym twierdzeniu. W analogiczny sposób wyrazić możemy warunki nakładane na zbiór postulatów P_2 w przypadku pozostałych definicji klasy M_2^*, tj. definicji D2 i D3. Podobnie jak poprzednio, ich realizacja gwarantuje niepustość klasy M_2^*. A oto kolejne warunki odpowiadające definicjom D1, D2, D3:

C1. Dla każdego modelu $\mathfrak{M}_1 \in M(P_1)$ istnieje model $\mathfrak{M}_2 \in M(P_2)$ taki, że $\mathfrak{M}_1 = \mathfrak{M}_2|_1$.

C2. Dla każdego modelu $\mathfrak{M}_1 \in M(P_1)$ istnieje model $\mathfrak{M}_2 \in M(P_2)$ taki, że $\mathfrak{M}_1 < \mathfrak{M}_2|_1$.

C3. Dla każdego modelu $\mathfrak{M}_1 \in M(P_1)$ istnieje model $\mathfrak{M}_2 \in M(P_2)$ taki, że $\mathfrak{M}_1 \subset \mathfrak{M}_2|_1$.

Sformułowane jak wyżej, warunki te mają najwyraźniej charakter semantyczny. Powstaje zatem pytanie, czy — podobnie jak w przypadku wielu innych pojęć semantycznych — można znaleźć ich odpowiedniki syntaktyczne. Okazuje się, że zachodzi pod tym względem istotna różnica między warunkiem C1 a warunkami pozostałymi. W przeciwieństwie do warunku C1, warunki C2 i C3 dają się sformułować w sposób czysto syntaktyczny. I tak, warunek C2 jest równoważny warunkowi:

C2′. $Cn(P_2) \cap Z_1 \subset Cn(P_1)$,

gdzie Z_1 symbolizuje zbiór zdań języka L_1. Warunek C3 natomiast jest równoważny warunkowi:

C3′. $Cn(P_2) \cap A_1 \subset Cn(P_1)$,

gdzie A_1 symbolizuje zbiór czysto uniwersalnych zdań języka L_1 (tj. zdań języka L_1 równoważnych logicznie takim zdaniom o postaci normalnej, które nie zawierają kwantyfikatorów szczegółowych)[7]. Warunki C2′ i C3′ mają, jak widać, charakter czysto

6) Wywód ten ma charakter szkicowy i intuicyjny. Można by go przedstawić w sposób pełny i precyzyjny za cenę pewnej rozbudowy naszych rozważań. Trzeba by było mianowicie w tym celu scharakteryzować dokładnie metajęzyk opisywanego przez nas języka L_2, co pozwoliłoby na eksplikację owego zagadkowego nieco pojęcia „prawdziwości niezależnej od doświadczenia". Twierdzenie prawdziwe niezależnie od doświadczenia — to, z grubsza mówiąc, takie, które daje się udowodnić na gruncie metajęzyka. Bez trudu można by wówczas okazać, iż twierdzenie (*) wtedy tylko daje się udowodnić na gruncie metajęzyka, gdy daje się udowodnić twierdzenie C1.

7) Równoważność warunków C2 i C2′ stwierdzona jest m.in. w książce J. Shoenfielda *Mathematical Logic*, 1967; jej dowód nie nastręcza trudności. Równoważność warunków C3 i C3′ okazana została po raz pierwszy w pracy J. Łosia „On the extending of models I", *Fundamenta Mathematicae* 42, 1955.

syntaktyczny. Pierwszy z nich żąda, aby wszystkie konsekwencje logiczne postulatów P_2 należące do języka L_1 wynikały z postulatów P_1; drugi żąda tego samego od tych tylko konsekwencji logicznych postulatów P_2, które są zdaniami czysto uniwersalnymi języka L_1.

Warunek C1 nie ma, jak wspominałem, odpowiednika czysto syntaktycznego; ściślej — nie ma takiego odpowiednika odnoszącego się do języka L_2. Nie jest nim w szczególności warunek C2′ — standardowy warunek nietwórczości zbioru P_2 ze względu na zbiór P_1. C1 pociąga C2′, lecz nie na odwrót. Można okazać, iż istnieją takie zbiory zdać P_1 i P_2, które spełniają warunek C2′, a nie spełniają warunku C1[8]. Jest to fakt ważny i ciekawy. Jeśli zbiór P_2 jest nietwórczy ze względu na zbiór P_1, nie nakłada on na interpretację terminów języka L_1 żadnych innych warunków wyrażalnych w języku L_1, prócz tych, które nakłada zbiór P_1. Okazuje się jednak, że może on nakładać na interpretację tych terminów pewne warunki różne od P_1 niewyrażalne w języku L_1! Stąd też istnieć mogą w takiej sytuacji modele zbioru P_1, które nie są fragmentami żadnego z modeli zbioru P_2. Warunek C2′ pociąga warunek C1 jedynie w pewnych przypadkach szczególnych. Oto niektóre z nich:

1. C2′ implikuje C1, gdy P_2 jest zbiorem zdań czysto uniwersalnych.[9]
2. C2′ implikuje C1 wtedy, gdy każdy model zbioru P_1 jest modelem o uniwersum skończonym (a więc wtedy, gdy konsekwencją P_1 jest twierdzenie o istnieniu co najwyżej n przedmiotów).

W przypadkach, w których realizacja warunku nietwórczości nie gwarantuje realizacji warunku C1, użytkownik języka L_2 nie dysponuje żadnym dostępnym sobie kryterium pozwalającym stwierdzić, czy warunek C1 został spełniony. Nie może tego rozstrzygnąć pozostając «wewnątrz» swego języka; aby to uczynić, musi wyjść poza jego granice. Stąd — znaczenie przywiązywane do istnienia kryteriów synktatycznych dla warunków takich, jak rozważane obecnie, i problematyczna nieco przydatność w praktyce naukowej warunków, które kryteriów takich nie posiadają.

2. Analiza semantyczna danego języka empirycznego zakłada konieczność określenia zbioru postulatów dla jego terminów pozalogicznych. Analizując język L_2, wyróżnić musimy w szczególności zbiór postulatów P_2 dla nowo wprowadzonych predykatów $q_1, ..., q_m$. W jaki sposób możemy to uczynić? W skład procedury wzbogacania danego języka empirycznego o nowe terminy teoretyczne wchodzi zawsze akt uznania przez użytkowników tego języka określonego zbioru zdań charakteryzujących terminy wprowadzane. Czy ów zbiór zdań może być uznany za zbiór językowych postulatów? Załóżmy, iż zbiorem zdań uznanych w procedurze wzbogacania języka L_1 o predykaty $q_1, ..., q_m$ jest zbiór T_2. Czy zbiór ten może być utożsamiony ze zbiorem P_2 — postulatów wyznaczających w przedstawiony wyżej sposób klasę modeli właściwych języka L_2? Zbiór P_2, jak widzieliśmy, nie może być zbiorem dowolnym. Jeśli

8) Por. m.in. J. Shoenfield, wyd. cyt.

9) Por. J. Łoś, wyd. cyt.

interpretacja języka L_2 ma być zgodna z przyjętymi założeniami semantycznymi, zbiór P_2 musi spełniać określone warunki. Musi to być, mówiąc najogólniej, zbiór w pewnym sensie nietwórczy (ze względu na zbiór P_1). W zależności od tego, jak określamy klasę M_2^* modeli właściwych języka L_2 — zgodnie z definicją D1, D2, czy D3, ów warunek nietwórczości przybiera odpowiednio sprecyzowaną postać — C1, C2, lub C3. Otóż nie jest bynajmniej rzeczą wykluczoną, że zbiór T_2 zdań faktycznie uznanych warunku takiego nie spełnia. Sądzę, że jest to sytuacja nie tylko możliwa, lecz często w rzeczywistej praktyce naukowej spotykana. Nie uzasadniając tutaj tego przypuszczenia, chcę zwrócić jedynie uwagę na fakt, że w naukach empirycznych terminy teoretyczne bywają nierzadko charakteryzowane przez ogół aksjomatów pewnej teorii empirycznej, wśród których nie sposób na drodze pragmatycznej wyróżnić twierdzeń definicyjnych i hipotez rzeczowych. Jest rzeczą jasną, że ogół aksjomatów takiej teorii warunku nietwórczości spełniać nie może, gdyż byłoby to niezgodne z jej empirycznym charakterem. Toteż dokonując logicznej rekonstrukcji języka takiej teorii nie możemy uznać ogółu jej aksjomatów za zbiór postulatów dla terminów wprowadzanych. Wyodrębnienie tego zbioru staje się zadaniem logika, który tej rekonstrukcji dokonuje. Jak zadanie to można rozwiązać?

Niech T_2 symbolizuje, jak wyżej, zbiór zdań uznanych przez użytkowników języka L_2 przy wprowadzaniu predykatów $q_1, ..., q_m$. Załóżmy, iż rodzaj interpretacji języka L_2 podpada pod schemat odpowiadający którejś z uwzględnionych przez nas definicji klasy M_2^* — D1, D2 lub D3. Zbiór zdań języka L_2, który uznany być może za występujący w tych definicjach zbiór postulatów P_2, spełniać musi odpowiedni warunek nietwórczości — C1, C2 lub C3. Jeśli zbiór T_2 warunek taki spełnia, za zbiór P_2 uznamy po prostu zbiór T_2. W wypadku przeciwnym, w ten sposób postąpić nie możemy. Zbiór T_2 traktować musimy jako zbiór, który prócz postulatów dla predykatów q_1, ..., q_m obejmuje również pewne twierdzenia rzeczowe, i jako zbiór P_2 przyjąć musimy zbiór odpowiednio słabszy. Jakie określać go mają warunki? Musi to być przede wszystkim zbiór, który spełnia odpowiedni warunek nietwórczości — C1, C2 lub C3. Musi to być ponadto zbiór, który „odpowiada" zbiorowi T_2, który, innymi słowy, obejmuje zawarte w tym właśnie zbiorze postulaty. Jest to żądanie dość nieokreślone, dopuszczające różne eksplikacje. Podam tu jedną z nich, wyjaśniając naprzód jej sens w zastosowaniu do schematu odpowiadającego definicji D2.

Warunek nietwórczości, jaki spełniać ma zbiór P_2, przybiera w tej sytuacji postać warunku C2, lub równoważnego mu warunku C2'. Zbiór P_2 nie może tu pociągać żadnych zdań języka L_1, nie będących konsekwencjami zbioru P_1:

$$Cn\,(P_2) \cap Z_1 - Cn\,(P_1) = \emptyset.$$

Jednocześnie, jeśli zbiór T_2 warunku takiego nie spełnia, stwierdzić musimy, iż:

$$Cn\,(T_2) \cap Z_1 - Cn\,(P_1) \neq \emptyset.$$

Zbiór powyższy reprezentuje ogół tych warunków, które na interpretację języka L_1 nakłada zbiór T_2, a których nie może nakładać zbiór P_2. Wydaje się więc, iż zbiór P_2

powinien być zbiorem na tyle mocnym, aby po dołączeniu zbioru powyższego stawał się równoważny logicznie zbiorowi T_2. Takie właśnie żądanie formułuje warunek:
A. $Cn\,[(Cn\,(T_2) \cap Z_1 - Cn\,(P_1)) \cup P_2] = Cn\,(T_2)$.
Zbiór P_2 możemy więc określić jako zbiór, który spełnia warunki C2′ i A.

Ten sam warunek A służyć może do określenia zbioru P_2 w sytuacji odpowiadającej definicji D1. Tutaj co prawda zbiór P_2 spełniać musi warunek nietwórczości C1, mocniejszy od warunku C2′. Mówiąc swobodnie, zbiór P_2 nie może nakładać na interpretację języka L_1 żadnych warunków różnych od tych, które nakłada zbiór P_1 — zarówno wyrażalnych w języku L_1, jak i niewyrażalnych w tym języku. Ale dlatego też, zbiór reprezentujący te wszystkie — wyrażalne w języku L_1 — warunki, które na interpretację tego języka nakłada zbiór T_2, a których nie może nakładać zbiór P_2, jest zbiorem tym samym co poprzednio:

$$Cn\,(T_2) \cap Z_1 - Cn\,(P_1).$$

Niczego więcej nie wykraczającego poza nasz język L_1 do zbioru P_2 i tu dodać nie możemy. W rezultacie, zbiór P_2 określony może zostać w tym przypadku jako zbiór spełniający warunki C1 i A.

Inaczej natomiast sformułować musimy omawiany warunek dla schematu odpowiadającego definicji D3. Zbiór P_2 spełniać musi częściowy tylko warunek nietwórczości, C3, który w swej wersji synktatycznej C3′ żąda, aby:

$$Cn\,(P_2) \cap A_1 — Cn\,(P_1) = \emptyset,$$

tj, aby zbiór P_2 nie pociągał żadnych czysto uniwersalnych zdań języka L_1 nie będących konsekwencjami zbioru P_1. Wszelkie zatem ograniczenia, które na interpretację języka L_1 nakłada zbiór T_2 , a których nie może nakładać zbiór P_2, reprezentuje zbiór:

$$Cn\,(T_2) \cap A_1 - Cn\,(P_1).$$

Zbiór ten w połączeniu ze zbiorem P_2 powinien więc tworzyć zbiór równoważny logicznie zbiorowi T_2:
B. $Cn\,[(Cn\,(T_2) \cap A_1 - Cn\,(P_1)) \cup P_2] = Cn\,(T_2)$.
W sytuacji tej zbiór P_2 określamy więc jako zbiór czyniący zadość warunkom C3′ i B.

Zestawmy wyniki tych rozważań, podając warunki charakteryzujące zbiór P_2 przy każdej z wyróżnionych tu interpretacji języka L_2. Zbiór P_2 ma być zbiorem zdań języka L_2, który (dla danych zbiorów P_1 i T_2) spełnia:

(1) w przypadku definicji D1 — warunki C1 i A;

(2) w przypadku definicji D2 — warunki C2′ i A;

(3) w przypadku definicji D3 — warunki C3′ i B.

Zauważmy przede wszystkim, że wtedy, gdy zbiór T_2 sam spełnia odpowiedni warunek nietwórczości, zbiór P_2 staje się na mocy powyższych określeń równoważny logicznie zbiorowi T_2. Ten ostatni więc może być uznany w takich przypadkach za zbiór postulatów P_2. W przypadkach pozostałych musi być to zbiór odpowiednio słabszy. Warto może podkreślić fakt, że w przypadkach tych przyjęte określenia nie wyznaczają zbioru P_2 w sposób jednoznaczny. Łatwo podać można przykłady takich nierównoważnych logicznie zbiorów P_2' i P_2'', z których każdy spełnia (dla ustalonych zbiorów P_1 i T_2)

warunki wyszczególnione pod (1), a tym samym — słabsze od nich — warunki wyszczególnione pod (2)[10]. Różnica między takimi zbiorami wydaje się z semantycznego punktu widzenia nieistotna. Każdy z nich uważany być może za zbiór postulatów odpowiadający zbiorowi T_2. O wyborze między nimi decydować muszą względy natury pragmatycznej.

Najważniejszym problemem, jaki nasuwają przyjęte określenia, jest problem istnienia zbiorów spełniających wyszczególnione w nich warunki. Czy dla dowolnych zbiorów P_1 i T_2 istnieje zawsze zbiór P_2, który warunki takie spełnia? Różnie na to pytanie wypadnie odpowiedzieć w zależności od tego, o które warunki chodzi. W przypadku (1) i (2) odpowiedź jest przecząca, w przypadku (3) — twierdząca. Problem istnienia zbiorów spełniających warunki C1 i A oraz zbiorów spełniających warunki C2′ i A, był przedmiotem szczegółowych badań prowadzących m.in. do następujących wyników:

1. Dla pewnych zbiorów P_1 i T_2 nie istnieje ani taki zbiór P_2, który spełnia warunki C1 i A, ani taki, który spełnia warunki C2′ i A.
2. Dla pewnych zbiorów P_1 i T_2 nie istnieje zbiór P_2 spełniający warunki C1 i A, istnieje natomiast zbiór P_2 spełniający warunki C2′ i A[11].

Problem istnienia zbiorów spełniających warunki C3′ i B ma, jak wspomniałem, rozstrzygnięcie pozytywne:

Dla dowolnych zbiorów P_1 i T_2 istnieje zawsze zbiór P_2 spełniający warunki C3′ i B.

Ponieważ problem ten, w przeciwieństwie do poprzednich, do tej pory rozważany nie był, podam szkic dowodu powyższego twierdzenia. Przypominam w tym celu nałożone na zbiór P_2 warunki:

C3′. $Cn\,(P_2) \cap A_1 \subset Cn\,(P_1)$;

B. $Cn\,[(Cn\,(T_2) \cap A_1 - Cn\,(P_1)) \cup P_2] = Cn\,(T_2)$.

Jeśli $Cn\,(T_2) \cap A_1 - Cn\,(P_1) = \emptyset$, przyjmiemy, iż $P_2 = T_2$ Rozpatrzmy przypadek, gdy $Cn\,(T_2) \cap A_1 - Cn\,(P_1) \neq \emptyset$, i załóżmy, że zdanie $\alpha \in Cn\,(T_2) \cap A_1 - Cn\,(P_1)$. Jako zbiór P_2 przyjmiemy w tym przypadku zbiór wszystkich zdań warunkowych otrzymanych ze zdań zbioru T_2 przez poprzedzenie każdego z nich wspólnym poprzednikiem α:

$$P_2 = \{\alpha \to \beta_i\}_{\beta_i \in T_2}$$

Należy okazać, że tak określony zbiór P_2 spełnia warunki C3′ i B. Fakt spełniania warunku B jest oczywisty. Chcąc okazać, iż P_2 spełnia warunek C3′, załóżmy, iż $\gamma \in$ $\in Cn(P_2) \cap A_1$. Skoro $\gamma \in Cn\,(P_2) \cap A_1$, γ jest konsekwencją pewnego skończonego podzbioru zbioru zdań P_2, a więc: $\gamma \in Cn\,(\{\alpha \to \beta_1, ..., \alpha \to \beta_k\})$, dla pewnego k. To zaś równoważne jest stwierdzeniu, iż $\gamma \in Cn\,(\{\alpha \to \beta_1 \wedge ... \wedge \beta_k\})$. Łatwo zauważyć, iż tak

10) Przykłady takich zbiorów konstruowane były dla pewnych przypadków szczególnych, w których zbiór T_2 stanowiła para tzw. zdań redukcyjnych. Por. m.in. *The Logic of Empirical Theories*, wyd. cyt.

11) Problem ten rozważany był w pracy M. Przełęckiego i R. Wójcickiego „Inessential parts of extensions of first-order theories", *Studia Logica* 28, 1971. Pojęcia, których owe rozważania dotyczą, różnią się nieco od przyjętych tu określeń (1) i (2). Dowody podanych wyżej twierdzeń można jednak otrzymać z dowodów zawartych w tamtej pracy przez nieznaczną tylko modyfikację tych ostatnich.

może być tylko wtedy, gdy $\gamma \in Cn(\{\sim\alpha\})$. Skoro α jest z założenia zdaniem czysto uniwersalnym, $\sim\alpha$ jest zdaniem czysto egzystencjalnym. Ponieważ γ ma być również zdaniem czysto uniwersalnym, łatwo można okazać (na podstawie twierdzenia o interpolacji), że wynika ono ze zdania $\sim\alpha$ tylko wtedy, gdy bądź $\alpha \in Cn(\emptyset)$ bądź $\gamma \in Cn(\emptyset)$. Ale przypadek pierwszy jest niemożliwy, bo z założenia $\alpha \notin Cn(P_1)$. Zachodzić zatem musi ewentualność druga. Tym samym $\gamma \in Cn(P_1)$, a to właśnie należało okazać.

Tak więc wtedy, gdy interpretacja języka L_2 podpada pod schemat definicji D3 i, co za tym idzie, zbiór jego postulatów P_2 określony zostaje przez przewidziane dla tej sytuacji warunki (3), mamy gwarancję, że żbiór tak określony istnieje i znamy sposób jego konstrukcji. W sytuacjach pozostałych gwarancji takiej nie mamy. Okazać się więc może w pewnym konkretnym przypadku, że zbiór P_2 określony przez przewidziane dla danej sytuacji warunki (1) lub (2) w ogóle nie istnieje. Chcąc wyróżnić i tu pewien zbiór postulatów dla języka L_2, musimy określić go nieco inaczej. Możemy tego dokonać osłabiając w odpowiedni sposób charakteryzujący ów zbiór warunek A.[12]

3. Na koniec — parę uwag w sprawie stopnia ogólności przedstawionych tu konstrukcji. Języki, do których ograniczyliśmy nasze rozważania — to języki niezmiernie proste: oparte na węższym rachunku predyktatów (z identycznością) i nie zawierające innych terminów pozalogicznych prócz predyktatów. Powstaje wobec tego wątpliwość, czy rozważania nasze stosować się mogą do rzeczywistych języków empirycznych, w szczególności do języków typowych teorii empirycznych, takich jak teorie fizykalne. Rysem charakterystycznym tych teorii jest to, iż posługują się one rozbudowanym aparatem matematycznym. Czy można z niego zdać sprawę w ramach przyjętych przez nas konstrukcji? Czy charakter rozważanego tu języka L i jego interpretacji M^* dopuszcza taką możliwość? Odpowiedź na to pytanie zależy w dużej mierze od tego, jak bogaty jest ów aparat. Jeśli wykracza on poza środki elementarne angażując pojęcia takie, jak ogólne pojęcie zbioru, czy funkcji, nie możemy zdać z niego sprawy bez istotnego rozszerzenia przyjętych tu konstrukcji, gdyż są to konstrukcje z założenia ograniczone do języków elementarnych. Wiele z istniejących teorii fizykalnych taki nieelementarny aparat bez wątpienia stosuje. Ale istnieje również, jak się zdaje, wiele interesujących teorii fizykalnych, które nie wykraczają poza aparat elementarny; co ciekawsze , niektóre — z pozoru nieelementarne teorie fizykalne — dają się bez istotnego zubożenia sformułować w języku elementarnym.[13] Otóż zastosowanie zarysowanych przez nas konstrukcji do języków elementarnych teorii fizykalnych nie wydaje się przedstawiać jakichś trudności zasadniczych. Wymaga jednak pewnych

12) Pewne sugestie w tej sprawie zawiera praca M. Przełęckiego i R. Wójcickiego „Inessential parts of extensions of first-order theories", cytowana wyżej, oraz praca tych samych autorów pt. „The problem of analyticity", *Synthese* 19, 1969

13) R. Montague w pracy „Deterministic Theories", zamieszczonej w zbiorze *Decisions, Values and Groups*, 1962, pokazuje szczegółowo, jak można sformułować w ten sposób klasyczną mechanikę punktu materialnego.

wyjaśnień i zastrzeżeń. Zarówno bowiem te języki, jak i ich interpretacje, różnią się na pierwszy rzut oka wyraźnie od rozważanych do tej pory.

Język takiej teorii fizykalnej — nazwijmy go językiem L_f — ujmowany bywa z reguły jako tzw. język wielotypikalny i zawierający symbole funkcyjne. Wyróżnia się w nim, w przypadku najprostszym, dwa typy zmiennych: $x_1, x_2, \ldots$ i $y_1, y_2, \ldots$, a jako stałe pozalogiczne symbole funkcyjne dwóch rodzajów: $g_1, \ldots, g_k$ oraz $f_1, \ldots, f_l$. Modele takiego języka L_f przybierają postać tzw. modeli dwuzakresowych:

$$\langle U_1, U_2; G_1, \ldots, G_k, F_1, \ldots, F_l\rangle,$$

gdzie U_1 - to zbiór wartości zmiennych $x_1, x_2, \ldots, U_2$ —to zbiór wartości zmiennych $y_1, y_2, \ldots, G_1, \ldots, G_k$ — to funkcje o argumentach i wartościach ze zbioru U_1, a $F_1, \ldots, F_l$ — to funkcje o argumentach ze zbioru U_2, a wartościach ze zbioru U_1. Zgodnie z zamierzoną interpretacją języka L_f, U_1 ma być zbiorem liczb (najczęściej zbiorem liczb rzeczywistych), U_2 — pewnym zbiorem przedmiotów fizycznych, $G_1, \ldots, G_k$ — pewnymi operacjami matematycznymi, a $F_1, \ldots, F_l$ — pewnymi wielkościami fizycznymi. Zbiór U_1 i funkcje $G_1, \ldots, G_k$ reprezentują więc zakładany w danej teorii aparat matematyczny.

Mimo wyraźnej odmienności tego języka i jego modeli od naszego języka L i jego modeli (T) — różnice te mają w istocie charakter jedynie techniczny. Wiadomo bowiem, że wszystko, co da się wyrazić w wielotypikalnym języku z funkcjami, da się też wyrazić, mówiąc swobodnie, w odpowiednio dobranym jednotypikalnym języku bez funkcji, a więc w języku takim, jak język L. Nie mogę tu precyzować, ani uzasadniać tego — dobrze zresztą znanego — twierdzenia (czyni to każdy obszerniejszy podręcznik logiki). Poprzestanę zatem na paru uwagach wyjaśniających.

Przejście od języka z funkcjami do odpowiadającego mu języka bez funkcji sprowadza się pod względem synktatycznym do zastąpienia każdego k–argumentowego symbolu funkcyjnego k+1–argumentowym predykatem. Otrzymujemy w ten sposób zamiast języka z symbolami funkcyjnymi $g_1, \ldots, g_k, f_1, \ldots, f_l$ język zawierający $k+l$ odpowiadających im predykatów $r_1, \ldots, r_k, q_1, \ldots, q_l$. Przejście z kolei od języka dwutypikalnego do języka jednotypikalnego polega na przyjęciu jednego typu zmiennych i wprowadzeniu zamiast dwóch poprzednich typów zmiennych dwóch dodatkowych jednoargumentowych predykatów s_1, s_2. W rezultacie dochodzimy na tej drodze do języka tego samego rodzaju, co języki przez nas rozważane — zawierającego jeden tylko typ zmiennych i $n = 2+k+l$ predykatów $s_1, s_2, r_1, \ldots, r_k, q_1, \ldots, q_l$. Nazwijmy go, jak poprzednio, językiem L. Zakładamy, iż jego modele, czyli układy typu:

$$\langle U; S_1, S_2, R_1, \ldots, R_k, Q_1, \ldots, Q_l\rangle$$

związane są z modelami języka L_f w sposób następujący:

1. Uniwersum U jest sumą obu zakresów poprzednich: $U=U_1 \cup U_2$.
2. Podzbiory S_1, S_2 pokrywają się odpowiednio z zakresami U_1, U_2: $S_1=U_1, S_2=U_2$.
3. Każda relacja $R_i (i=1, \ldots, k)$ odpowiada funkcji G_i zgodnie ze schematem:

$$R_i (x_1, \ldots, x_{k_i}, y) \equiv G_i (x_1, \ldots, x_{k_i}) = y;$$

w ten sam sposób każda relacja Q_i $(i=1, \ldots, l)$ odpowiada funkcji F_i.

Założenia te czynią intuicyjnie zrozumiałym sposób, w jaki to wszystko, o czym mówi się w języku L_f, wyrażone może być w języku L. (Wystarczy w tym celu, mówiąc swobodnie, zastąpić w danym zdaniu języka L_f zwroty z symbolami funkcyjnymi przez odpowiednie predykaty, a zmienne związane ujednolicić i zrelatywizować do predykatu s_1 lub s_2).

Wspominaliśmy, iż zgodnie z zamierzoną interpretacją języka L_f pewne jego terminy otrzymują interpretację matematyczną. Przy przejściu do języka L rolę tych terminów przejmują predykaty s_1, r_1, ..., r_k. Otóż stać można na stanowisku, że dana teoria wyposażona zostaje w stosowny aparat matematyczny wtedy tylko, gdy predykaty te denotują jednoznacznie określone, z góry ustalone twory matematyczne. I tak, predykat s_1 denotować musi zbiór liczb rzeczywistych — oznaczmy go przez $\mathbf{S}_1$, a predykaty r_1, ..., r_k — określone relacje zachodzące między tymi liczbami (odpowiadające operacji dodawania, mnożenia itp.) — oznaczmy je przez $\mathbf{R}_1$, ..., $\mathbf{R}_k$. Założenie takie wyrażalne jest oczywiście w przyjętej przez nas terminologii. Żąda ono, aby do klasy M^* modeli właściwych rozważanego języka L należały takie tylko modele (1):

$$\langle U; \mathbf{S}_1, S_2, \mathbf{R}_1, ..., \mathbf{R}_k, Q_1, ..., Q_l \rangle,$$

w których predykaty s_1, r_1, ..., r_k otrzymują ową ustaloną interpretację matematyczną: $\mathbf{S}_1$, $\mathbf{R}_1$, ..., $\mathbf{R}_k$. Nazwijmy je modelami standardowymi języka L. Przy takim jednak warunku nakładanym na klasy modeli właściwych rozważanych w tej pracy języków, pewnej modyfikacji musiałyby ulec niektóre wywody zawarte w jej części drugiej. Oparte one bowiem były na założeniu, iż klasa modeli właściwych M^* wyodrębniona została z klasy modeli postulatów $M(P)$ za pomocą bezpośrednich zabiegów interpretacyjnych typu definicji ostensywnej. Był to jedyny rodzaj bezpośrednich procedur interpretacyjnych, do jakich odwoływaliśmy się w pracy obecnej zgodnie z jej empirystycznymi założeniami. W przypadku rozważanym, ten rodzaj interpretacji bezpośredniej nie wchodzi oczywiście w grę. Wyodrębnienie klasy modeli standardowych języka L odwoływać się musi do interpretacji bezpośredniej innego rodzaju i inne, wskutek tego, pociągającej konsekwencje.

Sądzę jednak, że fakt wyposażenia teorii fizykalnej w stosowny aparat matematyczny pojmowany może być również w sposób bardziej liberalny, w pełni harmonizujący z konstrukcjami zarysowanymi w pracy obecnej i nie wymagający jakiejkolwiek ich modyfikacji. Można mianowicie terminy matematyczne traktować jako terminy dopuszczające — jak wszelkie terminy teoretyczne — interpretację jedynie pośrednią, a więc daną przez określony zbiór postulatów. Zbiór ten wchodziłby po prostu w skład zbioru postulatów P i w ten sposób wyznaczał interpretację terminów matematycznych języka L. Chcąc, aby interpretacja ta zdeterminowana była jak najściślej, musimy jako zbiór postulatów dla predykatów s_1, r_1, ..., r_k przyjąć pewien zbiór maksymalny. Może nim być po prostu ogół prawd matematycznych wyrażalnych w języku L. Oznaczamy go symbolem Mt i definiujemy, jak następuje:

$\alpha \in Mt$ wtedy, gdy α jest zdaniem prawdziwym w każdym modelu standardowym języka L.

Warunek żądający, aby każdy model właściwy języka L był modelem zbioru zdań Mt, jest oczywiście warunkiem słabszym, niż warunek żądający, aby był to model standardowy. Warto jednak zwrócić uwagę na fakt, że w wielu zastosowaniach — ważnych z logicznego punktu widzenia — różnica ta nie odgrywa istotnej roli. Ilustracją tego może być zależność następująca:

Jeśli zbiór zdań X jest zbiorem skończonym, zdanie α jest prawdziwe w każdym modelu standardowym zbioru X wtedy i tylko wtedy, gdy $\alpha \in Cn\,(Mt \cup X)$.[14]

Przemawia to za tym, iż ów bardziej liberalny sposób interpretacji terminów matematycznych pozwala również zdać sprawę z aparatu matematycznego zakładanego przez niektóre języki empiryczne. I one, co za tym idzie, objęte zostają naszymi rozważaniami.

14) Twierdzenie to przytacza R. Montague w cytowanej wyżej pracy.

Filozofia Nauki
Rok I, 1993, Nr 2-3

Dwa podejścia do logicznej struktury teorii naukowych: teoriomnogościowe *versus* teoriomodelowe*
Kilka uwag o książce J. Sneeda
The logical structure of mathematical physics.

Monografia Sneeda *The logical structure of mathematical physics*, bez wątpienia jedna z najważniejszych prac o tej tematyce, które ukazały się ostatnio, zasługuje na szczegółową i dokładną analizę. Przedstawione przeze mnie uwagi z całą pewnością nie będą w stanie oddać w pełni bogactwa i doniosłości poruszanych w niej zagadnień. Skoncentrują się one na tym, co stanowi główną ideę dzieła: na koncepcji logicznej struktury twierdzeń empirycznych występujących w teoriach fizycznych, rozwiniętej w pierwszych pięciu rozdziałach książki. Zagadnienia poruszane w pozostałych trzech rozdziałach omówione zostaną tylko pobieżnie. W szczególności nie będę zajmować się problemem relacji równoważności i redukowalności, zachodzących między teoriami fizycznymi, pomimo że zagadnienie to zostało opracowane w szczególnie wnikliwy i oryginalny sposób. Opis logicznej struktury teorii fizycznych został przeprowadzony w kilku krokach, z których każdy jest pewnym udoskonaleniem poprzednich. Przedstawię tutaj w bardzo schematyczny i uproszczony sposób te propozycje, a następnie porównam je z pewnym rodzajem opisu „tradycyjnego" (w sensie Sneeda). O ile podejście autora można nazwać teoriomnogościowym, o tyle podejście tradycyjne, które chcę przedstawić, można nazwać teoriomodelowym (albo semantycznym). Aby to porównanie przeprowadzić w prosty i przejrzysty sposób, nie będę zachowywał orginalnej symboliki autora przy przedstawianiu jego idei.

---

* Artykuł ten jest rozszerzoną wersją odczytu, wygłoszonego przez autora podczas Seminarium Metodologicznego Sekcji Logiki IFiS PAN, 9 grudnia 1972 roku.

I

Wszystkie twierdzenia autora są oparte na oryginalnej metodzie aksjomatyzacji teorii naukowych, po raz pierwszy wprowadzonej przez Suppesa, mianowicie aksjomatyzacji za pomocą definicji predykatu teoriomnogościowego. Predykat teoriomnogościowy ma charakteryzować strukturę matematyczną danej teorii fizycznej. Predykat ten zostaje następnie użyty do budowy twierdzeń empirycznych tej teorii. Zilustrujmy, za Sneedem, to pojęcie przy pomocy predykatu teoriomnogościowego zbliżonego do tych, które występują w rzeczywistych teoriach fizycznych, ale nieco mniej skomplikowanego:

x jest S zawsze i tylko, gdy istnieje D, n i t takie, że:

(1) $x = \langle D, n, t \rangle$;

(2) D jest skończonym, niepustym zbiorem;

(3) n i t są funkcjami odwzorowującymi D w zbiór liczb rzeczywistych;

(4) dla każdego $y \in D$, $t(y) > 0$;

(5) $\sum_{y \in D} n(y) \cdot t(y) = 0$.

Obiekty, które są S, będziemy nazywać modelami dla S. Mają one pewną swoistą strukturę matematyczną. Wszystko, o czym «z sensem» można orzec, że jest S, musi mieć pewną minimalną strukturę teoriomnogościową, scharakteryzowaną przez warunki (1) - (3):

x jest S_0 zawsze i tylko, gdy istnieje takie D, n i t, że:

(1) $x = \langle D, n, t \rangle$;

(2) D jest skończonym, niepustym zbiorem;

(3) n i t są funkcjami odwzorowującymi D w zbiór liczb rzeczywistych.

Obiekty, o których można orzec S_0 będziemy nazywać możliwymi modelami dla S.

Pamiętając o tym przykładzie założymy, że z każdą teorią fizyczną T związana jest klasa modeli dla T, M_T, oraz klasa możliwych modeli dla T, $\mathcal{M}$ (Oczywiście $M_T \subseteq \mathcal{M}$.) Elementy $\mathcal{M}$ będziemy oznaczać przez $\mathfrak{m}$, podklasy $\mathcal{M}$ — przez M. Najprostsza propozycja dotycząca tworzenia tez empirycznych za pomocą predykatu teoriomnogościowego może być sformułowana jako zdanie o postaci:

(1a) $\mathfrak{m}^* \in M_T$,

gdzie $\mathfrak{m}^*$, pewien możliwy model dla teorii T, reprezentuje układ fizyczny, do którego teoria ma się stosować, tzn. jej zamierzone zastosowanie. Ponieważ, zgodnie z założeniem autora, zwykle istnieje nie jedno zastosowanie $\mathfrak{m}^*$, lecz cała klasa takich zastosowań M^*, całkowita empiryczna zawartość teorii T będzie reprezentowana przez zdanie o postaci:

(1b) $M^* \subseteq M_T$.

To, czy zdanie (1b) jest zdaniem empirycznym, zależy od tego, jak jest scharakteryzowana klasa M^*. (1b) może być zdaniem empirycznym tylko wtedy, gdy spełniony jest następujący warunek:przynajmniej niektóre elementy M^* muszą być opisane w sposób, który nie zakłada, że pewien element M^* jest M_T. Jest to jednak warunek, który w typowej teorii fizyki matematycznej nie jest spełniony. Wynika to z teoretycznego charakteru pewnych funkcji występujących w takiej teorii. To zrelatywizowane do teorii pojęcie funkcji teoretycznej jest scharakteryzowane ogólnie przez autora w następujący sposób. Niech t będzie funkcją pojawiającą się w możliwym modelu $\mathfrak{M}$ teorii T, i niech M^* będzie klasą jej zamierzonych zastosowań. Funkcja t jest teoretyczna ze względu na T zawsze i tylko, gdy dla każdego $\mathfrak{M}_i^* \in M^*$ każda metoda pomiaru wartości funkcji t dla pewnych elementów z $\mathfrak{M}_i^*$ zakłada, że $\mathfrak{M}_j^* \in M_T$, dla pewnego $\mathfrak{M}_j^* \in M^*$. Jest dość oczywiste, że obecność jakiejkolwiek funkcji T-teoretycznej w teorii T powoduje, że (1b) nie jest twierdzeniem empirycznym. Każda próba uzasadnienia tego twierdzenia na gruncie empirycznym prowadzi z konieczności do błędnego koła lub do *regressus ad infinitum*.

Wszystkie pozostałe propozycje konstrukcji twierdzeń empirycznych za pomocą predykatu teoriomnogościowego są oparte na założeniu, że modele, będące zamierzonymi zastosowaniami teorii fizycznej T, nie zawierają żadnej funkcji T-teoretycznej. Jeżeli w przykładzie podanym wyżej założymy, że funkcja n nie jest, a t jest teoretyczna ze względu na S, to klasa takich modeli jest zdefiniowana w następujący sposób:

x jest P_0 zawsze i tylko, gdy istnieją D i n takie, że:

(1) $x = \langle D, n \rangle$;

(2) D jest skończonym, niepustym zbiorem;

(3) n jest funkcją odwzorowującą D w zbiór liczb rzeczywistych.

Obiekty, o których można orzec P_0, będziemy nazywać możliwymi modelami częściowymi dla S.

Aby dokonać tego rozróżnienia w sposób ogólniejszy, wprowadzimy kilka kolejnych symboli. Klasę możliwych modeli częściowych dla teorii T, tzn. modeli zawierających wyłącznie funkcje, które nie są T-teoretyczne, będziemy oznaczać przez $\mathcal{M}_0$; jej elementy przez $\mathfrak{M}_0$; jej podklasy przez M_0. Dla każdego $\mathfrak{M} \in \mathcal{M}$, $\mathfrak{M}|_0$ będzie oznaczać element M_0 otrzymany z $\mathfrak{M}$ przez usunięcie z niego wszystkich funkcji T-teoretycznych. Jeżeli $\mathfrak{M}_0 = \mathfrak{M}|_0$, $\mathfrak{M}_0$ będziemy nazywać obcięciem modelu $\mathfrak{M}$, a $\mathfrak{M}$ — przedłużeniem modelu $\mathfrak{M}_0$ (zgodnie z powszechnie przyjętą terminologią teoriomodelową). Dla każdego $M \subseteq \mathcal{M}$, $M|_0$ będzie oznaczać klasę wszystkich obcięć elementów M. (Oczywiście $\mathcal{M}|_0 = \mathcal{M}_0$.) Druga propozycja rekonstruowania twierdzeń empirycznych formułowanych w teorii T może być teraz przedstawiona w następujący sposób:

(2a) $\mathfrak{M}_0^* \in M_T|_0$,

gdzie $\mathfrak{M}_0^*$, jest możliwym modelem częściowym dla T, reprezentującym jedno z jej zamierzonych zastosowań. Pełna treść empiryczna teorii T będzie wyrażona przez zdanie postaci:

(2b) $M_0^* \subseteq M_T|_0$,

gdzie M_0^* oznacza klasę wszystkich zamierzonych zastosowań.

Chociaż zdaniu (2b), w przeciwieństwie do zdania (1b), można by przyznać charakter empiryczny, jest ono, w opinii autora, wciąż wadliwe pod innymi względami. (2b) stwierdza, że każde zamierzone zastosowanie teorii T może być przedłużone do modelu dla T przez dodanie pewnych funkcji teoretycznych. Jednak funkcje te, chociaż związane z różnymi zastosowaniami, nie powinny być od siebie niezależne. Musimy zatem wprowadzić, jako dodatkową część empirycznej zawartości teorii, założenie, że cała klasa funkcji teoretycznych podlega pewnym ograniczeniom. Zgodnie z tymi ograniczeniami, pomiędzy wartościami funkcji teoretycznych, użytych w różnych zastosowaniach, zachodzą pewne relacje. Szczegółowy charakter tych ograniczeń może być różny dla różnych teorii. Ogólna postać takiego ograniczenia może być scharakteryzowana przy pomocy następującej definicji:

C jest ograniczeniem dla klasy M możliwych modeli dla T zawsze i tylko, gdy:

(1) C jest klasą podklas M;

(2) dla każdego $\mathfrak{M} \in M$, $\{\mathfrak{M}\} \in C$.

(Warunek (2) stwierdza, że ograniczenie nie wyklucza żadnej konkretnej funkcji, lecz jedynie pewne ich kombinacje.) Za pomocą pojęcia ograniczenia C, autor formułuje swoją trzecią propozycję, rekonstruującą empiryczną treść teorii T jako zdanie o postaci:

(3) Istnieje klasa M taka, że: $M|_0 = M_0^*$, $M \subseteq M_T$ oraz $M \in C$.

Chociaż zdanie o postaci (3) wydaje się adekwatnie oddawać empiryczną treść pewnych teorii fizycznych, nie jest ono adekwatne względem wszystkich takich teorii. Istnieją teorie fizyczne, w których zakłada się, że funkcje teoretyczne mają specjalną postać dla pewnych zastosowań tych teorii. Sposób analizy takich teorii jest, ogólnie rzecz biorąc, następujący. Należy zdefiniować dodatkowe predykaty, będące «restrykcjami» podstawowego predykatu. Za pomocą tych predykatów buduje się zdanie, które głosi, że istnieją funkcje teoretyczne, które przekształcają wszystkie zamierzone zastosowania w modele dla predykatu podstawowego, niektóre zdefiniowane podklasy zamierzonych zastosowań w modele dla predykatów dodatkowych i spełniają pewne ograniczenia. Sformułujmy zdanie tego typu dla najprostszego przypadku, w którym występuje jeden dodatkowy predykat, określający podklasę M_{T_1} klasy modeli M_T, oraz jedna podklasa $M_{0_1}^*$ klasy zamierzonych zastosowań M_0^*. Przy tych założeniach empiryczna treść teorii T może być przedstawiona przez zdanie o postaci:

(4) Istnieje klasa M taka, że: $M|_0 = M_0^*$, $M \subseteq M_T$, $M \in C$ oraz dla każdego $\mathfrak{M} \in M$, jeśli $\mathfrak{M}|_0 \in M_{0_1}^*$, to $\mathfrak{M} \in M_{T_1}$.

Jest to ostatnia rozważana w monografii Sneeda propozycja, dotycząca tworzenia tez empirycznych za pomocą predykatów teoriomnogościowych. Propozycja ta jest zilustrowana przez logiczną rekonstrukcję faktycznie istniejącej teorii fizycznej — mechaniki Newtona. Jej empiryczna treść okazuje się wyrażalna przez zdanie o postaci (4) (a mówiąc ściślej — przez pewne jego uogólnienie).

Wszystkie rozważane do tej pory propozycje odnosiły się do twierdzeń empirycznych głoszonych przez daną teorię fizyczną. Jakiego rodzaju obiektem jest jednak sama teoria? Sneed daje raczej niekonwencjonalną odpowiedź na to pytanie. Teoria fizyczna T zostaje utożsamiona z parą uporządkowaną, zbudowaną, mówiąc swobodnie, z formalizmu teorii F oraz z jej zamierzonych zastosowań I:

$$T = \langle F, I \rangle.$$

Formalizm F jest scharakteryzowany przez następujące czynniki: klasę M możliwych modeli dla T, klasę M_0 możliwych modeli częściowych dla T, klasę M_T modeli dla T, oraz ograniczenie C nałożone na M:

$$F = \langle M, M_0, M_T, C \rangle.$$

Klasa I zamierzonych zastosowań jest utożsamiona z opisaną wyżej klasą M_0^*:

$$I = M_0^*,$$

lub, zgodnie z inną koncepcją, z podklasą klasy M_0^*, składającą się ze wszystkich tzw. «zastosowań paradygmatycznych».

To, co przedstawiłem wyżej, jest jedynie zarysem koncepcji przedstawionej w monografii, która analizuje to zagadnienie szczegółowo i dokładnie. Znaczna jej część poświęcona jest próbie precyzyjnego ustalenia, „czym jest teoria fizyki matematycznej i w jaki sposób odróżnić jedną tego typu teorię od innej". Wiele uwagi poświęca się też w niej pewnym problemom związanym z rozwojem teorii fizycznych — „jak się je tworzy i jak się je obala". Nie mogę podjąć tutaj dyskusji na te tematy. Ograniczę się jedynie do analizy naszkicowanych wyżej rozważań dotyczących logicznej struktury teorii fizycznych. Zrobię to, porównując te rozważania z pewnymi tradycyjnymi. próbami logicznej rekonstrukcji teorii fizycznych, w szczególności — próbami opartymi na aparacie pojęciowym teorii modeli.[1]

II

Główna różnica między podejściem teoriomnogościowym a teoriomodelowym polega na sposobie, w jaki określone są istotne składniki logicznej struktury teorii. Te teoriomnogościowe obiekty są, w ujęciu Sneeda, zdefiniowane w sposób bezpośredni, przez pewne warunki teoriomnogościowe, podczas gdy tradycyjnie są one określone w sposób pośredni, przez ich teoriomodelowe relacje do pewnych obiektów językowych. Odnosi się to, w szczególności, do głównego składnika T — klasy modeli dla T, M_T. Klasa ta, zdefiniowana wyżej przez pewien predykat teoriomnogościowy, jest na ogół traktowana jako klasa modeli (w ścisłym, teoriomodelowym sensie) pewnego zbioru

1) Próby tego typu są zawarte np. w pracy R. Montague *Deterministic theories, Decisions, Values and Groups II* (Oxford 1962) oraz w artykułach R. Wójcickiego: „Semantyczne pojęcie prawdy w metodologii nauk empirycznych" (*Studia Filozoficzne* 3/1969) i „Metody formalne w problematyce teoriopoznawczej" (*Studia Filozoficzne* 1/1972). Monografia *The logic of empirical theories* (London 1969; polska wersja — *Logika teorii empirycznych*, Warszawa 1988) i niektóre moje artykuły również odwołują się do teoriomodelowego podejścia do logicznej struktury teorii empirycznych, ale są to teorie bardziej elementarne niż teorie fizyki matematycznej.

zdań — aksjomatów teorii T. Równoważność tych dwóch charakterystyk można łatwo wykazać dla wszystkich teorii elementarnych, tzn. teorii pierwszego rzędu. Jako ilustrację rozważmy, za Sneedem, pewną wyjątkowo prostą teorię matematyczną, mianowicie teorię grup. Odpowiedni predykat teoriomnogościowy jest zdefiniowany w sposób następujący:

x jest grupą zawsze i tylko wtedy, gdy istnieją D i $\circ$ takie, że:

(1) $x = \langle D, \circ \rangle$;

(2) D jest niepustym zbiorem;

(3) $\circ$ jest funkcją odwzorowującą $D \times D$ w D;

(4) dla każdych $a, b, c \in D$, $a \circ (b \circ c) = (a \circ b) \circ c$;

(5) dla każdych $a, b \in D$ istnieje $e \in D$ takie, że $a = b \circ e$;

(6) dla każdych $a, b \in D$ istnieje $e \in D$ takie, że $a = e \circ b$.

Obiekty, które spełniają ten predykat, noszą nazwę modeli dla teorii grup. Te obiekty, które spełniają pierwsze trzy z powyższych warunków, noszą nazwę możliwych modeli dla tej teorii. Te same klasy obiektów mogą być określone pośrednio, w sposób teoriomodelowy. Język teorii grup może być utożsamiony z językiem logiki predykatów pierwszego rzędu z dwuargumentowym symbolem funkcyjnym $\circ$, będącym jedyną stałą pozalogiczną. Aksjomaty teorii grup wyrażone w tym języku mają następującą postać:

(A1) $(x)(y)(z)(x \circ (y \circ z) = (x \circ y) \circ z)$;

(A2) $(x)(y)(\exists z)(x = y \circ z)$;

(A3) $(x)(y)(\exists z)(x = z \circ y)$.

Jest oczywiste, że obiekty poprzednio nazywane możliwymi modelami dla teorii grup są po prostu modelami tego języka (w innej terminologii — jego semimodelami, realizacjami), podczas gdy obiekty nazywane modelami dla teorii grup są modelami (w ścisłym, teoriomodelowym sensie) aksjomatów (A1)-(A3).

W wypadku teorii nieelementarnych relacje te stają się bardziej skomplikowane. Dla prostoty rozważmy zatem teorie formalizowalne w języku logiki predykatów pierwszego rzędu, chociaż większość teorii fizycznych nie ma tej własności. Niech L będzie językiem logiki predykatów pierwszego rzędu, zaś M — klasą jego modeli. Teorię T utożsamimy teraz z pewnym zbiorem zdań w języku L. Jak się powszechnie przyjmuje, jest ona zbiorem konsekwencji logicznych pewnego skończonego zbioru A aksjomatów teorii T:

$$T = Cn(A).$$

Klasa modeli dla T, M_T, może być utożsamiona z klasą modeli zbioru A:

$$M_T = M(A).$$

Przy powyższych założeniach, możemy zbudować teoriomodelowe odpowiedniki niektórych z wymienionych powyżej propozycji.

Należy zauważyć, że tak rozumiana teoria może mówić coś o świecie — może formułować twierdzenia empiryczne — tylko wtedy, gdy została jej nadana odpowiednia interpretacja. Przy podejściu teoriomodelowym, ta zamierzona interpretacja może być traktowana jako klasa modeli języka teorii, która, ogólnie rzecz biorąc, odpowiada temu, co, poprzednio zostało nazwane jej klasą zamierzonych zastosowań. Modele należące do tej klasy będziemy nazywali zamierzonymi (lub właściwymi) modelami teorii T. Rozważmy najpierw wypadek, gdy interpretacja teorii T jest wyznaczona jednoznacznie, tzn. gdy istnieje tylko jeden model zamierzony $\mathcal{M}^*$. Stwierdzenie, że teoria T jest prawdziwa znaczy, że jest prawdziwa w modelu $\mathcal{M}^*$, a to, z kolei, jest równoważne stwierdzeniu, że $\mathcal{M}^*$ jest modelem aksjomatów A:

(1′a) $\mathcal{M}^* \in M(A)$.

Zdanie to może być traktowane jako odpowiednik zdania (1a). Ale interpretacja teorii T nie musi być wyznaczona jednoznacznie. Istnieją powody, aby przypuszczać, że żadna teoria empiryczna nie ma jednoznacznie wyznaczonej interpretacji. Wszystkie takie teorie charakteryzują się pewnego rodzaju wielością interpretacji. Ich interpretację zatem należy traktować jako klasę modeli zamierzonych M^*, zawierającą więcej niż jeden element. Teorię T, interpretowaną w ten sposób, nazwiemy prawdziwą, jeżeli jest prawdziwa we wszystkich modelach M^*, lub — co na jedno wychodzi — jeżeli:

(1′b) $M^* \subseteq M(A)$.

Zdanie to można traktować jako odpowiednik zdania (1b).

Status metodologiczny zdań (1′a) lub (1′b) zależy od sposobu, w jaki scharakteryzowany jest model $\mathcal{M}^*$ lub klasa modeli M^*. Dochodzmy w tym miejscu do rozważanego przez Sneeda problemu terminów teoretycznych. Istnieje «tradycyjne» podejście do tego zagadnienia, które bardzo przypomina ujęcie Sneeda. Używa się w nim podobnego, zrelatywizowanego do teorii, pojęcia terminu teoretycznego, które jednak jest rozumiane w nieco inny, bardziej «uogólniony» sposób. Mówiąc niezbyt ściśle, termin t języka L jest teoretyczny ze względu na teorię T zawsze i tylko, gdy jego zamierzona interpretacja zależy od zbioru aksjomatów A teorii T; innymi słowy — gdy t ma być interpretowany w taki sposób, aby zdania A były prawdziwe. Jest to oczywiście silniejsze znaczenie tego pojęcia niż przedstawione poprzednio. (Jak się wydaje, pokrywa się ono ze znaczeniem tego pojęcia występującym u Sneeda w przypadku, gdy M^* obejmuje tylko jeden model.) Oparte jest ono jednak na podobnej idei i wypływają z niego podobne wnioski. Występowanie w teorii T terminów T-teoretycznych powoduje, że zdania (1′a) i (1′b) nie są zdaniami empirycznymi. Istnieje zbliżony do poprzedniego sposób ominięcia tej trudności.

Niech L_0 będzie podjęzykiem języka L zawierającym jedynie te terminy z L, które nie są teoretyczne ze względu na T. Niech M_0 będzie klasą modeli języka L_0. $\mathcal{M}|_0$ będzie teraz oznaczać obcięcie modelu $\mathcal{M}$ do języka L_0 (w ścisłym, teoriomodelowym sensie), zaś $M|_0$ klasę takich obcięć elementów M. Wydaje się, że istnieją różne sposoby rekonstrukcji drugiej propozycji Sneeda przy przyjętych założeniach. Jeżeli założymy, że zamierzoną interpretacją języka L_0 (tzn. terminów nieteoretycznych teorii T) jest model $\mathcal{M}_0^*$, to zamierzoną interpretację całego języka L (tzn. wszystkich terminów teorii T) możemy utożsamić z klasą modeli M^* zdefiniowaną w następujący sposób :

$$M^* = \{\mathfrak{M}: \mathfrak{M}|_0 = \mathfrak{M}_0^* \text{ i } \mathfrak{M} \in M(A)\}.$$

Twierdzenie, że teoria T jest prawdziwa przy tej interpretacji, tzn. że

$$M^* \subseteq M(A),$$

nie jest twierdzeniem empirycznym: jego prawdziwość jest znana *a priori*. Empiryczne twierdzenie odpowiadające teorii T jest, przy tej rekonstrukcji, wyrażalne przez zdanie stwierdzające niepustość klasy M^*:

$$M^* \neq \emptyset$$

Jest to zdanie równoważne zdaniu:

(2'a) $\mathfrak{M}_0^* \in M(A)|_0$,

które jest oczywistym odpowiednikiem (2a). Rekonstrukcja ta staje się nieco bardziej skomplikowana, jeżeli zamierzona interpretacja języka L_0 jest wyznaczona przez klasę modeli M_0^*. Klasa M^*, dostarczająca interpretacji językowi L, musi być teraz scharakteryzowana przez dwa warunki:

(C1) $M^* = \{\mathfrak{M}: \mathfrak{M}|_0 \in M_0^* \text{ i } \mathfrak{M} \in M(A)\}$;

(C2) $M_0^* \subseteq M^*|_0$

Jeśli tylko klasa spełniająca te warunki istnieje, prawdziwość teorii T lub — równoważnie — zdania:

$$M^* \subseteq M(A)$$

jest z góry zapewniona. Empiryczna treść teorii T zawarta jest tutaj w twierdzeniu orzekającym istnienie klasy M^*:

Istnieje klasa spełniająca warunki (C1) *i* (C2).

Można wykazać, że twierdzenie to redukuje się do zdania:

(2'b) $M_0^* \subseteq M(A)|_0$

będącego teoriomodelowym odpowiednikiem (2b).

Wydaje się jednak, że przedstawiona wyżej logiczna rekonstrukcja teorii fizycznej posiada pewne wady; zakłada ona bowiem aprioryczny charakter zdania mówiącego o prawdziwości aksjomatów teorii. Znane są różne próby uniknięcia takich konsekwencji. Wszystkie one wskazują na niejednorodność zbioru aksjomatów wszelkich teorii fizycznych. Aksjomaty spełniają podwójną funkcję. Po pierwsze, stwierdzają pewien fakt empiryczny dotyczący funkcji nieteoretycznych; po drugie, ustalają interpretację terminów teoretycznych. Powstaje problem rozbicia zbioru aksjomatów A na dwie części: faktualną, spełniającą jedynie pierwszą funkcję, oraz definicyjną, spełniającą jedynie drugą funkcję. Nazywa się je odpowiednio syntetycznym i analitycznym składnikiem A — i oznacza odpowiednio przez A_0 i A_1. Taką ogólną charakterystykę można sprecyzować na wiele sposobów[2]. Wspomnimy tu o jednym z nich. Zgodnie z nim, owe dwa składniki A są scharakteryzowane przez następujące warunki:

(1) $M(A_0) = M(A)|_0$;

(2) $M_0 \subseteq M(A_1)|_0$;

2) Por. np. artykuły M. Przełęckiego i R. Wójcickiego „The problem of analyticity" (*Synthese* 19/1969) i „Inessential parts of extensions of first-order theories" (*Studia Logica* 28/1971).

(3) $M(A_0 \cup A_1) = M(A)$.

Warunki (1)-(3) nie są tożsame z definicjami zbiorów A_0 i A_1. Ani istnienie, ani jedyność takich zbiorów nie jest tu zapewniona. Można wykazać, że jeżeli żądamy, aby zbiory A_0 i A_1 były zbiorami zdań pierwszego rzędu, to zbiory takie nie zawsze istnieją. Jednakże gdy zbiory te istnieją, wówczas teoria T może być zrekonstruowana według poniższego schematu. Jeżeli zamierzona interpretacja L_0 jest dana przez model $\mathfrak{M}_0^*$, wówczas klasa M^*, która określa zamierzoną interpretację L, jest zdefiniowana w następujący sposób:

$$M^* = \{\mathfrak{M}: \mathfrak{M}|_0 = \mathfrak{M}_0^* \text{ i } \mathfrak{M} \in M(A_1)\},$$

gdzie A_1 jest analitycznym składnikiem A. Zdanie

$$M^* \neq \emptyset$$

równoważne zdaniu:

$$\mathfrak{M}_0^* \in M(A_1)|_0,$$

jest teraz prawdziwe *a priori*. Z drugiej strony, zdanie:

$$M^* \subseteq M(A),$$

równoważne zdaniu:

$$M^* \subseteq M(A_0),$$

gdzie A_0 oznacza syntetyczny składnik A, ma charakter empiryczny i reprezentuje twierdzenia empiryczne teorii T. Jest to po prostu zdanie stwierdzające prawdziwość teorii T. Jeśli zamierzona interpretacja L_0 jest dana poprzez klasę modeli M_0^*, definicja klasy M^* przyjmuje postać:

$$M^* = \{\mathfrak{M}: \mathfrak{M}|_0 \in M_0^* \text{ i } \mathfrak{M} \in M(A_1)\},$$

i ma te same konsekwencje co poprzednio.

Dobrze wiadomo, że dla każdego skończonego A zawsze istnieją składniki A_0 i A_1, scharakteryzowane przez warunki (1)-(3), zbudowane ze zdań nieelementarnych, tzn. zdań drugiego rzędu. Niech A_R będzie tzw. zdaniem Ramseyowskim dla A, tzn. domknięciem egzystencjalnym formuły otrzymanej przez właściwe równoczesne podstawienie odpowiednich zmiennych za wszystkie terminy T-teoretyczne w A. Składniki A_0 i A_1 można teraz zdefiniować w następujący sposób:

$$A_0 = \{A_R\}$$
$$A_1 = \{A_R \to A\}.$$

Zobaczmy, jak zaproponowane wyżej ogólne rozwiązanie stosuje się do tego szczególnego przypadku. Łatwo zauważyć, że dla A_0 i A_1, zdefiniowanych jak powyżej, definicja:

$$M^* = \{\mathfrak{M}: \mathfrak{M}|_0 = \mathfrak{M}_0^* \text{ i } \mathfrak{M} \in M(A_1)\}$$

przyjmuje postać:

$$M^* = \{\mathfrak{M}: \mathfrak{M}|_0 = \mathfrak{M}_0^* \text{ i : jeśli } \mathfrak{M}|_0 \in M(A)|_0, \text{ to } \mathfrak{M} \in M(A)\},$$

a zdanie:

$$M^* \subseteq M(A_0)$$

redukuje się do zdania:

(2'a) $\mathfrak{M}_0^* \in M(A)|_0$,

Jest to oczywiście teoriomodelowy odpowiednik zdania (2a) w rekonstrukcji Sneeda. Podobnie definicja:

$$M^* = \{\mathfrak{M}: \mathfrak{M}|_0 \in M_0^* \text{ i } \mathfrak{M} \in M(A_1)\},$$

przyjmuje postać:

$$M^* = \{\mathfrak{M}: \mathfrak{M}|_0 \in M_0^* \text{ i : jeśli } \mathfrak{M}|_0 \in M(A)|_0, \text{ to } \mathfrak{M} \in M(A)\},$$

a zdanie:

$$M^* \subseteq M(A_0)$$

staje się równoważne zdaniu:

(2'b) $M_0^* \subseteq M(A)|_0$

które nie jest niczym innym, jak tylko teoriomodelowym odpowiednikiem zdania (2b). Przy takiej rekonstrukcji, empiryczne twierdzenia teorii fizycznej stają się równoważne zdaniom stwierdzającym prawdziwość aksjomatów tej teorii. Jest to wniosek dość intuicyjny i wydaje się przemawiać za słusznością tego podejścia.

III

Przeprowadzona przez nas dotąd teoriomodelowa rekonstrukcja oparta była na nierealistycznym założeniu, że rozważana teoria jest formalizowalna za pomocą języka rachunku predykatów pierwszego rzędu. Teorie fizyczne, które posługują się pojęciami teorii mnogości i matematyki, nie spełniają tego wymogu. Spróbujmy więc go uniknąć. Można to zrobić na kilka sposobów. Wydaje się, że najprostszym z nich jest następujący. Zakłada się, że język teorii fizycznej poza terminami logicznymi i specyficznie fizycznymi zawiera pewne terminy teorii mnogości i matematyki wzięte w ich standardowej interpretacji. Niech L^+ będzie takim językiem, i niech L będzie jego podjęzykiem, zawierającym oprócz terminów logicznych jedynie specyficzne terminy fizyczne teorii T. Z czysto syntaktycznego punktu widzenia, L^+ może być zrekonstruowany jako język pierwszego rzędu. Z semantycznego punktu widzenia jednakże musi on być traktowany jako język nieelementarny, ponieważ we wszystkich jego modelach symbole matematyczne i teoriomnogościowe mają taką samą, standardową interpretację. Modele tego rodzaju będziemy nazywać standardowymi modelami języka L^+. Przez klasę M^+ będziemy teraz rozumieć klasę składającą się wyłącznie z modeli standardowych dla L^+. M będzie klasą obcięć tych modeli do języka L. Podobnie, dla każdego zbioru A^+ zdań języka L^+, klasa $M^+(A^+)$ będzie zdefiniowana jako klasa modeli standardowych zbioru A^+, a $M(A^+)$ — jako klasa ich obcięć do języka L. Tak rozumiane, pojęcia te umożliwiają rekonstrukcję, w sposób przedstawiony powyżej, dowolnej teorii fizycznej, zaksjomatyzowanej za pomocą definicji pewnego predykatu teoriomnogościowego. Klasom obiektów spełniających predykaty S_0 i S odpowiadać będą pewne klasy M i $M(A^+)$ (w obecnym znaczeniu), dla odpowiednio wybranego języka L^+ i L, i zbioru zdań A^+.

Skoro elementarny charakter teorii okazuje się nieistotny dla jej teoriomodelowej rekonstrukcji, powstaje pytanie, skąd bierze się tradycyjna chęć zajmowania się elementarnymi językami i teoriami. Odpowiedź brzmi: aksjomatyzowanie teorii empirycznej w języku logiki predykatów pierwszego rzędu pozwala formułować wiele ważnych problemów metodologicznych, trudno uchwytnych w inny sposób. Jako przykład można podać problem eliminacji terminów teoretycznych. W ujęciu teoriomnogościowym Sneeda nadaje się temu problemowi duże znaczenie, ale — o ile mogę to ocenić — problem ten jest jasno postawiony tylko wtedy, gdy się go ograniczy właśnie do języków i teorii elementarnych. Niech T będzie teorią sformalizowaną w języku L, zaś A niech będzie zbiorem jej aksjomatów. Niech t będzie T-teoretycznym terminem, a L_0 — podjęzykiem L zawierającym tylko T-nie-teoretyczne terminy. Powiemy teraz, że

> t jest eliminowalny z T zawsze i tylko, gdy istnieje zbiór X zdań języka L_0 taki, że $M(X) = M(A)|_0$.

Jak łatwo zauważyć, jeśli taki zbiór istnieje, musi on być równy zbiorowi $Cn(A) \cap L_0$ logicznych konsekwencji A należących do języka L_0. Tak więc definicja powyższa redukuje się do następującej postaci:

> t jest eliminowalny z T zawsze i tylko, gdy $M(Cn(A) \cap L_0) = M(A)|_0$.

Jak powszechnie wiadomo, istnieją teorie i terminy, dla których powyższy warunek nie jest spełniony (Sneed daje interesujące przykłady takich teorii i terminów). Jeśli jednak w powyższej definicji opuścimy żądanie, aby X był zbiorem zdań elementarnych, taki zbiór będzie zawsze istniał: będzie on składał się z Ramseyowskiego zdania dla A, A_R, które oczywiście spełnia warunek:

$$M(A_R) = M(A)|_0.$$

W wypadku teorii nieelementarnych musimy jednakże usunąć owo żądanie. To, co taka teoria mówi o świecie, jest wyrażalne tylko w nieelementarnym języku L^+. Nie tylko aksjomaty A^+, ale również ich nieteoretyczne konsekwencje są zdaniami, w których używa się pewnych pojęć teoriomnogościowych i matematycznych. I to, co mówią one o nieteoretycznych obiektach, może być nie wyrażalne w żadnym języku elementarnym. Weźmy jako przykład naszą prostą teorię S. Twierdzi ona, bez użycia żadnych funkcji teoretycznych, że zbiór indywiduów D jest skończony — warunek, który nie może być sformułowany równoważnie przez żaden zbiór zdań elementarnych. W takim wypadku trudno żądać, aby pewien zbiór X zdań elementarnych spełniał warunek:

$$M(X) = M(A^+)|_0.$$

A jeśli usuniemy ograniczenie do zdań elementarnych, każdy termin teoretyczny t każdej teorii T będzie traktowany jako eliminowalny. Rozróżnienie to traci wtedy jakikolwiek sens.

Ten wniosek, jak się wydaje, stosuje się również do przykładu Sneeda. Zgodnie z jego teoriomnogościowym podejściem, mówi się raczej o eliminowalności funkcji, a nie terminów. Mówiąc niezbyt ściśle, teoretyczna funkcja t jest eliminowalna z teorii T zawsze i tylko, gdy klasa $M_T|_0$ jest definiowalna «bez użycia tej funkcji». Autor sugeruje, że pojęcie to może być uściślone przez aksjomatyzację danej teorii w pewnym języku formalnym. Jeśli jednak to zrobimy w sposób przedstawiony wyżej, to otrzymamy, jak się wydaje, nihilistyczny wniosek: skoro teoria fizyczna rozważana przez autora może być zrekonstruowana tylko jako nieelementarna, wszystkie terminy teoretyczne staną się z niej eliminowalne. W sposób oczywisty przeczy to twierdzeniu autora, że pewne funkcje teoretycznych są nieeliminowalne. Intuicja autora jest prawdopodobnie następująca. Chociaż w Ramseyowskim zdaniu A_R nie używa się terminu teoretycznego t, to stosuje się — w pewnym sensie — funkcję t: „mówi się o czymś podobnym" do niej, „odsyła się" do niej (aby zacytować pewne sformułowania autora). To prawda, że zdania, które wchodzą w skład zbioru X, muszą z reguły być zdaniami nieelementarnymi. Jednocześnie jednak muszą to być „zdania o wartościach funkcji nieteoretycznych", które mówią coś o tych funkcjach „bez wprowadzania żadnego dodatkowego aparatu pojęciowego", opisują je „w prosty, bezpośredni sposób". Muszę przyznać, że intuicja, która leży u podstaw tych wyjaśnień, pozostaje dla mnie niejasna. Pewne przykłady omawiane przez Sneeda zdają się sugerować, że niektóre ograniczenia nakładane na zdania z X są związane z rekursywnym charakterem odpowiednich relacji. Przedstawione w powyższy sposób pojęcie eliminowalności, zastosowane do teorii nieelementarnych, na pewno wymaga dalszych wyjaśnień.

Jak widzieliśmy, omawiane do tej pory podejście teoriomnogościowe, tzn. propozycje typu (1) i (2), mają proste odpowiedniki teoriomodelowe. Czy takie odpowiedniki mają również pozostałe propozycje? Obie — tzn. propozycje typu (3) i (4) — odwołują się w sposób istotny do pojęcia ograniczenia. Jest to pojęcie, które nie było dotąd rozważane w tradycyjnym ujęciu, i które sprawia pewne trudności, gdy próbuje się teoriomodelowo zrekonstruować powyższe propozycje. Aby to zrobić, musimy zdefiniować klasę modeli zamierzonych M^*, które wyznaczają interpretację języka L. Jeśli jednak nałożymy na M^*pewne ograniczenia C, nie wyznaczymy jej jednoznacznie: to, co zdefiniujemy, nie będzie klasą modeli zamierzonych dla L, tylko raczej klasą takich klas. Jakie znaczenie ma stwierdzenie, że teoria T sformułowana w takim języku L jest prawdziwa? Pojęcie prawdy dla języka tak zinterpretowanego wymaga przedefiniowania.

Nie jestem jednak w pełni przekonany, że nie możemy sobie poradzić bez pojęcia ograniczenia. Co więcej, pewne zastosowania tego pojęcia, sugerowane przez Sneeda, wydają mi się trudne do przyjęcia. Jest to problem zbyt skomplikowany, aby omawiać go w tym miejscu, więc ograniczę się jedynie do kilku uwag. Niech symbol funkcyjny t będzie terminem teoretycznym teorii T, a funkcje t_i i t_j będą jego dwiema interpretacjami w dwóch modelach zamierzonych $\mathfrak{M}_i^*$ i $\mathfrak{M}_j^*$, należących do klasy M^*. Ponieważ zarówno o t_i, jak i t_j, zakłada się, że spełniają ten sam zbiór aksjomatów, w którym wszystkim nieteoretycznym terminom zapewnia się ich zamierzoną interpretację, nie są one bynajmniej od siebie niezależne. Funkcje t_i i t_j, chociaż związane określonymi relacjami, nie muszą być identyczne. Co więcej, nie muszą być identyczne nawet wtedy, gdy dziedziny modeli $\mathfrak{M}_i^*$ i $\mathfrak{M}_j^*$ składają się z tych samych indywiduów i funkcje te są miarą tych samych «istotnych» cech tych indywiduów. Może to mieć miejsce z powodu wieloznaczności, z jaką jest określona interpretacja terminu t. Aksjomaty, które wiążą go z terminami nieteoretycznymi, mogą określać — i zwykle określają — jego interpretację dla zamierzonej interpretacji terminów nieteoretycznych w sposób niejednoznaczny. To, co faktycznie określają, jest klasą interpretacji dla t. Każdej takiej interpretacji jest przyporządkowany jeden model zamierzony z klasy M^*. Tak więc wartość t dla każdego indywiduum x może być różna w różnych modelach z klasy M^* ($t_i(x)$ nie musi równać się $t_j(x)$) — nie dlatego, że t nie jest miarą «istotnych» własności, lecz dlatego, że jest wyposażone w niejednoznaczną interpretację. Nakładanie na wszystkie funkcje teoretyczne przyporządkowane «istotnym» własnościom ograniczenia typu $\langle =, = \rangle$ nie pozwala zdać sprawy z nieostrości terminów teoretycznych — cechy, która wydaje mi się cechą nieodłączną wszystkich wyrażeń empirycznych. Wniosek ten nie stosuje się do innych rodzajów ograniczeń, mniej ścisłych niż $\langle =, = \rangle$,

Istnieje jednak punkt widzenia, zgodnie z którym samo pojęcie ograniczenia wydaje się zbędne. Zgodnie z nim, zakłada się, że istnieje tylko jedna dziedzina przedmiotów, do której odnosi się teoria fizyczna — dziedzina maksymalna, równa sumie wszystkich dziedzin zakładanych w poprzednich rozważaniach. Sneed wspomina o tym stanowisku, ale uważa że jest ono trudne do utrzymania[3]. Nie podzielam tego sceptycyzmu. Uważam, że logiczna rekonstrukcja teorii fizycznej opierająca się na takim założeniu jest w zasadzie realizowalna. Jest ona prawdopodobnie bardziej skomplikowana pod pewnymi technicznymi względami; wydaje mi się jednak bardziej zadowalająca z filozoficznego punktu widzenia. Problem ten jest zbyt fundamentalny i zbyt trudny, by można go było rozważać w tym artykule. Chciałbym tylko dodać jedną uwagę w tej sprawie. W obrębie naszego teoriomodelowego podejścia do logicznej struktury teorii fizycznej, rozważany teraz pogląd jest równoważny założeniu, że wszystkie modele zamierzone teorii T, tzn. wszystkie modele należące do klasy M^*, mają tę samą dziedzinę indywiduów. Nie znaczy to jednak, że istnieje tylko jeden taki model. Z powodu występującej zawsze nieostrości typowych terminów teoretycznych, M^* będzie z zasady zawierać wielość modeli odpowiadających różnym interpretacjom tych terminów. Jeżeli jednak wszystkie modele mają identyczną dziedzinę, nie ma potrzeby nakładać na funkcje teoretyczne jakichkolwiek ograniczeń. Wszystkie konieczne związki między ich wartościami będą ustalone przez aksjomaty teorii. W konsekwencji, gdyby ten punkt widzenia był możliwy do obrony, propozycje typu (3) i (4) okazałyby się zbyteczne, a treść empiryczna każdej teorii fizycznej mogłaby być adekwatnie przedstawiona za pomocą propozycji (2).

Propozycja ta, przedstawiona w monografii w języku teorii mnogości, może być, jak to pokazano wyżej, przełożona na język teorii modeli. Co zyskujemy w ten sposób? Chciałbym tutaj wymienić tylko dwa argumenty, które moim zdaniem, przemawiają na

3) Podobny punkt widzenia został niezależnie przedstawiony przez R. Wójcickiego. W swym artykule, wymienionym w przypisie 1, dowodzi on nieuniknionej wielości zastosowań każdej teorii fizycznej.

rzecz wyższości podejścia teoriomodelowego. Wydaje się ono wygodniejsze niż teoriomnogościowe, i zarazem właściwsze z punktu widzenia pewnego stanowiska filozoficznego. Jest ono wygodniejsze przez to, że pozwala mówić w sposób prosty o wyrażeniach teorii fizycznej — o jej terminach i aksjomatach. Przy teoriomnogościowym podejściu można to zrobić tylko w sposób okrężny. Jeżeli aksjomatyzujemy teorię T za pomocą predykatów teoriomnogościowych, nie używamy żadnych specyficznych terminów teorii T. Klasa M_T modeli dla T jest zdefinowana jedynie za pomocą ogólnych terminów teoriomnogościowych. Weźmy jako przykład teorię T. Przy teoriomnogościowym podejściu do tej teorii nie istnieje coś takiego, jak termin t. Co więcej, nie istnieje coś takiego, jak pewna określona funkcja f, chociaż niektóre sformułowania mogłyby to sugerować. Mówiąc ściśle, wszystkie zdania o funkcji t z teorii S powinny być zrekonstruowane jako pewne zdania warunkowe o postaci:

jeśli $\langle D, n, t\rangle$ *jest* S (*albo* S_0), *to* t *jest* ...,

co jest dość niewygodnym sposobem mówienia.

Z filozoficznego punktu widzenia podejście teoriomnogościowe można ocenić jako instrumentalistyczne raczej niż realistyczne. Oczywiście nie jest to instrumentalizm w skrajnej postaci. O teorii fizycznej T zakłada się, że ma ona pewną treść empiryczną, która jest wyrażalna przez — prawdziwe lub fałszywe — zdanie:

(2b) $M_0^* \subseteq M_T|_0$,

Ale jest to jedyne twierdzenie formułowane przez teorię T. Zauważmy, że nie zawiera ono żadnego terminu teoretycznego teorii T. Jedynymi specyficznymi terminami teorii T pojawiającymi się w (2b) są terminy denotujące nieteoretyczne funkcje, które tworzą zamierzone zastosowania należące do M_0^*. Tak więc tylko nieteoretyczne terminy z T mogą być tutaj traktowane jako wyrażenia zinterpretowane, i, w konsekwencji, tylko nieteoretyczne zdania z T mogą być uznane za twierdzenia prawdziwe lub fałszywe. Przy podejściu teoriomodelowym, wszystkim terminom teorii T, i wszystkim jej zdaniom, jest przyporządkowana pewna interpretacja. Interpretacja ta jest określona przez M^* — klasę modeli zamierzonych dla T. To prawda, że w ten sposób terminy teoretyczne nie zawsze uzyskują jednoznaczną interpretację: z zasady są one interpretowane wieloznacznie. Nie wyklucza to jednak, że przynajmniej niektóre zdania, zawierające terminy teoretyczne, mają wartość logiczną: są prawdziwe lub fałszywe. Tak więc ten sposób podejścia wydaje się właściwszy z realistycznego punktu widzenia. Nie chciałbym jednak przeceniać tego argumentu. Ponieważ wydaje się, że oba podejścia — teoriomnogościowe i teoriomodelowe — są wzajemnie «przetłumaczalne», różnica między nimi ma w dużym stopniu charakter werbalny.

Podsumowując te uwagi dotyczące pewnych aspektów monografii Sneeda, chciałbym jeszcze raz podkreślić jej ogromne zalety. O ile wiem, jest to jedyna praca na ten temat, która skutecznie radzi sobie z realizacją podstawowych celów każdych rozważań metodologicznych: analizą ważnych problemów metodologicznych dotyczących rzeczywistych teorii naukowych. Monografia Sneeda zajmuje się rzeczywistymi teoriami fizycznymi, a nie jakimiś uproszczonymi fikcyjnymi bytami. Główny problem zaś, któremu jest poświęcona — problem treści empirycznej teorii fizycznej — należy według mnie do najważniejszych problemów w metodologii teorii fizycznych. W pełni zgadzam się z *credo* autora. „Ostatecznym wytworem teorii fizycznych są twierdzenia empiryczne. Nie da się w pełni zrozumieć, czym zajmuje się fizyka matematyczna, jeśli się nie zobaczy, jak dochodzi ona do swoich twierdzeń empirycznych. Jest to cel, ku któremu zmierza jej logiczna rekonstrukcja." Monografia Sneeda wnosi wybitny wkład w realizację tego celu.

Tłumaczyła Anna Lissowska

ANALITYCZNOŚĆ

Filozofia Nauki
Rok I, 1993, Nr 2-3

O pojęciu zdania analitycznego*

I

1. Pojęcie zdania analitycznego należy do zasobu pojęć niezbędnych do uprawiania logicznej teorii nauki. Nie będą uzasadniał tutaj tego przekonania. Badania należące do tej dyscypliny świadczą, moim zdaniem, wymownie o roli, jaką w nich odgrywa odróżnienie zdań analitycznych od zdań syntetycznych. Ale nie wszyscy dzielą to przekonanie. W ostatnich zwłaszcza latach wzmogła się wyraźnie krytyka takiego stanowiska. Krytyka ta opiera się na argumentach kwestionujących w ogóle możliwość przeprowadzenia granicy pomiędzy tymi dwoma rodzajami twierdzeń. Samo pojęcie zdania analitycznego uznane zostało za wysoce zagadkowe. Niewątpliwie krytyka ta zawiera szereg trafnych uwag, które przyczyniły się do wyjaśnienia pewnych zagadnień związanych z pojęciem zdania analitycznego. Ale jej ostateczne konkluzje wydają się zbyt pochopne i nieprzekonujące. Trudności, na jakie natrafiają próby precyzacji pojęcia zdania analitycznego, nie są chyba większe od tych, z którymi mamy do czynienia w przypadku wielu innych pojęć logicznej teorii nauki, takich jak pojęcie definicji, terminu spostrzeżeniowego, teoretycznego itp., i nie stanowią dostatecznej podstawy do rezygnacji z tego pojęcia. Wskazują one jednak na konieczność pewnych ograniczeń, jakim zadanie takie musi zostać poddane.

Większość owych argumentów krytycznych dotyczy pojęcia zdania analitycznego na gruncie języka potocznego. Pojęcie to ma istotnie charakter wysoce nieokreślony, będący rezultatem nieokreśloności samego języka potocznego. Próby precyzacji tego

* Rezultaty przedstawione w pracy niniejszej pokrywają się częściowo z wynikami osiągniętymi przez R. Wójcickiego w pracy: „Analityczne komponenty definicji arbitralnych", *Studia Logica*, t. XIV. Dotyczy to głównie rezultatów zawartych w części pierwszej. Chciałbym podkreślić, iż wyniki R. Wójcickiego uzyskane zostały całkowicie niezależnie od rezultatów przedstawionych w pracy obecnej.

pojęcia bez uprzedniej precyzacji języka potocznego wydają się beznadziejne. Inaczej jednak wygląda sytuacja na gruncie tzw. języków sztucznych. Języki takie traktować możemy jako rezultat precyzacji języka potocznego, lub raczej pewnych jego części. Precyzacja ta polega na wyraźnym sformułowaniu reguł językowych, ustalających zasób wyrażeń danego języka (reguły składniowe) oraz ich znaczenia czy denotacje (reguły semantyczne). Kodyfikacja reguł językowych umożliwia sformułowanie ogólnikowych intuicji wyrażanych w określeniu zdania analitycznego jako zdania prawdziwego na mocy znaczenia, w sposób bardziej uchwytny: zdanie analityczne języka *J* — to zdanie prawdziwe na mocy reguł języka *J*. To określenie stało się dla Carnapa punktem wyjścia dla konstrukcji szeregu znanych definicji pojęcia zdania «logicznie prawdziwego», utożsamianego przez niego początkowo z pojęciem zdania analitycznego. Definicje te jednak uważać można za częściowe tylko rozwiązanie naszego zagadnienia. Zdania logicznie prawdziwe nie wyczerpują ogółu zdań analitycznych. Z dwóch klasycznych przykładów zdań analitycznych:

(1) Żaden człowiek nieżonaty nie jest żonaty

(2) Żaden kawaler nie jest żonaty

tylko zdanie (1) należy do zdań logicznie prawdziwych. Zdania logicznie prawdziwe — to, w innej terminologii, tautologie logiczne (prawa logiki i ich substytuty). Problem ich definicji uważać można w zasadzie za rozwiązany. Gorzej natomiast przedstawia się sprawa definicji tych zdań analitycznych, które nie mają charakteru tautologii. W swych pracach późniejszych Carnap próbował rozwiązać i to zadanie[1]. Próby te nie wyszły jednak na ogół poza szkicowe i fragmentaryczne rozważania. W podobnym w zasadzie kierunku idą usiłowania innych autorów[2]. I oni, podobnie jak Carnap, nawiązują do pojęcia postulatu znaczeniowego, stanowiącego uogólnienie pojęcia definicji projektującej. Zdanie (2) uważa się bowiem na ogół za zdanie analityczne dlatego, że zdanie to stanowi konsekwencję logiczną definicji występującego w nim terminu „kawaler". Nie wszelkie jednak terminy wprowadzane zostają do języka za pomocą definicji. Rolę definicji pełnią niekiedy postulaty znaczeniowe nie spełniające stawianego definicjom warunku przekładalności. Ich konsekwencje logiczne mogą mieć również charakter zdań analitycznych.

Opartą na takich właśnie sugestiach definicję zdania analitycznego podaje Ajdukiewicz. Propozycja ta stanowi najdojrzalszą próbę sformułowania omawianej koncepcji. Pojęcie postulatu i zdania analitycznego (w sensie semantycznym) określa Ajdukiewicz w sposób następujący: „Zdanie *Z* zawierające termin λ jest postulatem języka *J*, jeżeli w języku *J* obowiązuje semantyczna konwencja terminologiczna, ustanawiająca, że termin λ ma denotować taki przedmiot, który spełnia na miejscu terminu λ zdanie Z... W

1) Por. artykuły: „Meaning Postulates", *Philosophical Studies* 3 (1952); „Beobachtungssprache und theoretische Sprache", *Dialectica* 12 (1958).

2) Por. K. Ajdukiewicz: „Le problème du fondement des propositions analytiques", *Studia Logica*, t. VIII (1958); M. Kokoszyńska: „Deduction as a Method of Proof", Atti del XII Congresso Internazionale di Filosofia, Vol. 5, 1960.

sensie semantycznym zdaniem analitycznym języka *J* jest każdy postulat tego języka i konsekwencje logiczne postulatów". Spróbujmy określenie to poddać pewnej interpretacji, powołując się na propozycje zawarte w pracach Carnapa i Kokoszyńskiej. Język *J* utożsamić możemy z systemem semantycznym w rozumieniu Carnapa. Pod względem strukturalnym język ten traktować będziemy — o ile wyraźnie nie założymy czegoś przeciwnego — jako język prostej teorii typów (lub pewnego jej fragmentu), wzbogacony o proste terminy deskryptywne. Przyjmiemy, iż metajęzyk *M* zawiera język *J* jako swoją część oraz założymy, iż jest to pod względem logicznym język dostatecznie bogaty na to, aby zawierać zmienne związane tych wszystkich typów logicznych, do których należą terminy deskryptywne języka *J*. Sens terminów logicznych języka *J* uważać będziemy za ustalony. Sens terminów deskryptywnych języka *J* wyznaczają sformułowane w metajęzyku *M* reguły semantyczne ustalające ich denotacje[3]. Semantyczną konwencję terminologiczną dla terminu „λ" utożsamić możemy obecnie z regułą denotowania dla terminu „λ". Załóżmy, iż reguła ta ma postać następującą:

(3) $$\bigwedge_x (\text{„}\lambda\text{" denotuje } x \rightarrow \Phi(x)),$$

orzekającą, iż termin „λ" denotuje tylko taki przedmiot, który spełnia sformułowany w języku *J* warunek Φ[4]. Przy takich założeniach zdanie:

(4) $$\Phi(\lambda),$$

powstające z wyrażenia „Φ (*x*)" przez podstawienie na miejsce zmiennej *x* terminu „λ" stanowi, zgodnie z definicją Ajdukiewicza, postulat języka *J*, a ogół konsekwencji zdania (4) — zbiór zdań analitycznych języka *J*.

Postulaty znaczeniowe i ich konsekwencje nazywa się niekiedy twierdzeniami definicyjnymi. Istotnie, gdy „Φ (*x*)" jest funkcją jednostkową, np. w przypadku najprostszym:

(5) $$x = \lambda_1,$$

postulat „Φ (λ)" stanowi definicję terminu „λ" w ścisłym tego słowa znaczeniu. Pojęcie postulatu uważać więc można za naturalne uogólnienie pojęcia definicji. Posługując się tą terminologią koncepcję Ajdukiewicza wyrazić możemy jako propozycję utożsamiania zdań analitycznych z ogółem twierdzeń definicyjnych. Koncepcja ta odznacza się dużą prostotą i intuicyjnością. Niewątpliwie chwyta ona jedno ze znaczeń, jakie wiążemy z pojęciem zdania analitycznego. Jednocześnie jednak koncepcja ta pociąga konsekwencje niezgodne z innymi intuicjami dotyczącymi zdań analitycznych. Konsekwencje te wydają się świadczyć o tym, iż jest to koncepcja zbyt szeroka, zaliczająca do zdań analitycznych pewne twierdzenia, których byśmy tam skądinąd zaliczyć nie chcieli.

---

3) Przyjmiemy dla uproszczenia, iż dla każdego terminu deskryptywnego istnieje dokładnie jedna reguła denotowania. Nie wyklucza to oczywiście istnienia reguł ustalających denotacje nie dla pojedynczych terminów, lecz dla szeregu terminów jednocześnie. W dalszych wywodach posługiwać się będziemy regułami najprostszymi, dotyczącymi jednego terminu deskryptywnego.

4) Taką postać reguł denotowania sugeruje Kokoszyńska.

2. Rozpatrzmy sytuację następującą. Język J zawiera jako jedyne terminy deskryptywne terminy „λ_1", ..., „λ_n". Reguły denotowania tych terminów są takie, iż zdanie:

(6) $$\Phi_1 (\lambda_1, ..., \lambda_n)$$

należy na ich gruncie do zdań syntetycznych języka J. Może nim być np. dowolne przyrodzone prawo. Wzbogacamy język J o nowy termin „λ" dołączając następującą regułę denotowania:

(7) $$\bigwedge_x (\text{„}\lambda\text{" denotuje } x \rightarrow \Phi_1 (\lambda_1, ..., \lambda_n) \wedge \Phi_2 (x)) .$$

Postulatem staje się więc obecnie zdanie:

(8) $$\Phi_1 (\lambda_1, ..., \lambda_n) \wedge \Phi_2 (\lambda) ,$$

a zdanie (6) — jako konsekwencja tego postulatu — zdaniem analitycznym. Wprowadzenie zatem do języka J nowego terminu sprawia, iż pewne zdanie, w którym ten termin w ogóle nie występuje, zmienia swój charakter: z twierdzenia syntetycznego, rzeczowego staje się twierdzeniem analitycznym, definicyjnym. Wydaje się, iż tego rodzaju zmiana możliwa jest tylko przy zmianie znaczenia danego zdania. Wprowadzenie nowego terminu musiałoby więc pociągnąć zmianę znaczenia terminów dotychczasowych. Z drugiej jednak strony, nowa reguła denotowania ustala tylko denotację terminu wprowadzanego, nie nakładając żadnych warunków na terminy pozostałe. Nie zmienia więc chyba ich znaczenia. Tak mogłoby być tylko wtedy, gdyby — wbrew naszemu założeniu — reguła denotowania dla terminu „λ" dotyczyła również terminów pozostałych, głosząc np., iż:

(9) $$\bigwedge_{x_1} ... \bigwedge_{x_n} \bigwedge_y (\text{„}\lambda_1\text{", ..., „}\lambda_n\text{", „}\lambda\text{" denotują } x_1, ..., x_n, y \rightarrow \Phi_1 (x_1, ..., x_n) \wedge \Phi_2 (y)).$$

Powyższe konsekwencje występują oczywiście nie tylko w przypadku postulatów postaci (8), które z pewnością nie reprezentują typowej postaci faktycznie spotykanych postulatów. Nie jest rzeczą konieczną, aby dany postulat zawierał jako czynnik koniunkcji jakieś zdanie syntetyczne języka dotychczasowego. Wystarczy, jeśli takie zdanie z danego postulatu logicznie wynika. A bywa tak w przypadku postulatów znajdujących niewątpliwie zastosowanie przy budowie teorii naukowych. Należą do nich przede wszystkim postulaty o postaci tzw. definicji cząstkowej. Na tę ich własność wielokrotnie zwracano uwagę[5]. Weźmy pod uwagę najprostszą postać definicji cząstkowej terminu „Q":

(10) $$\bigwedge_x ((P (x) \rightarrow Q (x)) \wedge (R (x) \rightarrow \sim Q (x))),$$

gdzie „P" i „R" są terminami o ustalonych uprzednio znaczeniach. Wypowiedź (10) jest postulatem dla terminu „Q", gdy reguła denotowania dla tego terminu brzmi:

(11) $$\bigwedge_X (\text{„}Q\text{" denotuje } X \rightarrow \bigwedge_x ((P (x) \rightarrow X (x)) \wedge (R (x) \rightarrow \sim X (x)))).$$

W tej sytuacji zarówno postulat (10), jak i wszelkie jego konsekwencje logiczne, zaliczyć musimy, zgodnie z obecną koncepcją, do zdań analitycznych. Ale wśród owych konsekwencji znajduje się twierdzenie:

---

5) Czyniłem to m.in. w pracy: „Pojęcia teoretyczne a doświadczenie", *Studia Logica*, t. XI (1961).

(12) $$\bigwedge_x (P(x) \rightarrow \sim R(x)),$$

które terminu definiowanego w ogóle nie zawiera i które w języku dotychczasowym ma z reguły charakter syntetyczny. W każdym razie w tych sytuacjach, w których definicje cząstkowe najczęściej się stosuje, postulaty znaczeniowe dla terminów „*P*" i „*R*" zależności (12) nie implikują. Trudno przy tym tutaj mówić, aby owemu «uanalitycznieniu» twierdzenia (12) towarzyszyła zmiana znaczenia terminów „*P*" i „*R*". W sytuacjach, o których mowa, idzie właśnie o to, aby tamte terminy swoje dotychczasowe znaczenie zachowały, tak iż wypowiedź (11) wydaje się właściwym sformułowaniem reguły denotowania dla terminu „*Q*".

Podobne konsekwencje pojawiają się również w przypadku definicji zupełnych. Ajdukiewicz sam dostarcza tutaj interesującego przykładu. Definicja „grama" głosząca, iż:

(13) gram jest to masa 1 cm^3 wody o temperaturze 4^oC

implikuje twierdzenie głoszące, iż:

(14) wszystkie cm^3 wody o temperaturze 4^oC mają tę samą masę.

Twierdzenie (14) traktuje się w fizyce jako zdanie wyrażające pewną doświadczalną zależność. Ma ono więc w języku fizyki nie zawierającym definicji grama charakter zdania syntetycznego. Skoro jednak definicję (13) uznajemy za zdanie analityczne, musimy w języku fizyki wzbogaconym o tę definicję uznać za zdanie analityczne również twierdzenie (14). Czy zmieniło ono obecnie swe znaczenie? Czy też, nie zmieniając znaczenia, zmieniło tylko swój logiczny charakter? Żadna z tych ewentualności nie wydaje się dostatecznie przekonująca.

Ajdukiewicz skłania się, jak można przypuszczać, do stanowiska, w myśl którego «uanalitycznienie» danego zdania nie pociąga zmiany jego znaczenia — w każdym razie nie pociąga takiej zmiany znaczenia, która by pozbawiała to zdanie charakteru doświadczalnego. Przytaczane przez nas twierdzenia: (6), (12), czy (14) pozostają — jako zdania analityczne języków wzbogaconych o terminy: „λ", „*Q*" czy „gram" — twierdzeniami doświadczalnymi, podległymi kontroli doświadczenia i narażonymi na obalenie przez dane doświadczalne. Fakt ten służy Ajdukiewiczowi do poparcia tezy głoszonej, iż zdania analityczne wymagają dla swego uzasadnienia odwołania się do doświadczenia. Istotnie, zaliczenie do zdań analitycznych twierdzeń typu: (6), (12) czy (14) taką konkluzję narzuca. W konsekwencji i postulaty typu: (8), (10), czy (13) stają się zależne od doświadczenia. Postulaty te bowiem są prawdziwe tylko pod tym warunkiem, iż prawdziwe są ich doświadczalne konsekwencje: (6), (12), czy (14). Gdyby któraś z nich okazała się na podstawie doświadczenia fałszem, nie moglibyśmy uznać za prawdę i odpowiedniego postulatu. Same konsekwencje terminologiczne nie wystarczają dla zagwarantowania im prawdziwości. Mówiąc ogólnie, reguła denotowania postaci (3) gwarantuje prawdziwość postulatu (4) tylko wtedy, gdy istnieje przedmiot, który spełnia warunek Φ. Gdy żaden przedmiot tego warunku nie spełnia, nie może spełniać go i przedmiot λ. W tej sytuacji postulat (4) nie może być uznany za zdanie prawdziwe. Otóż gdy postulat (4) pociąga konsekwencje, które okazują się w świetle doświadczenia fałszem, żaden przedmiot nie spełnia warunku Φ. A więc i postulat (4) nie może być uznany za prawdę.

Ta konsekwencja koncepcji Ajdukiewicza ma, podobnie jak poprzednia, posmak pewnego paradoksu. Raził nasze intuicje fakt, iż wprowadzenie nowego terminu przekształca w zdania analityczne twierdzenia, w których ów termin w ogóle nie figuruje. Podobnie wydaje się sprzeczny z naszymi intuicjami związanymi z pojęciem zdania analitycznego fakt, iż prawdziwość zdań analitycznych może być zależna od doświadczenia, iż charakter analityczny danego zdania może iść w parze z jego charakterem doświadczalnym. Wszystko to wydaje się świadczyć o tym, iż nasze intuicje dotyczące zdań analitycznych nie są ze sobą zgodne. Z jednej strony istnieje tendencja do utożsamiania zdań analitycznych z twierdzeniami definicyjnymi. Z drugiej strony istnieją intuicje wiążące pojęcie zdania analitycznego z pojęciem twierdzenia prawdziwego na mocy znaczenia, a więc niezależnie od doświadczenia. Wywody Ajdukiewicza okazują, iż koncepcje te nie pokrywają się ze sobą. Istnieją twierdzenia definicyjne, których prawdziwość uzależniona jest od doświadczenia. Definicja Ajdukiewicza realizuje koncepcję pierwszą. Koncepcja pozostała zasługuje jednak na realizację co najmniej w tym samym stopniu. Odpowiada ona temu pojmowaniu zdań analitycznych, które dominowało w dziejach logiki i filozofii. Ale nie tylko względy tradycji przemawiają za takim ujęciem. Podział ogółu twierdzeń na tak rozumiane zdania analityczne i syntetyczne jest podziałem doniosłym z punktu widzenia współczesnej metodologii. Wspominałem już o tym na wstępie. Toteż warto, jak mi się zdaje, podjąć próbę takiego określenia zdań analitycznych, które by respektowało wspomniane intuicje.

3. Określenie czyniące zadość powyższemu żądaniu musi być węższe od poprzedniego. Nie każde twierdzenie definicyjne uznamy za zdanie analityczne; nie każdą więc konsekwencję postulatu znaczeniowego zaliczymy do takich zdań. Musimy w jakiś sposób wyodrębnić klasę tych konsekwencji, które są zdaniami syntetycznymi; ta ostatnia może być zresztą w poszczególnych przypadkach klasą pustą.

Zadanie wyróżnienia analitycznych i syntetycznych konsekwencji postulatów znaczeniowych podejmowane było do tej pory przede wszystkim w stosunku do tych postulatów, które mają postać definicji cząstkowych. Rozwiązanie tego zadania, proponowane przeze mnie w innym miejscu[6], polegało na przedstawieniu definicji cząstkowej jako koniunkcji dwóch składników: składnika analitycznego i syntetycznego. Konsekwencje logiczne składnika analitycznego mają charakter zdań analitycznych; konsekwencjom logicznym składnika syntetycznego dana definicja analityczności nie gwarantuje. W przypadku definicji cząstkowej postaci (10), składnik syntetyczny jest tożsamy z twierdzeniem (12), a składnik analityczny — z twierdzeniem następującym:

(15) $$\bigwedge_x ((P(x) \wedge \sim R(x) \rightarrow Q(x)) \wedge (R(x) \wedge \sim P(x) \rightarrow \sim Q(x))).$$

6) „Pojęcia teoretyczne a doświadczenie", op. cit.

Adekwatność tej propozycji uzasadniana była tym, iż 1° twierdzenie (12) nie zawiera terminu definiowanego i pociąga logicznie te wszystkie zdania nie zawierające terminu definiowanego, które wynikają z definicji (10); 2° twierdzenie (15) nie pociąga logicznie żadnych zdań nie zawierających terminu definiowanego (różnych od tautologii); 3° koniunkcja twierdzenia (12) i (15) jest równoważna logicznie definicji (10).

Powstaje pytanie, jak uogólnić to rozwiązanie. W jaki sposób «rozbić» postulat o dowolnej postaci na takie składniki, które by spełniały sformułowane wyżej warunki? Odpowiedź na to pytanie zawarta jest w zasadzie we wspomnianej pracy Carnapa[7]. Spróbujmy przede wszystkim sformułować w postaci ogólnej składnik syntetyczny danego postulatu:

(4) $$\Phi(\lambda).$$

Ma to być zdanie nie zawierające terminu „λ" i aksjomatyzujące zbiór tych wszystkich konsekwencji logicznych postulatu (4), które nie zawierają terminu „λ". Czy zdanie takie zawsze daje się sformułować? Odpowiedź na to pytanie zależy od rodzaju języka, którego postulatem jest twierdzenie (4). Przy pewnych założeniach co do własności tego języka otrzymujemy odpowiedź twierdzącą. Bezpośrednie zastosowanie znajduje tu wynik uzyskany przez Ramsey'a[8]. Okazał on, iż zdaniem spełniającym wspomniany warunek jest wypowiedź:

(16) $$\bigvee_{x} \Phi(x),$$

powstała z postulatu (4) przez zastąpienie terminu „λ" zmienną odpowiedniego typu (różną od wszystkich zmiennych występujących w „Φ") i związanie jej kwantyfikatorem szczegółowym. Zdanie (16) implikuje dokładnie te same twierdzenia nie zawierające terminu „λ", co postulat (4). Możemy więc uważać je za składnik syntetyczny owego postulatu. Oczywiście przy założeniu, że zdanie (16) należy do wyrażeń tego języka, którego postulatem jest twierdzenie (4). Aby to mogło mieć miejsce, ów język *J* zawierać musi zmienne związane tego samego typu, co termin „λ". Warunek taki nie zawsze jest spełniony. Załóżmy, iż *J* jest językiem elementarnym, odpowiadającym pod względem logicznym węższemu rachunkowi funkcyjnemu, a „λ" — pewnym predykatem. Wówczas zdanie (16) nie należy do wyrażeń języka *J*. Nie może więc pełnić w nim roli składnika syntetycznego postulatu (4). Czy w takim języku *J* ów składnik w ogóle daje się sformułować? Na pytanie to wypadnie odpowiedzieć przecząco. Nie ma ogólnej metody sformułowania żądanego twierdzenia. Zadanie takie sprowadza się do aksjomatyzacji zbioru wszystkich konsekwencji logicznych postulatu (4) nie zawierających terminu „λ", przy pomocy skończonej liczby aksjomatów sformułowanych w języku *J* niewzbogaconym o termin „λ". Zadanie to jest niewykonalne. Badania Craiga[9]

7) „Beobachtungssprache und theoretische Sprache", op. cit.

8) *The Foundations of Mathematics and Other Essays*, 1931; propozycję Ramsey'a omawiałem w cytowanej wyżej pracy.

9) „On Axiomatizability within a System", *Journal of Symbolic Logic* 18 (1953); wspominałem o nich w cytowanej wyżej pracy.

dostarczają co prawda metody aksjomatyzacji owego zbioru, ale przy pomocy nieskończonej liczby aksjomatów. W pewnych przypadkach specjalnych daje się oczywiście sformułować twierdzenie spełniające żądane warunki; tak jest np. w przypadku postulatu o postaci definicji cząstkowej (10), gdzie twierdzenie (12) stanowi żądaną aksjomatykę dla omawianego zbioru. Ale w przypadku ogólnym twierdzenia takiego sformułować nie możemy. Musimy wobec tego zadowolić się wypowiedzią typu (16) jako ogólną formą składnika syntetycznego postulatu (4) i przyjąć, że na gruncie pewnych języków ów składnik syntetyczny wyrazić się nie daje. Przypuśćmy, iż J jest takim właśnie językiem. Przejście do języka J^*, który by pozwalał na sformułowanie wypowiedzi (16), polega na wzbogaceniu języka J o zmienne związane tego typu logicznego, który reprezentuje termin „λ", a więc wyłącznie o logiczne środki wyrazu. Jest to rozszerzenie zawsze wykonalne i — z pewnego punktu widzenia — nieistotne. Zgodnie z przyjętymi założeniami, warunek taki spełnia m.in. nasz metajęzyk M. Możemy zatem stwierdzić, iż wypowiedź (16) przedstawia ogólną postać składnika syntetycznego dowolnego postulatu (4) języka J — bądź w samym języku J, bądź w stanowiącym jego logiczne rozszerzenie języku J^*.

Podanie ogólnej postaci składnika analitycznego nie nastręcza obecnie większych trudności. Ma to być, jak wiemy, zdanie, które nie pociąga logicznie żadnych zdań nie zawierających terminu „λ" (różnych od tautologii), a w koniunkcji ze składnikiem syntetycznym (16) tworzy zdanie równoważne logicznie postulatowi (4). Idąc za sugestiami Carnapa, możemy przyjąć w tym charakterze wypowiedź:

$$(17) \qquad \bigvee_{x} \Phi(x) \rightarrow \Phi(\lambda) .$$

Wypowiedź ta spełnia, jak łatwo okazać, warunki stawiane przez nas składnikowi analitycznemu postulatu (4). Każda konsekwencja twierdzenia (17) zawiera termin „λ", a koniunkcja twierdzenia (16) i (17) jest równoznaczna postulatowi (4). Wypowiedź (17) reprezentować więc może składnik analityczny postulatu (4) w języku J, o ile daje się w tym języku sformułować. Tutaj powtórzyć można rozważania przeprowadzone w związku z wypowiedzią (16). Prowadzą one do wniosku, iż w pewnych przypadkach składnik analityczny postulatu języka J wyrazić się daje jedynie w języku J^* stanowiącym logiczne rozszerzenie języka J.

Wyodrębnienie składnika syntetycznego i analitycznego definicji cząstkowej (10) okazuje się obecnie pewnym uszczegółowieniem rozwiązania ogólnego. Składnik syntetyczny w postaci twierdzenia (12) jest, jak się łatwo przekonać, równoważny logicznie składnikowi typu (16):

$$(18) \qquad \bigvee_{X} \bigwedge_{x} ((P(x) \rightarrow X(x)) \wedge (R(x) \rightarrow \sim X(x))) ,$$

a składnik analityczny w postaci twierdzenia (15) — składnikowi typu (17):

$$(19) \qquad \bigvee_{X} \bigwedge_{x} ((P(x) \rightarrow X(x)) \wedge (R(x) \rightarrow \sim X(x))) \rightarrow$$
$$\rightarrow \bigwedge_{x} ((P(x) \rightarrow Q(x)) \wedge (R(x) \rightarrow \sim Q(x))) .$$

Potwierdza to adekwatność proponowanego rozwiązania.

W odróżnieniu od koncepcji Ajdukiewicza zaliczającej do zdań analitycznych ogół konsekwencji logicznych postulatu (4), koncepcja obecna przyznaje charakter analityczny tylko pewnym spośród nich: konsekwencjom logicznym składnika analitycznego (17). Ich analityczność wydaje się nie ulegać wątpliwości. Nie ma wśród nich żadnego twierdzenia, które by mogło wyrażać jakikolwiek fakt doświadczalny. Wszelkie takie twierdzenia należą wyłącznie do konsekwencji logicznych składnika syntetycznego (16). Wypowiedź (16) stwierdza, iż istnieje przedmiot, który spełnia określony warunek, sformułowany w terminach o ustalonych uprzednio znaczeniach. To może być istotnie stwierdzeniem pewnego doświadczalnego faktu. Natomiast wypowiedź (17) ogranicza się do stwierdzenia, iż jeśli taki przedmiot istnieje, to jest nim przedmiot λ, gdzie „λ" jest terminem nie posiadającym żadnego uprzednio ustalonego sensu. Nie może to być, co za tym idzie, twierdzenie doświadczalne.

To, że koncepcja obecna zalicza do zdań analitycznych tylko takie konsekwencje danego postulatu, których analityczność nie budzi wątpliwości, wydaje się w świetle tych rozważań faktem uzasadnionym. Czy zachodzi również zależność odwrotna? Czy wszelkie takie konsekwencje zaliczone zostają do zdań analitycznych? Czy wśród konsekwencji składnika syntetycznego nie ma takich zdań, które nie są konsekwencjami składnika analitycznego, a mimo to mają niewątpliwie charakter analityczny? Oczywiście, zdania takie mogą istnieć. Ale wówczas ich charakter analityczny zapewniony będzie przez inne postulaty znaczeniowe. Dany postulat ich analityczności nie gwarantuje. Zbiór konsekwencji logicznych składnika analitycznego danego postulatu obejmuje wszystkie zdania analityczne, które zawdzięczają swą analityczność owemu postulatowi. Chcąc określić ogół zdań analitycznych danego języka, musimy odwołać się do ogółu jego postulatów. Powiemy więc, iż zdanie analityczne języka *J* — to zdanie będące konsekwencją logiczną dowolnego zbioru składników analitycznych postulatów języka *J*. Określenie to wymaga jeszcze pewnej modyfikacji. Nie każdy składnik analityczny należeć musi, jak pamiętamy, do zdań języka *J*. Toteż jako zdania analityczne języka *J* traktować będziemy tylko te spośród wspomnianych konsekwencji, które są zdaniami języka *J*. Zgodnie z przyjętym określeniem, również postulat (4) należeć może do zdań analitycznych języka *J*, jeśli tylko jego składnik syntetyczny (16) okaże się na mocy innych postulatów języka *J* zdaniem analitycznym. Tę samą konsekwencję otrzymujemy oczywiście wtedy, gdy składnik syntetyczny (16) ma charakter tautologii logicznej. Wówczas składnik analityczny (17) jest po prostu logicznie równoważny postulatowi (4). Z taką sytuacją mamy do czynienia w przypadku pewnych typów definicji zupełnej.

Koncepcja obecna unika, jak widać, jednego z paradoksów koncepcji poprzedniej. Wprowadzenie do języka *J* terminu „λ" przez dołączenie reguły denotowania ustalającej wyłącznie denotację terminu „λ" oraz przez przyjęcie opartego na tej regule postulatu nie może przekształcić żadnego zdania syntetycznego języka *J* w zdanie analityczne języka *J* wzbogaconego o termin „λ", gdyż musiałoby to być zdanie, w którym termin

„λ" w ogóle nie figuruje, a w myśl obecnej koncepcji postulat dla terminu „λ" żadnemu zdaniu tego typu analityczności nie gwarantuje. Ten fakt wydaje się przemawiać na korzyść proponowanej definicji. Czy definicja ta unika również drugiego ze wspomnianych paradoksów? Odpowiedź na to pytanie nie jest — wbrew pozorom — łatwa. Chcąc jej udzielić, zmuszeni będziemy poruszyć pewne podstawowe problemy logicznej teorii języka.

II

1. Rozważania części pierwszej zmierzające do okazania analitycznego charakteru wyróżnionych przez nas zdań wydają się przemawiać za ich niezależnością od doświadczenia. W przeciwieństwie do postulatu:

(4) $$\Phi(\lambda),$$

składnik analityczny:

(17) $$\bigvee_x \Phi(x) \rightarrow \Phi(\lambda)$$

nie pociąga żadnych zdań, które by mogły mieć charakter doświadczalny. Jakżeż więc doświadczenie mogłoby okazać jego fałszywość? Sprawa jednak nie przedstawia się tak prosto. Skoro wypowiedź (4) jest postulatem dla terminu „λ", regułą denotowania dla tego terminu jest reguła:

(3) $$\bigwedge_x (\text{„}\lambda\text{" denotuje } x \rightarrow \Phi(x)).$$

Przypuśćmy, iż składnik syntetyczny:

(16) $$\bigvee_x \Phi(x)$$

jest zdaniem syntetycznym języka J, i że doświadczenie okazuje jego fałszywość:

(20) $$\sim \bigvee_x \Phi(x).$$

Wówczas reguła denotowania (3) prowadzi do wniosku, iż:

(21) $$\sim \bigvee_x (\text{„}\lambda\text{" denotuje } x).$$

W tej sytuacji zatem reguła denotowania dla terminu „λ" nie wyposaża tego terminu w żadną denotację[10].

Jakiż stąd możemy wysnuć wniosek co do wartości logicznej postulatu (4) w języku J wzbogaconym o termin „λ" — nazwijmy go językiem J'? Jedno jest niewątpliwe: postulatu tego nie możemy uznać za zdanie prawdziwe języka J'. Ale ta decyzja pozostawia jeszcze dwie możliwości. Możemy przyjąć 1° iż postulat (4) jest zdaniem fałszywym języka J', lub 2° iż postulat (4) nie jest w ogóle zdaniem języka J'. Rozwiązanie pierwsze uprzywilejowuje wyraźnie obecną koncepcję zdania analitycznego w porównaniu z koncepcją Ajdukiewicza. Wobec (20) postulat (4) okazuje się fałszem, natomiast jego składnik analityczny (17) pozostaje prawdą. Tym samym ów składnik i wszelkie jego konsekwencje — czyli, w myśl obecnej koncepcji, wszelkie zdania analityczne — pozostają niewrażliwe na kontrolę doświadczenia, prawdziwe na mocy

10) Podobnie byłoby oczywiście wtedy, gdyby składnik syntetyczny (16) był zdaniem kontradyktorycznym.

samej terminologicznej konwencji. Rozwiązanie drugie stawia pod omawianym względem obie koncepcje w tej samej sytuacji. Skoro, wobec (20), postulat (4) nie jest w ogóle zdaniem języka J', nie jest nim również jego składnik analityczny (17), który zawiera ów postulat jako swoją część. Nie będąc zaś zdaniami języka J', nie są też oczywiście zdaniami prawdziwymi. A zatem sama konwencja terminologiczna nie wystarcza do zagwarantowania im prawdziwości. Doświadczenie może zmusić do odrzucenia zarówno postulatu (4), jak i jego składnika analitycznego (17) — co prawda nie jako zdań fałszywych, lecz jako wyrażeń nie należących do danego języka. Ewentualność pierwsza jest bez wątpienia prostsza i z wielu względów atrakcyjniejsza. Czy daje się jednak konsekwentnie utrzymać? Pytanie to, jak zwykle w takich razach, wymaga niezbędnego uzupełnienia: w jakim języku? W zależności bowiem od rodzaju danego języka wypadnie wybrać pierwszą lub drugą ewentualność.

W naszych dotychczasowych rozważaniach język J traktowaliśmy pod względem logicznym jako język prostej teorii typów, lub pewnych jej fragmentów, idąc w tym za przykładem prac poświęconych problemowi zdań analitycznych (w tym podstawowych prac Carnapa i Ajdukiewicza). Otóż wydaje się, iż tego rodzaju język dopuszcza wyłącznie ewentualność drugą. Jest to widoczne zwłaszcza wtedy, gdy J jest dostatecznie bogaty w logiczne środki wyrazu, zawierając zmienne związane tych wszystkich typów logicznych, do których należą jego terminy deskryptywne. Czy może być terminem deskryptywnym takiego języka J termin „λ", który niczego nie denotuje? Czy może być, co za tym idzie, zdaniem języka J wyrażenie zawierające taki termin? W języku J wszelkie terminy deskryptywne mają charakter nazw, czyli wyrażeń stanowiących możliwe podstawienia zmiennych związanych. Innymi słowy, w języku J dla dowolnego terminu deskryptywnego „λ" mają walor takie reguły inferencji, jak reguła podstawiania:

(22) $$\bigwedge_x \Phi(x) \vdash \Phi(\lambda)$$

oraz reguła egzystencjalnej generalizacji:

(23) $$\Phi(\lambda) \vdash \bigvee_x \Phi(x).$$

W konsekwencji zdanie:

(24) $$\bigvee_x (x = \lambda)$$

jest tautologią logiczną języka J. Zdanie to stwierdza, iż istnieje przedmiot będący λ. Prawdą musi być więc również jego metajęzykowy odpowiednik:

(25) $$\bigvee_x (\text{„}\lambda\text{" denotuje } x),$$

stwierdzający, iż istnieje przedmiot denotowany przez termin „λ"[11]. A zatem, zaliczenie terminu „λ" w poczet terminów deskryptywnych języka J zmusza do uznania, iż termin „λ" jakiś przedmiot denotuje. Skoro więc stwierdzenie (20) prowadzi na gruncie reguły

11) Zależność głosząca, iż „λ" denotuje przedmiot p w języku J wtedy, gdy zdanie „$p = \lambda$" jest prawdziwe w języku J — wydaje się konsekwencją wszelkiej adekwatnej eksplikacji pojęcia denotowania.

denotowania dla terminu „λ" postaci (3) do sprzecznej z twierdzeniem (25) konkluzji (21) głoszącej, iż termin „λ" nie denotuje niczego, nie możemy terminu tego uważać w tej sytuacji za jeden z terminów deskryptywnych języka J.

Mówiąc dokładniej, terminami deskryptywnymi języka J mogą być wyłącznie terminy takie, które denotują jeden i tylko jeden przedmiot. Tautologią logiczną języka J jest bowiem również zdanie:

(26) $$\bigwedge_x \bigwedge_y (x = \lambda \wedge y = \lambda \rightarrow x = y),$$

głoszące, iż co najwyżej jeden przedmiot jest λ. Jego metajęzykowy odpowiednik:

(27) $$\bigwedge_x \bigwedge_y (\text{„}\lambda\text{" denotuje } x \wedge \text{„}\lambda\text{" denotuje } y \rightarrow x = y)$$

stwierdza, iż termin „λ" denotuje co najwyżej jeden przedmiot. Ale jeśli regułą denotowania dla terminu „λ" pozostaje reguła (3), żadne dane doświadczenia nie mogą stać w sprzeczności z twierdzeniem (27). Tak mogłoby być jedynie wtedy, gdyby reguła denotowania dla terminu „λ" nakładała nie tylko niezbędne, ale i wystarczające warunki na denotację tego terminu:

(28) $$\bigwedge_x (\text{„}\lambda\text{" denotuje } x \equiv \Phi(x)).$$

Wówczas ewentualne twierdzenie doświadczalne:

(29) $$\bigvee_x \bigvee_y (x \neq y \wedge \Phi(x) \wedge \Phi(y)),$$

orzekające, iż co najmniej dwa przedmioty spełniają warunek Φ, prowadziłoby do sprzeczności z twierdzeniem (27), pozbawiając z tej racji wyrażenie „λ" charakteru terminu deskryptywnego języka J.

Sytuacja zatem przedstawia się następująco. Termin „λ" tylko wtedy uznać możemy za termin deskryptywny języka J, gdy istnieje dokładnie jeden przedmiot stanowiący jego denotację:

(30) $$\dot{\bigvee_x} (\text{„}\lambda\text{" denotuje } x).$$

Jest to konsekwencją tego, iż zdanie:

(31) $$\dot{\bigvee_x} (x = \lambda)$$

jest tautologią języka J. Jeśli regułą denotowania dla terminu „λ" jest reguła (3), warunek (30) jest spełniony wtedy, gdy prawdą jest twierdzenie (16). Gdyby prawdą była jego negacja: (20), warunek (30) nie byłby spełniony. Nie moglibyśmy wówczas terminu „λ" uznać za termin deskryptywny języka J, a wyrażeń, w których ów termin figuruje — (4) czy (17) — za zdania tego języka.

Wprowadzenie do języka J terminu „λ" stanowi w istocie przejście do nowego języka J'. Przy założeniu prawdziwości twierdzenia (20) wprowadzenie takie okazuje się jednak, jak widzimy, zabiegiem nieskutecznym. Jakież stąd wnioski wysnuć można co do języka J'? Czy język taki w ogóle istnieje? Odpowiedź na to pytanie zależy od tego, jak pojmujemy ową procedurę wzbogacenia języka J o nowe terminy deskryptywne. Jeśli procedura ta wyraża się bezwarunkowym przyjęciem metajęzykowego twierdzenia głoszącego, iż „λ" należy do terminów deskryptywnych języka J', to tak scharakteryzowany język J' nie istnieje, gdyż jest tworem sprzecznym. Twierdzenie

(20) i reguła denotowania (3) pociągają bowiem — na gruncie języków o omawianej strukturze — negację twierdzenia powyższego. Skoro język J' nie istnieje, wyrażenia (4) czy (17) nie mogą być jego zdaniami — z powodów aż nadto oczywistych. Ale ową procedurę wzbogacenia języka J o nowe terminy deskryptywne można pojmować liberalniej. Jej wyrazem może być przyjęcie metajęzykowego twierdzenia:

(32) $\bigvee_{x} \Phi (x) \rightarrow$ „λ" należy do terminów deskryptywnych języka J',

zaliczającego „λ" do języka J' tylko pod warunkiem prawdziwości twierdzenia (16). Takie pojmowanie owej procedury wydaje się właściwsze w przypadku języków o omawianej strukturze i tak też pojmować je będziemy w dalszym ciągu rozważań. Otóż tak określony język J' okazuje się po prostu identyczny z wyjściowym językiem J. Twierdzenie (20) i reguła denotowania (3) pociągają bowiem, jak widzieliśmy, konsekwencję głoszącą, iż „λ" nie należy do terminów deskryptywnych języka J', niesprzeczną bynajmniej z twierdzeniem (32). Oczywiście i w tym przypadku wyrażenia (4) czy (17) nie należą do zdań języka J'.

Omawiana własność języków o takiej strukturze logicznej, jak język J, zasługuje na szczególną uwagę. To, czy wyrażenie „λ" należy do terminów deskryptywnych, inaczej nazw, języka J, czy też do nich nie należy — a więc jest w języku J, swobodnie mówiąc, nonsensem — nie jest sprawą czysto syntaktyczną, lecz zależy od tego, czy wyrażenie „λ" denotuje dokładnie jeden przedmiot. To zaś na gruncie pewnych metod denotowania dla terminu „λ" może być sprawą doświadczenia. Doświadczenie więc decydować może o tym, czy „λ" jest nazwą języka J, czy nonsensem; a zatem i o tym, czy jakieś wyrażenie zawierające termin „λ" jest zdaniem języka J, czy nonsensem. Jeśli termin „λ" okazuje się, jako termin pozbawiony denotacji, nonsensem, nonsensem jest każde zdanie, w którym ów termin figuruje. A więc zarówno zdania zaliczane przez nas do analitycznych: składnik analityczny postulatu (17), czy tautologia (24), jak i zdania syntetyczne, dajmy na to zdanie:

(33) $$\Psi (\lambda).$$

Czy te dwa typy zdań są więc w równym stopniu zależne od doświadczenia? Czy zdania analityczne niczym się pod tym względem nie różnią od zdań pozostałych? Otóż tak nie jest. Stopień zależności do doświadczenia jest w obu przypadkach inny. W przypadku zdań analitycznych doświadczenie rozstrzyga tylko o tym, czy dane wyrażenie, np. (17), jest zdaniem, czy nie. Może ono prowadzić do wniosku, iż wyrażenie to jest nonsensem. Ale gdy prowadzi do wniosku, iż wyrażenie (17) jest zdaniem, prowadzi tym samym do wniosku, iż jest zdaniem prawdziwym. Pewien fakt doświadczalny może być niezbędnym warunkiem sensowności takiego wyrażenia. Ale ów niezbędny warunek sensowności stanowi zarazem wystarczający warunek jego prawdziwości. Doświadczenie nie jest w stanie okazać fałszywości zdania analitycznego. Inaczej rzecz się ma ze zdaniami syntetycznymi. I tutaj doświadczenie może rozstrzygać o sensowności danego wyrażenia, np. (33). Ale w tym przypadku rola doświadczenia na tym się nie kończy. Wyrażenie (33) może być zdaniem, ale zdaniem fałszywym. Fakt doświadczalny, będący niezbędnym warunkiem sensowności wyrażenia (33), nie stanowi bynajmniej wystarczającego warunku jego prawdziwości. Inne fakty doświadczalne decydują o wartości logicznej takiego wyrażenia. Doświadczenie może więc okazać fałszywość zdania syntetycznego.

Rozważania te pokazują, iż proponowana przez nas definicja zdania analitycznego wyodrębnia mimo wszystko klasę zdań, które również pod względem swego stosunku do doświadczenia wyróżniają się spośród ogółu zdań. W pewnym sensie są od doświadczenia niezależne. Jeśli tylko uzyskują postulowane znaczenie, muszą być zdaniami prawdziwymi. Mając to na myśli można by utrzymywać, iż są zdaniami prawdziwymi na mocy samego znaczenia. Ta teza, przy bardziej rygorystycznym rozumieniu, utrzymać się jednak, jak widzieliśmy, nie daje. Nie można w szczególności twierdzić, iż wyróżnione przez nas zdania analityczne są zdaniami prawdziwymi na mocy samej terminologicznej konwencji. Nasza językowa umowa nie wystarcza do zagwarantowania im prawdziwości, skoro nie zawsze zapewnia im sensowność. Pozostaje wobec tego pytanie, czy nie jest rzeczą możliwą wyodrębnienie klasy takich zdań, których nie tylko prawdziwość, ale i sensowność nie podlegałyby kontroli doświadczenia, a więc zdań, które by były od doświadczenia niezależne w najbardziej rygorystycznym rozumieniu. Gdyby zaś to okazało się możliwe, czy nie było by rzeczą stosowną takim właśnie zdaniom przyznać miano zdań analitycznych, zachowując w ten sposób pewne intuicje, które z tym rodzajem zdań wiążemy? To nie jest zresztą sprawa istotna. Niezależnie od tego, jakim zdaniom nadamy miano zdań analitycznych, wyodrębnienie wspomnianej klasy zdań wydaje się zadaniem ważnym z punktu widzenia logicznej teorii nauki. Możemy je w odróżnieniu od zdań zdefiniowanych poprzednio nazwać zdaniami analitycznymi w sensie węższym. Podanie ich definicji nie jest zadaniem łatwym. Toteż zmuszony będę ograniczyć się do zarysowania w sposób daleki od precyzji pewnych prób rozwiązań, opatrzonych szeregiem znaków zapytania.

2. Chcąc otrzymać na gruncie języka *J* o strukturze logicznej prostej teorii typów klasę takich zdań, których nie tylko prawdziwość, ale i sensowność nie zależałaby od doświadczenia, musimy wyodrębnić ze zbioru zdań określonych przez nas jako analityczne pewien podzbiór, zawężając w ten sposób jeszcze bardziej wyjściowe pojęcie zdania analitycznego reprezentowane przez Ajdukiewicza. Załóżmy, jak poprzednio, iż zdanie:

(4) $$\Phi(\lambda)$$

jest postulatem języka *J*, a więc, iż w języku tym reguła denotowania dla terminu „λ" ma postać:

(3) $$\bigwedge_x (\text{„}\lambda\text{" denotuje } x \rightarrow \Phi(x))\,.$$

Ajdukiewicz do zdań analitycznych języka *J* zaliczał wszelkie konsekwencje logiczne postulatu (4); my proponowaliśmy zaliczenie jedynie konsekwencji logicznych jego składnika analitycznego:

(17) $$\bigvee_x \Phi(x) \rightarrow \Phi(\lambda)\,.$$

A jakie wypowiedzi zaliczymy do zdań analitycznych obecnie? Na pytanie to trudno odpowiedzieć w sposób równie prosty, jak poprzednio. Ustalmy więc stopniowo, jakie rodzaje wypowiedzi na pewno do klasy tak rozumianych zdań analitycznych się kwalifikują.

Punktem wyjścia będzie dla nas stwierdzenie, iż charakter taki przysługuje prawom logiki, tj. tautologiom logicznym nie zawierającym żadnych terminów deskryptywnych. Zakładamy, że ani ich sensowność, ani prawdziwość od doświadczenia zależeć nie może. Na tej podstawie do zdań analitycznych języka J zaliczymy przede wszystkim te postulaty języka J, których składnik syntetyczny jest prawem logiki. Tak więc, jeśli składnik syntetyczny postulatu (4):

(16) $$\bigvee_{x} \Phi (x)$$

należy do praw logiki, postulat (4) jest zdaniem analitycznym. Postulat ten jest wówczas równoważny logicznie swemu składnikowi analitycznemu (17). Jest rzeczą widoczną, iż jest on całkowicie od doświadczenia niezależny. Prawo logiki gwarantuje nam, iż istnieje przedmiot spełniający warunek Φ. W tej sytuacji reguła denotowania dla terminu „λ" (3) istotnie wyposaża ten termin w żądaną denotację i zapewnia tym samym sensowność i prawdziwość opartemu na niej postulatowi (4).

Postulaty, których składnik syntetyczny należy do praw logiki, charakteryzują terminy wprowadzane wyłącznie za pomocą terminów logicznych. Jeśli postulat (4) ma mieć taki charakter, zawierać może poza terminem „λ" jedynie zmienne i stałe logiczne. Nazwijmy postulaty takie dla skrótu postulatami aksjomatycznymi. Prócz nich istnieją jednak postulaty charakteryzujące terminy wprowadzane przy pomocy innych, uprzednio wprowadzonych, terminów deskryptywnych. Postulat (4), traktowany jako postulat tego typu, zawierać może np. poza terminem „λ" terminy deskryptywne: „λ_1", ..., „λ_n". Takie też postulaty mieliśmy głównie na uwadze w poprzednich rozważaniach. Nazwijmy je z kolei postulatami definicyjnymi. Otóż postulaty definicyjne należeć mogą również do zdań analitycznych języka J. Gdy postulat (4) jest postulatem definicyjnym, jego składnik syntetyczny (16) nie może być co prawda prawem logiki, skoro zawiera pewne terminy deskryptywne: „λ_1", ..., „λ_n". Ale składnik ten sam należeć może do zdań analitycznych języka J. Jeśli tak jest, postulat (4) jest zdaniem całkowicie od doświadczenia niezależnym. I on więc uznany być może za zdanie analityczne języka J. Tak więc do zdań analitycznych języka J zaliczymy i te postulaty języka J, których składnik syntetyczny jest zdaniem analitycznym języka J. Przyjąć możemy wreszcie, iż do zdań analitycznych języka J należą konsekwencje logiczne dowolnego zbioru zdań analitycznych języka J zawierające jako jedyne terminy deskryptywne terminy występujące w zdaniach tego zbioru. Jest rzeczą oczywistą, iż nie wzbogacimy w ten sposób klasy zdań analitycznych języka J o zdania, które by w jakikolwiek sposób zależeć mogły od doświadczenia.

Charakteryzując w sposób powyższy zdania analityczne nie popadamy, wbrew pozorom, w błędne koło. Sformułowane ściśle, nasze określenie przybierałoby postać definicji indukcyjnej. Jej warunek wyjściowy głosiłby, iż postulaty języka J, których składniki syntetyczne są prawami logiki, należą do zdań analitycznych języka J; natomiast warunki indukcyjne stwierdzałyby, iż 1° postulaty języka J, których składniki syntetyczne są zdaniami analitycznymi języka J, należą do zdań analitycznych języka J; 2° konsekwencje logiczne dowolnego zbioru zdań analitycznych języka J, zawierające jako jedyne terminy deskryptywne terminy występujące w zdaniach tego zbioru, należą do zdań analitycznych języka J. Określenie to wymaga jeszcze pewnej modyfikacji ze względu na podkreślany już przez nas fakt, iż składniki syntetyczne postulatów języka J same nie zawsze w tym języku dają się wyrazić. W takich przypadkach zamiast żądać, jak to czyni pierwszy warunek indukcyjny, aby do zdań analitycznych języka J należały owe składniki, musimy ograniczyć się do żądania, aby do zdań takich należały wszelkie — będące zdaniami języka J — konsekwencje logiczne danego postulatu, zawierające wyłącznie uprzednio wprowadzone terminy deskryptywne.

To skomplikowane nieco określenie prowadzi istotnie, jak się zdaje, do wyróżnienia w języku J żądanej klasy zdań: zdań analitycznych w sensie węższym. Rozpatrzmy niektóre jego konsekwencje. Otóż tak wyróżniona klasa zdań obejmuje spośród tautologii logicznych prawa logiki oraz te ich substytuty, które zawierają wyłącznie terminy deskryptywne wprowadzone przez postulaty będące zdaniami analitycznymi języka J. Wydaje się to zgodne z naszymi intencjami, skoro, jak się łatwo przekonać, są to wszystkie i tylko te tautologie, których sensowność, a więc i prawdziwość, nie zależy od wyników doświadczenia. A jakie twierdzenia, poza tautologiami, wchodziłyby w skład tak określonych zdań analitycznych? Przyjrzyjmy się paru prostym przykładom.

Weźmy pod uwagę postulat języka J o postaci definicji zupełnej. Niechaj będzie nią definicja równoważnościowa predykatu „Q" odwołująca się do predykatów „P" i „R", np. definicja:

(34) $$\bigwedge_x (Q(x) \equiv P(x) \wedge R(x))\,.$$

Definicja (34) jest postulatem języka J, gdy reguła denotowania dla „Q" ma postać:

(35) $$\bigwedge_X (\text{„}Q\text{" denotuje } X \rightarrow \bigwedge_x (X(x) \equiv P(x) \wedge R(x)))\,.$$

Składnikiem syntetycznym postulatu (34) jest zdanie:

(36) $$\bigvee_X \bigwedge_x (X(x) \equiv P(x) \wedge R(x))\,.$$

Zdanie to nie jest prawem logiki, bo zawiera terminy deskryptywne „P" i „R", ale jako tautologia logiczna zawierająca wyłącznie te terminy jest zdaniem analitycznym języka J, jeśli tylko postulaty dla tych terminów są zdaniami analitycznymi języka J. A tak jest np. wtedy, gdy są to postulaty aksjomatyczne języka J. Wówczas, zgodnie z naszym określeniem, postulat (34) należy także do zdań analitycznych języka J. Tak więc, przy pewnych założeniach, wszelkie definicje równoważnościowe predykatów języka J zaliczyć możemy do zdań analitycznych tego języka w sensie węższym.

Inaczej przedstawia się sytuacja przytaczanej przez nas poprzednio definicji cząstkowej (10). Jej składnik syntetyczny (18), równoważny logicznie twierdzeniu (12), z reguły nie ma charakteru zdania analitycznego języka J. Nie jest tautologią; nie jest też na ogół nietautologiczną konsekwencją postulatów wprowadzających terminy „P" i „R". Dlatego też, nawet przy założeniu, iż postulaty te są zdaniami analitycznymi języka J, np. postulatami aksjomatycznymi tego języka, postulat (10) nie staje się zdaniem analitycznym języka J. Nie jest nim również składnik analityczny (19), równoważny logicznie twierdzeniu (15). Żadne zdanie zawierające wprowadzony przez definicję cząstkową (10) termin „Q" nie uzyskuje w języku J charakteru zdania analitycznego w sensie węższym.

Ten stan rzeczy może ulec zmianie tylko wtedy, gdy postulat wprowadzający termin „Q" poddamy pewnej modyfikacji, modyfikując tym samym równocześnie regułę denotowania dla terminu „Q". Zamiast definicji cząstkowej (10) możemy przyjąć jako postulat jej składnik analityczny (19), np. w uproszczonej postaci (15). Reguła denotowania dla terminu „Q" przybiera wówczas postać:

(37) $$\bigwedge_X (\text{„}Q\text{" denotuje } X \rightarrow \bigwedge_x ((P(x) \wedge \sim R(x) \rightarrow X(x)) \wedge \wedge (R(x) \wedge \sim P(x) \rightarrow \sim X(x)))).$$

Składnik syntetyczny postulatu (15):

(38) $$\bigvee_X \bigwedge_x ((P(x) \wedge \sim R(x) \rightarrow X(x)) \wedge (R(x) \wedge \sim P(x) \rightarrow \sim X(x)))$$

jest tautologią zawierającą wyłącznie terminy „P" i „R". Toteż przy założeniu analityczności postulatów wprowadzających te terminy jest zdaniem analitycznym języka J. A więc i postulat (15) możemy zaliczyć do zdań analitycznych języka J.

Powyższy przykład ilustruje ogólną metodę zapewniania analityczności postulatom językowym. Metoda ta polega na zastępowaniu danego postulatu jego składnikiem analitycznym, co pociąga oczywiście odpowiednią zmianę reguły denotowania. Zamiast postulatu:

(4) $$\Phi(\lambda)$$

przyjmujemy postulat:

(17) $$\bigvee_x \Phi(x) \rightarrow \Phi(\lambda),$$

zastępując regułę denotowania:

(3) $$\bigwedge_x (\text{„}\lambda\text{" denotuje } x \rightarrow \Phi(x))$$

regułą:

(39) $$\bigwedge_x (\text{„}\lambda\text{" denotuje } x \rightarrow (\bigvee_x \Phi(x) \rightarrow \Phi(x))).$$

Składnik syntetyczny postulatu (17):

(40) $$\bigvee_x (\bigvee_x \Phi(x) \rightarrow \Phi(x))$$

jest tautologią. Jeśli tylko zawiera takie terminy deskryptywne, które wprowadzone zostały przez postulaty analityczne, sam jest zdaniem analitycznym. Jest więc nim również, w myśl naszego określenia, postulat (17) oraz jego logiczne konsekwencje. Gdybyśmy zatem wszystkie postulaty języka J zastąpili ich składnikami analitycznymi, otrzymalibyśmy język, w którym wszystkie postulaty i ich logiczne konsekwencje — czyli zdania analityczne w rozumieniu Ajdukiewicza — należą do zdań analitycznych w obecnym rygorystycznym rozumieniu.

Warto tutaj zwrócić uwagę na pewną szczególną własność tak wyróżnionych zdań analitycznych. Wydaje się mianowicie, iż wszelkie figurujące w nich terminy deskryptywne muszą być terminami pozbawionymi sensu empirycznego, jakkolwiek by się ten ostatni określiło. Postulaty wprowadzające te terminy odwoływać się bowiem muszą w ostatecznej instancji do terminów wprowadzonych przez postulaty aksjomatyczne. A postulaty aksjomatyczne, nakładające na owe terminy pierwotne warunki sformułowane wyłącznie za pomocą terminów logicznych, a więc warunki wyłącznie strukturalne, nie są w stanie wyposażyć je w jakikolwiek sens empiryczny[12]. A zatem i terminy pochodne wprowadzone przez postulaty definicyjne odwołujące się do tego rodzaju terminów pierwotnych pozbawione być muszą sensu empirycznego. Jest to niewątpliwie konsekwencja niepożądana, ograniczająca stosowalność pojęcia zdania analitycznego w metodologii nauk empirycznych.

Nasuwa się pytanie, czy owej klasy zdań analitycznych w sensie węższym, a więc zdań całkowicie niezależnych od doświadczenia, nie można zdefiniować w sposób nie budzący tego rodzaju zastrzeżeń. Odpowiedź na to pytanie uzależniliśmy na wstępie tych rozważań od rodzaju języka, którego ta definicja ma dotyczyć. Do tej pory poszukiwaliśmy jej dla języków o strukturze logicznej prostej teorii typów. Istotną dla naszego problemu własnością tych języków było to, iż dane wyrażenie uważane być mogło za termin deskryptywny, czyli nazwę (dowolnego typu) takiego języka tylko wtedy, gdy wyrażenie to denotowało dokładnie jeden przedmiot. W przeciwnym razie — do terminów deskryptywnych takiego języka nie należało, a zatem było w nim wyrażeniem bezsensownym. Czy nie ma jednak języków, które by pod tym względem różniły się od języków omawianych? Problem ten jest przedmiotem wielu rozważań i dyskusji, gdyż dotyka on zasadniczych spraw z dziedziny logicznej teorii języka. Idzie głównie o to, iż uzależnienie sensowności danego wyrażenia od tego, czy istnieje przedmiot będący jego denotacją, a więc w pewnych przypadkach — od wyników doświadczenia, uważa się za fakt niepożądany. Doświadczenie winno decydować jedynie o tym, czy dane zdanie jest prawdą, czy fałszem, a nie o tym, czy jest zdaniem, czy nonsensem. Kryteria sensowności wszelkich wyrażeń winny być czysto strukturalne[13]. Rozważania nad problemem języków o czysto strukturalnych kryteriach sensowności doprowadziły do interesujących rezultatów. Nawiązując do nich, rozważymy raz jeszcze zagadnienie zdań analitycznych. Tym razem jednak poprzestać musimy na niezmiernie szkicowych i dyskusyjnych uwagach i propozycjach.

3. Uzależnienie sensowności terminu „λ" od posiadania przez ten termin denotacji traktowaliśmy jako konsekwencję tego, iż w języku omawianym tautologią logiczną było zdanie:

(24) $$\bigvee_{x} (x = \lambda)\,.$$

12) Zwracałem na ten fakt uwagę w pracy: „Interpretacja systemów aksjomatycznych", *Studia Filozoficzne* 6 (1960).

13) Por. m.in. znane wywody Quine'a.

Zdanie to interpretowaliśmy jako zdanie stwierdzające istnienie przedmiotu będącego λ. Dlatego też jako jego metajęzykową konsekwencję traktowaliśmy zdanie:

(25) $$\bigvee_x („\lambda\text{” denotuje } x),$$

stwierdzające istnienie denotacji terminu „λ". Czy jednak taka interpretacja tautologii (24) stanowi jedyną możliwą interpretację tego twierdzenia? Jest to niewątpliwie interpretacja najczęstsza, ale nie jedyna. Jest ona wynikiem zwykłego, «przedmiotowego» rozumienia kwantyfikatora. Istnieje jednak i inne rozumienie kwantyfikatora, które takiej interpretacji nie pociąga. Mam na myśli «językowe» rozumienie Leśniewskiego[14], zgodnie z którym kwantyfikator szczegółowy nie stwierdza istnienia jakiegokolwiek przedmiotu. Tautologii (24) nie interpretujemy wówczas jako zdania głoszącego, że istnieje przedmiot będący λ. Toteż tautologia ta nie pociąga twierdzenia metajęzykowego (25) głoszącego, iż istnieje przedmiot denotowany przez termin „λ". To, czy wyrażenie „λ" jest nazwą tak interpretowanego języka *J*, nie zależy zatem od posiadania przez „λ" jakiejkolwiek denotacji. Może być sprawą czysto strukturalną. Każde wyrażenie o takim a takim kształcie zaliczyć możemy do nazw języka *J*, i dla każdego z nich zdanie (24) będzie prawdą. A więc zarówno dla nazwy „Sokrates", jak i dla nazwy „Pegaz". W języku takim postulat (4) nie będzie zatem w przypadku nieistnienia denotacji terminu „λ" nonsensem. Konsekwencją tego jest fakt, iż podana przez nas w części pierwszej definicja zdania analitycznego wyróżnia na gruncie takiego języka klasę zdań całkowicie od doświadczenia niezależnych, czyli pokrywającą się z klasą zdań analitycznych w sensie węższym.

Rozwiązanie to trudno jednak uznać za zadowalające. A to z powodu problematycznego charakteru owej «językowej» interpretacji kwantyfikatora. Jej sens nie jest bynajmniej jasny. Co więcej, odbiega wyraźnie od tych intuicji, jakie z pojęciem kwantyfikatora wiąże się na ogół w logice współczesnej. Ogromna większość jej przedstawicieli opowiada się za «przedmiotową» interpretacją kwantyfikatora, która słusznie uchodzić może za interpretację klasyczną. Przy niej zaś tautologia (24) posiada walor egzystencjalny i ogranicza, jak widzieliśmy, zasób nazw danego języka do terminów denotujących. Chcąc więc przy zachowaniu klasycznej interpretacji kwantyfikatora pozbyć się owego kłopotliwego ograniczenia, musimy nasz język poddać pod względem logicznym określonej modyfikacji. Modyfikacja ta musi pociągać odrzucenie tautologicznego charakteru twierdzenia (24). Musi więc reformować w odpowiedni sposób zakładany przez nas do tej pory rachunek logiczny.

Najprostszym a zarazem najbardziej drastycznym rozwiązaniem jest redukcja owego rachunku logicznego do węższego rachunku predykatów (z identycznością). Językiem odpowiadającym takiemu rachunkowi jest język, który prócz stałych logicznych (spójników rachunku zdań, kwantyfikatorów, znaku identyczności) obejmuje jeden tylko typ zmiennych: zmienne indywiduowe, i jeden typ terminów

14) Reprezentuje je m.in. C. Lejewski w artykule: „Logic and Existence", *British Journal for the Philosophy of Science* 5, 18 (1954).

deskryptywnych: predykaty pierwszego rzędu. Rzeczą istotną dla naszych rozważań jest to, iż w języku takim żaden termin deskryptywny nie ma charakteru nazwy, tj. nie należy do możliwych podstawień zmiennych związanych. Nie ma bowiem w tym języku nazw indywiduowych — jedynych terminów nadających się do podstawiania za zmienne indywiduowe, i nie ma zmiennych wyższego rzędu, których podstawieniami mogłyby być predykaty. W rezultacie dla żadnego terminu deskryptywnego takiego języka zdanie (24) nie może mieć waloru; zdanie to w języku takim po prostu nie daje się sformułować. Niechaj predykat „*Q*" będzie terminem wprowadzanym do tak scharakteryzowanego języka *J*. Skoro język *J* nie zmusza nas do uznania, iż istnieje przedmiot będący *Q*, nie musimy w metajęzyku *M* przyjmować, iż istnieje przedmiot będący denotacją terminu „*Q*". Możemy uważać „*Q*" za predykat — a więc wyrażenie sensowne języka *J* — niezależnie od tego, czy termin ten cokolwiek denotuje. Na mocy reguł czysto syntaktycznych „*Q*" należeć będzie do predykatów języka *J*, a poprawnie zbudowane zdania ów termin zawierające — do zdań języka *J*.

Tak scharakteryzowany język *J* nasuwa jednak natychmiast pytanie o możliwość jego rozszerzenia. Rozszerzeniem takim musi być przede wszystkim jego metajęzyk *M*. Nie ulega wątpliwości, iż musi to być język pod względem logicznym istotnie bogatszy od języka *J*. Inaczej sformułowanie, zgodnie z dotychczasowym schematem, reguł denotowania dla terminów języka *J* przedstawiałoby zadanie niewykonalne. Ale i z innych powodów możliwość rozszerzenia języka *J* ma znaczenie decydujące. *J* jest jako język przedmiotowy zbyt ubogi pod względem logicznym na to, aby mógł służyć do formułowania bogatszych teorii naukowych. Jako język elementarny nie pozwala na mówienie o dowolnych zbiorach, czy rodzinach zbiorów, o których w takich teoriach mowa. Ale czy wzbogacenie pod względem logicznym języka *J* nie pozbawi go owej, tak ważnej w naszych oczach, zalety: strukturalnych kryteriów sensowności? Tak byłoby niewątpliwie wtedy, gdyby owo wzbogacenie polegało na dobudowaniu do obowiązującego w języku *J* rachunku logicznego wyższych pięter teorii typów. Przywracalibyśmy bowiem w ten sposób charakter tautologii zdaniu typu (24), stwierdzającemu istnienie *Q*, a to zmuszałoby nas z kolei do uznania zdania typu (25), stwierdzającego istnienie denotacji terminu „*Q*". Ale nie jest to bynajmniej droga jedyna. Wzbogacenie języka *J* nie musi zakładać prostej teorii typów. Może być dokonane na gruncie innego aparatu logicznego. Rachunek logiczny języka *J* pozostaje wówczas w zasadzie ten sam, obejmując jeden rodzaj zmiennych i jeden rodzaj terminów deskryptywnych, nie stanowiących podstawień owych zmiennych. Nie ograniczamy jednak tutaj wartości zmiennych do określonego typu logicznego; mogą być nimi zarówno indywidua, jak i zbiory dowolnego typu. Do tak ujętego rachunku logicznego dołączamy teorię mnogości lub odpowiedni jej fragment, wprowadzając w ten sposób pojęcie zbioru scharakteryzowane odpowiednio dobranym układem aksjomatów. Przechodzimy tym samym na teren języka nieelementarnego, mogącego służyć do formułowania dowolnie bogatych teorii naukowych. Otóż takie wzbogacenie języka *J* nie pociąga zdań typu (24). Przedmiotem denotowanym przez predykat „*Q*" może być tylko pewien zbiór. A nie dla każdego predykatu tezą języka J jest zdanie głoszące, iż predykatowi temu odpowiada jakiś zbiór. Termin „Q" może być uznany za predykat — a więc wyrażenie sensowne języka J — mimo iż termin ten nie denotuje niczego. Tego rodzaju rozszerzenie języka J spełnia więc żądane przez nas warunki. W szczególności nadaje się do roli metajęzyka M.

Załóżmy, iż:

(41) $$\Phi(Q)$$

jest postulatem języka J. Regułę denotowania dla terminu „Q" sformułować można w metajęzyku M o opisanej strukturze w sposób następujący:

(42) $$\bigwedge_x (\text{„}Q\text{" denotuje } x \rightarrow Z(x) \wedge \Phi(x)),$$

gdzie „$Z(x)$" jest stwierdzeniem, iż x jest zbiorem[15]. Pozwala to na sformułowanie składnika syntetycznego postulatu (41), zgodnie z ogólnym schematem, jako wypowiedzi:

(43) $$\bigvee_x (Z(x) \wedge \Phi(x)),$$

a składnika analitycznego — jako wypowiedzi:

(44) $$\bigvee_x (Z(x) \wedge \Phi(x)) \rightarrow \Phi(Q).$$

Jeśli okazuje się teraz, iż nie ma zbioru spełniającego warunek Φ, termin „Q" nie denotuje żadnego przedmiotu. Skoro pozostaje mimo to predykatem języka J, postulat (41) możemy uznać za fałsz, a nie za nonsens. Składnik analityczny (44) pozostaje więc prawdą. Jeśli zdania analityczne określimy jako konsekwencje logiczne owego składnika, będą to zdania całkowicie od doświadczenia niezależne, analityczne w sensie węższym. Oczywiście, składnik analityczny (44) wyrazić się daje w języku J tylko wtedy, gdy jest to język nieelementarny.

Omawiane przez nas ostatnio języki predykatów — elementarne i nieelementarne — odznaczają się istotnie dużą logiczną prostotą. Czysto strukturalny charakter obowiązujących w nich kryteriów sensowności sprawia, iż problem należenia jakiegoś wyrażenia do danego języka nie jest nigdy sprawą doświadczenia. Definicja zdania jest definicją syntaktyczną. Te niewątpliwe zalety «teoretyczne» okupuje jednak język taki pewnymi wadami «praktycznymi». Dotkliwie zwłaszcza daje się odczuć przy próbach formułowania szeregu twierdzeń i teorii naukowych brak wyrażeń o charakterze nazw. Ale wprowadzenie nazw, a więc wyrażeń stanowiących podstawienia zmiennych związanych, pozbawia, jak widzieliśmy, dany język wspomnianej zalety. To bowiem, czy jakieś wyrażenie jest nazwą danego języka, nie jest sprawą czysto strukturalną. Próbo-

15) „$\Phi(x)$" powstaje z „$\Phi(Q)$" przez podstawienie wyrażeń typu „$y \in x$" na miejsce wyrażeń typu „$Q(y)$" — gdy „Q" jest predykatem jednoargumentowym; analogicznie — wyrażeń typu „$\langle y, z\rangle \in x$" na miejsce wyrażeń typu „$Q(y, z)$" — gdy „Q" jest predykatem dwuargumentowym, itp.

wano temu zaradzić zastępując nazwy odpowiednio zdefiniowanymi deskrypcjami[16]. Ale to prowadzi do znacznej komplikacji aparatu logicznego i z tego znów powodu okazuje się niewygodne. Zaproponowano również wyjście inne, o którym chciałbym na zakończenie wspomnieć w paru słowach[17].

Rozwiązanie to polega na dopuszczeniu do języka J nazw, tj. terminów stanowiących możliwe podstawienia zmiennych związanych, przy jednoczesnej modyfikacji klasycznego rachunku logicznego, tak by w charakterze nazw mogły bez sprzeczności figurować terminy nie denotujące żadnych przedmiotów. A więc nie tylko „Sokrates", ale i „Pegaz". W tak zmodyfikowanym rachunku logicznym nie może być dla dowolnej nazwy „a" tautologią zdanie:

(45) $$\bigvee_x (x = a).$$

Muszą więc pewnym ograniczeniom ulec wymienione poprzednio reguły inferencji: reguła podstawiania (22) i reguła egzystencjalnej generalizacji (23), które wspomnianą tautologię rodzą. Ograniczenia te przybierać mogą oczywiście taką lub inną postać zależnie od takiej lub innej postaci klasycznego rachunku logicznego. W każdym razie w ich rezultacie klasyczna reguła egzystencjalnej generalizacji:

(46) $$\Phi(a) \vdash \bigvee_x \Phi(x)$$

ustępuje miejsca regule słabszej:

(47) $$\bigvee_x (x = a), \Phi(a) \vdash \bigvee_x \Phi(x),$$

pozwalającej na wyprowadzenie wniosku, iż istnieje przedmiot spełniający warunek Φ, na podstawie tego, iż a spełnia warunek Φ, tylko wtedy, gdy istnieje przedmiot będący a, a więc gdy nazwa „a" posiada denotację. Podobnemu ograniczeniu ulega jednocześnie reguła podstawiania. Zdanie (44) nie ma w tak zmodyfikowanym rachunku logicznym charakteru tautologii. Jest w zasadzie pewnym twierdzeniem rzeczowym; prawdziwym, gdy „a" coś denotuje, fałszywym — w wypadku przeciwnym. W języku J odpowiadającym takiemu rachunkowi logicznemu określenie kategorii nazw — podobnie jak kategorii predykatów — ma charakter czysto syntaktyczny. Kryteria sensowności wyrażeń takiego języka nie są więc sprawą doświadczenia. Ceną, jaką za to płacimy, jest ograniczenie klasycznego rachunku logicznego. A to jest zawsze przedsięwzięciem ryzykownym. Można więc słusznie zapytywać, czy nie jest to cena zbyt wysoka. Jeśli idzie o problem zdań analitycznych, język obecny dzieli wszystkie zalety poprzedniego języka predykatów. W porównaniu z tym ostatnim posiada jeszcze i tę zaletę, iż w przypadku postulatów wprowadzających terminy o charakterze nazw pozwala zawsze na sformułowanie składników analitycznych owych postulatów w ich najogólniejszej postaci. Określenie zdań analitycznych języka J jako tych zdań języka

16) Por. znaną propozycję Quine'a.

17) Por. H. Leblanc i T. Hailperin: „Nondesignating Singular Terms", *The Philosophical Review* 68, 2 (1959); J. Hintikka: „Existential Presuppositions and Existential Commitments", *Journal of Philosophy* 56, 3 (1959).

J, które wynikają z dowolnego zbioru składników analitycznych postulatów języka *J* — wyodrębnia na gruncie takiego języka klasę zdań analitycznych w sensie węższym. Odpowiada tym samym naszym intuicjom związanym z pojęciem zdania analitycznego. Oczywista, nie wszystkim i nie bez zastrzeżeń. Takiego warunku nie spełnia żadne z określeń podawanych przez nas w toku tych wywodów. Traktować je można jedynie jako przybliżone i szkicowe sformułowania poszukiwanych rozwiązań.

Filozofia Nauki
Rok I, 1993, Nr 2-3

O pojęciu nieistotnego występowania terminów (przyczynek do słownika logicznego)

Mówimy, że dany termin w pewnych zdaniach występuje w sposób istotny, a w innych w sposób nieistotny. Odróżnienie to znajduje zastosowanie przy rozważaniu wielu problemów semantycznych. Co więcej, pojęcie nieistotnego występowania terminów wydaje się ważne nie tylko z logicznego, ale i z filozoficznego punktu widzenia. To, o czym mówimy naprawdę, a nie z pozoru tylko, wyznaczone jest przez coś więcej niż samą obecność takich a nie innych słów w naszych wypowiedziach. Chcąc stwierdzić, o czym naprawdę jest mowa, musimy sięgnąć nierzadko pod «powierzchnię» języka. W szczególności, tylko wtedy gdy pewien termin występuje w danym twierdzeniu w sposób istotny, twierdzenie to faktycznie głosi coś o tym, do czego się ów termin odnosi. Toteż odróżnianie istotnego i nieistotnego występowania terminów stanowi jeden z prostych, lecz skutecznych sposobów uwalniania się od «ułud» języka. Innym środkiem służącym temu samemu celowi jest postulowane przez reizm odróżnianie literalnego i zastępczego rozumienia zwrotów językowych.

Pojęcie nieistotnego występowania terminu Q w zdaniu Z określane bywa różnie. Najczęściej spotykamy dwa typy takich określeń. Wedle jednego z nich, termin Q występuje w zdaniu Z nieistotnie, gdy zdanie Z równoważne jest pewnemu zdaniu nie zawierającemu terminu Q. Drugi rodzaj eksplikacji nosi wyraźnie charakter semantyczny. W sformułowaniu swobodnym głosi on, że termin Q występuje w zdaniu Z nieistotnie, gdy wartość logiczna zdania Z niezależna jest od sposobu interpretacji terminu Q. Określenia te okazują się zresztą — przy odpowiednich założeniach — sformułowaniami równoważnymi. Chcąc jednakże rozpatrzyć ich konsekwencje i wzajemne zależności, musimy je przede wszystkim poddać uściśleniu. W swej obecnej postaci są one jawnie niedookreślone i dopuszczają w rezultacie różne interpretacje. O jakiej równoważności mowa w określeniu pierwszym? Logicznej, «analitycznej», czy «faktycznej»? W określeniu drugim niedopowiedzenia dotyczą przede wszystkim rodzaju interpretacji pozostałych terminów występujących w zdaniu Z. Czy warunek, jaki ta definicja formułuje, zachodzić ma dla dowolnej interpretacji pozostałych terminów, czy tylko dla ich interpretacji «dopuszczalnej», lub «właściwej»? Każda z tych ewentualności prowadzi do innego nieco pojęcia nieistotnego występowania terminu Q w zdaniu Z. Każde z tych pojęć odpowiada, jak się zdaje, pewnym intuicjom wiązanym z owym zwrotem i każde znajduje zastosowanie w stosownej dziedzinie rozważań. Spróbujmy zatem dokonać — częściowej bodaj — eksplikacji i typologii głównych wersji omawianego pojęcia.

Musimy się w tym celu oprzeć na określonym aparacie pojęciowym i określonych założeniach semantycznych. Dostarczy nam ich współczesna semantyka logiczna, pojęta jako teoria modeli języków sformalizowanych. Przedmiotem naszych rozważań będzie zatem standardowo sformalizowany język J. Jego interpretacje utożsamiać będziemy z jego modelami. Chcąc uprościć dalsze sformułowania i ilustrujące je przykłady, ograniczymy się do pewnego przypadku szczególnego. Założymy, że nasz język J zawiera — prócz stałych logicznych obejmujących spójniki zdaniowe i kwantyfikatory — następujące wyrażenia pozalogiczne: zmienne nazwowe $x_1, x_2, ...$, nazwy $a_1, ..., a_k$ oraz predykaty $P_1, ..., P_m, Q$. Przyjmiemy dla prostoty, że predykaty P_1 i Q są jednoargumentowe. Modele języka J oznaczać będziemy symbolem $\mathfrak{M}$ (lub $\mathfrak{M}'$). Model taki, jak wiadomo, ma postać układu:

$$\mathfrak{M} = \langle U, b_1, ..., b_k, R_1, ..., R_m, S\rangle,$$

złożonego z niepustego zbioru U, z indywiduów $b_1, ..., b_k$ wyróżnionych ze zbioru U i z relacji (zbiorów) $R_1, ..., R_m, S$ określonych w zbiorze U. Każdy model języka J wyznacza jedną z interpretacji tego języka: jego zmiennym nazwowym przyporządkowuje jako zbiór wartości zbiór U, a nazwom i predykatom jako ich denotacje odpowiednie indywidua i relacje (zbiory). Pojęcie modelu $\mathfrak{M}$ języka J pozwala zdefiniować w znany dobrze sposób pojęcie zdania języka J prawdziwego w modelu $\mathfrak{M}$. Mówiąc swobodnie, zdanie Z jest prawdziwe w modelu $\mathfrak{M}$, gdy jest tak, jak głosi zdanie Z w interpretacji wyznaczonej przez model $\mathfrak{M}$. Przykładowo: zdanie $P_1\, a_1$ jest prawdziwe w modelu $\mathfrak{M}$, gdy przedmiot b_1 należy do zbioru R_1. Ponieważ w dalszym ciągu interesować nas będzie status terminu Q, wyróżnimy w języku J część nie obejmującą tego predykatu. Będzie nią język J_0 zawierający jako jedyne stałe pozalogiczne nazwy $a_1, ..., a_k$ i predykaty $P_1, ..., P_m$. Modele języka J_0 oznaczać będziemy symbolem $\mathfrak{M}_0$. Są to układy postaci:

$$\mathfrak{M}_0 = \langle U, b_1, ..., b_k, R_1, ..., R_m\rangle,$$

Symbolem $\mathfrak{M}\,|_0$ oznaczać będziemy tzw. redukt modelu $\mathfrak{M}$ do języka J_0, tj. model $\mathfrak{M}_0$ języka J_0 otrzymany z modelu $\mathfrak{M}$ języka J przez usunięcie z tego ostatniego denotacji predykatu Q, czyli zbioru S. Warunek: $\mathfrak{M}\,|_0 = \mathfrak{M}'\,|_0$ stwierdza zatem, że modele $\mathfrak{M}$ i $\mathfrak{M}'$ języka J różnią się co najwyżej denotacją predykatu Q; zarówno ich *universa*, jak i denotacje nazw i predykatów pozostałych są identyczne.

Wprowadzony aparat pojęciowy pozwala na wyraźne sformułowanie wspomnianych na wstępie definicji nieistotnego występowania terminów. Jako punkt wyjścia przyjmiemy definicję drugą, odwołującą się do kryterium semantycznego. Jej pierwotne ogólnikowe sformułowanie daje podstawę do skonstruowania co najwyżej pewnego schematu definicyjnego, a nie konkretnej definicji. Schemat ten głosi co następuje:

(I) termin Q występuje nieistotnie w zdaniu Z, gdy dla dowolnego modelu $\mathfrak{M}_0$ należącego do klasy K_0 oraz dla dowolnych modeli $\mathfrak{M}$ i $\mathfrak{M}'$ takich, że $\mathfrak{M}|_0 = \mathfrak{M}'|_0 = \mathfrak{M}_0$: Z jest prawdziwe w $\mathfrak{M}$ wtedy i tylko wtedy, gdy Z jest prawdziwe w $\mathfrak{M}'$.

(I) jest schematem jedynie, gdyż odwołuje się do bliżej nieokreślonej klasy modeli języka J_0, K_0. Dopiero ustalenie, o jaką to klasę chodzi, czyni z (I) konkretną definicję. Mówiąc inaczej, (I) jest definicją pewnego pojęcia relatywnego: termin Q występuje nieistotnie w zdaniu Z ze względu na klasę K_0. Przejście do definicji odpowiedniego pojęcia absolutnego wymaga zastąpienia K_0 nazwą określonej klasy modeli języka J_0. W zależności od tego, co to będzie za klasa, otrzymamy takie a nie inne pojęcie absolutne. Ideę, którą ujmuje schemat (I), można wyrazić następująco. Termin Q występuje w zdaniu Z w sposób nieistotny, gdy to, jak zinterpretujemy termin Q — jaką przypiszemy mu denotację — nie ma wpływu na wartość logiczną zdania Z; wartość ta wyznaczona jest jednoznacznie przez interpretację pozostałych terminów zdania Z — tj. przez model języka J_0 — jeśli tylko jest to model należący do klasy K_0.

Przy przyjętych przez nas założeniach łatwo można okazać, że schemat (I) równoważny jest następującemu schematowi, który odpowiada pierwszej z wymienionych przez nas eksplikacji pojęcia nieistotnego występowania terminów:

(1) termin Q występuje nieistotnie w zdaniu Z, gdy istnieje zdanie Z_0 nie zawierające terminu Q takie, iż dla dowolnego modelu $\mathfrak{M}_0$ należącego do klasy K_0 oraz dla dowolnego modelu $\mathfrak{M}$ takiego, że $\mathfrak{M}|_0 = \mathfrak{M}_0$: zdanie $Z \equiv Z_0$ jest prawdziwe w $\mathfrak{M}$.

I w tym przypadku przechodzimy ze schematu (1) do określonej definicji przez jednoznaczne ustalenie klasy K_0.

Najmocniejszą wersję definiowanego pojęcia otrzymujemy utożsamiając w powyższych schematach klasę K_0 z klasą wszystkich modeli języka J_0; oznaczmy ją przez $\mathbf{K}_0$:

(A) $K_0 = \mathbf{K}_0$.

Założenie (A) prowadzi na gruncie schematu (I) do definicji nieistotnego występowania terminów, którą oznaczać będziemy przez (IA) (takiej samej konwencji trzymać się będziemy i w dalszych przypadkach). Jeśli termin Q występuje w zdaniu Z nieistotnie w sensie definicji (IA), to jakąkolwiek przyjmiemy interpretację pozostałych terminów zdania Z, interpretacja ta przesądza wartość logiczną tego zdania w sposób jednoznaczny. Założenie (A) prowadzi w zastosowaniu do schematu (1) do szczególnie prostego sformułowania kryterium nieistotności: termin Q występuje w zdaniu Z nieistotnie, gdy zdanie to jest równoważne logicznie jakiemuś zdaniu nie zawierającemu terminu Q. Jako przykłady zdań tego rodzaju wymienić można — prócz wszystkich tautologii i kontrtautologii, a więc zdań typu: $Qa_1 \vee \sim Qa_1$ czy $Qa_1 \wedge \sim Qa_1$ — zdania typu: $P_1a_1 \wedge (\sim P_1a_1 \rightarrow Qa_1)$ itp.

Mimo że definicje (IA) i (1A) chwytają najczęściej chyba spotykane rozumienie nieistotności, są to — z pewnego punktu widzenia — definicje zbyt rygorystyczne. Odwołują się one do całkowicie dowolnych interpretacji języka J_0 — do wszystkich bez wyjątku jego modeli. Skoro jednak J_0 ma być językiem sensownym, wyposażonym w określone znaczenie, nie każda jego interpretacja będzie z tym znaczeniem zgodna. Możemy więc w przypadku każdego takiego języka wyróżnić klasę jego interpretacji «dopuszczalnych» (lub «możliwych»). Jakie modele języka J_0 będą do owej klasy należeć? Klasyczny sposób określenia klasy modeli dopuszczalnych odwołuje się do pojęcia postulatu znaczeniowego. Zakłada się, że znaczenie terminów języka J_0 wyznaczone jest, między innymi, przez pewien zbiór zdań tego języka, nakładający określone warunki na sposób interpretacji poszczególnych terminów i ich wzajemne związki. Zdania te nazywamy postulatami znaczeniowymi języka J_0. Niech P_0 symbolizuje ich zbiór. Otóż klasa modeli dopuszczalnych — to, przy tym ujęciu, klasa tych modeli języka J_0, w których prawdziwe są postulaty P_0; krótko — klasa modeli postulatów P_0, symbolicznie — $\mathbf{K}_0(P_0)$. Z nią właśnie utożsamić możemy klasę K_0 ze schematów (I) i (1):

(B) $K_0 = \mathbf{K}_0(P_0)$.

Pomijamy tym samym przy kryterium nieistotnego występowania terminu Q te interpretacje terminów pozostałych, które nie są interpretacjami dopuszczalnymi. Do tego, aby termin Q występował w zdaniu Z nieistotnie w sensie definicji (IB), potrzeba i wystarcza, by każda dopuszczalna interpretacja terminów pozostałych determinowała jednoznacznie wartość logiczną zdania Z. Zgodnie ze sformułowaniem (1B), termin Q występuje w zdaniu Z nieistotnie, gdy zdanie to jest równoważne «analitycznie» — czyli na gruncie postulatów znaczeniowych P_0 — jakiemuś zdaniu języka J_0. Prócz przykładów podanych poprzednio do zdań takich należeć będzie np. zdanie $P_1a_1 \vee Qa_1$, jeśli P_1a_1 jest prawdziwe w każdym modelu klasy $\mathbf{K}_0(P_0)$, lub — co na jedno wychodzi — jeśli P_1a_1 jest konsekwencją logiczną zbioru P_0; podobnie — zdanie $P_1a_1 \wedge Qa_1$, w przypadku gdy P_1a_1 jest fałszywe w każdym modelu klasy $\mathbf{K}_0(P_0)$, a więc gdy ze zbioru P_0 wynika logicznie negacja P_1a_1.

Jeśli J_0 jest językiem empirycznym, znaczenia jego terminów nie wyczerpuje najbogatszy nawet układ postulatów znaczeniowych. Znaczenie to ustalane jest również przy pomocy pewnych środków pozawerbalnych (takich jak definicja ostensywna). W ten sposób dopiero wyznaczone zostaje to, co stanowi interpretację «właściwą» (lub «zamierzoną») języka J_0. Klasa owych modeli właściwych nie pokrywa się z klasą modeli dopuszczalnych, choć jest w niej oczywiście zawarta. Oznaczymy ją więc odrębnym symbolem $\mathbf{K}_0^*$. Otóż jedna z możliwych konkretyzacji schematów (I) i (1) polega na utożsamieniu klasy K_0 z klasą modeli właściwych języka J_0:

(C) $K_0 = \mathbf{K}_0^*$.

Dochodzimy w ten sposób do jeszcze luźniejszego pojęcia nieistotnego występowania terminu Q. Wystarcza, jeśli każda interpretacja właściwa terminów pozostałych jednoznacznie determinuje wartość logiczną zdania Z, abyśmy mogli uznać, że termin Q występuje w tym zdaniu w sposób nieistotny w sensie definicji (IC).

Na szczególną uwagę zasługuje pewien przypadek skrajny sytuacji rozważanej obecnie. Mamy z nim do czynienia wtedy, gdy interpretacja właściwa języka J_0 określona jest w sposób jednoznaczny. Klasa modeli właściwych jest wówczas klasą jednostkową: obejmuje jeden model właściwy, $\mathfrak{M}_0^*$. Założenie (C) przybiera w tym przypadku postać szczególną:

(D) $K_0 = \{\mathfrak{M}_0^*\}$.

Zgodnie z definicją (ID), termin Q występuje w zdaniu Z nieistotnie, gdy wartość logiczna tego zdania zdeterminowana jest przez sam model $\mathfrak{M}_0^*$, a więc bez względu na to, jaką denotację przypiszemy terminowi Q. Ten sam warunek sformułować możemy, zgodnie z definicją (1D), jako warunek żądający, aby zdanie Z było równoważne «faktycznie» — tj. w każdym modelu języka J o redukcie $\mathfrak{M}_0^*$ — jakiemuś zdaniu nie zawierającemu terminu Q. Jakie zdania języka J spełniają taki warunek? Prócz zdań wymienionych poprzednio, należeć do nich będzie m.in. zdanie $P_1a_1 \vee Qa_1$, gdy P_1a_1 jest prawdziwe w $\mathfrak{M}_0^*$, lub zdanie $P_1a_1 \wedge Qa_1$, gdy P_1a_1 jest fałszywe w $\mathfrak{M}_0^*$. Widać na tych przykładach, że w przeciwieństwie do kryteriów (IA) i (IB) kryteria (IC) i (ID) odwołują się do pewnych faktów empirycznych; taki charakter bowiem może mieć fakt prawdziwości zdania P_1a_1 w modelu $\mathfrak{M}_0^*$. Stwierdzenie, iż termin Q występuje nieistotnie w zdaniu Z, może więc w przypadku definicji (IC) i (ID) wymagać odwołania się do doświadczenia.

Pojęcie ostatnio zdefiniowane jest najsłabszym z pojęć dotąd uwzględnionych. Zależności, jakie zachodzą między wyróżnionymi klasami modeli języka J_0:

$$\{\mathfrak{M}_0^*\} \subset \mathbf{K}_0^* \subset \mathbf{K}_0(P_0) \subset \mathbf{K}_0$$

pociągają określone zależności między zdefiniowanymi przy ich pomocy pojęciami nieistotnego występowania terminów. Jeśli symbole IA, IB, IC i ID potraktujemy jako nazwy odpowiednich relacji nieistotnego występowania terminu Q w zdaniu Z, to zależności między tymi relacjami przedstawić możemy następująco:

$$\mathrm{IA} \subset \mathrm{IB} \subset \mathrm{IC} \subset \mathrm{ID}.$$

Pojęcia powyższe nie wyczerpują jednak tych możliwości, jakie zawiera w sobie idea nieistotnego występowania terminów. Wszystkie definicje dotychczasowe były konkretyzacjami tego samego schematu (I) (lub równoważnego mu schematu (1)). Wydaje się jednak, że nie jest to jedyny schemat definicyjny dla omawianego pojęcia. Na uwagę zasługuje również schemat będący istotnym osłabieniem poprzedniego (powstający z tamtego przez zastąpienie kwantyfikatora ogólnego „dla dowolnego modelu $\mathfrak{M}_0$” kwantyfikatorem szczegółowym „dla pewnego modelu $\mathfrak{M}_0$”):

(II) termin Q występuje nieistotnie w zdaniu Z, gdy dla pewnego modelu $\mathfrak{M}_0$ należącego do klasy K_0 oraz dla dowolnych modeli $\mathfrak{M}$ i $\mathfrak{M}'$ takich, że $\mathfrak{M}|_0 = \mathfrak{M}'|_0 = \mathfrak{M}_0$: Z jest prawdziwe w $\mathfrak{M}$ wtedy i tylko wtedy, gdy Z jest prawdziwe w $\mathfrak{M}'$.

Można okazać, podobnie jak w przypadku poprzednim, że schemat (II) równoważny jest schematowi stanowiącemu analogiczne osłabienie schematu (1):

(2) termin Q występuje nieistotnie w zdaniu Z, gdy istnieje zdanie Z_0 nie zawierające terminu Q takie, iż dla pewnego modelu $\mathfrak{M}_0$ należącego do klasy K_0 oraz dla dowolnego modelu $\mathfrak{M}$ takiego, że $\mathfrak{M}|_0 = \mathfrak{M}_0$: zdanie $Z \equiv Z_0$ jest prawdziwe w $\mathfrak{M}$.

Podobnie też jak poprzednio, uzyskujemy ze schematów powyższych konkretne definicje nieistotnego występowania terminu Q w zdaniu Z, identyfikując klasę modeli K_0 kolejno z klasami: $\mathbf{K}_0$, $\mathbf{K}_0(P_0)$, $\mathbf{K}_0^*$, czy $\{\mathfrak{M}_0^*\}$ i dochodząc w ten sposób do pojęć: (IIA), (IIB), (IIC) i (IID). Zależności między nimi są odwrotne niż w przypadku poprzednim; (IIA) jest pojęciem najsłabszym, (IID) — najmocniejszym (identycznym z (ID)):

$$\mathrm{ID} = \mathrm{IID} \subset \mathrm{IIC} \subset \mathrm{IIB} \subset \mathrm{IIA}.$$

Pojęcia tak zdefiniowane są w rezultacie pojęciami bardzo słabymi, a stąd bardzo szerokimi. Czy odpowiadają im mimo to jakieś intuicje wiązane z nieistotnym występowaniem terminów?

Warto przede wszystkim zwrócić uwagę na fakt, że mimo swego liberalizmu każde z tych pojęć wyłącza pewne rodzaje zdań ze swego zakresu. Dotyczy to również pojęcia najsłabszego, (IIA). Zgodnie z definicją (IIA), termin Q występuje nieistotnie nie tylko we wszystkich przypadkach przytaczanych dotychczas, ale i w dowolnych zdaniach typu: $P_1a_1 \vee Qa_1$, czy $P_1a_1 \wedge Qa_1$. Z łatwością jednak można podać przykłady zdań, w których termin Q występuje w sposób istotny. Należą do nich zdania takie jak: $\bigvee x\, Qx$, Qa_1, $P_1a_1 \equiv Qa_1$ itp. To samo dotyczy *a fortiori* pojęć pozostałych. Zauważmy przy tym, że im słabsze jest pojęcie nieistotnego występowania terminów, tym mocniejsza jest jego negacja: pojęcie istotnego występowania terminu Q w zdaniu Z. Walor pojęć definiowanych przez schemat (II) upatrywać więc można w tym, że dostarczają one rygorystycznych definicji istotnego występowania terminów, a to pojęcie nierzadko tak właśnie bywa rozumiane i, jak się wydaje, w pewnych kontekstach tak rozumiane być powinno. Chcemy niekiedy o istotnym występowaniu terminu Q w zdaniu Z mówić dopiero wtedy, gdy przy każdej (ewentualnie, przy każdej dopuszczalnej, lub właściwej) interpretacji terminów pozostałych jest tak, że interpretacja terminu Q ma wpływ na wartość logiczną zdania Z. A tę właśnie ideę realizują definicje odpowiadające schematowi (II).

Jeśli by można było mówić o najwłaściwszym rozumieniu pojęcia nieistotnego (czy istotnego) występowania terminów, to rozumienie takie skłonny byłbym wiązać z tymi wersjami tego pojęcia, które odwołują się do klasy dopuszczalnych interpretacji języka J_0, $\mathbf{K}_0(P_0)$ — a zatem z pojęciami (IB) i (IIB). One to chwytają ideę nieistotnego i istotnego występowania terminów w sposób, w moim przekonaniu, najtrafniejszy.

Na koniec parę słów o możliwości uogólnienia powyższych rozważań. Uogólnienie na terminy o innych kategoriach syntaktycznych (predykaty k-argumentowe, nazwy, symbole funkcyjne) nie przedstawia żadnych trudności. Pewne problemy powstają natomiast wtedy, gdy chcemy mówić o nieistotnym występowaniu w zdaniu Z nie jednego terminu Q, lecz kilku terminów równocześnie: $Q_1, ..., Q_n$. Dwie zarysowują się tutaj możliwości. Jedna z nich sprowadza się do tego, aby o nieistotnym występowaniu tych terminów mówić wtedy, gdy każdy z nich z osobna występuje nieistotnie ze względu na wszystkie pozostałe terminy języka J. Kryterium nieistotnego występowania terminu Q_i ($i = 1, ..., n$) odwoływać się więc będzie jako do języka J_0 do tej części języka J, która jako stałe pozalogiczne zawiera terminy następujące: $a_1, ..., a_k$, $P_1, ..., P_m$, $Q_1, ..., Q_{i-1}$, $Q_{i+1}, ..., Q_n$. Druga możliwość polega na tym, aby mówić o nieistotnym występowaniu układu terminów $\langle Q_1, ..., Q_n\rangle$ ze względu na pozostałe terminy języka J. Ponieważ owe terminy pozostałe to terminy: $a_1, ..., a_k, P_1, ..., P_m$, język J_0 jest tutaj rozumiany tak samo jak poprzednio. Łatwo sprawdzić, że te dwa pojęcia się nie pokrywają. Każdy z terminów $Q_1, ..., Q_n$ może występować nieistotnie, a ich układ $\langle Q_1, ..., Q_n\rangle$ — istotnie w pewnym zdaniu Z. Od konkretnej dziedziny zastosowań zależeć będzie wybór któregoś z owych pojęć. Podobnie jak poprzednio, tak i tutaj zadaniem naszym było jedynie zarysowanie istniejących w tej sprawie możliwości.

Filozofia Nauki
Rok I, 1993, Nr 2-3

Analityczność i syntetyczność

1. Pojęcie analityczności, podobnie jak syntetyczności, odnosi się do zdań lub — wedle dawniejszych koncepcji — do sądów pojmowanych jako znaczenia zdań.

Prawdziwość zdania, mówiąc najogólniej, zależeć może zarówno od sensu terminów, które w nim występują, jak i od faktów, o których mówi. Jeśli zależy ona istotnie od obu tych czynników, zdanie nazywane jest syntetycznym; jeśli natomiast wyznaczona jest przez sam sens terminów, zdanie zaliczane jest do analitycznych. Prawdy syntetyczne są opisami świata nas otaczającego, prawdy analityczne — jedynie «ubocznymi produktami» języka, za pomocą którego ów świat opisujemy.

Rozróżnienie analitycznych i syntetycznych zdań (lub sądów), choć kwestionowane z różnych powodów przez niektórych filozofów (ich najgłośniejszym współczesnym przedstawicielem jest Quine), przez innych — w szczególności tych, którzy reprezentują różne odmiany filozofii logicznego empiryzmu — uważane jest za podstawowy element logicznej teorii nauki.

Rozróżnienie to związane jest z innym, bardziej fundamentalnym podziałem ogółu prawd na prawdy aprioryczne i empiryczne: niezależne i zależne od doświadczenia. Prawdy analityczne, będąc twierdzeniami uzasadnialnymi bez odwołania się do danych doświadczenia, należą do prawd apriorycznych. Stanowią one najmniej kontrowersyjny rodzaj sądów *a priori*, w przeciwieństwie do sądów syntetycznych *a priori*, których istnienie jest jednym z najbardziej spornych problemów filozoficznych. Dyskusja tocząca się wokół prawd analitycznych nie kwestionuje na ogół tego, że poznanie nasze zawiera składnik językowy, którego prawdziwość zagwarantowana jest przez sam sens wyrażeń danego języka. W dyskusji tej idzie głównie o to, czy składnik ten daje się wyodrębnić w postaci ściśle określonej klasy zdań.

Pojęcie analitycznych i syntetycznych zdań (ściślej sądów) wprowadzone zostało w sposób wyraźny przez Kanta. W innej terminologii rozróżnienie to pojawiło się w dziejach filozofii już wcześniej. Bezpośrednim poprzednikiem Kanta był pod tym względem Leibniz, który wyróżniał prawdy rozumowe i prawdy faktyczne, określając pierwsze jako prawdziwe we wszelkich możliwych światach, drugie jako prawdziwe tylko w świecie rzeczywistym. Kant definiował sąd analityczny jako taki, w którym pojęcie orzecznika zawarte jest (w sposób ukryty) w pojęciu podmiotu. Podczas gdy sądy syntetyczne zawierają pewną informację o podmiocie, sądy analityczne wyjaśniają jedynie pojęcie, pod które podmiot podpada. Określenia te dalekie są od jasności i precyzji. Próbowano je więc uściślać na różne sposoby, dążąc do zbudowania ogólnej i konsekwentnej teorii prawd analitycznych. Próby te nawiązywały zarówno do koncepcji Kanta jak i Leibniza.

2. Definicja Kanta ograniczona jest ponadto do zdań szczególnego rodzaju: twierdzeń podmiotowo-orzecznikowych. Najczęstszą próbą jej uogólnienia jest charakterystyka zdań analitycznych jako zdań prawdziwych na mocy samego znaczenia terminów w nich występujących. Jest to określenie ogólnikowe, dopuszczające różne eksplikacje. Wstępny krok do jego uściślenia stanowić może zakresowa charakterystyka pojęcia analityczności — ustalenie głównych rodzajów zdań, jakie podpadać mają pod to pojęcie. Wśród tych, którzy posługują się obecnie pojęciem analityczności, istnieje prawie powszechna zgoda co do tego, że do zdań takich należą wszystkie prawdy logiczne: twierdzenia logiki i ich podstawienia, zwłaszcza gdy logika utożsamiona zostaje z elementarnym rachunkiem logicznym (węższym rachunkiem kwantyfikatorów). Z uwagi na to, że pojęcie prawdy logicznej doczekało się w logice współczesnej zadowalającej definicji, ten rodzaj zdań analitycznych nie budzi poważniejszych wątpliwości formalnych czy merytorycznych. Nie wyczerpuje on jednak ogółu zdań analitycznych ani w języku potocznym, ani w języku nauki.

Prawdy logiczne to zdania prawdziwe na mocy znaczenia samych stałych logicznych, które w nich występują. Istnieją jednak zdania, których prawdziwość wydaje się zagwarantowana przez znaczenie nie tylko stałych logicznych, ale i terminów pozalogicznych, w nich zawartych. Podczas gdy zdanie: „Żaden nieżonaty mężczyzna nie jest żonaty", może być przykładem zdania prawdziwego na mocy znaczenia samych stałych logicznych, takich jak „żaden", „nie", czy „jest", zdanie: „Żaden kawaler nie jest żonaty", wypada uznać za prawdziwe na mocy znaczenia wszystkich terminów w nim zawartych, w tym terminów pozalogicznych, takich jak „kawaler" czy „żonaty". Ten rodzaj zdań analitycznych próbowano definiować za pomocą różnych, ściślejszych nieco kryteriów, najczęściej jako zdania otrzymywalne z prawd logicznych przez podstawienie wyrazów równoznacznych (np. wyrażenia „kawaler" zamiast „nieżonaty mężczyzna"), a więc jako zdania równoznaczne z prawdami logicznymi.

Podkreśla się jednak, że pojęcie równoznaczności wymaga wyjaśnienia w takim samym stopniu, co pojęcie analityczności (gdyż nawet samo za pomocą tego ostatniego bywa definiowane). Powyższe określenie zastępuje się więc często kryterium bardziej uchwytnym, identyfikującym zdania analityczne z logicznymi konsekwencjami definicji terminów pozalogicznych w nich zawartych. Kryterium to ujawnia jednak wyraźnie swój ograniczony charakter. Definicje nie stanowią na pewno jedynego sposobu ustalania znaczeń terminów, jeśli nawet ograniczyć się do sposobów wyłącznie werbalnych (pomijając takie niewerbalne procedury, jak tzw. definicja ostensywna).

Twierdzenia będące środkiem ustalania znaczeń terminów przyjmują niejednokrotnie formy różniące się od definicji w ścisłym tego słowa znaczeniu. Przykład mogą stanowić tzw. zdania redukcyjne wprowadzające pojęcia dyspozycyjne lub układy aksjomatów charakteryzujące pojęcia pierwotne teorii naukowych. Wszystkie takie twierdzenia, niezależnie od ich postaci, nazywane są postulatami znaczeniowymi. Bliższa eksplikacja tego rodzaju zdań odwołuje się do pojęcia reguły semantycznej (lub konwencji terminologicznej). Język sensowny pojmowany jest jako całość charakteryzowana nie tylko przez reguły składniowe, ale i przez reguły semantyczne określające znaczenie (lub denotacje) prostych wyrażeń językowych. Zdanie $F(t)$ nazywane jest postulatem znaczeniowym dla terminu t, jeśli obowiązuje w danym języku reguła semantyczna żądająca, aby termin t rozumiany był w sposób taki, który zapewnia prawdziwość zdania $F(t)$ lub — w innym nieco ujęciu — aby termin t denotował przedmiot taki, który spełnia warunek $F(x)$. Dane zdanie może być oczywiście postulatem znaczeniowym dla więcej niż jednego terminu, a dany termin może być określany przez więcej niż jeden postulat. Pojęcie postulatu znaczeniowego pozwala zdefiniować zdania analityczne jako logiczne konsekwencje zbioru wszystkich postulatów znaczeniowych danego języka. Tak określone zdania analityczne obejmują, oprócz prawd logicznych, również zdania prawdziwe na mocy znaczenia terminów pozalogicznych, ponieważ znaczenia tych ostatnich mają być zdeterminowane właśnie przez zbiór postulatów znaczeniowych danego języka.

3. Pojęcie postulatu znaczeniowego umożliwia również prostą eksplikację pojęcia „świata możliwego", a tym samym Leibnizowskiej koncepcji prawd analitycznych («prawd rozumu»). W jednym ze znaczeń tego terminu to, co się nazywa — ze względu na dany język — „światem możliwym", może być utożsamione z możliwą interpretacją tego języka, rozumianą jako dowolna dziedzina, do której się tak określony język może odnosić. Aby stanowić interpretację możliwą, dziedzina ta musi być modelem postulatów znaczeniowych danego języka, tj. dziedziną, w której postulaty te są prawdziwe. Zdania analityczne, rozumiane jako zdania prawdziwe we wszystkich możliwych światach, mogą być wobec tego zdefiniowane jako zdania prawdziwe we wszystkich modelach postulatów znaczeniowych, tj. we wszystkich dziedzinach, w których postulaty te są prawdziwe.

W zastosowaniu do języków opartych na elementarnym rachunku logicznym definicja ta sprowadza się do poprzedniej, identyfikującej zdania analityczne z logicznymi konsekwencjami postulatów znaczeniowych. Utożsamianie «świata możliwego» z modelem postulatów znaczeniowych jest jednak pewnym uproszczeniem nie zawsze dopuszczalnym. W przypadku języków bogatszych teorii naukowych nie wszystkie związki znaczeniowe między ich terminami wyrażane są za pomocą postulatów formułowanych w danym języku, co uniemożliwia traktowanie dowolnego modelu tych postulatów jako możliwej interpretacji języka. Interpretacja taka, a zatem i «świat możliwy» muszą być określone w tych przypadkach innym sposobem.

4. Definicja zdań analitycznych odwołująca się do pojęcia postulatu znaczeniowego stwarza specyficzne problemy przy próbach jej stosowania do rzeczywistych systemów językowych, w szczególności do języków istniejących teorii naukowych. Powstają bowiem pytania dotyczące rodzaju twierdzeń pełniących w tych systemach rolę postulatów znaczeniowych oraz odróżnienia ich od pozostałych twierdzeń danej teorii. Nie są to systemy o wyraźnie określonych regułach semantycznych, zatem odpowiedzi na te pytania zależeć będą od takiej czy innej dopuszczalnej kodyfikacji owych reguł.

Na ogół zakłada się, że istnieje pod rozważanym względem zasadnicza różnica między teoriami formalnymi (logicznymi i matematycznymi) a teoriami empirycznymi. Wszystkie twierdzenia teorii formalnej mają, zgodnie z tym poglądem, taki sam charakter semantyczny. Ci, którzy w ogóle posługują się pojęciem analityczności, przypisują im z reguły charakter prawd analitycznych. Jeśli ujmuje się daną teorię formalną w postaci aksjomatycznej, jej aksjomaty traktowane są na gruncie takiego stanowiska jako postulaty znaczeniowe dla wszystkich terminów specyficznych, które zawierają.

Problem wyróżnienia postulatów znaczeniowych, a w konsekwencji zdań analitycznych, w języku teorii empirycznej jest bardziej złożony i kontrowersyjny. Z jednej strony, teoria empiryczna jako taka zawierać musi pewne twierdzenia empiryczne, a więc zdania nieanalityczne. Jej aksjomaty nie mogą pełnić zatem jedynie funkcji postulatów znaczeniowych. Z drugiej strony, w typowej teorii empirycznej znaczenie niektórych przynajmniej jej terminów jest, jak się wydaje, wyznaczone właśnie przez układ jej aksjomatów (rozumiany tu zgodnie z przyjętą praktyką tak szeroko, że obejmuje m.in. również definicje). Trudność ta rozwiązywana bywa różnie.

Najbardziej znane jest rozwiązanie oparte na pewnych założeniach empirystycznych. Zakłada ono, że terminy specyficzne typowej teorii empirycznej dzielą się na dwie klasy: klasę terminów — ze względu na daną teorię — nieteoretycznych (w skrajnym przypadku — obserwacyjnych) oraz klasę terminów teoretycznych. Pierwsze wyposażone są w znaczenie z góry, w sposób niezależny od aksjomatów danej teorii. Drugie uzyskują znaczenie tylko dzięki aksjomatom danej teorii, które wiążą je z terminami poprzednimi. Typowa rekonstrukcja logiczna mechaniki klasycznej zalicza np. termin „położenie" do klasy pierwszej, terminy „masa" i „siła" do klasy drugiej: w przeciwieństwie do pierwszego znaczenie tych ostatnich wydaje się wyznaczone właśnie przez naczelne zasady mechaniki.

Aksjomaty teorii empirycznej pełnią zatem funkcję dwojaką. Z jednej strony stwierdzają pewne fakty empiryczne, m.in. ogólne prawidłowości empiryczne, wyrażalne za pomocą terminów nieteoretycznych, z drugiej zaś determinują znaczenie terminów teoretycznych (w szczególności — ich denotacje). Powstaje w związku z tym problem wyodrębnienia tych dwóch funkcji: rozbicia ogółu aksjomatów teorii empirycznej na dwa składniki — faktualny, spełniający tylko funkcję pierwszą, i konwencjonalny, spełniający tylko drugą. Zbiór postulatów znaczeniowych dla terminów teoretycznych winien być utożsamiony nie z ogółem aksjomatów danej teorii empirycznej, tylko z ich składnikiem konwencjonalnym. Rozwiązanie tego problemu (sformułowane po raz pierwszy przez Carnapa, a potem rozwijane i modyfikowane przez innych autorów) przedstawia się w skrócie i uproszczeniu, jak następuje. Niech

$$F(o_1, ..., o_m, t_1, ..., t_n)$$

będzie koniunkcją aksjomatów teorii empirycznej T, zawierającej terminy nieteoretyczne $o_1, ..., o_m$ oraz teoretyczne $t_1, ..., t_n$. Składnik faktualny aksjomatów zostaje utożsamiony ze zdaniem egzystencjalnym:

$$\bigvee x_1, ..., \bigvee x_n \; F(o_1, ..., o_m, x_1, ..., x_n),$$

powstającym z powyższej koniunkcji przez podstawienie w miejsce terminów teoretycznych odpowiednich zmiennych i związanie ich kwantyfikatorami egzystencjalnymi. Składnik konwencjonalny aksjomatów może być wówczas reprezentowany przez zdanie warunkowe:

$$\bigvee x_1, ..., \bigvee x_n \; F(o_1, ..., o_m, x_1, ..., x_n) \rightarrow F(o_1, ..., o_m, t_1, ..., t_n).$$

Składnik faktualny tak zdefiniowany pociąga wszystkie nieteoretyczne konsekwencje aksjomatów i tym samym ujmuje ich zawartość empiryczną. Składnik konwencjonalny natomiast nie pociąga (poza prawdami logicznymi) żadnych twierdzeń nieteoretycznych; jego jedyną funkcją jest interpretacja terminów teoretycznych. Koniunkcja obu składników równoważna jest logicznie ogółowi aksjomatów. W tej sytuacji postulat znaczeniowy dla terminów teoretycznych teorii T może być utożsamiony ze składnikiem konwencjonalnym, a zdania analityczne języka teorii T — z konsekwencjami logicznymi tego składnika.

Rozwiązanie to przyjmuje postać prostszą i bardziej intuicyjną dla pewnych przypadków szczególnych. Należą do nich typowe i ważne w praktyce naukowej procedury definiowania terminów dyspozycyjnych za pomocą tzw. zdań redukcyjnych. W przypadku najprostszym aksjomaty charakteryzujące termin dyspozycyjny t za pomocą terminów obserwacyjnych o_1, o_2 przybierają formę następującą:

$$\bigwedge x\,(o_1(x) \rightarrow t(x)) \wedge \bigwedge x\,(o_2(x) \rightarrow \sim t(x)).$$

Składnik faktualny tych aksjomatów redukuje się do elementarnego twierdzenia empirycznego:

$$\bigwedge x\,(o_1(x) \rightarrow \sim o_2(x)),$$

a składnik konwencjonalny do odpowiedniego zdania warunkowego:

$$\bigwedge x\,(o_1(x) \rightarrow \sim o_2(x)) \rightarrow \bigwedge x\,(o_1(x) \rightarrow t(x)) \wedge \bigwedge x\,(o_2(x) \rightarrow \sim t(x))$$

Ten ostatni uznany być może za postulat znaczeniowy dla terminu t, a jego konsekwencje — za zdania analityczne języka zawierającego ów termin.

Przedstawiona tu próba wyodrębnienia z języka teorii empirycznej klasy zdań analitycznych narażona jest na zarzut pewnej arbitralności. Wydaje się to jednak nieuniknione wobec faktu notorycznej nieokreśloności semantycznej języków rzeczywistych teorii naukowych. Jednocześnie rozwiązanie to unika głównych obiekcji wysuwanych w stosunku do dychotomii zdań analitycznych i syntetycznych. W szczególności zdaje sprawę z dwoistego charakteru aksjomatów teorii empirycznej, łączących funkcje praw empirycznych i postulatów znaczeniowych, w czym upatrywano zasadniczą trudność wyróżnienia klasy zdań analitycznych. Aksjomaty teorii empirycznej zaliczone tu zostają do zdań syntetycznych. Zawierają one jednak komponent analityczny, który może zostać wyodrębniony — w sposób wskazany wyżej — z klasy ich logicznych konsekwencji.

Tak zdefiniowana klasa zdań analitycznych tworzy odrębną kategorię semantyczną i epistemologiczną. Jeśli nawet zgodzić się z tezą (tak mocno podkreśloną przez Quine'a), iż każde twierdzenie naukowe może być zrewidowane w konfrontacji z doświadczeniem, stwierdzić trzeba, że odrzucenie zdania syntetycznego stanowi jedynie rewizję systemu naszych przekonań, podczas gdy odrzucenie zdania analitycznego oznacza rewizję systemu naszych pojęć. Nie tylko inaczej opisujemy świat, ale i w innym czynimy to języku.

5. Zdanie, którego negacja jest zdaniem analitycznym, nazywane jest zdaniem kontradyktorycznym. Zgodnie z definicją analityczności jest to zdanie, którego negacja wynika logicznie ze zbiorów postulatów znaczeniowych danego języka. Można więc zdanie kontradyktoryczne określić jako zdanie fałszywe na mocy samego znaczenia terminów w nim zawartych. Zdanie, które nie jest ani zdaniem analitycznym, ani kontradyktorycznym, nazywane jest zdaniem syntetycznym. Miano zdań syntetycznych rezerwuje się niekiedy dla węższej klasy zdań, np. dla tych tylko zdań nieanalitycznych i niekontradyktorycznych, którym przysługiwać ma — różnie zresztą rozumiana — empiryczna sensowność.

BIBLIOGRAFIA

[1] Ajdukiewicz K., „Zagadnienie uzasadniania zdań analitycznych", [w:] Ajdukiewicz K., *Język i poznanie*, t. 2, Warszawa 1956.

[2] Carnap R., „Beobachtungssprache und theoretische Sprache", *Dialectica*, 1958, 12.

[3] Carnap R., „Meaning Postulates", *Philosophical Studies*, 1952, 3.

[4] Kemeny J. G., „Analyticity versus Fuzziness", *Synthese*, 1963, 15.

[5] Nowaczyk A., „Analityczność i aprioryzm", [w:] Pelc J. (red.), *Semiotyka polska 1894–1969*, Warszawa 1971.

[6] Przełęcki M., Wójcicki R., „The Problem of Analyticity", *Synthese*, 1968–1969. 19.

[7] Quine W v. O., „Dwa dogmaty empiryzmu", [w:] Quine W. v. O., *Z punktu widzenia logiki. Eseje logiczno-filozoficzne*, Warszawa 1969.

ONTOLOGIA

Filozofia Nauki
Rok I, 1993, Nr 2-3

W sprawie istnienia przedmiotów teoretycznych

Przedmiotami teoretycznymi nazywa się odpowiedniki ontologiczne terminów teoretycznych. Terminy teoretyczne mogą być jednak pojmowane szerzej lub węziej, a przez ich odpowiedniki ontologiczne można rozumieć bądź ich denotacje, bądź desygnaty. Do terminów teoretycznych zalicza się bowiem zarówno terminy odnoszące się (wyłącznie lub między innymi) do pewnych przedmiotów postrzegalnych i przypisujące im pewne własności teoretyczne — przykładem może być termin „magnes" — jak i terminy odnoszące się wyłącznie do pewnych przedmiotów niepostrzegalnych, a więc terminy takie, jak: „elektron", „molekuła", czy „gen". Tylko te ostatnie będę w zasadzie nazywał tutaj terminami teoretycznymi. Ograniczę się przy tym do terminów o charakterze predykatów, pomijając dla uproszczenia tak ważny rodzaj terminów teoretycznych, jak terminy funkcyjne. Przedmiotami teoretycznymi będę nazywał z kolei nie denotacje, lecz desygnaty owych predykatów. A więc nie klasę elektronów (czy też własność bycia elektronem), lecz poszczególne elektrony. Ten rodzaj terminów teoretycznych nasuwa swoiste problemy metodologiczne. Powstaje pytanie, jaką rolę odgrywają takie terminy w języku teorii empirycznej. W szczególności, czy odpowiednie teorie naukowe pociągają twierdzenia o istnieniu tak rozumianych przedmiotów teoretycznych? Jeśli tak, to stwierdzenie ich istnienia jest uzasadnione w stopniu co najmniej tym samym, co owe teorie. Wydawać by się mogło, że praktyka naukowa narzuca na to pytanie odpowiedź twierdzącą. Istnieją elektrony, bo tak głoszą dobrze potwierdzone teorie współczesnej fizyki. Czy jednak odpowiedź taka istotnie nie może podlegać dyskusji? Analiza metodologiczna teorii naukowej musi, co prawda, wychodzić od tego, co dana teoria głosi, lecz na tym nie może poprzestać. Jej zadaniem jest interpretacja zastanych sformułowań mająca na celu sprecyzowanie pod względem logicznym ich sensu. Otóż taka rekonstrukcja logiczna empirycznych teorii naukowych doprowadza niektórych do zakwestionowania owego obiegowego poglądu. Przykładem takiego stanowiska może być pogląd Mehlberga przedstawiony w jego znanej książce *The Reach of Science*. Pogląd, wedle którego żadna teoria empiryczna nie może zawierać twierdzeń o istnieniu przedmiotów teoretycznych, gdyż jej język nie pozwala w ogóle na ich sformułowanie. Sposoby wprowadzania do języka teorii terminów wyposażonych w sens empiryczny nie pozwalają, zdaniem autora, na wprowadzenie jakiegokolwiek predykatu desygnującego przedmioty niepostrzegalne. Ten prowokujący pogląd wymaga bliższej analizy. Obierzemy go za punkt wyjścia dla naszych rozważań, przeciwstawiając mu w dalszym ciągu stanowisko odmienne, dopuszczające w języku teorii empirycznej predykaty desygnujące przedmioty niepostrzegalne, a tym samym — twierdzenia głoszące istnienie takich przedmiotów. Rozpatrzymy w tym celu pewne rodzaje postulatów służących do wprowadzania terminów teoretycznych, starając się prześledzić niektóre konsekwencje płynące z użycia określonego typu postulatów. Chciałbym przy tym problemy te ująć w sposób nieco ściślejszy, umożliwiający zastosowanie aparatu pojęciowego współczesnej semantyki logicznej. Referat niniejszy nawiązuje pod tym względem do mojej pracy pt. „Z semantyki pojęć otwartych" [1] i odwołuje się do pewnych jej wyników.

1. Spróbujemy przede wszystkim scharakteryzować krótko z interesującego nas punktu widzenia język teorii empirycznej. Będzie to oczywiście daleko idące uproszczenie i idealizacja faktycznego stanu rzeczy — niezbędne jednakże, jak sądzę, dla uchwytnego przedstawienia wybranych zagadnień. Ograniczymy się do języka możliwie najprostszego — opartego o węższy rachunek predykatów (z identycznością) i zawierającego jako jedyne terminy pozalogiczne predykaty k-argumentowe[2]. Założymy, iż ogół przedmiotów rozpatrywanych w naszym języku (czyli będących wartościami zmiennych) — jego universum U — stanowi zbiór wszelkich przedmiotów fizycznych (materialnych), nie precyzując jednak bliżej tego niejasnego, jak wiadomo, pojęcia. W zbiorze tym wyróżnimy pewien podzbiór właściwy, U_1, obejmujący wszelkie przedmioty postrzegalne. Jednocześnie wśród ogółu predykatów naszego języka wyróżnimy pewien ich rodzaj: predykaty postrzeżeniowe $O_1, ..., O_l$. Język J zawierający wyłącznie predykaty: $O_1, ..., O_l$, nazwiemy językiem postrzeżeniowym. Zarówno pojęcie postrzegalnego przedmiotu, jak i postrzeżeniowego predykatu wymagają pewnych wyjaśnień w celu uniknięcia możliwych nieporozumień.

Mówiąc o postrzegalności pewnego przedmiotu mam na myśli możliwość jego bezpośredniego postrzeżenia zagwarantowaną przez prawa przyrody. Dany przedmiot jest postrzegalny, jeśli przysługuje mu pewna własność tego rodzaju, iż do praw przyrody należy twierdzenie głoszące, że ktokolwiek (w odpowiednich warunkach) spojrzy na

1) *Studia Logica*, XV, 1964.
2) W dalszych rozważaniach odwoływać się będziemy niekiedy do wyrażeń zawierających nazwy indywiduowe. Będziemy to jednak czynili głównie dla celów ilustracyjnych, tak że w zasadzie całość tych wywodów ograniczona być może do języka predykatów.

jakiś przedmiot o tej własności, ten go spostrzeże. To ogólnikowe wyjaśnienie nie ma być, oczywiście, definicją postrzegalności. Ma ono jedynie zwrócić uwagę na fakt, że do tak rozumianych przedmiotów postrzegalnych należeć będą pewne spośród przedmiotów tak odległych w czasie czy przestrzeni, iż nikt ich nigdy nie postrzegał ani nie postrzeże, gdy tymczasem inne spośród nich postrzegalnymi nie będą. Przedmiotem postrzegalnym będzie dinozaur, niepostrzegalnym — gen zawarty w jego organizmie. Tak pojęty podział przedmiotów fizycznych na postrzegalne i niepostrzegalne utożsamić można z pewnym przybliżeniem z podziałem na makro- i mikroobiekty. Przedmioty postrzegalne — to przedmioty dostatecznie duże, a niepostrzegalne — zbyt małe na to, aby można je było dostrzec[3]. Jest to oczywiście podział bardzo nieostry. Ale to, w którym dokładnie miejscu przeprowadzi się tutaj granicę, nie ma dla naszych rozważań większego znaczenia. Ważne jest przede wszystkim to, że pewne przedmioty należeć będą w każdym razie do przedmiotów postrzegalnych, a pewne inne — do niepostrzegalnych. Do tych ostatnich zaliczymy niewątpliwie takie obiekty, jak cząstki elementarne, atomy czy molekuły.

Predykaty postrzeżeniowe określa się często przez odwołanie się do pojęcia własności (czy stosunków) „postrzegalnych". Chcąc uniknąć tego dość zagadkowego pojęcia, odwołam się tutaj raczej — tak jak to uczyniłem m.in. w pracy „O definiowaniu terminów spostrzeżeniowych" [4] — do sposobu interpretacji tego rodzaju predykatów, określając je jako terminy zinterpretowane bezpośrednio. Interpretację taką stanowić może tzw. definicja ostensywna, ustalająca denotację danego predykatu przez wskazanie pewnych przedmiotów jako jej elementów, a pewnych innych — jako elementów jej dopełnienia. Procedurę tę analizowałem szczegółowo we wspomnianej pracy. Tutaj zwrócę tylko uwagę na te sprawy, które będą miały decydujące znaczenie dla dalszych rozważań.

Niech takim ostensywnie definiowanym predykatem będzie 1-argumentowy predykat O_1. Przedmioty wskazywane jako jego desygnaty (lub desygnaty jego negacji) muszą na pewno należeć do przedmiotów postrzegalnych. Gdyby denotacja predykatu O_1 była ustalona tylko w stosunku do przedmiotów wskazanych, pozostawałaby tym samym całkowicie nieustalona w stosunku do wszelkich przedmiotów niepostrzegalnych. Przyjmuje się jednak zwykle, iż przedmioty wskazywane pełnią tylko rolę wzorców. Przez ich wskazanie ustala się denotację predykatu O_1 również w stosunku do pewnych przedmiotów innych. Procedura ta ma pod względem logicznym charakter

3) Powstać może wątpliwość, czy wszelkie przedmioty makroskopowe uważać możemy za postrzegalne. Potoczne pojęcie (makroskopowego) przedmiotu fizycznego wydaje się jednak zbyt nieokreślone, aby pozwalało rozstrzygnąć tę kwestię. Czy dwie molekuły gazu oddalone od siebie o jeden metr tworzą jeden (makroskopowy) przedmiot fizyczny? Jeśli tak, byłby to przykład makroskopowego przedmiotu niepostrzegalnego. Rozstrzygnięcie powyższej wątpliwości wydaje się możliwe tylko pod warunkiem sprecyzowania owego pojęcia potocznego.

4) *Rozprawy Logiczne*, Warszawa 1964.

dość zagadkowy[5]. Wydaje się jednak, iż jakkolwiek by się próbowało zdać z niej sprawę, można przyjąć założenie, iż w ten sposób nie da się ustalić denotacji predykatu O_1 w stosunku do żadnego przedmiotu niepostrzegalnego. Denotacja predykatu zdefiniowanego wyłącznie ostensywnie — a więc bez użycia jakiegokolwiek innego predykatu pozalogicznego — pozostaje całkowicie nieustalona w zbiorze przedmiotów niepostrzegalnych. Założenie to wydaje się dość przekonywające. Jedynym kryterium przynależności jakiegoś przedmiotu do denotacji takiego predykatu jest jego podobieństwo pod względem wyglądu do któregoś z przedmiotów wskazanych. Trudno zaś mówić w przypadku przedmiotu niepostrzegalnego, iż wygląda on tak, jak któryś z przedmiotów wzorcowych (pozytywnych, czy negatywnych), skoro przedmiotu takiego w ogóle wyobrazić sobie nie można (można sobie wyobrazić co najwyżej pewien przedmiot inny, odpowiednio większy od danego). Można oczywiście kwestionować powyższe założenie — przy odmiennych nieco rozumieniach predykatu postrzeżeniowego czy definicji ostensywnej, dopuszczających np. posługiwanie się innymi predykatami o już ustalonej interpretacji. Wydaje się jednak, iż rozumienie nasze stanowi jedną z dopuszczalnych eksplikacji pojęcia predykatu postrzeżeniowego i w tym też rozumieniu będziemy się tym pojęciem konsekwentnie posługiwali.

Jak zatem przedstawia się w rezultacie interpretacja predykatu takiego jak O_1? Określony podzbiór A_1 zbioru przedmiotów postrzegalnych U_1 zawierać się ma w jego denotacji, a inny taki podzbiór B_1 — w jej dopełnieniu. Denotacja ta jest zatem scharakteryzowana jednoznacznie tylko w zbiorze $A_1 \cup B_1$, zawartym w U_1. Gdyby zbiór $A_1 \cup B_1$ był identyczny z U_1, denotacja ta byłaby wyznaczona jednoznacznie w zbiorze wszelkich przedmiotów postrzegalnych. Tak też się niekiedy dla uproszczenia przyjmuje, chociaż na ogół nie odpowiada to rzeczywistości. Z reguły zbiór $A_1 \cup B_1$ jest różny od U_1 i denotacja O_1 jest scharakteryzowana wieloznacznie nawet w zbiorze przedmiotów postrzegalnych. Na tym polega notoryczna nieostrość predykatów postrzeżeniowych. W każdym zaś razie denotacja ta jest scharakteryzowana wieloznacznie w zbiorze wszelkich przedmiotów fizycznych U, przyjętym przez nas za universum języka J. W zawartym w U zbiorze wszelkich przedmiotów niepostrzegalnych denotacja O_1 jest całkowicie nieustalona. Nie gwałcąc interpretacji predykatu O_1 możemy dowolny przedmiot z tego zbioru zaliczyć do jego denotacji lub go z niej wyłączyć. Nie dysponujemy żadnymi kryteriami orzekania predykatu O_1 lub jego negacji o jakimkolwiek przedmiocie niepostrzegalnym. Zdania, które takie orzeczenia stwierdzają, są zasadniczo nierozstrzygalne. Taki sam charakter przysługuje pozostałym predykatom postrzeżeniowym. Gdy O_i jest predykatem k-argumentowym, zdanie, które orzeka ten predykat lub jego negację o k przedmiotach, z których choćby jeden jest przedmiotem niepostrzegalnym, jest zdaniem zasadniczo nierozstrzygalnym.

5) Por. pracę cytowaną.

Istotne z logicznego punktu widzenia własności języka J wyrazić można ściślej w terminologii współczesnej semantyki logicznej. Modelami języka J o predykatach: O_1, ..., O_l są wszelkie dziedziny typu:

$$\langle U; X_1, ..., X_l \rangle,$$

gdzie X_i jest relacją o tylu członach, ile argumentów ma predykat O_i. Interpretacja bezpośrednia predykatów języka J sprowadza się do wyznaczenia określonej rodziny $\mathbf{M}_1$ modeli języka J. Universum każdego z nich stanowi zbiór U_1. Jest to z reguły rodzina zawierająca więcej niż jeden model, w czym znajduje wyraz nieostrość predykatów O_1, ..., O_l. Rodzinę $\mathbf{M}$ modeli języka J dostarczającą właściwej (zamierzonej) interpretacji tego języka określamy, jak następuje: $\mathfrak{M}$ należy do $\mathbf{M}$ wtedy i tylko wtedy, gdy $\mathfrak{M}$ jest modelem, którego universum stanowi zbiór U, i który jest rozszerzeniem pewnego modelu rodziny $\mathbf{M}_1$[6].

Właściwym modelem języka J może być więc dowolne rozszerzenie do universum U jakiegokolwiek modelu należącego do rodziny $\mathbf{M}_1$. Interpretacja języka J jest zawsze wieloznaczna — dana nie przez określony model, lecz przez obszerną ich rodzinę. W konsekwencji wyróżnić możemy dwa rodzaje zdań języka J:

(1) zdania, które we wszystkich modelach rodziny $\mathbf{M}$ mają tę samą wartość logiczną, a więc we wszystkich są prawdziwe lub we wszystkich fałszywe — nazwijmy je zdaniami rozstrzygalnymi; oraz

(2) zdania, które w pewnych modelach rodziny $\mathbf{M}$ są prawdziwe, a w innych fałszywe — zwane w dalszym ciągu zdaniami nierozstrzygalnymi[7].

Do zdań rozstrzygalnych należą oczywiście wszelkie tautologie logiczne i ich negacje. Należeć też będą do nich m.in. zdania orzekające predykat O_1 o pewnych przedmiotach ze zbioru U_1. Natomiast każde zdanie orzekające predykat O_1 o jakimkolwiek przedmiocie nie należącym do zbioru U_1 będzie zdaniem nierozstrzygalnym. Znajdziemy bowiem zawsze pośród modeli rodziny $\mathbf{M}$ takie, w których ów przedmiot będzie należał do denotacji O_1 i takie, w których do niej należał nie będzie — jak to z określenia rodziny $\mathbf{M}$ wyraźnie wynika.

Język J stanowi część postrzeżeniową języka teorii empirycznej. Ten ostatni obejmuje oprócz predykatów postrzeżeniowych: O_1, ..., O_l, pewne predykaty niepostrzeżeniowe, a więc — w szerokim tego słowa znaczeniu — teoretyczne: T_1, ..., T_m. Przyjmujemy w stosunku do omawianego języka następujące założenie: jedynymi terminami tego języka interpretowanymi bezpośrednio są predykaty postrzeżeniowe. A zatem wszelkie predykaty pozostałe mają wyłącznie interpretację pośrednią. Interpretacja ta polega na przyjęciu dla predykatów teoretycznych pewnego układu postulatów i na scharakteryzowaniu denotacji tych predykatów jako relacji, które czynią zadość

6) Model $\mathfrak{M}' = \langle U'; X_1', ..., X_l' \rangle$ jest rozszerzeniem (nadmodelem) modelu $\mathfrak{M} = \langle U; X_1, ..., X_l \rangle$, gdy $U \subset U'$ oraz $\bigwedge_{x_1,...,x_k \in U} (X_i'(x_1, ..., x_k) \equiv X_i(x_1, ..., x_k))$ dla każdego $i = 1, ..., l$.

7) Pojęcia te rozważam szczegółowo w pracy „Z semantyki pojęć otwartych".

warunkom nałożonym na nie przez owe postulaty — przy właściwej dla języka J interpretacji predykatów postrzeżeniowych.

Przedstawmy tę sytuację w sposób nieco ściślejszy dla przypadku najprostszego: języka J', który powstaje z języka J przez dołączenie do predykatów: $O_1, ..., O_l$ predykatu T_1. Modelami języka J' są wszelkie dziedziny typu:

$$\langle U; X_1, ..., X_l; Y_1\rangle .$$

Określmy przede wszystkim rodzinę $\mathbf{M}'$ modeli języka J' obejmującą wszystkie wzbogacenia modeli języka J należących do rodziny $\mathbf{M}$. A zatem: $\mathfrak{M}'$ należy do $\mathbf{M}'$ wtedy i tylko wtedy, gdy $\mathfrak{M}'$ jest modelem stanowiącym wzbogacenie pewnego modelu rodziny $\mathbf{M}$[8]. Gdyby na denotację predykatu T_1 nie były nałożone żadne warunki, rodzina $\mathbf{M}'$ reprezentowałaby właściwą interpretację języka J'. Zakładamy jednak, że predykat T_1 wprowadzony został przy pomocy pewnego postulatu A_1. Przy tym założeniu właściwej interpretacji języka J' dostarcza nam pewna podrodzina rodziny $\mathbf{M}'$, obejmująca spośród niej tylko te modele, w których spełniony jest postulat A_1. Nazwiemy ją rodziną $\mathbf{M}'(A_1)$ i zdefiniujemy, jak następuje: $\mathfrak{M}'$ należy do $\mathbf{M}'(A_1)$ wtedy i tylko wtedy, gdy $\mathfrak{M}'$ należy do $\mathbf{M}'$ i A_1 jest prawdziwe w $\mathfrak{M}'$. Rodzina $\mathbf{M}'(A_1)$ obejmuje z reguły więcej niż jeden model. W każdym razie tylko takie przypadki rozpatrywać będziemy w dalszym ciągu, ograniczając tym samym rodzaj postulatów A_1 (np. do postulatów nietwórczych). Interpretacja języka J' jest więc zawsze wieloznaczna. Wobec tego i w stosunku do języka J' znajdują zastosowanie pojęcia zdania rozstrzygalnego i nierozstrzygalnego.

(A) Z jest zdaniem rozstrzygalnym wtedy i tylko wtedy, gdy Z jest prawdziwe w każdym modelu rodziny $\mathbf{M}'(A_1)$ lub Z jest fałszywe w każdym modelu tej rodziny.

A zatem:

(A') Z jest zdaniem nierozstrzygalnym wtedy i tylko wtedy, gdy Z jest prawdziwe w pewnym modelu rodziny $\mathbf{M}'(A_1)$, a fałszywe w innym modelu tej rodziny.

2. Opierając się na powyższej rekonstrukcji języka teorii empirycznej spróbujmy przedstawić krótko wspomniane na wstępie stanowisko Mehlberga w sprawie terminów teoretycznych. Stanowisko to związane jest z poglądem autora na rodzaj postulatów nadających się do wprowadzania terminów do języka teorii empirycznej. Postulaty te muszą mieć wedle niego postać tzw. definicji cząstkowych — bądź dla terminu wprowadzanego, bądź dla jego negacji:

(1) $$\bigwedge_x (\Phi x \rightarrow T_1 x) \quad \text{lub}$$

(1′) $$\bigwedge_x (\Psi x \rightarrow {\sim} T_1 x)\,,$$

przy czym formuły: Φx i Ψx zawierać muszą jako terminy pozalogiczne wyłącznie predykaty postrzeżeniowe: $O_1, ..., O_l$, a co najmniej jedna z nich musi być formułą spełnioną przez pewien przedmiot. Wtedy i tylko wtedy, predykat T_1 będzie terminem

8) Model $\mathfrak{M}' = \langle U'; X_1', ..., X_l'; Y_1'\rangle$ jest wzbogaceniem modelu $\mathfrak{M} = \langle U; X_1, ..., X_l\rangle$, gdy $U' = U$ oraz $X_i' = X_i$ dla każdego $i = 1, ..., l$.

empirycznie sensownym — zgodnie z reprezentowaną przez Mehlberga rygorystyczną koncepcją empirycznego sensu. Otóż Mehlberg stara się okazać, iż w ten sposób nie można wprowadzić żadnego predykatu, który by desygnował (nie tylko — wyłącznie, ale nawet — między innymi) pewne przedmioty niepostrzegalne. Predykat T_1 desygnuje bowiem tylko te przedmioty, które desygnuje występujący w postulacie (1) predykat Φ. A skoro ten ostatni jest predykatem postrzeżeniowym, jego desygnatami są wyłącznie przedmioty postrzegalne.

Wywód ten zawiera, po pierwsze, pewne luki. Wyrażenie Φ nie musi być przecież jednym z predykatów postrzeżeniowych: $O_1, ..., O_l$. Może być dowolnym wyrażeniem złożonym z tych predykatów. Po drugie, wywód ten opiera się na swoistej koncepcji desygnowania, reprezentowanej przez Mehlberga we wspomnianej książce. Jest to koncepcja wyraźnie kontrowersyjna, odbiegająca od innych, tradycyjnych koncepcji desygnowania. Wszystkie te koncepcje omawiam w pracy „Z semantyki pojęć otwartych". Tutaj ograniczyć się muszę do paru słów wyjaśnienia. Niechaj P będzie dowolnym predykatem opisanego języka J', którego interpretację stanowi rodzina modeli $\mathbf{M}'(A_1)$. Wedle Mehlberga, predykat P desygnuje tylko te przedmioty, które spełniają Px w każdym modelu rodziny $\mathbf{M}'(A_1)$. Wobec tego istotnie żaden predykat postrzeżeniowy: $O_1, ..., O_l$ nie może desygnować jakiegokolwiek przedmiotu niepostrzegalnego; nie może też tego uczynić predykat T_1 wprowadzony za pomocą postulatu A_1:

(1*) $$\bigwedge_x (O_i x \rightarrow T_1 x) \text{ dla pewnego } i = 1, ..., l.$$

Ale inne koncepcje desygnowania ujmują to pojęcie mniej rygorystycznie. Predykat P nie może desygnować tylko tych przedmiotów, które nie spełniają Px w żadnym modelu rodziny $\mathbf{M}'(A_1)$. Jeśli jakiś przedmiot spełnia Px w pewnym modelu tej rodziny, może być desygnatem P — przy czym wedle niektórych koncepcji przedmiot taki jest istotnie desygnatem P, wedle innych — to, czy nim jest, czy nie, pozostaje sprawą nierozstrzygalną. Otóż przy takim ujęciu desygnowania wywód powyższy traci walor. Zarówno predykaty postrzeżeniowe, jak i predykat taki jak T_1, mogą być uważane za predykaty desygnujące pewne przedmioty niepostrzegalne.

Wydaje się jednak, iż nawet przy takich założeniach tkwi w przedstawionym poglądzie coś słusznego. Jeśli predykat T_1 ma desygnować wyłącznie przedmioty niepostrzegalne, nie może być wprowadzony do języka postrzeżeniowego J nie tylko za pomocą postulatu takiego jak (1*), ale i ogólniejszego odeń postulatu (1) Predykat taki byłby terminem naukowo nieprzydatnym. Spróbujmy obecnie zdać sprawę z tego, na czym owa nieprzydatność ma polegać.

Zakładamy więc, iż postulat A_1 dla predykatu T_1 ma postać:

(1) $$\bigwedge_x (\Phi x \rightarrow T_1 x),$$

gdzie Φx jest formułą o jednej zmiennej wolnej x należącą do języka J, a więc zawierającą wyłącznie predykaty: $O_1, ..., O_l$. Zakładamy również, iż T_1 ma być predykatem desygnującym wyłącznie przedmioty niepostrzegalne, a więc nie należące do zbioru U_1. Wobec tego predykat Φ nie może desygnować żadnych przedmiotów postrzegalnych, tj. należących do U_1. Interpretację właściwą języka J stanowi, jak wiemy, rodzina modeli **M**. A zatem warunek, jakiemu czynić ma zadość formuła Φx, sformułować możemy, jak następuje: dla dowolnego modelu $\mathfrak{M}$ należącego do rodziny **M** jest tak, iż żaden przedmiot należący do zbioru U_1 nie spełnia Φx w $\mathfrak{M}$. Mówiąc swobodniej: nie ma takiej dopuszczalnej interpretacji języka J, przy której pewien przedmiot postrzegalny spełniałby formułę Φx. Ale warunek powyższy pociąga logicznie — w sposób dość oczywisty[9] — następującą konsekwencję: istnieje taki model $\mathfrak{M}$ należący do rodziny **M**, iż żaden przedmiot należący do zbioru U nie spełnia Φx w $\mathfrak{M}$. Inaczej mówiąc: istnieje taka dopuszczalna interpretacja języka J, przy której żaden przedmiot (a więc nie tylko postrzegalny, ale i niepostrzegalny) nie spełnia formuły Φx.

Jakie konsekwencje wynikają stąd dla interpretacji predykatu T_1 wprowadzonego przez postulat (1)? Odpowiadając na to pytanie postaram się wskazać jedną z takich konsekwencji, odwołującą się w swoim sformułowaniu do pojęć, które wydają mi się interesujące z ogólnoteoretycznego punktu widzenia i przydatne w różnych analizach logicznych języka teorii empirycznej. Będą to pojęcia istotnego i nieistotnego występowania terminu T_1 w zdaniach języka J'. Mówimy zazwyczaj, iż zdanie Z zawiera termin T_1 w sposób nieistotny, gdy jest równoważne logicznie pewnemu zdaniu Z^* nie zawierającemu terminu T_1. Gdy tak nie jest, Z zawiera T_1 w sposób istotny. Można jednak oprócz tego pojęcia «absolutnego» wprowadzić pojęcie zrelatywizowane do określonej interpretacji pozostałych terminów pozalogicznych występujących obok T_1 w zdaniu Z. W przypadku rozważanego przez nas języka J' interpretacja owych terminów pozostałych wyznaczona jest przez rodzinę modeli **M**. Toteż pojęcie to w zastosowaniu do języka J' zdefiniować można w sposób następujący:

(*B*) Z zawiera T_1 w sposób istotny wtedy i tylko wtedy, gdy dla dowolnego modelu $\mathfrak{M}$ należącego do rodziny **M** istnieją takie modele $\mathfrak{M}'_1$ i $\mathfrak{M}'_2$ należące do rodziny **M′** i stanowiące wzbogacenia modelu $\mathfrak{M}$, iż Z jest prawdziwe w $\mathfrak{M}'_1$, a fałszywe w $\mathfrak{M}'_2$.

A zatem:

(*B′*) Z zawiera T_1 w sposób nieistotny wtedy i tylko wtedy, gdy istnieje taki model $\mathfrak{M}$ należący do rodziny **M**, iż dla dowolnych modeli $\mathfrak{M}'_1$ i $\mathfrak{M}'_2$ należących do rodziny **M′** i stanowiących wzbogacenia modelu $\mathfrak{M}$ zachodzi zależność następująca: Z jest prawdziwe w $\mathfrak{M}'_1$ wtedy i tylko wtedy, gdy Z jest prawdziwe w $\mathfrak{M}'_2$.

A więc, jeśli predykat T_1 występuje w zdaniu Z w sposób nieistotny, można zawsze znaleźć taką dopuszczalną interpretację predykatów pozostałych, iż to, jak zinterpretujemy predykat T_1, nie będzie miało żadnego wpływu na wartość logiczną zdania Z; jeśli zdanie to jest prawdziwe (resp. fałszywe) przy pewnej interpretacji T_1, pozostaje prawdziwe (resp. fałszywe) przy dowolnej innej interpretacji tego predykatu.

9) Dowód tej ogólnej zależności przeprowadzam w innej, przygotowanej do druku pracy.

Należy tutaj zwrócić uwagę na fakt, iż wprowadzone obecnie pojęcie (*B*) nie jest jedynym możliwym pojęciem istotnego występowania terminu T_1 w zdaniach języka J' zrelatywizowanym do interpretacji terminów pozostałych. Wspomniane pojęcie «absolutne» prowadzi wprost do nieco innego, słabszego niż (*B*), pojęcia relatywnego.

(*C*) *Z* zawiera T_1 w sposób istotny wtedy i tylko wtedy, gdy istnieje model $\mathfrak{M}$ należący do rodziny **M** oraz modele $\mathfrak{M}'_1$ i $\mathfrak{M}'_2$ należące do rodziny **M′** i stanowiące wzbogacenia modelu $\mathfrak{M}$ takie, iż *Z* jest prawdziwe w $\mathfrak{M}'_1$, a fałszywe w $\mathfrak{M}'_2$.

Odpowiednio:

(*C′*) *Z* zawiera T_1 w sposób nieistotny wtedy i tylko wtedy, gdy dla dowolnego modelu $\mathfrak{M}$ należącego do rodziny **M** oraz dla dowolnych modeli $\mathfrak{M}'_1$ i $\mathfrak{M}'_2$ należących do rodziny **M′** i stanowiących wzbogacenia modelu $\mathfrak{M}$ zachodzi zależność następująca: *Z* jest prawdziwe w $\mathfrak{M}'_1$ wtedy i tylko wtedy, gdy *Z* jest prawdziwe w $\mathfrak{M}'_2$.

Zgodnie z tym ostatnim określeniem, jeśli predykat T_1 występuje w zdaniu *Z* w sposób nieistotny, to przy każdej dopuszczalnej interpretacji predykatów pozostałych jest tak, iż sposób interpretacji predykatu T_1 nie ma żadnego wpływu na wartość logiczną zdania Z. W przypadku skrajnym, gdy rodzina **M** obejmuje tylko jeden model, a więc gdy interpretacja języka *J* jest jednoznaczna, pojęcia (*B*) i (*C*) pokrywają się ze sobą. Otrzymujemy w ten sposób pojęcie istotnego występowania terminu T_1 w zdaniu *Z* zrelatywizowane do określonego modelu $\mathfrak{M}_0$, wyznaczającego jednoznacznie interpretację pozostałych terminów pozalogicznych występujących w zdaniu Z.

Z wprowadzonych obecnie pojęć, pojęcie (*B*) wydaje się lepiej odpowiadać naszym potrzebom i nim też posługiwać się będziemy w dalszym ciągu rozważań. Zajęliśmy poprzednio w sprawie desygnowania stanowisko możliwie najmniej restryktywne. W rezultacie wszelkie modele rodziny **M** traktować musimy jako dopuszczalne interpretacje języka *J*. Zgodnie z pewną koncepcją desygnowania, jeden z tych modeli stanowi właściwy model języka *J*; przy przyjętej charakterystyce tego języka pozostaje jednak sprawą nierozstrzygalną, o który to model chodzi. Jeśli chcemy więc stwierdzić z pewnością, iż wartość logiczna pewnego zdania *Z* zależy od takiego czy innego sposobu interpretacji terminu T_1, musi to być, jak widać, zdanie, które ów termin zawiera w sposób istotny w sensie (*B*). W związku z tym pozostaje też następująca własność tego pojęcia. Niechaj **M*** będzie dowolną rodziną modeli języka *J* zawartą w rodzinie **M**. Jeśli zdanie *Z* zawiera T_1 w sposób istotny ze względu na **M**, zdanie to zawiera T_1 w sposób istotny również ze względu na **M***. Natomiast w przypadku pojęcia (*C*) otrzymujemy zależność odwrotną: jeśli zdanie *Z* zawiera T_1 w sposób nieistotny ze względu na **M**, zdanie to zawiera T_1 w sposób nieistotny również ze względu na **M***. Rodzinę **M*** traktować możemy jako rezultat dalszej precyzacji terminów języka *J*. Ich interpretacja scharakteryzowana zostaje bardziej jednoznacznie. Otóż żadne zdanie, które zawierało termin T_1 w sposób istotny w sensie (*B*), nie może stać się w wyniku takiej precyzacji zdaniem, które by zawierało ów termin w sposób nieistotny. Mamy tu do czynienia z własnością zdań trwałą w procesie precyzacji języka. Inaczej jest w przypadku pojęcia (*C*). Klasa zdań zawierających T_1 w sposób istotny w sensie (*C*) zwęża się — a nie rozszerza — w miarę precyzacji terminów języka *J*. Mówiąc w dalszym ciągu o istotnym czy nieistotnym zawieraniu się predykatu T_1 w zdaniach języka *J'*, będziemy mieli na myśli wyłącznie pojęcie (*B*).

Wracając do głównego toku naszych rozważań, spróbujemy odpowiedzieć na postawione przedtem pytanie: na czym polega nieprzydatność terminu takiego jak T_1, tj. wprowadzonego do języka *J* za pomocą postulatu A_1 postaci (1) i nie desygnującego żadnych przedmiotów postrzegalnych. Jedną z możliwych odpowiedzi fomułuje następujące twierdzenie:

Żadne zdanie języka *J'* zawierające T_1 w sposób istotny nie jest rozstrzygalne.

Twierdzenie to jest prostą konsekwencją definicji występujących w nim pojęć (*A*) i (*B*). Załóżmy, iż *Z* zawiera T_1 w sposób istotny: dla każdego $\mathfrak{M}$ należącego do $\mathbf{M}$ istnieją $\mathfrak{M}'_1$ i $\mathfrak{M}'_2$ należące do $\mathbf{M}'$ i stanowiące wzbogacenia $\mathfrak{M}$ takie iż *Z* jest prawdziwe w $\mathfrak{M}'_1$, a fałszywe w $\mathfrak{M}'_2$. Wiemy ponadto, iż istnieje $\mathfrak{M}$ należący do $\mathbf{M}$ taki, iż żaden przedmiot nie spełnia Φx w $\mathfrak{M}$. Niechaj modelem tym będzie $\mathfrak{M}_0$. Istnieją zatem $\mathfrak{M}'_1$ i $\mathfrak{M}'_2$ należące do $\mathbf{M}'$ i stanowiące wzbogacenia $\mathfrak{M}_0$ takie, iż *Z* jest prawdziwe w $\mathfrak{M}'_1$, a fałszywe w $\mathfrak{M}'_2$. Ale skoro w $\mathfrak{M}_0$ żaden przedmiot nie spełnia Φx — poprzednika postulatu A_1 o postaci (1), postulat A_1 jest prawdziwy przy dowolnej interpretacji T_1, czyli w każdym wzbogaceniu modelu $\mathfrak{M}_0$. A więc również w $\mathfrak{M}'_1$ i w $\mathfrak{M}'_2$. Istnieją, co za tym idzie, modele $\mathfrak{M}'_1$ i $\mathfrak{M}'_2$ należące do rodziny $\mathbf{M}'(A_1)$ takie, iż *Z* jest prawdziwe w $\mathfrak{M}'_1$, a fałszywe w $\mathfrak{M}'_2$. Tym samym, *Z* jest zdaniem nierozstrzygalnym.

Wyrażona w powyższym twierdzeniu właściwość terminu T_1 dyskredytuje ten termin jako składnik języka nauki. Zdania, które w badaniu naukowym odgrywają jakąś rolę, muszą należeć do zdań rozstrzygalnych. Tylko takie zdania mogą zostać uzasadnione lub obalone. Nie można okazać, że zdanie Z jest prawdziwe (*resp.* fałszywe) — a więc nie można go uzasadnić (*resp.* obalić) — jeśli zdanie to można w danym języku rozumieć zarówno tak, iż będzie prawdą, jak i tak, iż będzie fałszem. A taki charakter mają właśnie wszelkie zdania nierozstrzygalne. Ale w żadnym zdaniu rozstrzygalnym termin T_1 nie może wystąpić w sposób istotny. W każdym takim zdaniu można pozostałe terminy rozumieć w danym języku tak, iż sposób interpretacji terminu T_1 będzie najzupełniej obojętny. Wartość logiczna takiego zdania będzie zdeterminowana całkowicie przez interpretację terminów pozostałych. To, czy przez T_1 będziemy rozumieli zbiór pusty, czy całe universum, czy też jakikolwiek jego podzbiór — nie będzie miało żadnego wpływu na prawdziwość (resp. fałszywość) takiego zdania. Termin T_1 o tego rodzaju własnościach jest w języku nauki po prostu zbędny.

Uzasadniona wydaje się przeto konkluzja, iż języka postrzeżeniowego *J* nie można wzbogacić o żaden termin T_1, desygnujący wyłącznie przedmioty niepostrzegalne za pomocą postulatów typu (1). Nie można zatem wprowadzić w ten sposób terminów teoretycznych takich, jak: „elektron", „molekuła" czy „gen". Oprócz postulatów typu (1) stoją oczywiście do naszej dyspozycji — nawet przy ograniczeniu się do definicji cząstkowych — postulaty typu (1′):

(1′) $$\bigwedge_x (\Psi x \rightarrow \sim T_1 x)\,.$$

Postulaty te — w przeciwieństwie do poprzednich — mogą być wykorzystane przy wprowadzaniu terminów takich jak „elektron". Istnieją bowiem przedmioty postrzegalne, które nie są elektronami; a zatem predykat Ψ należący do języka postrzeżeniowego J może tutaj — w przeciwieństwie do Φ — desygnować pewne przedmioty postrzegalne. Mogą więc w rezultacie pojawić się tutaj zdania rozstrzygalne zawierające termin „elektron" w sposób istotny. Czyż nie jest jednak rzeczą jasną, iż przy wprowadzaniu terminów takich jak „elektron" nie możemy się do tego rodzaju postulatów ograniczyć? Nie można właściwej interpretacji predykatu „elektron" wyznaczyć wyłącznie przez ustalenie, jakich przedmiotów ten predykat nie desygnuje. Taka interpretacja terminu „elektron" nie wyczerpuje na pewno faktycznej interpretacji tego terminu w języku nauki. W szczególności nie zapewnia ona charakteru rozstrzygalnego twierdzeniu o istnieniu elektronów: $\bigvee_x T_1 x$.

Jak zatem pogodzić powyższe wnioski z bezspornym faktem występowania w języku wielu teorii empirycznych terminów takich jak „elektron"? Mehlberg usiłuje to osiągnąć, traktując terminy takie jako wyrażenia synkategorematyczne, tj. nie jako samodzielne predykaty, lecz jako pozbawione samodzielnego znaczenia części składowe takich predykatów. W języku empirycznie sensownym nie może więc być, co prawda, samodzielnym predykatem termin „elektron"; może nim być jednak termin „zawierający-elektron". Jest to bowiem predykat desygnujący między innymi przedmioty postrzegalne. Predykat taki może być wobec tego wprowadzony za pomocą postulatu typu (1):

(1) $$\bigwedge_x (\Phi x \rightarrow T_1 x)\,,$$

gdzie Φ jest wyrażeniem języka postrzeżeniowego J — np. za pomocą postulatu głoszącego, iż: „dla każdego x: jeżeli x jest komorą Wilsona znajdującą się w takim a takim stanie, to x zawiera w swym wnętrzu (wolne) elektrony". Możemy bowiem tutaj przyjąć, iż istnieją przedmioty postrzegalne spełniające Φx w każdym modelu rodziny **M**. A wobec tego możemy sformułować zdania języka J' rozstrzygalne i zawierające predykat T_1 w sposób istotny. Niech a_1 będzie pewnym przedmiotem postrzegalnym spełniającym formułę Φx przy wszelkiej dopuszczalnej interpretacji języka J. Jak się łatwo przekonać, zdania: $T_1 a_1$, czy $\bigvee_x T_1 x$ będą zdaniami rozstrzygalnymi, a zarazem zawierającymi predykat T_1 w sposób istotny. W tak pojętym języku teorii empirycznej jest więc miejsce na terminy teoretyczne w sensie szerszym: predykaty desygnujące między innymi przedmioty postrzegalne i przypisujące im pewne własności teoretyczne. Nie ma natomiast miejsca na terminy teoretyczne w sensie węższym: predykaty desygnujące wyłącznie przedmioty niepostrzegalne. Nie ma wobec tego możliwości formułowania twierdzeń o istnieniu przedmiotów teoretycznych. Zdanie: $\bigvee_{x} T_1 x$ może głosić co najwyżej, że istnieją przedmioty zawierające elektrony, ale nie, że istnieją elektrony.

3. Jeśli chcemy zająć w sprawie powyższej stanowisko odmienne, bardziej, jak się wydaje, zgodne z istniejącym w nauce stanem rzeczy, musimy odrzucić któreś ze sformułowanych wyżej założeń. Jeśli chcemy, w szczególności, wzbogacić o terminy teoretyczne w sensie węższym scharakteryzowany tak, jak wyżej, język postrzeżeniowy J, musimy odrzucić założenie ograniczające rodzaj postulatów wprowadzających terminy teoretyczne do postulatów o postaci definicji cząstkowych. Postulatami nadającymi się do wprowadzania tak rozumianych terminów teoretycznych wydają się postulaty o terminach teoretycznych «kontrolowanych» przez kwantyfikator egzystencjalny. Idzie tu o postulaty, które doprowadzone do postaci normalnej odznaczają się tym, iż pewne zmienne będące argumentami terminu wprowadzanego związane są kwantyfikatorem egzystencjalnym lub znajdują się w zasięgu takiego kwantyfikatora[10].

Prostym i najczęściej spotykanym przykładem takiego postulatu może być postulat:

(2) $$\bigwedge_{x} [\Phi x \rightarrow \bigvee_{y} (T_1 yx \wedge T_2 y)],$$

gdzie Φ jest wyrażeniem języka postrzeżeniowego J, a T_1 i T_2 — terminami teoretycznymi, przy czym T_2 jest tutaj terminem teoretycznym w sensie węższym, desygnującym wyłącznie przedmioty niepostrzegalne. Postulat (2) uważać można za rezultat traktowania w postulacie (1) predykatu teoretycznego takiego jak „zawierający-elektron" jako predykatu złożonego z dwóch samodzielnych predykatów teoretycznych: T_1 — „jest zawarty w" i T_2 — „elektron". Wobec tego ów predykat złożony: $\bigvee_{y} (T_1 yx \wedge T_2 y)$, możemy traktować jako predykat desygnujący pewne przedmioty postrzegalne; a tym samym możemy tak traktować i predykat języka postrzeżeniowego Φ. Możemy zatem przyjąć, iż pewne przedmioty postrzegalne spełniają Φx w każdym modelu rodziny **M** — w przeciwieństwie do sytuacji rozważanej uprzednio. Postulat (2) nakłada jednak — w przeciwieństwie do postulatu (1) — bardzo słabe warunki na interpretację terminu T_2. Powstaje więc i tu pytanie, czy wprowadzony za pomocą takiego postulatu termin teoretyczny w sensie węższym, T_2, może być terminem naukowo przydatnym; czy może w szczególności występować w sposób istotny w zdaniach rozstrzygalnych. Okazuje się, jak zobaczymy, iż termin T_2 spełnia istotnie ten warunek. Aby jednak móc tę konkluzję należycie sformułować i uzasadnić, musimy rozpatrzyć szereg możliwych ujęć procedury wprowadzania terminu T_2 za pomocą postulatu (2). Możliwości te

10) Pojęcie postulatu o terminach teoretycznych «kontrolowanych» przez kwantyfikator egzystencjalny wprowadzone zostało przez Stopes-Roe w artykule „Some Considerations Concerning Interpretative Systems", *Philosophy of Science*, 25, 1958. Autor pokazuje, iż taki charakter ma każdy postulat sformułowany w języku opartym o węższy rachunek predykatów, który nie jest równoważny logicznie pewnej definicji cząstkowej (zwykłej lub uogólnionej). Artykuł zawiera sugestię o nieobserwowalnym charakterze przedmiotów denotowanych przez terminy wprowadzone przez tego rodzaju postulaty.

powstają w związku z różnym pojmowaniem interpretacji występującego w postulacie (2) terminu T_1. Czy predykat ten ma już jakąś niezależną od postulatu (2) interpretację? A jeśli tak, to jaką? Czy też postulat (2) formułuje jedyne warunki, jakie ma spełniać jego denotacja? Rozpatrzmy naprzód tę ostatnią, najbardziej liberalną koncepcję.

Język J'', który stanowić będzie przedmiot dalszych rozważań, powstaje z dołączenia do scharakteryzowanego tak, jak poprzednio, języka postrzeżeniowego J dwóch predykatów teoretycznych: T_1 i T_2. Jego modelami są wszelkie dziedziny typu:

$$\langle U; X_1, ..., X_l; Y_1, Y_2\rangle .$$

Niechaj $\mathbf{M}''$ będzie rodziną tych modeli języka J'', które stanowią wzbogacenia modeli rodziny $\mathbf{M}$. Zakładamy obecnie, iż postulat (2) jest jedynym postulatem wprowadzającym predykaty T_1 i T_2; oznaczymy go przez A_{12}. Interpretację właściwą języka J'' wyznacza więc rodzina $\mathbf{M}''(A_{12})$, obejmująca te i tylko te modele z rodziny $\mathbf{M}''$, w których prawdziwy jest postulat A_{12}. Do tak zinterpretowanego języka J'' możemy zastosować wprowadzone uprzednio pojęcie zdania rozstrzygalnego. Jego definicja, analogiczna do definicji (A), odwołuje się oczywiście do rodziny $\mathbf{M}''(A_{12})$. Komplikuje się natomiast w zastosowaniu do języka J'' sprawa pojęcia odpowiadającego pojęciu (B). A to ze względu na fakt, iż mamy tu do czynienia ze wzbogaceniem języka J nie o jeden termin, lecz o parę terminów jednocześnie. W rezultacie możemy wyróżnić trzy różne pojęcia odpowiadające pojęciu (B). Pierwsze z nich, (B_1), dotyczy nie samego terminu T_2, lecz pary terminów $\langle T_1, T_2\rangle$. To, że Z zawiera $\langle T_1, T_2\rangle$ w sposób istotny, określimy tak, jak to czyni definicja (B), odwołując się tylko zamiast do rodziny $\mathbf{M}'$, do rodziny $\mathbf{M}''$. Dwa pozostałe pojęcia dotyczą już nie pary terminów $\langle T_1, T_2\rangle$, lecz interesującego nas tutaj terminu T_2. Pojęcie (B_2) — zawierania przez zdanie Z terminu T_2 w sposób istotny — definiujemy również zgodnie z definicją (B) z tą tylko różnicą, że tam, gdzie tamta odwołuje się do rodzin $\mathbf{M}$ i $\mathbf{M}'$, definicja (B_2) odwołuje się odpowiednio do rodzin $\mathbf{M}'$ i $\mathbf{M}''$. Natomiast definicję pojęcia (B_3) sformułować można, jak następuje:

(B_3) Z zawiera T_2 w sposób istotny wtedy i tylko wtedy, gdy dla dowolnego modelu $\mathfrak{M}''$ należącego do rodziny $\mathbf{M}''(A_{12})$ istnieją takie modele $\mathfrak{M}_1''$ i $\mathfrak{M}_2''$ należące do rodziny $\mathbf{M}''$ i różniące się od $\mathfrak{M}''$ jedynie interpretacją T_2, iż Z jest prawdziwe w $\mathfrak{M}_1''$, a fałszywe w $\mathfrak{M}_2''$.

Różnica między dwoma ostatnimi pojęciami — (B_2) i (B_3) — wymaga paru słów wyjaśnienia. W obu — fakt istotnego, czy nieistotnego występowania w zdaniu Z terminu T_2 zrelatywizowany jest do interpretacji pozostałych terminów pozalogicznych zawartych w zdaniu Z. W obu — interpretacja predykatów postrzeżeniowych: O_1, ..., O_l rozumiana jest tak samo — jako wyznaczona przez rodzinę modeli $\mathbf{M}$. Inaczej natomiast pojmowana jest interpretacja pozostałego predykatu teoretycznego T_1. Pojęcie (B_2) odwołuje się do całkowicie dowolnej interpretacji tego predykatu — danej przez rodzinę $\mathbf{M}'$; pojęcie (B_3) tylko do takiej interpretacji, która jest zgodna z postulatem A_{12} — stąd odwołanie się do rodziny $\mathbf{M}''(A_{12})$. Różnica ta ma pewne znaczenie dla rozważanej przez nas sprawy.

Przede wszystkim jednak stwierdźmy, iż zgodnie z każdym z podanych wyżej określeń istnieją zdania rozstrzygalne języka J'' zawierające termin T_2 (ewentualnie $\langle T_1, T_2\rangle$) w sposób istotny. Należy do nich niewątpliwie zdanie: $\bigvee_y T_2 y$ — np. „istnieją elektrony". Jego rozstrzygalność jest konsekwencją tego, iż pewne przedmioty postrzegalne spełniają Φx w każdym modelu rodziny **M**. Jest przy tym rzeczą oczywistą, iż zdanie to zawiera termin T_2 (ewentualnie $\langle T_1, T_2\rangle$) w sposób istotny — w którymkolwiek z wyróżnionych przez nas znaczeń. Różnice między nimi ujawniają się dopiero w stosunku do zdań zawierających ponadto termin T_1. Do takich, niewątpliwie ważnych w teorii empirycznej zdań należą zdania typu $\bigvee_y (T_1 y a_1 \vee T_2 y)$ — np. „a_1 zawiera elektrony" — gdzie a_1 jest przedmiotem postrzegalnym spełniającym Φx w każdym modelu rodziny **M**. Są to na pewno zdania rozstrzygalne — prawdziwe w każdym modelu rodziny $\mathbf{M}''(A_{12})$. Czy zawierają one przy tym termin T_2 (ewentualnie $\langle T_1, T_2\rangle$) w sposób istotny? Jak się łatwo przekonać, odpowiedź na to pytanie wypadnie twierdząco przy definicji (B_1) i (B_3), przecząco przy definicji (B_2). W przypadku (B_1) odpowiedź jest oczywista. W przypadku (B_2) można z łatwością wskazać taką interpretację T_1 — np. relację pustą — przy której sposób interpretacji T_2 będzie bez wpływu na wartość logiczną całego zdania. Wszystkie takie interpretacje są z kolei wyłączone w przypadku (B_3) jako niezgodne z postulatem A_{12}.

Widać stąd, iż zakres zdań rozstrzygalnych zawierających termin T_2 w sposób istotny zwiększa się przy ograniczaniu swobody interpretacji terminu T_1. Braliśmy do tej pory pod uwagę wyłącznie ograniczenia nakładane przez postulat A_{12}. Ujęcie takie nie odpowiada jednak na pewno faktycznej praktyce naukowej. Wzorcem dla terminu T_1 był dla nas predykat „jest zawarty w", pokrywający się znaczeniowo z predykatem „jest częścią". Otóż jest rzeczą jasną, iż predykat taki użyty np. w postulacie wprowadzającym termin „elektron" ma z góry ustaloną interpretację. Musimy go więc traktować jako predykat, który został już przedtem wprowadzony do naszego języka za pomocą odrębnych, różnych od (2), postulatów. Jakie to mogą być postulaty? Wydaje się, iż w przypadku predykatu „jest częścią" postulaty takie obejmować muszą w każdym razie zdania dwóch rodzajów: (1) postulaty charakteryzujące strukturalne własności relacji bycia częścią — np. aksjomaty elementarnej mereologii; (2) postulaty ustalające związki tego predykatu z predykatami postrzeżeniowymi, w rodzaju rozważanych przez nas definicji cząstkowych, czy postulatów o terminach «kontrolowanych» przez kwantyfikator egzystencjalny. Niech A_1 oznacza obecnie koniunkcję ogółu postulatów dla predykatu T_1, a J' — język powstały przez wprowadzenie do języka postrzeżeniowego J predykatu T_1 za pomocą postulatu A_1. Jego interpretację właściwą stanowi, określona jak poprzednio, rodzina modeli $\mathbf{M}'(A_1)$. Tak scharakteryzowany język J' wzbogacamy następnie o predykat T_2 wprowadzony za pomocą postulatu (2), który oznaczać będziemy obecnie przez A_2. Niechaj $\mathbf{M}''(A_1)$ będzie rodziną modeli tak otrzymanego języka J'' obejmującą każdy i tylko taki model tego języka, który stanowi wzbogacenie pewnego modelu rodziny $\mathbf{M}'(A_1)$. Interpretację właściwą języka J'' stanowić będzie rodzina $\mathbf{M}''(A_1, A_2)$, obejmująca te i tylko te modele z rodziny $\mathbf{M}''(A_1)$, w których prawdziwy jest postulat A_2. Zdania rozstrzygalne obecnego języka J'' określamy jak poprzednio, odwołując się oczywiście tym razem do rodziny $\mathbf{M}''(A_1, A_2)$. Zwrot: zdanie Z zawiera termin T_2 w sposób istotny — definiujemy zgodnie z definicją (B), zastępując w niej rodzinę $\mathbf{M}$ przez $\mathbf{M}'(A_1)$, a $\mathbf{M}'$ przez $\mathbf{M}''(A_1)$. Łatwo okazać przy tych ustaleniach, iż istnieją zdania rozstrzygalne języka J' zawierające termin T_2 w sposób istotny — w szczególności zdanie: $\bigvee_y T_2 y$. A rozważane poprzednio zdania: $\bigvee_y (T_1 y a_1 \vee T_2 y)$? Odpowiedź na to pytanie zależy od rodzaju postulatów A_1. Wydaje się jednak, iż jeśli postulaty te mają choć w przybliżeniu odpowiadać np. faktycznej interpretacji predykatu „jest częścią", muszą obejmować zdania tak ograniczające interpretację tego predykatu, iż zapewni ona rozważanym zdaniom pożądany charakter. Wystarczy, jeśli do postulatów A_1 należeć będzie np. postulat:

$$\bigwedge_x (\Phi x \rightarrow \bigvee_y T_1 y x)\,,$$

w którym Φ reprezentuje to samo wyrażenie, co w postulacie A_2. A postulat taki ma dla predykatu „jest częścią" charakter trywialny. Głosi on jedynie, iż każdy przedmiot spełniający postrzeżeniowy warunek Φ posiada jakąś część (coś zawiera).

Widzimy przeto, iż w przeciwieństwie do postulatów o postaci definicji cząstkowych typu (1) pewne postulaty o terminach teoretycznych „kontrolowanych" przez kwantyfikator egzystencjalny nadają się do wprowadzania do języka postrzeżeniowego J terminów teoretycznych w sensie węższym — predykatów desygnujących tylko przedmioty niepostrzegalne. Rozważaliśmy z tego punktu widzenia postulat postaci (2) jako typowego reprezentanta tego rodzaju zdań. Najprostszym ich przykładem może być postulat:

$$\bigwedge_x (\Phi x \rightarrow \bigvee_y T_2 y)\,,$$

stanowiący chyba najsłabszą z dopuszczalnych odmian takich postulatów. Przykłady pewnych form bardziej złożonych podamy nieco niżej. Wszystkie te postulaty nadają się do wprowadzania terminów teoretycznych w tym mianowicie sensie, że pozwalają na formułowanie zdań rozstrzygalnych zawierających owe terminy w sposób istotny. Trzeba podkreślić, iż jest to warunek bardzo słaby, wysuwający w stosunku do wprowadzanych terminów minimalne wymagania. Mają one zapewnić tylko to, aby tak wprowadzony termin nie był w teorii empirycznej całkowicie zbędny. Wiadomo jednak, iż od terminów teorii empirycznej wymaga się spełnienia i innych warunków. Podstawowym z nich jest posiadanie przez te terminy empirycznego sensu. Ten właśnie wzgląd decydował, jak wiemy, o ograniczeniu się przez Mehlberga do postulatów o postaci definicji cząstkowych. Jak zatem przedstawia się sprawa empirycznej sensowności terminów teoretycznych w sensie węższym, wprowadzonych przez postulaty w rodzaju postulatu (2)? Nie spełniają one oczywiście przyjętego przez Mehlberga kryterium empirycznej sensowności, jak to widać z określenia tego kryterium przytaczanego przez nas uprzednio. Pod tym względem są one w sytuacji takiej samej jak terminy teoretyczne w sensie węższym, wprowadzone przez postulaty typu (1). Ale kryterium to wydaje się zbyt rygorystyczne. A pewne bardziej liberalne — a więc i bardziej adekwatne — kryteria empirycznej sensowności są przez termin taki, jak T_2, wprowadzony za pomocą postulatu (2) — spełnione. Tak jest w szczególności w przypadku ostatniego kryterium Carnapa[11]. Istnieje zdanie zawierające wyłącznie termin T_2: $\sim\bigvee_y T_2 y$, które wraz z postulatem (2) pociąga logicznie pewne zdanie języka postrzeżeniowego: $\sim\bigvee_x \Phi x$, z samego tego postulatu nie wynikające. Nawiasem mówiąc, kryterium to spełniają również terminy teoretyczne w sensie węższym wprowadzone przez postulaty typu (1). Cała ta sprawa nie ma jednak większego znaczenia, gdyż jest rzeczą widoczną, iż postulaty typu (2) nie stanowią jedynych postulatów ustalających związki rozważanych przez nas terminów z terminami postrzeżeniowymi. Obok nich wchodzą również w grę postulaty innych rodzajów, w tym postulaty o postaci definicji cząstkowych dla negacji danego terminu teoretycznego — a więc rozważane już przez nas postulaty typu (1′) — a te zapewniają owym terminom sens empiryczny, pojmowany bardzo rygorystycznie.

Tak więc otrzymaliśmy w wyniku naszych wywodów pewną rekonstrukcję logiczną języka teorii empirycznej, dopuszczającą istnienie w tym języku terminów teoretycznych desygnujących wyłącznie przedmioty niepostrzegalne i tym samym pozwalającą na formułowanie twierdzeń o istnieniu przedmiotów teoretycznych: $\bigvee_y T_2 y$. A skoro, zgodnie z założeniem, istnieją przedmioty postrzegalne spełniające warunki Φ, istnieć też muszą odpowiednie przedmioty teoretyczne. Twierdzenie o istnieniu tych przedmiotów stanowić może zresztą samodzielną hipotezę teoretyczną.

* * *

Na koniec chciałbym zwrócić uwagę na pewien ogólniejszy aspekt opisanej rekonstrukcji języka teorii empirycznej. W naszych rozważaniach stosowaliśmy ją do procedury wzbogacania języka postrzeżeniowego o terminy teoretyczne w sensie węższym. Takie zastosowanie było wynikiem określonej interpretacji universum języka U oraz jego podzbioru U_1. U utożsamialiśmy ze zbiorem wszelkich przedmiotów fizycznych, a U_1 — ze zbiorem wszelkich przedmiotów postrzegalnych. Rezultaty naszych wywodów są jednak w gruncie rzeczy niezależne od jakiejkolwiek określonej interpretacji zbiorów U i U_1. Istotne dla nich jest tylko to, iż wszelkie predykaty języka J mają w zbiorze przedmiotów nie należących do U_1 interpretację całkowicie dowolną, gdy tymczasem predykaty wprowadzane do języka J mają desygnować właśnie przedmioty nie należące do U_1. A skoro tak, to osiągnięte rezultaty mogą mieć zastosowanie również do innych sytuacji. Obszerną ich klasę stanowią sytuacje powstające przy przejściu od teorii opisujących jeden rodzaj obiektów do teorii wyjaśniających tamte teorie przez odwołanie się do obiektów, z których składają się obiekty badane poprzednio; a więc przy przejściu od przedmiotów codziennego otoczenia do świata molekuł,

11) „The Methodological Character of Theoretical Concepts", *Minnesota Studies...*, Vol. I. 1956.

od niego do świata atomów, a wreszcie cząstek elementarnych; a w naukach biologicznych — od organizmów do komórek, od komórek do chromosomów czy genów. Sytuacje te wymagają wprowadzenia do języka zawierającego terminy o interpretacji określonej w stosunku do pewnego typu przedmiotów — terminów, które odnosić się mają do przedmiotów innego rodzaju. Wydaje się więc, iż co najmniej niektóre z takich sytuacji można podciągnąć pod zarysowany przez nas schemat, interpretując odpowiednio zbiory U i U_1, np. U_1 — jako zbiór obiektów nie mniejszych niż atomy, a U — jako zbiór obejmujący prócz tego cząstki elementarne. We wszystkich takich procedurach występują, jak sądzę, postulaty, w których terminy wprowadzane «kontrolowane» są przez kwantyfikator egzystencjalny. W każdym z tych postulatów istotną rolę odgrywa predykat „jest częścią". Zilustrujemy to na paru przykładach — niezmiernie oczywiście uproszczonych. Zapiszemy je w postaci symbolicznej, aby uwidocznić w ten sposób rozmaitość form, jakie przybierać mogą postulaty omawianego typu. Relację bycia częścią oznaczać będziemy przez C, a zdefiniowaną przy jej pomocy relację posiadania części wspólnej („zachodzenia na") przez Z: $Zxy \equiv \bigvee_z (Czx \wedge Czy)$.

(1) „Każda porcja gazu jest zbiorem molekuł"
symbolicznie:

$$\bigwedge_x \bigwedge_y [Gx \wedge Cyx \rightarrow \bigvee_z (Mz \wedge Czx \wedge Zyz)],$$

gdzie G symbolizuje gaz, a M — molekułę, przy czym M jest terminem wprowadzanym.

(2) „Każdy atom wodoru składa się z protonu i elektronu"
symbolicznie:

$$\bigwedge_x \bigwedge_y \{Hx \wedge Ay \wedge Cyx \rightarrow \bigvee_u \bigvee_v [Pu \wedge Ev \wedge Cuy \wedge Cvy \wedge \bigwedge_z (Czy \rightarrow Zzu \vee Zzv)]\},$$

gdzie H symbolizuje wodór, A — atom, P — proton, a E — elektron, przy czym P i E są terminami wprowadzanymi.

(3) „Każda gameta organizmu homozygotycznego ze względu na cechę F zawiera gen na tę cechę"
symbolicznie:

$$\bigwedge_x \bigwedge_y [H(F)x \wedge Gy \wedge Cyx \rightarrow \bigvee_z (G(F)z \wedge Czy)],$$

gdzie $H(F)$ symbolizuje homozygotę ze względu na cechę F, G — gametę, a $G(F)$ — gen na cechę F, przy czym $G(F)$ jest terminem wprowadzanym. Podobnych przykładów można by podać wiele. Poprzestaniemy na przytoczonych, gdyż ilustrują one chyba dostatecznie omawiany ostatnio typ sytuacji. Każdy z nich budzi niewątpliwie szereg pytań i wątpliwości, wymagających szczegółowej analizy. Analiza taka jest jednak zadaniem wykraczającym poza ramy tego referatu. Tutaj ograniczyć się muszę do powyższych szkicowych uwag, których głównym celem jest podkreślenie szerokiego i różnorodnego zasięgu rozważanej problematyki.

Filozofia Nauki
Rok I, 1993, Nr 2-3

O świecie rzeczywistym i światach możliwych

Aparatura pojęciowa teorii modeli stosowana jest dziś szeroko poza domeną logiki matematycznej, do której należy ze swej genezy i istoty. Próbuje się przy jej pomocy ujmować pewne problemy natury filozoficznej — pewne zagadnienia metodologii, teorii poznania, ontologii. Są to na pewno próby warte podejmowania, stwarzające szanse precyzacji tych notorycznie niejasnych i trudno uchwytnych spraw. Jednocześnie jednak — jak wszelkie próby przenoszenia pewnego aparatu pojęciowego w dziedziny nowe — nie są one wolne od pewnych niebezpieczeństw. Przy tego rodzaju ekspansji danego aparatu pojęciowego istnieje zawsze groźba jego zniekształcenia, niewłaściwego zrozumienia, obarczenia obcymi treściami i intuicjami. Taka też sytuacja zachodzi, moim zdaniem, w przypadku niektórych zastosowań aparatury teorio–modelowej. W jej pojęcia wkłada się przy tej okazji swoisty sens «filozoficzny», niezgodny z ich znaczeniem pierwotnym. Dwojakie wydają się źródła tego stanu rzeczy. Pojęcia teorio–modelowe są — tak, jak wszelkie pojęcia logiczne — pojęciami «formalnymi». Różnią się pod tym względem zasadniczo od wszelkich pojęć pozalogicznych, w szczególności filozoficznych. Sprawia to komuś, kto logikiem nie jest, trudności w uchwyceniu ich właściwego sensu, w zrozumieniu, że dane pojęcie to «tylko» oznacza, że «nic więcej» poza nim się nie kryje. Stąd — tendencje do doszukiwania się w takich pojęciach pewnej treści bogatszej, do wkładania w nie pewnych intuicji «pozaformalnych». Sprzyja temu jednocześnie często spotykany sposób prezentowania tej aparatury przez logików. Prócz oficjalnych określeń i definicji posługują się oni z upodobaniem sformułowaniami metaforycznymi — efektownymi i sugestywnymi, ale łatwo mylącymi, jeśli się o ich metaforyczności zapomina i traktuje na serio. Do takich sformułowań w teorii modeli należą przede wszystkim zwroty mówiące o przedmiotach tej teorii — o modelach języków sformalizowanych — jako o «światach rzeczywistych i możliwych». Przenosi się ten sposób mówienia na określenia definiowanych na gruncie tej teorii pojęć modalnych: pojęć możliwych i koniecznych zdań, lub stanów rzeczy. Możliwe zdanie — to zdanie prawdziwe w pewnym «świecie możliwym»; możliwy stan rzeczy — to stan rzeczy zachodzący w którymś z takich «światów». Brzmią te sformułowania tak samo, jak znane definicje filozoficzne — na przykład określenia Leibniza. A jaki jest ich właściwy, teorio–modelowy sens? Czy usprawiedliwia on te «filozoficzne» sformułowania? Próbą odpowiedzi na te pytania będą poniższe uwagi. Zanim do nich przejdę, spróbuję naszkicować w paru słowach owe podstawowe teorio–modelowe konstrukcje, które tu wchodzą w grę. Ograniczę się przy tym do ich maksymalnie uproszczonej wersji; jej ewentualne komplikacje nie mają dla naszych rozważań większego znaczenia.

Przedmiotem naszych uwag uczynimy język najprostszego typu — tzw. standardowo sformalizowany język J, zawierający prócz stałych logicznych, obejmujących spójniki zdaniowe, kwantyfikatory i predykat identyczności, wyrażenia pozalogiczne następujących rodzajów: zmienne indywiduowe x_1, x_2, ..., predykaty (o dowolnej liczbie argumentów) P_1, P_2, ..., oraz nazwy indywiduowe a_1, a_2, ... W schemat taki wtłoczyć się zresztą dają języki większości teorii naukowych, a także znaczne fragmenty języka potocznego. Dla teorio–modelowego ujęcia danego języka istotne jest rozbicie go na dwa komponenty: formalny i treściowy. Pierwszy — to język niezinterpretowany, scharakteryzowany tylko syntaktycznie, drugi — to jego interpretacja, wyznaczona przez teorio–mnogościowy twór zwany modelem. W przypadku języka J model taki przybiera postać układu

$$\mathfrak{m} = \langle U, R_1, R_2, ..., b_1, b_2, ...\rangle$$

złożonego z niepustego zbioru U, z relacji R_1, R_2..., zachodzących między elementami zbioru U i z indywiduów b_1, b_2, ... wyróżnionych w zbiorze U. Każdy model $\mathfrak{m}$ języka J wyznacza jedną z możliwych interpretacji tego języka: jego zmiennym przyporządkowuje jako zbiór wartości zbiór U, a predykatom P_1, P_2, ... i nazwom a_1, a_2, ... jako ich denotacje odpowiednio relacje R_1, R_2, ... i indywidua b_1, b_2, ... Zakładamy, że w przeciwieństwie do wyrażeń pozalogicznych, stałe logiczne języka J mają we wszystkich modelach tego języka tę samą, standardową interpretację (nie wymieniamy ich z tej racji wśród składników modelu $\mathfrak{m}$). Pojęcie modelu $\mathfrak{m}$ języka J pozwala zdefiniować w znany dobrze sposób pojęcie zdania języka J prawdziwego w modelu $\mathfrak{m}$. Mówiąc swobodnie, zdanie α jest prawdziwe w modelu $\mathfrak{m}$, gdy jest tak, jak głosi zdanie α w interpretacji wyznaczonej przez model $\mathfrak{m}$. Przykładowo: zdanie $P_1 (a_1, a_2)$ jest prawdziwe w modelu $\mathfrak{m}$, gdy relacja R_1 (stanowiąca w modelu $\mathfrak{m}$ denotację predykatu P_1) zachodzi między indywiduami b_1 i b_2 (stanowiącymi w modelu $\mathfrak{m}$ denotacje nazw a_1 i a_2), tj. gdy $R_1 (b_1, b_2)$. Gdy język J jest językiem sensownym, wyposażonym w określone znaczenie, istnieje wśród jego interpretacji możliwych interpretacja właściwa, odpowiadająca temu zastanemu znaczeniu. Istnieje, co za tym idzie, wśród jego modeli możliwych model właściwy $\mathfrak{m}^*$, wyznaczający ową interpretację:

$$\mathfrak{m}^* = \langle U^*, R_1{}^*, R_2{}^*, ..., b_1{}^*, b_2{}^*, ...\rangle.$$

Jeśli model możliwy języka J określimy jako to, o czym można mówić w języku J (ze względu na jego właściwości synktatyczne), model właściwy będzie tym, o czym faktycznie mówi się w języku J. Odwołując się do tego pojęcia, definiujemy dla zdań języka J «absolutne» pojęcie prawdziwości: zdanie α jest prawdziwe, gdy jest prawdziwe w modelu $\mathfrak{m}^*$.

Ta prosta aparatura pojęciowa służy m.in. do eksplikacji pojęć modalnych: pojęcia możliwości i konieczności w różnych ich odmianach. Przede wszystkim definiuje się przy jej pomocy modalne pojęcia metalogiczne — orzekane bezpośrednio o zdaniach, a nie o stanach rzeczy. To pojęcia możliwego i koniecznego zdania języka J. Ograniczymy się tutaj do omówienia pojęcia możliwości. Pojęcie konieczności jest, jak wiadomo, definiowane przez to ostatnie: zdanie α jest konieczne, gdy zdanie nie–α nie jest możliwe. Pojęcie zdania możliwego występuje w trzech co najmniej odmianach, odpowiadających trzem wyróżnianym zwykle rodzajom możliwości: możliwości logicznej, językowej (inaczej analitycznej) i fizycznej. Ich teorio–modelowe definicje odwołują się do trzech klas modeli języka J — oznaczymy je symbolicznie przez M , M_A i M_F. Klasa M — to, w najprostszym ujęciu, klasa wszystkich modeli języka J. Klasy M_A i M_F — to pewne jej podklasy. Określić je możemy najprościej, wyróżniając dwa podzbiory zdań języka J: zbiór jego postulatów znaczeniowych (inaczej aksjomatów) A, oraz zbiór praw fizycznych F. Klasę M_A definiujemy jako klasę tych modeli języka J, w których prawdziwe są wszystkie zdania zbioru A, a klasę M_F — jako klasę tych modeli języka J, w których ponadto prawdziwe są wszystkie zdania zbioru F. W rezultacie:

$$M_F \subset M_A \subset M\,.$$

Owe trzy rodzaje zdań możliwych języka J definiujemy za pomocą następujących definicji:

(D1) Zdanie α jest możliwe logicznie, gdy istnieje taki model $\mathfrak{m}$ należący do klasy M, że α jest prawdziwe w $\mathfrak{m}$.

(D2) Zdanie α jest możliwe językowo, gdy istnieje taki model $\mathfrak{m}$ należący do klasy M_A, że α jest prawdziwe w $\mathfrak{m}$.

(D3) Zdanie α jest możliwe fizycznie, gdy istnieje taki model $\mathfrak{m}$ należący do klasy M_F, że α jest prawdziwe w $\mathfrak{m}$.

Przy przyjętych przez nas upraszczających założeniach określenia powyższe sprowadzają się do znanych i prostych stwierdzeń. Zdanie możliwe logicznie — to po prostu zdanie niesprzeczne; zdanie możliwe językowo — to zdanie niesprzeczne z ogółem postulatów znaczeniowych danego języka; wreszcie zdanie możliwe fizycznie — to zdanie niesprzeczne z ogółem postulatów znaczeniowych i praw fizycznych. Sens tych określeń zależny jest oczywiście od zawartości zbiorów A i F. Wiadomo dobrze, że problem ich wyróżnienia jest sprawą niełatwą i dyskusyjną. Mówiąc ogólnikowo, pierwszy z nich ma obejmować wszelkie twierdzenia pozalogiczne o charakterze definicyjnym; drugi — wszelkie uniwersalne twierdzenia fizyczne o charakterze empirycznym. Nie możemy tu wchodzić w sprawę bliższej precyzacji tych określeń. Warto jedynie zwrócić uwagę na to, że nasze założenie utożsamiające owe zbiory ze zbiorami zdań języka J stanowi daleko idące uproszczenie. Gdy J jest językiem elementarnym, ogół zależności znaczeniowych zachodzących między jego wyrażeniami, a w szczególności ogół prawidłowości fizycznych, bywa niewyrażalny w samym języku J. Nie możemy wówczas zakładać, że klasy M_A i M_F — to klasy elementarnie definiowalne. Definicje (D2) i (D3) stają się w takim przypadku niesprowadzalne do przytaczanych wyżej sformułowań prostszych.

Zilustrujmy wprowadzone pojęcia na paru przykładach — z konieczności naiwnych i uproszczonych. Załóżmy, że język J — to fragment języka polskiego obejmujący m.in. predykaty „jest większy od", „przyciąga", oraz nazwy „Ziemia" i „Księżyc". Przyjmijmy też, że do postulatów znaczeniowych tego języka należy zdanie „dla każdego x_1 i x_2: jeśli x_1 jest większe od x_2, to x_2 nie jest większe od x_1", a do praw fizycznych twierdzenie „dla każdego x_1 i x_2: jeśli x_1 przyciąga x_2, to x_2 przyciąga x_1". Rozpatrzmy na gruncie tych założeń status modalny zdań następujących:

(1) Księżyc jest większy od Ziemi i Księżyc nie jest większy od Ziemi;

(2) Księżyc jest większy od Ziemi i Ziemia jest większa od Księżyca;

(3) Ziemia przyciąga Księżyc, a Księżyc nie przyciąga Ziemi;

(4) Księżyc jest większy od Ziemi.

Jest rzeczą widoczną, że zdanie (1) jest niemożliwe logicznie, a tym samym językowo i fizycznie; zdanie (2) jest możliwe logicznie, ale niemożliwe językowo i fizycznie; zdanie (3) jest niemożliwe jedynie fizycznie, a zdanie (4) jest możliwe w każdym z wyróżnionych przez nas sensów.

Zarysowana tu definicja zdania możliwego nie jest jedyną wersją tego pojęcia znaną z literatury przedmiotu. Bywa ona, w zależności od celów, jakim ma służyć, w różny sposób modyfikowana. O pewnych takich modyfikacjach chcę wspomnieć jeszcze w paru słowach. Sprowadzają się one do przyjęcia dodatkowych, coraz to mocniejszych założeń:

(i) Wszystkie modele klasy M (a tym samym klas M_A i M_F) mają wspólne uniwersum U^*;

(ii) Język J zawiera nazwy wszystkich elementów uniwersum U^*;

(iii) Nazwy elementów uniwersum U^* interpretowane są tak samo jak we wszystkich modelach klasy M (a tym samym, klas M_A i M_F); denotacją nazwy a_1 jest we wszystkich tych modelach to samo indywiduum b_1^*.

Przyjęcie tych założeń (wszystkich, lub tylko niektórych) prowadzi na gruncie definicji (D1)–(D3) do pewnych pojęć możliwości mocniejszych od wyjściowych. Wszystkie one jednak oparte są na tych samych podstawowych intuicjach i formalnych konstrukcjach.

Otóż, jak wspomniałem na wstępie, owe teorio–modelowe konstrukcje prezentowane bywają w swoistej szacie słownej nawiązującej do tradycyjnej terminologii filozoficznej. Zabiegiem podstawowym jest tu utożsamianie modeli języka omawianego ze «światami możliwymi», a modelu właściwego ze «światem rzeczywistym». Pozwala to z kolei na określenie zdania możliwego jako zdania prawdziwego w pewnym «świecie możliwym». Czy jest to jednak terminologia uprawniona? Utożsamianie modelu właściwego m* języka *J* ze «światem rzeczywistym» sprzeciwów w zasadzie nie budzi, choć wymaga pewnych wyjaśnień i zastrzeżeń. m* — jako pewien układ złożony ze zbioru przedmiotów i określonych w tym zbiorze relacji — jest tym, co najtrafniej chyba nazywa się dziedziną rzeczywistości. Ogólnie biorąc, jest to pewien fragment świata rzeczywistego, a nie «świat rzeczywisty». «Światem» można by nazwać co najwyżej fragment — w jakimś sensie — «maksymalny»; na przykład model właściwy języka «maksymalnego», pojętego jako język całej naszej wiedzy o rzeczywistości. Byłaby to jednak, po pierwsze, «maksymalność» tylko względna, zrelatywizowana do aktualnego etapu naszej wiedzy; po drugie, nie pokrywająca się bynajmniej z «uniwersalnością», gdyż języków w pełni «uniwersalnych» współczesna semantyka logiczna, jak wiadomo, nie dopuszcza. Ponieważ obiektami semantyki teorio–modelowej są z reguły języki «niemaksymalne», np. języki poszczególnych teorii naukowych, ich modele właściwe zasługują jedynie na nazwę fragmentów świata rzeczywistego. Warto podkreślić, że są to zawsze fragmenty świata w określony sposób «artykułowane». Nie pokrywają się z prostymi zbiorami przedmiotów; obejmują również pewne wyróżnione w tych zbiorach relacje (podzbiory, funkcje, indywidua itp.). Ów sposób «artykulacji» wyznaczony jest przez zasób wyrażeń języka, którego dany fragment jest korelatem. W rezultacie ograniczoność takiego fragmentu polega nie tylko na niepełności jego uniwersum, ale i na jednostronności jego «artykulacji».

Jeśli — przy wspomnianych zastrzeżeniach — nazywanie modelu właściwego języka *J* „światem rzeczywistym" (przez ten język opisywanym) jest w zasadzie procedurą uprawnioną, określanie innych jego modeli jako „światów możliwych" jest terminologią wprowadzającą w błąd. Błędna jest w szczególności sugestia, jakoby w przeciwieństwie do «świata rzeczywistego» owe «światy możliwe» zbudowane były z jakichś przedmiotów fikcyjnych — nieistniejących a możliwych. Wszystkie modele języka *J* zbudowane są z takich tylko przedmiotów, jakie zakłada ontologia logika, te modele opisującego, w jakich istnienie ów logik wierzy. Ta ontologia — to, mówiąc w uproszczeniu, teoria mnogości (z reguły tzw. naiwna teoria mnogości). Zarzuca się jej niekiedy «mnożenie bytów ponad potrzebę», nieuprawnione «rozdymanie» naszego świata. Obejmuje ona istotnie prócz przedmiotów konkretnych nieskończoną hierarchię przedmiotów abstrakcyjnych (zbiory przedmiotów konkretnych, zbiory takich zbiorów itp). Nie obejmuje jednak nigdy przedmiotów fikcyjnych; nie obejmuje w szczególności niczego takiego, jak nieistniejące a możliwe indywidua! Toteż wszelkie konteksty, w których o takich indywiduach z pozoru mowa, rekonstruuje logik z reguły w sposób taki, który żadnych przedmiotów fikcyjnych nie zakłada — np. przez utożsamienie tego rodzaju przedmiotów z pewnymi tworami abstrakcyjnymi[1]. W rezultacie

1) Spotykamy niejednokrotnie wypowiedzi logików mówiące o możliwych zdarzeniach (stanach, czynach itp.); w istocie mowa w nich nie o konkretnych zdarzeniach, lecz o typach zdarzeń, a więc o pewnych przedmiotach abstrakcyjnych.

uniwersa wszystkich modeli języka J — to zbiory przedmiotów rzeczywistych, relacje w nich wyróżniane — to relacje między przedmiotami rzeczywistymi, itp. Status ontologiczny wszystkich modeli języka J jest dokładnie taki sam. Żaden z tych modeli nie jest «bardziej rzeczywisty» od innych. Dotyczy to również modelu właściwego, który pod względem ontologicznym niczym się od pozostałych modeli nie wyróżnia. Te rzekome «światy możliwe» są równie realne, jak ów «świat rzeczywisty»; są to — podobnie jak on — różne fragmenty tej samej rzeczywistości. Swój wyróżniony charakter zawdzięcza model właściwy względom nie ontologicznym, lecz językowym. Jest to ten fragment rzeczywistości, który reprezentuje właściwą interpretację języka J. Język ten, rozumiany jako twór syntaktyczny, może być interpretowany na różne sposoby — odpowiadają im różne modele języka J. Jeśli jednak J jest językiem sensownym, jeden z tych sposobów wyróżniony jest jako właściwy — jemu właśnie odpowiada model właściwy języka J. Dany fragment rzeczywistości jest tu więc kwalifikowany jako «rzeczywisty» lub «możliwy» ze względu na dany język, na jego zastane znaczenie. Logiczne pojęcia „świata rzeczywistego" i „świata możliwego" — to pojęcia nie ontologiczne, lecz semantyczne[2].

Do czegóż zatem sprowadza się owo filozoficznie brzmiące określenie zdania możliwego jako zdania prawdziwego w pewnym «świecie możliwym»? Jest to, jak wiemy, swobodna parafraza definicji (D1), utożsamiającej zdanie możliwe (logicznie) ze zdaniem prawdziwym w pewnym modelu języka J. Stwierdzenie, iż zdanie α jest możliwe, to nic innego zatem jak stwierdzenie, iż zdanie to jest prawdziwe przy pewnej interpretacji języka J, lub — co na jedno wychodzi — przy pewnej interpretacji wyrażeń pozalogicznych występujących w zdaniu α. Gdy zdanie α jest przy swej interpretacji właściwej fałszywe, fakt jego możliwości sprowadza się do istnienia pewnej interpretacji innej, przy której zdanie to staje się prawdziwe. Skoro jednak jest to interpretacja inna, zdanie α tak zinterpretowane zachowuje tylko swoje dotychczasowe własności syntaktyczne, a traci semantyczne: inny ma obecnie sens, o czym innym mówi, co innego stwierdza. Aby uwyraźnić tę myśl, nazwijmy zdanie wyposażone w określony sens — stwierdzeniem. Cóż zatem znaczy powiedzenie, że dane stwierdzenie jest co prawda fałszywe, lecz możliwe? Nie znaczy ono bynajmniej, że stwierdzenie to jest prawdziwe w jakimś różnym od rzeczywistego «świecie możliwym». Znaczy tylko

2) Zamiast pojęcia „świata możliwego" używa się niekiedy we współczesnej semantyce logicznej pojęcia „opisu świata możliwego" i na nim opiera konstrukcje modalne. Ów «opis świata możliwego» — to pewien zbiór zdań języka J, a więc pojęcie ewidentnie językowe, a nie ontologiczne. Tylko dzięki związkom semantycznym z modelami języka J ów zbiór zdań uważany być może za opis jakiegokolwiek «świata». Reprezentowane bywają w tej sprawie różne koncepcje. Za «opis świata możliwego» uważa się bądź zbiór wszystkich zdań języka J prawdziwych w pewnym modelu $\mathfrak{m}$, bądź pewne jego podzbiory; np. tzw. zbiór modelowy („*model-set*") Hintikki, pojęty jako dowolny «dostatecznie obszerny» podzbiór ogółu zdań języka J prawdziwych w modelu $\mathfrak{m}$, lub tzw. opis stanu („*state-desctription*") Carnapa, obejmujący wszystkie zdania atomowe i negacje zdań atomowych języka J prawdziwe w modelu $\mathfrak{m}$ (przy wymienionych wyżej założeniach (i)–(iii)). Wracamy tym samym do problemu poprzedniego. Mamy tu bowiem do czynienia z «opisami światów możliwych» o tyle tylko, o ile na nazwy takich «światów» zasługują modele języka.

tyle, że prawdziwe jest pewne inne stwierdzenie, tożsame z tamtym pod względem syntaktycznym.

Zilustrujmy tę konkluzję na przykładzie omawianego uprzednio zdania języka polskiego „Księżyc jest większy od Ziemi", które uznaliśmy za zdanie możliwe. Przy swej interpretacji właściwej, przy której nazwa „Księżyc" denotuje Księżyc, nazwa „Ziemia" — Ziemię, a predykat „jest większy od" — relację większości, jest to zdanie fałszywe. Łatwo jednak znaleźć interpretację taką, przy której stanie się ono prawdą. Będzie nią np. interpretacja przyporządkowująca predykatowi „jest większy od" jako denotację relację mniejszości (lub interpretacja zamieniająca interpretacje właściwe nazw „Księżyc" i „Ziemia"). Dlatego właśnie zdanie to uznajemy za możliwe. Mówiąc paradoksalnie, zdanie „Księżyc jest większy od Ziemi" jest możliwe, bo prawdziwe jest zdanie „Księżyc jest mniejszy od Ziemi" (lub zdanie „Ziemia jest większa od Księżyca"). Inna to na pewno idea, niż ta, którą wiąże się zwykle z pojęciem możliwości i wyraża słowami: zdanie „Księżyc jest większy od Ziemi" jest możliwe, bo jest tak, jak to zdanie głosi, w pewnym «świecie możliwym».

Mówimy do tej pory w zasadzie o możliwości logicznej. Ale zupełnie podobnie ma się sprawa z możliwością językową i fizyczną. Jedyna różnica — to dodatkowy warunek nakładany tu na ową interpretację, która ma okazać możliwość zdania α: ma to być interpretacja spełniająca postulaty znaczeniowe danego języka, ewentualnie prawa fizyczne. Zdanie α jest możliwe językowo (fizycznie), gdy zdanie to jest prawdziwe przy pewnej interpretacji języka J, przy której prawdziwe są zarazem postulaty znaczeniowe A (oraz prawa fizyczne F). Gdy α jest przy swej interpretacji właściwej zdaniem fałszywym, interpretacja okazująca jego możliwość językową (fizyczną) musi się oczywiście różnić od tamtej. Nadaje ona tym samym zarówno zdaniu α, jak i postulatom A (oraz prawom F) inny sens. Otrzymujemy w rezultacie tę samą konsekwencję, co poprzednio. Nie jest bynajmniej tak, jakoby dane stwierdzenie — fałszywe w świecie rzeczywistym — było prawdziwe w pewnym «świecie możliwym», w którym prawdziwe są ponadto postulaty znaczeniowe (oraz prawa fizyczne) wyposażone w swój zwykły sens. Jest raczej tak, że pewne inne stwierdzenie i inaczej rozumiane postulaty (oraz prawa) okazują się prawdziwe w naszym świecie rzeczywistym. Omawiany wyżej przykład służyć może za ilustrację również tej sytuacji, bo dotyczy zdania możliwego nie tylko logicznie, ale także językowo i fizycznie.

Pojęcia możliwości uwzględnione przez nas do tej pory — to modalne pojęcia metalogiczne: odnoszące się do zdań określonego języka i wyróżniające pewne ich rodzaje. Można by zatem sądzić, że w tym leży przyczyna nieadekwatności tych pojęć w stosunku do wiązanych z nimi zwykle filozoficznych intuicji. Te bowiem dotyczą nie tyle zdań — pewnych tworów językowych, ile stanów rzeczy — pewnych tworów ontologicznych: to nie samo zdanie, ale to, o czym zdanie mówi, ma być możliwe, czy konieczne. Logika stworzyła liczne systemy tak rozumianych pojęć modalnych. Pojęcia możliwości i konieczności występujące w tzw. logikach modalnych — to pojęcia nie metalogiczne, lecz logiczne. Z syntaktycznego punktu widzenia są to jednoargumentowe spójniki zdaniowe: możliwe, że α, konieczne, że α. W odróżnieniu od zwrotów rozważanych poprzednio, symbol α występuje tu w roli zwykłej, a nie cudzysłowowej: jako nazwa pewnego stanu rzeczy, a nie zdania. Spróbujmy wobec tego przyjrzeć się bliżej znaczeniom wiązanym z owymi spójnikami modalnymi. Wystarczy, jak poprzednio, ograniczyć się do spójnika możliwości, bo spójnik konieczności definiowalny jest przy jego pomocy: konieczne, że α — to tyle, co — niemożliwe, że nie α. Uważa się niekiedy, że sens takich spójników wyznaczony jest przez układ aksjomatów danego systemu modalnego. Gdy jednak interesuje nas filozoficzna zawartość pojęć modalnych, nie możemy poprzestać na ich aksjomatycznej charakterystyce. Ta ustala jedynie pewne formalne własności tych pojęć i związki między nimi zachodzące — wspólne pojęciom o różnej treści i różnej zawartości filozoficznej. O właściwym sensie pojęć modalnych decyduje semantyczna interpretacja danego systemu modalnego. Toteż chcąc ów sens uchwycić, odwołać się musimy do semantyki logik modalnych. Zadowalająca semantyka logik modalnych jest osiągnięciem stosunkowo niedawnym. Nie mogę tu jej przedstawić w sposób szczegółowy, ale i nie ma potrzeby, bo nawet pobieżna jej charakterystyka pozwala zauważyć, że nie wnosi ona do dyskusji dotychczasowej żadnych zasadniczo nowych elementów. Jest to semantyka teorio-modelowa, nadająca logicznym pojęciom modalnym sens w istocie taki sam, jaki przysługiwał im w ich wersji metalogicznej.

Rozważmy przypadek najprostszy — taki, w którym logiczny spójnik możliwości występuje wyłącznie przed zdaniami (tj. formułami zdaniowymi nie zawierającymi zmiennych wolnych). Sens tego spójnika charakteryzuje reguła semantyczna definiująca pojęcie prawdziwości dla tego rodzaju zdań możliwościowych:

(R) Zdanie „możliwe, że α" jest prawdziwe w modelu $\mathfrak{m}^*$, gdy istnieje taki model $\mathfrak{m}$ «alternatywny» względem $\mathfrak{m}^*$, że zdanie α jest prawdziwe w $\mathfrak{m}$.

Język, do którego należą wspomniane zdania — to opisany poprzednio język J wzbogacony o spójnik możliwości. Tak samo, jak poprzednio, rozumiane jest też pojęcie modelu, oraz pojęcie prawdziwości w modelu (w zastosowaniu do zdań języka J nie zawierających spójnika możliwości). Pojęciem wymagającym wyjaśnienia jest natomiast występująca w powyższej regule relacja «alternatywności», zachodząca między modelami danego języka. W semantyce logik modalnych jest to pojęcie pierwotne, charakteryzowane tylko przez pewne własności formalne (zwrotność, symetria, przechodniość) — różne zresztą dla różnych systemów logiki modalnej. Ale dopiero pewna konkretyzacja tego pojęcia determinuje określony sens pojęcia możliwości. Można tego dokonać na wiele sposobów. Owa klasa modeli alternatywnych względem $\mathfrak{m}^*$ może być określana różnie. W rozważanym przez nas przypadku może to być — tak, jak poprzednio — bądź klasa wszystkich modeli M, bądź jej podklasy M_A i M_F — w zależności od tego, czy idzie nam o możliwość logiczną, językową czy fizyczną. Dla naszej argumentacji ważne jest to, że w każdym z tych przypadków klasa modeli «alternatywnych» względem $\mathfrak{m}^*$ obejmuje pewne modele języka J różne od $\mathfrak{m}^*$. Fakt ten pociąga analogiczne konsekwencje dotyczące sensu pojęcia możliwości, jak te, na które zwracaliśmy uwagę poprzednio.

Sens spójnika możliwości określony przez regułę (R) wyrażany bywa metaforycznie, jak następuje: dany stan rzeczy jest możliwy, gdy zachodzi w pewnym «świecie możliwym». Ale jest to sformułowanie mylące. Owe «światy możliwe» — to nic innego jak interpretacje pewnego języka. Reguła (R) głosi tylko tyle, że zdanie „możliwe, że α" jest prawdziwe, gdy istnieje taka interpretacja jego wyrażeń pozalogicznych, przy której prawdziwe jest zdanie α. Jeżeli zdanie α jest fałszywe przy swej interpretacji właściwej, prawdziwe może być tylko przy pewnej interpretacji innej — inny, co za tym idzie, nadającej mu sens. Nie jest więc tak, że dany stan rzeczy — nie zachodzący w świecie rzeczywistym — jest możliwy, gdy zachodzi w pewnym «świecie możliwym». W myśl reguły powyższej, jest możliwy wtedy, gdy pewien inny stan rzeczy — tożsamy z nim pod względem strukturalnych — zachodzi w świecie rzeczywistym. Formułując rzecz paradoksalnie powiemy, że to, że Księżyc jest większy od Ziemi, jest stanem rzeczy możliwym, bo faktem jest m.in. to, że Księżyc jest mniejszy od Ziemi. Istotna z tego punktu widzenia wspólność między owymi stanami sprowadza się do tego, że oba scharakteryzowane być mogą jako zachodzenie dwuczłonowej relacji między dwoma indywiduami.

Daleko tu jesteśmy od tych filozoficznych intuicji, które na ogół wiąże się z pojęciem możliwości, i które często podkłada się pod teorio–modelowe konstrukcje modalne. Konstrukcje te — nie tylko w swej wersji metalogicznej, ale i logicznej — tego rodzaju intuicji nie chwytają i chwytać nie mogą. Zwróćmy uwagę na to, że wszystkie przytaczane do tej pory eksplikacje wyjaśniają sens pojęć modalnych w metajęzyku, który żadnych takich pojęć nie zawiera. Zachodzi pod tym względem, nawiasem mówiąc, wyraźna różnica w porównaniu z klasycznymi pojęciami logicznymi: ich eksplikacje formułowane są w metajęzyku, który zawiera pojęcia logiczne tak samo rozumiane. Co więcej — metajęzyk służący do eksplikacji pojęć modalnych nie zawiera w gruncie rzeczy żadnych pojęć «pozaformalnych». Nic dziwnego, że okazuje się językiem zbyt ubogim na to, aby wyrazić pewne «pozaformalne» treści filozoficzne wiązane z pojęciem możliwości. Ich uchwycenie wymaga odwołania się do bogatszego aparatu pojęciowego. Co to ma być za aparat? Jak wyglądać mają owe filozoficzne eksplikacje? Aby móc odpowiedzieć na te pytania, trzeba dysponować określoną koncepcją filozoficznego pojęcia możliwości, trzeba żywić dostatecznie wyraźne intuicje dotyczące jego zawartości. Nie spełniając tego warunku, nie potrafię wysunąć w tej sprawie żadnej propozycji. Ograniczę się do wskazania przykładu pewnej koncepcji odmiennej od dotychczasowych — koncepcji, która wydaje się wolna od paradoksalnych rysów koncepcji teorio–modelowej.

Mam na myśli koncepcję możliwości w trójwartościowej logice Łukasiewicza. Decydująca dla naszych rozważań jest nie tyle jej formalna charakterystyka, co intuicyjna interpretacja. Jest to, jak wiadomo, koncepcja oparta na określonym założeniu filozoficznym — tezie indeterminizmu. Teza ta zakłada istnienie pewnych niezdeterminowanych stanów rzeczy. Przyjmuje, iż zajście pewnych stanów przyszłych nie jest wyznaczone w sposób jednoznaczny przez ogół zdarzeń dotychczasowych — zdarzenia te mogą równie dobrze zajść, jak i nie zajść. To właśnie — stany rzeczy możliwe. A zdania opisujące te stany — to zdania możliwe. Z logicznego punktu widzenia istotne jest to, że są to zdania o niezdeterminowanej wartości logicznej — zdania, które nie są ani prawdziwe, ani fałszywe. To odróżnia tę koncepcję od dotychczasowych konstrukcji teorio-modelowych, wedle których każde zdanie danego języka ma określoną wartość logiczną w każdym modelu tego języka: jest w nim prawdziwe lub fałszywe. Dotyczy to również modelu właściwego $\mathfrak{m}^*$: każde zdanie jest prawdziwe w $\mathfrak{m}^*$, lub fałszywe w $\mathfrak{m}^*$, innymi słowy — jest prawdziwe, lub fałszywe. W rezultacie, chcąc traktować pewne zdanie fałszywe jako możliwe, musieliśmy abstrahować od jego właściwej interpretacji, od jego zastanego znaczenia. Tutaj do tego zmuszeni nie jesteśmy. Jako zdania możliwe traktujemy zdania logicznie niezdeterminowane, zdania, które nie są «jeszcze» ani prawdą, ani fałszem; zdanie, które jest «już» fałszem, nie jest zdaniem możliwym. Ujmując rzecz ontologicznie powiemy, iż stanami rzeczy możliwymi są stany fizycznie niezdeterminowane; stan rzeczy, który nie zaszedł, nie jest stanem możliwym. Nic nas tu zatem nie skłania do tego, aby możliwość danego stanu rzeczy eksplikować jako zachodzenie pewnego innego stanu rzeczy — jak to czyniliśmy poprzednio. Jest rzeczą jasną, że mamy tu do czynienia z dwoma różnymi pojęciami możliwości. Eksplikacja obecna chwyta inne intuicje, niż te, które usiłowały uchwycić eksplikacje poprzednie. To, że Księżyc jest większy od Ziemi, nie jest na pewno stanem rzeczy możliwym w sensie rozważanym obecnie; to samo dotyczy statusu modalnego zdania opisującego ów stan. Jako przykłady zdań możliwych przytacza Łukasiewicz prognozy typu „za rok od dziś będę w Warszawie", itp.

Koncepcja Łukasiewicza daleka jest od jasności — zwłaszcza w swej warstwie interpretacyjnej. Nie tylko dlatego, że opiera się na wysoce wątpliwym założeniu filozoficznym, ale i dlatego, że posługuje się pojęciami niedość jasnymi i ścisłymi. Jej precyzacja wymagałaby wprowadzenia wyraźnej relatywizacji do czasu; posługiwania się, w szczególności, zrelatywizowanym do czasu pojęciem prawdziwości: pojęciem zdania prawdziwego w chwili t. Muszę się przyznać, że sens takiego zwrotu pozostaje dla mnie dość zagadkowy. Toteż raz jeszcze podkreślić chciałbym, że koncepcji Łukasiewicza nie traktuję bynajmniej jako zadowalającej kontrpropozycji w stosunku do rozwiązań rozważanych poprzednio. Jest to po prostu przykład propozycji innej. Sformułowanie przekonującej filozoficznie eksplikacji pojęcia możliwości jest zadaniem nie dla logika, lecz filozofa. Jako logik chciałem jedynie zwrócić uwagę na ograniczoność — z filozoficznego punktu widzenia — logicznej aparatury teorio-modelowej i przestrzec przed zbytnimi nadziejami wiązanymi z jej filozoficznymi zastosowaniami. Aparatura ta, stanowiąca podstawowe narzędzie współczesnej semantyki logicznej, znakomicie dostosowane do jej zadań, okazuje się wbrew oczekiwaniom narzędziem niedostosowanym do analizy pewnych — zbliżonych z pozoru do logicznych — zagadnień filozoficznych. Jest na to zbyt uboga treściowo, zbyt «pusta». Przykład pojęć modalnych wydaje się potwierdzać tę sceptyczną opinię.

Filozofia Nauki
Rok I, 1993, Nr 2-3

O tym, czego nie ma
(Na marginesie „Sofisty" Platona)*

Lektura słynnego ustępu z Platońskiego *Sofisty*, poświęconego problemom fałszu i niebytu, nie przestaje być źródłem zdumienia i podziwu. Ten tekst sprzed lat ponad dwóch tysięcy mówi o sprawach, które nie straciły do dziś nic ze swej aktualności, i czyni to w sposób, który prowokuje nie do historycznej analizy, lecz do merytorycznej dyskusji — tak jakby to był artykuł opublikowany na łamach któregoś ze współczesnych filozoficznych czasopism. Nie będąc historykiem filozofii greckiej, tylko filozofem zainteresowanym tymi problemami, które interesowały filozofów greckich, tak właśnie chciałbym potraktować tekst *Sofisty*: jako aktualną wypowiedź filozoficzną w sprawie, którą uważam za ważną i interesującą. Uwagi, które mi ten tekst nasuwa, nie mają więc — i mieć nie mogą — charakteru egzegetyczno-historycznego. Nie są w szczególności próbą możliwie najwierniejszego uchwycenia Platońskiej myśli — odpowiedzi na pytanie, co Platon rzeczywiście miał na myśli, pisząc owe znane ustępy swego późnego dialogu. Jest to raczej próba rekonstrukcji niektórych argumentów dialogu w języku filozofii współczesnej — sformułowania współczesnego wariantu roztrząsanych przez Platona problemów. Kryje się oczywiście za tym przekonanie, że wariant taki odpowiada jednej z możliwych interpretacji Platońskiego tekstu

---

* Artykuł ten zawdzięcza swe powstanie seminarium prowadzonemu wspólnie przez doc. dra Bogusława Wolniewicza i przeze mnie w Instytucie Filozofii UW w roku akademickim 1978/79. Seminarium to, którego temat brzmiał *Pojęcie istnienia w filozofii i logice*, było poświęcone w części lekturze *Sofisty* Platona.

— tekstu, jak wszelkie „wielkie" teksty filozoficzne, niewątpliwie niejednoznacznego, dopuszczającego, co za tym idzie, interpretacje różne.[1]

Owa najdonioślejsza filozoficznie treść *Sofisty*, zawarta w jego części drugiej (w szczególności w ustępach 236 E — 264 B), obejmuje problematykę dość różnorodną — ontologiczną, teoriopoznawczą, semantyczną. To jednak, co dla całości rozważań wydaje się sprawą centralną, ma charakter pewnego fundamentalnego problemu semantycznego. Na nim też skupię się w swych uwagach, starając się problem ów wyrazić w aparacie pojęciowym współczesnej semantyki logicznej. Ale współczesna semantyka logiczna nie stanowi bynajmniej jakiejś teorii jednolitej. Punktem wyjścia dla swoich rozważań uczynić chcę zatem tę semantykę w jej najbardziej klasycznej wersji: w postaci standardowej semantyki teoriomodelowej. Jakim to więc problemom na gruncie tak pojętej semantyki odpowiadać mają sprawy dyskutowane w Platońskim dialogu przez jego dwóch protagonistów: Gościa z Elei i Teajteta?

Problem główny postawiony zostaje w toku poszukiwań adekwatnej definicji sofistyki, które doprowadzają rozmówców w rezultacie do takiej cechy definicyjnej, jak „umiejętność stwarzania złudnych wyglądów". Ale owo pojęcie złudnego wyglądu, jak również pokrewne mu pojęcie fałszywego zdania czy sądu okazują się pojęciami zagadkowymi — bezsensownymi lub wewnętrznie sprzecznymi. „To, żeby coś mogło wyglądać i wydawać się czymś, a nie być, i żeby można mówić coś, ale prawdy nie — to wszystko jest pełne kłopotów zawsze; i dawniej, i teraz. Bo jak się trzeba wyrazić, żeby przyznać istnienie fałszywym zdaniom albo fałszywym sądom i nie popaść w tym wyrażeniu w sprzeczność — to jest strasznie trudna rzecz" (236 E). A ma tak być dlatego, „bo to wyrażenie ośmiela się zakładać, że istnieje to, co nie istnieje. Przecież inaczej nie mógłby powstać fałsz" (237 A). Wyjaśnienie tego stwierdzenia przynosi dalszy fragment dialogu. „Sąd fałszywy to będzie ten, który stwierdza coś przeciwnego niż to, co jest ...Sąd fałszywy stwierdza to, czego nie ma..., stwierdza, że ...jakoś istnieje to, czego w żaden sposób nie ma". A zatem, jak konkluduje Teajtet, „istnieć jakoś muszą niebyty, jeżeli się czasem myli ktoś w czymś" (240 D — 240 E).

Ta konkluzja jednak nie wydaje się nieuchronna. Wynika ona z poprzedzających ją założeń tylko przy pewnej, bynajmniej nie jedynej, ich interpretacji. Nie chodzi tu, jakby sądzić można, o to, że sąd fałszywy — to sąd egzystencjalny, który głosi, że istnieje coś, co faktycznie nie istnieje. Nie wiadomo bowiem wówczas, dlaczego to właśnie fałszywość takiego twierdzenia miałaby pociągać istnienie owych faktycznie nie istniejących bytów. Nieuprawnione wydaje się poza tym ograniczenie wszelkich sądów do sądów egzystencjalnych; spośród innych ich rodzajów co najmniej negacje sądów egzystencjalnych wymagają odrębnego uwzględnienia — jako sądy niesprowa-

1) Taki charakter niniejszych uwag pozwalał oprzeć je na przekładzie oryginalnego tekstu dialogu. W artykule korzystałem z przekładu W. Witwickiego, który konfrontowałem z niemieckim przekładem O. Apelta i angielskim przekładem F. Cornforda. Wszystkie cytaty pochodzą z przekładu W. Witwickiego: *Platon, Sofista. Polityk*, Warszawa 1956.

dzalne do sądów egzystencjalnych. Tak też czyni sam Platon, definiując w innym miejscu swego dialogu zdanie fałszywe za pomocą formuły identycznej ze słynną formułą Arystotelesowską: „i to twierdzenie ...będziemy uważali za fałszywe, które mówi, że nie istnieje to, co istnieje, i to, które mówi, że istnieje to, czego nie ma" (241 A). Ale fałszywość sądu, który głosi, że nie istnieje coś, co faktycznie istnieje, na pewno konkluzji Teajteta o istnieniu niebytów w żaden bezpośredni sposób nie pociąga. Toteż właściwa interpretacja przyjętych założeń wydaje się odmienna. Nie chodzi w niej o to, co dane zdanie *explicite* głosi, lecz raczej o to, co stwierdza, do czego się odnosi, co przedstawia. Różnych dla określenia tego semantycznego stosunku używa się terminów, wspólne jednak dla tych określeń jest to, że owym korelatem semantycznym zdania — tym, co dane zdanie stwierdza, do czego się odnosi, co przedstawia — ma być pewien stan rzeczy lub, ogólniej mówiąc, pewna sytuacja.[2] Tym, co przedstawia zdanie *p*, jest sytuacja, którą nazywać można zwrotem: to, że *p*. Każde zdanie sensowne, każdy sąd, stwierdza zatem określony stan rzeczy, określoną sytuację. Dotyczy to zarówno sądów prawdziwych, jak i fałszywych. Różnica między tymi dwoma rodzajami sądów polegać ma na tym, że sądy prawdziwe przedstawiają sytuacje istniejące, sądy fałszywe — nie istniejące. Prawdziwe zdanie „Teajtet siedzi" stwierdza pewien stan rzeczy istniejący: to, że Teajtet siedzi. W przeciwieństwie do tego stan rzeczy stwierdzany przez fałszywe zdanie „Teajtet lata": to, że Teajtet lata — jest czymś nie istniejącym, jest niebytem. A więc, „istnieć jakoś muszą niebyty, jeżeli się czasem myli ktoś w czymś".

Platon przeprowadza w tym punkcie pewną analogię pomiędzy sądami a wizerunkami przedmiotów. Sądom prawdziwym odpowiadać mają podobizny: wierne wizerunki przedmiotów istniejących. Sądom fałszywym — złudne wyglądy, które określić by można jako niewierne wizerunki przedmiotów istniejących. Ale ową niewierność odwzorowania Platon pojmuje w sposób szczególny. To, co niewiernie odwzorowuje przedmiot istniejący, nie jest w istocie niczym innym jak wiernym wizerunkiem pewnego przedmiotu nie istniejącego. Złudny wygląd — to wygląd czegoś, czego nie ma. I to pojęcie więc, podobnie jak pojęcie fałszywego sądu, „ośmiela się zakładać, że istnieje to, co nie istnieje". I jedno, i drugie odwołuje się do czegoś nie istniejącego.

Wkraczamy w ten sposób, chcąc nie chcąc, w ową niebezpieczną dziedzinę niebytu, przed którą z naciskiem przestrzega dwukrotnie przez Platona przytaczany dwuwiersz Parmenidesa. Rozważania o niebycie mają w dialogu charakter całkowicie ogólny — dotyczą pojęcia niebytu w ogóle, niezależnie od tego, czy jest to niebyt stanowiący odpowiednik fałszywych sądów — a głównym problemem jest zagadnienie natury semantycznej: kwestia sensowności zdań dotyczących niebytu. „Staraj się skupić, jak tylko potrafisz najbardziej, i ...powiedzieć o niebycie cokolwiek poprawnie" (239 B) — wzywa swego rozmówcę Gość z Elei. Ale to właśnie okazuje się niewykonalne. Pro-

2) Posługuję się tu terminologią wittgensteinowską, wprowadzoną przez B. Wolniewicza; por. np. *Sytuacje jako korelaty semantyczne zdań*, „Studia Filozoficzne" 2 (147), 1978.

blem dyskutowany w dialogu ma swój wyraźny odpowiednik na gruncie współczesnej semantyki logicznej, a niektóre stwierdzenia Platońskich rozmówców brzmią zadziwiająco podobnie do znanych sformułowań Russella czy Quine'a. Trudność, na jaką natrafiamy, gdy chcemy „powiedzieć o niebycie cokolwiek poprawnie", ma swe źródło w strukturze logicznej zdań, za pomocą których coś o czymś orzekamy. Ta struktura pojmowana jest w *Sofiście* w ten sam sposób, co na gruncie współczesnego rachunku logicznego (w jego klasycznej wersji). Najprostsze zdanie typu: *a* jest *A* wymaga dla swej sensowności, aby *a* było nazwą indywiduową, tj. nazwą określonego, istniejącego przedmiotu. Nie można więc w szczególności orzekać czegokolwiek o czymś, co nie istnieje. Tak samo stawia sprawę Platon. „Zdanie, jeżeli już jest, to musi czegoś dotyczyć, a nie może dotyczyć niczego" (262 E). „Gdyby nie dotyczyło niczego, to w ogóle nie byłoby zdaniem" (263 C). Nie możemy, co za tym idzie, traktować słowa „niebyt" jako pewnej nazwy indywiduowej — *n*, i wyrażać naszych sądów o niebycie przy pomocy zdań typu: *n* jest *A*. Mówiąc tak bowiem — stwierdzając, że niebyt jest taki czy inny — zakładamy tym samym, że niebyt istnieje, „doczepiamy mu bycie" (239 A). Sprzeczności takiej przy tym, twierdzi Platon, uniknąć nie potrafimy. Konkluzję swoich rozważań o niebycie formułuje Gość z Elei właśnie w postaci zdania: *n* jest *A*. „Niebyt jest bez sensu i nie do wypowiedzenia, i nie do wyrażenia" (239 A). Czyniąc tak, wbrew woli „doczepia mu bycie". Warto może w tym miejscu zauważyć, że ta akurat sprzeczność nie wydaje się nieunikniona. Współczesne odróżnienie języka od metajęzyka pozwala na sformułowanie owej konkluzji jako pewnego stwierdzenia metajęzykowego, mówiącego nie o niebycie, lecz o słowie „niebyt" i jego znaczeniu (tj. o pojęciu niebytu). Jest to jednak przypadek szczególny. Inne wypowiedzi o niebycie trudno traktować jako ukryte wypowiedzi metajęzykowe. Czy można je mimo to sformułować poprawnie?

Za pozytywne rozwiązanie tej kwestii uważa się zwykle Russellowską teorię deskrypcji. Zgodnie z nią pojęciu niebytu odpowiadać winna w języku nie nazwa indywiduowa, lecz deskrypcja — określona: „to, czego nie ma" („to, co nie istnieje"), lub nieokreślona: „coś, czego nie ma" („coś, co nie istnieje"). Są to zresztą zwroty, których odpowiednikami posługują się Platońscy rozmówcy. Otóż zdania z deskrypcją w podmiocie nie wymagają, jak wiadomo, dla swej sensowności istnienia przedmiotu podpadającego pod daną deskrypcję. Możemy więc w takim języku mówić sensownie o tym, czego nie ma, nie „doczepiając" przez to istnienia do niebytu. Jeszcze radykalniejszym rozwiązaniem jest propozycja Quine'a wyrugowania z języka wszelkich nazw i deskrypcji na rzecz predykatów jako jedynych terminów pozalogicznych. Pojęciu niebytu odpowiadałby tu więc predykat *N*: „nieistniejący" („nie istnieje"). Ponieważ sensowność predykatu nie wymaga tego, aby istniało cokolwiek, co pod ten predykat podpada, nic nie stoi na przeszkodzie formułowaniu sensownych, a nawet prawdziwych zdań zawierających predykat *N*, a więc zdań mówiących poprawnie i bez sprzeczności o czymś, co nie istnieje. Innego jeszcze rozwiązania tego problemu dostarczają systemy tzw. logiki wolnej (*free logic*), stanowiące pewne modyfikacje klasycznego rachunku logicznego. Rozszerzają one kategorię nazw indywiduowych o tzw. nazwy puste — wyrażenia mogące sensownie figurować jako podmioty zdań typu: *a* jest *A*, ale nie denotujące żadnych przedmiotów istniejących. „Niebyt" może być zatem jedną z takich nazw pustych — *n*. Możemy w związku z tym orzekać o nim sensownie pewne własności za pomocą zdań: *n* jest *A*, gdyż zdania te nie implikują tu żadnych konsekwencji egzystencjalnych.

Wydawać by się mogło po tym, co zostało powiedziane, że Platoński problem poprawnego mówienia o niebycie znalazł na gruncie współczesnej semantyki logicznej zadowalające rozwiązanie: „niebyt jest do wypowiedzenia" — i to na różne sposoby. Wbrew pozorom jednak żaden z tych sposobów nie może być uznany za adekwatny. Poprawność, jaką gwarantują naszym wypowiedziom o niebycie, jest poprawnością formalną jedynie, a nie merytoryczną. Pozwala ona mówić o tym, czego nie ma, w sposób sensowny, ale nie pozwala wyrazić tego, co by się wyrazić chciało. Jest tak, najogólniej mówiąc, dlatego, że — na gruncie wszystkich niemal proponowanych rozwiązań — wartość logiczna zdań o niebycie zostaje z góry przesądzona w sposób arbitralny. Jeśli, zgodnie z rozwiązaniem Russella, termin „niebyt" potraktujemy jako deskrypcję: „to, co nie istnieje", będzie to oczywiście deskrypcja pusta, skoro żaden przedmiot *x* nie spełnia warunku: *x* nie istnieje. W konsekwencji każde zdanie postaci: to, co nie istnieje, jest *A*, które przypisuje niebytowi jakikolwiek predykat prosty, będziemy musieli uznać za zdanie fałszywe. Jeśli natomiast, zgodnie z propozycją Quine'a, ograniczymy się wyłącznie do predykatu „nieistniejący" — *N*, jako odpowiednika terminu „niebyt", zdania orzekające o niebycie jakąś własność *A* przybiorą postać zdań ogólnych: dla każdego *x*, jeśli *x* jest *N*, to *x* jest *A*, lub szczegółowych: dla pewnego *x*, *x* jest *N* i *x* jest *A*. Ponieważ z założenia żaden przedmiot *x* nie spełnia warunku: *x* jest *N*, wszystkim zdaniom danego typu wypadnie przyznać tę samą wartość logiczną: dowolne zdanie ogólne będzie prawdą, dowolne zdanie szczegółowe — fałszem, niezależnie od tego, jaką to własność *A* przypisują one temu, co nie istnieje.

Rozwiązania oparte na systemach logicznych z nazwami pustymi nie mają tak jednolitego charakteru, bo też systemy te różnią się między sobą co do swych założeń semantycznych.[3] Jeśli — jak to ma miejsce w większości takich systemów — nie przyporządkowują one nazwom pustym żadnych przedmiotów jako ich denotacji, otrzymujemy konsekwencje podobne do wyżej wymienionych: wszystkie zdania typu: *n* jest *A* zaliczone zostają w sposób arbitralny do tej samej klasy — wszystkie zostają uznane za zdania fałszywe lub wszystkie za zdania o nieokreślonej wartości logicznej (a więc ani prawdziwe, ani fałszywe). Z odmienną sytuacją mamy do czynienia tylko w tych systemach, które prócz dziedziny przedmiotów istniejących wprowadzają dziedzinę przedmiotów nie istniejących, lecz możliwych, i nazwom pustym przypo-

3) Por. np. następujące ujęcia: K. Lambert i B. van Fraassen, „Meaning Relations, Possible Objects, and Possible Words"; P. Woodruff, „Logic and Truth Value Gaps"; oraz D. Scott, „Advice on Model Logic", [w:] *Philosophical Problems in Logic* (wyd. K. Lambert), Dordrecht 1970.

rządkowują takie właśnie «byty» jako ich denotacje. Ponieważ interpretacja poszczególnych predykatów zostaje tu określona również w dziedzinie owych przedmiotów możliwych, zdania typu: *n* jest *A* mogą zostać zróżnicowane pod względem swej wartości logicznej: dla pewnych predykatów *A* mogą przybierać wartość prawdy, dla innych — fałszu. A zatem pewne własności mogą niebytowi przysługiwać, a inne — nie. Konstrukcja ta jednak, opierająca się na zagadkowym pojęciu przedmiotów nie istniejących, budzi zasadnicze wątpliwości filozoficzne. Niezależnie zresztą od tych wątpliwości generalnych przypisywanie nazwie „niebyt" jako jej denotacji pewnego przedmiotu możliwego wydaje się nieuprawnione. W przeciwieństwie do takich nazw pustych, jak „Pegaz", sens nazwy „niebyt" wyklucza, w moim przekonaniu, taką ewentualność: niebyt nie istnieje z konieczności, na mocy samego swego znaczenia. I to więc stanowisko nie przynosi zadowalającego rozwiązania Platońskiego dylematu. Nadal pozostaje on problemem otwartym.

Swoista paradoksalność pojęcia niebytu przejawia się również w fakcie wyraźnie stwierdzonym przez Platona, a najdobitniej sformułowanym przez Quine'a.[4] Polega on na tym, iż właściwą odpowiedzią na pytanie „Co istnieje?", może być jedynie odpowiedź „Wszystko". Jej negacja bowiem, równoważna stwierdzeniu „Jest coś, co nie istnieje", jest wewnętrznie sprzeczna. Fakt ten znajduje potwierdzenie na gruncie współczesnego klasycznego rachunku logicznego. Predykat „*x* istnieje", który utożsamiony być może z predykatem „*x* jest czymś", definiowany bywa w elementarnym rachunku logicznym za pomocą warunku „*x* jest identyczny z czymś": *x* istnieje, gdy dla pewnego y, $x = y$. W konsekwencji, w rachunku tym zdanie ogólne „dla wszelkiego *x*, *x* istnieje" okazuje się tautologią, a jego negacja „dla pewnego *x*, *x* nie istnieje" — twierdzeniem kontradyktorycznym. W niektórych systemach tzw. logiki wolnej unika się co prawda tej konsekwencji, ale czyni się to za cenę założenia co najmniej równie wątpliwego: wprowadzenia dziedziny przedmiotów nie istniejących. Rację ma więc istotnie Gość z Elei, stwierdzając, że „to wszystko jest pełne kłopotów zawsze; i dawniej, i teraz".

Kłopotów tych rozmówcy z Platońskiego dialogu starają się uniknąć, wprowadzając na miejsce dotychczasowego pojęcia niebytu, rozumianego jako „przeciwieństwo bytu" (258 E,...) takie pojęcie niebytu, które by było wolne od paradoksów pojęcia poprzedniego, a zarazem pozwoliło zdać sprawę z semantycznej natury fałszywych zdań i sądów. Mówiąc najogólniej, jest to, w przeciwieństwie do «absolutnego» pojęcia niebytu rozważanego do tej pory, pojęcie o charakterze «relatywnym»: pojęcie niebytu, który „uczestniczy w bycie" (260 D) — niebytu, który istnieje. Teza, której broni Gość z Elei, głosi, „że z jednej strony to, co nie istnieje, istnieje jednak jakoś, i że to, co istnieje — jednak w pewnym sposobie nie istnieje" (241 D). Do konkluzji takiej dochodzą obaj

4) W. Quine, „O tym, co istnieje", [w:] *Z punktu widzenia logiki*, Warszawa 1969.

dyskutanci po obszernej analizie pojęcia bytu i jego rodzajów. Ten słynny ustęp, poświęcony problemowi związków logicznych pomiędzy podstawowymi rodzajami bytu, był od dawna przedmiotem uwagi filozofów i logików i doczekał się wnikliwych analiz.[5] Tutaj ograniczę się tylko do tego, co pozostaje w bezpośrednim związku z kluczowym dla naszych rozważań pojęciem niebytu. Dyskusja Gościa z Elei z Teajtetem koncentruje się na pięciu głównych rodzajach: byt, spoczynek, ruch, tożsamość, różność, a jej konkluzją jest przekonująco brzmiące stwierdzenie, iż, choć są to rodzaje różne, niektóre z nich łączą się ze sobą, a niektóre nie. Tak np. ruch łączy się z bytem, a nie łączy ze spoczynkiem, jeśli przez łączenie się dwóch rodzajów rozumieć, jak tego chce Platon, fakt, że coś może uczestniczyć zarazem w jednym i drugim. W całej tej dyskusji odczuwa się jednak szczególnego rodzaju zmaganie się z trudnościami związanymi z pewnymi podstawowymi pojęciowymi dystynkcjami. Są to trudności odróżnienia kategorii takich, jak: „byt" — rozumiany jako nazwa indywiduowa — *b*, ogółu (zbioru lub sumy) rzeczy i jako predykat uniwersalny — *B*, orzekany o poszczególnych rzeczach; słowo „jest" — rozumiane jako spójka: jest (taki a taki) i jako predykat identyczności: jest identyczny (z tym a tym); a wreszcie „rodzaj" — rozumiany jako własność i jako relacja. Ta ostatnia dystynkcja jest z pewnością dla owych rozważań o rodzajach bytu najdonioślejsza. Powiedzielibyśmy bowiem dziś, że niektóre z porównywanych tu rodzajów — to własności, inne — to dwuczłonowe relacje. A jeśli tak jest istotnie, to trudno je porównywać — pod niektórymi przynajmniej z rozważanych względów, np. pod względem ich zakresu. Platon, któremu brak wyraźnej świadomości tej zasadniczej ontologicznej różnicy, ujmuje jednak w konkretnych przypadkach sprawę właściwie i dochodzi z reguły do konkluzji trafnych. Porównując np. pod względem zakresu własność A z relacją R, porównuje w istocie własność A z pewną inną własnością B, odpowiadającą temu, co we współczesnej terminologii nazywa się dziedziną owej relacji R. Porównane więc zostają warunki tego samego logicznego typu: x jest A, oraz dla pewnego y, x jest R od y. Przykładem mogą być zwroty: x jest w ruchu oraz x jest tożsame z czymś.

Otóż podobna konstrukcja zastosowana zostaje przy eksplikacji owego uwolnionego od sprzeczności i paradoksów pojęcia niebytu. Byt i niebyt określone zostają za pomocą analogicznych formuł. Byt — to tyle co: bycie czymś (jakimś), niebyt — tyle co: niebycie czymś (jakimś). Dokładniej: x jest bytem („uczestniczy w bycie"), jeśli x jest jakieś, ma jakąś własność, a więc jeśli dla pewnego A, x jest A; x jest niebytem („uczestniczy w niebycie"), jeśli x nie jest jakieś, nie ma jakiejś własności, a więc jeśli dla pewnego A, x nie jest A.[6] Owa relacja «niebycia» nazywana jest również w dialogu relacją różności; to, że x nie jest A, wyrażane jest przez zwrot: x jest różne od A. Niebyt zatem, w sformułowaniu Gościa z Elei, to wszystko to, co jest od czegoś różne. To

5) Por. w szczególności B. van Fraassen, „Logical Structure in Plato's 'Sophist'", *The Review of Metaphysics* 22, 1969.

6) Podobnie odczytuje ten tekst *Sofisty* B. van Fraassen, wyd. cyt.

szerokie rozumienie niebytu przeciwstawione zostaje rygorystycznemu rozumieniu dotychczasowemu, traktującemu niebyt jako „przeciwieństwo bytu". W rozumieniu tamtym *x* jest niebytem tylko wtedy, gdy *x* nie jest niczym, gdy *x* nie ma żadnej własności; a więc gdy dla wszelkiego *A*, *x* nie jest *A*. Tak pojmowany niebyt, rzecz jasna, nie istnieje. W przeciwieństwie do tego, istotną cechą owego liberalnego pojęcia niebytu jest to, że „niebyt... jest zgoła nie mniej niż sam byt czymś istniejącym" (258 B). Coś, co nie ma pewnej własności, ma zawsze jakąś własność inną; coś, co nie jest jakimś *A*, jest zawsze jakimś *B* — a zatem jest pewnym bytem. Argumentacja Platona w tym punkcie opiera się na założeniu, iż coś, co nie jest *A*, jest tym samym nie-*A*, a bycie nie-*A* jest w równym stopniu bytem, co bycie *A*. „Nazwy przybierające na początku «nie»... oznaczają coś z rzeczy innych niż... te rzeczy, których dotyczą nazwy następujące po przeczeniu" (257 C). Myśl ta poparta zostaje szeregiem przykładów. „W tym, co różne, istnieje jakaś jego cząstka przeciwstawiona pięknu" — „to, co... nazywamy niepięknym". „To niepiękne, to jest przeciwstawienie jednego bytu innemu bytowi" (257 E). „I tak samo trzeba przyjąć, że to, co niesprawiedliwe w stosunku do tego, co sprawiedliwe, zgoła nie więcej istnieje jedno od drugiego". „I tak samo powiemy w innych wypadkach" (258 A). „Natura tego, co różne, jest rozdrobniona na wszystkie byty... i każdą jej cząstkę przeciwstawioną bytowi ośmieliliśmy się nazwać niebytem" (258 E). „Niebyt jest rodzajem jednym z wielu i jest rozsiany po wszystkich bytach" (260 B).

Jeśli pominąć sprawę swoistej terminologii, pogląd Platona nie stracił wiele ze swej aktualności. Jego stanowisko wydaje się przekonujące i dziś — choć z punktu widzenia współczesnej semantyki logicznej wymaga pewnych zastrzeżeń i ograniczeń. Gdy *A* jest predykatem, odnoszącym się do pewnej własności (czy rodzaju) przedmiotów, takim samym predykatem jest jego negacja — nie-*A*. Inaczej pod pewnym względem jest tylko w tym przypadku, gdy *A* jest predykatem uniwersalnym. Wówczas nie-*A* jest predykatem pustym, a więc odnoszącym się do własności, która niczemu nie przysługuje. Powstaje zatem pytanie, czy ona również uważana być może za rodzaj bytu. A może wtedy jedynie zmuszeni jesteśmy odmówić tego miana, gdy nie-*A* jest predykatem kontradyktorycznym — odnoszącym się do własności, która nie tylko faktycznie niczemu nie przysługuje, ale i niczemu przysługiwać nie może? Inna trudność powstaje wtedy, gdy bierzemy pod uwagę zakres danego predykatu rozumiany jako zbiór w sensie teoriomnogościowym. Aby zakres predykatu nie-*A* był zbiorem dobrze określonym, nie może to być, jak wiadomo, zbiór czegokolwiek, co nie jest *A*. Musi to być zawsze podzbiór pewnego ustalonego, traktowanego jako *universum*, zbioru *U*: a więc zbiór tych elementów *U*, które nie są *A*. Toteż w rozważaniach tych będziemy z reguły takie *universum* milcząco zakładali.

Na gruncie wprowadzonych w ten sposób pojęć bytu i niebytu oczywista staje się teza broniona przez Gościa z Elei: „to, co nie istnieje, istnieje jednak jakoś... to, co istnieje, jednak w pewnym sposobie nie istnieje". Czy tak określone pojęcie niebytu pozwala również zdać sprawę z tego, co stanowi główny przedmiot rozważań

Platońskich rozmówców — z istoty fałszywych zdań i sądów? Rozmówcy dochodzą w tej sprawie do odpowiedzi pozytywnej. Określiwszy strukturę zdań najprostszego typu (261 D — 262 E), charakteryzują zdanie fałszywe jako zdanie, które o danym przedmiocie „mówi coś innego, niż jest"; „więc o czymś, czego nie ma, mówi tak, jakby to istniało" — dosłownie: „mówi niebyty jako byty" (263 B). Zakładają przy tym, że idzie tu o niebyt w tym właśnie niekontrowersyjnym sensie, który przyjęli uprzednio. Czy tak jest naprawdę — trudno stwierdzić w sposób stanowczy. Tekst daleki jest od jednoznaczności i dopuszcza interpretacje różne. Bronić chcę jednak tezy, iż jest wśród nich interpretacja, która powyższy warunek spełnia. Jest to, w moim przekonaniu, interpretacja bliska tej, jaką dostarcza współczesna semantyka teoriomodelowa. Wyjaśnia ona fałszywość danego zdania, nie odwołując się do żadnych nie istniejących przedmiotów — ani sytuacji, ani rzeczy. Eksplikację tę najkrócej można by ująć tak: zdanie „*a* jest *A*" jest fałszem wtedy, gdy nie jest tak, że *a* jest *A*; a to z kolei ma miejsce wtedy, gdy *a* jest nie-*A*. W tym ostatecznym sformułowaniu mówimy wyłącznie o przedmiotach istniejących: o indywiduum *a* i własności (rodzaju) nie-*A*, która jest niebytem tylko w sensie Platońskim (a więc tym samym jest jednym z rodzajów bytu). Zilustrujmy to na przykładzie rozważanym przez Platońskich rozmówców. To, że zdanie „Teajtet lata" jest fałszem, możemy wyrazić po prostu za pomocą stwierdzenia: Teajtet nie-lata, przypisującego pewnej istniejącej osobie: Teajtetowi, pewien «relatywny» jedynie, a więc istniejący niebyt: nie-latanie. Eksplikacja taka odpowiada w istocie definicji fałszywości zdań atomowych formułowanej w semantyce teoriomodelowej. Główna różnica dotyczy ontologicznego charakteru owego bytu odpowiadającego predykatowi nie-*A*, który w klasycznej semantyce tego typu utożsamiany bywa z reguły ze zbiorem, co nie stanowi z pewnością adekwatnej interpretacji Platońskiego «rodzaju». To, że w podanej eksplikacji mowa tylko o zdaniach najprostszego typu, tzw. zdaniach atomowych, nie stanowi jakiegoś istotnego ograniczenia. Definicja prawdy i fałszu ma w semantyce współczesnej, jak wiadomo, postać definicji rekurencyjnej, która prawdziwość i fałszywość dowolnego zdania sprowadza w rezultacie do prawdziwości czy fałszywości zdań atomowych.

Chciałbym na koniec zwrócić uwagę na pewne generalne rysy takiego ujęcia. Jest to ujęcie wolne od tych — wspomnianych na wstępie — założeń, które stanowiły główne źródło dalszych paradoksów i trudności. Jednym z tych założeń była równoważność, którą wyrazić można w sposób następujący: dane zdanie fałszywie opisuje to, co istnieje, wtedy i tylko wtedy, gdy zdanie to prawdziwie opisuje coś, co nie istnieje. Rozwiązanie proponowane takiej równoważności nie zakłada. Zdanie fałszywe, podobnie jak zdanie prawdziwe, odnosi się do tego samego, istniejącego świata. A to, że ów świat opisuje fałszywie, nie polega bynajmniej na tym, że prawdziwie opisuje jakiś świat nieistniejący. Podobnie jak niewierność wizerunku jakiegoś przedmiotu rzeczywistego nie sprowadza się, wedle takiego stanowiska, do tego, że jest to wierny wizerunek jakiegoś przedmiotu fikcyjnego. W rezultacie nie ma potrzeby wprowadzania jakichś sytuacji nie istniejących jako odpowiedników zdań fałszywych. Co więcej, nie zakłada się tu w ogóle tego, jakoby jakiekolwiek zdanie — prawdziwe lub fałszywe — przedstawiało, czy stwierdzało, jakąś sytuację. Świat, który takie zdanie opisuje, pojmowany jest raczej jako pewien układ rzeczy — konkretnych lub abstrakcyjnych przedmiotów, takich jak indywidua, zbiory, własności, relacje itp. Poszczególne składniki zdania o charakterze nazwowym (nazwy indywiduowe, predykaty) niektóre z tych przedmiotów denotują; natomiast zdanie jako całość, nie będąc nazwą, niczego też nie denotuje, ani nie przedstawia. To, co dane zdanie stwierdza, wyrażane jest przez samo to zdanie, a nie przez nazwę jakiejkolwiek sytuacji.

W języku takiej właśnie «semantyki rzeczy» próbowałem wyrazić konkluzję Platońskich rozważań, zdając sobie oczywiście sprawę z tego, że nie jest to z pewnością interpretacja jedyna; co więcej — że nie jest to, być może, interpretacja w stosunku do Platońskiej myśli najwierniejsza. Nie sądzę też, aby całkowita rezygnacja z «semantyki sytuacji» była możliwa. Idzie raczej o to, aby — zgodnie z zasadą Ockhama — uciekać się do niej tylko tam, gdzie jest naprawdę niezbędna. A nie wydaje mi się, aby to miało miejsce w przypadku problemu prawdziwości i fałszywości zdań. Pojęcie sytuacji (faktu, stanu rzeczy) jest na pewno potrzebne tam, gdzie się o sytuacjach (faktach, stanach rzeczy) wyraźnie mówi. Mówi się zaś o nich tak nie wtedy, kiedy się używa dowolnych zdań, tylko wtedy, kiedy się używa odpowiednich nazw — nie zdania p, tylko nazwy: to, że p. Uzasadnione wydaje się przekonanie, że są konteksty, w których nazwy takie występują w sposób istotny. Semantyka tego rodzaju wypowiedzi wymaga więc odpowiedniej «sytuacyjnej» ontologii. Semantyka zdań wolnych od tego typu zwrotów może się, w moim przekonaniu, bez takiej ontologii obejść. Przykład *Sofisty* — odczytany tak jak wyżej — ma potwierdzać takie przypuszczenie.

Filozofia Nauki
Rok I, 1993, Nr 2-3

Zasada wyłączonego środka a zagadnienie idealizmu

Pod nazwą zasady wyłączonego środka kryją się twierdzenia różnego rodzaju i różnej treści: ontologiczne i semantyczne, logiczne i metalogiczne. W swej wersji ontologicznej (lub logicznej) zasada ta głosi, mówiąc swobodnie, że „z dwóch stanów rzeczy sprzecznych zawsze jeden istnieje", w wersji semantycznej (lub metalogicznej) stwierdza, że „z dwóch zdań sprzecznych zawsze jedno jest prawdziwe". Niezależnie jednak od sposobu jej sformułowania, zasadzie tej przypisuje się określony ładunek filozoficzny, sprawiający, iż jest to twierdzenie nieobojętne wobec pewnych podstawowych kontrowersji filozoficznych. Zasada ta, charakterystyczna dla pewnych stanowisk filozoficznych, ma być niezgodna z innymi. Najlepiej może znana jest jej rola w sporze między klasycznym a intuicjonistycznym (czy konstruktywistycznym) stanowiskiem w filozofii matematyki. Stanowiąc znamienny element koncepcji klasycznej, odrzucana jest przez wszystkie kierunki intuicjonistyczne (i konstruktywistyczne). Jednakże nie na tym, wielokrotnie dyskutowanym problemie, chciałbym się skoncentrować w tych rozważaniach. Poświęcone one będą roli zasady wyłączonego środka w innym sporze filozoficznym: w fundamentalnej filozoficznej kontrowersji między realizmem a idealizmem, w szczególności tak zwanym idealizmem obiektywnym (logicznym lub transcendentalnym). Sprawa ta stała się jednym z głównych przedmiotów filozoficznych dociekań Kazimierza Ajdukiewicza, który poświęcił jej wiele uwagi i namysłu zarówno w swoich pracach dawniejszych, jak i powojennych. Bezpośrednio zagadnieniu temu poświęcona jest w całości rozprawa z roku 1937 pt. „Problemat transcendentalnego idealizmu w sformułowaniu semantycznym", ale sprawa ta powraca również w jego publikacjach późniejszych, w szczególności w rozprawie „Epistemologia i semiotyka" z roku 1948 i w polemicznej wypowiedzi „W sprawie artykułu prof. A. Schaffa o moich poglądach filozoficznych” z roku 1953.[1]

Reprezentowane w tych pracach stanowisko przyjmiemy jako punkt wyjścia dla dalszych rozważań, których głównym celem jest ustosunkowanie się do przedstawionej tam argumentacji. Argumentacja ta zmierza do okazania, że zasada wyłączonego środka jest w sposób istotny związana ze stanowiskiem realistycznym. Zasada ta ma być nie do pogodzenia z tezą idealizmu. Co więcej, samo odrzucenie tej zasady ma świadczyć już o zajęciu stanowiska idealistycznego.

Przypomnijmy w wielkim skrócie główny tok myśli autora, opierając się w tym przede wszystkim na jego podstawowej pracy z roku 1937.[2] Przedmiotem analizy autorskiej jest ta odmiana idealizmu, której sztandarowa teza głosi, iż „świat rzeczywisty jest tylko korelatem podmiotu transcendentalnego”. W jednej z wersji tej tezy, ów zagadkowy podmiot transcendentalny utożsamiony zostaje z systemem tak zwanych norm transcendentalnych — reguł (czy kryteriów) wyróżniających określony zbiór sądów. W wersji tej, teza idealizmu sprowadza się do twierdzenia, że prawdą jest tylko sąd podyktowany przez normy transcendentalne. Tę właśnie epistemologiczną tezę poddaje Ajdukiewicz w swym artykule semantycznej parafrazie. Odwołuje się w tym celu do swojej koncepcji języka, zgodnie z którą język pojęty jako zbiór wyrażeń o ustalonych znaczeniach utożsamiony zostaje ze zbiorem wyrażeń o ustalonych związkach bezpośredniego wynikania pomiędzy zdaniami, wyznaczonych przez (dedukcyjne, aksjomatyczne i empiryczne) reguły bezpośredniego wynikania zwane regułami języka (lub regułami sensu). Otóż w semantycznej parafrazie tezy idealizmu, normom transcendentalnym odpowiadać mają reguły języka nauk przyrodniczych (a raczej wszelkich nauk empirycznych). Sądy podyktowane przez owe normy — to nic innego jak zdania wyznaczone przez reguły języka nauk empirycznych, czyli tezy tego języka. „W tej semantycznej parafrazie naczelne twierdzenie transcendentalnego idealizmu, głoszące, że prawdziwe są tylko sądy podyktowane przez normy transcendentalne, przełoży się na zdanie głoszące, że w języku nauk (empirycznych) prawdziwe są tylko [...] te zdania, które są tezami tego języka”.

Ta interpretacja zagadnienia idealizmu bliska jest, w przekonaniu autora, intuicji, które mieli transcendentalni idealiści. Zarazem formułuje ona to zagadnienie w sposób, który pozwalać ma na proste jego rozwiązanie przez odwołanie się do pewnych wyników metalogiki, rozumianej jako teoria systemów dedukcyjnych. Jest to możliwe dlate-

1) O tym, jaką wagę przywiązywał Ajdukiewicz do tego nurtu swej twórczości, świadczy m.in. jego wypowiedź z 1953 r. Pisząc w niej o swej pracy o transcendentalnym idealizmie, stwierdza co następuje. „Praca ta jest w moim osobistym rozwoju punktem zwrotnym, od którego poczynając wszystkie niemal moje filozoficzne poczynania są zwrócone przeciw idealizmowi. Ten front zachowują też wszystkie moje publikacje powojenne. Wszystkie one atakują idealizm z różnych stron”.
Wszystkie cytowane w artykule prace Ajdukiewicza zamieszczone są w wyborze jego pism *Język i poznanie*, Warszawa 1960 (t. I), 1965 (t. II).

2) Wszystkie cytaty z pism Ajdukiewicza przytoczone bez podania źródła pochodzą z tej właśnie pracy.

go, że zgodnie ze wspomnianą koncepcją Ajdukiewicza, język pojęty jako zbiór wyrażeń o ustalonych znaczeniach jest tym samym pewnym systemem dedukcyjnym, gdyż jest zbiorem wyrażeń, dla których ustalone zostały związki bezpośredniego wynikania między zdaniami, a w tym właśnie upatruje się istotę systemu dedukcyjnego. Do tak rozumianego języka stosuje się więc bezpośrednio wszystko to, co metalogika mówi o systemach dedukcyjnych. Na dwa jej twierdzenia powołuje się Ajdukiewicz w postulowanym przez siebie rozwiązaniu zagadnienia idealizmu. Jedno z nich to twierdzenie o niezupełności wszelkich systemów dedukcyjnych zawierających system arytmetyki. Ponieważ język nauk empirycznych obejmuje jako swoją część język arytmetyki, system dedukcyjny wyznaczony przez reguły języka nauk empirycznych jest systemem niezupełnym: istnieją w tym języku pary zdań sprzecznych, z których żadne nie jest tezą systemu. Drugie z owych twierdzeń metalogiki — to, będąca przedmiotem naszej dyskusji, metalogiczna zasada wyłączonego środka. Powołując się na słynną pracę Tarskiego o pojęciu prawdy, traktuje Ajdukiewicz tę zasadę jako konsekwencję klasycznej definicji prawdy. Orzeka ona, że z dwóch zdań sprzecznych jedno musi być zdaniem prawdziwym. Prostym wnioskiem z tych twierdzeń jest konkluzja głosząca, że w systemach niezupełnych nie każda prawda jest tezą. Więc i w systemie wyznaczonym przez reguły języka nauk empirycznych nie każde zdanie prawdziwe jest tezą tego języka. Stanowi to bezpośrednie zaprzeczenie naczelnego twierdzenia idealizmu w jego wersji semantycznej.

To rozwiązanie zagadnienia idealizmu nasuwa wiele wątpliwości i zastrzeżeń, o których sam autor lojalnie wspomina w swym artykule. Zbyt rygorystyczne, na przykład, wydać się może ograniczenie systemów wyznaczanych przez reguły języka do systemów dedukcyjnych z finitystycznym pojęciem tezy (czy konsekwencji) itp. Jednak najistotniejsze, z uwagi na interesujący nas tu problem, wątpliwości — to te, które dotyczą waloru metalogicznej zasady wyłączonego środka dla języków (systemów) niezupełnych. Mówiąc w tym kontekście o językach niezupełnych, charakteryzuje je Ajdukiewicz jako „języki zawierające tzw. terminy o nieostrym znaczeniu". Odwołując się do klasycznych przykładów takich terminów („młody", „łysy" itp.), określa je autor w sposób ogólny jako „terminy, których wstawienie w kontekst rozstrzygalny wedle reguł języka może uczynić ten kontekst wedle reguł języka nierozstrzygalnym". Takim terminem będzie, na przykład, fikcyjna nazwa „abra" wprowadzona do języka polskiego za pomocą dwóch aksjomatów: „każdy abra jest człowiekiem" i „każdy mężczyzna jest abra". Zdanie „Ewa jest abra", podobnie jak jego negacja, służyć może jako przykład kontekstu nierozstrzygalnego wedle reguł tak wzbogaconego języka polskiego. Otóż zasada wyłączonego środka każe jedno lub drugie z tych dwu zdań uznać za prawdziwe. „Czy to orzeczenie — pyta Ajdukiewicz — nie wydaje się paradoksalne?" „Jakże może być prawdziwe zdanie zaliczające Ewę do zakresu nazwy «abra» lub zdanie wykluczające ją z tego zakresu, skoro znaczenie, jakie nazwa ta otrzymała, jest tak ułomne, iż zakresu tej nazwy nie ustala!" „Czyż nie jesteśmy raczej skłonni przyjąć, że z obu tych sprzecznych zdań żadne nie jest prawdziwe?" Odpowiadając na tę wątpliwość, Ajdukiewicz zwraca uwagę na fakt, iż metalogiczna zasada wyłączonego środka równoważna jest na gruncie klasycznej definicji prawdy logicznej (ontologicznej) zasadzie wyłączonego środka. „Kto by pod wpływem wyłożonych wątpliwości zdecydował się odrzucić metalogiczną zasadę wyłączonego środka, musiałby albo odrzucić zasadę logiczną lub też zrezygnować z klasycznego pojęcia prawdy, zastępując je pojęciem tezy języka. Obranie drugiej z obu tych dróg byłoby równoważne z zajęciem stanowiska idealistycznego".

Ten wywód zamyka Ajdukiewiczowską krytykę transcendentalnego idealizmu. Jak ustosunkować się do zawartej w nim argumentacji? Ajdukiewicz wyraźnie zdaje sobie sprawę z jej szkicowości i niekompletności: „uwagi te rzucam na tym miejscu w formie zupełnie surowej, zdając sobie sprawę z tego, że poruszone kwestie zasługują na głębszą analizę". Wraca też do tej sprawy w późniejszych swych publikacjach. W jednej z nich[3] uchyla stanowczo wspomniane wątpliwości co do waloru metalogicznej zasady wyłączonego środka w stosunku do kontekstów nierozstrzygalnych, pisząc, że przytaczana na ich poparcie argumentacja „popełnia ten zasadniczy błąd, iż miesza niemożność rozstrzygnięcia między dwoma zdaniami sprzecznymi z tym, że żadne z tych dwu zdań sprzecznych nie jest prawdziwe". Jednak i tutaj dodaje, że „rzecz wymagałaby obszerniejszego omówienia". Sprawie tej — ważnej nie tylko w związku z krytyką idealizmu — warto się przyjrzeć w sposób bardziej wnikliwy i szczegółowy.

Zwróćmy przede wszystkim uwagę na pewną niejasność pojęcia niezupełności stosowanego do empirycznych systemów językowych. Istotną rolę pełnią w nich empiryczne reguły językowe, nakazujące — mówiąc swobodnie — uznawanie pewnych zdań wobec pewnych danych doświadczenia. Odpowiedników takich reguł nie znajdujemy, ściśle biorąc, w metalogicznym pojęciu systemu dedukcyjnego. To, co w rzeczywistości odpowiada empirycznemu systemowi językowemu, określić należałoby raczej jako zbiór tez podyktowanych przez reguły języka na gruncie ogółu danych doświadczenia. System taki jest niezupełny, jeśli wśród tak rozumianych tez brak zarówno jakiegoś zdania, jak i jego negacji. Upatrując źródło niezupełności języka nauk empirycznych w nieostrym znaczeniu pewnych terminów empirycznych (takich jak „młody" czy „łysy"), stawiamy sprawę inaczej niż wtedy, gdy niezupełność tę przypisujemy faktowi zawierania się w tym języku niezupełnego systemu arytmetyki. Wiemy, że rozszerzenie tego ostatniego do systemu zupełnego jest logicznie niemożliwe. W przeciwieństwie do tego, nie wydaje się rzeczą logicznie niemożliwą takie «zaostrzenie» znaczenia terminów empirycznych, aby wzbogacony w ten sposób język empiryczny stał się systemem zupełnym. Jeśli okazuje się to niewykonalne, to jest to raczej niemożliwość «fizyczna». Ze względu na naturę rzeczywistości empirycznej i charakter naszych procedur językowych, empiryczne reguły języka ustalające znaczenie terminów empirycznych wyznaczają ich zakres w sposób niejednoznaczny, pozostawiając zawsze pewien obszar nieokreśloności. Niezależnie jednak od takich czy innych możli-

3) „Zmiana i sprzeczność" z 1948 r.

wości «zaostrzenia» sensu terminów empirycznych, stwierdzić trzeba wyraźnie, że «zaostrzenie» to stanowi zawsze zmianę dotychczasowego znaczenia tych terminów. Język w ten sposób wzbogacony jest po prostu językiem innym. Prawdą pozostaje więc stwierdzenie, że istniejący język nauk empirycznych jest językiem niezupełnym — nie tylko dlatego, że obejmuje niezupełny język arytmetyki, ale i dlatego, że zawiera terminy empiryczne o nieostrym znaczeniu. Problem stosowalności metalogicznej zasady wyłączonego środka do zdań tego języka pozostaje wobec tego w pełni aktualny.

Mimo zgłaszanych wątpliwości, Ajdukiewicz skłonny jest, jak widzieliśmy, rozstrzygać go pozytywnie, uważając, że inna decyzja zmuszałaby do rezygnacji z założeń, z których rezygnować się nie powinno. Należą do nich logiczna zasada wyłączonego środka i klasyczna definicja prawdy, których konsekwencją ma być zasada metalogiczna. Jak motywuje Ajdukiewicz decyzję utrzymania w mocy obu tych założeń? Pewne uwagi na temat logicznej zasady wyłączonego środka znaleźć możemy we wspomnianym artykule z roku 1953. Powstaje przede wszystkim pytanie, jak tę — logiczną, czy ontologiczną — zasadę należałoby sformułować. Pojmowana jako ogólne twierdzenie ontologiczne wyrażana jest zwykle za pomocą powiedzeń takich jak: „z dwóch stanów rzeczy sprzecznych zawsze jeden istnieje", „zawsze jakoś jest lub tak właśnie nie jest" itp. Formalizacja tych zwrotów wymaga takich środków logicznych jak zmienne zdaniowe wiązane kwantyfikatorami, za pomocą których owo ogólne twierdzenie ontologiczne sformułować można jako tezę: dla każdego *p*, *p* lub nie-*p*. Ci, którzy mają skrupuły natury filozoficznej wobec traktowania schematów zdaniowych jako zmiennych wiązanych kwantyfikatorami, muszą poprzestać na poszczególnych podstawieniach schematu: *p* lub nie-*p*, reprezentujących te «konkretyzacje» ogólnej ontologicznej zasady wyłączonego środka, które wyrażalne są w przyjętym języku. Różnice między tymi ujęciami są jednak w rozważanym przez nas zagadnieniu nieistotne. Decydujące jest to, że przy wszystkich tych sformułowaniach ontologiczna zasada wyłączonego środka jest zdaniem (lub zbiorem zdań) języka przedmiotowego, a więc tej części, m. in., języka polskiego, która mówi o otaczającym nas świecie, a nie o jego językowym obrazie. Jako zdanie tego języka musi stosować się do jego reguł. Reguły te, zdaniem Ajdukiewicza, nakazują nam stanowczo jej uznanie. W niektórych ze swych prac[4] Ajdukiewicz kwalifikuje ową zasadę jako aksjomat języka przedmiotowego, czyli zdanie, którego odrzucić niepodobna, nie zmieniając przez to sensu użytych w nim wyrazów; w innych[5] — jako zdanie niespornie potwierdzone przez praktykę i doświadczenie. W każdym z tych ujęć ontologiczna zasada wyłączonego środka zaliczana jest do prawdziwych zdań języka przedmiotowego. Przesądza to o jej nienaruszalności. Idealista bowiem, jak podkreśla Ajdukiewicz, utrzymuje, że jego idealistyczna interpretacja pojęć takich jak istnienie, czy prawda, nie zmienia wartości

---

4) M. in. w artykule „Logika a doświadczenie" z 1947 r.
5) Np. w wypowiedzi „W sprawie artykułu prof. A. Schaffa..."

logicznej żadnych zdań języka przedmiotowego. „Przyjmując [...] tezę idealistyczną, ludzie niczego nie będą potrzebowali zmieniać ze swych poglądów, do których ich praktyka i doświadczenie doprowadziły".[6] Do takich właśnie twierdzeń należeć ma ontologiczna zasada wyłączonego środka.

Jeśli godzimy się na tę konkluzję, jedyną kwestionowalną przesłanką metalogicznej zasady wyłączonego środka, pozostaje klasyczna definicja prawdy. Argumentacja Ajdukiewicza skierowana przeciwko rezygnacji z tej definicji głosi, że rezygnacja ta prowadzi do zastąpienia klasycznego pojęcia prawdy pojęciem tezy języka, a to „równoważne jest z zajęciem stanowiska idealistycznego". Dziwnie nieco brzmi to stwierdzenie jako argument w polemice z idealizmem. Istota zarzutu polega, jak sądzić można, na tym, iż definicyjne utożsamienie pojęcia prawdy z pojęciem tezy języka czyni z naczelnego twierdzenia idealizmu transcendentalnego, wedle którego prawdziwe mają być tylko te zdania, które są tezami języka, pustą tautologię. W stosunku do tej argumentacji powstaje jednak pytanie, dlaczego to rezygnując z klasycznej definicji prawdy utożsamiać musimy tym samym pojęcie prawdy, z pojęciem tezy języka. Czy jest to istotnie jedyna modyfikacja klasycznego pojęcia prawdy, która utrzymując w mocy logiczną zasadę wyłączonego środka pozwalałaby odrzucić zasadę metalogiczną? W swych dalszych wywodach uzasadnić chcę negatywną odpowiedź na to pytanie: podać taką modyfikację klasycznej definicji prawdy, która zachowuje logiczną zasadę wyłączonego środka, uchyla metalogiczną, a jednocześnie wolna jest od zarzutu idealizmu. Co więcej, definicja ta zachowywać ma to, co stanowi istotę definicji klasycznej: semantyczne rozumienie prawdy jako zgodności z rzeczywistością.

Definicja prawdy spełniająca te warunki wysunięta została we współczesnych badaniach nad semantyką wyrażeń nieostrych. Jedno z ujęć tej semantyki stanowi pewną modyfikację standardowej semantyki teoriomodelowej, dostosowaną do języków zawierających terminy nieostre. Semantyka teoriomodelowa utworzona była z myślą o językach wolnych od zjawisk takich jak zjawisko nieostrości — mówiąc ogólniej, językach jednoznacznie zinterpretowanych. Interpretacja semantyczna takiego języka J utożsamiona jest z jego modelem $\mathfrak{m}$ — określoną strukturą teoriomnogościową, przyporządkowującą wyrażeniom prostym języka J w sposób jednoznaczny odpowiednie twory teoriomnogościowe. I tak, jednoargumentowemu predykatowi P przyporządkowany zostaje w modelu $\mathfrak{m}$ określony podzbiór jego universum, $P^{\mathfrak{m}}$, a nazwie indywiduowej a — określony element tego universum, $a^{\mathfrak{m}}$. To, co się nazywa językiem zinterpretowanym, reprezentowane więc być może przez parę $(J, \mathfrak{m})$, złożoną z określonego syntaktycznie języka J i jego semantycznego modelu $\mathfrak{m}$. Dla tak pojętego języka formułuje się następnie klasyczną definicję prawdy w postaci definicji indukcyjnej. Zawarty w niej warunek prawdziwości dla zdań prostych języka $(J, \mathfrak{m})$ podpada pod ogólny schemat:

(1) „z" jest prawdą, gdy $z^{\mathfrak{m}}$,

6) „W sprawie artykułu prof. A. Schaffa..."

w którym z symbolizuje zdanie języka J, „z" nazwę zdania z, a $z^{\mathfrak{m}}$ interpretację zdania z wyznaczoną przez model $\mathfrak{m}$. W przypadku, gdy zdaniem z jest najprostsze zdanie typu $P(a)$, schemat (1) przybiera postać:

(1.1) „$P(a)$" jest prawdą, gdy $a^{\mathfrak{m}} \in P^{\mathfrak{m}}$,

gdyż tym, co głosi zdanie $P(a)$ w interpretacji wyznaczonej przez model $\mathfrak{m}$, jest stwierdzenie, iż przedmiot denotowany przez nazwę a w modelu $\mathfrak{m}$ należy do zbioru denotowanego przez predykat P w tymże modelu. Uzupełniwszy tę definicję klasycznymi warunkami prawdziwości dla zdań złożonych, łatwo możemy okazać, że na jej gruncie ma walor zasada wyłączonego środka zarówno w swej wersji logicznej, jak i metalogicznej. Dla dowolnego zdania z języka (J, $\mathfrak{m}$) prawdą jest zarówno cała alternatywa: z lub nie-z, jak i jeden z jej sprzecznych członów.

Konstrukcja ta nie znajduje jednak bezpośredniego zastosowania do języków zawierających terminy nieostre, nie można bowiem języków takich uważać za języki jednoznacznie zinterpretowane przy przyjętym w tej semantyce pojęciu interpretacji. Nieostrość terminów danego języka nie pozwala utożsamić interpretacji tego języka z jego standardowym modelem, gdyż nie pozwala utożsamić denotacji owych terminów z odpowiednimi teoriomnogościowymi składnikami takiego modelu. Nie sposób traktować denotacji nieostrego predykatu P jako określonego zbioru — w ścisłym, teoriomnogościowym sensie tego terminu — skoro dla przedmiotów z tzw. zakresu nieostrości predykatu P jego znaczenie nie ustala żadnych kryteriów przynależności do owej denotacji; każdy z tych przedmiotów możemy do denotacji predykatu P równie dobrze zaliczyć, jak go z niej wykluczyć, nie gwałcąc znaczenia tego predykatu. Można zatem powiedzieć, że zbiór, który by miał stanowić denotację predykatu P, wyznaczony jest przez sens tego predykatu w sposób niejednoznaczny, lub — co na jedno wychodzi — że w sposób jednoznaczny wyznaczona jest jedynie pewna klasa zbiorów, odpowiadających wszystkim możliwym sposobom klasyfikacji przedmiotów z zakresu nieostrości tego predykatu na przedmioty będące P i nie-P. W ten sposób otrzymujemy w rezultacie jako interpretację języka J zawierającego terminy nieostre nie jeden model, $\mathfrak{m}$, lecz pewną klasę modeli, M. Tak więc, każdy język zinterpretowany w sposób niejednoznaczny reprezentowany być może przez parę (J, M), składającą się z określonego syntaktycznie języka J i klasy jego semantycznych modeli M.

Jaki sens nadać możemy pojęciu prawdy dla języka tak zinterpretowanego? Rozwiązaniem najczęściej wysuwanym i najbardziej przekonującym wydaje się to, które znalazło wyraz w tzw. teorii «super-prawdy» (inaczej semantyce «super-waluacji»). Odwołując się do klasycznego rozumienia pojęcia prawdy i fałszu w danym modelu $\mathfrak{m}$, określoną wartość logiczną przypisujemy tu tylko tym zdaniom języka (J, M), których wartość logiczna w modelach klasy M jest niezmienna. Te, które są prawdziwe we wszystkich modelach klasy M, zaliczone zostają do zdań prawdziwych, te, które są fałszywe we wszystkich tych modelach — do zdań fałszywych. Zdania, które w pewnych modelach klasy M są prawdziwe, w innych fałszywe, zostają uznane za zdania pozbawione wartości logicznej. Odpowiadający tej teorii warunek prawdziwości dla zdań prostych języka (J, M) podpadać więc będzie pod ogólny schemat:

(2) „z" jest prawdą, gdy dla każdego $\mathfrak{m} \in M$, $z^{\mathfrak{m}}$.

Konkretyzacją schematu (2) dla zdań typu $P(a)$ będzie tu warunek:

(2.1) „$P(a)$" jest prawdą, gdy dla każdego $\mathfrak{m} \in M$, $a^{\mathfrak{m}} \in P^{\mathfrak{m}}$.

Określenie zdania fałszywego nie różni się od tradycyjnego: zdanie z jest fałszem, gdy negacja zdania z jest prawdą. Na gruncie takiej teorii prawdy widać wyraźnie, że w przypadku gdy przedmiot a należy do zakresu nieostrości predykatu P, ani zdanie $P(a)$, ani negacja tego zdania, nie-$P(a)$, nie są zdaniami prawdziwymi. Jednocześnie alternatywa tych zdań ma zapewnioną prawdziwość. Łatwo też pokazać w sposób ogólny, że proponowana definicja prawdy zachowuje logiczną zasadę wyłączonego środka, a uchyla metalogiczną. Dla dowolnego zdania z języka (J, M) prawdą musi być alternatywa: z lub nie-z, mimo że żaden z jej członów może nie być zdaniem prawdziwym. Zilustrujemy to na fikcyjnym przykładzie dyskutowanym przez Ajdukiewicza. Stwierdzenie, iż Ewa jest abra lub nie jest abra, uznać musimy za prawdziwe, bo, mówiąc swobodnie, pozostaje ono prawdziwe przy wszelkim możliwym «zaostrzeniu» terminu „abra". Natomiast zdaniu „Ewa jest abra", podobnie jak jego negacji, odmówić musimy prawdziwości, gdyż ich wartość logiczna zależy od tego, jak termin „abra" sprecyzujemy (to jest, który ze zbiorów zgodnych z regułami przyjętymi dla terminu „abra" przypiszemy mu jako denotację).

Czy przyjęcie takiej definicji prawdy może być uznane za „równoważne z zajęciem stanowiska idealistycznego"? Skłonny jestem odpowiedzieć na to przecząco i bronić proponowanej definicji przed zarzutem idealizmu. Na pytanie, czy definicja ta utożsamia prawdę z tezą, trudno dać odpowiedź jednoznaczną, bo sprawa nie jest prosta. Najkrócej rzecz można ująć tak: jeśli z utożsamieniem takim mamy tu istotnie do czynienia, mowa może być jedynie o tożsamości zakresowej, a nie treściowej. Można, jak sądzę, okazać — przy pewnych założeniach dodatkowych — że zbiór zdań, którym omawiana teoria prawdy przypisuje określoną wartość logiczną, pokrywa się faktycznie ze zbiorem zdań rozstrzygalnych wedle reguł danego języka, a w szczególności, że zbiór prawd pokrywa się ze zbiorem tez języka. Problem to jednak niełatwy do konkluzywnego rozstrzygnięcia — z tego przede wszystkim powodu, że w systemach językowych, o których tu mowa, istotną rolę odgrywają empiryczne reguły języka, reguły nie występujące w tych systemach dedukcyjnych, które stanowią przedmiot klasycznych badań metalogicznych. W sprawę tę jednak wnikać bliżej nie będziemy, ponieważ nie wydaje się ona dla interesującego nas tutaj problemu decydująca. Nawet gdyby miało się okazać, że na gruncie omawianej definicji zbiór prawd pokrywa się w pewnych warunkach ze zbiorem tez języka, definicja ta pojęcia prawdy z pojęciem tezy bynajmniej nie identyfikuje. Nie tylko bezpośrednio nie definiuje ona pojęcia prawdy przez pojęcie tezy, ale i pośrednio nie odwołuje się przy definiowaniu zdania prawdziwego do pojęcia tezy języka, czy też ogólniejszego od niego pojęcia zdania rozstrzygalnego wedle reguł języka. Warunek prawdziwości stanowiący prawą stronę równoważności (2), czy (2.1), sformułowany jest w zasadzie w języku przedmiotowym (a raczej, w tej części metajęzyka, która stanowi przekład języka przedmiotowego). Warunek ten mówi więc coś o rzeczywistości, do której odnosi się język przedmiotowy, nie różniąc się pod tym względem niczym od warunku prawdziwości charakterystycznego dla klasycznej definicji prawdy postaci (1), czy (1.1). Podobnie jak tamta, definicja proponowana jest definicją semantyczną, wyrażającą pewien stosunek między językiem a rzeczywistością, do której się ten język odnosi. Podobnie jak tamta, stwierdza ona, że zdanie danego języka jest prawdziwe, gdy jest tak, jak to zdanie głosi w interpretacji przysługującej temu językowi. Tak jak definicja klasyczna, i ta istotę prawdziwości upatruje w zgodności z rzeczywistością. To, co ją od tamtej różni — to charakterystyka owej rzeczywistości. Pod względem formalnym różnica ta przejawia się w tym, że rzeczywistość, o której mowa, reprezentowana jest nie przez jedną strukturę teoriomnogościową (jeden model języka), m, tylko przez klasę takich struktur (klasę takich modeli), *M*. Czy jest to zgodne z realistycznym pojmowaniem rzeczywistości? Czy nie pociąga konsekwencji idealistycznych? Wątpliwości te uważam za bezpodstawne. Zwróćmy uwagę na to, iż owa klasa struktur *M* reprezentować ma nie rzeczywistość «samą w sobie», rzeczywistość «po prostu», lecz tę rzeczywistość, do której odnosi się język *J*, to, o czym mówimy w języku *J*, inaczej — korelat semantyczny języka *J*. Postulowana przez nas nieostrość tego języka polega właśnie na tym, iż to, o czym mówimy za jego pomocą, wyznaczone jest w sposób nieostry. A więc, z punktu widzenia «ostrej» ontologii teoriomnogościowej, wyznaczone jest w sposób niejednoznaczny. Dlatego to owemu korelatowi semantycznemu języka *J* odpowiadać musi na gruncie tej ontologii nie jeden model języka *J*, tylko klasa takich modeli. Nie dostrzegam w tym jeszcze żadnych elementów idealizmu.

W związku z ową nieokreślonością korelatu semantycznego języków nieostrych wysuwa się niekiedy następującą wątpliwość. Nieokreśloność ta, jak widzieliśmy, odpowiedzialna jest za pojawienie się w danym języku tzw. zdań nie zdeterminowanych, to jest zdań pozbawionych wartości logicznej. Zdaniem takim będzie zdanie *P*(*a*) przypisujące predykat *P* przedmiotowi z zakresu jego nieostrości. Nie jest ono ani prawdą ani fałszem. Czy fakt ten nie zmusza nas do przyznania, że nie zdeterminowana pod danym względem jest również sama rzeczywistość? Czy ów fakt nie zdeterminowania rzeczywistości jest do przyjęcia dla konsekwentnego realisty? Niewłaściwe wydaje mi się takie stawianie sprawy. Niejednoznacznie zdeterminowane jest to, do czego odnosi się dany język, w tym przypadku — to, do czego odnosi się predykat *P*; odpowiada mu nie jeden zbiór przedmiotów, lecz wiele różniących się między sobą zbiorów. To nie świat jest nie zdeterminowany pod danym względem; nie zdeterminowany jest ów «wzgląd», o którym mowa. Nie określone jest pojęcie bycia *P* i dlatego na pytanie, czy *a* jest *P*, nie można odpowiedzieć ani tak, ani nie. Mówiąc nieco metaforycznie: to my zadajemy światu nieokreślone pytanie i dlatego świat daje nam nieokreśloną odpowiedź.

Nie widzę w rezultacie żadnych racji dla przypisywania rozważanej przez nas definicji prawdy charakteru idealistycznego. Nie ma takiego charakteru ani samo pojęcie interpretacji języka nieostrego (czy — ogólniej — niezupełnego), ani pojęcie prawdziwości określone dla zdań języka tak zinterpretowanego. Przywiązuję do tej konkluzji pewne znaczenie, ponieważ podzielam przekonanie Ajdukiewicza co do tego, iż istotnym elementem stanowiska idealistycznego jest charakterystyczna dla tego stanowiska teoria prawdy. Różnicę między realistyczną a idealistyczną teorią prawdy formułuje Ajdukiewicz ogólnie w sposób następujący. Pierwsza definiuje prawdziwość jako zgodność z rzeczywistością, druga — jako zgodność z kryterium prawdy („rozumie się przy tym przez owo kryterium prawdy jakąś taką własność zdań, którą można sformułować nie mówiąc niczego o rzeczach, których te zdania dotyczą").[7] Definicja realistyczna jest więc definicją semantyczną; istotę prawdziwości upatruje w stosunku zdania do rzeczywistości. Definicję idealistyczną określa się niekiedy jako definicję «epistemiczną»; istotę prawdziwości upatruje ona w stosunku zdania do rezultatu określonych zabiegów poznawczych. Podobne przeciwstawienie występuje również w kontekście innych kontrowersji filozoficznych. Tak na przykład Andrzej Grzegorczyk, nawiązując do sporu między klasyczną a intuicjonistyczną (i konstruktywistyczną) filozofią matematyki, wyróżnia dwa podstawowe stanowiska filozoficzne i dwa związane z nimi pojęcia wartości: „wartość pojmowaną ontologicznie i wartość pojmowaną empirycznie". Prawda jako „wartość pojmowana ontologicznie" — to „to, co jest zgodne z rzeczywistością"; prawda jako „wartość pojmowana empirycznie" — to „to, co jest uznawane we wszelkich warunkach".[8] We wszystkich tych przeciwstawieniach dochodzi do głosu pewna podstawowa opozycja filozoficzna, którą skłonny byłbym uważać za istotę sporu między realizmem a idealizmem. Przedmiot kontrowersji sformułować można najogólniej tak oto. Czy to, co istnieje, zależne jest od tego, co jest poznawalne? Czy to, co jest prawdą, zależy od tego, co jest rozstrzygalne? Realizm daje na te pytania odpowiedź przeczącą, idealizm — twierdzącą. Gdzie mieści się w tym schemacie nasza teoria «super-prawdy»? Starałem się okazać, że jej miejsce jest po stronie realistycznej. To, o czym mówimy, jest co prawda zależne od naszych możliwości językowych; stąd bierze się owa nieokreśloność przedmiotu naszej myśli. Ale to, co jest prawdą o tym przedmiocie, a co fałszem, co istnieje, a czego nie ma — określone jest w sposób niezależny od naszych rezultatów i możliwości poznawczych. Przejawem tego jest stosunek owej teorii prawdy do zasady wyłączonego środka: przy uchyleniu zasady metalogicznej, zachowanie zasady logicznej. Dzięki tym własnościom teoria ta dostarcza argumentu przeciwko krytyce, jaką wobec transcendentalnego idealizmu wysunął Ajdukiewicz. Utrzymywał on, jak pamiętamy, iż — wbrew idealiście — prawdy nie można utożsamiać z tezą, bo zbiór prawd jest zupełny,

7) „Epistemologia i semiotyka".
8) „Klasyczne, relatywistyczne i konstruktywistyczne sposoby uznawania twierdzeń", *Studia Logica* 27, 1971.

a zbiór tez niezupełny. Tym, co stwierdza zupełność zbioru prawd, jest właśnie metalogiczna zasada wyłączonego środka. Jeśli ją odrzucimy, argumentacja powyższa traci swą moc. Zbiór prawd — i to prawd rozumianych w sposób nie odbiegający od klasycznego — okazuje się zbiorem niezupełnym, jeśli mają to być prawdy języka empirycznego, a więc notorycznie nieostrego.

Tak wypadałoby ocenić «oficjalną» argumentację Ajdukiewicza, zawartą przede wszystkim w jego pracy o transcendentalnym idealizmie. W pracach późniejszych jednakże, w szczególności w rozprawie „Epistemologia i semiotyka", znajdujemy również inną nieco wersję tej argumentacji, odwołującą się nie do metalogicznej, lecz do logicznej zasady wyłączonego środka. Ponieważ Ajdukiewicz traktował obie te zasady jako równoważne, nie przykładał wagi do tej modyfikacji. Nabiera ona jednak znaczenia na gruncie naszej definicji prawdy, która odrzucając zasadę metalogiczną utrzymuje w mocy zasadę logiczną. Warto więc na koniec przyjrzeć się i tej wersji Ajdukiewiczowskiej krytyki. Poprzednia rekonstrukcja naczelnego twierdzenia idealizmu transcendentalnego głosiła, iż „z" jest prawdą zawsze i tylko, gdy „z" jest tezą (lub — w innym nieco ujęciu — gdy „z" spełnia kryterium prawdy). Obecnie twierdzenie to formułuje się jako równoważność: z zawsze i tylko, gdy „z" spełnia kryterium prawdy. Otóż równoważność ta — mówi Ajdukiewicz — jest jawnie fałszywa, skoro z jednej strony dla dowolnego zdania z ma walor logiczna zasada wyłączonego środka: z lub nie-z, a z drugiej strony „w każdym języku przedmiotowym odpowiednio bogatym znajdą się zdania wedle kryterium nierozstrzygalne, a więc takie, że ani ono samo, ani jego negacja nie spełnia kryterium". Wobec utrzymania w mocy przez proponowaną tu definicję prawdy logicznej zasady wyłączonego środka, argumentacja ta, w przeciwieństwie do poprzedniej, zachowuje swoją ważność. Powstaje jednak pytanie, czym uzasadnić takie właśnie sformułowanie twierdzenia idealistycznego. Okazuje się, że w dojściu do takiej formuły decydującą rolę odgrywa jako krok pośredni słynna „konwencja Tarskiego", charakteryzująca jego definicję prawdy:

(T) „z" jest prawdą zawsze i tylko, gdy z.

Ona to pozwala przejść od formuły poprzedniej do obecnej. Krok ten jednak traci swe uzasadnienie na gruncie tej wersji klasycznej definicji prawdy, którą rozważamy obecnie: teoriomodelowej definicji prawdy dla języków niejednoznacznie zinterpretowanych. Warunek prawdziwości dla zdania z w podanej uprzednio równoważności (2) przybiera postać: dla każdego $\mathfrak{m} \in M$, $z^{\mathfrak{m}}$. Warunek prawdziwości dla negacji zdania z — dla każdego $\mathfrak{m} \in M$, nie-$z^{\mathfrak{m}}$ — nie stanowi tu bynajmniej negacji warunku poprzedniego. Toteż nie znajduje wobec nich zastosowania logiczna zasada wyłączonego środka. Żaden z tych warunków może nie być spełniony. I tak też jest w przypadku zdań nie zdeterminowanych. Nie jest to na pewno jedyna droga, na jakiej dojść można do omawianego obecnie sformułowania naczelnego twierdzenia idealizmu. Można je próbować uzasadniać bezpośrednio, a nie pośrednio — poprzez analizę pojęcia prawdy. Pewne uwagi Ajdukiewicza wskazują na taką możliwość. Ale ta droga wydaje się mało przekonująca, a proponowana w jej wyniku eksplikacja tezy idealistycznej — mało intuicyjna.

W związku z omawianą sprawą warto może raz jeszcze wyraźnie sformułować założenie, jakie leży u podłoża przedstawionej tu argumentacji. Założenie to dotyczy rodzaju ontologii zakładanej przez przyjętą przez nas teorię prawdy. Jak wszelka semantyka teoriomodelowa, i ta teoria oparta jest na ontologii teoriomnogościowej. Przedmioty, do których odnoszą się wyrażenia danego języka, utożsamione zostają z określonymi tworami teoriomnogościowymi. To, co głosi zdanie takiego języka, wyrażone więc zostaje ostatecznie w języku «stosowanej» teorii mnogości. Dlatego to klasyczny schemat Tarskiego (*T*) zastąpiony zostaje teoriomodelowym schematem (1) — nawet w przypadku języków jednoznacznie zinterpretowanych. Jeśli warunek prawdziwości zdania *z* wyrażony ma być w teoriomnogościowym aparacie pojęciowym, nie zawsze rolę tę pełnić może samo zdanie *z*. Metajęzyk, w którym formułujemy definicję prawdy, zawierać więc musi nie tyle sam język przedmiotowy, ile jego teoriomnogościowy przekład. Oparcie się na teoriomnogościowej ontologii prowadzi z kolei do traktowania języków nieostrych (czy niezupełnych) jako języków niejednoznacznie zinterpretowanych. Ich interpretacja utożsamiona zostaje w konsekwencji z klasą modeli, a prawdziwość zdefiniowania zgodnie ze schematem (2), do którego w sposób istotny odwołuje się przedstawiona przez nas argumentacja. Co przemawia za przyjęciem takiej właśnie teorii ontologicznej? Po prostu to, iż jest to jedyna formalna teoria rzeczywistości o dostatecznym stopniu precyzji i ogólności. Dzięki tym własnościom ułatwia ona uściślenie, a tym samym i rozstrzygnięcie niektórych zagadnień filozoficznych, „których rozwiązanie napotykało trudności z powodu niedostatecznej precyzji aparatu pojęciowego, który służył do ich sformułowania".

Filozofia Nauki
Rok I, 1993, Nr 2-3

Argumentacja reisty

I

Celem tych rozważań jest analiza argumentów, jakie na poparcie tezy reizmu przytaczają jego zwolennicy — a wśród nich sam jego twórca. Podejmowane już były próby charakterystyki tej argumentacji. Zawarte są one zarówno w wypowiedziach Kotarbińskiego, pochodzących z różnych „faz rozwojowych" jego poglądów, jak i w uwagach krytyków i komentatorów tej doktryny — przede wszystkim w wnikliwych analizach Ajdukiewicza i Kotarbińskiej (Ajdukiewicz 1930, 1934; Kotarbińska 1967). Zasadnicze myśli dotyczące uzasadnienia tezy reistycznej, jakie w pracach tych można znaleźć, uważam za trafne. To, co mam w tej sprawie do dodania, sprowadza się do pewnych uwag szczegółowych — uzupełniających i porządkujących zagadnienia — oraz do pewnych refleksji uogólniających. Zagadkowy bowiem i frapujący wydaje mi się problem ogólny, którego przykładem tylko jest omawiany tu przypadek. Jaki jest sposób uzasadniania tez filozoficznych, w szczególności filozoficznych tez ontologicznych? I jaki jest — ściśle zależny od owego sposobu — charakter metodologiczny takich tez? Czy sposób ich uzasadniania nie wykracza poza metody stosowane w nauce? Czy ich status metodologiczny mieści się jakoś w tradycyjnych podziałach twierdzeń naukowych na zdania analityczne i syntetyczne, aprioryczne i empiryczne? Przy rozważaniu tego zagadnienia doktryna reizmu zasługuje na szczególną uwagę — jako oryginalny pogląd wielkiego filozofa współczesnego, wyróżniającego się w dodatku niezwykle wysoką samoświadomością metodologiczną.

Uprzedzając wynik dalszych rozważań, chciałbym na wstępie przedstawić ich zasadniczą ideę. Zwykło się, od czasu znanej krytyki Ajdukiewicza, wyróżniać co najmniej dwie wersje tezy reistycznej: wersję ontologiczną i wersję semantyczną. Ontologiczna teza reizmu jest pewnym twierdzeniem o świecie. Głosi ona, iż istnieją tylko rzeczy. Teza semantyczna jest pewnym twierdzeniem o języku, za pomocą którego ów świat opisujemy. Głosi ona w swobodnym sformułowaniu, iż naprawdę mówimy tylko o rzeczach. Jakiż jest wzajemny stosunek tych tez? W swej późniejszej refleksji na ten temat Kotarbiński traktuje tezę semantyczną jako «punkt wyjścia» koncepcji reistycznej, mając na myśli przede wszystkim jej pierwszeństwo genetyczne (Kotarbiński 1958). Ale ogół tekstów autora poświęconych wykładowi jego doktryny mówi coś więcej. Pozwala sądzić, że teza semantyczna poprzedza tezę ontologiczną nie tylko w porządku genetycznym, ale i w porządku metodologicznym. Trzeba odróżnić ten ostatni od porządku logicznego. Teza T_1 poprzedza tezę T_2 w sensie logicznym, gdy T_1 jest racją logiczną dla T_2. Teza T_1 poprzedza tezę T_2 w sensie metodologicznym, gdy T_1 stanowi uzasadnienie dla T_2. Porządki te nie pokrywają się w przypadku wspomnianych tez reistycznych. Teza semantyczna nie pociąga logicznie tezy ontologicznej, ale do niej to właśnie odwołuje się podstawowy sposób uzasadnienia tezy ontologicznej. Mimo iż teza ontologiczna jest twierdzeniem o świecie, a nie języku, uzasadniana jest przez powołanie się na tezę semantyczną, a więc na pewne twierdzenie o języku. Nie sądzę przy tym, aby to był fakt swoisty dla reistycznej jedynie tezy ontologicznej. Skłonny byłbym w nim upatrywać rys symptomatyczny dla sposobu uzasadniania i innych tez ontologicznych — tez należących do obszernej i doniosłej klasy twierdzeń filozoficznych. Warto więc temu sposobowi uzasadniania przyjrzeć się nieco bliżej. Skoro «punktem wyjścia» jest w nim teza semantyczna, od niej wypada rozpocząć naszą analizę.

II

1. Z góry trzeba stwierdzić, że semantyczna teza reizmu występuje w dziełach jej autora w różnych — nierównoważnych logicznie — wersjach. Jedną z tych wersji sformułować można na podstawie tej prezentacji idei reistycznej, która poprzedza wszystkie pozostałe. Tam, gdzie po raz pierwszy pojawia się owa idea, czytamy co następuje. „Wszelkie zdania, w których wypowiada się coś pozornie o innym jakimś przedmiocie, nie o rzeczy jakiejś, traktujemy jako zwroty zastępcze dla zdań innych, rozumianych już literalnie, a orzekających wyłącznie o rzeczach" (Kotarbiński 1929). Spróbujmy wyjaśnić niektóre występujące w tym powiedzeniu zwroty. Rzecz — to, wedle jednej z definicji podanych przez autora, tyle co przedmiot będący kiedyś, będący gdzieś i fizykalnie jakiś. Pojęcie zdania orzekającego coś o przedmiotach danego rodzaju zastąpić można, zgodnie z sugestią autora, pojęciem zdania, które formalnie pociąga za sobą egzystencję takich przedmiotów. Wprowadźmy w związku z tym następujące skróty. Zdaniem niereistycznym nazwiemy zdanie pociągające za sobą formalnie egzystencję przedmiotów innych niż rzeczy, zdaniem reistycznym — zdanie, które egzystencji takich przedmiotów nie pociąga. Pozwala to sformułować omawianą tezę jak następuje:

(T_s) Wszelkie sensowne zdanie niereistyczne jest przekładalne na pewne zdanie reistyczne.

Ale występujące w tym sformułowaniu pojęcie przekładalności wymaga odrębnego wyjaśnienia. Idzie w nim bowiem nie tyle o „przekład" w ścisłym znaczeniu tego słowa, co o „parafrazę". Pojęcie przekładu jest tu stosowane do wyrażeń, których właściwy sens jest sensem nie dosłownym, lecz — w szerokim tego słowa znaczeniu — przenośnym. Toteż mówiąc, że zdanie niereistyczne Z jest przekładalne na zdanie reistyczne Z_r, mamy na myśli to, iż właściwy sens zdania Z jest sensem nie dosłownym, lecz przenośnym, i że sens ten jest identyczny z sensem dosłownym zdania Z_r. Zawarte w tezie T_s ograniczenie owej przekładalności do „sensownych" zdań niereistycznych motywowane jest tym, iż dopuszcza się istnienie zdań pozornie tylko sensownych, a więc wyrażeń, które będąc zdaniami w sensie gramatycznym pozbawione są w istocie określonego sensu. Jako prosty przykład sensownego zdania niereistycznego przytoczmy za autorem wypowiedź:

(Z) Białość przysługuje śniegowi,

która, rozumiana literalnie, pociąga twierdzenie o istnieniu białości. Jej sens właściwy oddaje, zdaniem Kotarbińskiego, reistyczna parafraza:

(Z_r) Śnieg jest biały,

która, zakładając istnienie śniegu i rzeczy białych, nie zakłada bynajmniej istnienia białości.

Podane wyżej sformułowanie tezy T_s jest wyraźnie niepełne. Jako teza semantyczna wymaga ona relatywizacji do określonego języka J lub określonego rodzaju języków. W jakim języku J zachodzić ma postulowana w niej przekładalność? Rozważania autora odnoszą się bezpośrednio do języka polskiego. Ale osobliwości tego języka etnicznego nie wydają się dla dyskutowanego problemu istotne. Ważne wydaje się natomiast to, że jest to język naszej wiedzy. Wykład reizmu niejednokrotnie odwołuje się do tego pojęcia. Wyróżniony zostaje zbiór zdań należycie uzasadnionych i uznanych z tej racji za prawdziwe jako zbiór zdań składających się na ogół naszej wiedzy W. Uzasadnione wydaje się przypuszczenie, iż, zgodnie z intencją autora, język J — to język wiedzy W, w szczególności — język nauki. A więc nie tylko język polski, ale i wszelki inny język, w którym wiedza W daje się sformułować. Teza semantyczna T_s ma mieć walor dla każdego języka spełniającego ten warunek.

2. W wywodach twórcy reizmu spotykamy również inną nieco wersję tezy semantycznej. W dotychczasowym ujęciu teza ta głosiła przekładalność na język reistyczny każdego sensownego zdania niereistycznego. W ujęciu, o którym mowa obecnie, postuluje ona przekładalność każdego prawdziwego zdania niereistycznego:

(T_{sa}) Wszelkie prawdziwe zdanie niereistyczne jest przekładalne na pewne zdanie reistyczne.

Oto jedno ze sformułowań autora wyrażających tę ideę. „Wszelka wypowiedź pociągająca za sobą formalnie egzystencję [przedmiotów innych niż rzeczy] może być prawdziwa o tyle tylko, o ile ma ona jakiś taki charakter zastępczy, nieliteralny, jakieś takie rozumienie wtórne, przy którym wspomniany dowód egzystencji nie mógłby znaleźć miejsca" (Kotarbiński 1930/31). Zgodnie z tym ujęciem, istnienie reistycznej parafrazy ma być warunkiem niezbędnym prawdziwości danego zdania niereistycznego, a nie samej jego sensowności. Formalnie rzecz biorąc, jest to więc sformułowanie logicznie słabsze od poprzedniego: teza T_s pociąga tezę T_{sa}, lecz nie na odwrót. To jednak, jaki jest w istocie rzeczy stosunek między tymi tezami, zależy od pewnych okoliczności dodatkowych: od struktury logicznej języka, o którym mowa, i od rodzaju reguł przekładu, który tezy te zakładają. Ta ostatnia sprawa ma znaczenie decydujące, reguły te bowiem mogą mieć postać taką, że zapewniając przekładalność wszystkim prawdziwym zdaniom niereistycznym zapewniają ją przez to samo już wszystkim zdaniom fałszywym. I tak też na ogół bywa, a przykłady reguł przytaczane przez reistę przemawiają za taką ewentualnością i w przypadku przekładalności przez niego postulowanej. Prócz przekładu konkretnych zdań niereistycznych podaje się reguły przekładu schematów takich zdań, gwarantujące przekładalność wszystkim zdaniom pod nie podpadającym — niezależnie od ich prawdziwości czy fałszywości. Mówi się więc, jak przełożyć na język reistyczny nie tylko konkretne zdanie niereistyczne: „Białość przysługuje śniegowi", ale i jego ogólny schemat:

($S(Z)$) X-owość przysługuje Y-owi,

którego reistyczną parafrazą ma być schemat:

($S(Z_r)$) Y jest X-owy.

Ponieważ jednak pełnego zestawu takich reguł przekładowych nikomu jak dotąd, nie udało się podać, problem równoważności tez T_s i T_{sa} traktować musimy jako otwarty, a tezy te jako sformułowania wyrażające dwie różne wersje tezy semantycznej.

3. Teza semantyczna przyjmuje również postać mocniejszą od tezy T_s. Wersja taka pojawia się już w *Elementach* i uchodzić może za najbardziej reprezentatywną postać tej doktryny. W porównaniu z wersjami poprzednimi operuje ona bardziej rygorystycznym pojęciem zdania reistycznego. W ujęciu dotychczasowym zdaniem takim mogło być dowolne zdanie, które nie pociąga egzystencji przedmiotów innych niż rzeczy. Wedle koncepcji omawianej musi to być zdanie języka o ściśle określonych własnościach syntaktycznych i semantycznych. Nazwijmy taki język językiem reistycznym. Jakie warunki nakłada nań autor? Po pierwsze — pewne warunki strukturalne. Język reistyczny — to język o strukturze logicznej ontologii Leśniewskiego. W języku takim, jak wiadomo, zdaniami podstawowymi, do których sprowadzalne są wszystkie inne, są zdania typu „A jest B". Otóż na zdania te nakłada reista pewien warunek dodatkowy. Rozumiane literalnie, zdanie „A jest B" ma być sensowne tylko wtedy, gdy „A" i „B" są nazwami rzeczy. Do tak rozumianego języka reistycznego odwołuje się omawiana obecnie wersja tezy semantycznej:

(T_{sb}) Wszelkie sensowne zdanie nie będące zdaniem języka reistycznego jest przekładalne na pewne zdanie języka reistycznego.

Analogiczną wersję otrzymujemy w przypadku tezy T_{sa}. Tak sformułowane tezy są logicznie mocniejsze od poprzednich, gdyż każde zdanie języka reistycznego jest zdaniem reistycznym, ale nie na odwrót. Prostym przykładem sformułowanym w języku ontologii Leśniewskiego może być negacja zdania Z: „Nieprawda, że białość jest przysługująca śniegowi". Nie będąc zdaniem języka reistycznego (bo zawiera nazwę nie-rzeczy), jest zdaniem reistycznym (bo, na gruncie ontologii, nie pociąga istnienia białości).

Nie sposób nie zauważyć, że zawarty w charakterystyce języka reistycznego warunek sensowności zdań typu „*A* jest *B*" ma z punktu widzenia naszych poczuć językowych charakter dość arbitralny. Tym ważniejsze stają się próby jego uzasadnienia, jakie znaleźć można w wykładzie doktryny reistycznej. Sprowadzają się one do następującej argumentacji. Sens spójki „jest" w zdaniu „*A* jest *B*" zależny jest od metodologicznego charakteru tego zdania. Spójka ta ma sens literalny tylko wtedy, gdy zdanie „*A* jest *B*" jest zdaniem spostrzeżeniowym („wypowiedzią pośrednią sądu spostrzegawczego"), a tak z kolei jest tylko wtedy, gdy *A* i *B* są przedmiotami zasadniczo spostrzegalnymi; takimi zaś przedmiotami są wyłącznie rzeczy (Kotarbiński 1929, 1949). Trudno argumentację tę uznać za przekonującą. Uzasadnione wydaje się przypuszczenie, że zdanie „Białość jest przysługująca śniegowi" ma taki sam charakter metodologiczny, jak zdanie „Śnieg jest biały"; oba wydają się zdaniami spostrzeżeniowymi — zdaniami, o których prawdziwości decydują wyniki naszych spostrzeżeń. Natomiast niewątpliwie analityczne zdanie „Białość jest cechą" różni się pod względem metodologicznym od obu z nich. Spójka „jest" winna by więc mieć taki sam sens w zdaniu pierwszym i drugim, a różny w pierwszym i trzecim — wbrew twierdzeniom reisty. Sądzę jednak, że — ogólnie biorąc — charakter metodologiczny zdania „*A* jest *B*" związany jest z sensem słów „*A*" i „*B*", a nie z sensem spójki „jest". Ta ostatnia może więc mieć sens taki sam w zdaniach o różnym typie metodologicznym. Można się zgodzić z tym, że w przeciwieństwie do zdania „Śnieg jest biały" zdanie „Białość jest przysługująca śniegowi" ma sens nie literalny, lecz przenośny. Ale ten przenośny sens całości wydaje się rezultatem przenośnego charakteru zwrotu „przysługująca śniegowi", a nie słowa „jest", które w obu tych zdaniach może mieć sens identyczny. Tak więc owa rygorystyczna wersja tezy semantycznej T_{sb} wydaje się bardziej problematyczna od poprzednich. Ale też żąda ona zbyt wiele, skoro główny cel reisty wydaje się spełniony już przez to, co gwarantuje najsłabsza wersja tej tezy, T_{sa}. Na niej więc przede wszystkim wypada się skupić w dalszych rozważaniach.

4. Jak uzasadnia twórca reizmu głoszoną przez siebie tezę semantyczną? W późniejszej refleksji na ten temat sposób jej uzasadnienia nazywa uzasadnieniem „naiwno-intuicyjnym i pospolicie indukcyjnym" (Kotarbiński 1958). I jest to charakterystyka trafna. Teza semantyczna jest pewnym twierdzeniem ogólnym uzasadnianym przez autora przez powołanie się na poszczególne jego przypadki. Przesłanki, na których się opiera to uzasadnienie, dotyczą konkretnych zdań niereistycznych lub schematów takich zdań, i stwierdzają ich przekładalność na określone zdania lub schematy reistyczne. Pamiętając, co rozumie się tu przez „przekładalność", możemy przesłanki te sformułować jako twierdzenia głoszące, iż właściwy sens danego zdania niereistycznego Z (*resp.* jego schematu $S(Z)$) jest sensem przenośnym, identycznym z sensem dosłownym zdania reistycznego Z_r (*resp.* schematu $S(Z_r)$). Proste przykłady takich

twierdzeń przytaczaliśmy wyżej. Nie ulega wątpliwości, że przypadki, na które w swej argumentacji autor się powołuje, nie wyczerpują ogółu zdań objętych zakresem tezy semantycznej. Nie zmienia tej sytuacji fakt uwzględnienia pewnych klas zdań — wszystkich zdań podpadających pod dany schemat. Teza semantyczna jest uogólnieniem wykraczającym daleko poza zbiór wszystkich stwierdzonych przypadków. Toteż jest uogólnieniem uzasadnionym jedynie częściowo, przez zwykłą indukcję niewyczerpującą — zgodnie z tym, co stwierdza sam jej autor.

Metodologiczny charakter tezy semantycznej zależny jest w rezultacie od metodologicznego charakteru jej przesłanek. Tej sprawie musimy poświęcić obecnie nieco uwagi, bo nie jest to sprawa oczywista. Przesłanki tezy semantycznej należą do zdań stwierdzających równoznaczność pewnych zwrotów językowych. Są to więc twierdzenia podpadające pod schemat następujący:

(a) Zdanie Z_1 znaczy tyle, co zdanie Z_2 w języku J.

Jaki charakter metodologiczny możemy przypisać takim twierdzeniom? Na gruncie semantyki logicznej zwykło się je uważać za zdania analityczne. Wydaje się to szczególnie przekonujące przy takiej eksplikacji sensu tych zdań, która zdanie (a) traktuje jako odpowiednik zwrotu:

(b) Zdanie: „Z_1 zawsze i tylko, gdy Z_2" jest analityczne w języku J.

Jeśli analityczność w języku J — to tyle, co prawdziwość na mocy reguł języka J, to uzasadnione wydaje się przekonanie, że zdanie (b) metajęzyka MJ języka J przypisujące analityczność pewnemu zdaniu języka J samo jest analityczne w metajęzyku MJ. Ale tak jest tylko przy określonych założeniach dotyczących własności metajęzyka MJ, w szczególności — tego, w jaki sposób zdefiniowany jest w nim język J. Sprawą decydującą jest to, czy definicja języka J zawiera wyszczególnienie jego (syntaktycznych i semantycznych) reguł, czy też nie zawiera. Jeżeli język J określony jest jako język o takich to a takich regułach, stwierdzenie równoznaczności dwóch zdań języka J ma charakter metajęzykowego zdania analitycznego. Jeśli natomiast język J zidentyfikowany zostaje w inny sposób, stwierdzenie takie nie musi być zdaniem analitycznym. A z taką właśnie sytuacją mamy, jak sądzę, do czynienia w przypadku omawianym. Język J jest tu określony jako „język polski" („język, którym mówią Polacy") lub jako „język naszej wiedzy" („język, w którym sformułowana jest nasza wiedza"). Rekonstrukcja reguł tak określonego języka jest sprawą empirii; tylko doświadczenie może nam powiedzieć, jakie to reguły obowiązują w języku tak zdefiniowanym. W konsekwencji charakter empiryczny mają też przesłanki typu (a) lub (b), stwierdzające równoznaczność pewnych wyrażeń takiego języka. Wydaje się, że sytuacja nie zmieniłaby się w sposób zasadniczy nawet wtedy, gdyby język J został określony jako „język, którym ja mówię". I w tym przypadku odtworzenie przeze mnie reguł języka J wymagałoby odwołania się do doświadczenia — z tą jedynie różnicą, że odpowiednie dane empiryczne objęłyby tu również pewne dane introspekcyjne. Inaczej byłoby tylko wtedy, gdyby język J nie był językiem „zastanym", lecz językiem przeze mnie skonstruowanym — m.in. przez konwencjonalne przyjęcie określonych reguł. Ale taka ewentualność nie odpowiada intencjom reisty, który chce, aby jego teza miała walor dla języków istniejących, w szczególności dla istniejącego języka nauki. A jeśli tak, to przesłankom, na których opiera on swą tezę, należy przypisać charakter twierdzeń empirycznych. Charakter empirycznego uogólnienia będzie więc miała i sama teza semantyczna, uzasadniana indukcyjnie za pomocą owych twierdzeń.

5. Doktryna reizmu wkrótce po jej prezentacji poddana została wnikliwej krytyce, która przekonująco, jak sądzę, przemawia za koniecznością pewnej modyfikacji reistycznej tezy semantycznej. Ajdukiewicz słusznie zwrócił uwagę na to, że język potoczny, a więc i język polski, i język taki, jak język naszej wiedzy, są językami «chwiejnymi» — nie określonymi jednoznacznie ani pod względem syntaktycznym ani semantycznym. Dopuszczają zatem różne precyzacje, które różnie determinują ich własności syntaktyczne i semantyczne. „Filozofowie precyzując język, którym mówią, wjeżdżają, nie gwałcąc mowy potocznej, na jeden z kilku torów, które mowa zostawia otworem". I tak też, zdaniem Ajdukiewicza, czyni reista. „Jeśli się powie, że tak [jak głosi reista] jest w każdym języku dopuszczalnym jako możliwa precyzacja mowy potocznej, twierdzenie to jest, jak się zdaje, fałszywe" (Ajdukiewicz 1934). Teza reistyczna może być prawdziwa tylko przy *pewnej* dopuszczalnej precyzacji języka polskiego. Jej sformułowanie zatem wyraźnie takie ograniczenie powinno zawierać. Zamiast prostej relatywizacji do „języka J" winna w niej występować relatywizacja do „pewnej dopuszczalnej precyzacji języka J". Oznaczmy tę „Ajdukiewiczowską" modyfikację tezy T_s symbolem T_s^A.

(T_s^A) Wszelkie sensowne zdanie niereistyczne jest przekładalne, przy pewnej dopuszczalnej precyzacji języka *J*, na pewne zdanie reistyczne.

Analogiczne modyfikacje otrzymujemy w przypadku pozostałych wersji tezy semantycznej, w szczególności — wersji T_{sa}. Jeśli intuicyjny sens tezy semantycznej w jej wersji oryginalnej oddawał zwrot: „Naprawdę mówimy tylko o rzeczach", to sens jej wersji zmodyfikowanej wyrazić można powiedzeniem: „Możemy mówić tak, aby mówić tylko o rzeczach".

Czy taka modyfikacja tezy semantycznej wpływa na jej charakter metodologiczny? Krytyka Ajdukiewicza sugeruje odpowiedź twierdzącą. Dopóki w sposób konwencjonalny nie zostanie dokonana odpowiednia precyzacja języka potocznego, teza reistyczna pozostaje twierdzeniem nierozstrzygalnym. „Reizm może być więc najwyżej rzeczą konwencji" (Ajdukiewicz 1934). Nie wydaje się jednak, aby wniosek ten stosował się do semantycznej tezy reizmu sformułowanej i rozumianej tak, jak z tego zdaje sprawę twierdzenie T_s^A. Twierdzenie to zachowuje, podobnie jak poprzednie, charakter empirycznego uogólnienia. To prawda, że przesłanki, na których się opiera, różnią się nieco od przesłanek tamtych uogólnień. Podpadają one nie pod omawiany uprzednio schemat (a), lecz pod schemat trochę osłabiony:

(c) Zdanie Z_1, w jednym ze znaczeń, jakie ma w języku *J*, znaczy tyle, co zdanie Z_2.

Mimo tej różnicy przesłanki te jednak nie przestają być twierdzeniami empirycznymi. Aby to sobie uświadomić, wystarczy wziąć pod uwagę przykład analogiczny:

Słowo „zamek" w jednym ze znaczeń, jakie ma w języku polskim, znaczy tyle, co „obronna budowla mieszkalna".

Nie ma chyba wątpliwości, że jest to twierdzenie empirycznie rozstrzygalne. Taki sam charakter mają przesłanki, do których odwołujemy się uzasadniając tezę T_s^A, i taki sam wobec tego charakter ma sama ta teza.

Występujące w sformułowaniu tezy T_s^A pojęcie dopuszczalnej precyzacji języka J może być scharakteryzowane nieco wyraźniej przez prosty warunek nałożony na taką precyzację. W każdym niejednoznacznie określonym języku J są zdania, których sensowność i wartość logiczna pozostaje ta sama — niezależnie od takiego czy innego znaczenia, jakie zdania te przybierać mogą w języku J. Precyzacja języka J może być uznana za dopuszczalną wtedy tylko, gdy te własności owej klasy zdań w pełni respektuje. Zdanie, które przy wszelkim swym znaczeniu w języku J jest sensowne (*resp.* bezsensowne) i prawdziwe (*resp.* fałszywe), musi takim pozostać przy wszelkiej dopuszczalnej precyzacji tego języka. Warto zauważyć, że zdania, które składają się na naszą wiedzę, należą do tej właśnie klasy zdań. Konsekwencją tego jest fakt, że postulowany przez tezę T_s^A reistyczny przekład «zachowuje» wiedzę W: zdanie Z należy do wiedzy W wtedy i tylko wtedy, gdy należy do niej jego reistyczny przekład Z_r.

6. Przedstawiony wyżej sposób uzasadnienia reistycznej tezy semantycznej nie jest jedynym, z jakim spotykamy się w wykładzie tej doktryny. Znajdujemy tam argumentację, która nie powołuje się na przesłanki semantyczne, lecz ontologiczne. Dla uzasadnienia semantycznej tezy reizmu odwołuje się do jego tezy ontologicznej. „Reista waży się na tezę, która [...] nosi piętno idei ontologicznej. Twierdzi on, że każdy obiekt jest rzeczą [...] i ten pogląd uzasadnia w jego oczach wszystko, co pragnęliśmy wyżej wyjaśnić" — m.in. reistyczną tezę semantyczną (Kotarbiński 1949). Jest rzeczą oczywistą, że mamy tu do czynienia z uzasadnieniem tezy semantycznej pod tym tylko warunkiem, iż teza ontologiczna uzasadniona została w sposób niezależny. To, czy założenie takie jest uprawnione, jest sprawą otwartą, do której rozważenia przejdziemy w drugiej części tych uwag. Teraz spróbujmy odpowiedzieć na pytanie, na czym by właściwie miało polegać owo uzasadnienie — przy założeniu zasadności tezy ontologicznej. Zapytajmy przede wszystkim, jaki stosunek logiczny łączy obie tezy. Odpowiedź zależy od tego, którą wersję tezy semantycznej mamy na uwadze. Teza ontologiczna pociąga logicznie jej wersję najsłabszą; nie pociąga — bez jakichś dodatkowych założeń — wersji mocniejszych. Teza ontologiczna — to twierdzenie, które głosi, iż każdy przedmiot jest rzeczą, lub — równoważnie — iż:

(T_o) Istnieją tylko rzeczy.

Przy założeniu, że T_o jest zdaniem tego języka J, do którego odnosi się teza semantyczna T_{sa} (lub — inaczej — że język J jest językiem, którym mówi reista), teza T_o pociąga logicznie tezę T_{sa}. Jeśli istnieją tylko rzeczy, to zdanie niereistyczne może być prawdziwe tylko pod warunkiem, iż jest rozumiane w sposób nieliteralny, reistyczny; w przeciwnym bowiem razie stwierdzałoby istnienie przedmiotów nie będących rzeczami. Teza T_o nie pociąga bezpośrednio tez semantycznych mocniejszych od T_{sa} — T_s, a tym bardziej T_{sb}. Może je jednak pociągać przy założeniu stwierdzającym, że każdy sposób zapewnienia prawdziwości tezie T_{sa} (przez podanie odpowiednich reguł przekładu reistycznego) zapewnia tym samym prawdziwość owym tezom mocniejszym — co, nawiasem mówiąc, nie wydaje się dalekie od prawdy w przypadku tezy T_s. Warto natomiast podkreślić bezpośredni związek logiczny tezy T_o z tezą semantyczną T_{sa} zmodyfikowaną zgodnie z sugestią Ajdukiewicza, tj. tezą $T_{sa}{}^A$. Będąc pewnym osłabieniem tezy T_{sa} wynika ona tym samym z tezy T_o.

Skoro teza semantyczna T_{sa} (podobnie jak $T_{sa}{}^A$) jest konsekwencją logiczną tezy ontologicznej T_o, uzasadnienie tej pierwszej przez tę drugą ma charakter wnioskowania dedukcyjnego. Pozwala to na stwierdzenie, że teza semantyczna T_{sa} jest uzasadniona w stopniu tym samym, co teza ontologiczna T_o. W rezultacie rozumowanie to przesuwa punkt ciężkości w tej sprawie z uzasadnienia tezy semantycznej na uzasadnienie tezy ontologicznej. Do rozważenia tego ostatniego zagadnienia przechodzimy obecnie.

III

1. We wstępie do tych uwag wyrażony został pogląd, iż właściwe uzasadnienie tez reistycznych przebiega w sposób odwrotny od ostatnio omawianego. To nie teza semantyczna jest uzasadniana przez tezę ontologiczną, lecz na odwrót: teza ontologiczna jest uzasadniana przez tezę semantyczną. Z pozoru, co prawda, może wyglądać to inaczej. Teza ontologiczna jest pewnym twierdzeniem o świecie, nie zaś o języku, a sposób jej uzasadnienia sam twórca reizmu charakteryzuje jako opierający się na „całym dotychczasowym doświadczeniu", na „wiedzy globalnej" itp. (Kotarbiński 1952). Ale bliższe wejrzenie w rzeczywistą argumentację pokazuje, że „wiedza", do której się autor odwołuje, to swoista „wiedza językowa", a dane doświadczenia, na których się opiera, to pewne fakty semantyczne. Toteż bezpośrednio uzasadniają one reistyczną tezę semantyczną, a pośrednio dopiero — tezę ontologiczną. Do czego bowiem sprowadza się zasadnicza argumentacja reisty? Polega ona na obalaniu rzekomych kontrargumentów wysuwanych przeciwko tezie ontologicznej. Kontrargumenty te — to przykłady należących do naszej wiedzy, a więc uznawanych za prawdziwe, zdań niereistycznych, czyli takich, które, rozumiane dosłownie, zakładają istnienie nie-rzeczy. Argumentacja reisty zmierza do okazania, że właściwy sens tych zdań jest sensem niedosłownym i reistycznym; tak rozumiane, nie stwierdzają one istnienia przedmiotów innych niż rzeczy i nie popadają wobec tego w sprzeczność z ontologiczną tezą reizmu. Argumenty reisty przybierają, jak widać, postać zdań stwierdzających przekładalność konkretnych zdań niereistycznych (lub ich schematów) na pewne zdania reistyczne (lub ich schematy). A zatem są to te same dokładnie twierdzenia, które pełniły rolę przesłanek przy indukcyjnym uzasadnianiu tezy semantycznej. Ich przykłady przytaczane były poprzednio. Tym więc, co reista uzasadnia

bezpośrednio za pomocą swych argumentów, jest teza semantyczna. Powstaje problem, jak przejść od niej do tezy ontologicznej.

2. W przypadku tezy semantycznej w jej słabej — i co za tym idzie, najmniej kontrowersyjnej — postaci T_{sa}, związek między nią a tezą ontologiczną jest jednostronny: z T_o wynika logicznie T_{sa}, ale nie na odwrót. Nie pozwala to na dedukcyjne wyprowadzenie T_o z T_{sa}; umożliwia jednak wnioskowanie redukcyjne. I z możliwości tej istotnie reista korzysta. Jeden ze sposobów uzasadnienia tezy ontologicznej, jakie zawiera wykład tej doktryny, polega na wyprowadzeniu jej z tezy semantycznej w drodze wnioskowania redukcyjnego. Przyjmując, iż teza T_{sa} została już uzasadniona w sposób niezależny, uzasadniamy tezę T_o, która jest jej logiczną racją, jako hipotezę tłumaczącą tezę T_{sa}. Jest to sposób, w jaki uzasadnia się w praktyce naukowej hipotezy empiryczne, choć wartość uzasadniająca takiego wnioskowania pozostaje sprawą wysoce dyskusyjną. W przypadku przez nas rozważanym będzie to wartość raczej niska, toteż reistyczna teza ontologiczna w ten tylko sposób uzasadniona musi pozostać, jak stwierdza sam jej autor, „ryzykownym domysłem" (Kotarbiński 1958). Nasuwa się w związku z tym pytanie, czy nie zachodzi między tymi dwiema tezami jakiś ściślejszy związek logiczny, pozwalający na wyprowadzenie tezy ontologicznej z tezy semantycznej w sposób mniej zawodny.

3. Zanim przejdziemy do rozważenia tezy semantycznej w jej wersji liberalnej, T_{sa}, stanowiącej główny przedmiot naszej analizy, zwróćmy uwagę na szczególny charakter związku tezy ontologicznej T_o z tezą semantyczną w jej wersji najbardziej rygorystycznej, T_{sb}. Przy wspomnianym wyżej założeniu zaliczającym tezę T_o do zdań tego języka J, o którym mówi teza T_{sb}, teza T_{sb} pociąga logicznie tezę T_o. Semantyczna teza T_{sb} stwierdza, mówiąc skrótowo, przekładalność języka J (języka polskiego, czy języka nauki) na język reistyczny. A ten ostatni zdefiniowany został w sposób taki, który z góry zapewnia prawdziwość tezie ontologicznej T_o. Teza ta głosząca, iż każdy przedmiot jest rzeczą, jest prawdziwa na mocy samych reguł języka reistycznego, zawierających, jak pamiętamy, regułę, która przypisuje sensowność (a więc i prawdziwość) zdaniom typu „*A* jest *B*" pod tym tylko warunkiem, że „*A*" i „*B*" są nazwami rzeczy. Uzasadnienie tezy ontologicznej T_o sprowadza się więc w tym przypadku całkowicie do uzasadnienia tezy semantycznej T_{sb}; jeśli tylko język J — język polski, czy język nauki — jest przekładalny na język reistyczny, teza ontologiczna ma w nim z góry zagwarantowaną prawdziwość. Cóż z tego jednak, kiedy uzasadnienie tej właśnie wersji tezy semantycznej jest, jak widzieliśmy, mało przekonujące, a sama teza — wysoce problematyczna. Jest to teza dość mocna na to, aby pociągać za sobą tezę ontologiczną, ale zbyt mocna na to, aby być prawdziwą. Warto jednocześnie zauważyć, że sugerowane przez krytykę Ajdukiewicza osłabienie tezy T_{sb} do postaci $T_{sb}{}^A$ daje w rezultacie taką wersję tezy semantycznej, która co prawda ma większe szanse realizacji, ale która sama tezy ontologicznej nie pociąga. Jeżeli przekładalność na język reistyczny ma zachodzić tylko dla pewnej dopuszczalnej precyzacji języka J, teza T_o nie ma w J z góry zagwarantowanej prawdziwości.

4. Wróćmy zatem do rozważanej uprzednio tezy semantycznej T_{sa} i zapytajmy o możliwości wyprowadzenia z niej tezy ontologicznej T_o. Ponieważ teza T_{sa}, będąc konsekwencją logiczną tezy T_o, sama tezy T_o logicznie nie pociąga, potrzebne są jakieś dodatkowe założenia do tego, aby z tezy T_{sa} móc wyprowadzić tezę T_o. Rolę tego rodzaju założenia pełnić może — wedle spotykanego niekiedy poglądu — zasada zwana „brzytwą Ockhama". Zasada ta, głosząca iż „nie należy mnożyć bytów ponad konieczność", traktowana bywa jako założenie umożliwiające przejście od pewnych twierdzeń semantycznych, takich jak teza T_{sa}, do pewnych twierdzeń ontologicznych, takich jak teza T_o. Na czym owo przejście ma polegać? Rozpatrzmy tę sprawę najpierw dla tezy T_{sa}, a potem dla jej „Ajdukiewiczowskiej" modyfikacji T_{sa}^A, która nasuwa pod tym względem dodatkowe problemy i komplikacje.

Zasada Ockhama ma charakter pewnej metodologicznej dyrektywy, która głosi, iż nie należy przyjmować istnienia jakichś przedmiotów, jeśli to nie jest konieczne. Powstaje więc natychmiast pytanie, o jakiej tu konieczności mowa. Jedna z możliwych eksplikacji tego pojęcia relatywizuje je do ogółu naszej wiedzy W. Przyjęcie istnienia przedmiotów danej kategorii ontologicznej ma być konieczne wtedy, gdy wiedza W zakłada istnienie takich przedmiotów, tj. gdy pociąga logicznie twierdzenie o ich istnieniu. Tak rozumiana zasada Ockhama znajduje istotnie zastosowanie w omawianej przez nas sytuacji. Jeżeli prawdą jest to, co głosi reistyczna teza semantyczna T_{sa}, właściwy sens twierdzeń składających się na naszą wiedzę W jest sensem reistycznym: właściwie rozumiane, twierdzenia te nie zakładają istnienia przedmiotów innych kategorii ontologicznych prócz kategorii rzeczy. A zatem przyjęcie istnienia przedmiotów innych niż rzeczy nie jest konieczne. A skoro tak, to w myśl zasady Ockhama ich istnienia przyjmować nie należy. Narzuca się jednak od razu pytanie: dlaczego? Jakie może być uzasadnienie takiej dyrektywy? Odpowiedź na to pytanie nie jest jednoznaczna. Jeżeli dyrektywa ta ma dostarczać uzasadnienia reistycznej tezie ontologicznej, musi opierać się na pewnym założeniu ontologicznym. Najprościej można je sformułować, jak następuje:

(Z_o) Istnieją tylko takie przedmioty, których istnienie zakłada wiedza W.

W przypadku omawianym prowadzi ono istotnie do wniosku, iż istnieją tylko rzeczy. Ale samo jest założeniem arbitralnym, pozbawionym przekonującego uzasadnienia. Dlaczego istnieć miałyby tylko takie przedmioty, o których istnieniu wiemy? Jeśli z pozoru założenie to wydaje się intuicyjne, to jest tak dlatego, że skłonni jesteśmy utożsamiać je z twierdzeniem innym, o charakterze nie ontologicznym, lecz metodologicznym:

(Z_m) Tylko o takich przedmiotach można zasadnie twierdzić, że istnieją, których istnienie zakłada wiedza W.

Założenie Z_m jest istotnie twierdzeniem nie budzącym wątpliwości. Jeżeli wiedza W (wraz ze swymi logicznymi konsekwencjami) obejmować ma ogół uzasadnionych twierdzeń, to założenie Z_m trzeba uznać za prawdziwe na mocy samej definicji wiedzy. Będąc — w przeciwieństwie do poprzedniego — założeniem w pełni uzasadnionym, nie stanowi ono jednak — jak tamto — przesłanki wystarczającej do wyprowadzenia z tezy semantycznej T_{sa} tezy ontologicznej T_o. Prowadzi do wniosku nie ontologicznego, lecz metodologicznego:

(T_m) Tylko o rzeczach można zasadnie twierdzić, że istnieją.

Tezę T_m nazwać można — w odróżnieniu od tezy semantycznej i ontologicznej — reistyczną tezą metodologiczną.

5. Do podobnych, jak powyższe, konkluzji prowadzi analiza rozumowań wychodzących od tezy semantycznej w jej osłabionej wersji $T_{sa}{}^A$ choć w tym przypadku mamy do czynienia z sytuacją nieco bardziej złożoną. Na gruncie tezy $T_{sa}{}^A$ nie mamy prawa utrzymywać, że istnieje jeden właściwy sens twierdzeń składających się na wiedzę *W*, z którego sprawę zdaje ich przekład reistyczny. Ze względu na nieokreśloność języka *J* twierdzenia te dopuszczają interpretacje różne, a przekład reistyczny odpowiada tylko jednej z nich. W tej sytuacji niektórzy skłonni są twierdzić, że fakt przekładalności naszej wiedzy na język reistyczny pozbawiony jest jakichkolwiek konsekwencji ontologicznych. „Okoliczność, że zdania zawierające „nazwy pozorne" [tj. nazwy nie-rzeczy] dają się zastąpić przez równoznaczne, a nie zawierające nazw pozornych, nie dowodzi niczego jak tylko, że (teoretycznie przynajmniej) można się obejść bez nazw pozornych" (Ajdukiewicz 1930). Otóż zasada Ockhama wydaje się dyrektywą, która pozwala uchylić powyższą konkluzję. Na jej podstawie przekład reistyczny wiedzy *W* zostaje wyróżniony spośród pozostałych jako szczególnie uprzywilejowany — jeśli nie pod względem ontologicznym, to metodologicznym. Ale aby nadawać się do tej roli, zasada Ockhama musi być zinterpretowana nieco inaczej niż poprzednio. Dotyczy to przede wszystkim występującego w niej pojęcia konieczności, zrelatywizowanego do wiedzy *W*, a więc zależnego od charakteryzujących ją założeń. Założenie głoszące, iż wiedza *W* dopuszcza różne sposoby sformułowania, różniące się od siebie pod względem swych konsekwencji egzystencjalnych, sugeruje następującą eksplikację owej konieczności: Przyjęcie istnienia przedmiotów danej kategorii ontologicznej można uważać za konieczne wtedy, gdy każdy sposób sformułowania wiedzy *W* zakłada istnienie takich przedmiotów. Zasada Ockhama zabrania przyjmować istnienia przedmiotów nie spełniających tego warunku. Podobnie jak poprzednio, dwojakie może być uzasadnienie takiej dyrektywy: ontologiczne lub metodologiczne. Nie wolno przyjmować wspomnianych twierdzeń egzystencjalnych albo dlatego że są to twierdzenia fałszywe, albo dlatego że są nieuzasadnione. Założenie ontologiczne, do którego odwołuje się pierwszy rodzaj uzasadnienia, przybiera tu postać następującą:

($Z_o{}^A$) Istnieją tylko takie przedmioty, których istnienie jest zakładane przez każdy sposób sformułowania wiedzy *W*.

Założenie metodologiczne, na którym się opiera drugi typ uzasadnienia, głosi co następuje:

($Z_m{}^A$) Tylko o takich przedmiotach można zasadnie twierdzić, że istnieją, których istnienie jest zakładane przez każdy sposób sformułowania wiedzy *W*.

Tak zinterpretowana zasada Ockhama znajduje zastosowanie w sytuacji, w której ma walor semantyczna teza reizmu T_{sa}^{A}. Postulowany przez nią reistyczny sposób sformułowania wiedzy *W* zakłada tylko istnienie rzeczy. Przy założeniu, że pozostałe sposoby sformułowania wiedzy *W* nie czynią słabszych założeń egzystencjalnych, zasada Ockhama nie pozwala przyjmować istnienia przedmiotów innych niż rzeczy. Wedle założenia Z_{o}^{A} mamy tak czynić dlatego, że tylko rzeczy istnieją; wedle założenia Z_{m}^{A} — dlatego, że tylko o rzeczach można zasadnie twierdzić, że istnieją. Dochodzimy więc do podobnych wniosków, jak poprzednio. Zasada ontologiczna Z_{o}^{A} umożliwia nam wyprowadzenie z tezy semantycznej T_{sa}^{A} reistycznej tezy ontologicznej T_{o}. Ale sama ta zasada ma charakter wysoce problematyczny. Dlaczego istnieć miałyby tylko takie przedmioty, do których uznania zmusza nas nasza wiedza? Przekonująco za to brzmi to, co głosi zasada metodologiczna Z_{m}^{A}. Przy pewnej, dość intuicyjnej, eksplikacji pojęcia twierdzenia zasadnego wynika ona z samej charakterystyki wiedzy *W*. Ta zasada jednak nie pozwala na wyprowadzenie tezy ontologicznej. Służyć może jedynie do uzasadnienia na podstawie semantycznej tezy T_{sa}^{A} reistycznej tezy metodologicznej T_{m}.

Podsumowując powyższą analizę sposobów uzasadniania reistycznej tezy ontologicznej przez tezę semantyczną, przypomnijmy dwa rodzaje wyróżnionych procedur. Oba wychodzą od liberalnej wersji tezy semantycznej — w jej postaci oryginalnej T_{sa} lub zmodyfikowanej T_{sa}^{A}. Jeden z nich polega na wyprowadzeniu z niej tezy ontologicznej (która jest jej logiczną racją) za pomocą wnioskowania redukcyjnego. Drugi jest próbą wyprowadzenia tezy ontologicznej z tezy semantycznej za pomocą entymematycznego wnioskowania dedukcyjnego, w którym rolę przesłanki uzupełniającej pełnić ma pewna wersja zasady Ockhama. Okazuje się przy tym, iż najbardziej przekonująca wersja tej zasady umożliwia nam jedynie wyprowadzenie reistycznej tezy metodologicznej. Wyprowadzenie tezy ontologicznej wymaga oparcia się na wersji znacznie bardziej kontrowersyjnej.

6. Jak zaznaczałem na wstępie, podstawowy sposób uzasadnienia ontologicznej tezy reizmu powołuje się na racje natury semantycznej. I takim też sposobem uzasadniania zajmowaliśmy się do tej pory. Ale choć jest to sposób podstawowy, nie jest to sposób jedyny. W wykładzie doktryny reistycznej — zwłaszcza jej fazy późniejszej (Kotarbiński 1958) — znajdujemy argumentację, która zawiera próbę uzasadnienia ontologicznej tezy reizmu odwołującą się do racji nie semantycznych, lecz epistemologicznych. Wypada więc owej argumentacji epistemologicznej przyjrzeć się nieco bliżej. Jej istotę można, w upraszczającym skrócie, wyrazić za pomocą następującego, formalnie poprawnego sylogizmu:

Tylko rzeczy są poznawalne.
Istnieją tylko przedmioty poznawalne.
A zatem istnieją tylko rzeczy.

Wnioskowanie to uzasadnia konkluzję w stopniu takim, w jakim uzasadnione są jego przesłanki. W tekstach autora tej argumentacji spotykamy próbę uzasadnienia pierwszej przesłanki powyższego sylogizmu. Uzasadnienie to sprowadza się do dwóch tez:

Tylko przedmioty spostrzegalne są poznawalne.

Tylko rzeczy są spostrzegalne.

Teza pierwsza (która, nawiasem mówiąc, ową spostrzegalność ujmuje bardzo szeroko, utożsamiając ją z czysto logiczną możliwością spostrzeżenia) jest wyrazem określonego stanowiska teoriopoznawczego, będącego swoistą odmianą sensualizmu. Stoją więc za nią pewne racje, przytaczane nieraz przez jej zwolenników. Teza druga jest konsekwencją określonego rozumienia pojęcia spostrzegania; może być więc uważana za prawdziwą na mocy samego sensu tego terminu (o tym, co to jest za sens, będzie jeszcze mowa w dalszym ciągu tych rozważań). Nie wchodząc w tej chwili w ocenę stopnia uzasadnienia obu tez, a tym samym i pierwszej przesłanki powyższego sylogizmu, stwierdzić musimy, że w przeciwieństwie do takich czy innych prób uzasadnienia przesłanki pierwszej nie znajdujemy w tekście autora żadnej próby uzasadnienia przesłanki drugiej. Trudno się też domyślić, na czym by taka próba miała polegać. Stwierdzenie, że istnieją tylko przedmioty poznawalne, pozostaje w tej sytuacji założeniem całkowicie arbitralnym. Nie może więc ono w jakimkolwiek stopniu uzasadniać konkluzji tego sylogizmu: ontologicznej tezy reizmu.

Nasuwa się jednak możliwość pewnej modyfikacji owej przesłanki drugiej — podobna do tej, jaka miała miejsce w przypadku założeń leżących u podstawy zasady Ockhama. Może i tu tym, co się w istocie rzeczy ma na myśli, jest pewne twierdzenie metodologiczne, nie zaś ontologiczne. Tak zmodyfikowana przesłanka druga głosiłaby, co następuje:

Tylko o przedmiotach poznawalnych można zasadnie twierdzić, że istnieją.

Twierdzenie takie wydaje się analitycznie prawdziwe: jeżeli o przedmiocie jakimś mogą zasadnie twierdzić, że istnieje, musi to być przedmiot w pewnym przynajmniej stopniu poznawalny. Ale zastąpiwszy w naszym sylogizmie ontologiczną przesłankę drugą jej wersję metodologiczną, nie mamy prawa wyprowadzić jako konkluzji reistycznej tezy ontologicznej T_o. Jedyną konkluzją uprawnioną będzie wtedy teza metodologiczna T_m:

Tylko o rzeczach można zasadnie twierdzić, że istnieją.

Mamy tu zatem do czynienia z podobnym dylematem, jaki zarysował się przy próbach uzasadnienia rozważanych poprzednio.

7. Jedną z najdonioślejszych filozoficznych konsekwencji ontologicznej tezy reizmu jest teza radykalnego realizmu. Głosi ona, mówiąc swobodnie, że nie istnieją treści wyobrażeń (w innej terminologii — obrazy immanentne), w szczególności — tzw. jakości zmysłowe. Mówiąc nieco ostrożniej, zaprzecza ona temu, aby mogły być zdaniami prawdziwymi twierdzenia o istnieniu takich przedmiotów rozumiane w sposób literalny. Na gruncie cytowanej przez nas definicji rzeczy, określającej ją jako przedmiot będący kiedyś, gdzieś i fizykalnie jakiś, teza radykalnego realizmu jest niewątpliwie konsekwencją tezy reizmu, ponieważ treści naszych wyobrażeń, zgodnie

z ich zwykłym rozumieniem, nie mogą być utożsamione z jakimikolwiek rzeczami. Warto jednak przy okazji zauważyć, że podział na tak określone rzeczy i nie-rzeczy nie pokrywa się z tradycyjnie rozumianym podziałem na konkrety i abstrakty. Zgodnie z tym rozumieniem, abstrakty — to przedmioty, którym w ontologii teoriomnogościowej odpowiadają zbiory. Są to więc przedmioty, które ani nie są kiedyś, ani nie są gdzieś, ani nie są fizykalnie jakieś. Przeciwstawić im można konkrety jako przedmioty, którym przysługuje co najmniej jedna z owych charakterystyk. Treści wyobrażeń, nie będąc rzeczami (bo nie są ani gdzieś, ani fizykalnie jakieś), są konkretami, a nie abstraktami (bo są kiedyś). Gdyby przyjąć takie pojęcie konkretu, a nie utożsamiać go, jak to czyni twórca reizmu, z pojęciem rzeczy, teza konkretyzmu głosząca, że istnieją tylko konkrety, byłaby tezą słabszą od tezy reizmu. W tej sytuacji teza realizmu radykalnego, będąc konsekwencją tezy reizmu, nie byłaby konsekwencją tezy konkretyzmu; stanowiłaby w stosunku do tej ostatniej pewne założenie dodatkowe. Nasuwa się przypuszczenie, że tak rozumiany konkretyzm odpowiadał ontologicznemu stanowisku Leśniewskiego. Tłumaczyłoby to fakt nazwania przez Kotarbińskiego jego realizmu radykalnego „dziką nadbudówką" do systemu ontologii Leśniewskiego (Kotarbiński 1958). Jako pogląd odrębny od poglądu konkretyzmu wymagałby on wtedy odrębnego uzasadnienia. I istotnie, uzasadnienie takie zawarte jest w oryginalnym wykładzie doktryny realizmu radykalnego (Kotarbiński 1929, 1930, 1935, 1958). Oprócz ogólnych argumentów na rzecz reizmu, którego konsekwencją jest realizm radykalny, znajdujemy tam argumentację specyficzną, przemawiającą w sposób swoisty za tezą realizmu radykalnego.

Argumentacja ta przybiera postać analogiczną do argumentacji poprzednio rozpatrywanej: stanowi próbę obalenia kontrargumentów wytaczanych przeciwko tezie realizmu radykalnego. Główny z owych kontrargumentów głosi, iż istnieją treści wyobrażeń, bo treści są tym, co jest nam bezpośrednio dane. Krytyka tego argumentu opiera się na określonej eksplikacji zwrotu „*B* jest bezpośrednio dane osobie *A*", traktującej go jako równoznaczny ze zwrotem „osoba *A* spostrzega *B*" i nadającej swoisty sens występującemu w nim pojęciu spostrzegania. Warunkiem sensowności (a więc i prawdziwości) owego zwrotu jest to, aby „*B*" było nazwą rzeczy. Spostrzeganie jest tu pojmowane jako bezpośredni stosunek między człowiekiem a rzeczą; spostrzegać bowiem możemy tylko to, co oddziałuje na nasze narządy zmysłowe, a tym mogą być tylko rzeczy. Powiedzenie, iż spostrzegamy treści, jest nie tylko fałszywe, lecz nonsensowne. A zatem bezpośrednio dane są nam rzeczy, a nie treści. Przyjmowanie istnienia treści wyobrażeń (obrazów immanentnych, jakości zmysłowych) jest niczym nie uprawnioną spekulacją. Postulowana przez reistę eksplikacja terminu „spostrzegać" nie ma być oczywiście propozycją arbitralną. Ma odpowiadać znaczeniu (a raczej jednemu ze znaczeń), jakie termin ten ma w języku polskim — w szczególności w tej jego części, którą stanowi język teorii poznania. Pozwalać ma więc na sformułowanie tych problemów i twierdzeń teoriopoznawczych, które reista uznaje za sensowne. Choć większość twierdzeń na temat treści wyobrażeń ma swoje reistyczne parafrazy, są i takie, które ich nie mają i muszą być przez reistę zakwalifikowane jako bezsensowne; ich przykładem mogą być twierdzenia o takiej czy innej lokalizacji przestrzennej elementów treści. Dodajmy, iż trudno się nie zgodzić z reistą co do takiej ich kwalifikacji.

IV

1. Kończąc na tym przegląd argumentów przytaczanych na poparcie tez reistycznych, zapytajmy, czego one w rezultacie dowodzą. Czy za ich pomocą udało się reiście okazać prawdziwość swych twierdzeń? Czy którąkolwiek z tez reistycznych można uważać za dostatecznie uzasadnioną? Ograniczmy się do oceny tezy semantycznej jako tezy metodologicznie podstawowej. Ocena jej uzasadnienia musi brać pod uwagę jej logiczną strukturę. Teza ta — zarówno w postaci T_s jak i T_{sa} — jest twierdzeniem zawierającym dwa kwantyfikatory: ogólny i szczegółowy:

> *Wszelkie* sensowne (*resp.* prawdziwe) zdanie niereistyczne jest przekładalne na *pewne* zdanie reistyczne.

(Osłabiona wersja tej tezy, T_s^A czy $T_{sa}{}^A$, zawiera oprócz tego jeszcze jeden kwantyfikator szczegółowy.) Jest to więc rodzaj zdania, którego nie można ani udowodnić, ani obalić w sposób konkluzywny za pomocą skończonej liczby konkretnych przykładów. A tak właśnie, jak widzieliśmy, reista usiłuje uzasadnić swą tezę. Podejmowane przez niego próby obalenia rzekomych kontrprzykładów są próbami okazania przekładalności pewnych konkretnych zdań (lub schematów takich zdań) na język reistyczny. W moim przekonaniu są to próby udane; proponowane parafrazy reistyczne brzmią przekonująco. Nie da się jednak zaprzeczyć, że są to tylko sukcesy cząstkowe. Dotyczy to w szczególności tej dziedziny wiedzy, która z założenia ma charakter abstrakcyjny, nie zaś konkretny: matematyki, wraz z jej teoriomnogościowymi podstawami i empirycznymi zastosowaniami. Próby reistycznej parafrazy zdań matematycznych mają charakter fragmentaryczny i nie w pełni zadowalający; stanowią — zdaniem samego twórcy reizmu — nieznaczny jedynie „wyłom w zewnętrznym ogrodzeniu" gmachu matematyki (Kotarbiński 1958). Toteż uzasadnienie semantycznej tezy reizmu jest w najlepszym razie uzasadnieniem częściowym, pozwalającym przypisać jej niewielki stosunkowo stopień pewności. Obawiam się jednak, że w rzeczywistości sytuacja tej tezy przedstawia się gorzej — że są racje przemawiające za jej fałszywością. Racjami takimi są właśnie pewne kontrprzykłady z dziedziny matematycznej. Warto zwrócić uwagę na strukturę logiczną zdań stanowiących kontrprzykłady tezy semantycznej. Są to zdania zawierające kwantyfikator ogólny i głoszące, co następuje:

> *Z* jest sensownym (*resp.* prawdziwym) zdaniem niereistycznym nieprzekładalnym na *żadne* zdanie reistyczne.

Zdanie takie łatwo jest obalić, ale niełatwo udowodnić. Nie wystarczy do tego żaden skończony zbiór konkretnych przykładów. Dowód prawdziwości takiego zdania możliwy jest tylko na gruncie jakiejś uznanej za prawdziwą ogólnej teorii. Nic więc dziwnego, że dowody takie udało się podać tylko dla pewnych zdań matematycznych, bo te pozwalały na odwołanie się do odpowiedniej teorii matematycznej. Przykładem może być zdanie głoszące, iż liczba *M-ów* jest skończona. Jest to zdanie niereistyczne, które — w przeciwieństwie do zdań typu: liczba *M*-ów jest równa *n* — jest nieprzekładalne na żadne zdanie reistyczne. Tak przynajmniej może być interpretowany znany fakt matematyczny polegający na tym, iż klasa modeli skończonych nie jest elementarnie definiowalna (lub — w innym ujęciu — na tym, iż pojęcie skończoności nie jest pojęciem predykatywnym). To tylko jeden z przykładów zdań omawianego rodzaju.

2. Są, jak z tego widać, powody do sądzenia, iż teza reizmu (semantyczna, a tym bardziej ontologiczna) jest tezą fałszywą. Różnie można się ustosunkować do tego faktu i różne z niego wyciągnąć konsekwencje.

(1) Zgadzając się z tym, że semantyczna teza wzięta w całej ogólności jest tezą fałszywą, podkreślać możemy fakt, że jest to teza prawdziwa w pewnym zakresie ograniczonym, i dążyć do wyraźnego sformułowania owych ograniczeń tak, aby tezie w ten sposób osłabionej zapewnić prawdziwość.

(2) Prawdziwość tezie semantycznej zapewnić możemy również w sposób inny: przez traktowanie tej tezy jako definicji występującego w niej (*explicite* lub *implicite*) pojęcia sensowności. Teza semantyczna stwierdza przekładalność sensownych zdań niereistycznych na język reistyczny. W dotychczasowym ujęciu odwoływała się ona do «zastanego» (potocznego lub naukowego) pojęcia sensowności. Można jednakże potraktować ją inaczej: jako formułującą pewien definicyjny warunek sensowności. Wówczas zdanie nieprzekładalne na język reistyczny byłoby na mocy definicji wyrażeniem niesensownym. Motywem do przyjęcia takiej tezy definicyjnej mogłaby być ontologiczna teza reizmu — pod warunkiem, iż przyjęta została niezależnie od tezy semantycznej (jako «credo hipotetyczne» itp.). Sensowność, o którą tu chodzi, to nie sensowność czysto syntaktyczna, lecz semantyczna, lub może raczej filozoficzna. Poprawne gramatycznie zdania niereistyczne nie posiadające parafraz reistycznych byłyby sensowne tylko syntaktycznie, ale nie semantycznie, i jako takie pozbawione by były klasycznie rozumianych wartości prawdy i fałszu.

Zgodnie z takim poglądem, abstrakcyjna (w szczególności matematyczna) część naszej wiedzy pełni funkcję instrumentalną jedynie, nie zaś poznawczą. Wartość poznawczą ma tylko baza empiryczna nauki, utożsamiana tu z bazą reistyczną. Podkreślić należy, że w skład tak rozumianej części abstrakcyjnej wchodzi nie tylko matematyka czysta, ale i zmatematyzowane partie nauk empirycznych. Przykładem może być teoria pomiaru wielkości empirycznych. Sensowność semantyczna, a więc i wartość poznawcza, przysługiwałaby tylko temu, co się nazywa bazą empiryczną pomiaru. O tym, że jest to baza wystarczająco bogata, świadczy fakt, iż wyznaczenie wartości danej wielkości empirycznej wymaga odwołania się wyłącznie do pojęć empirycznych o charakterze jakościowym, takich jak pewna relacja słabego porządku czy operacja konkatenacji. Przy takim ujęciu rola abstrakcyjnej nadbudowy nauki sprowadza się do zadania właściwego sformułowania i systematyzacji owej reistycznie pojmowanej bazy. Może być bowiem tak, że aby powiedzieć wszystko o rzeczach, trzeba mówić nie tylko o rzeczach.

(3) Można jednak na semantyczną tezę reizmu spojrzeć jeszcze inaczej. Można traktować postulowaną w niej przekładalność na język reistyczny nie jako warunek sensowności czy prawdziwości, ale po prostu jako kryterium określonego «zaangażowania ontologicznego». Charakteryzuje ono szczególny rodzaj ontologii nominalistycznej, najbardziej, być może, rygorystyczny: ontologię rzeczy. Toteż jest sprawą o wielkiej filozoficznej doniosłości to, jaki rodzaj dyskursu i jaka dziedzina wiedzy spełniają owo ontologiczne kryterium. Reizm w takim ujęciu pojmowany być może jako program przełożenia — a raczej okazania przekładalności — na język reistyczny tego wszystkiego, co się przełożyć da (i tak też pojmował go jego twórca w końcowej „fazie rozwojowej"). Tylko próby realizacji tego programu ustalić mogą granice owego «minimalistycznego» zaangażowania ontologicznego.

Nie ulega dla mnie wątpliwości, że najdonioślejszym rezultatem osiągniętym w tej dziedzinie jest dokonana przez samego twórcę reizmu i zawarta w jego dziełach reistyczna parafraza podstawowych problemów i twierdzeń takich fundamentalnych teorii filozoficznych, jak teoria języka, poznania i nauki. Mimo nieuniknionej szkicowości jest to parafraza niezmiernie wnikliwa i głęboka, konsekwentna i jednolita. W przeciwieństwie do skrajnych propozycji pozytywistycznych nie likwiduje ona żadnych istotnych problemów filozoficznych; jeśli dezawuuje jakieś problemy jako bezsensowne, są to z całą pewnością problemy pozorne. W rezultacie umożliwia przedstawienie owych teorii filozoficznych z rzadko spotykaną jasnością, precyzją i prostotą. Toteż jakiekolwiek by były dalsze losy programu reistycznego, reistyczna filozofia języka, poznania i nauki pozostanie osiągnięciem trwałym i imponującym — zasługującym na najwyższą uwagę i uznanie.

PRACE CYTOWANE

K. Ajdukiewicz: 1930, „Reizm", *Przegląd Filozoficzny*, 33.

K. Ajdukiewicz: 1934, „W sprawie «uniwersaliów»", *Przegląd Filozoficzny*, 37.

J. Kotarbińska: 1967, „Kłopoty z istnieniem. Rozważania z zakresu semantyki", [w:] *Fragmenty filozoficzne*, Seria trzecia, Warszawa.

T. Kotarbiński: 1929, *Elementy teorii poznania, logiki formalnej i metodologii nauk*, Lwów.

T. Kotarbiński: 1930, „Realizm radykalny", *Przegląd Filozoficzny*, 33.

T. Kotarbiński: 1930/31, „Uwagi na temat reizmu", *Ruch Filozoficzny* nr 1-10.

T. Kotarbiński: 1935, „Zasadnicze myśli pansomatyzmu", *Przegląd Filozoficzny*, 38.

T. Kotarbiński: 1949, „O postawie reistycznej, czyli konkretystycznej", *Myśl Współczesna*, z. 10(41).

T. Kotarbiński: 1952, „Odpowiedź", *Myśl Filozoficzna*, nr 2(4).

T. Kotarbiński: 1958, „Fazy rozwojowe konkretyzmu", *Studia Filozoficzne*, nr 4(7).

EPISTEMOLOGIA

Filozofia Nauki
Rok I, 1993, Nr 2-3

O pewnych filozoficznych konsekwencjach semantycznej definicji prawdy

Mówiąc o semantycznej definicji prawdy mam na myśli tę eksplikację klasycznego pojęcia prawdy — pojęcia utożsamiającego prawdziwość ze „zgodnością z rzeczywistością" — która, sformułowana po raz pierwszy przez Tarskiego, stała się odtąd podstawowym elementem współczesnej semantyki logicznej. Eksplikacja ta uchodzi powszechnie za filozoficznie neutralną, nie przesądzającą żadnych filozoficznych rozstrzygnięć. Upatrywano w tym zarówno jej zaletę, jak i wadę — zależnie od tego, czego od owej definicji oczekiwano. Wskazując pewne filozoficzne konsekwencje semantycznej definicji prawdy, zaprzeczam tym samym temu tradycyjnemu przeświadczeniu o jej filozoficznej neutralności.

Definicja ta formułowana bywa w dwóch co najmniej wersjach. Jedna z nich, charakterystyczna dla współczesnej semantyki języków sformalizowanych, pociąga, moim zdaniem, pewne filozoficzne konsekwencje dotyczące zakresu tych wypowiedzi językowych, którym przysługiwać ma prawda lub fałsz, rozumiane zgodnie z proponowaną w owej definicji eksplikacją. Konsekwencje te podważają mianowicie rozpowszechnione — zwłaszcza w środowisku logicznym — przekonanie o zasadniczej dychotomii ogółu zdań oznajmujących na zdania o charakterze opisowym, wyposażone w określoną wartość logiczną — a więc prawdziwe lub fałszywe, i zdania o charakterze oceniającym (lub normatywnym), którym żadna wartość logiczna nie przysługuje — które nie są ani prawdą, ani fałszem. Otóż wydaje się, że wspomniana definicja prawdy stawia oba rodzaje zdań w tej samej w zasadzie sytuacji. Istnieją z jednej strony zdania opisowe pozbawione wartości logicznej, z drugiej strony zdania oceniające, którym taką wartość możemy przypisać. Spróbuję naszkicować argumentację prowadzącą do takiej konkluzji.[1]

DWIE WERSJE SEMANTYCZNEJ DEFINICJI PRAWDY

Semantyczna definicja zdania prawdziwego sformułowana w sposób ścisły przybiera, jak wiadomo, postać tzw. definicji indukcyjnej. Jej warunek wyjściowy określa prawdziwość zdań prostych («atomowych»), a warunek indukcyjny prawdziwość zdań złożonych. Ten ostatni może być też rozumiany jako warunek określający własności semantyczne stałych logicznych służących do budowy zdań złożonych. Sposób interpretacji owego warunku indukcyjnego może być również przedmiotem dyskusji ujawniającej różnice stanowisk nieobojętne pod względem filozoficznym. Konsekwencje filozoficzne, o których chcę mówić obecnie, związane są jednak przede wszystkim ze sposobem interpretacji warunki wyjściowego. Do niego też ograniczę się w zasadzie w tych rozważaniach.

Istnieje pewien tradycyjny sposób prezentowania semantycznej definicji prawdy, wywodzący się od jej autora i stosowany do dziś w kontekstach dotyczących problemów pozaformalnych. Definicja ta, w szczególności jej warunek wyjściowy, przedstawiana bywa w sposób mniej więcej taki. Niech α będzie zdaniem prostym języka J, a 'α' jego nazwą w metajęzyku MJ. Jeśli zbiór zdań prawdziwych języka J oznaczymy przez Ver, warunek wyjściowy semantycznej definicji prawdy zapisujemy po prostu tak:

(1) $$`\alpha' \in Ver \text{ gdy } \alpha.$$

Weźmy dla przykładu najprostsze zdanie języka J składające się z jednoargumentowego predykatu P i nazwy a: $P(a)$. Warunek (1) głosi w tym przypadku, iż:

(1′) $$`P(a)' \in Ver \text{ gdy } P(a).$$

Przytaczam tu to znane dobrze sformułowanie po to, aby zwrócić uwagę na jedną jego właściwość. Otóż warunek (1) zakłada, że metajęzyk MJ, w którym formułujemy naszą definicję prawdy, obejmuje jako swoją część język J. Zdanie α (np. zdanie $P(a)$) stanowiące prawy człon równoważności (1) jest pewnym zdaniem języka J, którym posługujemy się tu jako zdaniem sensownym naszego metajęzyka MJ. Przyjmujemy tym samym z góry — nie poddając go jakiejkolwiek analizie — ten sposób interpretacji zdania α, który przysługuje mu w języku J. Jest to ta właściwość semantycznej definicji prawdy, która zapewnia jej, w omawianym ujęciu, filozoficzną neutralność. Każde i tylko takie zdanie języka J będzie zdaniem o określonej wartości logicznej, jakie ze względu na jego intuicyjny sens za takie uważamy — traktując je jako sensowny człon równoważności (1).

1) Konkluzja to zresztą nienowa. Istnieje zdań opisowych pozbawionych wartości logicznej postulował np. H. Mehlberg w książce *The Reach of Science*, 1958. Wartość logiczną przypisywali zdaniom oceniającym m.in. J. Kmita w artykule „Problem wartości logicznej ocen", *Studia Filozoficzne* 1(36), 1964, J. Vetulani w artykule „Wartość logiczna zdań wartościujących", tamże 2(45), 1966.

Ale semantyczna definicja prawdy przybiera również postać odmienną. Mam na myśli sposób, w jaki definicja ta ujmowana bywa na gruncie współczesnej semantyki języków sformalizowanych. Istota owego ujęcia leży, moim zdaniem, w tym, że samo pojęcie interpretacji języka *J* staje się tu przedmiotem analizy — i to analizy odwołującej się do ściśle określonego aparatu pojęciowego. Interpretacja sformalizowanego języka *J* pojęta jest jako pewien twór teoriomnogościowy $\mathfrak{M}$, zwany modelem języka *J*. Model taki utożsamiany bywa, jak wiadomo, z układem złożonym z niepustego zbioru, stanowiącego universum języka *J*, oraz z relacji określonych w tym zbiorze i wybranych jego elementów, stanowiących denotacje predykatów i nazw języka *J*. Otóż definicja zdania prawdziwego języka *J* zrelatywizowana jest, w tym ujęciu, do danej interpretacji tego języka, pojętej jako jego model $\mathfrak{M}$. Definiuje się więc przede wszystkim relatywne pojęcie zdania języka *J* prawdziwego w modelu $\mathfrak{M}$. Niech *Ver* ($\mathfrak{M}$) symbolizuje zbiór takich zdań. Warunek wyjściowy omawianej definicji możemy sformułować w postaci schematu:

(2a) $`\alpha' \in Ver(\mathfrak{M})$ gdy $\alpha(\mathfrak{M})$

gdzie $\alpha(\mathfrak{M})$ jest symbolicznym zapisem tego, co głosi zdanie α w interpretacji wyznaczonej przez model $\mathfrak{M}$. O co tu idzie, najłatwiej wyjaśnić na prostym przykładzie. Załóżmy, że denotacją predykatu *P* jest w modelu $\mathfrak{M}$ zbiór (relacja jednoczłonowa) **P**, a denotacją nazwy *a* przedmiot **a**. Szczególnym przypadkiem schematu (2a) będzie przy tych założeniach równoważność:

(2a′) $`P(a)' \in Ver(\mathfrak{M})$ gdy $\mathbf{a} \in \mathbf{P}$,

gdyż $P(a)$ ($\mathfrak{M}$) — to w tym przypadku tyle, co — $\mathbf{a} \in \mathbf{P}$.

Przejście od tego relatywnego pojęcia prawdy do tradycyjnego pojęcia absolutnego dokonuje się przez wyróżnienie spośród wszystkich możliwych interpretacji języka *J*, możliwych modeli tego języka, jego interpretacji właściwej, jego modelu właściwego, oraz przez utożsamienie prawdziwości dowolnego zdania języka *J* z jego prawdziwością w modelu właściwym. Oznaczając ów model właściwy języka *J* przez $\mathfrak{M}^*$, otrzymujemy w ten sposób jako warunek wyjściowy definicji prawdy równoważność następującą:

(2b) $`\alpha' \in Ver$ gdy $\alpha(\mathfrak{M}^*)$,

w której symbol $\alpha(\mathfrak{M}^*)$ ma sens ten sam, co poprzednio. I tak, jeśli przyjmiemy, że właściwą denotacją predykatu *P* jest zbiór $\mathbf{P}^*$, a nazwy *a* przedmiot $\mathbf{a}^*$, warunek prawdziwości zdania $P(a)$ przybierze postać:

(2b′) $`P(a)' \in Ver$ gdy $\mathbf{a}^* \in \mathbf{P}^*$.

Główna — z naszego punktu widzenia — różnica między warunkiem typu (2b) a warunkiem typu (1) polega na tym, że w przeciwieństwie do równoważności (1) prawy człon równoważności (2b) nie jest, ogólnie biorąc, identyczny ze zdaniem α: jest zdaniem, które głosi to, co głosi zdanie α w interpretacji wyznaczonej przez model $\mathfrak{M}^*$. Ta interpretacja wyrażona jest w pewnym ściśle określonym języku — przy pomocy pojęć teoriomnogościowych, takich jak pojęcie elementu, zbioru, relacji itp. Zdanie $\alpha(\mathfrak{M}^*)$ może być co najwyżej uważane za przekład zdania α na język teorii mnogości;

ściślej — teorii mnogości wzbogaconej o nazwy konkretnych indywiduów, zbiorów, relacji. Nie zakładamy tu więc, jak poprzednio, że nasz metajęzyk *MJ*, w którym formułujemy definicję prawdy, zawiera język *J* w jego oryginalnej postaci: *MJ* zawiera przekład języka *J* na język «stosowanej» teorii mnogości. Oczywiście, w pewnych przypadkach szczególnych sam język *J* może być językiem tego typu. Tak było właśnie w przypadku języka, dla którego Tarski konstruował po raz pierwszy swą semantyczną definicję prawdy: był to po prostu język pewnego fragmentu teorii mnogości. Dla języka takiego obie wyróżnione tu wersje tej definicji w gruncie rzeczy się pokrywają. Rozchodzą się one w sposób istotny dopiero wtedy, gdy w grę wchodzą języki odbiegające wyraźnie od języków teoriomnogościowego typu. I wtedy dopiero powstaje problem wyboru którejś z tych wersji. Za każdą z nich przemawiają określone racje. Nie zamierzam ich tu przytaczać i dyskutować, bo celem tych uwag nie jest bynajmniej uzasadnienie wyższości którejś z wyróżnionych wersji. Chcę jedynie zwrócić uwagę na pewne konsekwencje płynące z akceptacji wersji drugiej — nazwijmy ją umownie teoriomodelową. Warto z tej racji jednak uświadomić sobie pewne jej własności, które zarazem mogą być interpretowane jako argumenty przemawiające na jej korzyść.

Pojęcie prawdy — to podstawowe pojęcie semantyczne. Jako takie dotyczy ono pewnego stosunku zachodzącego pomiędzy wyrażeniami językowymi a rzeczywistością, do której się te wyrażenia odnoszą. W przypadku pojęcia prawdy ów stosunek — to, w myśl tradycyjnej formuły, zgodność zdania z rzeczywistością, o której w tym zdaniu mowa. Otóż ten semantyczny charakter pojęcia prawdy znajduje we współczesnej — teoriomodelowej — wersji jej definicji bezpośredni i precyzyjny wyraz. U podstawy tej definicji leży owo relatywne pojęcie prawdy, symbolizowane przez formułę: $\alpha \in Ver(\mathfrak{M})$ — zdanie α jest prawdziwe w modelu $\mathfrak{M}$. Jest to formuła o tej samej strukturze, co wszelkie formuły symbolizujące pojęcia semantyczne. Klasycznym przykładem takiego pojęcia może być pojęcie denotowania, wyrażalne m.in. przez zwrot: predykat *P* denotuje zbiór **P**. W zwrocie głoszącym, iż zdanie α jest prawdziwe w modelu $\mathfrak{M}$, oba człony stosunku semantycznego: wyrażenie językowe α i fragment rzeczywistości $\mathfrak{M}$ wymienione są *explicite*. Co więcej, oba zidentyfikowane są w sposób możliwie precyzyjny. Dotyczy to nie tylko członu językowego: zdania α, ale również członu pozajęzykowego: modelu $\mathfrak{M}$. Fragment rzeczywistości, o którym mówi zdanie α, zostaje tu opisany przy pomocy tych środków, jakie stoją do dyspozycji semantyka konstruującego swój metajęzyk *MJ*. Do środków tych należy bogaty i precyzyjny aparat pojęciowy współczesnej teorii mnogości, która nie jest niczym innym, jak ogólną formalną teorią rzeczywistości. Przy pomocy jej pojęć zidentyfikowany zostaje nie tylko fragment rzeczywistości, o którym mówi zdanie α — model właściwy języka *J*, $\mathfrak{M}^*$, ale i to, co o tym fragmencie rzeczywistości głosi zdanie α, stan rzeczy stwierdzany przez to zdanie — $\alpha(\mathfrak{M}^*)$. Ów stan rzeczy opisujemy tu przy pomocy tych wszystkich środków, którymi dysponujemy, w języku, którym mówimy sami, a nie — jak to ma miejsce w ujęciu tradycyjnym — wyłącznie przy pomocy środków, którymi dysponują użytkownicy języka J, wyłącznie w ich języku. Mówiąc obrazowo, semantyk budujący teoriomodelową definicję prawdy patrzy na oba człony stosunku semantycznego — na język i rzeczywistość — z zewnątrz. Semantyk reprezentujący jej tradycyjną wersję zajmuje taki punkt widzenia tylko w stosunku do członu językowego; rzeczywistość pozajęzykową widzi wyłącznie od wewnątrz, poprzez medium badanego przez siebie języka J. Różnie można oceniać tę charakterystykę. Ja skłonny byłbym dopatrywać się w niej racji preferujących teoriomodelową wersję semantycznej definicji prawdy. Niezależnie od tego sądzę, że jest to w każdym razie wersja, która stanowi — w stosunku do wersji tradycyjnej — alternatywną propozycję, w pełni zasługującą na rozpatrzenie. Spróbujmy zatem prześledzić niektóre jej konsekwencje. Są to konsekwencje związane z charakterystycznymi własnościami owego teorio-mnogościowego metajęzyka MJ. Mówiąc najkrócej, jest to język precyzyjny i ekstensjonalny. Każda z tych własności pociąga określone filozoficzne konsekwencje.

WARTOŚĆ LOGICZNA ZDAŃ O TERMINACH NIEOSTRYCH

Precyzja metajęzyka MJ polega, z interesującego nas tu punktu widzenia, na tym, iż język ten nie zawiera żadnych terminów niejednoznacznych i nieostrych. Jakże zatem zbudować w nim mamy definicję prawdy dla języków, które — jak większość języków empirycznych — obfitują w wyrażenia obu tych rodzajów? Jedynym wyjściem w przypadku obecności w języku J terminów wieloznacznych wydaje się rozbicie takiego języka przy jego rekonstrukcji logicznej na szereg języków jednoznacznych. Skoro bowiem zdanie α jest wypowiedzią wieloznaczną, żadne — z założenia jednoznaczne — zdanie $\alpha(\mathfrak{M}^*)$ nie może w sposób adekwatny spełniać warunku (2b). Obecność w języku J terminów nieostrych sugeruje inne rozwiązanie. Zwróćmy przede wszystkim uwagę na fakt, iż nieostrość pewnego terminu języka J, np. predykatu P, uniemożliwia utożsamienie interpretacji właściwej tego języka z określonym modelem $\mathfrak{M}^*$. Denotacją predykatu P w modelu $\mathfrak{M}^*$ musi być określony zbiór, a nieostrość predykatu P na tym właśnie polega, że predykat ten żadnego określonego zbioru nie denotuje. Aby przedmiot jakiś był zbiorem w sensie teoriomnogościowym, musi być ustalone dokładnie, co jest, a co nie jest jego elementem. Nie ma takiego przedmiotu, jak «zbiór nieostry». W tej sytuacji denotację predykatu P utożsamić wypada nie tyle z określonym zbiorem, co z określoną rodziną zbiorów. Każdy z tych zbiorów odpowiada pewnej możliwej klasyfikacji przedmiotów z zakresu nieostrości predykatu P na podpadające pod ten predykat i podpadające pod jego negację. Otrzymujemy w rezultacie jako interpretację właściwą języka J nie jeden model tego języka, $\mathfrak{M}^*$, lecz rodzinę takich modeli, M^*.

Istnieją dla tak zinterpretowanego języka J różne, teoretycznie możliwe sposoby zdefiniowania absolutnego pojęcia prawdy.[2] Najbardziej pod względem filozoficznym przekonujące wydaje się rozwiązanie następujące. Przy założeniu, że M^* stanowi

2) Rozważałem je w artykule „Z semantyki pojęć otwartych", *Studia Logica* 15, 1964.

rodzinę modeli właściwych języka J, prawdziwe zdania tego języka utożsamiamy ze zdaniami prawdziwymi w każdy modelu rodziny M^*, a fałszywe — ze zdaniami fałszywymi w każdym modelu tej rodziny. Przypomnijmy dla uproszczenia, że rodzina M^* obejmuje skończoną liczbę modeli języka J: $M^* = \{\mathfrak{M}^*_1, ..., \mathfrak{M}^*_n\}$. Posługując się wprowadzoną uprzednio symboliką, warunek wyjściowy definicji prawdy dla języka J sformułować możemy w tym przypadku jak następuje:

(3b) ‘α’ $\in$ *Ver* gdy $\alpha(\mathfrak{M}^*_1)$ i ... i $\alpha(\mathfrak{M}^*_n)$.

Pojęcie zdania fałszywego języka J wymaga obecnie odrębnej definicji. Jej warunek wyjściowy głosi:

(3c) ‘α’ $\in$ *Fls* gdy $\sim\alpha(\mathfrak{M}^*_1)$ i ... i $\sim\alpha(\mathfrak{M}^*_n)$.

Mówiąc krótko, zdanie α jest prawdziwe, gdy jest prawdziwe przy każdej dopuszczalnej interpretacji; jest fałszywe, gdy jest fałszywe przy każdej takiej interpretacji. Chcąc zilustrować te pojęcia na przykładzie najprostszego zdania $P(a)$ założymy, że nazwa a denotuje jedyny przedmiot $\mathbf{a}^*$ (jest więc nazwą ostrą), a predykat P denotuje rodzinę zbiorów składającą się z elementów $\mathbf{P}^*_1, ..., \mathbf{P}^*_n$. Przy tych założeniach warunki (3b) i (3c) sprowadzają się do równoważności:

(3b′) ‘$P(a)$’ $\in$ *Ver* gdy $\mathbf{a}^* \in \mathbf{P}^*_1 \cap ... \cap \mathbf{P}^*_n$;

(3c′) ‘$P(a)$’ $\in$ *Fls* gdy $\mathbf{a}^* \in \mathbf{P}^*_1 \cap ... \cap \mathbf{P}^*_n$,

gdzie $\mathbf{P}^*_i$ ($i = 1, ..., n$) symbolizuje dopełnienie zbioru $\mathbf{P}^*_i$ do universum modeli właściwych języka J (przyjmujemy dla prostoty, że wszystkie modele rodziny M^* mają wspólne universum).

Definicje te kwalifikują wszystkie zdania orzekające predykat P o przedmiotach spoza zakresu nieostrości tego predykatu jako zdania prawdziwe lub fałszywe. Zdania, które przypisują ten predykat przedmiotom należącym do jego strefy nieostrości, zostają, w myśl tych definicji, uznane za wypowiedzi pozbawione wartości logicznej. Zbiór *Ver* $\cup$ *Fls* zdań o określonej wartości logicznej nie wyczerpuje więc ogółu zdań języka J. W każdym języku zawierającym terminy nieostre istnieją zdania opisowe — czyli wyrażenia należące do syntaktycznej kategorii zdań, zbudowane wyłącznie z terminów opisowych — które nie są ani prawdziwe, ani fałszywe. Łatwo zauważyć, iż jeśli zdaniem takim jest zdanie α, jest nim również jego negacja $\sim\alpha$. Traci tym samym, w zastosowaniu do takich języków, walor metalogiczne prawo wyłączonego środka głoszące, iż z dwóch zdań sprzecznych co najmniej jedno jest prawdziwe. Konsekwencji tej skłonny byłbym przypisywać pewne znaczenie filozoficzne, bo znaczenie takie ma, w moim przekonaniu, samo zjawisko nieostrości. Widzę w nim nieusuwalną, istotną cechę wszelkiego języka empirycznego, uwarunkowaną właściwościami zarówno samej rzeczywistości, jak i naszego aparatu poznawczego. Zjawisko to przejawia się na różnych szczeblach poznania i w różnych warstwach języka — w jego warstwie obserwacyjnej i teoretycznej. Notoryczna nieostrość terminów obserwacyjnych jest odbiciem ciągłości charakterystycznej dla wszelkich zmian przestrzennych i czasowych. Nieostrość terminów teoretycznych, zwana ich „otwartością”, ma z kolei inne źródła. Terminy te odnoszą się z założenia do pewnych własności ukrytych, które przejawiają się w naszym doświadczeniu tylko w określonych okolicznościach. Stąd też ich związki definicyjne z terminami obserwacyjnymi mają z reguły charakter jedynie cząstkowy. Rezultatem owej cząstkowej definiowalności terminów teoretycznych jest ich swoista nieostrość, sprowadzająca się do opisanego wyżej faktu posiadania wielu, częściowo różnych interpretacji. Tak więc, uwagi nasze dotyczące wartości logicznej zdań o terminach nieostrych wydają się mieć znaczenie dla całości naszej wiedzy empirycznej.

WARTOŚĆ LOGICZNA ZDAŃ OCENIAJĄCYCH

Drugą — obok «ostrości» — wyróżnioną przez nas właściwością metajęzyka *MJ* jest jego ekstensjonalność. Jest to, mówiąc swobodnie, język zbiorów, a nie własności. Interpretacja wyrażeń języka *J* pojmowana jest na gruncie owego metajęzyka *MJ* jako przyporządkowanie tym wyrażeniom odpowiednich tworów teoriomnogościowych. I tak, interpretacja jednoargumentowego predykatu *P* wyznaczona jest, jak widzieliśmy, całkowicie przez przyporządkowanie mu jako denotacji określonego zbioru przedmiotów. Od tego wyłącznie, co to jest za zbiór, tj. z jakich jest złożony elementów, zależy prawdziwość czy fałszywość zdań utworzonych przy pomocy tego predykatu. A w każdym razie od tego wyłącznie zależy wartość logiczna zdań prostych typu $P(a)$. Ta właściwość metajęzyka *MJ*, w którym formułujemy teoriomodelową definicję prawdy, pociąga pewne filozoficzne konsekwencje, spośród których na szczególną uwagę zasługują konsekwencje dotyczące problemu wartości logicznej zdań oceniających (ocen i norm). W moim przekonaniu, ów ekstensjonalny charakter metajęzyka *MJ* zaciera różnice semantyczne między predykatami opisowymi a predykatami oceniającymi i, co za tym idzie, różnice między zdaniami opisowymi a zdaniami oceniającymi dotyczące ich wartości logicznej. Chciałbym tu bronić stanowiska traktującego predykaty oceniające jako szczególny rodzaj predykatów nieostrych i przyznającego zdaniom oceniającym wartość prawdy lub fałszu w tym samym sensie, w jakim przyznaliśmy ją zdaniom o terminach nieostrych.

Weźmy pod uwagę jako klasyczny przykład zdania oceniającego najprostszą ocenę etyczną, orzekającą np., że dany czyn jest moralnie dobry, a więc zdanie typu $P(a)$, w którym a jest nazwą konkretnego czynu, a P symbolizuje predykat „moralnie dobry". Nie widać, moim zdaniem, powodu, aby predykatowi takiemu odmawiać denotacji tego samego typu, jaką posiadają nieostre predykaty opisowe. Charakterystyczny dla języka polskiego sposób rozumienia predykatu „moralnie dobry" wyróżnia pewne konkretne czyny jako podpadające pod ten predykat, inne jako podpadające pod jego negację, nie przesądzając jednocześnie w stosunku do pozostałych jednoznacznych rozstrzygnięć. Wydaje się, że mamy na tej podstawie prawo przypisać predykatowi „moralnie dobry" jako jego denotację pewną rodzinę zbiorów, których elementami są konkretne czyny, podobnie jak to czyniliśmy w stosunku do wszelkich predykatów nieostrych. A jeśli tak, to prawdziwość i fałszywość zdań orzekających ten predykat możemy określić tak samo, jak określaliśmy je dla analogicznych zdań opisowych. Stanowisko takie zakłada więc, mówiąc ogólnie, że wszelki predykat oceniający *P* posiada denotację w postaci pewnej rodziny, zbiorów, złożonej — dajmy na to — z elementów: $\mathbf{P}_1^*$, ..., $\mathbf{P}_n^*$, a prawdziwość i fałszywość zawierających ten predykat zdań typu *P*(*a*) definiują warunki (3b′) i (3c′), podane uprzednio. Otóż, przy powyższych założeniach, niektórym zdaniom tego typu będzie przysługiwała określona wartość logiczna, o ile tylko któryś ze zbiorów: $\mathbf{P}_1^* \cap ... \cap \mathbf{P}_n^*$ lub $\mathbf{P}_1^* \cap ... \cap \mathbf{P}_n^*$ okaże się niepusty. Ale pustość obu tych zbiorów oznacza całkowitą nieostrość predykatu *P*. A zatem, jeśli tylko dany predykat oceniający nie jest terminem całkowicie nieostrym, istnieć muszą prawdziwe lub fałszywe zdania orzekające ten predykat. Na gruncie takiego stanowiska, odmówienie wszelkim zdaniom oceniającym określonej wartości logicznej staje się równoważne stwierdzeniu całkowitej nieostrości wszelkich predykatów oceniających.

Jest to stwierdzenie, które jak sugeruje przykład predykatu „moralnie dobry", wydaje się trudne do obrony. Chciałbym przy tej okazji zwrócić uwagę na konieczność każdorazowej relatywizacji tej kwestii do określonego języka. To, czy predykat *P* jest, czy nie jest całkowicie nieostry, zależy od tego, czy traktujemy go jako element tego czy innego języka, gdyż ten sam predykat może być na gruncie różnych języków różnie rozumiany. W przypadku predykatów oceniających szczególnie ważne, z tego punktu widzenia, jest wyróżnienie tzw. języków indywidualnych — języków pojmowanych jako języki poszczególnych osób. Nie ulega bowiem wątpliwości, że predykaty oceniające odznaczają się na gruncie danego języka indywidualnego znacznym stopniem ostrości, nie różniąc się pod tym względem zasadniczo od predykatów opisowych. Wątpliwości w tej sprawie mogą powstawać tylko w odniesieniu do pewnych języków grupowych — np. języków etnicznych. Tak też ujmowaliśmy problem nieostrości predykatu „moralnie dobry" — relatywizując go po prostu do języka polskiego. Wydaje się jednak, że i przy takiej relatywizacji teza o całkowitej nieostrości wszelkich predykatów oceniających nie znajduje uzasadnienia. Wrócimy jeszcze do sprawy nieostrości predykatów oceniających pod koniec tych rozważań. Ale wbrew pozorom nie ta teza stanowi decydujący argument dla tych, którzy kwestionują stosowalność określeń prawdy i fałszu do zdań oceniających. Nie dlatego odmawia się zdaniom orzekającym dany predykat oceniający określonej wartości logicznej, że uważa się ów predykat za termin całkowicie nieostry. Jakież zatem decydują o tym względy? Nie sposób tu przedstawiać, nawet w wielkim skrócie, przytaczanej na poparcie takiego stanowiska argumentacji. Argumentacja to bardzo rozbudowana, bardzo różnorodna, a w dodatku niejasna i trudno uchwytna.[3] Toteż ograniczę się tu do wskazania w sposób ogólnikowy i intuicyjny jednego tylko jej wątku, pozostającego w bezpośrednim związku z omawianą przez nas problematyką. A oto szkic owego rozumowania.

3) W sposób zwięzły przedstawia ją m.in. M. Fritzhand w artykule „Zagadnienie prawdy w etyce", *Studia Filozoficzne* 2(45), 1966.

Zdanie prawdziwe — to zdanie zgodne z rzeczywistością. Z jaką jednak rzeczywistością ma być zgodne prawdziwe zdanie oceniające? To, co zdanie stwierdza, i to, jaka jest rzeczywistość, charakteryzuje się, na gruncie referowanego stanowiska, w języku własności: jako układ takich a nie innych cech. Jakież to cechy rzeczywistości odzwierciedlają zdania oceniające? Do czego, w szczególności, odnosi się predykat oceniający? Dwie istnieją tu możliwości odpowiedzi. Możemy przyjąć, że predykat taki, podobnie jak predykaty opisowe, odnosi się do pewnych cech empirycznych, lub, jak się tu zwykło mówić, «naturalnych». Ale wówczas wpadamy w tzw. błąd naturalizmu, utożsamiając znaczenie predykatów oceniających ze znaczeniem pewnych predykatów empirycznych, czemu wydają się przeczyć nasze intuicje językowe. Możemy też postulować istnienie swoistych cech nieempirycznych, «pozanaturalnych» (danych nam w bezpośrednim poznaniu intuicyjnym), jako tych elementów rzeczywistości, do których odnoszą się predykaty oceniające. Ale wówczas narażamy się na zarzut nieodpowiedzialnej spekulacji. Warunkiem stosowalności klasycznego pojęcia prawdy do ocen moralnych ma być „wskazanie tego elementu obiektywnego rzeczywistości, z którym odnośna ocena moralna jest zgodna bądź niezgodna (...) I to takiego elementu rzeczywistości, że nikt nie mógłby rozsądnie przeczyć, iż jest on faktycznie jej elementem i że do niego to oceny moralne się odnoszą (...)”.[4] A to właśnie, jak widzimy, okazuje się niewykonalne. Ale niewykonalne tylko przy przyjętym sposobie identyfikacji elementów rzeczywistości. Przyjęcie w tej roli aparatu pojęciowego teorii mnogości pozwala bez trudu wskazać poszukiwane elementy. Rzeczywistość jest tu charakteryzowana w języku zbiorów, a nie własności. To, do czego odnosi się predykat oceniający — to jego denotacja, pojęta jako zbiór konkretnych indywiduów: rzeczy, osób, zdarzeń (a raczej jako rodzina takich zbiorów). Tylko od tego, jaki zbiór (czy rodzinę zbiorów) denotuje dany predykat, zależy prawdziwość zawierających ten predykat zdań. A to, z jakim zbiorem mamy do czynienia, zależy z kolei tylko od tego, z jakich składa się elementów. Ten sam zbiór przedmiotów charakteryzować można, jak wiadomo, przy pomocy różnych układów cech. Mogą to być cechy empiryczne lub nieempiryczne, «naturalne» lub «pozanaturalne». Taki czy inny sposób charakterystyki zbioru stanowiącego denotację danego predykatu oceniającego nie wpływa na wartość logiczną zdań orzekających ten predykat. Warunki prawdziwości zdań typu $P(a)$, sformułowane w ekstensjonalnym metajęzyku MJ, są «nieczułe» na zmiany sposobu charakterystyki zbioru $\mathbf{P}^*$ (czy rodziny zbiorów $\mathbf{P}^*_i$), jeśli tylko są to charakterystyki zakresowo równoważne.

To natomiast, co od sposobu takiej charakterystyki istotnie może być zależne — to znaczenie danego predykatu i zawierających go zdań, a nie ich wartość logiczna. Problem takiej a nie innej charakterystyki danej denotacji — to problem należący do teorii znaczenia, a nie teorii prawdy, zgodnie z przyjętym we współczesnej semantyce sposobem rozumienia tych określeń. Toteż w rozważaniach naszych, ograniczonych z

4) M. Fritzhand, tamże.

założenia do teorii prawdy, zagadnień dotyczących znaczenia predykatów oceniających omawiać nie zamierzam. Nie można jednak ominąć całkowicie pytania dotyczącego sposobu, w jaki denotacje predykatów oceniających mogą być wyznaczane — skoro istnienie takich denotacji zakładamy.

Nie mam tu zamiaru rozstrzygać, w jaki sposób owe denotacje faktycznie są wyznaczane. Chcę jedynie w paru słowach wspomnieć o zarysowujących się w tej sprawie możliwościach. Są one, moim zdaniem, analogiczne do tych, jakie istnieją w przypadku predykatów opisowych. Wyróżnia się tam w zasadzie dwa sposoby interpretacji danego predykatu, tj. wyznaczania jego denotacji: «pozawerbalny» i «werbalny». Interpretacja pozawerbalna znajduje zastosowanie wobec predykatów obserwacyjnych, przybierając na ogół postać tzw. definicji ostensywnej. Interpretacja werbalna wyznacza denotacje predykatów teoretycznych, wiążąc je poprzez tzw. postulaty znaczeniowe ze zinterpretowanymi predykatami obserwacyjnymi. Otóż sądzę, że w przypadku predykatów oceniających dałoby się wyróżnić podobne możliwości. Interpretacja pozawerbalna traktowałaby predykaty oceniające jako swego rodzaju predykaty obserwacyjne i wyznaczała ich denotację (w sposób z reguły wieloznaczny) przez demonstrowanie (niekoniecznie naoczne) odpowiednio dobranych przypadków wzorcowych — np. wzorców «moralnie dobrego» zachowania. Różnica między tą procedurą a definicją ostensywną polegałaby na tym, że w odróżnieniu od tej ostatniej procedura obecna apelowałaby nie do naszej zdolności postrzegania zmysłowego, lecz do naszej intuicji moralnej. Byłaby więc jakimś odpowiednikiem stanowiska intuicjonistycznego. Interpretacja werbalna z kolei ujmowałaby predykaty oceniające podobnie jak predykaty teoretyczne i wyznaczała ich denotację (też z reguły w sposób wieloznaczny) przy pomocy postulatów znaczeniowych wiążących je ze zinterpretowanymi predykatami empirycznymi. Mogłaby tym samym uchodzić za odpowiednik stanowiska naturalistycznego. Oczywiście, i w przypadku pozawerbalnej interpretacji predykatów oceniających dysponowalibyśmy pewnymi twierdzeniami ustalającymi związki między tymi predykatami a predykatami empirycznymi. Tyle tylko, że byłyby to twierdzenia faktyczne (zdania syntetyczne). Przy werbalnej interpretacji predykatów oceniających natomiast pewne z nich przybierałyby charakter twierdzeń definicyjnych (zdań analitycznych). Mimo to nie jestem pewien, czy stanowisku takiemu można by zarzucić błąd naturalizmu. Po pierwsze bowiem, owe związki definicyjne nie sprowadzałyby się na ogół do definicji predykatów oceniających w języku empirycznym, lecz stwierdzałyby pewne zależności luźniejsze (odpowiadające np. tzw. definicjom redukcyjnym) — tak, jak to ma miejsce w przypadku typowych predykatów teoretycznych. Po drugie, jako związki charakteryzujące denotację predykatu oceniającego mogłyby być interpretowane czysto ekstensjonalnie, a więc jako twierdzenia postulujące zawieranie się tej denotacji w denotacji określonego predykatu empirycznego itp. Nie wydaje się więc, aby przyjęcie takich postulatów znaczeniowych pociągało nieuchronnie tezę o równoznaczności predykatów i zdań oceniających z pewnymi predykatami i zdaniami empirycznymi.

Nie chciałbym tu ryzykować hipotezy przesądzającej, który z owych sposobów interpretacji predykatów oceniających jest sposobem właściwym (być może zresztą stosowane są oba). Sprawa wymaga na pewno szczegółowych i wnikliwych badań. Chcę natomiast dać wyraz swemu przekonaniu, że pewne tradycyjne problemy filozoficzne — takie, jak spór naturalizmu z intuicjonizmem — na tej płaszczyźnie znajdują właściwą interpretację. W zarysowanym tu schemacie pojęciowym dotyczą one, najogólniej mówiąc, tego, w jaki sposób zostaje wyznaczona rodzina modeli właściwych M^* dla badanego przez nas języka J. Istnienie takiej rodziny — i to rodziny spełniającej określone warunki (warunki wykluczające całkowitą nieostrość predykatów oceniających) — wydaje mi się faktem nie budzącym wątpliwości. Sprawą otwartą natomiast pozostaje problem właściwego sposobu jej charakterystyki.

Chciałbym na koniec wspomnieć o jednym jeszcze rodzaju argumentacji kwestionującej w pewnym sensie kwalifikację prawdziwościową zdań oceniających. Nie jest to argumentacja stwierdzająca zasadniczą niestosowalność pojęć prawdy i fałszu do zdań oceniających. Zakres tych zastosowań uważa się tu jednak za tak wąski, że trudno przypisywać mu jakąkolwiek — teoretyczną lub praktyczną — doniosłość. Mimo iż nie traktuje tu się predykatów oceniających jako terminów całkowicie nieostrych, uważa się je za terminy nieostre w stopniu tak znacznym, że ich zakres nieostrości obejmuje wszelkie niebanalne — teoretycznie bądź praktycznie interesujące — przypadki zastosowań. W rezultacie wartość logiczną przyznaje się tylko tym zdaniom orzekającym owe predykaty, które mają jawnie banalny, bo niekwestionowalny charakter. Wszystkie stwierdzenia niebanalne — w szczególności kontrowersyjne — będą wypowiedziami pozbawionymi wartości logicznej. Argumentem uzasadniającym tę konkluzję ma być właśnie kontrowersyjność takich wypowiedzi — rozbieżność praktyki stosowania danych predykatów oceniających do tych przypadków, o których w tych wypowiedziach mowa.

Myślę, że nie jest to argument zniewalający: że z faktu owej rozbieżności zastosowań można zdać sprawę w inny, w moim poczuciu, bardziej adekwatny sposób. Trzeba tylko przyjąć, że stosowanie predykatów oceniających nie jest procedurą niezawodną — podobnie zresztą jak stosowanie predykatów empirycznych. Toteż gdy w danym przypadku ktoś stosuje predykat P, a inny jego negację, nie musi to świadczyć o tym, że jest to przypadek z zakresu nieostrości tego predykatu; po prostu jedno z tych zastosowań jest błędne. Obaj rozmówcy — użytkownicy tego samego języka J, do którego należy predykat P — mogą rozumieć ów predykat w ten sam sposób, tę samą wiązać z nim denotację; przypadek rozważany może do tej denotacji istotnie należeć; mimo to jeden z nich może go — prawdziwie — kwalifikować jako P, drugi — fałszywie — jako nie-P. Ten błąd może mieć różne źródła — zależnie od tego, na jakiej drodze ów rozmówca doszedł do swego twierdzenia. Przyczyną błędu może być mylna intuicja. Nasze intuicje w sprawach wartości — m.in. intuicje moralne — nie są bynajmniej nieomylne. Prowadzą do prawdy tylko w pewnych sprzyjających (wewnętrznych i zewnętrznych) okolicznościach — podobnie jak to ma miejsce w przypadku spostrzeżeń zmysłowych. Ale tą przyczyną może być również oparcie się na fałszywym twierdzeniu faktycznym. Predykaty oceniające stosujemy niejednokrotnie w sposób pośredni, odwołując się do pewnych twierdzeń faktycznych wiążących je z predykatami empirycznymi. Otóż twierdzenia te (jak również pewne akceptowane przez nas twierdzenia empiryczne) mogą być, i często bywają, twierdzeniami fałszywymi — rezultatem pochopnych uogólnień, nieracjonalnych wierzeń, uprzedzeń itp. Tak więc, fakty kontrowersji w dziedzinie oceniania niekoniecznie świadczą o szczególnej nieostrości predykatów oceniających; mogą być po prostu świadectwem naszych błędów.

Przy rozważaniu sprawy nieostrości predykatów oceniających warto pamiętać o dwóch prostych wskazaniach. Po pierwsze, nie należy mylić nieostrości z wieloznacznością. Okazując nieostrość predykatu „moralnie dobry" przytaczamy przykłady czynów ocenianych przez jednych jako dobre, przez innych jako złe. Tymczasem nie ulega dla mnie wątpliwości, że w ocenach tych słowo „dobry" występuje niejednokrotnie w znaczeniach różnych — bynajmniej nie zawsze znacząc tyle, co „moralnie dobry". Tym samym przykłady te chybiają celu. Po drugie, nie należy osądzać pod względem nieostrości wszystkich predykatów oceniających *en bloc*. Różnią się one bowiem co do stopnia nieostrości ogromnie. Podobnie jak to ma miejsce w przypadku predykatów empirycznych, wyróżniają się pod tym względem szczególnie korzystnie predykaty porównawcze. O ileż ostrzejszy od klasyfikującego predykatu „moralnie dobry" okazuje się porównawczy predykat „moralnie lepszy"! W o ile mniejszym stopniu dotyczą go zarzuty wysuwane wobec tamtego! Tak więc, i z tego punktu widzenia stanowisko odmawiające zdaniom oceniającym wartości logicznej traci sporo ze swej pozornej oczywistości.

Broniąc w ten sposób tezy o stosowalności pojęć prawdy i fałszu do zdań oceniających, chciałbym się zastrzec przeciwko niewłaściwej jej interpretacji. Uważając niektóre przynajmniej zdania oceniające za prawdziwe lub fałszywe, nie przypisuję im tym samym bynajmniej charakteru twierdzeń naukowych. Nie każde zdanie prawdziwe zasługuje na miano twierdzenia naukowego. O naukowości danego zdania decyduje, mówiąc najogólniej, sposób jego uzasadniania, związany z kolei z rodzajem przysługującego mu znaczenia. A pod tym względem widzę między zdaniami oceniającymi a twierdzeniami naukowymi różnice zasadnicze. One to — a nie różnice w sprawach wartości logicznej — przesądzają o wartości poznawczej obu typów twierdzeń. Uwagi powyższe nie likwidują więc tej tradycyjnej i ważnej problematyki — umieszczają ją jedynie na innej nieco płaszczyźnie.

Filozofia Nauki
Rok I, 1993, Nr 2-3

Pojęcie prawdy w językach nauk empirycznych

Semantyczną teorię prawdy Tarskiego uważa się zgodnie za fundament współczesnej semantyki logicznej. Nie ma natomiast analogicznej zgody co do tego, jaki jest właściwy zakres stosowalności tej teorii i opartych na niej konstrukcji. Niektórzy zakres ten skłonni są ograniczać do języków i teorii matematycznych, za czym przemawia — nawiasem mówiąc — pierwotne brzmienie tytułu dzieła Tarskiego: *Pojęcie prawdy w językach nauk dedukcyjnych*. Inni rozciągają go na ogół języków o określonej strukturze syntaktycznej, niezależnie od tego, jakich dziedzin języki te miałyby dotyczyć — zgodnie z późniejszą, obcojęzyczną wersją owego tytułu: *Pojęcie prawdy w językach sformalizowanych*. Faktem jest, że — niezależnie od tego, jaki jest zakres możliwych zastosowań tej teorii — dotychczasowe jej zastosowania dotyczyły głównie języka nauk matematycznych. W tej dziedzinie koncepcja ta okazała się narzędziem niezmiernie skutecznym, umożliwiającym stworzenie tak ścisłej i bogatej teorii semantycznej, jaką jest współczesna teoria modeli. Próby wykroczenia poza ową dziedzinę — stosowania semantycznej teorii prawdy do języków innego, niematematycznego typu — datują się stosunkowo od niedawna i stale jeszcze są przedmiotem ożywionych kontrowersji. Dotyczy to również prób objęcia ową teorią języka nauk empirycznych. Te właśnie próby stanowić mają przedmiot naszych rozważań. Ograniczymy jednak ich zakres w sposób maksymalny. Chcąc skoncentrować się na tym, co jest charakterystyczne dla języka empirycznego jako takiego — a więc na tych własnościach, które wspólne są wszystkim językom empirycznym i odróżniają je od języków matematycznych — przedmiotem naszych rozważań uczynimy języki pewnych, dostatecznie prostych i ubogich, teorii empirycznych. Rezygnujemy więc z omawiania stosowalności semantycznej teorii prawdy do języka całej naszej wiedzy empirycznej. Rezygnujemy tym bardziej z omawiania tego zagadnienia w stosunku do całości języka naturalnego. Pomijamy tym samym szereg tych właściwości języków empirycznych, z którymi usiłowały się uporać znane z literatury próby budowania dostosowanej do tych języków semantycznej teorii prawdy. Należą tu przede wszystkim właściwości takie, jak: obecność w języku wyrażeń intensjonalnych (modalnych, deontycznych, epistemicznych itp.), okazjonalnych (np. temporalnych), wieloznacznych, występowania pustych nazw i deskrypcji, a wreszcie związana z uniwersalnością języka obecność wyrażeń metajęzykowych, w szczególności — odnoszących się do danego języka terminów semantycznych. Są to z pewnością ważne cechy języka naturalnego, a w konsekwencji i języka wielu teorii empirycznych, nie można ich jednak uważać za charakteryzujące wszelki dyskurs empiryczny. Istnieją niewątpliwie teorie empiryczne — fizyczne czy biologiczne — których język wolny jest od tego rodzaju «przypadłości» lub przynajmniej poprzez stosowną rekonstrukcję daje się od nich uwolnić. Co zatem charakteryzuje wszelki język empiryczny, odróżniając go od języka matematyki? Najkrócej można by powiedzieć, że jest to język, który mówi o świecie danym nam — bezpośrednio lub pośrednio — w doświadczeniu; który odnosi się do rzeczy danych nam w doświadczeniu i informuje o tym, jakie te rzeczy są. Mówiąc nieco dokładniej, wszelki język empiryczny zawiera terminy odnoszące się do rzeczy danych w doświadczeniu — krótko, terminy empiryczne, i zdania, których wartość logiczna zależna jest od doświadczenia — inaczej, zdania empiryczne. Obie te cechy odróżniają go od języka teorii matematycznych — języka matematyki «czystej». Toteż semantyka logiczna języka empirycznego, w szczególności semantyczna teoria prawdy dla takiego języka, nie może być mechanicznym przeniesieniem na tę dziedzinę teorii dotyczącej dziedzin matematycznych. Aby móc zdać sprawę z semantycznych odrębności języka empirycznego, ta ostatnia wymaga stosownej modyfikacji i rozwinięcia.

Semantyczna teoria prawdy utożsamiana tu jest z jej standardową wersją teoriomodelową. Opiera się ona na określonej teorii języka i określonej teorii rzeczywistości przez ten język opisywanej. Pierwsza — to logiczna składnia języków standardowo sformalizowanych, druga — to teoriomnogościowa ontologia. Pojęciem podstawowym jest pojęcie zdania prawdziwego danego języka sformalizowanego, zrelatywizowane do określonej interpretacji tego języka przyporządkowującej jego terminom odpowiednie twory teoriomnogościowe. Dla języka L najprostszego typu, zawierającego predykaty P_1, ..., P_n jako jedyne terminy pozalogiczne, struktura stanowiąca jego możliwą interpretację, zwana również modelem języka L, przybiera — jak wiadomo — postać układu:

$$\mathfrak{M} = \langle U, R_1, ..., R_n \rangle,$$

przyporządkowującego zmiennym języka L zbiór U jako ich zakres, a predykatom P_1, ..., P_n odpowiednie relacje R_1, ..., R_n, określone na zbiorze U jako ich denotacje. Zdefiniowane indukcyjnie pojęcie prawdy dla języka L zrelatywizowane do modelu $\mathfrak{M}$ głosi, swobodnie mówiąc, co następuje: zdania α języka L jest prawdziwe w modelu $\mathfrak{M}$, gdy jest tak, jak głosi α w interpretacji wyznaczonej przez $\mathfrak{M}$. Oprócz tego relatywnego pojęcia prawdy, Ver($\mathfrak{M}$), aparat pojęciowy semantyki logicznej zawiera również pojęcie prawdy, które nazwać można absolutnym, Ver. Jego definicja opiera się na założeniu, że wśród możliwych interpretacji danego języka istnieje interpretacja rzeczywista, inaczej — właściwa lub zamierzona. Zakłada się w konsekwencji, że wśród ogółu modeli języka L istnieje model zamierzony, $\mathfrak{M}^*$, wyznaczający tę interpretację, i definiuje zdanie prawdziwe jako zdanie prawdziwe w owym modelu zamierzonym:

$$\alpha \in \text{Ver}, \text{gdy } \alpha \in \text{Ver}(\mathfrak{M}^*).$$

Czy aparat powyższy znajduje zastosowanie w semantyce języka empirycznego? Odpowiedź, którą chciałbym tu przedstawić, jest w zasadzie odpowiedzią twierdzącą. Utrzymuje przede wszystkim, że relatywne pojęcia semantyczne, w tym relatywne pojęcie prawdy, stosują się bez jakichkolwiek istotnych zastrzeżeń do tych języków empirycznych, do których ograniczyliśmy tutaj nasze rozważania. Jako języki prostych teorii empirycznych dają się one wtłoczyć w schemat języków standardowo sformalizowanych, a to pozwala już zastosować do nich wspomniany aparat semantyczny, bo aparat ten jest związany z określoną strukturą syntaktyczną, wspólną dla języków dotyczących dziedzin różnych. Tak więc nic nie stoi na przeszkodzie, aby możliwa interpretacja języka empirycznego L omawianego typu utożsamiana była z jego modelem $\mathfrak{M}$, a prawdziwość zdania α języka L w modelu $\mathfrak{M}$ rozumiana zgodnie z jej ogólną definicją. To, co wedle reprezentowanego tu stanowiska różni w sposób istotny semantykę języka empirycznego od semantyki języka matematycznego, związane jest z absolutnym pojęciem prawdy — w szczególności z pojęciem interpretacji zamierzonej, do której się owo pojęcie prawdy odwołuje. To, jaka to jest interpretacja i jak zostaje wyznaczona, decyduje o empirycznym charakterze danego języka.

Na czym więc polega różnica w tym względzie między językiem matematycznym a empirycznym? Warto przede wszystkim zauważyć, że absolutne pojęcie prawdy pełni w semantyce języków matematycznych rolę ograniczoną. Najważniejsze problemy semantyczne dotyczące języków tego typu angażują jedynie pojęcie relatywne. Co więcej, w stosunku do języków wielu teorii matematycznych pojęcie absolutne pozbawione jest jakiegokolwiek zastosowania. Ma ono zastosowanie tylko wobec języków faktycznie zinterpretowanych, odnoszących się do określonych z góry dziedzin. A tego właśnie nie da się powiedzieć o obszernej klasie teorii matematycznych. Teoria grup może być tutaj typowym przykładem. Nie jest to teoria żadnej specyficznej dziedziny, raczej — pewnego ogólnego pojęcia. Jedyną odpowiedzią na pytanie, co teoria grup ma opisywać, jest odpowiedź tautologiczna: każdą dziedzinę, która jest grupą; innymi słowy — każdą strukturę, która jest modelem jej aksjomatów. Toteż jedynym — oprócz relatywnego — pojęciem prawdy, które znajduje w tym przypadku zastosowanie, jest pojęcie zdania prawdziwego w każdym modelu tej teorii (pokrywające się zakresowo z pojęciem jej twierdzenia). Nie wszystkie teorie matematyczne są teoriami tego typu. Istnieją wśród nich i takie, które mogą być traktowane jako teorie pewnych wyróżnionych dziedzin. Klasycznym przykładem jest arytmetyka liczb naturalnych. To, co teoria ta ma opisywać, może być określone jako system liczb naturalnych. System ten nie jest zdefiniowany jako struktura będąca modelem aksjomatów tej teorii. Wśród klasy takich modeli znajdują się również struktury inne, wyraźnie niezamierzone. Struktura zamierzona wyróżniona jest za pomocą pojęć teorii mnogości, a więc niezależnie od danej teorii. Wyznaczając zamierzoną interpretację języka arytmetyki, pozwala ona na wprowadzenie dla zdań tego języka absolutnego pojęcia prawdy: zdanie arytmetyczne jest po prostu prawdziwe, gdy jest prawdziwe przy tej właśnie interpretacji, tj. jako zdanie mówiące o systemie liczb naturalnych.

Otóż założeniem omawianego tu stanowiska jest przekonanie o tym, iż języki empiryczne przypominają język arytmetyki raczej niż teorii grup. Są to języki zinterpretowane, tj. wyposażone w interpretację zamierzoną, a odwołująca się do niej definicja absolutnego pojęcia prawdy stanowi istotny element ich semantyki. Pewne wątpliwości może budzić ta teza w stosunku do języków teorii, takich jak: cybernetyka czy teoria systemów, które w pewnym swym ujęciu przypominają teorię grup raczej niż arytmetykę. Ale czy w takim ujęciu są to istotnie teorie empiryczne? Czy może raczej teoriami empirycznymi są jedynie ich konkretne zastosowania? Wątpliwości takie nie powstają w stosunku do tych teorii empirycznych, do których ograniczyliśmy te rozważania. Ich język jest niewątpliwie językiem zinterpretowanym, a charakterystyka jego interpretacji zamierzonej stanowi jedno z głównych zadań jego semantyki.

Sprawa ta wymaga dodatkowych wyjaśnień, istnieje bowiem takie ujęcie owego zadania, przy którym staje się ono najzupełniej banalne. Ujęcie to odwołuje się do znanego schematu Tarskiego:

(T) $$\alpha \in \mathrm{Ver}, \text{ gdy } \alpha,$$

który traktowany być może nie tylko jako kryterium adekwatności semantycznej definicji prawdy, ale również jako część samej definicji, mianowicie jako jej warunek wyjściowy, definiujący pojęcie prawdy dla prostych, tj. atomowych, zdań języka L. Schemat (T) zakłada, że metajęzyk ML, w którym formułowana jest definicja prawdy dla języka przedmiotowego L, zawiera ów język przedmiotowy jako swoją część. Warunki prawdziwości dla zdań języka L wyrażone tu są w tymże języku L. Ta właściwość schematu (T) ma swój odpowiednik w teoriomodelowej wersji definicji prawdy, odwołującej się do pojęcia modelu zamierzonego. Polega tu na zdefiniowaniu modelu zamierzonego $\mathfrak{M}^*$ języka L za pomocą predykatów $P_1, ..., P_n$ samego języka L oraz uniwersalnego predykatu P_U, co do którego również założyć możemy, że należy do języka L (traktując $P_U(x)$ jako skrót dowolnej formuły uniwersalnej, np $x = x$):

$$\mathfrak{M}^* = \langle P_U, P_1, ..., P_n \rangle$$

Zgodnie z tym stanowiskiem, interpretacja zamierzona języka L przyjęta jest jako z góry dana, nie wymagająca żadnej rekonstrukcji czy eksplikacji. Język L wyposażony w tę interpretację traktowany jest przez semantyka jako jego język własny.

Zadania semantycznej teorii prawdy dla danego języka L mogą być jednak — i bywają — pojmowane w sposób szerszy. Nie przyjmuje się zamierzonej interpretacji języka L jako czegoś z góry danego, lecz ją właśnie czyni się przedmiotem analizy. Analiza ta musi spełniać określone warunki. Ma być przeprowadzona w metajęzyku ML, który — jako język semantyka, stanowiący jego własny środek komunikacji — musi czynić zadość określonym wymaganiom co do swego charakteru syntaktycznego i semantycznego. W szczególności, jest to zawsze język zakładający określoną ontologię i eksplikacja zamierzonej interpretacji języka L na gruncie tej właśnie ontologii ma być zrekonstruowana. Eksplikacja ta winna nadto spełniać pewne warunki pozaformalne. Powinna, najogólniej mówiąc, pozwalać na wyjaśnienie podstawowych własności semantycznych języka badanego. Tak też jest w podanym przykładzie arytmetyki liczb naturalnych. Model zamierzony języka tej teorii — system liczb naturalnych — zdefiniowany jest w ściśle określonym języku: za pomocą pojęć ogólnej teorii mnogości. Jego definicja stanowi tym samym eksplikację interpretacji zamierzonej na gruncie tej podstawowej, zakładanej przez semantyka, teorii. Zarazem eksplikacja ta zdaje sprawę z charakterystycznej własności semantycznej języka arytmetycznego — jego aprioryczności. Przy przyjętej interpretacji wszystkie twierdzenia arytmetyczne sprowadzają się do twierdzeń teoriomnogościowych. Zyskują więc taki sam status jak te ostatnie — a te na ogół uważa się za twierdzenia aprioryczne. Dotyczy to w szczególności aksjomatów arytmetyki liczb naturalnych wyrażonych w języku tak zinterpretowanym. Są to twierdzenia prawdziwe *a priori* (w jednym ze znaczeń tego wieloznacznego zwrotu). Ich prawdziwość wynika na gruncie ogólnej teorii mnogości z samej definicji ich interpretacji zamierzonej, inaczej mówiąc — z samego określenia tego, do czego się owe twierdzenia mają odnosić.

Analogicznie przedstawiają się zadania semantycznej definicji prawdy dla języków empirycznych. Interpretacja zamierzona języka empirycznego L, niezbędna dla zdefiniowania absolutnego pojęcia prawdy, winna zostać określona za pomocą tych środków formalnych i treściowych, którymi dysponuje metajęzyk ML, i to określona tak, aby zdać sprawę z empirycznego charakteru języka L. Język semantyki teoriomodelowej — a więc i metajęzyk ML — to język charakteryzujący się klasycznym rachunkiem logicznym i teoriomnogościową ontologią. Na gruncie tego właśnie aparatu pojęciowego ma być zrekonstruowana interpretacja zamierzona języka L. Sposób jej rekonstrukcji ma przy tym umożliwić wyjaśnienie semantycznego charakteru języka L: obecności w tym języku terminów i zdań empirycznych. Jeśli w szczególności zakładamy, iż język L jest językiem pewnej teorii empirycznej T o aksjomatach A: $T = Cn(A)$, to do zdań empirycznych języka L należeć muszą m.in. aksjomaty A. Muszą to być zdania, których prawdziwość jest sprawą doświadczenia; nie może zatem być przesądzona przez sam sposób ich interpretacji. Tylko wtedy teorię T uważać możemy za teorię empiryczną. Ale jeśli tak, to sposób określenia interpretacji zamierzonej języka empirycznego różnić się musi zasadniczo od sposobów właściwych językom matematycznym. Na czym więc sposób ten polega? Wiele z tego, co zrobiono we współczesnej semantyce teorii empirycznych — a nawet szerzej, w metodologii formalnej nauk empirycznych — związane jest mniej lub bardziej bezpośrednio z tym zagadnieniem. Nie osiągnięto, rzecz zrozumiała, w tej sprawie jakiegoś niekwestionowalnego rozstrzygnięcia. Wysunięto raczej szereg propozycji konkurencyjnych, różniących się między sobą zarówno założeniami formalnymi, jak i filozoficznymi. Te ostatnie zwłaszcza szczególnie wyraźnie determinują kierunek rozwiązań. Nie mogąc przedstawiać tutaj rozwiązań wszystkich, ograniczę się do szkicowej prezentacji jednego z nich.[1]

Rozwiązanie to oparte jest na założeniach filozoficznych, które opatruje się niekiedy mianem semantycznego empiryzmu. Stanowisko to sprowadza procedury interpretacji języka empirycznego do dwóch zasadniczych rodzajów: interpretacji bezpośredniej i pośredniej. To, do czego się dany termin ma odnosić, może być bądź pokazane wprost, bądź opisane za pomocą innych terminów. Interpretacja bezpośrednia ma zastosowanie tylko wobec terminów odnoszących się do przedmiotów obserwowalnych. Terminy odnoszące się do przedmiotów niedostępnych obserwacji mogą być interpretowane tylko pośrednio. W typowych teoriach empirycznych występują terminy obu rodzajów. Zakłada się więc, że pewne terminy języka L, tzw. o-terminy, uzyskują interpretację poprzez procedury ostensywne, inne, tzw. t-terminy — poprzez postulaty języka L wiążące je z terminami poprzednimi. Wyróżnia się w ten sposób w języku empirycznym L język L_o, zawierający wyłącznie o-terminy, którego model zamierzony $\mathfrak{M}_o^*$ wyznaczony jest w drodze interpretacji bezpośredniej. Model zamierzony $\mathfrak{M}^*$ całego języka L określony jest z kolei jako takie przedłużenie modelu $\mathfrak{M}_o^*$ (ewentualnie — jego rozszerzenia), które jest modelem postulatów P dla t-terminów języka L.

Taka koncepcja interpretacji zamierzonej języka L rodzi swoiste problemy i trudności. Pewne z nich związane są z wyodrębnieniem zbioru postulatów P. Sytuacją typową bowiem wydaje się przypadek, gdy interpretacja t-terminów teorii empirycznej T wyznaczona jest przez ogół aksjomatów A-tej teorii, gdy zatem $P = A$. Ale wówczas aksjomaty te tracą charakter zdań empirycznych. Nie mogą się okazać fałszywe na podstawie doświadczenia, bo wyklucza to sama definicja modelu zamierzonego języka L. Próby rozwiązania tej trudności podkreślają podwójną funkcję aksjomatów teorii empirycznej. Aksjomaty te stwierdzają pewne fakty doświadczalne, m. in. ogólne prawidłowości empiryczne, a zarazem «konstytuują znaczenie» t-terminów, wyznaczając ich interpretację zamierzoną. Powstaje w związku z tym problem rozbicia zbioru aksjomatów A na dwa składniki: faktualny A_F, spełniający tylko pierwszą funkcję, i konwencjonalny A_C, spełniający tylko drugą. Precyzacja tych pojęć i zbadanie możliwości wyodrębnienia tak sprecyzowanych składników stanowiły przedmiot szeregu prac kontynuujących klasyczne w tej dziedzinie badania Carnapa.[2] Doprowadziły one do wniosku, że omawiane pojęcia dopuszczają precyzacje nieco różne, bo różne nieco kryją się za nimi intuicje. Okazało się też, że chociaż istnieją pewne przypadki szczególne, w których wyodrębnienie tych dwóch składników (w którymkolwiek z ich

1) Rozwiązanie to przedstawione zostało szerzej w innych moich pracach; por. *The Logic of Empirical Theories*, London 1969; „Problem interpretacji języka empirycznego w ujęciu teoriomodelowym", *Studia Filozoficzne* nr 1, 1972.

2) Por. m. in. M. Przełęcki i R. Wójcicki, „The Problem of Analyticity", *Synthese*, nr 19, 1969; M. Przełęcki i R. Wójcicki, „Inessential Parts of Extensions of First-Order Theories", *Studia Logica*, nr 28, 1971; P. M. Williams, „On the Conservative Extensions of Semantical Systems: A Contribution to the Problem of Analyticity", *Synthese*, nr 25, 1973.

sensów) nie jest możliwe, w przypadkach pozostałych, które wydają się typowe dla rzeczywistych teorii empirycznych, daje się ono przeprowadzić w sposób zadowalający. Istnieje wówczas możliwość utożsamienia zbioru postulatów P dla t-terminów danej teorii empirycznej T ze składnikiem konwencjonalnym A_C jej aksjomatów A: $P = A_C$, co prowadzi, jak się zdaje, do adekwatnej charakterystyki interpretacji zamierzonej języka L teorii T. Jest to w każdym razie interpretacja nadająca aksjomatom tej teorii status zdań empirycznych, falsyfikowalnych na podstawie doświadczenia.

Inny rodzaj problemów, jakie nasuwa omawiana charakterystyka interpretacji zamierzonej języka L, związany jest z faktem jej niejednoznaczności. Zarówno model zamierzony języka L_o, jak i języka L, okazuje się wyznaczony w sposób wieloznaczny. W fakcie tym znajduje wyraz istotna cecha każdego języka empirycznego, różniąca go od języka matematyki: jego nieostrość lub — ogólniej — niedookreśloność. Jest to cecha nieodłącznie związana z empirycznością języka — z tym że odnosi się on nie do tworów abstrakcyjnych lecz do konkretnych, danych nam w doświadczeniu rzeczy. Na gruncie zakładanej w metajęzyku ML logiki i ontologii cecha ta przejawia się jako niejednoznaczność określenia modelu zamierzonego języka L, inaczej — jako wielość tak określonych modeli. Dotyczy to, choć z różnych niego względów, zarówno o-terminów, jak i t-terminów języka L. Analiza interpretacji bezpośredniej dowolnego o-predykatu P_i prowadzi do wniosku, iż interpretacja ta nie przyporządkowuje mu jako denotacji żadnego określonego zbioru w ścisłym, teoriomnogościowym sensie, bo dla pewnych elementów *universum* nie dostarcza żadnych kryteriów przynależności do owej denotacji. Jest to konsekwencja notorycznej nieostrości wszelkich o-terminów. W ramach teoriomnogościowej ontologii możemy zdać z tego faktu sprawę, przyporządkowując predykatowi P_i nie jeden zbiór, ale pewną klasę zbiorów, odpowiadających wszystkim możliwym sposobom klasyfikacji przedmiotów z zakresu nieostrości tego predykatu na P_i i nie-P_i. Otrzymujemy w rezultacie jako interpretację zamierzoną języka L_o nie jeden model $\mathfrak{M}_o^*$, ale pewną klasę modeli M_o^*. W konsekwencji, również interpretacja zamierzona całego języka L utożsamiona zostaje nie z jednym modelem $\mathfrak{M}^*$, lecz z pewną klasą modeli M^*, zgodnie z przyjętą poprzednio charakterystyką:

> $\mathfrak{M} \in M^*$, gdy $\mathfrak{M}$ jest przedłużeniem pewnego modelu $\mathfrak{M}_o \in M_o^*$ (ewentualnie jego rozszerzenia) spełniającym warunek $P \subset \mathrm{Ver}(\mathfrak{M})$,

gdzie P jest zbiorem postulatów dla t-terminów języka L. Biorąc pod uwagę charakter zbioru P (w szczególności jego nietwórczość ze względu na L_o), stwierdzić możemy, że klasa M^* jest co najmniej tak liczna jak klasa M_o^*. Z reguły jest to klasa znacznie liczniejsza, choćby dlatego, że typowe postulaty dla t-terminów nie sprowadzają się do ich definicji za pomocą o-terminów. Każdemu elementowi klasy M_o^* odpowiada w związku z tym więcej niż jeden element w klasie M^*. Na tym polega swoista niedookreśloność t-terminów. Jednym z jej przejawów jest podkreślany nieraz aproksymacyjny charakter wielkości fizycznych. W przeprowadzonej tu rekonstrukcji przejawia się on w tym, iż interpretacja zamierzona symbolu takiej wielkości fizycznej np. terminu „masa", przyporządkowuje temu terminowi nie jedną funkcję rzeczywistą, ale określoną klasę takich funkcji.

Zastąpienie modelu zamierzonego języka *L* klasą takich modeli stanowi sposób przezwyciężenia tego przeciwieństwa, jakie zachodzi między «ostrą» ontologią teoriomnogościową metajęzyka *ML* a «rozmytą» rzeczywistością doświadczalną opisywaną przez język *L*. Rozwiązując tę trudność, ujęcie to rodzi zarazem trudność inną: problem definicji absolutnego pojęcia prawdy dla języka *L* tak zinterpretowanego. Zamiast jedynego modelu zamierzonego $\mathfrak{M}^*$ — jak w sytuacji klasycznej — mamy tu klasę takich modeli M^*. Jaki sens możemy nadać stwierdzeniu głoszącemu, że zdanie języka *L* jest po prostu prawdziwe? Jak — innymi słowy — zdefiniować w tym przypadku zbiór Ver? W odpowiedzi na to pytanie wysunięto szereg propozycji, z których co najmniej dwie zasługują na szczególną uwagę.[3] Jedna z nich przypisuje absolutną wartość logiczną tylko tym zdaniom, których relatywna wartość logiczna pozostaje niezmienna we wszystkich modelach klasy M^*:

$\alpha \in$ Ver, gdy dla każdego $\mathfrak{M} \in M^*$, $\alpha \in$ Ver($\mathfrak{M}$),

$\alpha \in$ Fals, gdy dla każdego $\mathfrak{M} \in M^*$, $\alpha \in$ Fals($\mathfrak{M}$).

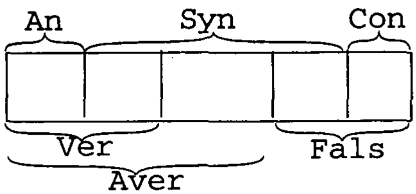
Zdania, które w pewnych modelach zamierzonych są prawdziwe, a w innych fałszywe, zostają tu uznane za zdania pozbawione wartości logicznej. Zgodnie z drugą, liberalniejszą propozycją, wszystkie te zdania uważane są za prawdziwe. Lepiej może, jak proponują niektórzy, mówić w tym przypadku o prawdziwości przybliżonej lub aproksymacyjnej, Aver. Jej definicja głosi zatem, co następuje:

$\alpha \in$ Aver, gdy dla pewnego $\mathfrak{M} \in M^*$, $\alpha \in$ Ver($\mathfrak{M}$).

Oba te pojęcia odpowiadają, jak się zdaje, pewnym istniejącym intuicjom i oba znajdują zastosowanie w pewnych sytuacjach.

Zarysowana tu konstrukcja pozwala również na wprowadzenie pojęć analitycznego, kontradyktorycznego i syntetycznego zdania języka *L*. Definicja klasy modeli zamierzonych M^* gwarantuje, jak widzieliśmy, prawdziwość wszystkich postulatów *P* i ich logicznych konsekwencji: Cn(*P*). Zdania te zaliczyć zatem możemy do zdań analitycznych języka *L*, An. Zdania, których negacje są zdaniami analitycznymi, nazwane być mogą z kolei zdaniami kontradyktorycznymi, Con. Ich fałszywość w dowolnym modelu klasy M^* jest z góry przesądzona. Te dwa rodzaje zdań wyczerpują klasę tych zdań języka *L*, których wartość logiczna zdeterminowana jest przez samą definicję interpretacji zamierzonej języka *L*. Wszystkie inne zdania języka *L* mogą być zaliczone do syntetycznych, Syn. Ich prawdziwość zależy od tego, jakimi okażą się modele klasy M_0 — to zaś, ze względu na sposób wyróżnienia tej klasy, uważane być może za sprawę doświadczenia. W rezultacie otrzymujemy następującą klasyfikację zdań języka *L*:

3) Por. m. in. M. Przełęcki, „Z semantyki pojęć otwartych". *Studia Logica*, nr 15, 1964; R. Wójcicki, „Semantyczne pojęcie prawdy w metodologii nauk empirycznych", *Studia Filozoficzne*, nr 3, 1969; R. Wójcicki, *Metodologia formalna nauk empirycznych*, Wrocław 1974.



M_0 — to zaś, ze względu na sposób wyróżnienia tej klasy, uważane być może za sprawę doświadczenia. W rezultacie otrzymujemy następującą klasyfikację zdań języka L:

Sądzić można, że uwzględnia ona ważniejsze typy zdań języka empirycznego wyróżniane pod względem ich wartości logicznej.

Filozofia Nauki
Rok I, 1993, Nr 2-3

Prawda

1. Pojęcie prawdy należy do podstawowych kategorii filozoficznych, a zagadnienie istoty prawdy stanowi jeden z naczelnych problemów filozoficznej teorii poznania. Wyrażenia „prawda", „prawdziwy" — podobnie jak „fałsz", „fałszywy" — używane bywają w różnych kontekstach i różnych znaczeniach. Niektóre tylko przypadki ich użycia mają charakter filozoficzny.

Podstawowy filozoficzny sens tych pojęć odnosi je do zdań lub sądów. Przez zdania rozumie się tu twierdzenia, tj. zdania wyposażone w określone znaczenie, a przez sądy — znaczenia takich zdań. Nie wszystkie jednak wypowiedzi stosujące termin „prawda" czy „fałsz" do tak rozumianych zdań lub sądów używają tych terminów w ich podstawowym filozoficznym sensie. Zwroty typu „prawdą jest, że *p*" lub „fałszem jest, że *p*", gdzie „*p*" symbolizuje dowolne zdanie, rozumiane bywają często jako zwroty równoznaczne odpowiednio z samym zdaniem „*p*" lub z jego zaprzeczeniem „nie *p*". Wyrażenia „jest prawdą" i „jest fałszem" pełnią tu rolę nie predykatów przypisujących pewne własności zdaniu „*p*", lecz jednoargumentowych spójników zdaniowych; pierwszy z nich odpowiada tzw. asercji zdania „*p*", drugi — jego negacji. Taka eksplikacja terminu „prawda" nazywana bywa niekiedy „nihilistyczną" teorią prawdy (Kotarbiński).

Filozoficzne eksplikacje tego pojęcia traktują prawdziwość jako autentyczną własność pewnych twierdzeń lub sądów. Zadaniem filozoficznej teorii prawdy jest odpowiedź na pytanie, na czym owa własność ma polegać. Rozmaite filozoficzne teorie różne na to pytanie dają odpowiedzi. Teorią dominującą, zarówno w dziejach filozofii, jak i wśród filozoficznych kierunków współczesności, jest teoria zwana korespondencyjną lub klasyczną. Mówiąc najkrócej, istotę prawdziwości upatruje ona w zgodności z rzeczywistością. Najwcześniejsze bodaj sformułowanie tej koncepcji zawarte jest w *Metafizyce* Arystotelesa: „jest fałszem powiedzieć o tym, co jest, że nie jest, lub o tym, co nie jest, że jest, jest prawdą powiedzieć o tym, co jest, że jest, lub o tym, co nie jest, że nie jest". Określenie Arystotelesa, ograniczone do zdań szczególnego rodzaju (mówiących o czymś, że to coś jest lub nie jest), uogólnione zostało na dowolne stwierdzenia w sposób wyżej wspomniany: twierdzenie prawdziwe — to twierdzenie zgodne z rzeczywistością. W znanym sformułowaniu scholastycznym, mówiącym o prawdziwości myśli raczej niż twierdzeń, definicja ta głosi: *veritas est adequatio rei et intellectus*.

2. Klasyczna koncepcja prawdy, odwołująca się do tego rodzaju określeń, stała się przedmiotem krytyki prowadzącej w efekcie do przeciwstawienia jej pewnych koncepcji nieklasycznych. Krytyka ta podkreślała metaforyczność sformułowań mówiących o zgodności zdania czy sądu z rzeczywistością i pytała, na czym właściwie owa zgodność zachodząca między tymi różnymi rodzajami bytów miałaby polegać. Nie widząc zadowalającej eksplikacji tego zagadkowego stosunku, postulowała zastąpienie go pojęciem bardziej uchwytnym i operatywnym. Takie miało być np. pojęcie zgodności danego zdania z ogółem zdań akceptowanych. Nie mogąc porównywać zdania z pozajęzykową rzeczywistością, o której zdanie to mówi, możemy porównywać je z innymi zdaniami, by stwierdzić np., czy zdanie to z nich wynika lub czy jest z nimi niesprzeczne. Otóż wedle tzw. koherencyjnej teorii prawdy na takiej właśnie zgodności polegać ma prawdziwość danego zdania. Niestety, charakter owej międzyzdaniowej zgodności nie został nigdy dokładniej określony. Najczęściej ma się na myśli zgodność z ogółem zdań, za którymi „opowiada się doświadczenie". Nie wiadomo jednak, czy zdanie prawdziwe to tylko takie, które z tamtych wynika, czy też i takie, które jest z nimi niesprzeczne. Każda z tych propozycji pociąga trudne do przyjęcia konsekwencje.

Inną nieklasyczną koncepcję prawdy przedstawia tzw. pragmatyczna teoria prawdy, w którejś ze swych licznych odmian. Teoria ta prawdziwość zdania, a raczej wyrażonego w nim przekonania, upatruje, mówiąc najogólniej, w jego pożyteczności. Według jednej z częściej spotykanych eksplikacji tej idei prawdziwość naszych przekonań polegać ma na tym, że prowadzą one do działań skutecznych, pozwalają osiągnąć zamierzony cel. Istotę prawdziwości utożsamiano również z oczywistością danego twierdzenia (różnie zresztą rozumianą) lub z powszechną na nie zgodą itp.

Wszystkie koncepcje nieklasyczne oparte są na wspólnej idei. Traktując zgodność zdania z rzeczywistością jako relację zagadkową i nieuchwytną, prawdziwość określają jako zgodność zdania z kryteriami stosowanymi w praktyce naukowej przy przyjmowaniu i odrzucaniu twierdzeń. Taka tylko relacja ma być dla nas dostępna i ona właśnie stanowić ma istotę prawdziwości. To, że dane zdanie jest prawdziwe, nie znaczy tu nic innego, jak to, że spełnia ono określone kryteria. W zależności od tego, co to są za kryteria i które z nich jest kryterium ostatecznym, otrzymujemy taką, a nie inną wersję nieklasycznej koncepcji prawdy. Kryterium prawdy staje się więc, wedle tej koncepcji, jej cechą definicyjną. To, po czym można poznać prawdziwość zdania, zostaje tu utożsamione z tym, na czym prawdziwość ma polegać. Na gruncie koncepcji klasycznej są to dwie sprawy różne. Definicja prawdy nie podaje jej kryteriów; istota prawdziwości jest określona niezależnie od sposobów jej poznawania.

3. Zgodność zdania czy sądu z rzeczywistością, stanowiąca istotę prawdziwości wedle koncepcji klasycznej, nie polega, rzecz jasna, na identyczności tych dwóch rodzajów bytów. Nie sprowadza się też, jak zarzucają niektórzy, do jakiegoś bliżej nieokreślonego ich podobieństwa. Wbrew krytykom teorii klasycznej relacja ta daje się określić w sposób dostatecznie jasny i przekonujący. Punkt wyjścia takiego określenia stanowić może pierwotne sformułowanie Arystotelesa, które w prosty sposób wyjaśnia, na czym owa zgodność zdania z rzeczywistością miałaby polegać. Mówiąc ogólnie i ogólnikowo zarazem, to, że jakieś twierdzenie jest zgodne z rzeczywistością, znaczy, że jest tak właśnie, jak to twierdzenie głosi. Klasyczną definicję prawdy wyrazić można zatem w sposób następujący: dane twierdzenie jest prawdziwe, gdy jest tak, jak to twierdzenie głosi.

Z takim sformułowaniem definicji prawdy związane są jednak pewne trudności natury logicznej, znajdujące swój wyraz w antynomiach semantycznych, z których tzw. antynomia kłamcy znana jest od czasów starożytnych. Antynomie te pokazują, że tego rodzaju określenie zdania prawdziwego, przyjmowane bez jakichkolwiek ograniczeń, prowadzi do sprzeczności.

Oto jedno ze sformułowań antynomii kłamcy: „Zdanie, które piszę w tej chwili, jest fałszywe". Jeżeli zdanie to jest prawdziwe, to jest tak jak ono głosi, czyli zdanie to jest fałszywe. Jeżeli zaś zdanie to jest fałszywe, to nie jest tak jak ono głosi, czyli zdanie to jest prawdziwe. Powstaje w ten sposób sprzeczność: zdanie powyższe jest prawdziwe wtedy i tylko wtedy, gdy jest fałszywe, a więc gdy nie jest prawdziwe.

Aby sprzeczności tej uniknąć, trzeba owo klasyczne określenie prawdziwości poddać precyzacji i stosownym ograniczeniom. Zadanie to po raz pierwszy zrealizował Tarski. W swej historycznej już dziś pracy [9], kładącej podwaliny współczesnej semantyki logicznej, zdefiniował w sposób precyzyjny i wolny od sprzeczności pojęcie zdania prawdziwego, które uważać można za eksplikację klasycznego pojęcia prawdy. Antynomii semantycznych unika Tarski dzięki odróżnieniu od danego języka, zwanego językiem przedmiotowym, jego metajęzyka, pojętego jako język służący do opisu języka przedmiotowego. Definicja prawdy dotyczy zdań języka przedmiotowego, sama zaś sformułowana jest w jego metajęzyku. Ten ostatni, obejmujący język przedmiotowy (lub jego przekład) jako swoją część, zawiera ponadto nazwy wyrażeń języka przedmiotowego, terminy odnoszące się do relacji syntaktycznych między tymi wyrażeniami i relacji semantycznych między wyrażeniami a tym, do czego się wyrażenia te odnoszą, oraz odpowiedni aparat logiczny. Zdanie, które przypisuje prawdziwość lub fałszywość zdaniu języka przedmiotowego, samo nie jest zdaniem języka przedmiotowego, lecz jego metajęzyka. Z punktu widzenia tych kryteriów zdanie, które o sobie samym orzeka prawdziwość lub fałszywość, nie jest zdaniem sensownym; nie powstaje więc w stosunku do niego problem prawdziwości i antynomia kłamcy traci swą podstawę. W podobny sposób unika się pozostałych antynomii semantycznych.

To, że semantyczna definicja prawdy stanowi eksplikację klasycznego pojęcia prawdy, znajduje wyraz w sformułowanym przez Tarskiego warunku jej merytorycznej trafności. Warunek ten żąda, aby konsekwencjami tej definicji były tzw. cząstkowe definicje prawdy, tj. metajęzykowe twierdzenia postaci: zdanie „*p*" jest prawdziwe wtedy i tylko wtedy, gdy *p*, dla każdego zdania „*p*" należącego do języka przedmiotowego. Stwierdzają one w sposób najprostszy to właśnie, co stanowi istotę klasycznego pojęcie prawdy: zdanie „*p*" jest prawdziwe, gdy jest tak, jak to zdanie głosi.

Merytorycznie trafną w sensie powyższym, a zarazem poprawną definicję prawdy zbudować można, jak pokazuje Tarski, dla każdego języka o ściśle ustalonej strukturze. Musi to być język, dla którego ustalony jest słownik wyrazów prostych oraz reguły budowania wyrażeń złożonych, w szczególności zdań, z wyrazów prostych — i to w sposób odwołujący się wyłącznie do kształtu wyrażeń. Kształt wyrażeń decydować ma też w sposób jednoznaczny o ich znaczeniu. Języki spełniające te warunki nazywane są językami sformalizowanymi.

Definicja prawdy dla języków sformalizowanych przybiera postać tzw. definicji rekurencyjnej. W językach takich prawdziwość zdań złożonych jest funkcją prawdziwości wchodzących w ich skład zdań prostych. W konsekwencji definicja taka składa się z dwóch kroków: pierwszy definiuje prawdziwość zdań prostych, drugi określa, w jaki sposób prawdziwość zdań złożonych zależy od prawdziwości zdań składowych. To, jaką konkretną postać przyjmują owe warunki definicyjne, zależy od rodzaju danego języka — od tego, z jakich wyrażeń zbudowane są jego zdania proste i za pomocą jakich sposobów utworzone są z tych ostatnich zdania złożone.

Przedmiotem współczesnej semantyki logicznej są przede wszystkim tzw. języki standardowo sformalizowane. Zdania proste tych języków zbudowane są z predykatów i odpowiedniej liczby argumentów nazwowych; w najprostszym przypadku może to być zdanie składające się z jednoargumentowego predykatu i jednej nazwy, a więc zdanie postaci „$P(a)$", (np. zdanie „Warszawa jest miastem"). Warunki określające prawdziwość zdań prostych przyjmują postać taką jak cząstkowe definicje prawdy. Dla zdania „$P(a)$" będzie to więc warunek:

(1) zdanie „$P(a)$" jest prawdziwe, gdy $P(a)$.

Warunki określające prawdziwość zdań złożonych ilustruje warunek dla negacji zdania „*p*":

(2) Zdanie „nie *p*" jest prawdziwe, gdy zdanie „*p*" nie jest prawdziwe.

4. Taka wersja semantycznej definicji prawdy, charakterystyczna dla cytowanej pracy Tarskiego, oparta jest na założeniu, że metajęzyk, w którym formułowana jest definicja prawdy dla danego języka przedmiotowego, zawiera ten ostatni jako swoją część. Założenie to zastępowane bywa we współczesnej semantyce logicznej warunkiem słabszym, żądającym, aby metajęzyk zawierał odpowiedni przekład języka przed-

miotowego. Metajęzyk, jako język teorii semantycznej, musi czynić zadość określonym wymaganiom co do swego charakteru syntaktycznego i semantycznego. W szczególności jest to język zakładający zawsze określoną ontologię. Toteż dane zdanie języka przedmiotowego, traktowane jako zdanie metajęzyka, winno być wyrażone za pomocą tych środków formalnych i treściowych, którymi dysponuje metajęzyk: winno być, w szczególności, sformułowane za pomocą aparatu pojęciowego zakładanej przez ten język ontologii. Nie będzie to więc, ogólnie biorąc, zdanie języka przedmiotowego w jego oryginalnej formie i interpretacji, lecz przekład tego zdania na tak określony metajęzyk.

Metajęzyk współczesnej semantyki logicznej — to język charakteryzujący się ontologią teoriomnogościową. Teoria mnogości stanowi ogólną formalną teorię rzeczywistości zakładaną przez semantykę logiczną; na język tej teorii przełożone tu zostaje każde zdanie języka przedmiotowego. Podstawowym pojęciem teorii mnogości jest pojęcie zbioru. Z jego pomocą definiuje się w konsekwencji to wszystko, do czego odnoszą się wyrażenia języka przedmiotowego. I tak, denotacje predykatów tego języka utożsamione zostają, na gruncie semantyki logicznej, z określonymi zbiorami — zbiorami indywiduów w przypadku predykatów jednoargumentowych, zbiorami n-elementowych układów indywiduów (czyli n-członowymi relacjami) w przypadku predykatów n-argumentowych. Denotacją predykatu „*P*" będzie wobec tego zbiór indywiduów *P*. Stąd teoriomnogościowym przekładem zdania „$P(a)$" będzie zdanie głoszące, iż przedmiot *a* jest elementem zbioru *P*: $a \in P$. Warunek określający prawdziwość zdania „$P(a)$" przybiera więc tutaj postać różną od klasycznej formuły Tarskiego:

(1′) zdanie „$P(a)$" jest prawdziwe, gdy $a \in P$.

W analogiczny sposób sformułowane zostają warunki określające prawdziwość innych zdań prostych języka przedmiotowego. W ten sposób całość tego, do czego odnosi się dany język przedmiotowy — dziedzina, o której mówi — zostaje utożsamiona z pewną teoriomnogościową strukturą. Jest nią układ złożony z uniwersum języka (tj. zbioru wszystkich przedmiotów, o których w danym języku mowa) oraz z denotacji jego wyrażeń prostych (np. odpowiednich zbiorów denotowanych przez jego predykaty). Struktura taka nazywana bywa modelem danego języka. Stąd też semantyka logiczna tak pojęta występuje zwykle pod nazwą teorii modeli, konstruowana zaś na jej gruncie definicja zdania prawdziwego nosi nazwę teoriomodelowej definicji prawdy.

5. Definicja ta uważana jest za konstrukcję, która w sposób możliwie precyzyjny, formalnie poprawny i zarazem merytorycznie adekwatny ujmuje istotę klasycznego pojęcia prawdy. Osiąga to jednak za cenę poważnych ograniczeń i uproszczeń dotyczących rodzaju rozważanych języków, a będących konsekwencją założeń, na których się ta konstrukcja w sposób istotny opiera. Powstaje w związku z tym problem stosowalności tak zdefiniowanego pojęcia prawdy do rzeczywistych systemów językowych, w szczególności — do języków istniejących teorii naukowych. Odróżnić z tego punktu widzenia trzeba przede wszystkim nauki formalne (matematyczne) i nauki empiryczne.

Język teorii matematycznych nie odbiega w zasadzie — ani pod względem swych właściwości syntaktycznych, ani semantycznych — od standardowo sformalizowanych języków semantyki logicznej, bo te ostatnie właśnie na wzór języków matematycznych zostały skonstruowane. Nie ma więc żadnych trudności formalnych ze stosowaniem do języków teorii matematycznych teoriomodelowej definicji prawdy. Powstaje natomiast w stosunku do pewnych teorii matematycznych wątpliwość, czy pojęcie prawdy w ogóle znajduje wobec nich zastosowanie. Po to, aby móc mówić o zdaniach danego języka, że są po prostu prawdziwe czy fałszywe, trzeba aby był to język zinterpretowany, mówiący o określonej dziedzinie rzeczywistości. Otóż istnieje obszerna klasa abstrakcyjnych teorii matematycznych — teoria grup może być ich przykładem — które trudno uznać za teorie jakiejś specyficznej dziedziny rzeczywistości. Są to raczej teorie pewnego ogólnego pojęcia. Jedyną odpowiedzią na pytanie, co teoria grup ma opisywać, jest odpowiedź tautologiczna: każdą strukturę, która jest grupą: innymi słowy — każdą dziedzinę, którą teoria ta prawdziwie opisuje! Skoro więc język teorii grup nie odnosi się do jakiejś określonej dziedziny rzeczywistości, nie jest to język zinterpretowany i trudno jego zdaniom przypisywać po prostu prawdziwość czy fałszywość. Może to być jedynie prawdziwość czy fałszywość relatywizowana do którejś z jego możliwych interpretacji. Każdej takiej interpretacji odpowiada jakaś teoriomnogościowa struktura, stanowiąca jeden z możliwych modeli owego języka. Pojęciem prawdy, które znajduje zastosowanie do języków matematycznych tego typu, jest więc pojęcie prawdy relatywizowane do możliwego modelu danego języka: pojęcia zdania prawdziwego w modelu *M*. Główne wyniki teorii modeli dotyczą owego pojęcia relatywnego. Obok niego występuje pojęcia zdania prawdziwego w każdym modelu, w którym prawdziwe są aksjomaty danej teorii. W przypadku teorii elementarnych pojęcie to pokrywa się zakresowo z pojęciem twierdzenia danej teorii (tj. zdania wyprowadzalnego z jej aksjomatów). Pojęcie zdania prawdziwego *tout court* traci natomiast w przypadku teorii takiej, jak teoria grup, jakiekolwiek zastosowanie. Zdanie po prostu prawdziwe — to zdanie prawdziwe w tym modelu danego języka, który stanowi jego model właściwy, tj. model odpowiadający jego właściwej interpretacji. W tym zaś właśnie przypadku żadna taka interpretacja — a zatem i żaden taki model — wyróżnione nie zostały.

Wśród teorii matematycznych istnieją i takie, które mogą być traktowane jako teorie pewnych wyróżnionych dziedzin. Klasycznym przykładem takiej konkretnej teorii matematycznej jest arytmetyka liczb naturalnych. To, co teoria ta ma opisywać, określone jest w sposób niezależny od samej tej teorii, za pomocą pojęć ogólnej teorii mnogości. Tak wyróżniona struktura, zwana systemem liczb naturalnych, stanowi właściwy model języka arytmetyki i wyznacza właściwą jego interpretację. Pozwala tym samym na wprowadzenie dla zdań tego języka zwykłego, nierelatywnego, pojęcia prawdy. Zdanie tego języka jest po prostu prawdziwe, gdy jest prawdziwe przy tej właśnie interpretacji, w tym właśnie wyróżnionym spośród innych modelu.

Stosowalność pojęcia prawdy do języka danej teorii matematycznej umożliwia postawienie podstawowego problemu dotyczącego stosunku zbioru zdań prawdziwych teorii do zbioru jej twierdzeń. Teorie matematyczne przedstawione być mogą w postaci sformalizowanych systemów aksjomatycznych. W teoriach takich pojęcie twierdzenia ma charakter czysto formalny. Jest to zdanie dowodliwe, czyli wyprowadzalne z ustalonych aksjomatów wedle ustalonych reguł, przy czym zarówno owe aksjomaty, jak i reguły scharakteryzowane zostają w sposób odwołujący się wyłącznie do kształtu wyrażeń. Dowodliwość jest więc pewną uchwytną własnością zdań, która służyć może jako kryterium ich prawdziwości. Kryterium temu łatwo zapewnić cechę niezawodności: przedstawić daną teorię w postaci takiego sformalizowanego systemu aksjomatycznego, aby każde zdanie dowodliwe w tym systemie było prawdziwe. Słynne twierdzenie Gödla pokazuje natomiast, że w przypadku bogatszych teorii matematycznych kryterium to nigdy nie jest pełne. Jeśli teoria jakaś zawiera w sobie arytmetykę liczb naturalnych, nie może być ujęta w żaden sformalizowany system aksjomatyczny, którego twierdzeniami byłyby wszystkie zdania prawdziwe tej teorii: istnieją zawsze w jej języku zdania prawdziwe, lecz niedowodliwe w danym systemie. Dowodliwość jest więc cząstkowym jedynie kryterium prawdziwości, toteż nie może jej zastąpić wbrew temu, co sugerują niektórzy przeciwnicy klasycznej koncepcji prawdy, nawet w dziedzinie nauk matematycznych.

6. Sprawa stosowalności teoriomodelowej definicji prawdy w dziedzinie nauk empirycznych przedstawia swoiste problemy i trudności. W przeciwieństwie do języka matematyki język nauk empirycznych odbiega znacznie od standardowo sformalizowanych języków teoriomodelowej semantyki logicznej. Jest to język, który w dużym stopniu pokrywa się z językiem naturalnym, dzieląc z nim jego charakterystyczne, lecz kłopotliwe z semantycznego punktu widzenia własności. Pod względem syntaktycznym jest to język znacznie bogatszy od standardowego języka rachunku predykatów, do którego w zasadzie ogranicza się semantyka logiczna. Ważniejsze są jednak różnice natury semantycznej. Obecność w języku naturalnym wyrażeń wieloznacznych, zwłaszcza okazjonalnych, sprawia, iż zarówno kategoria syntaktyczna danego wyrażenia, jak i jego znaczenie zależą od całości sytuacji, w jakiej wyrażenie to zostało użyte: nie są więc czymś wyznaczonym przez sam kształt wyrażenia. Język taki, w konsekwencji, nie poddaje się procedurze formalizacji — i to jakiejkolwiek, nie tylko standardowej. Tymczasem semantyczna definicja prawdy taką właśnie formalizację zakłada. Język naturalny, ponadto, jest z natury rzeczy językiem uniwersalnym: można w nim mówić o wszystkim, a więc i o nim samym — wbrew postulowanemu rozróżnieniu języka przedmiotowego i metajęzyka. Język ten, w szczególności, obejmuje zdania mówiące o swej własnej prawdziwości i fałszywości, czyli pewne zdania antynomialne.

Wszystkie te właściwości języka naturalnego stały się przedmiotem wnikliwych badań semantycznych. Usiłuje się tak zmodyfikować i rozwinąć klasyczną semantykę logiczną, aby objąć nią języki naturalne: dąży się zwłaszcza do tego, aby zbudować semantyczną teorię prawdy dostosowaną do języków tego typu. Osiągnięto w tej dziedzinie — i to różnymi sposobami — wiele godnych uwagi wyników, choć problem daleki jest jeszcze od zadowalającego rozwiązania. Upraszcza się on jednak znacznie, jeśli dotyczyć ma nie całości języka naturalnego, lecz języka danej nauki empirycznej lub ściślej — danej teorii empirycznej. Język taki stanowi fragment tylko języka naturalnego, i to fragment stosunkowo prosty. Nie pretendując do uniwersalności, wolny jest od jakichkolwiek zdań antynomialnych. Choć nie jest na ogół wolny od innych swoistości języka naturalnego zarówno syntaktycznych jak i semantycznych — może zostać od nich w zasadzie przez odpowiednią rekonstrukcję uwolniony. Można wyrugować z niego np. wyrażenia wieloznaczne, zwłaszcza okazjonalne, wprowadzając zamiast jednego wyrażenia wieloznacznego szereg wyrażeń jednoznacznych (zastępując np. wyraz „dzisiaj" określoną datą itp.).

Język takiej teorii empirycznej jest zasadniczo przekładalny na język tego typu, jaki reprezentują standardowo sformalizowane języki semantyki logicznej. Przekład taki, choć zniekształca dany język pod pewnymi względami, zachowuje jego funkcje czysto poznawcze: nie zmienia, w szczególności, wartości logicznej jego zdań. Istnieje więc na ogół teoretyczna możliwość sformułowania dowolnej teorii empirycznej w języku, który dopuszcza standardową formalizację. Powstają jednak specyficzne trudności przy próbach zastosowania do takiego języka semantycznej definicji prawdy w jej współczesnej wersji teoriomodelowej. Trudności te związane są z pojęciem interpretacji języka empirycznego, z określeniem dziedziny, do której się taki język odnosi. Na gruncie semantyki logicznej dziedzina ta utożsamiana jest z teoriomnogościową strukturą złożoną z uniwersum języka i denotacji jego terminów: są to, w przypadku języka predykatów, określone zbiory (indywiduów lub ich układów). Otóż tego, do czego się odnosi język empiryczny, nie można utożsamić z żadną strukturą rozważanego typu. Struktura taka jest na to zbyt «ostra», zbyt dokładnie wyznaczona. Interpretacja języka empirycznego nie przyporządkowuje mu w sposób jednoznaczny tak pojętej dziedziny. Nieostrość wszelkich terminów empirycznych jest tego wyraźnym przejawem. Wbrew założeniom semantyki logicznej predykat empiryczny nie denotuje żadnego określonego zbioru, bo dla pewnych elementów uniwersum, należących do tzw. zakresu nieostrości danego predykatu, jego sens nie dostarcza żadnych kryteriów przynależności do takiego zbioru. Zbiór mający stanowić denotację predykatu empirycznego zostaje wyznaczony w sposób wieloznaczny. W sposób jednoznaczny wyznaczona jest jedynie pewna klasa zbiorów i co najwyżej ona może być przyporządkowana takiemu predykatowi jako jego interpretacja.

W przypadku terminów oznaczających wielkości fizyczne nieostrość przejawia się w aproksymacyjnym charakterze owych wielkości. Denotacją terminu takiego, jak „masa", nie może być (wbrew założeniom semantyki logicznej) żadna określona funkcja rzeczywista, gdyż wyznaczająca interpretację tego terminu procedura pomiarowa nie pozwala na przypisanie danemu przedmiotowi jako wartości tej wielkości określonej liczby rzeczywistej; pozwala jedynie na przypisanie dowolnej liczby z określonego przedziału. I tu więc jako interpretację danego terminu przyporządkować można mu co najwyżej pewną klasę funkcji rzeczywistych.

Wszystko to sprawia, że właściwa interpretacja języka empirycznego winna być utożsamiana raczej z pewną klasą standardowych struktur teoriomnogościowych, a nie z pojedynczą taką strukturą (jak to miało miejsce w ujęciu klasycznej semantyki logicznej). W ten sposób przezwyciężone może być przeciwieństwo zachodzące między «ostrą» ontologią teoriomnogościową metajęzyka a «rozmytą» rzeczywistością empiryczną opisywaną przez język przedmiotowy. Zastąpienie jednego modelu właściwego języka przedmiotowego klasą takich modeli rodzi z kolei pytanie o sens twierdzenia głoszącego, że zdanie tak zinterpretowanego języka jest po prostu prawdziwe. Wśród odpowiedzi dominują propozycje przypisujące wartość logiczną tylko tym zdaniom, których relatywna wartość logiczna jest ta sama we wszystkich modelach właściwych danego języka. Zdanie jest prawdziwe, gdy jest prawdziwe w każdym modelu właściwym; jest fałszywe, gdy jest fałszywe w każdym takim modelu. Zdania, które w pewnym modelach właściwych są prawdziwe, w innych zaś fałszywe, zostają tu uznane za zdania pozbawione wartości logicznej.

7. Taka definicja prawdziwości nakazuje uznać za wypowiedzi pozbawione wartości logicznej liczne zdania, traktowane skądinąd jako uzasadnione twierdzenia istniejących teorii empirycznych. Należą do nich — charakterystyczne dla teorii fizycznych — twierdzenia postulujące dokładne wartości liczbowe różnych wielkości fizycznych, a więc twierdzenia, które ze względu na podkreślany wyżej aproksymacyjny charakter interpretacji danego języka mogą być prawdziwe w niektórych tylko jego modelach właściwych. Celowe więc wydaje się wprowadzenie oprócz poprzedniego pojęcia prawdziwości, które nazwać można ścisłym, pojęcia aproksymacyjnego: zdanie danego języka jest aproksymacyjnie prawdziwe, gdy jest prawdziwe w pewnym modelu właściwym tego języka. Podobnie daną teorię nazwać można aproksymacyjnie prawdziwą, gdy wszystkie jej twierdzenia są prawdziwe w pewnym modelu właściwym jej języka. Istniejące teorie empiryczne, zwłaszcza ilościowe, okazują się z reguły tylko aproksymacyjnie prawdziwe. Dotyczy to w szczególności teorii opartych na założeniach idealizacyjnych, które z natury rzeczy mogą być prawdziwe jedynie «z pewnym przybliżeniem».

Pojęcie prawdy aproksymacyjnej stanowi jedną z eksplikacji pojęcia prawdy względnej (lub cząstkowej), przeciwstawianego w filozofii nauki pojęciu prawdy absolutnej. Nie jest to eksplikacja jedyna. Pojęcie prawdy względnej kryje również idee odmienne. Z jednej strony, prawdami względnymi bywają nazywane pewne twierdzenia fałszywe. Chodzi tu z reguły o takie twierdzenia ogólne, które traktowane jako prawa uniwersalne, ważne zawsze i wszędzie, okazują się zdaniami fałszywymi, sformułowane natomiast z odpowiednimi ograniczeniami czasowo-przestrzennymi stają się (ściśle lub aproksymacyjnie) prawdziwe. Z drugiej strony, prawdami względnymi nazywa się również zdania (ściśle) prawdziwe, zwłaszcza zbiory takich zdań, stanowiące teorie naukowe. Tym ostatnim przypisuje się prawdziwość tylko względną, ponieważ nigdy nie wyczerpują one ogółu zdań prawdziwych w danej dziedzinie rzeczywistości. Ogromna większość teorii naukowych to teorie niezupełne: w ich języku sformułować się dają zdania prawdziwe, lecz nie będące ich twierdzeniami. Ale teorie zupełne są również w pewnym sensie niewyczerpujące, gdyż obejmują tylko te zdania prawdziwe, które sformułowane są w danym języku za pomocą określonego aparatu pojęciowego. Dotyczą więc zawsze tylko pewnego aspektu opisywanej rzeczywistości.

W polskiej literaturze filozoficznej [6] istnieje pewna eksplikacja pojęcia prawdy względnej, która odwołuje się do intuicji związanych z pojęciem istotności. Ma ona zastosowanie wobec twierdzeń o charakterze praw naukowych, stwierdzających zależności funkcjonalne między różnymi wielkościami. Miano praw względnych przypisuje się tu prawom stwierdzającym trafnie zależności zachodzące między daną wielkością a niektórymi istotnymi dla niej czynnikami. Prawdy absolutne mają uwzględniać wszystkie czynniki istotne, fałsz zaś — nie uwzględniać żadnych. Sens tych określeń zależny jest oczywiście od sensu nadawanego pojęciu czynnika istotnego.

Różnice zachodzące między zdaniami prawdziwymi ze względu na sposób i stopień ich uzasadnienia pozwalają wyróżnić prawdy analityczne i prawdy syntetyczne. Prawdziwość pierwszych zapewnia sam sens terminów w nich występujących; prawdziwość drugich jest sprawą doświadczenia. Metody uzasadniania prawd syntetycznych, obejmujące pewne procedury bezpośrednie (obserwacja) i pośrednie (wnioskowanie), składają się na to, co nazywa się ogólnie praktyką naukową i co uznawane jest na ogół za kryterium prawdy w dziedzinie nauk empirycznych.

Praktyka naukowa stanowi część praktyki społecznej, rozumianej jako wszelka działalność, przekształcająca rzeczywistość przyrodniczą i społeczną. W myśl filozofii marksistowskiej naczelnym kryterium prawdy jest raczej całość praktyki społecznej. Stojąc na gruncie klasycznej koncepcji prawdy, kryterium tego nie można jednak utożsamiać z definicją prawdy: nie mówi ono, co prawdziwość znaczy, lecz tylko, jak ją można poznać. Nie jest to przy tym kryterium niezawodne. W stosunku do zdań syntetycznych nie daje ono nigdy całkowitej pewności ich prawdziwości, lecz co najwyżej określone jej prawdopodobieństwo. Zdania syntetyczne zachowują zatem zawsze charakter hipotez. Ów hipotetyczny charakter tych zdań ma się też niekiedy na myśli, gdy przypisuje się im miano prawd względnych.

8. Wywodząca się od Tarskiego semantyczna definicja prawdy nie jest — wbrew temu, co się niekiedy twierdzi — filozoficznie neutralna. Stosowana do ogółu teorii naukowych zakłada określone stanowisko filozoficzne w sprawie poznawczej wartości nauki. Stwierdzenie, że danemu zdaniu przysługuje pojmowana zgodnie z tą definicją prawdziwość, oznacza, że zdanie to odnosi się do pewnej dziedziny rzeczywistości i że w tej dziedzinie jest tak, jak to zdanie głosi. Zakładając, że pojęcie prawdziwości tak rozumianej stosuje się do każdego w zasadzie twierdzenia naukowego i że każde takie

twierdzenie jest bądź prawdziwe (ściśle lub aproksymacyjnie), bądź fałszywe, opowiadamy się w sporze o wartość poznawczą teorii naukowych za stanowiskiem realizmu, który teoriom naukowym taką wartość przyznaje, a przeciwko formalizmowi i instrumentalizmowi, które traktują teorie naukowe (pierwszy — matematyczne, drugi empiryczne) jako swoiste «narzędzia» pozbawione wartości poznawczej. Tym samym nie przesądzamy jednak jeszcze żadnego rozstrzygnięcia filozoficznego sporu między materializmem a idealizmem. W tej sprawie semantyczna definicja prawdy pozostaje istotnie neutralna. Poza ogólnymi własnościami formalno-ontologicznymi definicja ta nie zakłada niczego bliższego o dziedzinie, którą dane zdanie prawdziwe ma opisywać. Dziedzina ta może się równie dobrze składać z materialnych, niezależnie od nas istniejących rzeczy, jak i z układów treści świadomości podmiotu poznającego. I w jednym, i w drugim przypadku może być reprezentowana przez podobną teoriomnogościową strukturę. Teoria prawdy odpowiada na pytanie, na czym polega natura stosunku zachodzącego między zdaniem prawdziwym a rzeczywistością. Pytanie, na czym polega natura owej rzeczywistości, musi być rozstrzygnięte niezależnie.

Swoiste i inne niż w nauce problemy nastręcza pojęcie prawdy w takich dziedzinach kultury, jak sztuka, moralność, religia itp.

BIBLIOGRAFIA

[1] Ajdukiewicz K, *Zagadnienia i kierunki filozofii* (*Teoria poznania. Metafizyka*), Warszawa 1949.

[2] Grzegorczyk A., *Zarys logiki matematycznej*, Warszawa 1981.

[3] Kotarbiński T., *Elementy teorii poznania, logiki formalnej i metodologii nauk*, Wrocław 1961.

[4] Kotarbiński T., „W sprawie pojęcia prawdy", *Przegląd Filozoficzny*, 1934, 37.

[5] Krajewski W., „O pojęciach prawdy względnej", *Studia Filozoficzne*, 1963, 3-4.

[6] Nowak L., „Prawda cząstkowa — prawda względna — prawda absolutna", *Poznańskie Studia z Filozofii Nauki*, 1976, 1.

[7] Przełęcki M., „Pojęcie prawdy w językach nauk empirycznych", *Studia Filozoficzne*, 1977, 6.

[8] Schaff A., *Z zagadnień marksistowskiej teorii prawdy*, Warszawa 1959.

[9] Tarski A., *Pojęcie prawdy w językach nauk dedukcyjnych*, Warszawa 1933.

[10] Tarski A., „The Semantic Conception of Truth", *Philosophy and Phenomenological Research*, 1944, 4.

[11] Wójcicki R., „Semantyczne pojęcie prawdy w metodologii nauk empirycznych", *Studia Filozoficzne*, 1969, 3.