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 $P_1, ..., P_n$. 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 $\Phi(P_1, ..., P_n)$ jest koniunkcją aksjomatów teorii T, a $\Psi(P_1, ..., P_n)$ — jednym z jej twierdzeń, to przy ujęciu formalnym zdaniom tym odpowiadają formuły ze zmiennymi wolnymi: $\Phi(X_1, ..., X_n)$ i $\Psi(X_1, ..., X_n)$. 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:

$$(X_1) ... (X_n) (\Phi(X_1, ..., X_n) \supset \Psi(X_1, ..., X_n));$$

bądź jako twierdzenia należące do metateorii o postaci:

Formuła „$\Psi(X_1, ..., X_n)$" wynika logicznie z formuły „$\Phi(X_1, ..., X_n)$".

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: $\mathfrak{M}' = \langle U', R_1', \ldots, R_n' \rangle$ i $\mathfrak{M}'' = \langle U'', R_1'', \ldots, R_n'' \rangle$. Dziedziny te są izomorficzne, gdy istnieje wzajemnie jednoznaczna relacja R przyporządkowująca elementom univer-

sum 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 $\Phi(P_1, ..., P_n)$ 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: $\mathfrak{M}' = \langle N^*, A', B', C', L', M', N' \rangle$ oraz $\mathfrak{M}'' = \langle N^*, A', B', C', L'', M'', N'' \rangle$, 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 $\mathfrak{M}'$ i $\mathfrak{M}''$ 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 $\Phi(P_1, ..., P_n)$, 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 $\mathfrak{A} = \langle U, R_1, ..., R_n, S_1, ..., S_m \rangle$ będzie modelem powyższej teorii T. Modelem częściowym odpowiadającym modelowi $\mathfrak{A}$ 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: $\mathfrak{M}'$ i $\mathfrak{M}''$. W modelach tych terminom elementarnym A, B, C przyporządkowane zostały jako denotacje klasy identyczne (zarówno w modelu $\mathfrak{M}'$, jak i w modelu $\mathfrak{M}''$ — klasy A', B', C'), a terminom teoretycznym L, M, N — różne (w modelu $\mathfrak{M}'$ — klasy L', M', N', w modelu $\mathfrak{M}''$ — 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 $\Phi(P_1, P_2, Q)$. 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.