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 M_0; jej elementy przez $\mathfrak{M}_0$; jej podklasy przez M_0. Dla każdego $\mathfrak{M} \in \mathsf{M}$, $\mathfrak{M}|_0$ 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 $M \subseteq \mathsf{M}$, $M|_0$ będzie oznaczać klasę wszystkich obcięć elementów M. (Oczywiście $\mathsf{M}|_0 = \mathsf{M}_0$.) 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 $\mathfrak{M}^*$. Stwierdzenie, że teoria *T* jest prawdziwa znaczy, że jest prawdziwa w modelu $\mathfrak{M}^*$, a to, z kolei, jest równoważne stwierdzeniu, że $\mathfrak{M}^*$ jest modelem aksjomatów *A*:

(1′a) $\mathfrak{M}^* \in M(A)$.

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 $\mathfrak{M}^*$ 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. $\mathfrak{M}|_0$ będzie teraz oznaczać obcięcie modelu $\mathfrak{M}$ 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 $\mathfrak{M}_0^*$, 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) $\quad \mathfrak{M}_0^* \in M(A)|_0$,

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) $\quad M^* = \{\mathfrak{M}: \mathfrak{M}|_0 \in M_0^* \text{ i } \mathfrak{M} \in M(A)\}$;

(C2) $\quad M_0^* \subseteq M^*|_0$

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) $\quad M_0^* \subseteq M(A)|_0$

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) $\quad M(A_0) = M(A)|_0$;

(2) $\quad M_0 \subseteq M(A_1)|_0$;

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:

$$A_0 = \{A_R\}$$
$$A_1 = \{A_R \rightarrow A\}.$$

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