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) $$\text{`}\alpha\text{'} \in Ver \text{ gdy } \alpha.$$

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′) $$\text{`}P(a)\text{'} \in Ver \text{ gdy } P(a).$$

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) $\text{‘}\alpha\text{’} \in Ver(\mathfrak{M})$ 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′) $\text{‘}P(a)\text{’} \in Ver(\mathfrak{M})$ 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) $\text{‘}\alpha\text{’} \in Ver$ 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′) $\text{‘}P(a)\text{’} \in Ver$ 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) $\text{'}\alpha\text{'} \in \mathit{Ver}$ 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) $\text{'}\alpha\text{'} \in \mathit{Fls}$ 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') $\text{'}P(a)\text{'} \in \mathit{Ver}$ gdy $\mathbf{a}^* \in \mathbf{P}_1^* \cap ... \cap \mathbf{P}_n^*$;

(3c') $\text{'}P(a)\text{'} \in \mathit{Fls}$ 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 $\mathit{Ver} \cup \mathit{Fls}$ 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: $\mathbf{P}_1^* \cap \ldots \cap \mathbf{P}_n^*$ lub $\mathbf{P}_1^* \cap \ldots \cap \mathbf{P}_n^*$ 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.