

LEKKO NIEODPOWIEDZIALNE I STRONNICZE
**WPROWADZENIE
DO FILOZOFII ANALITYCZNEJ**
RAFAŁ URBANIAK



Wydawnictwo



Academicon



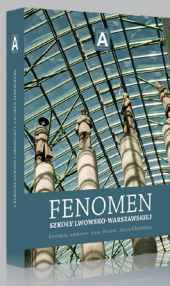
Wydawnictwo Academicon w wykazie wydawców MNiSW [100 pkt]!

Zapraszamy **AUTORÓW** monografii, prac doktorskich, habilitacyjnych i innych prac naukowych, popularnonaukowych i dydaktycznych do wydania książki w **nowoczesnym wydawnictwie**. Zapraszamy także do współpracy wydawniczej **REDAKTORÓW** czasopism, serii wydawniczych i prac zbiorowych.

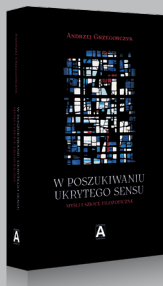
Publikuj z nami w open access!

[www: wydawnictwo.academicon.pl](http://www.wydawnictwo.academicon.pl), [e-mail: wydawnictwo@academicon.pl](mailto:wydawnictwo@academicon.pl), [tel.: 603 072 530](tel:603072530)

POLECANE



Red. nauk. Anna Brożek,
Alicja Chybińska
Fenomen
Sądków Lwowsko-
Warszawskiej



Andrzej Grzegorzczak
W poszukiwaniu
ukrytego sensu.
Myśli i szkice
filozoficzne



Natasha Szutta
Czy istnieje coś, co
zwiemy
moralnym charak-
terem i cnotą?



Artur Szutta
Moralne intuicje.
O poznaniu
dobra i zła

Formaty:



Lekko nieodpowiedzialne i stronnicze
wprowadzenie
do filozofii
analitycznej



Wydawnictwo Academicon

Rafał Urbaniak

Lekko nieodpowiedzialne i stronnicze
wprowadzenie
do filozofii
analitycznej

LUBLIN 2016

Korekta: Magdalena Korczyńska
Projekt typograficzny: Robert Kryński
Łamanie: Katarzyna Mikołajka
Projekt okładki: Olga Królak
Ilustracje: Olga Królak

© Copyright by Wydawnictwo Academicon, Lublin 2016

ISBN [ebook]: 978-83-62475-63-6

DOI: 10.52097/acapress.9788362475636

Wydawca: Wydawnictwo Academicon
ul. H. Modrzejewskiej 13, 20-810 Lublin
tel.: 603 072 530
e-mail: wydawnictwo@academicon.pl
www: <http://wydawnictwo.academicon.pl/>

Księgarnia online: ksiegarnia.academicon.pl

Analytic philosophy is who
you have lunch with

Gilbert Harman

Spis treści

1.	Uwagi techniczne	9
2.	Struktura tej książki	11
3.	Pojęcie filozofii analitycznej	12
4.	Podziękowania	14

CZĘŚĆ I WYBRANI DZIADKOWIE FILOZOFII ANALITYCZNEJ

5.	Bunt wczesnych filozofów analitycznych	17
6.	Moore <i>versus</i> McTaggart i Bradley	19
7.	Moore <i>versus</i> sceptycyzm	24
8.	Moore <i>versus</i> subiektywizm etyczny	25
9.	Moore <i>versus</i> naturalizm etyczny	26
10.	Problemy z antynaturalizmem Moore'a	30
11.	Dalsze problemy: analiza a informacja	32
12.	Dziedzictwo Moore'a	37
13.	Frege o wiedzy matematycznej	41
14.	Frege: sens i odniesienie	51
15.	Twierdzenie Cantora	58
16.	Paradoks Russella, teoria typów	62
17.	Russell i teoria deskrypcji	67
18.	Filozofia a składnia	78
19.	Carnap: pytania wewnętrzne i zewnętrzne	88

CZĘŚĆ II TROCHĘ BARDZIEJ MERYTORYCZNIE

20.	Mit zwrotu lingwistycznego	97
21.	Drogi i wyboje weryfikacjonizmu	100
22.	Zastosowanie: religia	117
23.	Zastosowanie: egoizm psychologiczny	124

CZĘŚĆ III PRAWDOPODOBIENSTWO

24.	Prawdopodobieństwo: podstawy	133
25.	Prawdopodobieństwo klasyczne	137
26.	Prawdopodobieństwo logiczne	139
27.	Częstotliwość i skłonność	145
28.	Subiektywizm i bayesianizm	147

CZĘŚĆ IV FILOZOFIA MATEMATYKI

29.	Teoria mnogości a arytmetyka	155
30.	Platonizm matematyczny	164
31.	Intuicjonizm matematyczny	167
32.	Formalizm matematyczny	169
33.	Neologicyzm	171
34.	Nominalizm	172
35.	Strukturalizm matematyczny	173

CZĘŚĆ V ODROBINA FILOZOFII JĘZYKA

36.	Problemy z deskryptywizmem	179
37.	Pojęcie operatora intensjonalnego	190
38.	Paradoksy implikacji a modalności	195

CZĘŚĆ VI ŚWIATY MOŻLIWE

39.	Światy możliwe: aparatura	203
40.	Światy możliwe: potrzeba zrozumienia	207
41.	Realizm modalny Davida Lewisa	208
42.	Problemy realizmu modalnego	212
43.	Jeśli nie realizm modalny, to co?	216
44.	Paradoksy maksymalności	221
45.	Zamiast zakończenia	235

1. Uwagi techniczne

Nie jestem historykiem filozofii, nie jestem historykiem filozofii analitycznej, a życie jest krótkie. Nie będę więc przejmował się szczegółami bibliograficznymi, szczegółową chronologią czy pierwszeństwem intelektualnym. Czytelnik spragniony takich informacji powinien skierować się gdzie indziej.

Niniejszy tekst nie jest monografią naukową. Raczej stanowić on będzie w miarę krótki opis (niektórych aspektów) tego, co uważam za kluczowe w rozwoju i naturze filozofii analitycznej. Opis ten będzie zawierał wiele uprzedzeń. Wybierać będę zagadnienia, które mnie interesują. Pomijać będę zagadnienia, o których wiem, ale mnie nie interesują. Pomijać też będę te, które mnie interesują, ale nie uważam ich za wystarczająco ogólne (lub ciekawe), żeby wzmiankować o nich we wprowadzeniu do filozofii analitycznej. Pomijać będę również zagadnienia, które są bardzo ważne, ale których w swojej niekompetencji nie dostrzegłem. Niestety, pomijać będę także te zagadnienia, o których chętnie bym napisał, gdybym miał na to czas, a książki nie trzeba było kiedyś skończyć.

Ocena zjawisk, o których będę mówił, również będzie czasami stronicza. Ze względu na zamierzoną krótkość tekstu niekiedy będę tylko wygłaszał pewne oceny lub uogólnienia, nie uzasadniając ich szczegółowo. Czasami będzie tak dlatego, że pogląd potrafię uzasadnić, ale uzasadnienie byłoby zbyt szczegółowe, żeby umieszczać je we wprowadzeniu do filozofii analitycznej. Czasami będzie tak dlatego, że twierdzenia te stanowić będą raczej wyraz moich intuicji niż rezultat dłuższych i szczegółowych przemyśleń.

Poza stylem prezentacji, doбором materiału oraz niektórymi przykładami nic w tej książce nie jest oryginalne, chyba że niechący. Pisząc, zaczytywałem się w różnych tekstach. O tych, które uznawałem za ciekawe, pisałem w sposób, który uznawałem za odpowiedni.

Teksty, które czytywałem w kółko, pisząc kolejne rozdziały, starałem się umieścić na liście tekstów polecanych. Jest jednak wiele argumentów, twierdzeń czy przykładów, które gdzieś kiedyś zasłyszałem, ale nie pamiętam gdzie, a zadanie przygotowania wykładów było ważniejsze niż przygotowanie przypisu z literaturą do każdego zdania. Pozwoliłem w takich wypadkach wygrać swojej sklerozie oraz niechłujności i z szacunku dla krótkości życia ludzkiego (zwłaszcza własnego) nie spędzałem zbyt dużo czasu na omawianiu tego typu szczegółów. Bardziej niż historia interesowały mnie same problemy i argumenty. To nie znaczy, że przypisałem sobie czyjkolwiek wynik lub argument. Jak rzekłem, nic nie jest tutaj oryginalne, chyba że niechcący. Ale i tekst ten nie jest tekstem badawczym, nie ma stanowić wkładu w naukę, a jedynie wprowadzenie do przedmiotu.

Niniejszy tekst jest szkicem kilku wykładów wygłoszonych na kursie dotyczącym wybranych zagadnień z filozofii analitycznej: bardziej niż wyjaśnienie wszystkich szczegółów ma on stanowić krótki wywód skłaniający do dalszych studiów. Jeżeli nie zadziała, jest to pewnie moja wina i z chęcią zapoznam się z uwagami (wolałbym osobiście, ale jeżeli nie ma takiej możliwości: rfl.urbaniak@gmail.com). Wyjątek stanowią uwagi, które nie są konstruktywne. Nieżyczliwy czytelnik, który takie uwagi ma, proszony jest o spisanie ich odręcznie własną krwią oraz spalenie tak przygotowanego materiału zamiast przysyłania go do mnie.

Krótką informacją o notacji. Cudzysłówów używam do przytaczania czyichś wypowiedzi, do tworzenia nazw wyrażeń lub ich treści lub do tworzenia nazw własności. Kursywy używam do tytułów prac, nazw wyrażeń znajdujących się w nazwach wyrażeń, cytatów w cytatach oraz do zapisu zwrotów obcojęzycznych. Tekstu wytłuszczonego używam do podkreślenia rzeczy, które uważam za szczególnie ważne, oraz do podkreślenia nowych terminów technicznych.

Struktura tej książki

2. Struktura tej książki

Zacznijmy od omówienia postaci jednego z prekursorów współczesnej filozofii analitycznej – George’a E. Moore’a. Mimo że kilka rozdziałów o nim będzie wyglądało historycznie, w istocie będziemy przyglądać się sposobowi, w jaki Moore formułował argumenty – będziemy te argumenty rekonstruować oraz oceniać.

Jednym z ważniejszych wątków w tej książce jest wiedza matematyczna, jej natura i przedmiot. Dlatego też po omówieniu też Moore’a przejdziemy do omówienia poglądów na matematykę, jakie żywił inny prekursor filozofii analitycznej – Gottlob Frege.

Po drodze wspomnimy też o innym wątku: Fregego teorii sensu i odniesienia, ponieważ wiąże się ona z innym istotnym w XX wieku zagadnieniem – filozofią nazw własnych, do której powrócimy w dalszej części książki.

Następnie zajmiemy się logicyzmem Bertranda Russella (poglądem, który miał stanowić ulepszenie teorii Fregego) oraz przeprowadzimy jego wstępną ocenę. Ponownie też poruszymy kwestię odniesienia i znaczenia, w takiej postaci, w jakiej jest ona obecna u Russella.

Później przejdziemy do zagadnienia sceptycyzmu wobec metafizyki i filozofii tradycyjnej w ogóle, obecnego od początku u wielu filozofów analitycznych. Skrajny sceptycyzm, jeśli chodzi o wartość poznawczą filozofii tradycyjnej, żywił na przykład Rudolf Carnap, i to na jego poglądach się skupimy. Doprowadzi nas to do omówienia różnych prób sformułowania kryterium sensowności wyrażeń. Z ich niepowodzenia można jednak wysnuć pewne istotne wnioski metodologiczne, których zastosowanie do filozofii religii i do egoizmu psychologicznego również zarysuję.

Istotną rolę we współczesnej filozofii analitycznej, a zwłaszcza filozofii formalnej (takiej filozofii, która posługuje się meto-

dami matematycznymi) odgrywa teoria prawdopodobieństwa. Dlatego też poświęcimy jej oraz jej interpretacji filozoficznej trochę miejsca.

Wreszcie wrócimy do filozofii matematyki: omówimy najpowszechniejsze podejście w podstawach matematyki, które posługuje się teorią mnogości, oraz dokonamy przeglądu sił i słabości kilku współczesnych filozoficznych teorii wiedzy matematycznej i przedmiotu matematyki.

Dokonawszy tego, wrócimy do kwestii związanych z sensem i odniesieniem i przyjrzymy się dominującemu obecnie pogładowi na temat odniesienia nazw własnych oraz temu, w jaki sposób wyłania się on z krytyki poglądów Fregego i Russella.

Wspólnym mianownikiem w niektórych aspektach filozofii matematyki oraz w teorii nazw własnych jest posługiwanie się pojęciami możliwości i konieczności. Ostatnia więc część tej książki stanowić będzie przegląd głównych strategii filozoficznego radzenia sobie z tymi pojęciami.

Każda grupa sekcji związanych z jednym tematem kończyć się będzie krótką **LISTĄ POLECANYCH LEKTUR**. Niektóre zagadnienia podsumowywane będą w krótkich „**KONSPEKTACH**”. Co jakiś czas w tekście znajdować się będą **KIERUNKOWSKAZY** przypominające czytelnikowi, na jakim etapie się znajdujemy.

3. Pojęcie filozofii analitycznej

W najszerszym sensie filozofem analitycznym jest każdy filozof, który:

- nie traktuje filozofowania jako wypowiedzania poetycko brzmiących wyznań o własnych stanach wewnętrznych,
- stara się przystępnie i zrozumiale wyjaśnić, co twierdzi,

Pojęcie filozofii analitycznej

- stara się argumentować jasno za tym, co twierdzi, i dlatego ważna jest dla niego znajomość matematycznych narzędzi pomagających kontrolować poprawność rozumowań – logiki,
- stara się jasno mówić, co jego argumenty zakładają i jak zamierzony wniosek ma z nich wynikać,
- uważa, że wyjaśnienie używanych terminów powinno odbywać się jasno i możliwie jednoznacznie w języku możliwie zbliżonym do języka potocznego, i jest to koniecznym wstępem do dalszych rozważań,
- unika posługiwania się niezdefiniowanymi uprzednio neologizmami.

Filozofią analityczną tak rozumianą zajmowali się już niektórzy filozofowie starożytni, a mistrzostwo w niej osiągnęli również pewni filozofowie średniowieczni.

W znaczeniu nieco węższym filozofowie analityczni wyłonili się na przełomie XIX i XX wieku i poza powyższymi aspektami charakteryzuje ich również zamiłowanie do logiki, niechęć do tworzenia wielkich systemów filozoficznych oraz zwracanie uwagi na język.

Nawet ujmując filozofię analityczną w tym węższym sensie, nie można jednoznacznie powiedzieć, że zaczęła się ona w jakimś konkretnym kraju albo w jakimś konkretnym czasie. Niektórzy przyjmują, że źródłem „zarazy” była Wielka Brytania, ale jest to obraz uproszczony. W zbliżonym czasie podobne zjawiska występowały w Polsce, Austrii i zapewne w wielu innych miejscach, o których nie mam pojęcia albo zapomniałem.

Jeżeli chodzi o logikę i filozofię, polska szkoła lwowsko-warszawska, która była jedną z najważniejszych na świecie szkół filozofii analitycznej w XX wieku, zasługuje na osobne omówienie; niestety nie będę miał okazji o niej powiedzieć za dużo w tej książce, głównie dlatego, że napisałem ją, próbując odpocząć od dość intensywnej pracy nad inną, jeszcze nudniejszą książką poświęconą jednej z głównych postaci tej szkoły (*Leśniewski's Systems of Logic and Foundations of Mathematics*,

2013). Polecam bardziej ogólną pracę o szkole lwowsko-warszawskiej autorstwa profesora Jana Woleńskiego, *Filozoficzna szkoła lwowsko-warszawska*.

4. Podziękowania

Winien jestem wdzięczność wszystkim, którym cierpliwość pozwoliła na zapoznanie się z wczesną wersją tej książki i którzy w swej życzliwości powiedzieli mi, co o niej myślą, w sposób nieskładający mnie do spalenia tych materiałów razem z laptopem.

W szczególności muszę wymienić te osoby, które należą do nielicznego grona studentów, którzy nie tylko czytają to, co jest obowiązkowe (a już ten warunek nie jest zbyt często spełniany), nie tylko czasem czytają to, co jest nadobowiązkowe, ale jeszcze, czytając, rozumieją to, co czytają, i myślą o tym samodzielnie. Wielu komentarzy dostarczyli Aleksandra Szulc, Zuzanna Gnatek, Tomasz Szutta, Karol Burka i Natalia Marcinowska.

Ponadto wdzięczny jestem moim bardziej kompetentnym kolegom, którzy znaleźli chwilę, by zmarnować ją na czytanie niniejszego materiału i dzielenie się ze mną uwagami o nim: Aleksandrze Samonek, Małgorzacie Stefaniak, Juliuszowi Doboszewskiemu, Michałowi Tomaszowi Godziszewskiemu, Maciejowi Sendłakowi i Jarosławowi Mrozkowi.

Jakość tego tekstu byłaby zdecydowanie gorsza, gdyby nie czujne oko i cierpliwość Kingi Beredy, która pomogła mi uniknąć wielu błędów i nie jest odpowiedzialna za żaden przeze mnie pozostawiony.

CZĘŚĆ



Wybrani dziadkowie filozofii analitycznej



KIERUNKOWSKAZ

Zaczniemy od opisanie buntu, który wcześni filozofowie analityczni podnieśli wobec współczesnego im stylu uprawiania filozofii, oraz od krytyki, jaką Moore sformułował pod adresem ówczesznie dominującego poglądu: idealizmu absolutnego. Następnie zarysujemy poglądy Moore'a na sceptycyzm oraz kwestie etyczne, przy okazji ukazując, jaki mają one związek z pozornie nieistotnymi zagadnieniami dotyczącymi języka, analityczności, teorii definicji itp. Zajmie nam to kilka następnych rozdziałów. ►

Bunt wczesnych filozofów analitycznych

5. Bunt wczesnych filozofów analitycznych

Wyobraźcie sobie wejście osoby, która nie miała nic wspólnego z filozofią, na wykład, powiedzmy, z klasycznej filozofii niemieckiej.

„[...] Również „tu” przeze mnie *wskazane*, którego się trzymam, jest takim *tym oto* „tu”, które w rzeczywistości *nie jest tym oto* „tu”, lecz jakimś „na przedzie” lub „z tyłu”, jakimś „w górze” lub „na dole”, jakimś „z prawej strony” lub „z lewej strony”. Z kolei to, co „w górze”, jest znowu wielością takich innobyków, jak „w górze”, „na dole” itp. „Tu”, które mieliśmy wskazać, zanika w innych „tu”, które z kolei również zanikają [...]. (Hegel)¹

„[...] jeżeli bowiem z empirycznych danych naocznych ciał i ich zmian (ruchu) usuniemy wszelki pierwiastek empiryczny, mianowicie to, co należy do wrażenia (*Empfindung*)², to pozostanie jeszcze przestrzeń i czas. Są one przeto czystymi danymi naocznymi, które stanowią *a priori* podstawę oglądów empirycznych i dlatego same nigdy nie mogą być usunięte. Przez to jednak właśnie, że są one czystymi danymi naocznymi, wykazują, że są tylko formami naszej zmysłowości, które muszą wyprzedzać wszelkie oglądanie empiryczne, tj. spostrzeżenie rzeczywistych przedmiotów, i zgodnie z którymi można przedmioty poznawać *a priori*, lecz co prawda tylko tak, jak się nam one przejawiają. (Kant)

- 1 Podkreślenia pochodzą od Hegla. Jest oczywiste, że pisanie wyrazu kursywą nadaje mu głęboki sens filozoficzny. Niestety nie znalazłem żadnego systematycznego studium filozoficznej doniosłości kursywy u Hegla, które wyjaśniałoby to zjawisko.
- 2 Dodawanie groźnie brzmiących niemieckich wyrazów w nawiasach jest wysoce skuteczną techniką, którą posługują się czasem tłumacze, by uczynić tekst zrozumiałym, a lekturę – łatwą (nie wspominając o skutkach ubocznych, jakimi są podkreślenie doniosłości intelektualnej tekstu i umożliwienie czytelnikom rzucania terminami oryginalnymi w późniejszych konwersacjach pod pretekstem dążenia do uniknięcia nieporozumień).

Można mieć wrażenie, że skomplikowany język jest tutaj przejawem zdecydowanej wyższości intelektualnej. Atmosfera tajemniczości i niezrozumiałości jest nieraz uznawana za wyraz doniosłości i głębi. Można też jednak twierdzić, że jest to tylko zabawa słowami (jak powiedział Jung: przechadzanie się w zbudowanej przez siebie celi z własnych słów). Ładnie opakowane pudełko nie musi zawierać cennego prezentu. Właśnie ta druga, mniej pretensjonalna postawa bliższa jest filozofii analitycznej (choć przeprowadzone osobiście badania wskazują, że rzucając wiadrem na konferencji filozofów analitycznych mamy ok. 5% szans na trafienie w pretensjonalnego filozofa analitycznego; 5% to jednak i tak znacznie mniejszy wskaźnik niż w przypadku całej populacji filozofów w ogóle)³.

W sytuacji podobnej do tej, w której wchodzimy do takiej sali wykładowej, znajdowali się niektórzy filozofowie u schyłku XIX wieku. W Europie filozofia była zdominowana przez dyskurs przepojony patetycznym i hermetycznym językiem, którego przykłady widzimy powyżej⁴.

Co więcej, z jakiegoś powodu niektórzy z tych wyrafinowanych filozofów wierzyli, że udało im się skonstruować jakiś całościowy i rozbudowany system wiedzy o świecie. Niestety, zazwyczaj trudno było w ich tekstach znaleźć przekonujące argumenty na rzecz prawdziwości tych przekonań.

Filozofia analityczna, przynajmniej we wczesnym okresie rozwoju swojej postaci współczesnej (i bardziej w Anglii niż na kontynencie), była buntem przeciwko takiej pewności siebie, przeciwko niechlujności argumentów i niepotrzebnej kom-

3 Na wypadek gdyby jaki(a)ś miłośnik(czka) filozofii niemieckiej chciał(a) się obrazić, dodam, że pomimo lekkiej ironii w przypisach powyżej, zgadzam się, że Kant i Hegel wielkimi filozofami byli. Ale to, co zrobili, było interesujące *pomimo* ich stylu, a nie *dzięki* niemu.

4 Na przykład Schopenhauer obserwuje: „Aby ukryć brak prawdziwych pomysłów, wielu konstruuje dla siebie imponującą aparaturę długich i złożonych wyrazów, skomplikowane kwieciste mowy i frazy, nowe i nienapotykanne wcześniej wyrażenia. Wszystko to razem wzięte stanowi skrajnie trudny żargon, który brzmi bardzo uczenie. Ale za pomocą tego wszystkiego nie przekazują żadnej treści”.

Moore *versus* McTaggart i Bradley

plikacji wyvodu filozoficznego. Jednym z takich wczesnych buntowników był GEORGE EDWARD MOORE (1873–1958) i od niego zaczniemy⁵. Przyjrzymy się najważniejszym argumentom, jakie sformułował, a przy okazji zapoznamy się trochę z jego stylem filozofowania. Zaczniemy od jego krytyki idealizmu, potem przyjrzymy się jego polemice ze sceptycyzmem, subiektywizmem i naturalizmem etycznym. Wskażemy słabości oraz zalety podejścia Moore’a.

6. Moore *versus* McTaggart i Bradley

W czasach wczesnych prac Moore’a filozofami dość popularnymi na Wyspach byli McTaggart i Bradley.

JOHN M.E. MCTAGGART (1866–1925) był idealistą metafizycznym (uważał, że natura rzeczywistości jest mentalna – to co się dzieje, nie jest niezależne od umysłu i podobne jest w swojej naturze do treści naszych snów) inspirowanym myślą Hegla⁶ i zasłynął głównie z głoszenia, że nie istnieje czas (tekst, w którym argumentował za nieistnieniem czasu, opublikowano w roku 1908)⁷.

5 Bertrand Russell (w *My philosophical development*) o swoim i G.E. Moore’a buncie mówi tak: "It was towards the end of 1898 that Moore and I rebelled against both Kant and Hegel. Moore led the way, but I followed closely in his footsteps... I felt a great liberation, as if I had escaped from a hot house into a wind swept headland... In the first exuberance of liberation, I became a naïve realist and rejoiced in the thought that grass really is green". Można to luźno przetłumaczyć jako: „Pod koniec 1898 r. Moore i ja zbuntowaliśmy się przeciwko Kantowi i Heglowi. Moore prowadził, ale byłem tuż za nim... poczułem wielkie wyzwolenie, tak jakbym uciekł z gorącego domu na wiatrami targany przylądek... w pierwszym zachłyśnięciu się wolnością zostałem naiwnym realistą i cieszyłem się myślą, że trawa naprawdę jest zielona”.

6 Napisał o niej dwie książki: *Studies in the Hegelian Dialectic* (1896) oraz *Studies in Hegelian Cosmology* (1901).

7 Więcej na ten temat poczytać można na przykład w stanfordzkiej Encyklopedii Filozofii: <http://plato.stanford.edu/entries/time/#McTArg>.

FRANCIS H. BRADLEY (1846–1924), chociaż tak jak McTaggart uznawał, że natura rzeczywistości jest czysto mentalna, kładł podejrzenie silny akcent na w sumie mało kogo obchodzącą kwestię dotyczącą tego, czy relacje są zewnętrzne, czy wewnętrzne⁸. Głosił, że wszystkie relacje są konieczne i wewnętrzne – nie istnieją relacje pomiędzy różnymi przedmiotami.

Obaj głosili więc tezy nieco zaskakujące, nawet jeżeli niekiedy zrozumiałe. Główną jednak tezę, co do której się zgadzali, było twierdzenie, że istnieje tylko jeden rzeczywisty przedmiot (ten pogląd krótko nazywa się **MONIZMEM**) i że nie istnieją przedmioty niezależne od umysłu (ten pogląd nazywa się **IDEALIZMEM**). Jeden z pierwszych tekstów Moore’a (*The Refutation of Idealism*, 1903) był właśnie atakiem na idealizm tych panów.

Atak Moore’a, dokładniej rzecz biorąc, nie był atakiem na główne twierdzenie idealizmu, wedle którego natura świata jest mentalna. Jego rozważania zupełnie nie rozstrzygały tej kwestii. Zamiast tego skoncentrował się na tym, co uznał za najważniejsze twierdzenie pomocnicze idealizmu: **być to być postrzeganym**. Żeby twierdzenie to brzmiało donioślej i wiarygodniej, czasem używa się wyrazów łacińskich: *esse* (bycie) to *percipi* (bycie postrzeganym) – w skrócie będziemy ten pogląd nazywać EP. Twierdzenie to odgrywa istotną rolę w przypadku wielu argumentów idealistycznych: jeżeli zgodzimy się, że bycie wymaga bycia postrzeganym, to łatwiej będzie nas prze-

8 Do dziś nie jest jasne, o co do końca chodziło. Wedle interpretacji Russella relacje wewnętrzne to te, które redukują się do własności przedmiotów, które pozostają w tej relacji, tak jak na przykład relacja „bycia taką samą figurą geometryczną” zachodzi ze względu na własności obiektów wziętych osobno, a relacja kopnięcia zachodząca między, powiedzmy, chomikiem a żabą, nie redukuje się do własności, które można zaobserwować, patrząc na chomika (czy na żabę) osobno. Wedle interpretacji Moore’a relacja jest wewnętrzna, jeżeli przedmioty w niej pozostające są w niej z konieczności, i tak relacja bycia większą liczbą zachodząca pomiędzy liczbą trzy i liczbą dwa jest relacją wewnętrzną, a relacja bycia liczbą, o której właśnie myślę, jest tylko przypadkową własnością liczby 42, i w związku z tym jest relacją zewnętrzną. Badacze Bradleya twierdzą czasem, że obie interpretacje są nietrafne, ale niestety nie udało mi się zrozumieć ich interpretacji. Na szczęście nie będziemy się zajmować tym rozróżnieniem w niniejszej książce.

konać o tym, że świat jest obiektem natury mentalnej, jak treści naszych snów⁹. Przyjrzymy się więc tezie EP.

Po pierwsze, „postrzeganie” w EP jest rozumiane szeroko: chodzi nie tylko o postrzeganie zmysłowe, lecz także o bycie przedmiotem innych działań umysłowych. Dzięki temu idealista może też mówić o istnieniu np. obiektów matematycznych, które nie mogą, dosłownie rzecz biorąc, być postrzegane zmysłowo.

Po drugie, Moore odrzuca pogląd, jakoby EP było definicją, a *esse* po prostu znaczyło dokładnie to samo, co *percipi*. Tak naprawdę robi to bez podania żadnego argumentu¹⁰. Można się jednak domyślać, że gdyby rzeczywiście EP było definicją, byłoby definicją nietrafną chociażby z tego prostego powodu, że istnieją wrażenia obiektów nieistniejących.

Doprecyzowana więc interpretacja EP głosi, że bycie postrzeganym jest koniecznym, ale niewystarczającym warunkiem istnienia, to znaczy – cokolwiek istnieje, musi być postrzegane, ale nie wszystko, co jest postrzegane, istnieje. Ta interpretacja unika problemu z przedmiotami złudzeń, które są w pewnym sensie postrzegane, ale nie istnieją.

Ponadto Moore uważa, że EP, jeżeli jest prawdziwe, musi być prawdziwe koniecznie. W przeciwnym wypadku nie mogłoby odgrywać tak ważnej roli w idealizmie¹¹. EP należy więc czy-

9 Niektórzy posługują się pełnym zwrotem łacińskim: *esse est percipi*. Być może w tej wersji zdanie to brzmi jeszcze dostojniej. Niemniej jednak ani Berkeley (od którego Moore ten zwrot zaczerpnął), ani Moore tego nie robili – u obu spójnik „jest” występował w języku angielskim.

10 „Myślę, że nie muszę dowodzić, że zasada *esse est percipi* nie jest więc zamierzona jedynie jako definicja słowa; ani też, że, gdyby była, byłaby skrajnie złą definicją”. („I think I need not prove that the principle *esse est percipi* is not thus intended merely to define a word; nor yet that, if it were, it would be an extremely bad definition”).

11 „[...] jeżeli twierdzenie, że *percipi* stanowi część całości rozumianej jako rzeczywistość, ma mieć jakąkolwiek doniosłość, musi ono znaczyć, że całość jest organiczna, przynajmniej w tym sensie, że inny składnik albo inne składniki nie mogą występować bez *percipi*, nawet jeżeli *percipi* może występować bez nich”. („[...] if the assertion that *percipi* forms part of the whole meant by reality is to have any importance, it must mean that the whole is organic, at least in this sense, that the other constituent or constituents of it cannot occur without *percipi*, even if *percipi* can occur without them”).

tać: „Jest konieczne, że cokolwiek istnieje, jest postrzegane”¹². Takie rozumienie ma co najmniej dwie dalsze interpretacje.

Pierwsza głosi, że to, iż *percipi* jest warunkiem koniecznym dla *esse*, zachodzi na mocy znaczenia samych wyrazów. Zdania prawdziwe na mocy znaczenia, takie jak „wszyscy kawalerowie są nieżonaci”, nazywa się zdaniami **ANALITYCZNIE PRAWDZIWYMI**. Pojęcie analityczności jest nieco problematyczne filozoficznie, ale na razie się tym nie przejmujemy. Opcję pierwszą Moore odrzuca, ponieważ gdyby EP było analityczne, to (wbrew idealistom) byłoby trywialne, nieinformatywne, a zatem i filozoficznie niedoniosłe.

Druga opcja głosi, że o ile związek pomiędzy *esse* i *percipi* jest konieczny, to nie jest konieczny na mocy znaczenia. To pozwala uniknąć zarzutu nieinformatywności. EP zdaniem Moore’a, jeżeli jest prawdziwe, powinno być konieczne, ale nie na mocy znaczenia. Zdania takie noszą miano zdań **KONIECZNYCH SYNTETYCZNYCH**¹³.

Zgódźmy się, że EP powinno być konieczne i syntetyczne. Kolejne ważne rozróżnienie sądów to rozróżnienie na sądy poznawalne bez pomocy żadnego specyficznego doświadczenia zmysłowego (*a priori*) oraz sądy, których poznanie (lub rozstrzygnięcie) wymaga jakichś specyficznych doświadczeń zmysłowych (*a posteriori*)¹⁴. Pojawia się więc pytanie: do której grupy wedle idealizmu zalicza się EP?

Moore uważa, że EP nie jest wedle idealizmu zdaniem *a posteriori*, ale nie mówi zbyt jasno, dlaczego. Podejrzewam,

¹² Uważa też, że ten związek konieczny można utożsamić z wyprowadzalnością *percipi* z *esse*. Z tego powodu nieraz używa raczej sformułowania: „Z tego, że coś istnieje, można poprawnie wywnioskować, że jest postrzegane”.

¹³ Zdania, których prawdziwości nie da się rozstrzygnąć po prostu na mocy ich znaczenia, nazywają się **SYNTETYCZNYMI** i przeciwstawiane są zdaniom **ANALITYCZNYM**.

¹⁴ „Specyficznego” znajduje się w tym określeniu nie bez powodu. Typowym zdaniem *a priori* jest zdanie „ $2 + 2 = 4$ ”. Poznanie jego prawdziwości nie zależy od żadnego specyficznego doświadczenia: nie ma żadnego określonego fragmentu rzeczywistości fizycznej, który trzeba by zaobserwować, żeby dowiedzieć się, że $2 + 2 = 4$. Nie znaczy to jednak, że można dowiedzieć się o prawdziwości tego zdania, nie posiadając żadnych w ogóle doświadczeń. Pewne doświadczenia będą niezbędne, żeby w ogóle nauczyć się języka itd.

że robi tak z dwóch powodów. Po pierwsze, zdania, których poznanie wymaga jakiegoś specyficznego doświadczenia, nie są według niego konieczne: doświadczenie dostarcza informacji tylko o stanach rzeczy w tym sensie przygodnych, czyli takich, które nie są ani konieczne, ani nie są koniecznie nierealizowalne. Po drugie, idealiści zasadniczo uważają, że ogólne prawdy filozoficzne nie są poznawalne za pomocą doświadczenia, a jedynie za pomocą rozumu – współgra to z ich poglądem, że świat w pewnym sensie jest obiektem o naturze mentalnej.

Założmy więc, że EP jest zdaniem koniecznym, syntetycznym i poznawalnym *a priori*. W jaki sposób jest ono poznawalne *a priori*? Jeden ze sposobów to bycie oczywistym. Moore uważa jednak, że nie jest to sposób, w jaki EP jest poznawalne *a priori* – chociażby dlatego, że uważa EP za zdanie fałszywe, a więc nienarzucające się z oczywistością¹⁵.

Jedynym pozostałym sposobem, na jaki EP może być prawdziwe *a priori*, jest zdaniem Moore'a wyprowadzalność z samej zasady sprzeczności¹⁶. To od razu stwarza poważną trudność: zdania wyprowadzalne z zasady sprzeczności powinny być analityczne, bo zasada sprzeczności jest prawdziwa na mocy tego, co przeczenie znaczy. W takim jednak razie idealizm zdaniem Moore'a próbuje głosić coś informatywnego (a zatem i syntetycznego), a zarazem musi uznać, że jest to coś analitycznego (bo wyprowadzalne z zasady sprzeczności). Żadne jednak zdanie nie może być zarazem analityczne i syntetyczne.

Argument ten jest problematyczny. Swoją strukturą przypomina inny argument, z którego sformułowania Moore jest sławny, a który omówimy niedługo. Jego ocenę odkładam więc na później.

15 „Jeżeli idealista wybiera twierdzenie, że jest to jedynie prawda oczywista, pozostaje mi tylko powiedzieć, że mi się taką nie wydaje. Sądzę jednak, że żaden idealista nigdy tak nie twierdził”. (“If the Idealist chooses to assert that it is merely a self-evident truth, I have only to say that it does not appear to me to be so. But I believe that no Idealist ever has maintained it to be so”).

16 To znaczy: z zasady, wedle której żadne zdanie nie może być prawdziwe jednocześnie ze swoim zaprzeczeniem.

7. Moore *versus* sceptycyzm

Idealizm dzielił ze sceptycyzmem (sceptycyzm w jakiejś kwestii to pogląd, że w tej kwestii nie można uzyskać wiedzy; sceptycyzm ogólny to pogląd, że w ogóle nie można uzyskać żadnej wiedzy) swoją wątpliwość co do istnienia przedmiotów niezależnych od doświadczenia (czyli takich, które mogą istnieć, gdy nie istnieje podmiot doświadczający ich). Część filozofów preferuje postawę sceptyczną: nie przyjmują przekonań zdroworozsądkowych na wiarę, żądają ich dowodu, a gdy takiego dowodu brak, uznają, że tak naprawdę nie stanowią one naszej wiedzy.

Moore do takich filozofów nie należał i jego opozycja wobec idealizmu i sceptycyzmu polegała (częściowo) na obróceniu tego wyzwania: założenia sceptycyzmu i idealizmu są na tyle skomplikowane i nieoczywiste, że uznanie ich i kwestionowanie zdrowego rozsądku wymaga silniejszej obrony niż uznanie samych przekonań zdroworozsądkowych. Uroczym przykładem posłużył się w *Some Main Problems of Philosophy* (1953)¹⁷: jeżeli jakaś teoria filozoficzna prowadzi do wniosku, że w czasie, kiedy siedzę w jadącym pociągu, nie mogę wiedzieć, że ten pociąg ma koła, to należy zastanowić się, co jest bardziej nieintuicyjne – przekonanie, że nie wiem, że pociąg ma koła, czy założenia tej teorii filozoficznej.

Taką samą strategię widać w *Proof of an External World* (1939). W odpowiedzi na żądanie dowodu istnienia przedmiotów w świecie Moore powiada: tu jest jedna ręka, tu jest druga. Zatem istnieją przynajmniej dwie ręce, które są przedmiotami istniejącymi w przestrzeni. Zatem istnieją co najmniej dwa przedmioty, które są zewnętrzne względem umysłu.

¹⁷ Książka ta była oparta na wykładach, które wygłosił w Londynie w latach 1910–1911.

Moore *versus* subiektywizm etyczny

Argument ten, zdaniem Moore'a, spełnia wszystkie wymogi poprawnego argumentu: przesłanki różnią się od wniosku, wniosek wynika z przesłanek oraz prawdziwość przesłanek jest znana.

Argument jest całkowicie analogiczny do argumentu za istnieniem trzech błędów w druku, polegającego na pokazaniu ich. Jest to argument całkowicie akceptowalny w życiu codziennym. Jeżeli jakiś filozof go odrzuca (na przykład mówiąc, że przesłanki wcale nie są znane), to znaczy, że posługuje się dużo bardziej wymagającym i nietypowym pojęciem wiedzy, odbiegającym od normalnych standardów przyjmowanych przez ludzi na co dzień. Za przyjęciem takiego bardziej wymagającego pojęcia wiedzy stoją zazwyczaj wyrafinowane teorie filozoficzne, które są bardziej wątpliwe niż twierdzenia o istnieniu rąk.

8. Moore *versus* subiektywizm etyczny

Jeden z głównych wariantów subiektywizmu etycznego (zwany ekspresywizmem) głosi, że gdy ktoś twierdzi, że coś jest dobre albo że coś jest złe, to nie stwierdza żadnego obiektywnego faktu, a jedynie mówi coś o swojej własnej postawie względem tego. Na przykład gdy mówię „Gryzienie żywych królików po nogach jest złe”, to wedle tego poglądu wyrażam po prostu swoją niechęć do gryzienia żywych królików po nogach i nie mówię wiele więcej niż „Gryzienie żywych królików po nogach? Fuuuuj!”.

Moore zauważył, że pogląd ten stoi w sprzeczności z obserwacją, że ludzie czasem nie zgadzają się w kwestiach etycznych. Jeżeli bowiem mówiąc „Gryzienie żywych królików po nogach jest złe”, mówię tylko, że mi się gryzienie żywych królików po nogach nie podoba, to nie mówię nic sprzecznego z kimś

innym, kto głosi, że gryzienie żywych królików po nogach jest dobre, i ma jedynie na myśli swoją własną pozytywną postawę emocjonalną względem ww. działalności.

Nie stanowi to ostatecznie powalającej krytyki subiektywizmu, ale stanowi dla niego pewne wyzwanie (które też podejmowali subiektywiści od czasów Moore'a). Subiektywista musi albo zaprzeczyć, że istnieją przypadki niezgody w etyce, i wyjaśnić, skąd bierze się nasze mylne przekonanie o ich istnieniu, albo przeformułować swoją subiektywistyczną analizę zdań wartościujących tak, aby dopuszczała możliwość niezgody mimo subiektywizmu. To, jak takie próby wyglądały, stanowi jednak opowieść na inną okazję¹⁸.

9. Moore *versus* naturalizm etyczny

Wiele teorii etycznych stara się zdefiniować dobro za pomocą sformułowania typu¹⁹:

[DEF] Przedmiot/zdarzenie/czyn x jest dobry wtedy i tylko wtedy, gdy x spełnia warunek φ ,

gdzie rozmaite teorie proponują różne sformułowania warunku φ . Teorie tego typu starają się zanalizować pojęcie dobra za pomocą pewnych prostszych pojęć.

¹⁸ Warto na przykład zacząć od rzucenia okiem na *The Emotive Meaning of Ethical Terms* C.L. Stevensona w *Mind* 46 (1937).

¹⁹ Ścisłe rzecz biorąc, Moore uważał, że czyn jest dobry lub zły ze względu na swoje konsekwencje i to one podlegać mają pierwotnej ocenie moralnej. Pogląd taki nazywany jest **KONSEKWENCJALIZMEM**.

Moore *versus* naturalizm etyczny

Na przykład **UTYLITARYZM** (w jednym z podstawowych wariantów) głosi:

Działanie *x* jest dobre wtedy i tylko wtedy, gdy *x* wytworzy większą ilość szczęścia niż jakikolwiek inny czyn, który dana osoba może wykonać zamiast *x*.

Moore w jednej z najważniejszych swoich prac, *Principia Ethica* (1903), podjął się krytyki tego typu prób zdefiniowania dobra za pomocą własności, które nazywał „naturalnymi”²⁰.

Uważał, że wszelkie próby podania takiej definicji skazane są na niepowodzenie, bo dobro jest niedefiniowalne.

Niedefiniowalność dobra miała wedle Moore’a dosyć poważną konsekwencję: żadne poprawne rozumowanie, które nie opiera się na przesłance stwierdzającej już, że coś jest dobre, nie może prowadzić do wniosku, że coś jest dobre. Innymi słowy, wychodząc od przesłanek, w które nie wbudowaliśmy już założeń o dobru, nie rozstrzygniemy, czy coś jest dobre. Na przykład z tego, że coś przyczynia się do największej możliwej ilości szczęścia, nie wynika zdaniem Moore’a, że jest to dobre.

Dlaczego właściwie niedefiniowalność miałaby mieć taką konsekwencję? Moore nie mówi tego wprost, ale podejrzewam, że intuicja jest tutaj następująca: aby móc ustalić za pomocą argumentu, że coś jest dobre, nie wbudowując już założeń o tym, co jest dobre, musimy wiedzieć, czego szukać „w świecie”. Musimy wiedzieć, jakie obserwacje lub jakie fakty potwierdzałyby twierdzenie, że coś jest dobre. Jeżeli jednak

²⁰ Co to znaczy, że własność jest naturalna? Moore uważał, że własności naturalne wyróżniają się tym, że mogą istnieć w czasie niezależnie od przedmiotów naturalnych (*Principia Ethica*, § 26). Pogląd ten opiera się nie tylko na niejasnym pojęciu istnienia samodzielnego własności w czasie, lecz także na nieco dziwnej metafizyce. Dlatego też dla naszych potrzeb możemy zapomnieć o tym sposobie definiowania naturalności i w tym kontekście uznać, że pojęcie to jest intuicyjnie zrozumiałe i że własność jest naturalna, jeżeli doświadczenie zmysłowe może dostarczyć w miarę bezpośrednich powodów, by uznać, że ona danemu przedmiotowi przysługuje lub nie przysługuje.

dobro jest niedefiniowalne, to nie możemy znać jego definicji za pomocą własności naturalnych, a zatem nie możemy wiedzieć, jakiego typu obserwacji dokonywać lub jakich faktów obserwowalnych szukać.

Wróćmy jednak na razie do definiowalności. Moore uważał, że w przypadku gdy dany predykat A jest definiowalny, definiowalny jest za pomocą wyrażenia typu:

Przedmiot x jest A wtedy i tylko wtedy, gdy x spełnia warunek φ ,

gdzie predykat A wyraża pewną własność złożoną, która jest tą samą własnością, co ta wyrażona za pomocą φ . Na przykład gdy zdefiniujemy predykat „jest wykładem” następująco:

x jest wykładem wtedy i tylko wtedy, gdy x jest procesem przekazywania informacji z ust wykładowcy do notatek studentów bez pośrednictwa aktów poznawczych ze strony studentów,

to predykat ten będzie wyrażał pewną złożoną własność: bycie procesem przekazywania informacji itd.

Moore uważał, że dobro nie jest definiowalne w tym sensie: nie jest własnością złożoną, którą zanalizować można za pomocą definicji. Jego zdaniem dobro jest własnością prostą i nieanalizowalną, której obecność lub brak oceniamy bezpośrednio i bez pomocy argumentów (mniej więcej tak, jak jego zdaniem bezpośrednio i bez pomocy argumentów wiemy, że mamy ręce).

Nie znaczy to, że uważał on, iż żadne zdanie o postaci [DEF] nie jest prawdziwe, a raczej, że nawet jeżeli jest prawdziwe, to nie jest prawdziwe jedynie na mocy znaczenia (analityczne) i nie jest poznawalne *a priori*.

Argument Moore’a za niedefiniowalnością dobra jest na tyle sławny, że dorobił się własnej nazwy: **ARGUMENT PYTANIA OTWARTEGO** (znaleźć go można w sekcji 13 *Principia Ethica*). Przebiega on następująco. Rozważmy dwa pytania:

Moore versus naturalizm etyczny

PYTANIE 1 Załóżmy, że x jest φ – czy x jest dobre?

PYTANIE 2 Załóżmy, że x jest φ – czy x jest φ ?

Przeprowadźmy następujące rozumowanie:

PRZESŁANKA 1 Jeżeli [DEF] jest poprawną definicją, to „dobry” i „ φ ” znaczą to samo.

PRZESŁANKA 2 Jeżeli „dobry” i „ φ ” znaczą to samo, to PYTANIE 1 wyraża to samo trywialne pytanie, co PYTANIE 2.

PRZESŁANKA 3 Dla każdej naturalnej własności φ PYTANIE 1 nie wyraża tego samego trywialnego pytania co PYTANIE 2. Innymi słowy: dla każdej naturalnej własności φ PYTANIE 1 jest nietrywialnym pytaniem otwartym.

WNIOSEK 1 Dla każdej naturalnej własności φ „dobry” i „ φ ” nie znaczą tego samego.

WNIOSEK 2 Dla każdej naturalnej własności φ [DEF] nie jest poprawną definicją.

PRZESŁANKA 1 wyraża dosyć intuicyjne przekonanie na temat tego, co powinna robić definicja. PRZESŁANKA 2 bierze się z przekonania, że znaczenie całego wyrażenia jest wyznaczone przez znaczenie jego składowych (a zatem, że zastąpienie jednej składowej inną o tym samym znaczeniu nie zmieni znaczenia całości)²¹. PRZESŁANKA 3 umotywowana jest, po pierwsze, intuicją, że pytania typu „czy dobro to przyczynianie się do największej ilości szczęścia” nie mają trywialnych odpowiedzi, a po drugie, faktem, że filozofowie potrafią się o odpowiedzi na takie pytania spierać, chociaż zazwyczaj nie spierają się odnośnie do tego, czy np. kawalerowie to mężczyźni (nigdy) nieżonaci (pytania pozbawione odpowiedzi tego typu nazywał Moore **PYTANIAMI OTWARTYMI**).

Podstawowy więc pomysł jest taki, że chociaż nie możemy sensownie wątpić, czy dobro to dobro, to możemy sensownie

²¹ Na przykład, jeżeli myślimy, że „oraz” znaczy to samo, co „i”, to wedle tej przesłanki zdanie „Predator jest cwany i szybki” powinno mieć to samo znaczenie, co „Predator jest cwany oraz szybki”.

wątpić, czy dobro to przyczynianie się do największej ilości szczęścia itp. Możemy zadać to pytanie, rozumieć je, a zarazem nie być przekonanym, czy odpowiedź na nie jest pozytywna. To, że możemy wątpić w taką identyfikację, wskazuje (według Moore'a) na to, że pojęcia te nie są synonimami, a więc jedno nie może definiować drugiego. Wszelkie próby utożsamienia własności dobra z jakimiś własnościami naturalnymi nazywał Moore „błędami naturalistycznymi” i uważał, że argument z pytania otwartego jasno wskazuje, że zasługiwały one na to miano.

Ocenę tego argumentu rozłożymy na raty. Najpierw przyjrzemy się drobniejszym kwestiom, jakie się pojawiają. Potem dopiero skoncentrujemy się na tym, co stanowi, moim zdaniem, istotę problemu z argumentem Moore'a.

10. Problemy z antynaturalizmem Moore'a

Po pierwsze, jest interesujące, że niezależnie od niedefiniowalności dobra Moore uważał, że istnieją syntetyczne zdania o dobru, które są oczywiste i poznawalne mocą intelektu: „przyjemność nie jest jedynym dobrem”, „obcowanie z pięknem jest dobre” itp. Tego typu zdania są, wedle Moore'a, prawdziwe, znane bezpośrednio, bez pomocy argumentu i oczywiste. Moore porównywał wręcz oczywistość zdań etycznych z oczywistością zdań matematycznych.

Nie jest jednak jasne, w jaki sposób przysługiwać im ma taka oczywistość, skoro pytania typu „czy obcowanie z pięknem jest dobre” powinny być, jego zdaniem, pytaniami otwartymi. Wyzwaniem, którego się nie podjął, było wyjaśnienie (w sposób niezależny od swojego argumentu z pytania otwartego), dlaczego zasady etyczne, w przeciwieństwie do matematycz-

Problemy z antynaturalizmem Moore'a

nych, mają być syntetyczne, skoro powinny charakteryzować się podobnym stopniem oczywistości.

Co więcej, jeżeli istnieją oczywiste i poznawalne *a priori* prawdy etyczne, to niemal bez znaczenia wydaje się kwestia ich analityczności, a z ich nieanalityczności nie wynika wiele na temat tego, czy możliwe są interesujące argumenty etyczne.

Innym źródłem napięć w filozofii etycznej Moore'a jest jego konsekwencjalizm. Uważał on bowiem, że czyny są dobre na mocy swoich konsekwencji i że jest tak z definicji²². To jednak wydaje się podejrzanе, ponieważ do takiej identyfikacji również stosuje się argument z pytania otwartego. Pytanie „czy dobre czyny to takie, które mają dobre konsekwencje” jest nietrywialne i otwarte, w przeciwieństwie do pytania „czy dobre czyny to dobre czyny”²³.

Kolejnym problemem jest to, że przejście od niedefiniowalności dobra do niemożności istnienia interesujących argumentów etycznych jest bardzo pośpieszne i zdecydowanie nietrywialne. Niedefiniowalność dobra nie wyklucza istnienia okresów warunkowych (zdań typu *gdyby...*, *to...*, lub *jeżeli...*, *to...*) albo nawet równoważności dotyczących moralności, które są konieczne, poznawalne *a priori*, ale nie są prawdziwe na mocy definicji.

W takim przypadku również możliwe byłyby argumenty za dobrocią czegoś pomimo niedefiniowalności dobra. (Tak jak na przykład mogę *a priori* wiedzieć, że jeżeli jakiś przedmiot w przestrzeni jest nieprzejrzysty, to jego powierzchnia jest kolorowa, nawet jeżeli zgodzę się, że bycie kolorowym jest niedefiniowalne, albo takiej definicji nie znam)²⁴.

22 "What I wish first to point out is that right does and can mean nothing but cause of a good result..." (*Principia Ethica*, § 89).

23 Tego typu krytykę Moore'a rozwinął David Ross w *The Right and the Good* (1930). Moore wycofał się jednak z tego twierdzenia o synonimiczności już w 1912 r. (*Ethics*): "I think, that the meaning of the two words is not the same; for, if it were, then it would be a mere tautology to say that it is always our duty to do what will have the best possible consequences".

24 Atak z tej perspektywy na pozycję Moore'a znaleźć można np. w pierwszym tomie *Philosophical Analysis in the Twentieth Century* Soamesa.

Każda z tych pobieżnie poruszonych kwestii zasługuje na szczegółowe rozważenie, którego w niniejszym tekście się nie podejmę. Wspomniawszy je więc tylko, przejdę do rozważenia tego, co stanowi moim zdaniem najważniejszy element argumentu Moore'a.

11. Dalsze problemy: analiza a informacja

U podstaw zarówno Moore'owskiej krytyki idealizmu, jak i naturalizmu etycznego leży ten sam chwyt:

- „*Esse est percipi*” nie może zdaniem Moore'a być analityczne, bo żeby było ciekawym twierdzeniem, musi być nietrywialne (założenie tutaj jest takie, że zdania analityczne możemy rozstrzygnąć trywialnie); pytanie, czy *esse* to *percipi*, musi być pytaniem otwartym (w sensie Moore'a),
- „Dobro to φ ” nie może być jego zdaniem analityczne, bo wszelka taka identyfikacja jest nietrywialna, a pytanie o jej prawdziwość zawsze będzie pytaniem otwartym.

Ruch ten zakłada, że o prawdziwość zdań analitycznie prawdziwych nie ma sensu pytać: takie pytania nie mogą być otwarte i każdy kompetentny użytkownik języka, gdy tylko takie zdanie usłyszy, uzna je za trywialnie prawdziwe. Założenie to nazwijmy **ZASADĄ PRZEJRZYSTOŚCI ZNACZENIA**.

Rzeczywiście tego typu intuicja wspierana jest przez przypadki zdań prawdziwych na mocy znaczenia, które są oczywiste, gdy tylko je usłyszymy: „kawaler to nieżonaty mężczyzna”²⁵, „trójkąt to figura geometryczna o trzech kątach” itp.

²⁵ W sensie: mężczyzna, który nigdy nie był żonaty. W innym wypadku oczywistym kontrprzykładem byłby wdowiec.

Dalsze problemy: analiza a informacja

Strategia ta jest podejrzana, bo bardzo łatwo można uogólnić ją na dowolny wynik analizy filozoficznej. Wynik taki bowiem, o ile będzie interesujący, będzie nieoczywisty, a zatem – na mocy zasady przejrzystości znaczenia – nieanalityczny. O ile będzie analityczny, na mocy zasady przejrzystości znaczenia będzie trywialny, a więc nieinteresujący. Osoby, które traktują to uogólnienie poważnie i się nim przejmują, nazywają ten dylemat **PARADOKSEM ANALIZY**. Jednak co właściwie jest z taką strategią nie tak?

Pogląd, że ludzie potrafią bez żadnych trudności rozpoznawać wszystko, co jest prawdziwe na mocy znaczenia, jest zbyt pośpieszny. Dla przykładu, dość niekontrowersyjna jest myśl, że to, czy z danych przesłanek wynika logicznie dany wniosek, jest kwestią znaczenia wyrażań.

Wynikanie logiczne jest w szczególności kwestią znaczenia operatorów logicznych występujących w przesłankach i we wniosku. Na przykład zdanie „Króliki są groźne” wynika ze zdania „Króliki są groźne, a żaby wesołe” ponieważ „a” (spójnik koniunkcji) znaczy to, co znaczy.

Ponieważ jednak tak jest, to zasada przejrzystości znaczenia wymagałaby, żeby każdy kompetentny użytkownik języka, który rozumie przesłanki i wniosek, z łatwością mógł rozpoznać, czy ów wniosek z tych przesłanek wynika, czy nie. Tak jednak nie jest: nie mamy wszechwiedzy logicznej, popełniamy błędy w rozumowaniach i w naszych ocenach cudzych rozumowań.

Na przykład od dłuższego czasu całe stado matematyków bezskutecznie próbują ustalić, czy hipoteza Goldbacha (która głosi, że każda liczba parzysta większa od dwóch jest sumą dwóch liczb pierwszych) wynika z aksjomatów arytmetyki Peana. Nawet jeżeli wynikanie takie zachodzi, zdecydowanie nie jest tak, że jego zachodzenie jest trywialnie dostrzegalne, mimo że jest relacją zachodzącą na mocy znaczenia wyrażań, i raczej mamy wrażenie, że wyrażenia arytmetyczne rozumiemy.

Można dowieść, że nawet dla stosunkowo ubożego języka (np. dla rachunku predykatów pierwszego rzędu) relacja konsekwencji nie jest algorytmicznie rozstrzygalna. Czytelnik

zainteresowany szczegółami może zajrzeć do niemal dowolnego podręcznika logiki formalnej i dowiedzieć się, co to jest logika predykatów pierwszego i drugiego rzędu, żeby wiedzieć, o czym mowa. Zasadniczo logika pierwszego rzędu, poza spójnikami takimi jak „i”, „lub” czy „wtedy i tylko wtedy, gdy”, zajmuje się też orzekaniem własności o przedmiotach („ a jest F ”) oraz kwantyfikacją po przedmiotach („Pewne x jest F ”, „Każde x jest F ”). Logika drugiego rzędu różni się od logiki pierwszego rzędu tym, że oprócz tego, o czym była mowa w poprzednim zdaniu, zajmuje się kwantyfikacją po własnościach pierwszego rzędu (np. „Dla pewnego F , a jest F ”). Różnica ta okazuje się dość istotna. O ile na przykład w języku pierwszego rzędu nie da się jednoznacznie opisać struktury liczb naturalnych, o tyle da się to zrobić w języku rzędu drugiego. Relacja wynikania zarówno dla logiki pierwszego rzędu, jak i dla logiki drugiego rzędu nie jest rozstrzygalna. Pozostawmy to jednak na marginesie.

To wskazuje, że o ile zasada przejrzystości znaczenia znajduje zastosowanie (być może) w przypadkach prostych, to w przypadkach bardziej złożonych zawodzi. Z samej nieoczywistości jakiejś kwestii nie wynika jej nieanalityczność.

Powstaje pytanie, czy kwestie, za których nieanalitycznością argumentował Moore, są na tyle skomplikowane, żeby ich rozstrzygnięcie mogło być nieoczywiste, nawet jeżeli jest analityczne.

Wygląda na to, że odpowiedź powinna być twierdząca. Jeżeli w ogóle cokolwiek jest skomplikowane, to na pewno skomplikowane są siatki pojęć, którymi posługujemy się na co dzień. Dla większości pojęć, których używamy, nie potrafimy na bieżąco podać interesującej analizy, bo nasza umiejętność **stosowania** tych pojęć (tj. posługiwania się nimi w języku potocznym i codziennych rozumowaniach) nie wymaga, żebyśmy potrafili je **zdefiniować**.

Być może jednak celem filozofa analitycznego wcale nie musi być podawanie definicji. Dobrze przeprowadzona analiza może po prostu podawać interesujące prawdy ogólne dotyczące pojęć analizowanych i dochodzić do nich w sposób zgodny z intuicjami, jakie mamy w poszczególnych przypadkach zastosowań

Dalsze problemy: analiza a informacja

(zobaczmy jeszcze przykłady takiej procedury uzyskiwania balansu pomiędzy twierdzeniami ogólnymi a intuicjami). Jaka część wyniku takiej „analizy” jest rzeczywiście analityczna, nie zawsze jest jasne i nie zawsze jest to ważne.

W przypadku pojęć interesujących filozoficznie analiza jest zazwyczaj trudna, ponieważ pojęcia te występują w naszych intuicyjnych sądach w wielu różnych kontekstach. Często nie jest w ogóle jasne, czy te poszczególne zastosowania mają ze sobą coś wspólnego i czy istnieją interesujące prawdy ogólne zgodne ze wszystkimi tymi konkretnymi intuicjami. Aby wyjaśnić, o co mi chodzi, posłużę się następującym przykładem.

Wielu logików filozoficznych stara się poddać analizie nasze potoczne rozumienie okresów warunkowych (zdań o postaci „Gdyby p , to q ”, które będziemy zapisywać: „ $p > q$ ”). Starają się oni skonstruować ogólną teorię warunków prawdziwości takich zdań oraz logikę opisującą związki wynikania między nimi. Materiał zaś, za pomocą którego swoje teorie kontrolują, to codzienne ludzkie intuicje dotyczące prawdziwości lub wynikania pewnych konkretnych okresów warunkowych.

Jedna z najprostszych i historycznie najwcześniejszych prób podania warunków prawdziwości dla okresu $p > q$ powiadała:

$p > q$ wtedy i tylko wtedy, gdy nie jest tak, że p i nie- q .

Taki warunek prawdziwości owocuje pewnymi relacjami wynikania. Łatwo na przykład zauważyć, że jeżeli nie jest tak, że p i nie- q , to nie jest tak, że $(p \text{ i } r)$ i nie- q dla dowolnego zdania r . Zatem z $p > q$, wedle tej analizy, wynika $(p \text{ i } r) > q$. Skrótowo zapisywać ten fakt można jako: $p > q \models p \wedge r > q$, gdzie będę używał symbolu $\models$ do wyrażania relacji wynikania²⁶.

²⁶ Technicznie rzecz biorąc, istnieje różnica między **dowodliwością**, czyli wyprowadzalnością z przesłanek za pomocą reguł sformułowanych w ramach kształtu wyrażań (np. „z p i q wyprowadzić można q ”) i **wynikaniem w sensie semantycznym**, tj. relacją, która zachodzi między przesłankami a wnioskiem, gdy w każdej możliwej sytuacji, w której prawdziwe są przesłanki, prawdziwy jest też wniosek. W kontekstach, gdy to

W tym jednak momencie wkracza porównanie z naszymi intuicjami w konkretnych przypadkach. Jeżeli nasza analiza jest poprawna, ten związek wynikania zachodzi, a więc nie powinniśmy trafić na żaden konkretny przypadek zdań p , q , r , dla których jest oczywiste, że z $p > q$ nie wynika $p \wedge r > q$.

Weźmy jednak zdanie „Gdyby była ładna pogoda, poszlibyśmy na spacer”. To zdanie (o postaci $p > q$) może być prawdziwe, nawet gdy fałszywe jest zdanie (o postaci $p \wedge r > q$): „Gdyby była ładna pogoda, a mnie przejechałby autobus, poszlibyśmy na spacer”. Okazuje się więc, że wyjściowa analiza była problematyczna i coś trzeba z nią zrobić: albo zaproponować inną analizę, albo w przekonujący i spójny sposób wyjaśnić, dlaczego nasza intuicja w tym konkretnym przypadku miała się mylić. Próby podania bardziej udanych analiz stanowią fascynującą historię, którą zostawić muszę na inną okazję.

Ogólny morał jest taki: analiza stara się zarysować interesujące prawdy (często ogólne) dotyczące pojęć, którymi się posługujemy. Działalność ta wymaga, po pierwsze, formułowania ogólnych hipotez, które wcale nie muszą być oczywiste, po drugie, sprawdzania, czy hipotezy te wyjaśniają nasze intuicje dotyczące konkretnych przypadków i zgadzają się z nimi. Nasza aparatura pojęciowa jest skomplikowana, a więc skomplikowane i niespodziewane mogą być analizy filozoficzne. To, że nie są one oczywiste, nie dowodzi ich nieanalityczności, a co najwyżej wskazuje na to, że złożoność naszej aparatury pojęciowej często nas przerasta i nie dysponujemy na jej temat wszechwiedzą.

Z tego wszystkiego nie wynika, że wedle filozofów analitycznych wszystkie prawdy filozoficzne są analityczne, czy że jedyną rzeczą, którą zajmować ma się filozofia, jest język. Co najwyżej, że ważnym zjawiskiem jest fakt, że technicznie poprawne rozważania filozoficzne wymagają między innymi dobrego zrozu-

rozdzielenie jest istotne, sybolu $\vdash$ używa się dla dowodliwości, a dla wynikania w sensie semantycznym używa się symbolu $\models$. Dla nas jednak nie ma to obecnie znaczenia.

mienia pojęć, którymi się filozof posługuje, i w miarę dobrego rozumienia relacji, jakie między nimi zachodzą.

12. Dziedzictwo Moore'a

Prace Moore'a okazały się ważne przynajmniej z następujących powodów.

- O ile argument etyczny z otwartego pytania oraz argument przeciwko idealizmowi nie były powalające, o tyle możliwość ich sformułowania wskazywała na **potrzebę dokładniejszego przemyślenia rozróżnień pomiędzy sędami analitycznymi i syntetycznymi, pomiędzy sędami oczywistymi a nieoczywistymi, pomiędzy sędami *a priori* i *a posteriori***. Istotna część rozwoju analitycznej filozofii języka w XX wieku to właśnie dochodzenie do ładu z tego typu pojęciami.
- O ile argument przeciwko sceptycyzmowi nie byłby w stanie przekonać sceptyka, o tyle wskazywał on na potrzebę dokładniejszego **rozważenia założeń filozoficznych sceptycyzmu**.
- Argument ten stanowił przykład strategii, która w XX wieku stała się dość popularna. Strategia ta polegała na **nieignorowaniu naszych codziennych sądów potocznych i ocenianiu naszych skomplikowanych teorii filozoficznych z perspektywy zgodności z nimi**. Gdy jakaś teoria wyklucza pewne sądy, które intuicyjnie uznajemy za trafne, należy całkiem na poważnie zastanowić się, co jest bardziej przekonujące: owe sądy czy założenia danej teorii.
- Tego typu podejście w pewnym sensie zarysowało **plan działania dla części epistemologów** (epistemologia to wywodząca się z języka greckiego – *episteme* znaczy *wiedza* – nazwa teorii poznania lub też teorii wiedzy, która zajmuje się opisem tego, jak wiedzę prawomocnie należy uzyskać

wać). Zamiast walczyć ze sceptycyzmem, część filozofów skupiła się raczej na **wyjaśnianiu, w jaki sposób wiemy to, co wiemy**, w sposób zgodny z większością standardowych użyciu pojęcia wiedzy.

- Argument z pytania otwartego oraz polemika ze sceptycyzmem stanowiły motywację dla szerszej debaty dotyczącej **statusu zasad etycznych**, która miała miejsce w XX wieku.
- **Sam styl Moore'a był nieco zaraźliwy.** Starał się on pisać w języku potocznym i unikać tworzenia nowych wyrazów czy używania wyrazów niejasnych. Unikał konstruowania ogólnych systemów światopoglądowych. Koncentrował się nie tylko na tezach filozoficznych, lecz także na tym, czy na pewno są one zrozumiałe, na rozumieniu ich treści oraz na dokładnym rozważeniu znanych argumentów za i przeciw.
- **Pewnym mitem jest to, że filozofowie analityczni zajmują się tylko językiem.** O ile później rozważymy, ile jest w tym prawdy o nieco bardziej współczesnych filozofach analitycznych (niewiele), o tyle widzimy, że nie jest to prawdą w przypadku Moore'a. Bronił on całkiem poważnych i treściwych tez filozoficznych: dobra jako prostej własności postrzeganej intelektem, konsekwencjalizmu w etyce, realizmu poznawczego i etycznego.



TEKSTY POLECANE

1. Sam tekst Moore'a *The Refutation of Idealism* w bezpłatnej wersji anglojęzycznej można łatwo wygooglować. Niestety, tekst ten nie jest ani wybitnie fascynujący, ani przystępny w stylu, ani tak jasny, jak mógłby być.
2. Innym tekstem klasycznym jest *Proof of an External World* Moore'a. Również łatwo wygooglować wersję oryginalną dostępną on-line za darmo. Tekst jest dużo przystępniej napisany niż *The Refutation of Idealism*.
3. Momentami miłe w lekturze (jeżeli nie liczyć długawych i nudnawych rozważań o danych zmysłowych) jest *Some Main Problems of Philosophy* Moore'a. Polskie wydanie nosi tytuł *Z głównych zagadnień filozofii*.

Dziedzictwo Moore'a

4. **Principia Ethica** Moore'a również można znaleźć w bezpłatnej wersji elektronicznej. Wersję polskojęzyczną pt. **Zasady etyki** znaleźć można na przykład tutaj: <http://sady.up.krakow.pl/antfil.moore.preth.htm>.
5. Podobnie sprawa ma się z nieco późniejszą pracą Moore'a **Ethics**.
6. Klasyczną polemikę z poglądami etycznymi Moore'a stanowi **Good and Evil** Petera Geacha (Analysis 17, 1956).
7. Warto też spojrzeć na **Zagadnienia etyki** M. Schlicka.

MOORE KONTRA IDEALIZM

1. **Percipi** jest rozumiane szeroko (obiekty matematyczne itd.).
2. EP nie jest definicją: **percipi** nie pociąga **esse**.
3. EP: **percipi** jest konieczne (ale niewystarczające) dla **esse**.
4. EP jest konieczne (bo filozoficznie doniosłe).
5. EP nie jest analityczne, bo nie jest trywialne.
6. EP nie jest **a posteriori**, bo (i) nie wymaga żadnego konkretnego doświadczenia oraz (ii) idealisci uważają prawdy filozoficzne za prawdy rozumu.
7. Jeżeli EP jest **a priori**, to albo jest oczywiste, albo dowodliwe (logicznie).
8. EP nie jest oczywiste.
9. Jeżeli EP jest logicznie dowodliwe, to jest analityczne i nieinformatywne – wbrew (5).



MOORE KONTRA SCEPTYCYZM

1. Ważenie intuicji: zdrowy rozsądek a sceptyczna teoria poznania.
2. Argument z **Proof of an External World**: Oto ręce – istnieją w przestrzeni, są więc przedmiotami zewnętrznymi względem umysłu.
3. Warunki poprawności argumentu są spełnione: przesłanki są znane, różnią się od wniosku i wniosek z nich wynika.
4. Odrzucenie (3) wymaga założeń filozoficznych mniej intuicyjnych niż same przesłanki.



MOORE KONTRA SUBIEKTYWIZM ETYCZNY

1. Subiektywizm tu rozważany: twierdzenia wartościujące tak naprawdę mówią tylko o postawach, nie o faktach.
2. Problem: ludzie nie zgadzają się co do wartościowań.



Wybrani dziadkowie filozofii analitycznej

3. Gdyby po prostu stwierdzali swoje postawy, nie byłoby przedmiotu niezgody: „Ja lubię x , ty nie lubisz x ”.



MOORE KONTRA NATURALIZM ETYCZNY

1. Naturalistyczne teorie próbują zdefiniować dobro za pomocą własności naturalnych.
2. Przykład utilitaryzmu.
3. Moore, *Principia Ethica*: dobro jest niedefiniowalne.
4. Skoro jest niedefiniowalne, nie istnieją interesujące argumenty w etyce.
5. Prawdy etyczne: jedynie przedmiot intuicji intelektualnej.
6. Nie wyklucza to prawdziwości zdań typu [DEF], jedynie ich analityczność i aprioryczność.
7. Argument pytania otwartego: „ x jest dobre” i „ x jest ϕ ” nie znaczą tego samego dla naturalnego ϕ , bo pytanie, czy dobre x jest ϕ , jest pytaniem nietrywialnym, otwartym.



TRUDNOŚCI ANTYNATURALIZMU MOORE’A

1. Moore sądził, że istnieją oczywiste i poznawalne intelektem prawdy o dobru.
2. Nie jest jasne, dlaczego są oczywiste, skoro odpowiadające im pytania także są otwarte.
3. Jeżeli istnieją takie prawdy, to pytanie, czy są analityczne, jest nieistotne, a argument pytania otwartego dotyczy analityczności.
4. Moore uważał (przez dłuższy czas), że konsekwencjalizm jest prawdziwy z definicji. Stosuje się jednak do tego poglądu argument pytania otwartego.
5. Przejście od niedefiniowalności do nieistnienia interesujących argumentów jest za szybkie: co ze zdaniem *a priori* prawdziwymi, ale prawdziwymi nie na mocy definicji?
6. Argument pytania otwartego jest analogiczny do argumentu przeciwko *esse est percipi*: identyfikacja nie może być analityczna, bo jest nietrywialna.
7. Założenie: przejrzystość znaczenia – każdy kompetentny użytkownik języka pozna od razu zdania prawdziwe analitycznie.
8. Przejrzystość znaczenia prowadzi do paradoksu analizy: skoro zdania analityczne są trywialne, analiza jest albo trywialna, albo nieanalityczna.

Frege o wiedzy matematycznej

9. Przejrzystość znaczenia jest za silna: przykład konsekwencji logicznej i twierdzeń matematycznych.
10. Często pojęcia potrafimy tylko stosować, a nie definiować; nasza aparatura pojęciowa jest skomplikowana.
11. Ważenie intuicji w analizie: trudne. Przykład okresów warunkowych.

KIERUNKOWSKAZ

Teraz właściwie należałoby przejść do omówienia tez kolegi Moore'a, Bertranda Russella. Żeby jednak zrozumieć to, co będę chciał o nim powiedzieć, dobrze najpierw mieć chociażby ogólne pojęcie o pewnym niemieckim prekursorze logiki matematycznej oraz filozofii analitycznej, który nazywał się Gottlob Frege.



13. Frege o wiedzy matematycznej

Zrozumienie, w jaki sposób możemy coś wiedzieć i jakie są źródła naszej wiedzy, jest ważne – lepiej rozumiejąc takie kwestie, trochę lepiej będziemy odróżniać przekonania uzasadnione od jedynie mniemań i przesądów oraz lepiej oceniać będziemy stopień uzasadnienia naszych przekonań. Ze względu na to, że matematyka wydaje się wzorcowym przykładem wiedzy, na szczególne zainteresowanie zasługują wiedza matematyczna oraz natura przedmiotu matematyki. Jednym z ważniejszych myślicieli zajmujących się statusem wiedzy matematycznej był GOTTLOB FREGE (1848–1925).

Za czasów Fregego matematykom udało się już sprowadzić wiele dyscyplin matematycznych do szeroko rozumianej arytmetyki: teorii liczb naturalnych i rzeczywistych²⁷. Wyglądało więc na to, że zrozumienie natury wiedzy matematycznej

²⁷ Obecnie nazwę „arytmetyka” rezerwuje się raczej dla teorii liczb naturalnych.

sprowadzało się do zrozumienia natury arytmetyki. Temu też zadaniu, z uporem maniaka, Frege poświęcił około 27 lat swoich badań, przy okazji zahaczając o wiele powiązanych z tym problemem szerszych kwestii filozoficznych.

W szczególności Frege przeciwstawiał się Kantowi, który uważał, że prawdy matematyczne są syntetyczne *a priori*. W przeciwieństwie do Kanta Frege uważał, że prawdy arytmetyczne w istocie są prawdami analitycznymi, prawdziwymi na mocy znaczenia terminów użytych. Co więcej, jako że w przypadku matematyki istotne jest tylko znaczenie wyrażen uznawanych za badane przez logikę (a więc np. spójników, takich jak *jeżeli...*, *to...*, kwantyfikatorów, takich jak *dla każdego*) czy, dla Fregego, takich wyrażen, jak ekstensja (odniesienie) pojęcia, dlatego powiadamy, że Frege uważał, że prawdy matematyki są „prawdami logicznymi”²⁸. Z Kantem jednak zgadzał się co do tego, że filozofia matematyki powinna wyjaśniać fakt, że matematyka wydaje się nam konieczna i *a priori*.

O ile prawdy logiczne są prawdziwe na mocy znaczenia, o tyle matematyka okazałaby się analityczna. Jeżeli zaś znaczenie wyrażen jest nam poznawczo dostępne *a priori*, matematyka w takim razie okazałaby się *a priori*. Jeśli prawdy logiczne są konieczne, rozmaite teorie matematyczne okazałyby się również konieczne. Chciałoby się też powiedzieć: o ile logika nie postuluje istnienia żadnych dziwnych przedmiotów, o tyle matematyka wymagałaby dosyć skromnej ontologii. Frege jednak na serio wierzył w istnienie obiektów matematycznych, takich jak liczby naturalne czy zbiory.

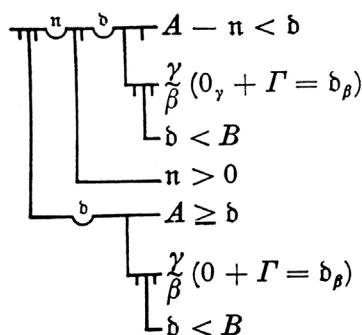
Również w przeciwieństwie do Kanta Frege uważał, że żywione przekonania filozoficzne trzeba popierać przekonującymi i zrozumiałymi argumentami. Diabeł tkwi w szczegółach: jeżeli chcemy uzasadnić twierdzenie, że prawdy arytmetyczne są prawdami logicznymi, to wypada, po pierwsze, powiedzieć, co rozumiemy przez logikę, po drugie, wyjaśnić, na czym nasz

²⁸ Frege zgadzał się natomiast z Kantem co do syntetyczności zdań geometrii.

Frege o wiedzy matematycznej

pomysł polega, i po trzecie, rzeczywiście wyprowadzić arytmetykę z twierdzeń logicznych.

Odpowiednio więc, jeśli chodzi o realizację pierwszego celu, pierwszym ważnym dziełem Fregego była nieco ceglasta książka zatytułowana *Begriffsschrift* (*Notacja pojęciowa*), wydana w roku 1879. Chociaż notacja, zasadniczo rzecz biorąc, była nieco nietypowa i ewidentnie nieprzyjazna dla zecera (por. ilustracja)²⁹,



to jednak Frege w ten sposób sformułował podstawowe systemy logiki klasycznej w ich niemal współczesnej formie, jeżeli pominąć kwestie czysto notacyjne. (Istnieje też aplikacja na iPada służąca do posługiwania się notacją z *Begriffsschrift*, nazywa się iBGS.)

Kolejne dzieło, *Die Grundlagen der Arithmetik* (*Podstawy arytmetyki*, 1884), zdecydowanie mniej sformalizowane, realizowało cel drugi i poświęcone zostało głównie filozoficznym poglądom Fregego na temat arytmetyki.

Trzecie dzieło, *Grundgesetze der Arithmetik* (*Podstawowe prawa arytmetyki*), poświęcone było realizacji celu trzeciego. Pierwszy tom ukazał się w roku 1893, a drugi w 1903. Tom pierwszy zawierał niepolemiczną ekspozycję logiki i pojęć używanych przez Fregego, tom drugi zaś formalne dowody wyprowadzające arytmetykę liczb naturalnych z tego, co Frege uważał za logikę. Frege miał zamiar napisać również tom trze-

²⁹ Ilustracja pochodzi ze strony <https://www.flickr.com/photos/15537756@Noo/1362372120>

ci, w którym prawdopodobnie planował dokończyć rozpoczętą w części drugiej analizę teorii liczb rzeczywiście. Jednak do rzeczy: na czym polegało podejście Fregego?

Kluczową rolę u Fregego odgrywały „pojęcia” (czy też „koncepty”). Pewne pojęcia są pojęciami pierwszego rzędu: stosują się do przedmiotów pozajęzykowych (np. pojęcie „krowy” jest pojęciem pierwszego rzędu). Pewne inne pojęcia są pojęciami drugiego rzędu: stosują się do pojęć pierwszego rzędu (np. pojęcie „pojęcia pierwszego rzędu” jest samo pojęciem drugiego rzędu). Co ważne, pojęcia nie muszą odpowiadać gramatycznie prostym wyrażeniom: nawet złożone opisy własności wyrażają pojęcia.

Jakiego rzędu są pojęcia związane z liczebnością, takie jak „dziesięć” w zdaniu „mam dziesięć palców”? Frege uważał, że pojęcie to nie stosuje się do samych przedmiotów, tylko do pojęcia palców. Nie liczymy bowiem przedmiotów jako takich, a raczej przedmioty ze względu na pojęcia: ten sam przedmiot może być jednym wiadrzem wody, pięcioma litrami wody, dwudziestoma szklankami wody itd. Pojęcie „dziesięć” jest więc pojęciem drugiego rzędu, które stosuje się do wszystkich pojęć pierwszego rzędu, które stosują się do dokładnie dziesięciu przedmiotów; pojęcie „zero” jest pojęciem drugiego rzędu, które stosuje się do wszystkich pojęć pierwszego rzędu, które nie stosują się do niczego (takich jak „jednorożec”, „filozofia bezzałożeniowa”, „uczciwy system podatkowy” itd.).

Drugim ważnym pojęciem jest pojęcie „ekstensji”. Zdaniem Fregego każdemu pojęciu odpowiada zbiór przedmiotów, do których się ono stosuje. Pojęciu „moje palce” odpowiada jego ekstensja: zbiór moich palców. Pojęciu „filozofia bezzałożeniowa” odpowiada zbiór pusty. Wedle Fregego każde pojęcie ma ekstensję. Pojęcie ekstensji kontrastuje się z pojęciem „intensji”, to znaczy treści danego wyrażenia, która tylko wyznaczać ma to, do czego to wyrażenie się odnosi.

Pojęcia „pojęcie” oraz „ekstensja” są, zdaniem Fregego, pojęciami logicznymi. Gdyby więc dzięki nim udało się zdefiniować pojęcie liczby, tak żeby za pomocą tych definicji oraz

Frege o wiedzy matematycznej

samej logiki dowieść zasadniczych prawd arytmetycznych, cel ukazania logiczności arytmetyki byłby zrealizowany. Tego więc zadania podjął się Frege.

Liczbę **ZERO** zdefiniował jako ekstensję pojęcia „zero”: to znaczy zbiór wszystkich pojęć, które do niczego się nie stosują. Liczbę **JEDEN** zdefiniował jako ekstensję pojęcia „jeden”: zbiór wszystkich pojęć, które stosują się do dokładnie jednego przedmiotu. Ogólnie każdą liczbę x utożsamiał ze zbiorem tych pojęć, które stosują się do dokładnie x przedmiotów.

Co więcej, zrobił to na tyle sprytnie, że pojęcie liczby nie występowało w jego definicjach. Na przykład zamiast powiedzieć, że dane pojęcie stosuje się do zera przedmiotów, wystarczy powiedzieć: nie istnieje takie x , do którego stosuje się to pojęcie. Zamiast powiedzieć, że pojęcie stosuje się do dokładnie jednego przedmiotu, można powiedzieć: istnieje takie x , do którego stosuje się to pojęcie, i każde y , do którego stosuje się to pojęcie, jest identyczne z tym x . Zamiast powiedzieć, że pojęcie stosuje się do dokładnie dwóch przedmiotów, można powiedzieć: istnieje takie x i takie y , że x nie jest identyczne z y , pojęcie stosuje się zarówno do x , jak i do y , oraz dla dowolnego z , jeżeli pojęcie stosuje się do z , to z jest albo identyczne z x , albo identyczne z y . I tak dalej.

Po podaniu definicji pozostawało wyprowadzić powszechnie przyjęte aksjomaty arytmetyki. W tym wypadku chodzi o arytmetykę Peana drugiego rzędu. Ma ona następujące aksjomaty:

Aksjomat	Znaczenie
$\forall x S(x) \neq 0$	0 nie jest następnikiem żadnej liczby.
$\forall x \forall y [S(x) = S(y) \rightarrow x = y]$	Dwie różne liczby nie mają tego samego następnika.
$\forall x (x + 0 = x)$	Część definicji dodawania.
$\forall x \forall y (x + S(y) = S(x + y))$	Część definicji dodawania.
$\forall x (x \times 0 = 0)$	Część definicji mnożenia.
$\forall x \forall y (x \times S(y) = (x \times y) + x)$	Część definicji mnożenia.
$\forall P [P(0) \wedge \forall x (P(x) \rightarrow P(S(x))) \rightarrow \forall x P(x)]$	Aksjomat indukcji.

Sens dwóch pierwszych aksjomatów jest dość jasny. Jeżeli chodzi o aksjomaty „definiujące” operacje dodawania i mnożenia, dość łatwo dostrzec, co one robią: określają daną operację dla przypadku „bazowego” (mnożenie przez zero, dodawanie zera), a potem określają operację dla przypadku $n+1$ za pomocą tej operacji dla przypadku n . I tak, jeżeli chcemy dowiedzieć się, jak wygląda wynik dodania liczby 2, najpierw dowiadujemy się, że $x + S(1) = S(x+1)$. Podobnie $x+1 = x + S(0) = S(x+0)$. Wiemy jednak, że $x+0 = x$, więc $x+1 = Sx$. Podstawiamy więc w pierwszym wzorze jedno za drugie i dostajemy: $x + S(1) = SSx = 2$.

Nieco bardziej skomplikowane zadanie wykonuje aksjomat indukcji. Dzięki niemu, jeżeli wiemy, że liczba zero posiada jakąś własność i że jeżeli jakaś liczba posiada tę własność, to jej następnik również ją posiada, to wiemy, że wszystkie liczby posiadają tę własność. W pewnym sensie aksjomat indukcji mówi nam, że nie ma żadnych innych liczb poza tymi, do których można „dotrzeć”, idąc od zera skończoną liczbę kroków po funkcji bycia następnikiem.

Ściśle rzecz biorąc, sam sposób sformułowania aksjomatów ukrywa jedno podstawowe założenie. Mianowicie, od razu zakładamy, że wyrażenie „0” odnosi się do czegoś, i że symbol „S” wyraża funkcję przyporządkowującą każdemu przedmiotowi jakiś przedmiot. Razem z założeniem, że dwie różne liczby nie mają tego samego następnika, pociąga to za sobą wniosek, że istnieje nieskończoność liczb naturalnych.

Zastanówmy się, jak posłużyć się definicją liczby podaną przez Fregego dla wyprowadzenia aksjomatów Peana. Dla przykładu: jak dowieść, że istnieje nieskończenie wiele liczb?

1. Bierzemy pierwsze lepsze pojęcie, które do niczego się nie stosuje, na przykład $x \neq x$ („niebycie identycznym z samym sobą”).
2. Wnioskujemy, że istnieje jego liczba (czyli ekstensja pojęcia „pojęcie, które stosuje się do tego samego co pojęcie $x \neq x$ ”).
3. Nazywamy tę liczbę liczbą zero.

Frege o wiedzy matematycznej

4. Bierzemy pojęcie $x = 0$ („bycie identycznym z liczbą zero”).
5. Zauważamy, że pojęcie to stosuje się do dokładnie jednego przedmiotu.
6. Wnioskujemy, że istnieje jego liczba.
7. Nazywamy ją liczbą 1.
8. Bierzemy pojęcie $x = 0 \vee x = 1$ („bycie identycznym z liczbą zero lub z liczbą jeden”).
9. Zauważamy, że pojęcie to stosuje się do dokładnie dwóch przedmiotów.
10. Wnioskujemy, że istnieje jego liczba.
11. Nazywamy tę ekstensję liczbą 2.
12. Kontynuujemy w ten sposób.

Dalsze szczegóły nie są dla nas ważne. Istotne są natomiast trzy rzeczy. Po pierwsze, sposób, w jaki Frege wprowadził pojęcie ekstensji do swojego systemu. Frege uważał, że aby wyjaśnić, co to jest ekstensja, wystarczy podać **KRYTERIUM IDENTYCZNOŚCI EKSTENSJI**: powiedzieć, kiedy dwie ekstensje są identyczne. Uczynił to za pomocą całkiem sławnej zasady zwanej **PODSTAWOWYM PRAWEM NUMER PIĘĆ** (w skrócie PPV):

Ekstensja pojęcia F jest identyczna z ekstensją pojęcia G wtedy i tylko wtedy, gdy dla dowolnego przedmiotu x : pojęcie F stosuje się do x wtedy i tylko wtedy, gdy do x stosuje się również pojęcie G .

PPV wygląda nieco bardziej przejrzyście, gdy zapiszemy je w notacji sformalizowanej:

$$Ekst(F) = Ekst(G) \equiv \forall x [F(x) \equiv G(x)]$$

Po drugie, bardzo ważną konsekwencją PPV jest istnienie ekstensji dla dowolnego pojęcia (tezę tę nazywa się **PEŁNĄ KOMPREENSIJĄ** albo **PEŁNYM WYRÓŻNIANIEM**). Wynika ona z PPV w sposób następujący:

1. PPV ma stosować się do dowolnych pojęć F i G . Stosuje się więc również, gdy G i F to to samo pojęcie:

$$Ekst(F) = Ekst(F) \equiv \forall x[F(x) \equiv F(x)]$$

2. Prawa strona mówi jednak, że F stosuje się do tych rzeczy, do których się stosuje, i jest prawdą logiczną.
3. Dostajemy więc lewą stronę:

$$Ekst(F) = Ekst(F)$$

4. Skoro ekstensja F jest identyczna z ekstensją F , to wynika z tego również, że coś jest ekstensją F :

$$\exists x(x = Ekst(F))$$

Po trzecie, Frege w swojej logice zaczął od PPV i wyprowadził z PPV inną bardzo ważną zasadę, która określa identyczność liczb, tak jak PPV określało identyczność ekstensji. Zasada ta nazywa się **ZASADĄ HUME'A** (ZH):

[ZASADA HUME'A] Liczba pojęcia F (liczba F -ów) jest identyczna z liczbą pojęcia G wtedy i tylko wtedy, gdy pomiędzy pojęciami F i G zachodzi relacja równoliczności.

Bardziej formalnie:

$$\#(F) = \#(G) \equiv F \sim G$$

Następnie, już nie posługując się PPV, w swojej logice Frege z ZH wyprowadził wszystkie aksjomaty arytmetyki Peana drugiego rzędu. Jak okaże się w rozdziale 33, jest to dość istotny fakt.

Co ważne, pojęcie równoliczności nie zakłada pojęcia liczby. Dwa pojęcia są równoliczne, gdy przedmioty, do których się stosują, można wzajemnie jednoznacznie sobie przyporządkować. Wyjaśnić to można w sposób następujący.

Frege o wiedzy matematycznej

Wyobraźmy sobie, że dwóch miłośników oranżady, Zbigniew i Włodzimierz, nabyło skrzynkę tegoż napoju (skrzynka zawierała 24 butelki) i postanowiło wypić ją razem, dzieląc się równo butelkami. Pierwsza strategia, jaka przysła im do głowy, była następująca:

Po prostu liczymy i pamiętajmy, ile kto wziął butelek, i pilnujemy, żeby każdy z nas wypił dwanaście.

Niestety, rozmawiając wesoło o filozofii matematyki, gdzieś w okolicy butelki numer sześć Zbigniew i Włodzimierz stracili rachubę, a po ósmej żaden z nich nie pamiętał nawet, że mieli cokolwiek liczyć. Następnego dnia rano nie byli w stanie rozstrzygnąć, czy skrzynką rzeczywiście podzielili się po równo.

Następnym razem w podobnych okolicznościach przyjęli więc inny plan działania:

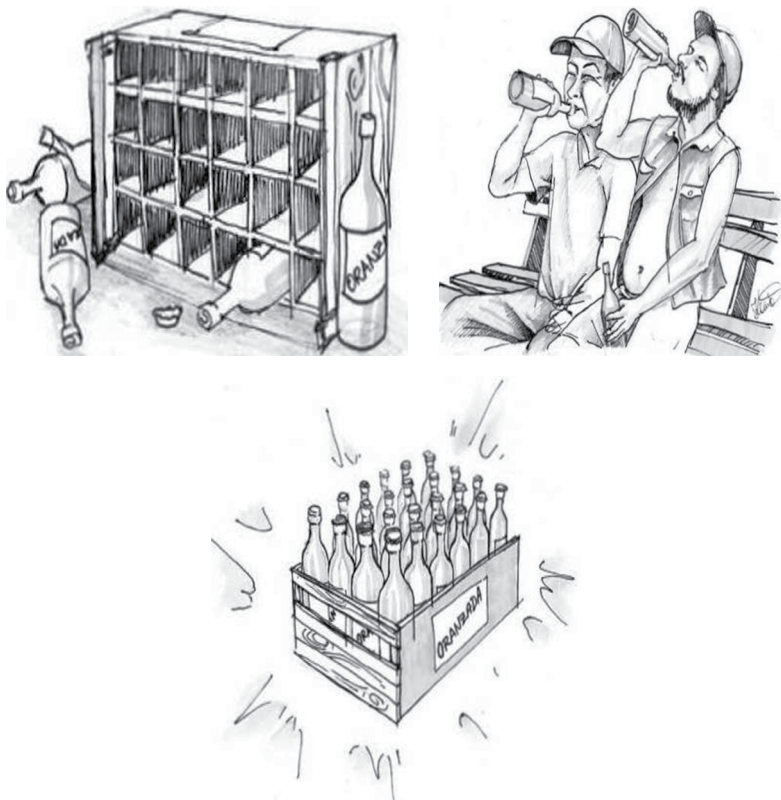
Zamiast starać się liczyć butelki i pamiętać, kto ile wypił, po prostu pilnujemy, żebyśmy butelki brali jednocześnie. Najpierw bierzemy razem po jednej butelce, pijemy, aż obaj je wykończymy, potem jednocześnie bierzemy po następnej butelce itd. W ten sposób musimy jedynie zwracać uwagę na to, żeby sięgać po butelki w tym samym czasie, a nie musimy ich tak właściwie liczyć.

W ten prosty sposób Zbigniew i Włodzimierz ustanowili równoliczność pomiędzy dwoma pojęciami: *butelki oranżady należące do pana Zbigniewa oraz butelki oranżady należące do pana Włodzimierza*, nie posługując się liczbami.

Otóż pomysł Fregego był bardzo podobny. By stwierdzić, że pojęcia F i G są równoliczne, zamiast liczyć, ile przedmiotów pod nie podpada i porównywać te liczby, można po prostu upewnić się, że każdemu przedmiotowi podpadającemu pod F odpowiada jednoznacznie jakiś przedmiot podpadający pod G , w taki sposób, że:

- dla każdego x podpadającego pod F istnieje dokładnie jeden przedmiot y podpadający pod G , który odpowiada x -owi,

- dla każdego y podpadającego pod G istnieje dokładnie jeden przedmiot x podpadający pod F , któremu y odpowiada.



Na pierwszy rzut oka (a dla Fregego nawet i na więcej rzutów oka, bo trochę czasu nad tym spędził) rzecz wygląda świetnie: mamy dobrze zdefiniowaną logikę, mamy dobrze zdefiniowane pojęcie klasy i liczby, dowiedliśmy aksjomatów arytmetyki, a zatem i pokazaliśmy, że arytmetyka jest w istocie logiką, więc i prawdy matematyczne są prawdami analitycznymi. Zwróćmy uwagę: teza, że prawdy arytmetyczne są prawdami logicznymi, jest silniejsza od tezy, że prawdy arytmetyczne są analityczne. Nie wszystkie prawdy analityczne muszą być dowodliwe w logice. Stąd nawet jeżeli okaże się, że prawdy arytmetyczne nie są

Frege: sens i odniesienie

prawdami logicznymi, nie wyniknie jeszcze, że są prawdami syntetycznymi. Jak jednak pokazał BERTRAND RUSSELL, sprawa jest nieco bardziej problematyczna, ponieważ system Fregego jest sprzeczny. Dlaczego tak jest, dowiemy się już niedługo, bo w rozdziale 16. Zanim do tego dojdzie, przyjrzyjmy się jednak jeszcze innemu aspektowi Fregowskiej filozofii, który okaże się istotny w rozwoju filozofii analitycznej w XX wieku.

14. Frege: sens i odniesienie

Za czasów Fregego dominujący obraz tego, jak funkcjonuje język, był psychologiczny: wedle poglądu psychologicznego język służy do komunikacji myśli pojmowanych jako obiekty mentalne zbudowane z IDEI. Pojęcie idei jest tutaj zazwyczaj nieco specyficzne, ale dla nas ważne jest to, że są one pewnego rodzaju obrazami mentalnymi; są one również prywatne – nikt nie może bezpośrednio zobaczyć idei nikogo innego. Zgodnie z tym podejściem używamy słów arbitralnie powiązanych z ideami (np. nic obiektywnego nie zmusza nas, byśmy z naszą ideą wiertarki wiązali właśnie wyraz „wiertarka”), aby zakomunikować innym ludziom, jakie myśli mamy na myśli. Słowa wyrażają idee, a zdania złożone ze słów wyrażają myśli złożone z idei.

Podejście psychologiczne jest problematyczne i nie wyjaśnia za dobrze zjawiska komunikacji. Załóżmy bowiem, że ktoś mówi do mnie „Wiertarki szczekają”. Żebym mógł to zrozumieć, muszę wiedzieć, jakie idee mówiący wiąże z wyrazami „wiertarka” i „szczekać”. Skoro jednak nie istnieje żadna obiektywna przyczyna, dla której wyraz „wiertarka” miałby wyrażać jakąś konkretną ideę (powiedzmy „wiertarka”), to nie mogę jedynie na podstawie samego wyrazu dowiedzieć się, o jaką ideę cho-

dzi mówiącemu. Musiałem wcześniej nauczyć się, że wyraz „wiertarka” nasz mówiący wiąże z ideą „wiertarki”.

W jaki sposób mogłem się jednak tego dowiedzieć? Idee są prywatne i mówiący nigdy nie mógł pokazać mi bezpośrednio swojej idei i powiedzieć: „Zobacz, tę ideę będę wiązał z wyrazem *wiertarka*”. Oczywiście mogł mi ten związek wytłumaczyć za pomocą innych wyrażen, ale wtedy pojawia się pytanie: w jaki sposób wiem, jakie idee wiąże on z tymi innymi wyrażeniami? Ściśle rzecz biorąc, tego argumentu Frege nie skonstruował; jest on raczej zbliżony do tzw. argumentu przeciwko językowi prywatnemu, jaki skonstruował Ludwig Wittgenstein, jeden z najbardziej przereklamowanych filozofów XX wieku.

Tego typu rozważania skłaniają do przyjęcia bardziej obiektywistycznej i niepsychologicznej teorii znaczenia; jedną z ważniejszych takich teorii skonstruował Frege. Odrzucał on twierdzenie, jakoby wyrażenia odnosiły się do idei, podając przynajmniej dwa argumenty.

Argument pierwszy głosi, że ludzie rozumieją się nawzajem, mimo że z używanymi wyrażeniami wiążą rozbieżne idee. I faktycznie, nasze wyobrażenia wiertarek, gdy będziemy używać wyrazu „wiertarka”, będą się różnić; różnić się też będziemy pojęciem, jakie mamy o wiertarkach, oraz stanami mentalnymi. To, co się dzieje w naszych umysłach, gdy używamy wyrazu „wiertarka”, nie jest dokładnie takie samo u wszystkich. Mimo to mamy silną intuicję, wedle której używamy tego wyrazu w jednym i tym samym znaczeniu – w innym wypadku nie bylibyśmy w stanie skutecznie się komunikować, a jednak dość często się to nam udaje (chyba że próbujemy komunikacji z jakimś urzędem). Znaczenia więc nie można utożsamiać z ideą.

Drugi argument powiada, że wyrażenia nie mogą odnosić się do idei, bo w większości przypadków, gdy rozprawiamy o czymkolwiek, to mówimy o świecie zewnętrznym, a nie o ideach. Gdy zajmujemy się matematyką, nie mówimy o naszych ideach liczb, tylko o liczbach; gdy mówimy o wiertarkach, to mówimy o rzeczywistych narzędziach, a nie o naszych ideach wiertarek.

Frege: sens i odniesienie

Do jakiego stopnia argumenty te są przekonujące, nie będziemy wnikali (osobiście uważam pierwszy za nieco lepiej sformułowany i bardziej przekonujący). Ważne jest to, że Frege zaproponował pewną alternatywę obiektywistyczną: wyrażenia muszą wyrażać coś obiektywnego i zewnętrznego w stosunku do naszych umysłów.

Co jednak mogą wyrażać wyrażenia, jeżeli nie idee? Być może po prostu odnoszą się one bezpośrednio do samych rzeczy? Przyjrzyjmy się, co by się stało, gdyby tak było w przypadku, powiedzmy, nazw i terminów jednostkowych.

Następujący problem nazywa się **ZAGADKĄ INFORMATYWNYCH IDENTYCZNOŚCI**. Jeżeli jedyne, co nazwy wnoszą do rozumienia wyrażen, w których występują, to przedmioty, do których się odnoszą, to zdanie:

[F1] Mark Twain to Mark Twain

wyraża dokładnie to samo, co zdanie:

[F2] Mark Twain to Samuel Langhorne Clemens.

To jednak wydaje się dosyć nieintuicyjne: o ile można nie wiedzieć, czy zdanie drugie jest prawdziwe, o tyle nie można nie wiedzieć, że prawdziwe jest zdanie pierwsze, jeżeli tylko się je rozumie. Musi więc istnieć jakaś różnica w znaczeniu między „Mark Twain” a „Samuel Langhorne Clemens”. Kojarzy się to może czytelnikowi z pytaniem otwartym u Moore’a? Słusznie.

Aby wyjaśnić tę różnicę, Frege postulował dwa aspekty roli semantycznej wyrażen³⁰:

- To, do jakich przedmiotów w świecie wyrażenie się **odnosi** (odniesienie Frege nazywał *Bedeutung*). Terminy jednostkowe odnoszą się do przedmiotów, które nazywają.

30 Właściwie to postulował tylko jeden wymiar w pracach wczesnych, takich jak *Begriffsschrift*, a pojęcie sensu wprowadził nieco później, w pracy *Über Sinn und Bedeutung*, pochodzącej z roku 1892.

Predykaty odnoszą się do funkcji, które przypisują przedmiotom wartości logiczne (na przykład predykat „wiertarka” odnosi się do funkcji, która przypisuje Prawdę każdej wiertarce i Fałsz każdej niewiertarce). Zdania odnoszą się do wartości logicznych: Prawdy lub Fałszu (Frege z powodów, których nie będziemy rozważać, uważał, że Prawda i Fałsz – wartości logiczne – są przedmiotami, do których zdania mogą się odnosić)³¹.

- **Sens (Sinn)** wyrażień: to, w jaki sposób odniesienie wyrażenia jest nam dane. Frege nie wyjaśnił nadzwyczaj klarownie, czym takie sensory są, ale wiadomo, że jeżeli możemy rozumieć dwa wyrazy i nie zdawać sobie sprawy, że odnoszą się do tego samego, to mają one różne sensory. Sensy zdań twierdzących nazywają się **MYŚLAMI**. Jeżeli można dwa zdania rozumieć, zajmując rozbieżne względem nich postawy (takie jak uznawanie), to zdania te wyrażają różne myśli.

Wedle Fregego zdania [F1] i [F2] różnią się, ponieważ wyrażają różne myśli. A wyrażają różne myśli, ponieważ „Mark Twain” ma inny sens niż „Samuel Langhorne Clemens”. Później teorie przypisujące sens nazwom własnym znajdują się pod obstrzałem – por. rozdz. 36, a teorie je krytykujące będą miały znaczenie dla filozofii umysłu; na to trzeba będzie jednak poczekać jakieś 50 lat.

Frege wykorzystał pojęcie sensu do paru innych rzeczy, wymieniając szybko tylko dwie. Po pierwsze, do wyjaśnienia zrozumiałości zdań z nazwami pustymi. Zdania takie jak „Kapitan Bomba dobrze strzela” nie mają, jego zdaniem, odniesienia, nie są prawdziwe lub fałszywe, bo mają składową, która nie ma odniesienia. A odniesienie całości, zdaniem Fregego, powinno być funkcją odniesienia składowych i sposobu, w jaki zdanie jest zbudowane – zasada ta nazywa się **ZASADĄ KOMPOZYCYJALNOŚCI ODNIESIENIA**. Zdanie to jest jednak w jakimś sensie zrozumiałe, ponieważ posiada sens.

³¹ Dla podkreślenia, że mówimy o wartościach logicznych jako przedmiotach Fregeowskich, piszemy **Prawda** i **Fałsz** od wielkiej litery.

Frege: sens i odniesienie

Po drugie, do wyjaśnienia różnic w odniesieniu (= wartości logicznej) zdań takich, jak:

- | | |
|-------|---|
| [F1'] | Zenobia wie, że Mark Twain to Mark Twain. |
| [F2'] | Zenobia wie, że Mark Twain to Samuel Langhorne Clemens. |

Jako że wygląda na to, że wszystkie składowe mają to samo odniesienie, oba te zdania powinny mieć tę samą wartość logiczną, jeżeli zaufać zasadzie kompozycjonalności odniesienia, która powiada, że odniesienie każdego wyrażenia złożonego jest wyznaczone przez odniesienie wyrażen składowych oraz sposób zbudowania tego wyrażenia. W rzeczywistości jednak, powiada Frege, w sytuacji, w której przypisujemy komuś zajmowanie jakiejś postawy względem jakiegoś zdania, to zdanie to nie pełni już swojej zwykłej codziennej funkcji wyrażania myśli. Raczej zachowuje się jak nazwa i odnosi się do swojego własnego sensu. Tak więc [F1'] stwierdza pewną relację pomiędzy Zenobią a sensem zdania [F1], a [F2'] stwierdza pewną relację pomiędzy Zenobią a sensem zdania [F2]. Skoro jednak myśli wyrażane przez [F1] i [F2] różnią się od siebie, to nie ma też problemu z sytuacją, w której Zenobia pozostaje w rozważanej relacji tylko do jednego z nich.

Teoria Fregego stanowi pewien postęp w stosunku do teorii psychologicznych: obiektywizuje znaczenia i nieco lepiej wyjaśnia fakt komunikacji między ludźmi. Jeżeli jednak się nad tym zastanowić, nie radzi sobie najlepiej z opisanym na początku problemem z prywatnością. Tak samo jak nie mogę nikomu pokazać idei znajdujących się w mojej głowie (jeżeli rzeczywiście mam jakieś idee w głowie), tak też i nikomu nie mogę wskazać palcem sensu, jaki wiąże z wyrazem „wertarka”, i zwolennik takiej teorii stoi przed wyzwaniem wyjaśnienia, w jaki sposób ludzie mogą uzyskać dostęp poznawczy do takich obiektów jak sensy i nauczyć się systematycznie wiązać te same sensy z tymi samymi wyrażeniami. W każdym razie przejdźmy do Russella i jego reakcji na podstawy arytmetyki skonstruowane przez Fregego.



TEKSTY POLECANE

1. *The Search for Certainty: a Philosophical Account of Foundations of Mathematics*, autorstwa Marcusa Giaquinto, zawiera bardzo przystępnie omówioną historię rozwoju podstaw matematyki i wiążących się z tą dziedziną problemów filozoficznych.
2. Sporo pism Fregego znaleźć można w języku polskim w tomie *Pisma semantyczne*, przetłumaczonym przez Bogusława Wolniewicza (Warszawa 1977).
3. Warto zajrzeć do *An Introduction to the Philosophy of Language* autorstwa Michaela Morrisa (Cambridge 2007).
4. Jednym z najklarowniejszych wprowadzeń do filozofii języka jest *Philosophy of Language. A Contemporary Introduction* Lycana (Routledge 2008). Jakości książki dowodzi fakt, że Lycan fajnie wygląda z wąsem: <http://www.unc.edu/~ujanel/>.



FREGE O WIEDZY MATEMATYCZNEJ

1. Cel Fregego: wykazanie logicznej natury arytmetyki dla pokazania jej analityczności (wbrew Kantowi), pewności i ontologicznej akceptowalności.
2. *Begriffsschrift* (1879): konstrukcja logiki matematycznej pierwszego i drugiego rzędu.
3. *Die Grundlagen der Arithmetik* (1884) i *Grundgesetze der Arithmetik* (1893, 1903): logiczna rekonstrukcja arytmetyki.
4. Rozróżnienie pojęć pierwszego i drugiego rzędu. Liczebność to pojęcie drugiego rzędu („dziesięć” przysługuje pojęciu palców, nie samym palcom).
5. Ekstensja: zbiór odpowiadający pojęciu. Każde pojęcie ma ekstensję (pełne wyróżnianie).
6. Liczba n to zbiór wszystkich pojęć, które stosują się do dokładnie n przedmiotów.
7. Dowodzenie nieskończoności liczb: zaczynamy od pojęcia pustego, wrzucamy jego liczbę 0 do dziedziny, bierzemy pojęcie bycia liczbą 0, wrzucamy jego liczbę do dziedziny itd.
8. Kluczowa rola w systemie – Podstawowe Prawo Numer Pięć (PPV): ekstensja jednego pojęcia jest identyczna z ekstensją drugiego wtedy i tylko wtedy, gdy te pojęcia stosują się do tych samych przedmiotów.

Frege: sens i odniesienie

9. PPV pociąga za sobą pełne wyróżnianie.
10. PPV pociąga też Zasadę Hume'a: liczby dwóch pojęć są identyczne wtedy i tylko wtedy, gdy pojęcia te są równoliczne (definicja równoliczności nie wymaga odniesienia do liczb).
11. Zasada Hume'a pociąga za sobą aksjomaty arytmetyki Peana (drugiego, a zatem i pierwszego rzędu).

FREGE: SENS I ODNIESIENIE



1. Przeciwnik: teorie psychologiczne, które mówią o ideach.
2. Teorie psychologiczne nie wyjaśniają możliwości komunikacji (skąd wiadomo, jaką ideę wiążę ze słowem „wierarka“?).
3. Zdaniem Fregego teorie psychologiczne nie wyjaśniają tego, że różni ludzie mają różne idee, ale przekazują to samo znaczenie. Nie wyjaśniają też tego, że język jest „nakierowany na świat”.
4. Rola semantyczna wyrażeń nie może jednak sprowadzać się tylko do odniesienia: wskazuje na to zagadka informatywnych identyczności. „Mark Twain to Samuel Langhorne Clemens” byłoby trywialne.
5. Stąd rozróżnienie na odniesienie (przedmioty, funkcje, wartości logiczne) i sens (sposób dania przedmiotu wyjaśniający informatywność identyczności).
6. Wiąże się z tym pogląd, że nazwy własne mają sens – okaże się on ważny później.
7. Zasada kompozycyjności odniesienia: odniesienie całego wyrażenia jest funkcją odniesienia poszczególnych jego składników.
8. Rozróżnienie sensu i odniesienia pomaga pogodzić kompozycyjność ze zrozumiałością zdań z nazwami pustymi: nie mają one odniesienia, ale mają sens.
9. Rozróżnienie sensu i odniesienia pomaga pogodzić różne postawy względem zdań o tym samym odniesieniu: w tych kontekstach zdania odnoszą się do swoich własnych sensów.



KIERUNKOWSKAZ

Za chwilę wrócimy do problemów z podstawami arytmetyki u Fregego: do paradoksu Russella i jego teorii typów. Zanim to zrobimy, przyjrzymy się, jak wyglądało jedno z ważniejszych twierdzeń matematycznych XIX wieku, przez analogię do którego Russell skonstruował swój paradoks. Gdy omówimy teorię typów Russella (a właściwie pewną uwpółcześnioną jej wersję), przejdziemy do jego teorii deskrypcji.

15. Twierdzenie Cantora

Zacznijmy od czegoś, co wygląda jak niewinne twierdzenie teorii mnogości. Zostało ono dowiedzione jeszcze przed Russellem przez dosyć czadowego matematyka GEORGA CANTORA. Twierdzenie to głosi, że **ZBIÓR POTĘGOWY** danego zbioru (to jest **ZBIÓR WSZYSTKICH** jego **PODZBIORÓW**) zawsze ma więcej elementów niż zbiór, którego jest **ZBIOREM POTĘGOWYM**. Co to znaczy?

Weźmy jakiś zbiór składający się z trzech elementów:

$$\{a, b, c\}$$

Ile ma on podzbiorów, czyli takich zbiorów A , o których jest prawdą, że jeżeli coś należy do A , to należy też do $\{a, b, c\}$?

$$A \subseteq \{a, b, c\} \equiv \forall x [x \in A \rightarrow x \in \{a, b, c\}]$$

Czytamy: A jest podzbiorem zbioru $\{a, b, c\}$ wtedy i tylko wtedy, gdy dla każdego x , jeżeli x jest elementem A , to x jest też elementem zbioru $\{a, b, c\}$.

Twierdzenie Cantora

- Po pierwsze, dosyć trywialnie podzbiorem każdego zbioru jest zbiór pusty ($\emptyset$), który nie ma żadnych elementów. Jest o nim bowiem prawdą, że jeżeli coś jest jego elementem, to jest też elementem dowolnego innego zbioru, bo implikacja *jeżeli p , to q* wedle standardowego matematycznego rozumienia jest prawdziwa, gdy tylko fałszywy jest poprzednik.
- Po drugie, podzbiory naszego zbioru są wszystkie zbiory jednoelementowe złożone z elementów naszego zbioru:

$$\{a\}, \{b\}, \{c\}$$

- Następnie mamy podzbiory dwuelementowe:

$$\{a, b\}, \{a, c\}, \{b, c\}$$

- W końcu, każdy zbiór jest swoim własnym podzbiorem:

$$\{a, b, c\}$$

Zbierzmy te wszystkie podzbiory do jednego zbioru:

$$\{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\}$$

To właśnie jest zbiór potęgowy zbioru $\{a, b, c\}$. Ogólnie: zbiorem potęgowym dowolnego zbioru B jest zbiór wszystkich jego podzbiórów, oznaczany $\mathcal{P}(B)$.

Twierdzenie Cantora powiada, że dla dowolnego zbioru B , $\mathcal{P}(B)$ ma więcej elementów niż sam zbiór B . Aby tego dowieść, wystarczy dowieść, że nie istnieje funkcja, która każdemu elementowi zbioru B przyporządkowuje jakiś element zbioru $\mathcal{P}(B)$ w taki sposób, że każdy podzbiór zbioru B (czyli każdy element zbioru $\mathcal{P}(B)$) jest przyporządkowany jakiemuś elementowi zbioru B . Gdyby bowiem $\mathcal{P}(B)$ nie miał więcej elementów niż B , to

dałoby się każdemu elementowi $\wp(B)$ przypisać osobny element zbioru B .

O takiej funkcji mówilibyśmy, że jest funkcją ze zbioru B na zbiór $\wp(B)$. Aby zaś dowieść, że nie istnieje taka funkcja, wystarczy pokazać, że jeżeli przyjmiemy istnienie takiej funkcji, otrzymamy sprzeczność.

Założmy więc, że istnieje funkcja $f: B \rightarrow \wp(B)$. To tylko szpanerski sposób zapisywania warunku, o którym była mowa powyżej. To znaczy, że dla każdego $x \in B$ istnieje $A \in \wp(B)$ takie, że $f(x) = A$ (funkcja jest określona na całym zbiorze B), i że dla każdego $A \in \wp(B)$ istnieje $x \in B$ takie, że $f(x) = A$ (każdy element zbioru $\wp(B)$ jest czemuś przypisany). Formalnie:

$$\forall x \in B \exists A \in \wp(B) f(x) = A \quad (1)$$

$$\forall A \in \wp(B) \exists x \in B f(x) = A \quad (2)$$

Możemy teraz zdefiniować zbiór tych elementów zbioru B , które nie należą do zbiorów im przypisywanych:

$$\{x \in B \mid x \notin f(x)\} = C \quad (3)$$

To znaczy: żeby sprawdzić, czy jakiś element x zbioru B należy do zbioru C , najpierw sprawdzamy, jaki zbiór to $f(x)$. Potem sprawdzamy, czy x należy do zbioru $f(x)$. Jeżeli nie, to wrzucamy go do C . Innymi słowy:

$$x \in C \equiv x \notin f(x) \quad (4)$$

Teraz zauważmy, że C jest podzbiorem zbioru B , a więc elementem zbioru $\wp(B)$. To znaczy, że stosuje się do niego warunek (2). To znaczy:

$$\exists x \in B f(x) = C \quad (5)$$

Twierdzenie Cantora

Nazwijmy takie x jakąś konkretną nazwą, powiedzmy „futrasty”. Dostajemy:

$$f(\text{futrasty}) = C \quad (6)$$

Nasz „futrasty” jest jednak elementem B i stosuje się do niego (4):

$$\text{futrasty} \in C \equiv \text{futrasty} \notin f(\text{futrasty}) \quad (7)$$

Problem jednak polega teraz na tym, że (6) identyfikuje f (futrasty) właśnie z C . Otrzymujemy zatem sprzeczność:

$$\text{futrasty} \in C \equiv \text{futrasty} \notin C \quad (8)$$

A zatem funkcja, o której mowa, nie istnieje i zbiór potęgowy dowolnego zbioru zawsze ma więcej elementów niż sam ten zbiór. (Wiemy też, że funkcja w drugą stronę – z danego zbioru w jego zbiór potęgowy – istnieje. Każdemu bowiem przedmiotowi x w zbiorze B odpowiada przynajmniej jeden element zbioru $\mathcal{P}(B)$, mianowicie $\{x\}$, zbiór jednoelementowy, którego jedynym elementem jest ten przedmiot).

Twierdzenie Cantora ma co najmniej dwie interesujące nas konsekwencje.

- Po pierwsze, wynika z niego, że **nie istnieje największa nieskończona liczba**: dla dowolnie dużego nieskończonego zbioru istnieje zbiór potęgowy tego zbioru, który posiada jeszcze więcej elementów. Na przykład: rozmiar zbioru liczb naturalnych jest najmniejszą liczbą nieskończoną (mówimy, że zbiór o takiej ilości elementów jest **przeliczalny**); jego zbiór potęgowy (czyli zbiór zbiorów liczb naturalnych) posiada więcej elementów – tyle samo, ile liczb rzeczywistych. Zbiór liczb rzeczywistych z kolei ma jeszcze więcej podzbiorów niż liczb rzeczywistych itd.

- Po drugie, wynika z niego, że **nie istnieje zbiór uniwersalny**: zbiór zawierający wszystkie zbiory. Gdyby bowiem istniał taki zbiór, jego elementami byłyby wszystkie jego podzbiory. To by znaczyło, że musiałby być przynajmniej tak duży jak własny zbiór potęgowy. Żaden jednak zbiór nie ma takiej własności.

16. Paradoks Russella, teoria typów

Dowodząc twierdzenia Cantora, założyliśmy istnienie funkcji spełniającej pewne warunki, po czym rozpatrzyliśmy zbiór zdefiniowany w sposób następujący:

$$x \in \mathcal{C} \equiv x \notin f(x)$$

Wyprowadzenie sprzeczności dowodziło nieistnienia takiej funkcji. Co by się stało, gdybyśmy definiując inny zbiór, $\mathcal{R}$, zamiast f posłużyli się funkcją, co do istnienia której nie mamy wątpliwości, na przykład funkcją identyczności, która każdemu przedmiotowi przypisuje sam ten przedmiot, $f(x) = x$? W takim razie zamiast $f(x)$ możemy po prostu pisać x :

$$x \in \mathcal{R} \equiv x \notin x \quad (9)$$

Definicja taka szybko prowadzi do sprzeczności, gdy zapytamy, czy $\mathcal{R}$ jest swoim własnym elementem:

$$\mathcal{R} \in \mathcal{R} \equiv \mathcal{R} \notin \mathcal{R} \quad (10)$$

Co poszło nie tak? Przecież nie zakładaliśmy po drodze istnienia żadnych dyskusyjnych funkcji!

Paradoks Russella, teoria typów

Problem pojawił się jeszcze wcześniej: w momencie, w którym założyliśmy, że z samego faktu, iż możemy napisać niby-definiujący warunek (9), wynika istnienie zbioru $\mathcal{R}$. Z tego samego też powodu paradoks ten (zwany paradoksem Russella) stanowił problem dla teorii Fregego. U Fregego istotną konsekwencją PODSTAWOWEGO PRAWA NUMER PIĘĆ **było pełne wyróżnianie**: teza, że każde pojęcie ma ekstensję. Gdyby tak było, pojęcie „niebycie swoim własnym elementem” również miałoby swoją własną ekstensję: zbiór wszystkich i tylko tych obiektów, które nie są swoimi własnymi elementami. Byłby to właśnie zbiór Russella, którego istnienie prowadziłoby do sprzeczności.

Russell napisał list do Fregego w roku 1902, tuż przed wydaniem przez Fregego drugiego tomu *Grundgesetze*. Wyzwanie dla Fregego było całkiem poważne: jak w sposób przekonujący zastąpić PPV inną zasadą, która z jednej strony uniknie paradoksu, z drugiej zaś strony dalej będzie wystarczająco silna, by pozwolić na zrekonstruowanie arytmetyki. Było za późno na zmodyfikowanie samej książki. Zamiast tego Frege dołączył do niej dodatek, w którym informował o problemie i sugerował pewne rozwiązanie.

Na czym ono polegało, nie będziemy się zastanawiać; wystarczy powiedzieć, że niedługo później polski logik Stanisław Leśniewski (a potem i szereg innych ludzi) wykazał, że rozwiązanie to nie pozwala na uniknięcie problemów. Leśniewski pokazał, że z tego rozwiązania wynika, że istnieją co najwyżej dwa przedmioty, co stanowi pewien problem dla rekonstrukcji arytmetyki, w której docelowo należałoby uznać istnienie więcej niż dwóch liczb. Frege uznał swoją porażkę i zajął się innymi sprawami. Kwestią naprawienia systemu zajęli się Russell z Whiteheadem w monumentalnym dziele *Principia mathematica* (1910–1913).

Diagnoza Russella (skonstruowana razem z francuskim matematykiem Henrim Poincarém) była stosunkowo prosta: definicja klasy Russella jest kolistą – zbiór ten jest zdefiniowany za pomocą kwantyfikacji po pewnych przedmiotach, co do których zakłada się, że sam zbiór Russella do nich należy.

Jak zaimplementować tę intuicję w logicznej rekonstrukcji arytmetyki, która uniknęłaby sprzeczności? Otóż zamiast postulować istnienie jednej dużej dziedziny, do której miałyby należeć zarazem klasa Russella, jak i indywidua, należy wedle Russella rozpatrywać obiekty w pewnej hierarchii typów. Russell, ściśle rzecz biorąc, uważał, że zbiory nie istnieją, i tłumaczył zdania o zbiorach na zdania o funkcjach; my dla prostoty prezentacji nie będziemy się tym przejmować i będziemy od razu mówić o zbiorach.

- Najpierw mamy typ 0 – typ indywiduów.
- Potem mamy typ 1 – typ zbiorów, których elementami mogą być tylko indywidua – elementy typu 0.
- Kontynuujemy budowę hierarchii, na poziomie każdego typu n umieszczając zbiory, których elementami mogą być tylko obiekty jeden schodek niżej – przedmioty z typu $n-1$.

Następnie modyfikujemy język teorii: zamiast jednego typu zmiennych teraz wprowadzamy osobno zmienne dla każdego typu: zmienne $x_0^0, x_1^0, x_2^0, \dots$ przebiegające typ 0, zmienne $x_0^1, x_1^1, x_2^1, \dots$ przebiegające typ 1 itd. Dodajemy też regułę składni, mówiącą, że „ $\in$ ” może mieć po bokach tylko zmienne różniące się o jeden typ: $x^i \in x^{i+1}$.

W takiej sytuacji paradoks Russella nie powstaje, bo nie istnieje zbiór wszystkich przedmiotów, które nie są swoim własnym elementem. Nie można w tym języku nawet takiego zbioru poprawnie zdefiniować.

Zamiast tego na każdym typie (oprócz zerowego) otrzymujemy co najwyżej małe, grzeczne zbiorki Russella. Na poziomie typu pierwszego dostajemy zbiór wszystkich przedmiotów typu 0, które nie są swoimi własnymi elementami: jest to po prostu zbiór wszystkich przedmiotów typu 0. Analogicznie rzecz się ma dla wyższych typów. Mówiąc ściśle, warunku niebycia swoim własnym elementem nie da się wewnątrz teorii nawet sformułować, bo będzie niegramatyczny.

Paradoks Russella, teoria typów

Dobrze, załóżmy, że wiemy, iż teoria unika paradoksu. Czy jest jednak wystarczająco silna, by pozwolić na rekonstrukcję arytmetyki? Zasadniczo strategia konstruowania definicji jest taka sama jak u Fregego: na danym typie k liczba n to zbiór zbiorów n -elementowych typu $k-1$. Jeżeli posłużyć się tego typu definicjami, rzeczywiście można otrzymać arytmetykę, ale niestety odbywa się to pewnym kosztem.

- Po drodze Russell z Whiteheadem posługują się pewnymi definicjami, które wyglądają na koliste: aby móc się nimi posłużyć, przyjmują aksjomat, że takie definicje definiują zbiory, które można zdefiniować bez kolistości, i nie jest jasne, czy taki aksjomat jest czysto logicznym twierdzeniem³².
- Pamiętamy, jak Frege dowodził nieskończoności liczb? Brał puste pojęcie, generował jego liczbę, wrzucał do dziedziny, brał pojęcie tej liczby, generował jego liczbę, wrzucał do dziedziny itd. Teraz to nie jest możliwe: liczba danego pojęcia będzie dwa typy wyżej, a typy się nie mieszają. Jeżeli nasz typ 0 zawiera skończoną liczbę elementów, każdy typ powyżej będzie zawierał tylko skończoną ilość liczb. Dlatego też Russell był zmuszony przyjąć **aksjomat nieskończoności**: musiał założyć, że istnieje nieskończona ilość indywiduów typu 0. W takiej jednak sytuacji pojawia się pytanie, do jakiego stopnia teoria, która postuluje istnienie nieskończoności indywiduów, dalej kwalifikuje się jako teoria czysto logiczna.
- Wprowadzenie wielości typów zaowocowało też rozmnożeniem liczb: zamiast jednej liczby zero mamy jedną liczbę zero na typie 2, jedną na typie 3, jedną na typie 4 itd. Zachodzić może więc podejrzenie, że teoria typów nie dostarcza jednolitej arytmetyki, takiej, jaką znamy.

32 Sprawa **aksjomatu redukowalności** niestety jest trochę techniczna i ze względu na styl tej książki więcej szczegółów omówić nie możemy.

- Inny problem to fakt, że teoria typów jest **bardzo** restrykcyjna. Wyklucza samą sensowność warunku nienależenia do samego siebie, wyklucza zbiory mieszane składające się z przedmiotów **różnych** niższych typów. Nie jest jasne, czy wprowadzając wszystkie ograniczenia mające na celu uniknięcie paradoksu Russella, nie wylano dziecka z kąpielą.
- Kolejna trudność pojawia się w momencie, gdy pytamy o sensowność teorii typów. Okazuje się, że samej hierarchii typów nie można opisać wewnątrz hierarchii typów (bo wewnątrz hierarchii nie możemy powiedzieć niczego o wszystkich jej poziomach naraz). Pojawia się pytanie, czy ograniczenia sensowności narzucone przez teorię typów nie powinny się również stosować do samej metody jej wprowadzania.

Mogłoby się więc wydawać, że projekt logicystyczny sprowadzenia matematyki do logiki, pomimo uratowania go od paradoksu Russella, leży w gruzach. Okazuje się jednak, że sprawa nie jest oczywista, bo od lat 80. projekt ożywa pod nową nazwą **NEOLOGICYZMU**, z nowym arsenałem środków i argumentów. Wrócimy do tej kwestii w rozdziale 33.

Kolejnym skutkiem ubocznym działań Russella i Whiteheada jest nurt **PREDYKATYWIZMU** w filozofii matematyki. Jedną z diagnoz paradoksów teoriomnogościowych (i nie tylko), zasugerowaną przez francuskiego matematyka Henriego Poincarégo oraz przez Russella, jest taka, że rozumowania paradoksalne posługują się definicjami, które zakładają, że obiekt definiowany znajduje się pośród obiektów, do których odnosi się definicja (które „znajdują się w zasięgu kwantyfikacji”). Definicje takie nazywają się definicjami niepredykatywnymi.

Na przykład gdy (1) powiadam, że dany zbiór x jest elementem zbioru Russella wtedy i tylko wtedy, gdy nie należy sam do siebie, a potem (2) podstawiam nazwę zbioru Russella za x , wnioskując, że zbiór Russella jest elementem zbioru Russella wtedy i tylko wtedy, gdy nie należy sam do siebie, to zakładam, że nazwę zbioru Russella prawomocnie mogę pod-

Russell i teoria deskrypcji

stawić za zmienną x w mojej definicji zbioru Russella. W tym sensie definicja zbioru Russella konstruowana w (1) jest niepredykatywna.

Predykatywiści skupili się na wymogu predykatywności definicji, ignorując większość szczegółów konstrukcji samej teorii typów. Badają głównie relacje pomiędzy różnymi technicznymi dookreśleniami pojęcia predykatywności a mocą teorii matematycznych, które można w rozważanych systemach uzyskać, nie posługując się definicjami niepredykatywnymi. O ile otrzymywane systemy nie zawsze są tak przydatne i tak silne jak matematyka klasyczna, o tyle podejmowane problemy są fascynujące (jeżeli ktoś jest wystarczająco dziwny).

Pewną trudność stanowi fakt, że definicje niepredykatywne często są niewinne i pożyteczne. Na przykład gdy wchodzę do pubu i mówię do barmana: „Poproszę najtańsze z waszych piw”, to moje określenie, czego chcę, jest niepredykatywne: najtańsze piwo znajduje się pośród piw, z których ma zostać wybrane. Tylko bardzo niezwykle barman byłby skłonny powiedzieć, że mojego opisu nie zrozumiał, bo nie był on predykatywny.

17. Russell i teoria deskrypcji

Bertrand Russell i Moore studiowali razem filozofię, pozostawali pod wzajemnym wpływem i uprawiali filozofię w podobnym duchu. O ile Moore atakował idealizm, subiektywizm etyczny i sceptycyzm, o tyle Russell skupił się na kwestiach nieco innych. Sporo czasu spędził, próbując zrozumieć naturę matematyki i starając się wykazać jej analityczność. Narzędzie jednak, którym się posługiwał przy tej okazji – logikę formal-

ną – starał się też stosować do bardziej ogólnych rozważań filozoficznych.

W szczególności Russell zasłynął ze swojej krytyki pewnego typu realizmu metafizycznego. Aby wyjaśnić, na czym on polegał, zacznijmy od rozważenia następujących zdań:

- | | |
|----------|--|
| ZDANIE 1 | Stół jest brudny. |
| ZDANIE 2 | Żyrafa sącząca martini szczerzy się do mnie. |
| ZDANIE 3 | Żyrafa sącząca martini nie nuci. |
| ZDANIE 4 | Żyrafa sącząca martini nie istnieje. |



Na pierwszy rzut oka zdanie pierwsze ma dość trywialną strukturę. W kontekście, w którym jest wypowiedzane, mowa o pewnym konkretnym stole, do którego odnosi się podmiot ZDANIA 1 „stół”, któremu przypisujemy pewną własność wyrażaną przez predykat „jest brudny”. Wygląda więc na to, że zdanie to ma strukturę podmiotowo-orzeczeniową. Tego typu podmioty w takich zdaniach nazywają się **DESKRYPCJAMI OKREŚLONYMI**, ponieważ posługują się opisem jakościowym (deskrypcją) dla wyznaczenia dokładnie jednego (określonego) przedmiotu. Stąd teoria Russella, o której będzie mowa, nazywa się **TEORIA DESKRYPCJI OKREŚLONYCH**.

Gramatycznie rzecz biorąc, pozostałe zdania (na pierwszy rzut oka) nie różnią się drastycznie od pierwszego: „żyrafa sącząca martini” jest podmiotem we wszystkich. W ZDANIU 2

Russell i teoria deskrypcji

przypisujemy jej pewną własność, w ZDANIU 3 odmawiamy jej pewnej własności, w ZDANIU 4 odmawiamy jej istnienia. Cały czas jednak mamy do czynienia z podmiotem i (mniej lub bardziej złożonym) orzeczeniem. A przynajmniej na to wygląda.

Spójrzmy teraz na następujący argument:

ZAŁOŻENIE 1	Zdania 1–4 są zdaniami podmiotowo-orzeczeniowymi. Zdania 3 i 4 są prawdziwe.
ZAŁOŻENIE 2	Zdanie podmiotowo-orzeczeniowe jest prawdziwe wtedy i tylko wtedy, gdy przedmiot, do którego odnosi się podmiot, posiada (nie posiada) własność, którą przypisuje (odmawia) mu orzeczenie.
WNIOSEK 1	Przedmiot, do którego odnosi się „żyrafa sącząca martini”, nie posiada własności nucenia i nie posiada własności istnienia.
WNIOSEK 2	Jest coś takiego jak przedmiot, do którego odnosi się „żyrafa sącząca martini”.

Rozumowanie tego typu prowadzi do uznania, że obiekty nieistniejące, jak żyrafy sączące martini, kwadratowe koła, złote góry itp., w jakimś sensie istnieją. Chcąc uniknąć paradoksalnego stwierdzenia, mówi się niekiedy o przedmiotach, które „są”, lecz które „nie istnieją”. Zwykle zarzuca się zwolennikom Meinonga³³, że ich teoria wymaga przeprowadzenia odróżnienia pomiędzy *byciem* oraz *istnieniem*, co jest kontrowersyjne samo w sobie i jak wielu argumentowało, może prowadzić do licznych problemów.

Pogląd, jakoby istniały różne typy istnienia, niektóre przysługujące takim obiektom, jak kwadratowe koła czy złote góry, żywił początkowo (w *The Principles of Mathematics*, 1903) Russell, który głosił, że każdy przedmiot myśli istnieje. Dużo bardziej

33 Alexius Meinong (1853–1920) był austriackim filozofem, znanym ze swojej teorii przedmiotów, w której systematycznie rozważał różne typy istnienia.

zdecydowanie głosił go Alexius Meinong, któremu pogląd ten przypisywał (krytykując go) Russell już w *On Denoting* (1905).

Istnieją powody, dla których ten typ realizmu przedmiotów myśli przestał się Russellowi podobać. Po pierwsze, samo twierdzenie, jakoby nieistniejące przedmioty w jakimś sensie istniały, brzmi nieco sprzecznie. Co więcej, Meinong dopuszczał również (w jakimś sensie) istnienie przedmiotów sprzecznych, takich jak kwadratowe koła, pomimo tego, że jednocześnie zgadzał się, że są one obiektami sprzecznymi. Po drugie, Russell uważał, że wyprowadzanie silnych twierdzeń o tym, co istnieje, jedynie na podstawie zjawisk językowych jest podejrzane. W *Introduction to Mathematical Philosophy* powiada:

» Wydaje mi się, że w tego typu [realistycznych, R.U.] teoriach brakuje poczucia rzeczywistości, które powinno zostać zachowane nawet w najbardziej abstrakcyjnych badaniach. Logika, uważam, nie bardziej może dopuścić istnienie jednorożców niż zoologia, bo logika zajmuje się światem rzeczywistym tak samo jak zoologia. [...] Powiedzieć, że jednorożce mają istnienie w heraldyce, literaturze czy wyobraźni, to tylko żalosny unik. To, co istnieje w heraldyce, nie jest zwierzęciem z krwi i kości poruszającym się i oddychającym z własnej woli. Istnieje tylko obraz lub opis słowny.

Co jednak z rozważanym przez nas argumentem? Jeżeli znamy poprawny logicznie argument, którego wniosek odrzucamy, dobrze byłoby, gdybyśmy byli w stanie wyjaśnić, którą z przesłanek odrzucamy i dlaczego. Russell odrzucił pierwszą część ZAŁOŻENIA 1: mimo że zdania takie gramatycznie wyglądają na zdania podmiotowo-orzecznikowe, uznał, że ich **FORMA LOGICZNA** wygląda nieco inaczej. To znaczy w swoim preferowanym języku formalnym podał warunki prawdziwości takich zdań, przy których te zdania, które intuicyjnie uznajemy za prawdziwe, są prawdziwe, ale ich prawdziwość nie wymaga istnienia sączących martini żyraf czy kwadratowych kół.

Russell i teoria deskrypcji

Odkładając na razie kwestię języka formalnego na bok, przyjrzymy się istocie tego ruchu. Otóż wedle Russella ZDANIE 1 głosi raczej:

Istnieje dokładnie jeden stół (o którym mowa), i ten stół jest brudny.

Analogicznie ZDANIE 2 powiada:

Istnieje dokładnie jedna żyrafa sącząca martini (o której mowa) i ta żyrafa szczyrzy się do mnie.

Jako że pierwsza połowa tak rozumianego ZDANIA 2 jest fałszywa, całe to zdanie jest fałszywe, zgodnie z naszymi intuicjami (a przynajmniej moją intuicją), że nie szczyrzy się do nas (a przynajmniej do mnie) żadna żyrafa sącząca martini.

ZDANIE 3 stanowi przypadek nieco bardziej skomplikowany: jego gramatykę zinterpretować można na co najmniej dwa sposoby (pretensjonalna nazwa tego zjawiska: zdanie to jest **SYNTAKTYCZNIE AMBIWALENTNE**). Po pierwsze, można rozumieć je tak, że prawie całe zdanie jest w zasięgu negacji:

Nie jest tak, że [istnieje dokładnie jedna żyrafa sącząca martini (o której mowa), i ta żyrafa nuci].

W tej interpretacji to zdanie jest prawdziwe. W drugiej interpretacji negacja znajduje się już w zasięgu deskrypcji:

Istnieje dokładnie jedna żyrafa sącząca martini (o której mowa) i nie jest tak, że ta żyrafa nuci.

W tej interpretacji zdanie to jest fałszywe. Bardziej intuicyjna jest interpretacja pierwsza; ZAŁOŻENIE 2 przy jednoczesnym zachowaniu przekonania o prawdziwości zdania sugeruje interpretację pierwszą.

Analogicznie ZDANIE 4 w bardziej przekonującej interpretacji głosi:

Nie jest tak, że istnieje dokładnie jedna żyrafa sącząca martini (o której mowa) i ta żyrafa istnieje.

Co sprowadza się do:

Nie jest tak, że istnieje dokładnie jedna żyrafa sącząca martini (o której mowa).

Reprezentacje treści zdań rozważanych można podać w bardziej jednoznacznym języku rachunków predykatów. „ P jest A ”, gdzie P jest deskrypcją określoną, interpretujemy jako „Istnieje dokładnie jedno P i to P jest A ” i zapisujemy w sposób następujący:

$$\exists x[P(x) \wedge \forall y(P(y) \rightarrow x = y) \wedge A(x)]$$

Pierwszy człon mówi, że istnieje jakieś x , które ma własność P , drugi mówi, że cokolwiek ma własność P , jest x -em (a więc gwarantuje, że co najwyżej jeden przedmiot jest P), a trzeci orzeka A o tym przedmiocie. Zdanie przeczące „ P nie jest A ” można interpretować na dwa sposoby wyrażające wspomnianą już wieloznaczność (zależnie od tego, czy negowana jest sama własność, czy całe zdanie):

$$\neg \exists x[P(x) \wedge \forall y(P(y) \rightarrow x = y) \wedge A(x)]$$

$$\exists x[P(x) \wedge \forall y(P(y) \rightarrow x = y) \wedge \neg A(x)]$$

Pierwsza formuła zaprzecza, jakoby istniał dokładnie jeden przedmiot, który jest P i zarazem A . Druga mówi, że istnieje dokładnie jeden przedmiot, który jest P , ale nie posiada on własności A . Sprawy nieco się komplikują, gdy przyjrzymy się podobnym zdaniom zawierającym nazwy własne:

Russell i teoria deskrypcji

ZDANIE 5 Kapitan Bomba nie istnieje.

W przypadku ZDAŃ 1–4 mieliśmy do czynienia z podmiotami zdań, które wyrażały własności. Dość łatwo mogliśmy te zdania interpretować jako wymagające, by istniał dokładnie jeden przedmiot posiadający te właśnie własności. Na przykład ZDANIE 2 wymaga istnienia przedmiotu, który ma własność bycia żyrafą sączącą martini. Nie ma tu problemu z istnieniem takiej własności, nawet jeśli nie ma takiej żyrafy. W przypadku nazw własnych sprawa jest nieco mniej oczywista, bo nie wiadomo, jaką własnością się posłużyć. Na czym miałyby polegać własność bycia Kapitanem Bombą i czy istnieje taka własność, gdy nie ma Kapitana Bomby?

Russell jest zdania, że (prawie) każda nazwa, która wygląda w języku potocznym na własną, stowarzyszona jest z pewnym opisem jakościowym, który wyraża jej treść (wyrafinowana nazwa tego poglądu: **DESKRYPCYJNA TEORIA NAZW**, będziemy o niej jeszcze mówić dość szczegółowo). I rzeczywiście, jeżeli z nazwą „Kapitan Bomba” stowarzyszymy jakąś deskrypcję określoną („Pogromca kosmitów” itd.), to ZDANIE 5 można interpretować jako:

Nie istnieje dokładnie jeden pogromca kosmitów itd.

Należy podkreślić: wprowadzenie teorii deskrypcji dla zdań typu 1–4 to jedna sprawa, a uznanie deskrypcyjnej teorii nazw to sprawa osobna. W szczególności deskrypcyjna teoria nazw głosi, że to, co uważamy za nazwy własne, nie odnosi się bezpośrednio do przedmiotu, a jedynie czyni to za pośrednictwem pewnej treści opisowej, wyrażanej przez stowarzyszoną z tą nazwą deskrypcję. Teoria deskrypcji dla zdań typu 1–4 nie mówi nic o nazwach własnych.

Warto zauważyć, że teoria deskrypcji ma pewną konsekwencję, którą nie wszyscy uznają za radosną. Mianowicie, wszelkie proste zdania orzekające coś o pustym podmiocie okazują się

falszywe, tak więc fałszem jest wedle tej teorii zarówno *Sherlock zjadł predatora*, jak i *Sherlock palił fajkę*, pomimo naszych intuicji, że jedno z tych zdań jest prawdziwe.

Russell, mimo że traktował nazwy własne jak deskrypcje, dopuszczał pewne wyrażenia, które uważał za **NAZWY LOGICZ-NIE WŁAŚCIWE**. Ich funkcją – jego zdaniem – było bezpośrednie odnoszenie się do przedmiotu bez pośrednictwa deskrypcji, a ich desygnat (przedmiot, do którego się odnoszą) był ich znaczeniem. Jednym z takich wyrażen (przynajmniej w pewnej fazie rozwoju poglądów Russella) było „to”. I rzeczywiście trudno powiedzieć prawdę, pokazując na jakiś przedmiot i mówiąc „To nie istnieje”. Sprawę komplikuje również fakt, że jedyne przedmioty, jakie można zdaniem Russella nazywać za pomocą nazw logicznie właściwych, to własne myśli, doświadczenia zmysłowe itp., do których mamy bezpośredni dostęp poznawczy (zagadnienie jest zbyt szczegółowe, żeby się nim teraz szerzej zająć). Takie nazwy własne, jak „Kapitan Bomba” czy „Rafał Urbaniak” nie są, zdaniem Russella, nazwami logicznie właściwymi, ponieważ rozumiemy zdania, w których występują, nawet gdy nie jesteśmy pewni, czy się do czegoś odnoszą³⁴.

Innym przykładem zastosowania teorii deskrypcji jest podejście Russella do **ZAGADKI IDENTYCZNOŚCI**. Wspominaliśmy już o tym zjawisku, mówiąc o Fregem. Przypomnijmy, na czym polega zagadka. Załóżmy nawet, że „Scott” jest nazwą logicznie właściwą. Rozważmy zdania:

ZDANIE 7 Król chciał wiedzieć, czy autor *Waverleya* to Scott.

ZDANIE 8 Król chciał wiedzieć, czy Scott to Scott.

³⁴ „Kiedykolwiek można założyć, że podmiot gramatyczny zdania nie istnieje, bez sprawienia, że zdanie będzie pozbawione znaczenia, jest jasne, że podmiot gramatyczny nie jest logicznie nazwą właściwą... we wszystkich takich przypadkach zdanie musi być analizowalne tak, że podmiot gramatyczny znika”. Wersja angielska: “Whenever the grammatical subject of a proposition can be supposed not to exist without rendering the proposition meaningless it is plain that the grammatical subject is not a [logically] proper name... in all such cases the proposition must be capable of being so analyzed that what was the grammatical subject shall have disappeared” [*Principia Mathematica*].

Jeżeli jedyną funkcją semantyczną wyrażen „Scott” oraz „autor *Waverleya*” jest odnoszenie się do przedmiotów, to ich funkcja semantyczna jest taka sama, a więc można zastępować jedno drugim w wyrażeniach złożonych przy zachowaniu sensu całości. Widać jednak jasno, że ZDANIA 7 i 8 mają zdecydowanie odmienne sensy. Wyjaśnienie Russella polegało na wskazaniu, że „autor *Waverleya*” nie jest terminem jednostkowym, tylko deskrypcją określoną, więc forma logiczna tych zdań jest dużo bardziej złożona.

Deskrypcyjna teoria nazw przydać się może również w radzeniu sobie z zagadkami identyczności, jeżeli są one sformułowane za pomocą nazw własnych. Mogę na przykład nie wiedzieć, czy Mark Twain to Samuel Langhorne Clemens (przypomnę: tak, to ten sam człowiek). Jeżeli jedyną rolą semantyczną nazw jest odnoszenie się, trudno to wyjaśnić. Jeżeli natomiast przyjmujemy, że te dwie nazwy stowarzyszone są z dwiema różnymi deskrypcjami określonymi, które wyrażają ich znaczenie, możliwość tego typu wątplenia wydaje się całkowicie zrozumiała.

TEKSTY POLECANE

1. Szczegóły konstrukcji logiczystycznej Russella i Whiteheada można znaleźć w napisanym przez nich monumentalnym dziele *Principia Mathematica* (1910–1913). To jedno z największych dzieł logiki XX wieku potrzebowało od autorów pięćdziesięciu funtów dopłaty, żeby je wydano.
2. Czytelnik bardziej zainteresowany udziałem Russella w rozwoju filozofii matematyki może zajrzeć do *The Search for Certainty. A Philosophical Account of Foundations of Mathematics* autorstwa M. Giaquinto.
3. Przyjemnie czyta się popularyzatorską książkę *Tajemnica Alefów* napisaną przez Amira Aczela. Traktuje ona w szczególności o różnych rodzajach nieskończoności.
4. Klasycznym tekstem, w którym Russell rozwija swoją teorię deskrypcji, jest *On denoting*. Tekst wyszedł w *Mind* 14 (1905), ale jest też publicznie dostępny on-line. Ukazał się też w języku polskim w tomie *Logika i język* w edycji J. Pelca.



Wybrani dziadkowie filozofii analitycznej

5. Innym tekstem Russella wartym lektury jest *Introduction to Mathematical Philosophy*.
6. Więcej o nazwach własnych i nazwach logicznie własnych przeczytać można w *Knowledge by Acquaintance and Knowledge by Description* w Proceedings of the Aristotelian Society 11 (1910–1911).



TWIERDZENIE CANTORA

1. Zbiór potęgowy zbioru $B - \mathcal{P}(B)$ – to zbiór wszystkich podzbiorów B .
2. Twierdzenie Cantora: $\mathcal{P}(B)$ zawsze ma więcej elementów niż B .
3. Wniosek 1: istnieje nieskończoność rodzajów nieskończoności.
4. Wniosek 2: nie istnieje zbiór uniwersalny (zbiór wszystkiego).



PARADOKS RUSSELLA, TEORIA TYPÓW

1. Paradoks Russella jest analogiczny do twierdzenia Cantora, powstaje przez rozważenie zbioru przedmiotów, które nie są swoimi elementami, i zapytanie, czy jest on swoim elementem.
2. Russell napisał list do Fregego. Ten wprowadził poprawkę w dodatku do *Grundgesetze*, ale okazała się ona nieskuteczna.
3. Teoria typów zamiast jednej wielkiej dziedziny wprowadza hierarchizację: indywidua, zbiory indywiduów, zbiory zbiorów indywiduów, ...*
4. Teoria też wprowadza ograniczenia gramatyczne: definicja zbioru Russella staje się bezsensowna (bo po lewej od $\in$ może być tylko coś, co jest o jeden typ niżej niż to, co jest po prawej).
5. Liczby definiowane są podobnie jak u Fregego: zbiory zbiorów równolicznych.
6. Wprowadzenie hierarchii uniemożliwia dowód nieskończoności liczb: trzeba założyć **aksjomat nieskończoności** indywiduów.
7. Trzeba też założyć, że niektóre koliste definicje dają się zastąpić.
8. Wprowadzenie typów prowadzi do rozmnożenia liczb. Mamy kopie każdej liczby na niemal każdym typie.
9. Teoria typów jest bardzo restrykcyjna. Wyklucza zbiory mieszane i samą sensowność wielu intuicyjnie zrozumiałych wypowiedzi.
10. Sam opis teorii typów nie jest sensowny wedle teorii typów.

* Ponownie dla uproszczenia mówię o zbiorach, a nie – jak Russell – o funkcjach.

Russell i teoria deskrypcji

RUSSELL I TEORIA DESKRYPCJI



1. Cel polemiki: krytyka silnego realizmu metafizycznego postulującego, że dla dowolnego zdania podmiotowo-orzecznikowego istnieje obiekt, do którego odnosi się podmiot tego zdania, i wyjaśnienie, jak tego typu zdania funkcjonują w dyskursie bez potrzeby uznawania istnienia obiektów sprzecznych, nieistniejących itp.
2. Deskrypcje określone: wyrażenia, które za pomocą opisu jakościowego mają na celu wyznaczenie dokładnie jednego przedmiotu.
3. Argument krytykowany: każda deskrypcja może wystąpić w podmiocie prawdziwego zdania podmiotowo-orzecznikowego. Aby takie zdanie było prawdziwe, podmiot musi się do czegoś odnosić. Zatem każda deskrypcja odnosi się do czegoś.
4. Rozwiązanie: forma logiczna takich zdań nie jest podmiotowo-orzecznikowa. Stwierdzamy w takim zdaniu raczej, że istnieje dokładnie jeden przedmiot posiadający pewną własność, i wtedy coś o nim orzekamy.
5. To pozwala na trafną analizę prawdziwych zdań egzystencjalnych przeczących oraz zdań orzekających własności o niespełnionych deskrypcjach.
6. Teoria jest rozszerzana na nazwy własne (w potocznym sensie): daje to deskrypcyjną teorię nazw.
7. Teoria deskrypcji radzi sobie z zagadką identyczności niezależnie od tego, czy są one sformułowane za pomocą deskrypcji, czy nazw własnych.

KIERUNKOWSKAZ



Na jakiś czas odkładamy na bok filozofię matematyki oraz teorię odniesienia. Zajmiemy się teraz skrzyżowaniem metametafizyki z filozofią języka. Przyjrzymy się jednej z najsilniejszych w XX wieku prób odmówienia sensowności tradycyjnym rozważaniom filozoficznym sformułowanej przez Rudolfa Carnapa. Przy okazji będziemy zmierzali powoli w stronę probabilistycznych narzędzi formalnych, które pobieżnie omówimy później.

18. Filozofia a składnia

„Zasadą wszechrzeczy jest woda”, „Zasadą wszechrzeczy jest ogień”, „Istnieje tylko jedna zasada rzeczywistości”, „Rzeczywistość ma naturę mentalną”, „Wszystko, co istnieje, jest materialne”. Tego typu zdania, głoszone na prawo i lewo, często bez głębszego wyjaśnienia czy uzasadnienia, łatwo znaleźć w pismach Wielkich Filozofów. Przeciwno szarżowaniu z wielkimi twierdzeniami tego typu buntowali się wcześnie filozofowie analityczni. Początkowo nie podejmowali się oni systematycznego rozważenia znaczenia czy metodologii obrony (lub krytyki) owych wypowiedzi. Na takie rozważania nie trzeba było jednak długo czekać.

Zasadniczy problem z Wielkimi Tezami Filozoficznymi jest taki, że wygląda na to, iż nigdy nikomu nie udało się żadnej przekonująco obronić. Czy jest to istnienie Boga, czy istnienie wolnej woli, czy Ostateczna Natura Rzeczywistości (pisana wielkimi literami brzmi dostojnie), żadna z takich kwestii nie doczekała się ostatecznych ustaleń.

Z jednej strony można na to reagować **nieuzasadnionym optymizmem**, mówiąc, że chociaż na razie sprawa nie jest jasna, to wiemy więcej, niż wiedzieliśmy, i że postęp w filozofii w końcu umożliwi udzielenie odpowiedzi na martwiące nas kwestie. Pewną trudnością związaną z nieuzasadnionym optymizmem jest to, że jest on nieuzasadniony.

Inną możliwą reakcją jest konstruowanie wyrafinowanych teorii filozoficznych mających na celu wyjaśnienie nam, dlaczego te kwestie, chociaż w zasadzie merytoryczne, nie są dostępne naszemu poznaniu (*vide* Kant). Jednak takie wyrafinowane teorie są często bardziej niezrozumiałe i nieuzasadnione niż kwestie, które są ich przedmiotem.

Jeszcze inną możliwą reakcją (w żadnym wypadku nie będę tutaj wymieniał wszystkich możliwych opcji) jest podejrziwość wobec sensowności podejmowanych problemów.

Filozofia a składnia

Być może Wielkie Tezy Filozoficzne to tylko mądrze brzmiący bełkot, a posługując się nimi, mamy tylko wrażenie głębi i doniosłości?

Problem uzasadnienia tego typu podejrzeń to jedna sprawa. Druga to wytłumaczenie, skąd bierze się złudzenie sensowności Wielkich Sporów Filozoficznych i na czym polega ich czysty werbalizm. Pewnej konstrukcji tego typu podjął się RUDOLF CARNAP (1891–1970), przedstawiciel austriackiego ugrupowania filozofów analitycznych rozmiłowanych w fizyce i filozofii nauki, zwanego **KOŁEM WIEDEŃSKIM**.

Pierwszym istotnym elementem poglądu Carnapa jest jego podział zdań na trzy rodzaje. Pierwszym rodzajem są **ZDANIA SYNTAKTYCZNE**, które dotyczą jedynie poprawności lub natury gramatycznej różnego rodzaju wyrażeń. Przykładami zdań syntaktycznych są: „Zdanie »Predator jest pacyfistą« jest zdaniem analitycznym” czy „Wyraz »predator« nie występuje w *Krytyce czystego rozumu*”.

Drugą grupę stanowią **RZETELNE ZDANIA PRZEDMIOTOWE**, które naprawdę dotyczą przedmiotów pozajęzykowych. Przykładami zdań tego typu są: „Wiertarki jadają żyrafy” czy „Tylko kosmici lubią logików”. Są to zdania przedmiotowe języka polskiego, bo dotyczą jedynie rzeczy, które nie są wyrażeniami języka polskiego: wiertarek, żyraf, jedzenia, kosmitów, logików itd.

Trzeci (najbardziej problematyczny) typ zdań to **ZDANIA PSEUDOPRZEDMIOTOWE**, które wyglądają jak rzetelne zdania przedmiotowe, ale których rzeczywista treść jest zasadniczo syntaktyczna. Przykładami takich sądów są wedle Carnapa:

Księżyc jest rzeczą

które należy rozumieć jako nietrafną reprezentację treści zdania syntaktycznego:

Wyraz „księżyc” jest rzeczownikiem, który zachowuje się tak, jakby odnosił się do przedmiotu pozajęzykowego

lub:

S.L. Clemens jest tożsamy z M. Twainem

które należy wedle Carnapa interpretować tak, jak:

Wyrażenia „S.L. Clemens” i „M. Twain” są synonimiczne.

Carnap uważa więc, że nie ma takiej (metafizycznie brzmiącej) własności jak tożsamość. Raczej gdy mówimy o tożsamości, to nie mówimy o samych przedmiotach, tylko o ich nazwach, i stwierdzamy, że pozostają one do siebie w pewnej relacji.

Kluczowym twierdzeniem Carnapa jest to, że sensowne twierdzenia filozoficzne są albo syntaktyczne, albo pseudoprzedmiotowe. O ile często uznaje on filozoficzne zdania syntaktyczne za wartościowe poznawczo, o tyle zdania pseudoprzedmiotowe uważa za zdradliwe: dotyczą bowiem języka, a udają, że mówią o samych rzeczach. Wygląda też na to, że Carnap stosuje zasadę: **zdania filozoficzne nieprzekładalne na zdania syntaktyczne pozbawione są sensu.**

Intuicja leżąca u podstaw niechęci do zdań pseudopredmiotowych nie jest całkiem bezpodstawna. Na przykład zdanie „Afryka opisywana jest przez Hemingwaya” w jakimś sensie nie stwierdza żadnej rzeczywistej własności Afryki (jak sugeruje Carnap, można wiedzieć wszystko o samej Afryce, nie mając pojęcia o tym, kto i gdzie o niej pisał). Raczej stwierdza ono pewną własność dzieł Hemingwaya – treścią tego zdania wydaje się „Wyraz *Afryka* występuje w jakimś tekście Hemingwaya”.

Zdania formułowane tak, jakby były o przedmiotach, nazywa Carnap także zdaniami **w TRYBIE MATERIALNYM**. Zdaniami w tym trybie są zarówno rzetelne zdania przedmiotowe, jak i zdania pseudopredmiotowe. Tryb, w którym zdanie jawnie orzeka coś o własnościach syntaktycznych jakichś wyrażen, nazywa Carnap **TRYBEM FORMALNYM**.

Problem z Wielkimi Twierdzeniami Filozoficznymi, jaki Carnap widzi, jest taki, że jeżeli w ogóle mają sens, to są one w trybie materialnym, chociaż wyrażają treści syntaktyczne. Udają, że mówią o rzeczach, a mówią o języku. Na przykład gdy sprawiamy wrażenie, że mówimy o cechach, tak naprawdę mówimy o wyrażeniach: gdy mówimy „Głupota jest cechą”, treścią jest „Wyrażenie »Jest głupi« jest predykatem jednoargumentowym” (czyli takim wyrażeniem, które orzeczono o jednej nazwie tworzy zdanie). Gdy mówimy „Miłość jest relacją”, tak naprawdę nie mówimy nic o samej miłości, a jedynie stwierdzamy, że wyrażenie „kocha” jest predykatem dwuargumentowym (że tworzy ono zdanie razem z dwoma nazwami).

Zagrożenie, jakie wiąże się z używaniem trybu materialnego zamiast trybu formalnego, jest takie, że mówiąc o języku, możemy nie zdawać sobie z tego sprawy i myśleć, że mówimy o samych rzeczach. To zaś, zdaniem Carnapa, prowadzić może do jałowych sporów filozoficznych, kiedy to myślimy, że z kimś się naprawdę nie zgadzamy, podczas gdy różnimy się tylko tym, że używamy nieco innego języka.

Jednym z przykładów nieużytecznych dyskusji filozoficznych jest wedle Carnapa dyskusja nad „prawdziwym sensem pojęcia możliwości”. Dyskusja ta pojawia się, gdy nie zdajemy sobie sprawy z tego, że zdania typu „Jest możliwe, że p ” tak naprawdę są zdaniami syntaktycznymi.

Na przykład jeżeli chodzi o możliwość logiczną, to „Jest możliwe, że p ” należy jego zdaniem rozumieć jako „Zdanie p nie jest logicznie sprzeczne”. Na przykład gdy mówię, że jest (logicznie) możliwe, że obaj jesteśmy w błędzie, to zdaniem Carnapa nie przypisuję żadnej tajemniczej własności możliwości żadnemu stanowi rzeczy (takiemu jak naszemu byciu w błędzie). Raczej przypisuję własność logicznej niesprzeczności zdaniu „Obaj jesteśmy w błędzie”.

Kolejny przykład: jeżeli chodzi o możliwość fizyczną, zdanie „Jest możliwe, że p ” należy rozumieć jako „Zdanie p nie

jest logicznie wykluczone przez prawa fizyki”. Gdy mówię „Lot na Marsa jest możliwy,” to nie mówię, że pewne wydarzenie (lot na Marsa) ma własność bycia możliwym, tylko że zdanie „Odbędzie się lot na Marsa” nie jest logicznie wykluczone przez prawa fizyki.

Analogiczne analizy można zdaniem Carnapa podawać dla innych modalności. Stąd też spór o prawdziwy sens możliwości pojawia się, gdy nie zdajemy sobie sprawy z tego, że zmiana rozumienia możliwości to po prostu zmiana języka, a nie zmiana teorii dotyczącej jakiejś jednej, niezmiennej i absolutnej istoty możliwości. A zmiana języka oceniana ma być nie na gruncie merytorycznym, tylko na gruncie pragmatycznym. Zamiast bezsensownie pytać, który język jest prawdziwy, powinniśmy pytać, który język jest bardziej przydatny do obecnych celów.

Inny przykład. Russell, jak pamiętamy, definiował liczby jako pewne zbiory zbiorów. Istnieje podejście alternatywne, w którym liczby są po prostu przedmiotami niezbudowanymi z innych przedmiotów (jak np. w podejściu Hilberta). Na poziomie pozornie materialnym można by się spierać, czy liczby są zbiorami zbiorów, czy obiektami podstawowymi. Carnap jednak uważa, że zamiast tego należy:

- po pierwsze, przełożyć te tezy na tryb formalny, otrzymując: (1) **wyrażenia liczbowe są wyrażeniami teorii zbiorów (drugiego rzędu)** oraz (2) **wyrażenia liczbowe to wyrażenia indywiduowe**;
- po drugie, należy zauważyć, że nie doprecyzowaliśmy, czy mówimy o tych samych językach. W rzeczywistości – wcale nie. Powinniśmy raczej otrzymać: (1') **wyrażenia liczbowe są wyrażeniami teorii zbiorów (drugiego rzędu) w języku systemu Russella** oraz (2') **wyrażenia liczbowe to wyrażenia indywiduowe w języku systemu Hilberta**.

Wtedy też dostrzeżemy, że spór był jałowy, bo między naszymi tezami nie było żadnej rzeczywistej sprzeczności.

Inny przykład sporu jałowego wedle Carnapa to spór, czym naprawdę są rzeczy. Z jednej strony niektórzy twierdzą, że (1) **rzecz jest jedynie układem danych zmysłowych**; z drugiej zaś inni twierdzą, że (2) **rzecz jest układem cząstek materii fizycznej**. Zdaniem Carnapa po przełożeniu na tryb formalny niezgoda znika i obie strony mają rację. W trybie formalnym bowiem poglądy te głoszą:

- (1) Każde zdanie, w którym występuje rzeczownik zachowujący się, jakby odnosił się do rzeczy zewnętrznej, jest równoważne zbiorowi zdań, w których taki rzeczownik nie występuje, a w którym występują jedynie rzeczowniki odnoszące się do danych zmysłowych.
- (2) Każde zdanie, w którym występuje rzeczownik zachowujący się jakby odnosił się do rzeczy zewnętrznej, jest równoważne jakiemuś zdaniu, które nie zawiera żadnych takich rzeczowników, a tylko współrzędne czasoprzestrzenne i funkcje fizyczne.

O ironio, okazało się, że (1) to jedno z bardziej wątpliwych zdań filozoficznych – trudno bowiem opisać stan rzeczy „na stole leży wiertarka” za pomocą listy **wszystkich możliwych** doznań zmysłowych, jakie mogą się z tym wiązać; lista bowiem wszystkich takich doznań byłaby nieskończona. Zwolennikom tezy (1) nie udało się też nigdy podać przekonującej reguły generowania takiego przekładu. Ale to zagadnienie jest dla nas uboczne, więc idźmy dalej.

Pewne tezy filozoficzne Carnap uważa za w miarę łatwo przekładalne na tryb formalny. Na przykład:

- ☞ Jakikolwiek byt realny, który odpowiadałby naszemu empirycznemu jego wyobrażeniu, musi wykraczać poza samo to wyobrażenie. Konstrukt nadbudowany na jakichś bezpośrednich danych świadomości nie może być tożsamy z realnym przedmiotem. Sama rzecz musi być bardziej konkretna; konstrukt w porównaniu z nią pozostaje czymś abstrakcyjnym

tłumaczy jako:

„Dla każdego wyrażenia zachowującego się, jakby było nazwą przedmiotu, i dla każdego skończonego zbioru Z zdań istnieją takie zdania zawierające tę nazwę, że ani potwierdzenie samych tych zdań, ani potwierdzenie ich negacji nie redukuje się całkowicie do potwierdzenia zbioru Z

i zgadza się z jego wersją w trybie formalnym.

Niestety przyznaje też, że formułowanie poglądów filozoficznych w trybie materialnym prowokuje również do stawiania tez i pytań wieloznacznych, których przekład na tryb formalny trudno znaleźć, lub po prostu bezsensownych. Przykłady takich pytań to: „Czy jaźń jest czymś więcej niż tylko jedną z treści doświadczenia, które nazywam moim?” czy pytanie o prawdziwą naturę liczby. Problemy, których nie można jasno przełożyć na tryb formalny, uznaje Carnap za pseudoproblemy.

Carnap jest przekonany o trafności swojej strategii. Sugeruje, że jeżeli w dyskusji filozoficznej przeciwnik odmawia przełożenia swoich poglądów na tryb formalny albo nie chce powiedzieć, do jakiego systemu językowego należy zrelatywizować jego tezę, to należy po prostu odmówić dalszej dyskusji, która byłaby tylko przelewaniem z pustego w próżne.

Wszystko byłoby pięknie i moglibyśmy wieść życie, nie przejmując się problemami filozoficznymi, gdyby nie to, że strategia Carnapa nie działa. Dlaczego? Jedną z metod oceny filozoficznych strategii argumentacyjnych jest ocena tego, czy nie prowadzą one do niepożądanych konsekwencji. A wygląda na to, że strategia Carnapa ma taką własność. Dlaczego właściwie mielibyśmy zgodzić się na to, że sensowne zdania filozoficzne mają treść zdań syntaktycznych? Carnap sugeruje, że jest tak dlatego, że są one na tryb formalny przekładalne. Załóżmy roboczo, że nawet wiemy, jak Carnap rozumie tutaj przekładalność. Jego argumentacja zdaje się zakładać, co następuje:

Filozofia a składnia

[CARNAPIZM] Jeżeli zdanie jest przekładalne na tryb formalny, to jego treść jest naprawdę syntaktyczna.

Jeżeli jednak zgodzimy się z Carnapem, że dobrym przekładem jest przekład następujący:

[TŁUMACZENIE 1]	
ZDANIE WYJŚCIOWE	TŁUMACZENIE
Księżyć jest rzeczą.	Wyraz „księżyc” jest rzeczownikiem, który zachowuje się tak, jakby się odnosił do przedmiotu pozajęzykowego.

to czemu nie mielibyśmy uznać za poprawne tłumaczenia następującego:

[TŁUMACZENIE 2]	
ZDANIE WYJŚCIOWE	TŁUMACZENIE
Wiertarka jest narzędziem.	Wyrazy „wiertarka” i „narzędzie” są rzeczownikami, które zachowują się tak, jakby pierwszy z nich nie odnosił się do niczego, do czego nie odnosi się drugi z nich.

Jeżeli jednak uważamy, że zdanie „Wiertarka jest narzędziem” z tego powodu jest zdaniem pseudoprzedmiotowym i nie jest tak naprawdę zdaniem o wiertarkach i narzędziach, mamy problem. Wygląda więc na to, że albo [CARNAPIZM] jest fałszywy, albo strategie translacyjne, którymi się posłużyliśmy, nie dają prawdziwych tłumaczeń, albo sama przetłumaczalność nie dowodzi pseudoprzedmiotowości.

Na czym jednak miałyby polegać bycie dobrym tłumaczeniem? Pierwsza intuicyjna propozycja to:

[PROPOZYCJA 1] Jedno wyrażenie jest tłumaczeniem drugiego wtedy i tylko wtedy, gdy wyrażenia te są synonimiczne.

[PROPOZYCJA 1] ma tę zaletę, że jeżeli ją przyjmiemy, [CARNAPIZM] wydaje się przekonujący. Problem polega jednak na tym, że ani

[TŁUMACZENIE 1], ani żadne z innych tłumaczeń tez filozoficznych proponowanych przez Carnapa nie wydaje się synonimiczne ze zdaniem wyjściowym. Zdania wyjściowe zawierają nazwy przedmiotów (takich jak księżyc), tłumaczenia takich nazw nie zawierają i mówią o wyrażeniach. Oczywiście można utrzymywać jakąś teorię spiskową, mówiąc, że większość ludzi, którzy twierdzą, że księżyc jest rzeczą, nie zdaje sobie sprawy z tego, że wcale nie mówią o księżycu. Ale strategia ta wydaje się mało przekonująca.

Być może należy więc propozycję osłabić i zamiast tego powiedzieć, co następuje:

[PROPOZYCJA 2] Jedno zdanie jest tłumaczeniem drugiego wtedy i tylko wtedy, gdy zdania te są logicznie równoważne (czyli zarazem prawdziwe lub zarazem fałszywe w każdej niesprzecznie opisywalnej sytuacji).

Problem jednak polega na tym, że [TŁUMACZENIE 1] nie jest poprawnym tłumaczeniem w sensie [PROPOZYCJI 2]. Może być przecież tak, że księżyc jest rzeczą, podczas gdy sam obiekt syntaktyczny „księżyc” stał się w języku polskim czasownikiem; może być również tak, że wyrażenie „księżyc” zachowuje się tak, jakby się odnosiło do rzeczy, podczas gdy księżyc przestał istnieć i nie jest rzeczą. Co więcej, [PROPOZYCJA 2] dalej nie wyjaśnia różnicy między [TŁUMACZENIEM 1] i [TŁUMACZENIEM 2].

Ogólnie rzecz biorąc, pojęcie tłumaczenia powinno być co najmniej tak silne jak [PROPOZYCJA 2] w tym sensie, że tłumaczenia powinny być co najmniej logicznie równoważne: wszelkie słabsze pojęcie nie będzie, intuicyjnie rzecz biorąc, miało wiele wspólnego z tłumaczeniem. Z drugiej jednak strony, osłabienie warunku aż do [PROPOZYCJI 2] jest problematyczne, bo w takim sensie słabszym od synonimiczności dostajemy na przykład wniosek, że dowolne dwa zdania konieczne (jak $2 + 2 = 4$ i $1 + 1 = 2$) są wzajemnie poprawnymi tłumaczeniami, bo są prawdziwe w tych samych możliwych okolicznościach.

Filozofia a składnia

Pójdźmy dalej: jeżeli chcemy utrzymać [CARNAPIZM] przy zachowaniu [PROPOZYCJI 2], to dojdziemy do wniosku, że każde konieczne zdanie jest przetłumaczalne na zdanie syntaktyczne; bo dowolne zdanie konieczne będzie prawdziwe w dokładnie tych samych warunkach, co zdanie syntaktyczne:

Wyrażenie „wierdarka”, jeżeli istnieje, jest równokształtne z wyrażeniem „wierdarka”.

Nie chcemy jednak raczej uznawać, że dowolne zdanie konieczne jest zdaniem czysto syntaktycznym. Zdanie $2 + 2 = 4$ jest zdaniem o liczbach, zdanie „Pada deszcz lub nie pada deszcz” jest zdaniem o deszczu itd.

Stoimy więc przed następującym dylematem: albo przyjmujemy [PROPOZYCJĘ 1], co sprawia, że [CARNAPIZM] jest przekonujący, ale zarazem sprawia, że rozważane niby-tłumaczenia nie są poprawnymi tłumaczeniami, albo przyjmujemy [PROPOZYCJĘ 2], która pozwala (być może przez chwilę) wierzyć, że mamy do czynienia z poprawnymi tłumaczeniami, ale sprawia, że [CARNAPIZM] jest kompletnie nieprzekonujący. Tak czy siak sytuacja wygląda słabo.

Uratowanie propozycji Carnapowskiej wymagałoby dostarczenia pojęcia tłumaczenia, które spełniałoby zarazem następujące warunki:

- (1) [CARNAPIZM] musi być dalej przekonujący.
- (2) Carnapowskie tłumaczenia tez filozoficznych, takie jak [TŁUMACZENIE 1], powinny być poprawnymi tłumaczeniami.
- (3) „Świrowe” tłumaczenia, takie jak [TŁUMACZENIE 2], nie powinny być poprawnymi tłumaczeniami.

Jak na razie nikt takiego pojęcia tłumaczenia nie wyeksplikował.

19. Carnap: pytania wewnętrzne i zewnętrzne

Nieco odmienne podejście do werbalności dyskusji metafizycznych prezentuje Carnap w *Empiricism, Semantics and Ontology*³⁵. Carnap twierdzi tam, że pytania o istnienie należy dzielić na **PYTANIA WEWNĘTRZNE i PYTANIA ZEWNĘTRZNE**.

Wewnętrzność (lub zewnętrzność) pytania oceniamy ze względu na daną aparaturę pojęciową (standardowy zwrot angielski to *linguistic framework*; w języku polskim funkcjonuje czasem tłumaczenie „schemat pojęciowy”; moje tłumaczenie pewnie też nie jest udane). Aparaturę pojęciową konstruujemy zawsze, gdy chcemy mówić o nowym rodzaju przedmiotów (powiedzmy o *X*-ach) i wprowadzamy sposoby i reguły mówienia o nich. Pytanie o istnienie jakichś konkretnych *X*-ów jest wewnętrzne ze względu na taką aparaturę pojęciową. Pytanie zaś o istnienie *X*-ów w ogóle jest pytaniem zewnętrznym względem tej aparatury.

Na przykład pytanie o istnienie liczby, która jest pierwiastkiem z liczby 287, jest wewnętrzne ze względu na aparaturę pojęciową arytmetyki, ponieważ pytamy o istnienie jakiejś konkretnej liczby, a samo posługiwanie się tą aparaturą nie gwarantuje odpowiedzi pozytywnej na to pytanie. Aby na nie odpowiedzieć, musimy przeprowadzić rozważania wewnętrznej teorii.

Dla kontrastu pytanie ogólne o istnienie liczb, sugeruje Carnap, jest pytaniem zewnętrznym i nie ma sensu w całej swojej ogólności, bo albo przyjmujemy aparaturę pojęciową arytmetyki i automatycznie mówimy tak, jakbyśmy zgadzali się na istnienie liczb (w ogóle), albo tej aparatury nie przyjmujemy, a w takim razie automatycznie mówimy tak, jakbyśmy istnienie liczb odrzucali.

³⁵ Tekst ukazał się po polsku w książce *Empiryzm, semantyka, ontologia* w tłumaczeniu Artura Koterskiego (2005).

Carnap: pytania wewnętrzne i zewnętrzne

Pytania wewnętrzne mają rzeczywiste odpowiedzi. Pytania zewnętrzne – nie; są one wedle Carnapa tylko kwestią wyboru aparatury pojęciowej. Wybór taki uważał on za kwestię pragmatyczną. Nie ma „prawdziwej aparatury pojęciowej” ani prawdziwych rozstrzygnięć ontologicznych, są tylko kwestie wyboru mniej lub bardziej przydatnych, wydajnych czy prostych narzędzi językowych. Zamiast więc pytać, czy istnieją liczby, sugeruje Carnap, powinniśmy raczej zapytać, czy pożytecznie jest posługiwać się aparaturą pojęciową, która liczby wprowadza i pozwala nam mówić tak, jakby one istniały.

Propozycja Carnapa jest dyskusyjna z różnych względów. Pierwszym problemem jest to, że Carnap nie wyjaśnia zbyt precyzyjnie, czym jest aparatura pojęciowa i (podążając za Carnapem) być może mamy jakieś (pragmatyczne?) powody, by nie przyjmować aparatury pojęciowej aparatur pojęciowych.

Po drugie (jak zauważa Chalmers w *Metametaphysics. New Essays on the Foundations of Ontology*, s. 112), za pomocą jednego i tego samego zdania można zadawać różne pytania, zarówno wewnętrzne, jak i zewnętrzne. Metafizyk może zapytać, czy istnieją liczby parzyste, i mieć na myśli pytanie ontologiczne i, zdaniem Carnapa, zewnętrzne, a matematyk może użyć tego samego zwrotu, by zadać dość trywialne pytanie wewnętrzne.

Po trzecie, o ile Carnap podkreśla, że wybór aparatury jest kwestią pragmatyczną, o tyle zupełnie przemilcza on obiektywny aspekt całej sprawy, który w sumie pozostaje mniej więcej tam, gdzie był. Jeżeli bowiem uważamy, że pewne aparatury są do pewnych określonych celów lepsze od innych, to dobrze by było wyjaśnić, dlaczego tak jest. Czy jest to po prostu pierwotny fakt, który możemy tylko zaobserwować i się z nim pogodzić, czy też jest raczej tak, że pewne aparatury są przydatne bardziej niż inne, dlatego, że bardziej adekwatnie ujmują jakiś aspekt rzeczywistości? Jeżeli nie ma rzeczywistych pytań ontologicznych i rzeczywistych na nie odpowiedzi, nie jest jasne, dlaczego przyjęcie aparatury mówiącej o liczbach miałoby być rzeczywiście przydatne.

Z drugiej strony z samej przydatności aparatury pojęciowej wprowadzającej jakieś obiekty nie wynika ich istnienie. Nie zmienia to jednak faktu, że taka przydatność dostarcza *prima facie* (czyli na pierwszy rzut wiertarką) powodu, by takie istnienie podejrzewać. Jeżeli zaś rzeczywiście używamy tej aparatury, a odrzucamy istnienie odpowiednich obiektów, stajemy przed wyzwaniem wyjaśnienia pożyteczności naszej aparatury w jakiś inny sposób.

Po czwarte, wszystko to ma wpływ na negatywną ocenę metafizyki, tylko jeżeli przyjmiemy, że decyzja o uznaniu jakiegoś metafizycznie istotnego rodzaju bytów (jak liczby, umysły czy obiekty fizyczne) to decyzja o przyjęciu aparatury pojęciowej. Kryterium jednostkowania aparatów pojęciowych nie jest jednak jasne, a z tego, że jakieś pytanie jest zewnętrzne względem jakiejś aparatury, nie wynika, że jest zewnętrzne **W OGÓLE** i że dlatego pozostaje kwestią czysto pragmatyczną.

Na przykład uznanie istnienia liczb parzystych może być postrzegane jako wewnętrzne względem uznania liczb. Może też jednak być postrzegane jako wprowadzenie nowej aparatury pojęciowej liczb parzystych. Dlaczego nie? Czy to świadczy o werbalności pytania o istnienie liczb parzystych? Może lepszy przykład: Czy pytanie o istnienie liczb urojonych jest kwestią zewnętrzną czy wewnętrzną? Względem czego?

Analogicznie uznanie istnienia liczb w ogóle może być postrzegane jako zewnętrzne (względem uznania aparatury liczb). Dlaczego jednak nie może być równocześnie postrzegane jako wewnętrzne względem aparatury pojęciowej przedmiotów w ogóle? Dlaczego sam fakt, że uznanie istnienia *X*-ów jest zewnętrzne względem jakiejś aparatury, ma pociągać za sobą to, że pytanie o istnienie *X*-ów jest czysto pragmatyczne i nie jest wewnętrzne względem żadnej innej aparatury?

Sam fakt, że zewnętrzność lub wewnętrzność pytania jest zrelatywizowana do wyboru aparatury, sprawia, że nie ma (chyba) pytań, które są absolutnie zewnętrzne czy absolutnie wewnętrzne. Jak na ironię to, czy dane pytanie jest zewnętrz-

Carnap: pytania wewnętrzne i zewnętrzne

ne czy wewnętrzne, samo staje się kwestią wyboru aparatury pojęciowej (o ile chcemy się bawić w żonglowanie aparaturą pojęciową aparatur pojęciowych).

Niezależnie od porażki ogólnego zamierzenia Carnapa warto mieć na uwadze jedną z kluczowych motywacji, jakie nim kierowały w konstrukcji jego poglądów. Carnap przy wielu okazjach promował postawę tolerancji dla odmiennych poglądów oraz podkreślał, że wiele z tych zjawisk, które postrzegamy jako różnice istotne, to tylko różnice słowne. O ile całościowa charakteryzacja wszystkich debat filozoficznych jako debat czysto werbalnych zawiodła, o tyle nie podważa intuicji, że przynajmniej w niektórych przypadkach coś jest na rzeczy i że rzeczywiście przynajmniej niektóre debaty filozoficzne (i nie tylko), mimo zacierzwienia obu stron, to tylko debaty słowne.

Szersza dyskusja na temat natury debat słownych i metodologii ich identyfikacji, nie popełniająca metodologicznych błędów Carnapa, musiała poczekać do końca XX wieku, kiedy to odżyła z nową siłą i jakością, bo ponownie podjęto kwestie werbalności niektórych debat filozoficznych z nowymi narzędziami w rękę (polecam gorąco *Metametaphysics* — szczególnie we wskazówkach bibliograficznych na stronie 115).

CARNAP O SYNTAKTYCZNOŚCI FILOZOFII

1. Zdania: syntaktyczne, rzetelne przedmiotowe i pseudopredmiotowe (takie, które są syntaktyczne, ale udają przedmiotowe).
2. Sensowne, niesyntaktyczne i doniosłe brzmiące zdania filozoficzne są przedmiotowe i wyrażają tak naprawdę sądy syntaktyczne.
3. Rozróżnienie trybu materialnego (jawnie o przedmiotach) i formalnego (jawnie o wyrażeniach) jest potrzebne: myślenie, że pseudopredmiotowe zdania filozofii są w trybie materialnym, prowadzi do jałowych sporów.
4. Problemy filozoficzne przełożone na tryb formalny stają się trywialne, ich rozwiązanie polega na relatywizacji do języka.
5. Problem: przekładalność na tryb syntaktyczny nie wystarcza do dowiedzenia syntaktyczności treści; pojęcie przekładalności jest zbyt niejasne.



Wybrani dziadkowie filozofii analitycznej

6. Próba pierwsza: przetłumaczalność to synonimiczność. Taka przekładalność wystarcza do dowiedzenia syntaktyczności treści, ale w większości przypadków, o których mówi Carnap, nie zachodzi.
7. Być może przetłumaczalność to logiczna równoważność? To też nie występuje tak często, jak Carnap by chciał.
8. Logiczna równoważność nie wystarcza do dowiedzenia syntaktyczności treści.
9. Nie sformułowano na razie pojęcia przetłumaczalności, które zarazem wystarcza do dowodzenia syntaktyczności i stosuje się do przypadków, o których mówi Carnap.



CARNAP: PYTANIA WEWNĘTRZNE I ZEWNĘTRZNE

1. Pytanie o istnienie jakichś specyficznych **X**-ów jest wewnętrzne ze względu na daną aparaturę pojęciową, o ile wprowadziła ona **X**-y w ogóle.
2. Pytanie o istnienie **X**-ów w ogóle jest zewnętrzne ze względu na aparaturę pojęciową, która dopiero wprowadza **X**-y w ogóle.
3. Wybór aparatury pojęciowej i tego, co ona ma wprowadzić, jest kwestią pragmatyczną.
4. Pytania zewnętrzne nie mają żadnej innej odpowiedzi niż odpowiedź o pożyteczność wprowadzenia tej, a nie innej aparatury pojęciowej.
5. Carnap nie wyjaśnia, co to jest aparatura pojęciowa i dlaczego mamy w aparaturę aparatur pojęciowych wierzyć.
6. Jedno i to samo zdanie może wyrażać pytanie zewnętrzne i pytanie wewnętrzne w różnych sytuacjach.
7. Carnap nic nie mówi o tym, dlaczego różne aparaty miałyby się różnić przydatnością.
8. Nie jest jasne, dlaczego pytania filozoficzne mają być zewnętrzne. Jasne, że są zewnętrzne względem pewnej aparatury pojęciowej, lecz także mogą być jednocześnie wewnętrzne względem innej. Co wtedy?

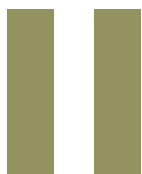
Carnap: pytania wewnętrzne i zewnętrzne

KIERUNKOWSKAZ



W świetle tak dużej koncentracji zainteresowań wczesnych filozofów analitycznych na języku łatwo rozwinął się mit, jakoby filozofowie analityczni zajmowali się tylko językiem. Celem najbliższej sekcji będzie wyjaśnienie paru kwestii związanych z tym mitem. Następnie wrócimy do kwestii kryteriów sensowności wyrażen, zanim przyjrzymy się kilku zastosowaniom narzędzi, na które trafiliśmy, w filozofii religii i etyki.

CZĘŚĆ



Trochę bardziej merytorycznie

20. Mit zwrotu lingwistycznego

W roku 1967 Richard Rorty opublikował zredagowaną przez siebie antologię *Zwrot lingwistyczny* (*The Linguistic Turn*). Głosił w niej, że miała miejsce rewolucja filozoficzna polegająca na pojawieniu się filozofii lingwistycznej, wedle której problemy filozoficzne mogą być zlikwidowane lub rozwiązane przez rozważania dotyczące języka.

Hasło się przyjęło i chętnie podchwycili je ci filozofowie, którzy nie przepadali zbyt za stylem, w jakim filozofowie analityczni uprawiali filozofię. Bardzo łatwo przeszli oni od „filozofii lingwistycznej” do „filozofii, która zajmuje się tylko językiem”, do „filozofii, która naiwnie myśli, że problemy filozoficzne są tylko problemami językowymi” i do „filozofii, która nie zajmuje się niczym ważnym”.

Każda nowość ma swój okres zachłyśnięcia się nią. Nastolatek odkrywa alkohol i często na początku z nim przesadza. Ja kupuję nowe **CALL OF DUTY** i prawie nie zajmuję się niczym innym przez kilka dni (przepraszam wszystkich studentów, którym nie oddałem na czas sprawdzonych prac). Podobnie filozofowie analityczni w okresie młodości tego stylu uprawiania filozofii nieco zachłysłni się tym podejściem i przesadzali w komentarzach na temat tego, co robili. Stąd takie patologiczne przypadki, jak niektóre uwagi Wittgensteina na temat problemów filozoficznych jako powstających wskutek niewłaściwego użycia języka, jak niektóre uwagi ALFREDA JULESA AYERA³⁶, który powiada, że filozof jako analityk nie zajmuje się własnościami rzeczy, czy jak komentarze MICHAELA DUMMETTA³⁷, który mówi, że celem filozofii jest analiza struktury myśli.

³⁶ Alfred Jules Ayer (1910–1989) był jednym z czołowych brytyjskich filozofów analitycznych, zasłynął z książki *Language, Truth, and Logic* (1936).

³⁷ Michael Dummett (1925–2011) jest jednym z ważniejszych filozofów analitycznych XX wieku. Poglądem, z którego jest znany, jest antyrealizm semantyczny, wedle

Trochę bardziej merytorycznie

Po pierwsze, nie da się zaprzeczyć, że w jakimś sensie filozofowie analityczni są zainteresowani językiem, ale zainteresowanie językiem nie jest cechą ich wyróżniającą. Językiem również interesowali się Martin Heidegger w swoich analizach etymologicznych, Jacques Derrida czy Hans-Georg Gadamer, a żaden z nich filozofem analitycznym nie jest.

Po drugie, niewielu filozofów analitycznych (a jeszcze mniej – żyjących filozofów analitycznych) rzeczywiście uważa, że problemy filozoficzne można zlikwidować i rozwiązać jedynie przez zajmowanie się językiem. Raczej uważają oni, że filozofia jest rzeczą skomplikowaną i jednym z istotnych elementów dobrego uprawiania filozofii jest dobre rozumienie i wyjaśnianie pojęć, którymi się posługuje, jasne przedstawianie argumentów, które się wysuwa, oraz ich struktury. Problemy filozoficzne same jako takie są trudne i nie ma potrzeby komplikowania ich pozersko skomplikowanym językiem i niewyjaśnionymi neologizmami.

Po trzecie, jest istotna różnica pomiędzy **zwracaniem uwagi na język a zajmowaniem się głównie językiem**. Właściciel firmy budowlanej będzie interesował się przydatnością poszczególnych modeli wiertarek do rozmaitych celów i pewien czas będzie spędzał na ocenie jakości sprzętu i materiałów, jakimi będą posługiwać się jego pracownicy. Z tego nie wynika, że głównym przedmiotem jego zainteresowań będzie budowa wiertarek oraz skład chemiczny różnych klejów do glazury. Byle jaki właściciel firmy budowlanej stwierdzi: „Oj tam, ważne, żeby budować”, kupi byle co i pobiegnie radośnie na budowę. Jednak dobry właściciel, zanim podejmie się budowy, odpowiednio dobierze narzędzia i materiały do wykonania tego zadania.

Podobnie, dobry filozof po drodze do argumentacji za jakimś twierdzeniem filozoficznym zastanowi się kilka razy nad tym, jakich narzędzi językowych do tego użyć, jakie one

którego zdania nierozstrzygalne pozbawione są znaczenia. Zasłynął również z prac historycznych dotyczących myśli Fregego.

Mit zwrotu lingwistycznego

pełnią funkcje i jakie konsekwencje ma ich dobór dla rozważanego zadania. W żadnym wypadku nie oznacza to jednak, że nie ma on nic lepszego do roboty niż myślenie o języku.

Być może ktoś zasugeruje, że uprawianie filozofii jest bardziej jak jeżdżenie na łyżwach: najlepiej się tym zajmować intuicyjnie, nie myśląc jednocześnie o takich rzeczach jak dobór i jakość sprzętu. W rzeczywistości jednak najlepsi łyżwiarze (a przynajmniej ich trenerzy) głęboko zastanawiają się nad doborem sprzętu. Jest to jeden z czynników, który odróżnia profesjonalnych łyżwiarzy od dzieci korzystających w niedzielę z lodowiska. Jedni i drudzy w pewnym podstawowym sensie jeżdżą na łyżwach. Ale pomiędzy tym, jak to robią, istnieje zasadnicza różnica jakości.

Poza nielicznymi i stosunkowo historycznymi przypadkami lekko świrowatego zachłyśnięcia się filozofią analityczną większość też głoszonych przez filozofów analitycznych w żadnym interesującym sensie nie dotyczy języka. Z nowymi narzędziami filozofowie ci podejmują wiele klasycznych problemów filozoficznych.

Zastanawiają się na przykład, jakie są podstawowe kategorie bytu. Oczywiście przy okazji rozważają, co to znaczy być kategorią bytu i jak taką, a nie inną kategoryzację można uzasadniać, ale ostatecznie chodzi im o kwestie metafizyczne.

Argumentują też na przykład za dualizmem w kwestii umysł – ciało lub przeciw niemu. Oczywiście po drodze niektóre argumenty zależeć mogą od takich szczegółów jak dobór teorii nazw własnych czy terminów naturalnorodajowych, ale to jedynie uczy, że w filozofii diabeł (gdy nie jest zajęty kształtowaniem polskiego systemu edukacji) tkwi w szczegółach, a nie że problem dualizmu jest problemem czysto lingwistycznym.

W szczególności podejście typu Carnapowskiego nie jest podejściem szeroko akceptowanym wśród filozofów analitycznych. Jest przypadkiem zachłyśnięcia się metodą. Oczywiście być może w jakimś sensie każde twierdzenie filozoficzne można „przełożyć” na twierdzenie o pojęciach czy o języku (wzoruując się na Carnapie), ale technika przekładu będzie zazwyczaj

Trochę bardziej merytorycznie

nieinteresująca i przekład nie będzie dowodził, że wyjściowy problem jest problemem tak naprawdę dotyczącym języka w jakimkolwiek interesującym sensie.

Nawet jeżeli zgodzimy się, że tezy brnione przez filozofów analitycznych są analityczne (a jest to bardzo dyskusyjny pogląd), to i tak rozważania w rozdziale 11 sugerują, że to, co filozofowie analityczni robią, jest dalekie od trywialności. Zasadniczo nie udało się sformułować takiej koncepcji analityczności, przy której zdania analityczne są trywialne; w tej kwestii warto zajrzeć do *The Philosophy of Philosophy* Timothy'ego Williamsona.

Stąd też, jeżeli chodzi o dobór tematów, nie da się określić filozofii analitycznej. Filozofowanie bowiem w tym stylu nie jest filozofowaniem na jakiś konkretny temat, czy jest to język, czy pojęcia, czy cokolwiek innego. Filozofów analitycznych jest dużo, zajmują się różnymi rzeczami, a tym, co ich łączy, jest raczej styl uprawiania filozofii niż wspólnie uznawane tezy filozoficzne czy jakaś ograniczona dziedzina filozofii. W tym sensie pytanie, jakimi problemami zajmują się filozofowie analityczni, jest jak pytanie „Z kim jadasz lunch?”. Pytanie to zakłada, że istnieje jedna osoba lub grupa, z którą jada się lunch, podczas gdy w rzeczywistości nic nie stoi na przeszkodzie, by jeść lunch z jedną osobą jednego dnia, z inną innego. I tak samo nic nie stoi na przeszkodzie, by filozof analityczny zajmował się etyką w poniedziałek, metafizyką we wtorek i epistemologią w środę. To też stanowi wyjaśnienie motta niniejszej książki.

21. Drogi i wyboje weryfikacjonizmu

Niezależnie od teorii zdań pseudopredmiotowych i syntaktycznej interpretacji sensownych twierdzeń filozoficznych filozofowie z klubu, do którego należał Carnap (**KOŁO WIE-**

Drogi i wyboje weryfikacjonizmu

DEŃSKIE), starali się wykluczyć szereg twierdzeń filozoficznych jako **poznawczo bezsensownych**, co miało niejako zapewnić rzetelność dyskusji filozoficznych. Aby to uczynić, starali się podać **KRYTERIUM SENSOWNOŚCI POZNAWCZEJ** zdań.

Ogólnie pomysł wydawał się całkiem prosty: zdania analitycznie prawdziwe i analitycznie fałszywe są sensowne, bo są prawdziwe lub fałszywe na mocy znaczenia. Zdania przygodne są sensowne, jeżeli pozostają w odpowiedniej relacji do empirycznie rozstrzygalnych zdań obserwacyjnych.

Gdy jednak przyrzeć się temu wszystkiemu z bliska, sprawa nie jest prosta. Uczynienie z powyższego ogólnego określenia interesującej poznawczo propozycji wymaga dookreślenia, jakie zdania są zdaniami obserwacyjnymi oraz na czym ma polegać ta relacja do zdań obserwacyjnych.

Pierwszym (w sumie pomniejszym) problemem było dookreślenie natury zdań obserwacyjnych. Problem polegał na tym, że członkowie Koła Wiedeńskiego, zwani też **NEOPOZYTYWISTAMI**, pracowali w cieniu nieco dziwnych pokartezjańskich pomysłów Russella, wedle których przedmiotem naszych własnych obserwacji są jedynie nasze własne dane zmysłowe.

Wedle tej koncepcji wiedza o tzw. świecie zewnętrznym ma być owocem tego typu obserwacji, ponieważ obiekty zewnętrzne są tak naprawdę konstruktami z danych zmysłowych. Pogląd niestety ma tę wadę, że jest absurdalny. Załóżmy bowiem, że z kolegą uciekam przed predatorem. Kolega mówi: „W krzakach obok jest predator”. Ja powiadam, „W krzakach obok nie ma predatora”. Jeżeli zdanie, które wypowiada mój kolega, odnosi się do konstruktu z jego danych zmysłowych, a moje zdanie odnosi się do mojego konstruktu z moich danych zmysłowych, nie zachodzi między nami żadna niezgoda. John McTiernan raczej by się nie zgodził.

Gdybyśmy chcieli uniknąć tego problemu, postulując, że predator jest konstrukcją z naszych wspólnych danych zmysłowych (a w sumie przez analogiczny argument z danych zmysłowych wszystkich pozostałych przy życiu towarzyszy

Trochę bardziej merytorycznie

ucieczki), popadniemy w jeszcze inny absurd. Wtedy bowiem, żeby wiedzieć, do czego się odnoszę, gdy mówię „predator”, muszę mieć dostęp do danych zmysłowych moich towarzyszy, co raczej wydaje się niemożliwe.

Pewną opcją jest porzucenie fascynacji danymi zmysłowymi i uznanie, że zdania obserwacyjne to po prostu bezczelne zdania o przedmiotach zewnętrznych, które napotykamy w życiu codziennym: o samym predatorze, o samych wiertarkach itd. O ile unikamy absurdów omówionych powyżej, to jednak ma to swój koszt epistemologiczny. Zdania obserwacyjne o obiektach zewnętrznych (takie jak „Predator jest teraz dwa metry od nas i biegnie w naszą stronę”) nie są bowiem ostatecznie weryfikowalne (w przeciwieństwie np. do „Wyduje mi się, że biegnie na mnie predator”, bo mój dostęp do tego, czy mi się coś wydaje, jest dużo bardziej bezpośredni niż do samego faktu dotyczącego tego, czy predator na mnie biegnie). Zawsze istnieje jakaś szansa, że przeprowadzone przeze mnie obserwacje są mylne, że jestem pod wpływem adrenaliny i myślę krowę z predatorem (to się zdarza częściej, niż człowiek by się spodziewał) itd.

Być może nie musimy się tym przejmować. Być może tak rozumiane zdania obserwacyjne o obiektach fizycznych są dostatecznie weryfikowalne i zgoda co do nich panująca jest wystarczająca, żeby się takimi problemami nie przejmować. Zasadniczo bowiem istnieje szeroka klasa zdań, które intuicyjnie rozpoznajemy jako zdania, których prawdziwość lub fałszywość możemy w zadowalający sposób ustalić za pomocą obserwacji, o ile odłożymy na bok zdziwaczałe przykłady i nieosiągalne standardy epistemiczne konstruowane przez wynudzonych filozofów. Dlatego też nie będziemy się w dalszej części martwić rozumieniem natury zdań obserwacyjnych i przyjrzymy się raczej temu, jaka relacja powinna zachodzić między danym zdaniem a zdaniami obserwacyjnymi, by zdanie to było poznawczo sensowne.

Jak więc sformułować kryterium poznawczej sensowności, nawet przy założeniu, że nasze rozumienie, czym są zda-

Drogi i wyboje weryfikacjonizmu

nia obserwacyjne, jest dostatecznie jasne? Możemy próbować zacząć od określenia, czym są ostateczna weryfikowalność i ostateczna falsyfikowalność, mówiąc:

Dane zdanie jest **ostatecznie weryfikowalne** (ostatecznie falsyfikowalne) wtedy i tylko wtedy, gdy wynika logicznie (gdy jego negacja wynika logicznie) z pewnego skończonego i niesprzecznego zbioru zdań obserwacyjnych.

- Warunek niesprzeczności występuje w tej definicji dlatego, że zbiór spreczny pociąga logicznie dowolne zdanie, a nie chcemy, żeby nasza definicja od razu rozstrzygała, że dowolne zdanie jest poznawczo sensowne.
- Warunek skończoności występuje w tej definicji dlatego, że chcemy, aby obserwacje rozstrzygające te zdania obserwacyjne można było przeprowadzić. Nie można zaś przeprowadzić nieskończonej ilości obserwacji.
- Warunek wynikania logicznego znajduje się w definicji dlatego, że chcemy, aby te obserwacje ostatecznie rozstrzygały o prawdziwości (fałszywości) tego zdania.

Definicja powyższa może nasuwać następującą próbę sformułowania kryterium:

[PRÓBA 1] Wyrażenie jest sensowne poznawczo wtedy i tylko wtedy, gdy jest ostatecznie weryfikowalne.

Problem jest taki, że jeżeli dziedzina przedmiotów rozważanych jest nieskończona, to wszelkie uogólnienia o tej dziedzinie nie będą ostatecznie weryfikowalne. Jeżeli na przykład ilość przestrzennych ciał fizycznych (przeszłych, obecnych i przyszłych) jest nieskończona, to żadna skończona liczba zdań obserwacyjnych nie pociągnie za sobą **logicznie** zdania „Wszystkie przestrzenne ciała fizyczne podlegają grawitacji”, bo żeby je ostatecznie zweryfikować, musielibyśmy zaobserwować **wszystkie** przestrzenne ciała fizyczne we wszystkich możliwych okolicznościach.

Trochę bardziej merytorycznie

Jako że uogólnienia są równoważne negacjom zdań egzystencjalnych, również negacje zdań egzystencjalnych o nieskończonych dziedzinach nie będą ostatecznie weryfikowalne. Na przykład zdania „Żadne przestrzenne ciało fizyczne nie jest predatorem” czy też „Nie istnieje przestrzenne ciało fizyczne, które jest predatorem” również nie będą wynikały logicznie z żadnego skończonego zbioru zdań obserwacyjnych, bo żeby je ostatecznie zweryfikować, musielibyśmy zaobserwować **każde** przestrzenne ciało fizyczne i ustalić, że nie jest ono predatorem.

Warto jednak zwrócić uwagę, że zdania obu typów są ostatecznie falsyfikowalne. Żeby sfalsyfikować ostatecznie zdanie ogólne, wystarczy zaobserwować jeden przypadek, który je wyklucza (na przykład wskazać chociaż jedno przestrzenne ciało fizyczne, które nie podlega grawitacji). Żeby sfalsyfikować negatywne twierdzenie egzystencjalne, wystarczy zaobserwować chociaż jeden przypadek pozytywny (na przykład zaobserwować chociaż jednego predatora).

To może zasugerować, że skoro [PRÓBA 1] nie działa, należy ją zastąpić próbą następującą:

[PRÓBA 2] Wyrażenie jest sensowne poznawczo wtedy i tylko wtedy, gdy jest ostatecznie falsyfikowalne.

Niestety w takim przypadku problem jest odwrotny. O ile nieweryfikowalne są pewne zdania ogólnotwierdzące (i przeczące egzystencjalne), o tyleiefalsyfikowalne są ich negacje.

Po pierwsze,iefalsyfikowalne ostatecznie są zdania egzystencjalne, o ile dziedzina jest nieskończona. Jeżeli ktoś twierdzi, że istnieje predator, to żadna skończona liczba obserwacji nie wykluczy **logicznie** prawdziwości tego zdania. Aby je sfalsyfikować, należałoby sprawdzić **wszystkie** elementy nieskończonej dziedziny i o każdym z nich stwierdzić, że nie jest predatorem.

Po drugie,iefalsyfikowalne ostatecznie są negacje zdań ogólnych, jeżeli dziedzina jest nieskończona. Jeżeli ktoś mówi,

Drogi i wyboje weryfikacjonizmu

że nie wszystkie przestrzenne ciała fizyczne podlegają grawitacji, to żeby to sfalsyfikować ostatecznie, muszę zaobserwować **każde** takie ciało i o każdym stwierdzić, że podlega grawitacji.

Ponadto zarówno [PRÓBA 1] jak i [PRÓBA 2] mają tę dziwną własność, że sprawiają, iż negacja zdania sensownego może nie być sensowna, bo jeżeli weryfikowalne (a więc sensowne wedle [PRÓBY 1]) jest zdanie egzystencjalne o dziedzinie nieskończonej, to jego negacja będzie nieweryfikowalna (a więc też i bezsensowna wedle [PRÓBY 1]).

Analogicznie, jeżeli falsyfikowalne (a więc sensowne wedle [PRÓBY 2]) będzie jakieś zdanie ogólnotwierdzące o jakiejś dziedzinie nieskończonej, to jego negacja będzieiefalsyfikowalna (a więc bezsensowna wedle [PRÓBY 2]).

Być może należy więc połączyć [PRÓBĘ 1] z [PRÓBĄ 2] w sposób następujący?

[PRÓBA 3] Wyrażenie jest sensowne poznawczo wtedy i tylko wtedy, gdy jest ostatecznie weryfikowalne lub ostatecznie falsyfikowalne.

Niestety, ten ruch nie pomoże za wiele, bo bezsensowne okażą się niektóre zdania z kwantyfikacją mieszaną. Jeżeli dziedzina superbohaterów jest nieskończona, to zdania typu:

[SUPERZDANIE] Każdy superbohater ma swojego arcywroga

nie będą sensowne w sensie [PRÓBY 3]. Zaczyna się od kwantyfikatora ogólnego, więc żeby je zweryfikować, musiałbym zaobserwować każdego superbohatera (a tych, założyliśmy, jest nieskończenie wielu). Zatem zdanie to jest nieweryfikowalne.

Żeby je sfalsyfikować, musiałbym znaleźć jakiegoś superbohatera, powiedzmy człowieka wiertarkę, i sfalsyfikować, co następuje:

Człowiek wiertarka ma swojego arcywroga.

Trochę bardziej merytorycznie

Aby to zdanie sfalsyfikować, to zweryfikować muszę tak naprawdę zdanie „Nie istnieje osoba, która jest arcywrogiem człowieka wiertarki”. W nieco nieładnym sformułowaniu, które jednak ukazuje strukturę logiczną tego zdania, widzimy, że jest to zdanie ogólne:

Dla każdej osoby, ta osoba nie jest arcywrogiem człowieka wiertarki.



To jednak znaczy, że zweryfikować muszę twierdzenie ogólne, które (przy założeniu nieskończoności dziedziny) weryfikowalne nie jest. [SUPERZDANIE] nie jest więc ani ostatecznie weryfikowalne, ani ostatecznie falsyfikowalne, a zatem jest poznawczo bezsensowne w świetle [PRÓBY 3]. Problem okazuje się jeszcze bardziej palący, gdy zauważymy, że dużo zdań z dziedziny fizyki ma tę samą formę logiczną, np. „Każda ciecz ma temperaturę wrzenia”.

Drogi i wyboje weryfikacjonizmu

Być może problem bierze się stąd, że oczekujemy **silnego potwierdzenia** – chcemy, by dla sensowności zdania jego prawdziwość lub fałszywość była **ostatecznie** rozstrzygnięta przez dane zmysłowe. Być może zamiast tego wystarczyłoby, gdybyśmy za sensowne uznawali takie zdania, które **robią obserwowalną różnicę**, niezależnie od tego, czy są ostatecznie potwierdzalne (lub obalalne) przez zdania obserwacyjne.

Wyzwanie teraz polega na tym:

[WYZWANIE] Musimy doprecyzować pojęcie robienia różnicy w taki sposób, który z jednej strony uchwyci intuicyjną treść tego pojęcia, z drugiej doprowadzi do uznania zdań intuicyjnie niewątpliwie sensownych za oficjalnie sensowne, a z trzeciej (jak mawiają hegliści, każdy kij ma trzy końce) nie prowadzi do uznania zdań intuicyjnie podejrzanych i bezsensownych za sensowne.

Podążając za Ayerem (kolegą Carnapa), spróbujmy następującego sposobu:

[PRÓBA 4] Wyrażenie p jest sensowne poznawczo wtedy i tylko wtedy, gdy p w koniunkcji z pewnymi innymi przesłankami $r_1, \dots, r_k$ pociąga za sobą logicznie pewne zdanie obserwacyjne q , które nie wynika z $r_1, \dots, r_k$.

Określenie to dopuszcza sytuację, w której $k = 0$. W takim przypadku po prostu rozważamy, czy p samo pociąga za sobą jakieś zdanie obserwacyjne.

Problem polega na tym, że [PRÓBA 4] nie stawia żadnych wymogów co do przesłanek pomocniczych $r_1, \dots, r_k$. To szybko prowadzi do trudności. Wystarczy bowiem, że r_1 stwierdza „Jeżeli p , to Rafał Urbaniak ma 132 zęby”. W takim przypadku samo r_1 nie prowadzi do wniosku, że Rafał Urbaniak ma 132 zęby, ale p razem z r_1 do tego wniosku prowadzi. I to niezależnie od p : dzieje się tak, nawet gdy p to „Istnieją niewykrywalne

Trochę bardziej merytorycznie

i niewidzialne latające żyrafy”. To zaś dyskwalifikuje [PRÓBĘ 4] jako kryterium sensowności.

Zasadniczo wiemy, skąd wziął się ten absurd: z użycia dziwnego r_1 , które samo nie było wybitnie poznawczo sensowne. Moglibyśmy więc próbować to naprawić, mówiąc:

[PRÓBA 5] Wyrażenie p jest sensowne poznawczo wtedy i tylko wtedy, gdy p w koniunkcji z pewnymi innymi **poznawczo sensownymi** przesłankami $r_1, \dots, r_k$ pociąga za sobą logicznie pewne zdanie obserwacyjne q , które nie wynika z $r_1, \dots, r_k$.

Być może pozwoliliby to uniknąć przesłanek pomocniczych, takich jak nasze r_1 . Trudność jednak polega na tym, że [PRÓBA 5] nie może służyć za definicję sensowności poznawczej, bo sama posługuje się pojęciem sensowności poznawczej! Żeby powiedzieć, że p jest sensowne poznawczo, musimy przecież już być w stanie ustalić, czy poznawczo sensowne są przesłanki pomocnicze!

Ayer podjął się też próby definicji niekolistej. Udało mu się podać jakąś odmianę [PRÓBY 5], która unikała zarzutu kolistości. Niemniej natrafiała ona na szereg innych trudności technicznych, których wyjaśnienie (choćby nieskomplikowane) zajęłoby trochę miejsca, które niestety w tej książce muszą ograniczać³⁸. Ważne jest to, że Ayerowi nie udało się z powodzeniem poradzić sobie z [WYZWANIEM].

Zasadniczo więc prace nad doprecyzowaniem relacji weryfikacji czy falsyfikacji, silnej czy słabej, jako próby podania kryterium sensowności porzucono. Spróbowano natomiast innego podejścia: zamiast mówić o potwierdzaniu, zaczęto mówić o przetłumaczalności:

[PRÓBA 6] Zdanie jest sensowne, jeżeli jest przetłumaczalne na język empiryczny.

³⁸ Warto jednak po szczegóły sięgnąć do rozdziału 13 pierwszego tomu *Philosophical Analysis in the Twentieth Century* Scotta Soamesa.

Drogi i wyboje weryfikacjonizmu

Przez taki język rozumiano język, który u podstaw posiada tylko logiczny predykat identyczności oraz predykaty wyrażające własności obserwowalne. Następnie wszystko, co można wyrazić za pomocą takich predykatów, standardowych spójników i kwantyfikacji, uznawano za sensowne.

[PRÓBA 6] radzi sobie z problemami, które napotykały próby wcześniejsze. Zdania ogólnotwierdzące mogą być nieweryfikowalne, ale mimo to będą sensowne, jeżeli sformułowane zostaną za pomocą empirycystycznie akceptowalnych predykatów (i analogicznie problem nie pojawia się dla innych typów zdań).

Za darmo dostajemy wniosek, że jeżeli dane wyrażenie jest sensowne, to sensowna jest też jego negacja oraz sensowne jest dowolne zdanie złożone ze zdań sensownych (a nie jest to wniosek, który łatwo było uzyskać przy próbach wcześniejszych).

Pytanie, jakie się pojawia, to czy [PRÓBA 6] nie wyklucza zbyt wielu zdań intuicyjnych. W szczególności okazało się, że istnieją całkiem poważne trudności w empirycystycznym przekładzie **TERMINÓW DYSPOZYCYJNYCH**, takich jak „łatwopalny”, „kruchy” czy „łamliwy”. Jeżeli bowiem spróbujemy powiedzieć:

x jest łamliwe wtedy i tylko wtedy, gdy (jeżeli x jest poddane danej niewielkiej sile y, to x się łamie)

mamy ten problem, że posłużyliśmy się po prawej stronie implikacją materialną, której prawdziwość wynika z fałszywości poprzednika lub prawdziwości następnika. To zaś znaczy, że o dowolnym przedmiocie, na który nigdy nie zadziałała siła y, będzie prawdą, że jeżeli jest poddany sile y, to się łamie (to samo będzie prawdą o dowolnym przedmiocie, który się łamie).

Na przykład założymy, że kręgosłup predatora nigdy nie został poddany sile y. Wtedy następujące rozumowanie byłoby poprawne:

Kręgosłup predatora nigdy nie został poddany sile y.

Jeżeli zatem kręgosłup predatora jest poddany sile y, to się łamie.

Kręgosłup predatora jest zatem łamliwy.

Trochę bardziej merytorycznie

Wniosek jest jednak fałszywy, wiemy bowiem niezależnie, że kręgosłup predatora jest mocniejszy od w miarę niełamliwego kręgosłupa Chucka Norrisa.



Problem być może polega na tym, że powinniśmy się posłużyć jakimś kontrfaktycznym okresem warunkowym języka potocznego, a nie implikacją materialną (por. uwagi w rozdziale 38). Sprawa nie jest do końca jasna. Nikt tego dotąd nie rozważył, bo w czasach, gdy było to problemem, implikacja materialna była w sumie jedynym narzędziem na rynku, a w czasach, gdy pojawiły się lepsze narzędzia, ludzie mieli inne problemy. To może też znaczyć, że ktoś to rozważył, ale ja w swojej ograniczoności o tym nie wiem³⁹.

W międzyczasie prace PIERRE'A DUHEMA (1861–1916), francuskiego fizyka i filozofa nauki, przekonały większość filozofów nauki, że nie ma czegoś takiego, jak weryfikacja czy falsyfikacja zdań wziętych w izolacji od uznawanej przez

³⁹ Pewne pesymistyczne uwagi na temat rozwoju tej strategii znaleźć można w 13 rozdziale Soamesa, ale nie uważam, żeby były powalająco przekonujące.

Drogi i wyboje weryfikacjonizmu

nas teorii. Na przykład gdy dokonujemy obserwacji za pomocą teleskopu, to weryfikacji czy falsyfikacji nie ulega zdanie obserwacyjne osobno wzięte (jak „Na Słońcu nie ma plam”). Raczej gdy obserwujemy Słońce (powiedzmy w XVII wieku) przez teleskop i rzeczywiście widzimy jakieś plamy, to możemy albo uznać nasze zdanie obserwacyjne za sfalsyfikowane, albo uznać, że plamy są raczej efektem wady teleskopu.

Ogólnie rzecz biorąc, w większości przypadków pomiarów i obserwacji w fizyce istotną rolę odgrywa cała teoria, która stoi za naszym przekonaniem, że techniki pomiaru czy obserwacji, którymi się posługujemy, są wiarygodne. I to ta cała teoria w koniunkcji z naszym zdaniem obserwacyjnym i zdaniami opisującymi okoliczności razem wzięte są przedmiotem naszej oceny. Teorie fizyczne zatem, uważał Duhem, mogą być jedynie testowane i oceniane jako całości, a nie jako zlepki osobno testowanych poszczególnych sądów.

Stanowisko tego typu być może da się uogólnić. Dlaczego tylko teleskopy czy encefalografy mają być podejrzane? Przecież za wiarygodnością obserwacji naocznej stoi cała teoria dotycząca tego, jak funkcjonuje nasz układ nerwowy i optyka gałek ocznych, a także cały opis warunków naszego eksperymentu czy naszej obserwacji. Tego typu uogólnienia poglądu Duhema na wszelkie teorie i wszelką wiedzę głosił WILLARD VAN ORMAN QUINE⁴⁰. Do jakiego stopnia miał rację, jest kwestią osobną, ale w jego tezach było coś na rzeczy. Przynajmniej tyle, żeby teza o tym, że poszczególne zdania nie mogą być testowane w izolacji od teorii, stała się popularna i zyskała miano **TEZY DUHEMA-QUINE’A**.

W każdym razie [PRÓBĘ 6] przestano uważać za kryterium ostateczne i przeniesiono ciężar znaczenia na całe teorie. Zamiast mówić o sensowności czy bezsensowności poznawczej zdań w izolacji, skupiono się na ocenie poznawczej wartości teorii

⁴⁰ Willard Van Orman Quine (1908–2000) zajmował się filozofią analityczną i logiką. Znany jest ze swojej krytyki rozróżnienia na zdania syntetyczne i analityczne oraz z holizmu metodologicznego, wedle którego teorie powinny być oceniane tylko w całości.

Trochę bardziej merytorycznie

jako całości. Teraz pytaniem było: Czy cała teoria robi jakąś empiryczną różnicę? Czy jest jakaś obserwacja, która by ją podważała albo potwierdzała? Jeżeli nie, to teoria, nawet jeżeli intuicyjnie posiada znaczenie, nie jest wartościowa poznawczo jako teoria o świecie. I to podejście do **EMPIRYCZNEJ WARTOŚCI POZNAWCZEJ** nie zostało podważone.

Niezależnie od problemów z eksplikacją pojęcia sensowności motywacja dla przejścia od wątpienia w sensowność pewnych zdań do rozważania po prostu ich wartości poznawczej jest następująca: żeby móc ocenić, czy istnieje jakiś sposób weryfikacji danego zdania, już musimy dysponować jakimś rozumieniem tego zdania. W przeciwnym przypadku jak moglibyśmy wiedzieć, czy coś kwalifikuje się jako jego weryfikacja lub falsyfikacja? Jest zasadnicza różnica pomiędzy „Istnieją niewidzialne i niewykrywalne żyrafy” a „Humwumumwy pląplują bleblacko”. W pierwszym przypadku wiemy, co zdanie głosi, i wiemy, że nie mamy metody ustalenia prawdziwości tego zdania. W przypadku drugim nie jesteśmy w stanie ustalić, czy zdanie jest weryfikowalne, bo nawet nie wiemy, co ono głosi. Nawet nie jesteśmy przekonani, że jest zdaniem. Tego typu różnica nie została jednak zbyt dobrze uchwycona przez propozycje neopozytywistyczne.

Wobec wszelkich neopozytywistycznych prób podania kryteriów sensowności wysunąć można też zarzut, który (jak słyszałem od kilku ludzi, ale zdanie to chyba jest nieweryfikowalne) zgłosił Roman Ingarden na kongresie w Pradze w latach 20. XX wieku. A nawet jeżeli to nie był on, to i tak warto zarzut rozważyć. Neopozytywiści głosili, że zdania sensowne albo są analityczne (analitycznie fałszywe lub analitycznie prawdziwe), albo spełniają pewien (zależny od wariantu, o którym mowa) warunek empirycznej sensowności. Należy więc postawić pytanie: Czy sam ten pogląd jest analityczny, czy też raczej spełnia warunek empirycznej sensowności?

Większość krytyków posługujących się tym chwytem raczej go nie przemysłała i zadowala się zrobieniem cwanej miny

Drogi i wyboje weryfikacjonizmu

po ogłoszeniu tego pytania. Oczywiście nie jest to strategia zwycięska, bo w rzeczywistości powinni oni również wyjaśnić, dlaczego każda z dwóch możliwych odpowiedzi miałaby być niezadowolająca.

Założmy roboczo, że empirycystyczne kryterium znaczenia samo nie jest empirycystyczne. Co z tego wynika? Albo to kryterium samo jest pozbawione znaczenia, albo jest analityczne. Dlaczego świat byłby zły, a neopozytywiści głupi, gdyby było pozbawione znaczenia? W ich rozumieniu pozbawione znaczenia poznawczego są wszelkie dyrektywy mówiące, że należy robić to, a nie tamto. Być może więc kryterium jest tego typu dyrektywą: **Za poznawczo sensowne i informatywne należy uznawać tylko takie zdania, które...**

A dlaczego miałaby być wykluczona opcja druga, dlaczego nie miałoby być to zdanie analityczne? Pewnie dlatego, że nie jest oczywiste! Co prawda można próbować oddalać ten zarzut, mówiąc, że zdania analityczne wcale nie muszą być oczywiste, bo zasada przejrzystości jest fałszywa. Nieprzejrzystość znaczenia ma jednak swoje granice: jeżeli jakieś rzekomo analityczne zdanie ewidentnie koliduje z pewnymi zdaniami, co do których mamy silne *a priori* intuicje, to szansa, że jest analityczne, jest raczej niska. Ale też neopozytywiści, próbując sformułować kryterium, zmieniali zdanie w obliczu przeciwnych ich stanowisku intuicji, aż do momentu, kiedy główne stanowisko przestało jawnie kolidować z prostymi intuicjami. Faktycznie trudno o przekonujące intuicyjne przykłady zdań sensownych, które nie są przekładalne na język empiryczny, czy które są przekładalne, ale bezsensowne.

Co z trzecią możliwością, wedle której kryterium sensowności jest hipotezą empiryczną? Główny problem jest taki, że społeczność użytkowników języka polskiego nie zgadza się co do sensowności pewnych wybranych zdań. W kwestiach filozoficznie interesujących intuicje przeciętnych użytkowników języka potocznego są równie pomocne, co siodło do ujeżdżania żyrafy w momencie, kiedy chce się jechać na krokodylu. Postulowanie więc empi-

Trochę bardziej merytorycznie

ryczności empirycznego kryterium sensowności wydaje się ryzykowne. Wesprzeć bowiem takie twierdzenie, rozumiane jako hipoteza empiryczna, należałoby empirycznymi badaniami na użytkownikach języka. Ale również nie jest to opcja wykluczona.

Podsumowując, zarzut z samoodniesienia kryterium nie jest powalający i pozostawia neopozytywistom pole manewru. Stanowi raczej wezwanie do doprecyzowania statusu ich tez. Ale też i Carnap sformułował pewne uwagi na temat metod filozofii, które bardzo dobrze stosują się do tej kwestii.

Jedną z metod, którą powinien posługiwać się filozof, jest zdaniem Carnapa technika **EKSPLIKACJI**. Eksplikacja polega na zastępowaniu rozmytego lub niejasnego pojęcia z języka potocznego pojęciem precyzyjniejszym, lepiej skonstruowanym i bardziej przydatnym. Na przykład niejasne pojęcie ciepła zostało w pewnym sensie w dyskursie naukowym zastąpione pojęciem temperatury. Zabieg eksplikacji jest więc do pewnego stopnia pragmatyczny: postulujemy, że lepiej będzie, jeżeli jakieś pojęcie zastąpimy innym; proponujemy zmianę naszego języka na bardziej precyzyjny przy zachowaniu istotnych aspektów pojęcia wyjściowego.

Z tej perspektywy formułowanie kryterium sensowności można postrzegać jako eksplikację: próbę zastąpienia niejasnego pojęcia sensowności precyzyjniejszym pojęciem sensowności poznawczej. W tym sensie jest to kryterium pozbawione sensu; to jednak nie stanowi większego zarzutu, bo z jednej strony sensowne jest już pytanie, czy pożytecznie jest to kryterium przyjąć (tak samo jak sensowne jest pytanie, czy warto mówić o temperaturze raczej niż o cieple), a z drugiej nie wszystko, co nie jest sensowne poznawczo, jest bezwartościowe. Wiele rzeczy bezsensownych poznawczo jest pożytecznych, jak np. robienie kanapek czy spacer do sklepu monopolowego. Zarzut Carnapa wobec metafizyki polegał nie na tym, że jest ona bezsensowna poznawczo, tylko na tym, że jest bezsensowna poznawczo **i udaje, że nie jest**.

Nie udało się ostatecznie wyeksplikować pojęć przekładalności empirycznej czy empirycznej doniosłości teorii. Musimy jed-

Drogi i wyboje weryfikacjonizmu

nak pamiętać, że niekoniecznie należy to traktować jako zarzut ostateczny wskazujący na to, że nie ma sensu mówić o takich czynnikach. W sumie nie udało się ostatecznie wyeksplikować pojęcia ostatecznej eksplikacji i (mam nadzieję) nikt się tym nie przejmuje. Niemożliwość dokładnej eksplikacji pojęcia osoby upośledzonej nie dowodzi nieistnienia osób upośledzonych. Niemożliwość ścisłej eksplikacji pojęcia pornografii nie dowodzi nieistnienia pornografii. Podobnie niemożliwość ścisłego zdefiniowania empirycznej doniosłości teorii nie dowodzi nieistnienia zjawiska empirycznej doniosłości. Jeżeli jakiejś teorii filozoficznej stawiamy wymogi z sufitu („Zdefiniuj dokładnie wszystkie pojęcia!”), to nie jest jasne, czy jeżeli ona tych wymogów nie spełnia, nie jest to problem raczej z wymogiem niż z teorią. Wymóg ścisłej, jednoznacznej i bezbłędnej eksplikacji wszędzie, gdzie się da, jest po prostu trochę przereklamowany (w końcu każda teoria ma pewne terminy pierwotne)⁴¹.

TEKSTY POLECANE



1. Poglądy Carnapa na temat trybu formalnego, zdań pseudoprzedmiotowych oraz sensowności zdań znaleźć można w jego *Filozofii i składni logicznej* oraz *Sprawdzalności i znaczeniu*, obie rzeczy wydane są po polsku w tłumaczeniu Andrzeja Zabłudowskiego. *Wprowadzenie do filozofii nauki* i *Pisma semantyczne* autorstwa Carnapa opublikowane zostały przez wydawnictwo Aletheia.
2. Z podejściem Carnapa do rozróżnienia pomiędzy pytaniami wewnętrznymi i zewnętrznymi zapoznać się można, czytając jego *Empiricism, Semantics and Ontology*. Tekst ukazał się w *Revue Internationale de Philosophie*, 4 (1950): 20–40, ale łatwo go znaleźć on-line, googlując.

⁴¹ W szczególności istnieją żyjący i racjonalni filozofowie, którzy ogólnie rzecz biorąc zgadzają się z empirycystycznym kryterium wartości poznawczej. Na przykład na rodzimym gruncie Bohdan Chwedeńczuk broni braku wartości poznawczej przekonań religijnych. Polecam jego *Przekonania religijne* jako całkiem prowokującą lekturę.

Trochę bardziej merytorycznie

3. Ważne próby sformułowania słabego warunku weryfikowalności znaleźć można u Ayera w jego *Language, Truth and Logic*, klasycznym manifestie neopozytywizmu.
4. Pierwszy tom *Philosophical Analysis in the Twentieth Century* Scotta Soamesa zawiera wiele fragmentów dotyczących interesujących mnie tutaj kwestii. W szczególności doniosły jest rozdział 13.
5. Więcej o zdaniach obserwacyjnych można przeczytać w tomie *Spór o zdania protokolarne* w tłumaczeniu Leszka Kopciucha.
6. Jedną z najnowszych antologii zawierających współczesne dyskusje dotyczące werbalności dysput filozoficznych jest *Metametaphysics. New Essays on the Foundations of Ontology*, zredagowane przez Chalmersa, Manleya i Wassermana (wydane przez Oxford University Press w 2009 r.).
7. Dobrym przykładem współczesnych zastosowań filozoficznych empirycystycznego podejścia do sensowności jest książka *Przekonania religijne* Bohdana Chwedeńczuka (Wyd. Aletheia, 2000).



DROGI I WYBOJE WERYFIKACJONIZMU

1. Ogólna strategia: zdania analityczne (lub analitycznie fałszywe) są sensowne i sensowne są te, które pozostają w odpowiedniej relacji do empirycznie rozstrzygalnych zdań obserwacyjnych.
2. Pierwszy problem: natura zdań obserwacyjnych.
3. Są one rozstrzygalne, tylko jeżeli dotyczą jedynie danych zmysłowych, ale wtedy nie są intersubiektywne.
4. Jeżeli są intersubiektywne, nie są ostatecznie rozstrzygalne.
5. Rozwiązanie tymczasowe: o przedmiotach, porzućmy **ostateczną** weryfikowalność i załóżmy, że sprawa jest raczej jasna.
6. Jaka relacja do zdań obserwacyjnych nadaje znaczenia zdaniu?
7. Weryfikacjonizm zawodzi w przypadku ogólnych zdań o dziedzinach nieskończonych.
8. Falsyfikacjonizm zawodzi w przypadku zdań szczegółowych (tj. w przypadku negacji zdań weryfikowalnych) o dziedzinach nieskończonych.
9. Wery- lub falsyfikacjonizm zawodzi w przypadku kwantyfikatorów mieszanych.
10. Próby określenia sensowności zdań w kategoriach ich relacji do całych teorii też zawodzą (Ayer).

Zastosowanie: religia

11. Przetłumaczalność na język empiryczny: trudności z terminami dyspozycyjnymi.
12. Trudność ogólna: teza Duhema-Quine'a (nie istnieje ostateczna weryfikacja/falsyfikacja, wszystkie oceny zdań zależą od przyjętej teorii).
13. To nie dowodzi bezsensowności pojęcia empirycznej doniosłości (jeżeli oddzielimy ją od kryterium sensowności); jedynie trudności w eksplikacji i rozmytości.
14. Zarzut z samozwrotności nie działa, można np. mówić o eksplikacji, która zawiera czynniki pragmatyczne. Bezsensowne nie znaczy pragmatycznie bezwartościowe.
15. Interesujący rezultat: docenienie (w miarę racjonalnej) falsyfikowalności nie jako kryterium sensu, ale jako kryterium wartości teorii.

22. Zastosowanie: religia

Jeżeli teoria dotyczy dziedziny nieskończonej, może się zdarzyć, że poszczególne obserwacje, technicznie rzecz biorąc, nie zwiększają prawdopodobieństwa danej teorii. Ale w takiej sytuacji pewne możliwe obserwacje, nawet jeżeli nie prowadzą do ostatecznej falsyfikacji, powinny obniżać jej wiarygodność.

Przez jakiś czas teorii, zgodnie z którą **wartościowa poznawczo** teoria powinna być falsyfikowalna, całkiem przekonująco bronił Karl Popper (dla zmylenia czytelnika pogląd Poppera nazywa się więc **FALSYFIKACJONIZMEM**). Problemy wiążące się z tezą Duhema-Quine'a doprowadziły do pewnego osłabienia falsyfikacjonizmu. Teoria empiryczna, aby być poznawczo wartościowa, powinna dopuszczać obserwacje, które (nawet jeżeli nie falsyfikują teorii całkowicie) mogą podważać wiarygodność tej teorii lub ją wspierać. Teorie, które w tym sensie nie wykluczają żadnego toku wydarzeń, nie mówią nic o świecie.

Trochę bardziej merytorycznie

Technika polemiki filozoficznej opierająca się na takich intuicjach nabrała impetu w pierwszej połowie XX wieku. Jedną z głównych faz rozwoju analitycznej filozofii religii to trwające do lat 60. debaty dotyczące sensowności języka religii i jej wartości poznawczej. Jednym z pierwszych myślicieli reprezentujących ten nurt był Ayer.

Zarzuty bezsensowności odeszły w zapomnienie wraz z empiryczną doniosłością jako kryterium znaczenia. Dyskusja przeszła w fazę, gdy dopuszczano jakąś sensowność języka religii, ale wątpiono w to, czy zawiera ona jakiegokolwiek treści empiryczne.

I tak Antony Flew w latach 50. opowiadał następującą przypiastkę (nie przedstawiam jej dosłownie):

„Dwaj podróżnicy trafiają w dżungli na coś, co wygląda jak zaniedbany ogród, w którym jest dużo kwiatów i chwastów. Jeden z nich twierdzi: „Gdzieś musi być ogrodnik, który się tym ogrodem zajmuje”. Drugi się z nim nie zgadza. Rozbijają więc obóz i czekają na ogrodnika, trzymając wartę. Nigdy go jednak nie widzą. Pierwszy podróżnik mówi więc: „Być może jest to niewidzialny ogrodnik!”. Ustawiają więc wokół ogrodu zasieki z druku kolczastego i ogrodzenie elektryczne (nie mam pojęcia, skąd podróżnicy w dżungli mieli tyle drutu kolczastego oraz płot i skąd wzięli prąd). Patrolują teren z psami (nie wiem, skąd wzięli psy). I czekają. Dalej nic. Pierwszy podróżnik dodaje: „Ten ogrodnik nie tylko jest niewidzialny, lecz także niewykrywalny i nieobserwowalny; kocha jednak swój ogród i o niego dba!”. Drugi zaś pyta: „Czym różni się niewidzialny i niewykrywalny ogrodnik kochający zaniedbany ogród od ogrodnika nieistniejącego?”.

Analogicznie niefalsyfikowalne wydają się pewne warianty teizmu – przynajmniej w takiej formie, w jakiej wyznaje go wielu ludzi. Odpowiednio nawiedzony ksiądz na pytanie o celowość dowolnego przykładu cierpienia, które nam wydaje się niesprawiedliwe i bezsensowne, odpowie po prostu, że

Zastosowanie: religia

zamiary Boga są niezbadane albo że na pewno ma on Plan, który pozostaje dla nas niepojęty.

Problem z tą strategią jest taki, że broniąc teizmu przed zarzutem z istnienia zła, bardzo łatwo znaleźć się w sytuacji, w której nic nigdy nie będzie w stanie sfalsyfikować teizmu. Z jednej strony fajnie, bo nikt nie podważy naszych poglądów. Z drugiej (rzadziej spotykanej) strony w tej sytuacji nasza teoria przestanie odgrywać jakąkolwiek poznawczą rolę empiryczną, bo nie będziemy nawet w stanie opisać możliwych obserwacji, które skłoniłyby nas do osłabienia naszej wiary.

Być może nic nie szkodzi. Być może istotą religii jest jej aspekt pozapoznawczy i nie jest ona nawet w zamiarze teorią empiryczną o świecie, tylko jakąś formą pozapoznawczej postawy względem świata. Takie pomysły na przykład czasami wyrażał Richard M. Hare. Jeżeli tak, teista nie ma się czym martwić, o ile wprost przyzna, że taki jest status jego przekonań⁴².

Wydaje się jednak, że większość teistów, głosząc, że Bóg istnieje, myśli, że rzeczywiście głosi, że Bóg istnieje. To znaczy przypisują swojemu teizmowi znaczenie faktywne, uważają, że teizm jest dosłownie prawdziwy i przypisują mu rolę eksplanacyjną. Wyzwanie, jakie stawia takim teistom Flew, jest następujące: **Jeżeli twoja teoria jest o świecie, to jakie obserwowalne sytuacje wyklucza?**

Teista tego typu powinien więc raczej rzeczywiście zgodzić się z tym, że istnienie zła stanowi świadectwo przeciwko istnieniu Boga, lecz także i świadectwo nierozstrzygające i nieostateczne, a może nawet i świadectwo przewyżnione przez pozytywne argumenty za teizmem. Na przykład Basil Mitchell, by wyjaśnić to podejście, posługuje się następującą przypowieścią:

42 Ogólnie niekognitywizm teologiczny to ciekawy nurt: jeżeli bowiem zdania religii nie stwierdzają stanów rzeczy, co właściwie robią, jaki jest tego mechanizm i czy robią to słusznie?

Trochę bardziej merytorycznie

» W czasie wojny w okupowanym kraju partyzant poznaje jakiegoś faceta, który wywiera na nim duże wrażenie. Ów człowiek mówi, że jest w opozycji, lecz także tłumaczy przekonująco, że jest szpiegiem i że będzie widziany podczas wykonywania zadania dla wroga. W rzeczywistości jednak wszystko, co będzie robił, będzie miało na celu pokonanie okupanta. Prosi partyzanta, żeby mu zaufał, niezależnie od tego, co się będzie działo. Z upływem czasu osoba ta jest widziana, gdy raz robi coś dobrego, raz coś złego.

Koledzy partyzanta zaczynają, w świetle tego świadectwa, wątpić w dobre intencje szpiega. Nasz partyzant mimo negatywnego świadectwa może racjonalnie utrzymywać, że szpieg jest po ich stronie.

Jest jasne, że dla Mitchella ten szpieg to Bóg w osobie Jezusa. Niemniej, Mitchell dopuszcza sytuację, w której partyzant w świetle natłoku świadectwa negatywnego porzuca przekonanie o niewinności szpiega. Po prostu taki natłok się nie pojawił. Podobnie Mitchell uważa, że zło przemawia na niekorzyść teizmu, ale nie ma tego zła na tyle dużo, żeby uzasadnić porzucenie teizmu.

Trafność analogii z partyzantem, zauważa Flew, jest dyskusyjna, bo o ile w przypadku partyzanta można próbować wyjaśnić podejrzanе zachowanie szpiega, mówiąc, że jest to jedyna metoda, którą może osiągnąć swój cel, to w przypadku wszechmocnego i wszechwiedzącego Boga nie wchodziłoby to raczej w grę.

William Hasker zauważa, że Flew, wdając się w tę polemikę, milcząco odrzuca swoje własne wyjściowe twierdzenie o bezsensowności języka religii. I rzeczywiście, Flew wypowiada się tak, jakby rozważane twierdzenia rozumiał. Niemniej nie jest to aż tak szkodliwe, jak sugerowałby Hasker. Rzecz bowiem rozbija się bardziej o wartość poznawczą teizmu niż o jego sensowność. Flew może rozumieć teizm, nie uważając go za poznawczo wartościowy. Może rozumieć wyrazy wchodzące

Zastosowanie: religia

w skład opisu poglądu religijnego i rozumieć sposób, w jaki wyrazy te zebrane są razem w zdania, jednocześnie zaprzeczając, jakoby teoria składająca się z tych zdań cokolwiek pomagała zrozumieć lub wyjaśnić.

Mitchell, dopuszczając falsyfikację, po prostu zgadza się z Flewem i stara się odpowiedzieć na jego wyzwanie. W rzeczywistości jednak nie robi tego zbyt skutecznie: z samej analogii nie można wywnioskować, w którym momencie racjonalnie byłoby przestać wierzyć szpiegowi. Tak samo powiedzenie, że zło na świecie stanowi świadectwo przeciwko istnieniu Boga, tyle że świadectwo niewystarczające, nie wyjaśnia, na czym polegałoby świadectwo wystarczające i dlaczego ta granica leży tam, a nie gdzie indziej.

Bardziej poważnie wyzwanie Flewa traktują Ian M. Crombie i John Hick. Powiadają oni, że chrześcijanie mają pewne konkretne oczekiwania co do przyszłości. Oczekują bowiem zmartwychwstania i życia po śmierci. W tym sensie, zauważa Hick, chrześcijaństwo jest weryfikowalne (choć może nie-falsyfikowalne): wystarczy umrzeć. Jeżeli chrześcijaństwo jest prawdziwe, przekonamy się o tym po śmierci.

Nie jestem pewien, do jakiego stopnia można tu mówić o ostatecznej weryfikacji. Stosuje się tu bowiem teza Duhema-Quine'a. Gdybym miał wrażenie, że umarłem i na końcu świata stoję przed sądem ostatecznym, być może wcale nie musiałbym się zgodzić z chrześcijaństwem. Równie dobrze (albo nawet z większym prawdopodobieństwem) mógłbym wywnioskować, że wcale nie umarłem, tylko przedawkowałem substancje psychodeliczne. Ale tego typu problemy można odłożyć na bok w świetle następującego poważniejszego problemu.

Założmy, że ktoś głosi, że czarne dziury są kosmicznymi jednokierunkowymi tunelami do innych światów, że zbliżywszy się do takiej dziury, zostajemy wessani i przeniesieni do jakiegoś zasadniczo niedostępnego w inny sposób świata, z którego nie mamy jak wrócić.

Trochę bardziej merytorycznie

W pewnym sensie hipoteza ta jest zrozumiała i wiemy, co głosi. Ale teoretycznie jest wysoce niesatysfakcjonująca, o ile nie powie się nam, jak może być uzasadniana lub podważana przez obserwacje.

Nieskuteczne jest powiedzenie: „Przecież hipoteza jest weryfikowalna, jeżeli jest prawdziwa! Wystarczy się zbliżyć do takiej czarnej dziury i dać się wessać!”. Potencjalna obserwacja, o której mowa, łamie wymóg intersubiektywności: nie można jej dokonać, a potem poinformować środowiska naukowego, które mogłoby eksperyment powtórzyć i o nim podyskutować. Zamiast więc brać taką hipotezę na poważnie, należy zapytać, jakie ma ona związki z obserwacjami, których faktycznie można dokonać.

Sytuacja weryfikacji eschatologicznej, obserwacji zbawienia jest analogiczna. Sama jej możliwość nie stawia teizmu w metodologicznie pozytywnym świetle.

Nie chcę wchodzić w rozległy i osobny temat tzw. problemu teodycei, na który natrafiają wszyscy, którzy wierzą w osobowego, wszechmocnego, wszechwiedzącego i dobrego Boga, gdy tylko uznają istnienie zła w świecie i chcą to istnienie zła w jakiś sposób z istnieniem tak rozumianego Boga pogodzić. Poprzestanę tylko na uwadze, że nie spotkałem się z udanym sformułowaniem warunku falsyfikacji (albo niemal ostatecznie przekonującej falsyfikacji) teizmu ze strony tych teistów, którzy uważają religię za zasadniczo falsyfikowalną, lecz także uważają zarazem, że taka ilość zła na świecie, jaka jest, jest usprawiedliwiona.

Albo unikają oni teodycei, popadając w kompletną impregnację wierzeń przed faktami, albo nie są przekonujący w wyjaśnianiu, dlaczego obecne zło nie spełnia jeszcze warunku falsyfikacji ich przekonań. To oczywiście nie dowodzi, że nie mają racji, a jedynie wskazuje na pewne początkowo nieoczywiste wyzwanie: godząc istnienie zła na świecie z istnieniem dobrego, wszechmocnego i wszechwiedzącego Boga, należy nie tylko usprawiedliwić istnienie zła, lecz także upewnić się, czy nasza

Zastosowanie: religia

metoda usprawiedliwiania nie jest taka, że usprawiedliwiłaby wszystko, co popadnie.

Można oczywiście nalegać, że istnienie zła nie stanowi argumentu przeciwko istnieniu Boga, bo nie wyklucza istnienia jakiejś nieznaney nam tajemniczej racji, dla której zło istnieje, a która uniesprzeczniałaby istnienie Boga z istnieniem zła. Nawet jeżeli zgodzimy się, że istnienie zła nie wyklucza się **logicznie** z istnieniem wszechmocnego, wszechwiedzącego i nieskończenie dobrego bytu, to taka odpowiedź nie wydaje się satysfakcjonująca.

Po pierwsze, stawia ona bardzo wysoki wymóg argumentom filozoficznym. Zakłada bowiem, że jeżeli nie mogę wykluczyć najdzikszego i najbardziej tajemniczego scenariusza, w którym moje przesłanki są prawdziwe, a wniosek fałszywy, to mój argument jest nieskuteczny. Jest to jednak dokładnie ten sam wymóg, który prowadzi do pewnych skrajnych wariantów sceptycyzmu, który głosi, że dopóki nie mogę wykluczyć scenariuszy, w których jestem zwodzony przez złego demona lub w których jestem mózgiem w próbówce, tak naprawdę nie posiadam wiedzy.

Gdy ktoś przyjmuje tego typu standardy, nie istnieje chyba skuteczny sposób odwiedzenia go od tego. Warto jednak zauważyć, że przyjęcie tego typu wymogu sprawia również, że nieprzekonujące stają się wszelkie argumenty za istnieniem Boga – nie wykluczają bowiem przekonująco rozmaitych, chociażby i najdzikszych scenariuszy, przy których przesłanki są prawdziwe, a wniosek fałszywy.

Po drugie, dość interesujące wydaje się, że niektórzy filozofowie, którzy próbują podawać racje za istnieniem Boga, w odpowiedzi na problem zła mówią: „Mamy tu do czynienia z tajemnicą, która przekracza ludzki rozum”. Pomińmy już rozważenie, czy stanowi to rzeczywiście wartościową poznawczo teorię czegokolwiek. Zauważmy zamiast tego, że w ten sposób, wydaje się, dopuszczają istnienie stanów rzeczy (przynajmniej dla nas) absurdalnych i niezrozumiałych, z drugiej jednak strony, gdy przychodzi do argumentacji za istnieniem Boga, gdy posługują się argumentem kosmologicznym („Bóg istnieje, bo

Trochę bardziej merytorycznie

świat musi mieć przyczynę”), powiadają, że nie mogą istnieć nieskończone ciągi przyczynowe i że przyczynowość nie może ciągnąć się wstecz w nieskończoność, bo takie ciągi byłyby absurdalne i niezrozumiałe.

Oczywiście, istnieje też strategia, wedle której kwestie religijne w ogóle nie podlegają ocenie rozumu. Zgodnie z tym stanowiskiem, religia jest „ponadracjonalna” (cokolwiek by to znaczyło) i dyskusja o racjonalnych argumentach za lub przeciw religii jest bezprzedmiotowa. O ile zajmowanie tego stanowiska może sprawiać pewną przyjemność i zapewniać pewną ekscytację osobnikom o dość poetyckiej mentalności, o tyle wygląda na to, że nie jest to najzdrowsza strategia przyjmowania przekonań. Jeżeli bowiem każdy ma prawo przyjmować dowolne przekonania religijne i nie musi ich bronić w żaden racjonalny sposób, to znaczyłoby to, że każdy ma prawo postępować wedle nakazów religii (i przywódców religijnych), którą sobie bez żadnego dobrego powodu wybrał. I jeżeli taka strategia jest trafna, nie ma sposobu, w który można by kogoś racjonalnie przekonać, że jednak myli się, przyjmując taki, a nie inny pogląd religijny lub quasi-religijny, który nakazuje mu, powiedzmy, zabijać innych, traktować swoją żonę jak wyposażenie gospodarstwa domowego albo budzić całą dzielnicę modłami puszczanymi przez głośniki o piątej rano.

23. Zastosowanie: egoizm psychologiczny

EGOIZM PSYCHOLOGICZNY to pogląd, wedle którego cokolwiek ludzie robią, robią to z powodów egoistycznych. Jeżeli nawet jakieś zachowanie sprawia wrażenie zachowania altruistycznego, to głębsza analiza motywów podmiotu działania ukazuje nieco głębiej ukryte motywy egoistyczne. Jeżeli nawet człowiek

Zastosowanie: egoizm psychologiczny

chce zrealizować jakiś cel nieegoistyczny, to w rzeczywistości traktuje go jako środek do dalszego, egoistycznego celu.

Tę deskrypcyjną teorię, mówiącą, **jakimi motywami ludzie się rzeczywiście kierują**, należy odróżniać od **EGOIZMU ETYCZNEGO**, który jest teorią normatywną, głoszącą, że wszelkie działania **powinny** być wykonywane z motywów egoistycznych.

Egoista psychologiczny musi wytłumaczyć naturę pozornie altruistycznych działań. Jego zdaniem działania takie podejmowane są z powodów całkiem prozaicznych i egoistycznych: podmiotowi działania sprawia ono przyjemność albo chce wyrzucić wrażenie na innych, albo przyjemnie mu się myśli o sobie jako o człowieku szlachetnym itd.

Jakimi rozważaniami można się kierować, uznając egoizm psychologiczny?

[MOTYWACJA 1] Wszelki motyw mojego działania jest moim motywem. Zawsze więc, kiedy działam, dążę do realizacji moich celów.

[MOTYWACJA 2] Zawsze gdy osiągam to, co chciałem, odczuwam przyjemność. Tak naprawdę zawsze chodzi mi o tę przyjemność.

[MOTYWACJA 3] Istnieją jasne przypadki, kiedy oszukujemy samych siebie, myśląc, że robimy coś dla dostojnego celu. Możliwe, że oszukujemy samych siebie zawsze, gdy myślimy, że robimy coś z powodów nieegoistycznych.

[MOTYWACJA 4] Gdy uczymy dzieci właściwego postępowania, głównym czynnikiem, który wykorzystujemy, jest to, że dążą one do przyjemności i unikają nieprzyjemności. Dzięki temu, posługując się nagrodami i karami, wyuczamy je stosowania się do społecznie przyjętych norm zachowania.

Trochę bardziej merytorycznie

[MOTYWACJA 1] jest nieprzekonująca. Z tego, że jakiś motyw jest moim motywem (w sensie: ja się nim kieruję), nie wynika nic na temat tego, jakiego rodzaju jest to motyw. A już z pewnością nie wynika, że mam na celu jedynie własną przyjemność. [MOTYWACJA 2] też nie jest trafna, bo myli skutek z celem. Zawsze po siłowni jestem zmęczony. Czy z tego wynika, że idę na tam **po to, żeby** czuć się zmęczonym? Z tego, że coś jest nieodłącznym skutkiem jakichś działań, nie wynika, że jest ich celem. Poza tym jest fałszem twierdzenie, że zawsze, gdy osiągamy to, czego chcemy, odczuwamy przyjemność. Czasem odczuwamy rozczarowanie, smutek, zmęczenie lub po prostu kaca.

Podstęp polega więc tutaj na tym, że z tego, iż z działań *prima facie* altruistycznych czerpiemy przyjemność, wnioskuje się, że to ta przyjemność jest rzeczywistym celem owych działań. Tego typu obraz sprawia, że motywy naszych działań wyglądają koliście: pragniemy przyjemności z realizacji naszego pragnienia. Jakiego pragnienia? Pragnienia realizacji pragnienia. Jakiego pragnienia? Pragnienia realizacji pragnienia realizacji... Bardziej wiarygodna wydaje się analiza, w której **najpierw** mamy jakiś cel, którym nie jest przyjemność z realizacji celu, a **potem** jest nam przyjemnie, gdy ten pierwszy cel zostaje zrealizowany.

Jeżeli wszystko, co [MOTYWACJA 3] sugeruje, to tylko możliwość powszechnego samooszukiwania się, to nie mamy się czym przejmować. Wiele różnych zdań może być prawdziwych na temat różnych rzeczy, ale z tego nie wynika nic na temat tego, czy rzeczywiście są prawdziwe. Powiedzenie, że możliwe jest powszechne samooszukiwanie się, to nie argument za powszechnym samooszukiwaniem się.

Jeżeli natomiast rzeczywiście ktoś stwierdza powszechne samooszukiwanie się, to zamiast zaufać takiemu wygłaszanemu z fotela twierdzeniu, należy zapytać o to, jakie dane psychologiczne za tym przemawiają i jakie jest uzasadnienie tego twierdzenia. A jeżeli [MOTYWACJA 3] ma odgrywać rolę argumentu za

Zastosowanie: egoizm psychologiczny

egoizmem, to egoista zapytany o uzasadnienie powszechnego samooszukiwania się nie może kuliście odpowiedzieć, że samooszukiwanie się jest powszechne, bo egoizm jest prawdziwy.

[MOTYWACJA 4] jest po prostu fałszywa. Dzieci, które dorastają, robiąc coś tylko ze strachu albo dla nagrody, nie są dobrze wychowywane. Raczej wygląda na to, że wychowanie dziecka wymaga rozmawiania z nim i uzasadniania decyzji na tyle, na ile tego typu kwestie dziecko może zrozumieć na danym etapie rozwoju.

Wiemy już zatem, że motywacje dla egoizmu psychologicznego są nieprzekonujące. To jeszcze nie dowodzi jego fałszywości. Ale właśnie w momencie, gdy zastanawiamy się, na czym polegałoby dowiedzenie fałszywości egoizmu psychologicznego, odkrywamy, na czym polega problem metodologiczny tej teorii.

Możemy próbować egoiście psychologicznemu podawać przykłady działań, które intuicyjnie uznajemy za nieegoistyczne. Egoista jednak pozostanie niewzruszony: egoistyczne motywy będzie przypisywał nawet męczennikom. To zaś sprawia, że jego teoria wygląda na nefalsyfikowalną. I rzeczywiście jest to teoria, która *a priori* przypisuje wszystkim działaniom ludzkim jeden motyw i dostarcza prostej sztuczki w stylu [MOTYWACJA 2], by dopasować opis dowolnego działania ludzkiego do teorii. Sama zaś teoria nie zostawia żadnego miejsca na przypadki, które by ją falsyfikowały.

Być może dałoby się sformułować jakiś interesujący poznawczo wariant egoizmu psychologicznego. Byłaby to jednak aposterioryczna teoria psychologiczna, która tezy swoje musiałaby opierać na rzeczywistych badaniach psychologicznych oraz (co kluczowe) na metodzie wykrywania „rzeczywistych motywów” niezależnej od uprzednio przyjętej teorii filozoficznej i niezależnej od samych zeznań podmiotów działania. Sformułowanie takiej metody i zastosowanie jej w tego typu szerokich badaniach, owocejących rezultatem zadowalającym egoistę psychologicznego, to zadanie, którego realizacja nie wydaje się zbyt prawdopodobna.

Trochę bardziej merytorycznie

Różne teorie przypisujące działaniom ludzkim jeden motyw nie tylko są zazwyczaj niefalsyfikowalne, lecz także przeczą sobie nawzajem (czy ja ostatecznie pragnę władzy, czy przyjemności?) oraz nie dostarczają przydatnego obrazu złożoności ludzkich motywów. Jeżeli wiem, że ktoś tak naprawdę chce przyjemności, to przez to, że przyjemność jest rozumiana tak szeroko, nie mam pojęcia, co on może zrobić i jakie decyzje podjąć. Informacja, że tak naprawdę chce on (w tym szerokim sensie) przyjemności, jest bezużyteczna. Żeby móc w ogóle w przybliżeniu zgadywać, jakie niektórzy ludzie będą podejmowali decyzje, muszę przyjrzeć się im na głębszym poziomie analizy. Na tym poziomie będziemy mówić o różnych poszczególnych pragnieniach i obawach, ale żadne z pożytecznych twierdzeń, które tam wystąpią, nie będzie dotyczyło tego jednego ostatecznego niefalsyfikowalnego motywu.

Co więcej, pogląd, jakoby jedynym i ostatecznym celem działań ludzkich było szczęście czy przyjemność, wcale nie jest taki oczywisty. Wyobraźcie sobie, że niczego nie pragniecie, nic was nie interesuje, niczego nie lubicie poza przyjemnością i szczęściem. Jakie macie szanse na osiągnięcie szczęścia, jeżeli nic poza szczęściem was nie cieszy? Wydaje się, że niewielkie. Jak powiada Joel Feinberg, przyjemność i szczęście mają tendencję do zakradania się do nas. Pojawiają się wtedy, gdy pochłonięci jesteśmy czymś innym, a najlepszą metodą uniknięcia szczęścia i przyjemności jest ograniczenie swoich pragnień jedynie do pragnienia szczęścia i przyjemności. Być może, aby osiągnąć szczęście, lepiej o nim zapomnieć.



TEKSTY POLECANE

1. Dobrym miejscem, by zacząć zapoznawanie się z analityczną filozofią religii, jest esej *Analytic Philosophy of Religion* Williama Haskera, znaleźć go można w *The Oxford Handbook of Philosophy of Religion*.

Zastosowanie: egoizm psychologiczny

2. Ciekawą antologią tekstów dotyczących falsyfikowalności teizmu jest *New Essays in Philosophical Theology* wydana w 1995 r., pod redakcją A. Flewa i A. MacIntyre'a.
3. Warto przeczytać też *Bóg a zło* K. Hubaczka.
4. Tekst dość dokładnie omawiający egoizm psychologiczny to *Psychological Egoism* Joela Feinberga (w antologii tekstów *Reason and Responsibility*, 2008).

KIERUNKOWSKAZ

Zajmiemy się teraz prawdopodobieństwem i jego filozofią. Najpierw przyjrzymy się trochę samej teorii prawdopodobieństwa, żebyśmy wiedzieli, o czym mówimy. Następnie zastanowimy się nad szeregiem interpretacji filozoficznych tej teorii i nad ich słabościami. Wreszcie przyjrzymy się bayesianizmowi, jednemu z obecnie ważniejszych nurtów w filozofii formalnej, która stosuje probabilistyczne metody modelowania siły przekonań do rozmaitych problemów filozoficznych.



CZĘŚĆ



Prawdopodo- bieństwo

24. Prawdopodobieństwo: podstawy

Przy okazji prób eksplikacji pojęcia potwierdzania częściowego teorii przez obserwacje, filozofowie nauki zainteresowali się teorią prawdopodobieństwa. Początkowo rozwijali ją tylko po to, żeby lepiej zrozumieć interakcje między teoriami a obserwacjami. Całkiem szybko jednak posługiwanie się formalnym aparatem teorii prawdopodobieństwa stało się dość powszechne wśród tych filozofów analitycznych, którzy w swoich rozważaniach stosują narzędzia formalne.

Obecnie nurt stosujący teorię prawdopodobieństwa w filozofii, zwłaszcza do epistemologii i analizy argumentów filozoficznych, jest jednym z głównych nurtów w filozofii analitycznej. Niestety dziedzina ta jest dość techniczna, więc trudno będzie mi w sposób przystępny w pełni opisać, co się w niej dzieje. Zamiast tego w sposób możliwie nieformalny postaram się wyjaśnić, skąd się wzięła, oraz wprowadzić aparaturę obecnie już zakładaną w dyskusjach w tej dziedzinie.

Zacznijmy od samych aksjomatów teorii prawdopodobieństwa w takiej postaci, w jakiej sformułowane zostały w roku 1933 przez Andrieja Kołmogorowa. Rozważymy wariant, w którym prawdopodobieństwo przypisywane jest zdaniom (a nie wydarzeniom i ich zbiorom). Prawdopodobieństwo dowolnego zdania A , $\Pr(A)$ musi spełniać następujące warunki⁴³:

- | | |
|--------------|---|
| [AKSJOMAT 1] | $\Pr(A) \geq 0$ dla dowolnego A . |
| [AKSJOMAT 2] | Jeżeli T jest tautologią logiki klasycznej, to $\Pr(T) = 1$. |
| [AKSJOMAT 3] | Jeżeli A i B wykluczają się logicznie, to $\Pr(A \vee B) = \Pr(A) + \Pr(B)$. |

43 Kołmogorow rozważał też przypadek przeliczalnie nieskończonej alternatywy zdań: wprowadził aksjomat, że dla zdań parami rozłącznych prawdopodobieństwo ich alternatywy jest równe sumie ich prawdopodobieństw. My jednak, w przeciwieństwie do matematyków, nie musimy się tym za bardzo przejmować.

Kluczowym pojęciem teorii prawdopodobieństwa jest pojęcie **PRAWDOPODOBIENSTWA WARUNKOWEGO**: jest to prawdopodobieństwo, które określa, na ile prawdopodobne jest dane zdanie, o ile prawdziwe jest pewne inne zdanie. Prawdopodobieństwo warunkowe zdania A ze względu na zdanie B zapisujemy jako i definiujemy przez:

$$\frac{\Pr(A \wedge B)}{\Pr(B)}.$$

Jest to ruch dość intuicyjny. Załóżmy, że rzucamy kością „szóstką”. Jaka jest szansa, że wypadnie liczba parzysta? $\frac{1}{2}$. Jaka jest szansa, że wypadnie liczba mniejsza od 3? $\frac{1}{3}$. Jaka jest szansa, że wypadnie liczba mniejsza od 3 i parzysta? Taka sama jak szansa, że wypadnie liczba 2, czyli $\frac{1}{6}$.

ZDANIE	PRAWDOPODOBIENSTWO
x jest parzyste	$\frac{1}{2}$
$x < 3$	$\frac{1}{3}$
$x < 3$ i x jest parzyste	$\frac{1}{6}$

Policzmy teraz, jaka jest szansa, że wypadnie liczba mniejsza od 3, o ile wypadnie liczba parzysta:

$$\frac{\Pr(x < 3 \wedge x \text{ jest parzyste})}{\Pr(x \text{ jest parzyste})} = \frac{1/6}{1/2} = 1/3$$

A jaka jest szansa, że wypadnie liczba parzysta, o ile wypadnie liczba mniejsza od 3?

$$\frac{\Pr(x < 3 \wedge x \text{ jest parzyste})}{\Pr(x < 3)} = \frac{1/6}{1/3} = 1/2$$

Niektórzy uważają nawet, że to pojęcie prawdopodobieństwa warunkowego powinno być pierwotne i że to za jego pomocą należy definiować prawdopodobieństwo bezwarunkowe.

Prawdopodobieństwo: podstawy

Z aksjomatów i definicji można wyprowadzać wiele fajnych rzeczy, o których tutaj nie będziemy mówić. Ale jedną rzeczą, o której trzeba na pewno wspomnieć, jest **TWIERDZENIE BAYESA**⁴⁴ (TB). Pozwala ono na wnioskowanie o prawdopodobieństwie hipotezy ze względu na dane w sytuacji, gdy dysponujemy raczej prawdopodobieństwem danych ze względu na hipotezę:

$$\Pr(A/B) = \frac{\Pr(B/A)\Pr(A)}{\Pr(B)}$$

Ważne spostrzeżenie jest takie – dokonujemy jakiejś obserwacji *B*. Chcemy wiedzieć, do jakiego stopnia wspiera ona naszą hipotezę *A*. Żeby móc to ustalić, potrzebujemy trzech informacji: wyjściowego prawdopodobieństwa hipotezy, wyjściowego prawdopodobieństwa wystąpienia obserwacji oraz informacji, na ile prawdopodobna byłaby obserwacja, gdyby hipoteza była prawdziwa.

Oglądaliście *Battlestar Galactica*? Przydatność TB można zobrazować za pomocą następującego przykładu. Wyobraźmy sobie, że cyloni (takie cyborgi) wyglądają tak jak ludzie i trudno ich od ludzi odróżnić. Założmy, że mamy test, który:

- jeżeli dana osoba jest cylonem, w 99,8% przypadków powie nam, że jest ona cylonem;
- jeżeli dana osoba jest człowiekiem, w 99,8% przypadków powie nam, że jest ona człowiekiem.

Testujemy więc jakąś osobę, test mówi nam, że jest cylonem. Czy to stanowi wystarczającą podstawę do wyrzucenia jej z naszego statku kosmicznego w próżnię?

Znamy dwa prawdopodobieństwa warunkowe: $\Pr(+/C)$ = $\Pr(\neg +/\neg C)$ = 0,998, gdzie *C* reprezentuje to, że dana osoba

⁴⁴ Thomas Bayes (1701–1761) był angielskim filozofem, który w chwilach wolnych od pisania traktatów o dobroci Boga zajmował się również statystyką.

jest cylonem, a + reprezentuje pozytywny wynik testu. Jeżeli podstawimy nasze stałe do twierdzenia Bayesa, dostaniemy:

$$\Pr(C / +) = \frac{\Pr(+ / C) \Pr(C)}{\Pr(+)} \quad (11)$$

To pokazuje nam, po pierwsze, że mamy za mało danych, bo potrzebujemy jeszcze znać $\Pr(C)$, czyli wiedzieć, jaka część społeczności to cyloni.

Po drugie, pokazuje, że musimy też jakoś dowiedzieć się, ile wynosi $\Pr(+)$. Na szczęście gdy już otrzymamy $\Pr(C)$, możemy $\Pr(+)$ wyprowadzić, posługując się wzorem:

$$\Pr(+)=\Pr(+/C)\Pr(C)+\Pr(+/\neg C)\Pr(\neg C) \quad (12)$$

Przyjmijmy na przykład, że cylonów jest dość mało: 0,1%. W takim razie razem z naszymi danymi (12) daje nam:

$$\Pr(+)=0,998\times0,001+\Pr(+/\neg C)\times0,999 \quad (13)$$

Skąd weźmiemy $\Pr(+/\neg C)$? To będzie to samo co $1-\Pr(\neg+/\neg C)$: 0,002. Dostajemy więc:

$$\begin{aligned} \Pr(+)&=0,998\times0,001+0,002\times0,999 \\ &=0,000998+0,001998 \\ &=0,002996 \end{aligned} \quad (14)$$

Zatem (11) daje nam:

$$\Pr(C / +) = \frac{\Pr(+ / C) \Pr(C)}{\Pr(+)} \quad (15)$$

$$= \frac{0,998 \times 0,001}{0,002996} \quad (16)$$

$$= \frac{0,000998}{0,002996} \quad (17)$$

$$\approx 0,3331 \quad (18)$$

Prawdopodobieństwo klasyczne

W takiej więc sytuacji pozytywny wynik testu wskazywałby tylko na to, że jest 33% szans, że osoba jest cylonem.

Morał ogólny: nawet jeżeli nasz test jest bardzo wydajny, to jeżeli zjawisko, które staramy się wykryć, jest bardzo nietypowe, sam fakt, że test był pozytywny, nie jest przekonującym świadectwem na rzecz prawdziwości hipotezy.

Tego typu metoda ma zastosowanie do wielu kwestii. Na przykład w rozważaniu problemów bioetycznych niezbędne jest zrozumienie związku pomiędzy danymi, jakie mamy, a hipotezami, które konstruujemy.

Czy wprowadzić tani powszechny test na HIV? To zależy nie tylko od jego skuteczności, lecz także od popularności HIV! Jeżeli HIV jest wystarczająco nietypowym zjawiskiem w rozważanej grupie, to duża część testów pozytywnych może być mylna, nawet jeżeli mamy dość trafny test. A to znaczy, że wprowadzenie takiego testu może mieć konsekwencję, że dość duża liczba ludzi będzie błędnie wierzyć, że ma HIV. Czy zaryzykujemy taką konsekwencję?

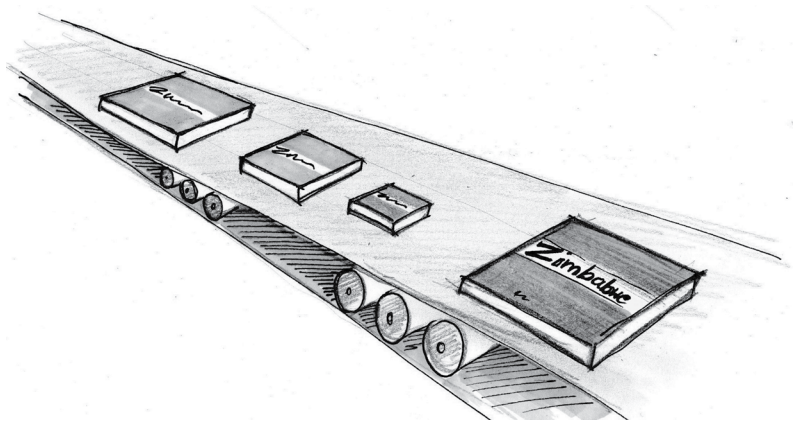
25. Prawdopodobieństwo klasyczne

Jak jednak rozumieć prawdopodobieństwo? Wedle **INTERPRETACJI KLASYCZNEJ** powinniśmy po prostu uznawać, że prawdopodobieństwo danej możliwości elementarnej jest równe ułamkowi, który w liczniku ma 1, a w mianowniku liczbę wszystkich możliwych wyników. Na przykład przy rzucie kością opcji jest 6, więc prawdopodobieństwo każdego z poszczególnych wyników wynosi $\frac{1}{6}$. Prawdopodobieństwo zaś opcji złożonej stanowi ułamek, który w liczniku ma liczbę wydarzeń elementarnych sprzyjających tej opcji, a w mianowniku liczbę wszystkich opcji. Na przykład prawdopodobieństwo przy rzucie kością wyrzucenia liczby parzystej wynosi $\frac{3}{6} = \frac{1}{2}$.

Pierwszy problem z taką interpretacją stanowi fakt, iż to podejście zakłada, że wszystkie możliwości elementarne są równie prawdopodobne. O ile może być to prawdą w przypadku rzucania kością, o tyle jest wysoce wątpliwe, żeby stosowało się do bardziej życiowych i skomplikowanych zjawisk.

Drugi, poważniejszy problem jest następujący⁴⁵. Wyobraźmy sobie fabrykę kwadratowych przewodników po Zimbabwie. Produkowane są one w różnych rozmiarach, pomiędzy 0 cm i 10 cm długości boku (te o boku 0 cm są dosyć tanie). Jakie jest prawdopodobieństwo, że dany przewodnik będzie miał długość boku pomiędzy 0 a 5 cm? Interpretacja klasyczna nakazywałaby podejrzewać, że 0,5.

No dobrze. Teraz rozważmy, co następuje. Minimalna powierzchnia to 0 cm². Maksymalna to 100 cm². Jaka jest szansa, że przewodnik będzie miał powierzchnię pomiędzy 0 a 25 cm²? No... podejście klasyczne wskazywałoby, że skoro jest to ¼ możliwych wielkości powierzchni, prawdopodobieństwo to wynosi 0,25.



⁴⁵ Poważniejszy, bo stanowi wyzwanie nie tylko dla interpretacji klasycznej, lecz także dla każdego, kto uznaje, że w przypadku całkowitego braku wiedzy należy po prostu prawdopodobieństwo rozdzielać po równo.

Prawdopodobieństwo logiczne

Teraz jednak pojawia się problem: posiadanie powierzchni pomiędzy 0 a 25 cm² to to samo, co posiadanie boku pomiędzy 0 a 5 cm. A jednak podejście klasyczne prowadzi do przypisania tym opcjom wykluczających się prawdopodobieństw! Wygląda więc na to, że interpretacja klasyczna nie jest wybitnie zadowalająca.

Trzecim problemem jest fakt, że twierdzenie, iż wydarzenia elementarne są równie prawdopodobne, wydaje się zakładać już jakieś rozumienie pojęcia prawdopodobieństwa, więc nie jest jasne, czy może z powodzeniem funkcjonować w interesującym wyjaśnieniu tego, czym jest prawdopodobieństwo.

26. Prawdopodobieństwo logiczne

Inną interpretacją jest **INTERPRETACJA LOGICZNA**. Wedle tej interpretacji wyjściowe aprioryczne prawdopodobieństwo zdań (którego sama teoria prawdopodobieństwa przecież nie dostarcza) da się ocenić poprzez analizę języka; dalsze reguły powiedzą nam, jak modyfikować to wyjściowe prawdopodobieństwo w świetle nowych danych.

Jednym z ważniejszych zwolenników tej interpretacji był Carnap, który starał się wyeksplikować zwrot „ x to stopień, do jakiego hipoteza H jest potwierdzona przez dane E ” zapisywany w skrócie jako $c(H, E) = x$.

Rozważmy jakiś prosty język zawierający skończoną ilość predykatów jednoargumentowych i skończoną ilość stałych indywiduowych.

LITERAL takiego języka to formuła, która albo jest atomiczna, albo jest negacją formuły atomicznej (np. Ga lub $\neg(Ga)$).

OPIS STANU RZECZY w takim języku to koniunkcja, która dla każdego predykatu i dla każdej stałej zawiera dokładnie jeden

literał z nich zbudowany. Intuicyjnie opis stanu rzeczy dla dowolnego przedmiotu i dowolnej własności wyrażonej przez predykat mówi, czy ten przedmiot ma tę własność.

OPIS STRUKTURY wygenerowany z danego opisu stanu rzeczy s to zbiór wszystkich tych opisów stanów rzeczy, które można otrzymać z s przez permutację stałych indywiduowych. Opis struktury w pewnym sensie zbiera razem te opisy stanu rzeczy, które zgadzają się co do ilości przedmiotów posiadających poszczególne własności.

Dla przykładu weźmy język z tylko jednym predykatem G i dwiema stałymi a i b . Istnieją cztery opisy stanu rzeczy:

- | | |
|---------------------------|--------------------------------|
| (a) $Ga \wedge Gb$, | (b) $Ga \wedge \neg Gb$, |
| (c) $\neg Ga \wedge Gb$, | (d) $\neg Ga \wedge \neg Gb$. |

i trzy opisy struktur:

- | | |
|---|---|
| (A) $\{Ga \wedge Gb\}$ | („wszystkie przedmioty mają własność G ”), |
| (BC) $\{Ga \wedge \neg Gb, \neg Ga \wedge Gb\}$ | („dokładnie jeden przedmiot ma własność G ”), |
| (D) $\{\neg Ga \wedge \neg Gb\}$ | („żaden przedmiot nie ma własności G ”). |

MIARA PRAWDOPODOBIEŃSTWA przypisuje prawdopodobieństwa do opisów stanów tak, że suma prawdopodobieństw wszystkich opisów stanów wynosi 1. Opisy stanów są rozłączne logicznie, więc prawdopodobieństwo alternatywy opisów stanów jest sumą prawdopodobieństw tych opisów. Każde zdanie jest równoważne alternatywie logicznie wykluczających się opisów stanów, więc dostajemy w ten sposób miarę, która przypisuje prawdopodobieństwo każdemu zdaniu.

Prawdopodobieństwo logiczne

Zakładając, że mamy już miarę prawdopodobieństwa m , $c(H, E)$ (stopień potwierdzenia H przez E) można zdefiniować za pomocą prawdopodobieństwa warunkowego:

$$c(H, E) = \frac{m(H \wedge E)}{m(E)} \quad (19)$$

Pozostaje więc określić miarę prawdopodobieństwa. Carnap rozpatrywał szereg opcji, przyjrzymy się dwóm najważniejszym.

Jeden sposób to podzielić prawdopodobieństwa równo na wszystkie opisy stanów. Jeżeli jest k opisów stanów i dokładnie n z nich logicznie pociąga za sobą H , prawdopodobieństwo H jest wtedy zdefiniowane jako $m^+(H) = n/k$.

Kontynuując nasz przykład, dochodzimy do wniosku, że każdy opis stanów dostaje wartość $m^+ = 1/4$. Zatem: $m^+(\neg Ga) = m^+(c) + m^+(d) = 1/4 + 1/4 = 1/2$. Jak się okazuje, stopień potwierdzenia $\neg Ga$ przez Gb – wedle określenia (19) – to także $1/2$:

$$c^+(\neg Ga, Gb) = \frac{m^+(\neg Ga \wedge Gb)}{m^+(Gb)} = \frac{m^+(c)}{m^+(a) + m^+(c)} = \frac{1/4}{1/2}$$

Podobne obliczenia pokazują, że $m^+(Ga) = c^+(Ga, Gb) = 1/2$.

To jednak dowodzi, że zaobserwowanie Gb nie ma wpływu na to, jakie prawdopodobieństwo powinniśmy przypisać Ga .

Problem ten można uogólnić. Nawet dla tysiąca przedmiotów $a_1, a_2, \dots, a_{1000}$ będzie prawdą, że nawet jak zaobserwujemy, iż 999 z nich posiada własność P , to nie zwiększy to potwierdzenia hipotezy, że ostatni też jest P . W takim jednak razie miara ta nie pozwala na modelowanie uczenia się z doświadczenia.

To skłoniło Carnapa do rozważenia innej miary prawdopodobieństwa, m^* . Metoda jest dość prosta: najpierw dzielimy prawdopodobieństwo 1 po równo na opisy struktur, zakładając, że każda struktura jest równie prawdopodobna. Potem dzieli-

my prawdopodobieństwo każdego opisu struktur na zawarte w nim opisy stanów.

W naszym przykładzie każdy z trzech opisów struktur dostaje prawdopodobieństwo $\frac{1}{3}$. Ponieważ (A) i (D) zawierają dokładnie jeden opis stanu, każdy z tych opisów dostaje prawdopodobieństwo $\frac{1}{3}$. Ponieważ (BC) ma dwa elementy, każdy z nich dostaje wartość $\frac{1}{6}$.

By zobaczyć, jak ta miara radzi sobie z poprzednią trudnością, porównajmy logiczne prawdopodobieństwo $\neg Ga$ z potwierdzeniem hipotezy $\neg Ga$ przez Gb (intuicyjnie, stopień potwierdzenia powinien być mniejszy od wyjściowego prawdopodobieństwa). $\neg Ga$ jest prawdziwe w (c) i (d) , a więc $m^*(\neg Ga) = m^*(c) + m^*(d) = \frac{1}{3} + \frac{1}{6} = \frac{1}{2}$. Z drugiej strony $c^*(\neg Ga, Gb)$ równe jest:

$$\frac{m^*(\neg Ga \wedge Gb)}{m^*(Gb)} = \frac{m^*(c)}{m^*(a) + m^*(c)} = \frac{1/6}{3/6} = 1/3$$

Tak jak się spodziewaliśmy, $c^*(\neg Ga, Gb) < m^*(\neg Ga)$. Podobnie $c^*(Ga, Gb) = \frac{2}{3} > m^*(Ga) = \frac{1}{2}$, a więc Gb (częściowo) potwierdza Ga i (częściowo) podważa $\neg Ga$.

Jednym z problemów, z którymi boryka się podejście Carnapa, jest nadmiar opcji. Okazuje się, że można jego podejście uogólnić i zdefiniować continuum miar prawdopodobieństwa pomiędzy m^+ a m^* , które zastosowane do konkretnych przypadków dadzą nam odmienne wyniki. Którą wybrać i dlaczego? Carnap zwykle preferował m^* , ale nigdy nie uzasadnił przekonująco tego wyboru.

Druga trudność jest taka, że stopień potwierdzenia hipotezy zależy od języka. Nie jest jednak jasne, jakie możemy mieć powody, by preferować ten, a nie inny język i w jaki sposób to podejście może wyjaśnić takie formy postępu w nauce, które wymagają zmiany języka.

Trzecia trudność jest taka, że nawet przy m^* żadna skończona liczba obserwacji nie jest w stanie zwiększyć stopnia

Prawdopodobieństwo logiczne

potwierdzenia hipotezy o dziedzinie (przynajmniej potencjalnie) nieskończonej. A to właśnie z tego typu hipotezami mamy zazwyczaj do czynienia w nauce. Pewne uogólnienie systemów Carnapa, które pozwala na niezerowe potwierdzenia tego typu, sformułował Jaakko Hintikka, ale systemy te są dość skomplikowane i mają swoje słabości.

Najgroźniejszy jednak dla syntaktycznego podejścia do prawdopodobieństwa jest paradoks **GRUE** (sformułowany przez Nelsona Goodmana)⁴⁶, który pokazuje, że jeżeli nie zwracamy uwagi na to, jakie predykaty dopuszczamy w naszym języku, rozumowania probabilistyczne dawać nam będą jednakowe pozytywne potwierdzenie sprzecznych hipotez.

Rozumowanie przebiega następująco. Wyobraźmy sobie, że zawodowo zajmujemy się polowaniem na predatory. Odkryliśmy, że nie potrafią one być niewidzialne, gdy sikają. W zasadzie zrobiliśmy karierę na czajeniu się na sikające predatory i zabijaniu ich w momencie, gdy stają się widzialne. Uznajemy więc zdanie następujące:

[PRED 1] Wszystkie zaobserwowane do teraz predatory były widzialne, gdy sikały.

I rzeczywiście było to prawdą dla całych wiader predatorów, które zabiliśmy. [PRED 1] wydaje się potwierdzać wniosek:

[PRED 2] Predator, na którego teraz poluję, będzie widzialny, gdy zacznie sikać.

Intuicyjnie ekstrapolowaliśmy własność „y jest widzialny, gdy sika” z przykładów zaobserwowanych na przypadek obecny. Nazwijmy tę własność własnością X.

⁴⁶ Nazwa pochodzi od uogólnień dotyczących kolorów kamieni szlachetnych: *green* i *blue*. Nasz przykład będzie nieco odmienny, ale jego istota będzie ta sama.

Wyobraźmy jednak sobie własność zdefiniowaną następująco:

[GRUE] Dany obiekt y jest grue wtedy i tylko wtedy, gdy: (a) y był zaobserwowany do teraz i posiadał własność X lub (b) y dopiero będzie zaobserwowany i nie posiada własności X .

Obserwacje, których do teraz dokonaliśmy, potwierdzają:

[PRED 1'] Wszystkie zaobserwowane do teraz predatory miały własność X .



Ale nie tylko! Potwierdzają również:

[PRED 3] Wszystkie zaobserwowane do teraz predatory miały własność grue.

Jeżeli więc [PRED 1'] wspiera:

[PRED 4] Następny predator będzie miał własność X

Częstotliwość i skłonność

to [PRED 3] wspiera:

[PRED 5] Następny predator będzie miał własność *grue*.

[PRED 5] znaczy jednak, że następny predator nie będzie miał własności *X*. Okazywałoby się więc, że doświadczenie w równym stopniu wspiera wniosek, że następny predator będzie *X*, jak i że nie będzie *X*.

Problem polega na tym, że dopuściliśmy całkiem nienaturalny predykat „*grue*”. Ale to właśnie wskazuje na to, że czysto syntaktyczne podejście nie rozwiąże problemu. Na poziomie bowiem syntaktycznym predykat to predykat i nie mamy podstaw, by jedno predykaty tolerować, a innych nie.

27. Częstotliwość i skłonność

Jeszcze inna interpretacja prawdopodobieństwa polega na tym, że prawdopodobieństwo danego wydarzenia jest względną częstotliwością wydarzeń tego typu w odpowiednio wybranej klasie odniesienia.

Brzmi dobrze, ale zastanówmy się, jakie ma to konsekwencje. Weźmy rzuty kostką. Prawdopodobieństwo wyrzucenia jedynki powinno być względną częstotliwością wyrzuceń jedynki w klasie odniesienia wszystkich rzutów kostką. Ale których wszystkich rzutów kostką? Nie ma sprzeczności w tym, że rzucę dziesięć razy i za każdym razem wyrzucę coś innego niż jedynkę. Czy to znaczy, że prawdopodobieństwo wyrzucenia jedynki będzie wynosić zero?

Być może chodzi o wszystkie do teraz wykonane rzuty kością na całym świecie. Ale nie jest zbyt wiarygodne, że dokładnie 1/6 z nich zaowocowała wyrzuceniem jedynki.

Może zatem zamiast mówić o klasie odniesienia jakichś rzeczywistych rzutów, powinniśmy mówić o częstotliwości granicznej w nieskończonej klasie potencjalnych rzutów? Może po prostu prawdopodobieństwo wyrzucenia jedynki wynosi $\frac{1}{6}$ dlatego, że **gdybym** rzucał w nieskończoność, to częstotliwość względna jedynek zmierzałaby do $\frac{1}{6}$, gdy liczba rzutów zmierzałaby do nieskończoności?

Trudność jest taka, że nie wiadomo, jaki obecny stan rzeczy miałby sprawiać, że częstotliwość względna nieistniejącego ciągu nieskończonego będzie taka, a nie inna. Nie wiadomo też, jakie miałyby być nasze źródło wiedzy o tego typu nieistniejących jeszcze ciągach. Co gorsza, nie ma obiektywnych powodów, by uważać, że każdy taki ciąg będzie miał granicę. Nawet pośród nieskończonych ciągów posiadających granicę może być tak, że zmieniając kolejność elementów ciągu, dostaniemy inną granicę – na którą permutację należy patrzeć?

Ponadto podejście częstotliwościowe ma duże trudności z wyjaśnieniem prawdopodobieństwa wydarzeń jednostkowych i nietypowych. Jeżeli chodzi o rzut kostką, wiadomo, że klasą odniesienia będą rzuty kością w ogóle. Ale co jest klasą odniesienia, powiedzmy, chrztu Polski? Na jaki potencjalny nieskończony ciąg powinienem patrzeć, żeby dowiedzieć się, jaka była jego szansa?

Opcją alternatywną jest uznanie, że przedmioty lub wydarzenia po prostu mają takie pierwotne i nieanalizowalne własności modalne, jak skłonność do wydarzania się z takim, a nie innym prawdopodobieństwem, czy skłonność do skutkowania takim, a nie innym wynikiem, z taką to a taką częstotliwością, w takich, a nie innych okolicznościach. W ten sposób unika się potrzeby wyjaśnienia tych własności za pomocą innych zjawisk oraz wyjaśniania się trudności, na które natrafiają różne teorie próbujące takie własności do innych zjawisk zredukować. Problem z tym podejściem jest taki, że po prostu traktuje ono prawdopodobieństwo jako nieanalizowalną własność pierwotną, nie wyjaśnia epistemologii prawdopodobieństwa i generalnie jest metafizycznie niejasne.

28. Subiektywizm i bayesianizm

Jedną z niewielu pozostałych opcji jest subiektywizm: pogląd, wedle którego prawdopodobieństwa to po prostu stopnie przekonania (tak rozumiane prawdopodobieństwa często nazywa się **KREDECJAMI**).

Standardowa analiza pojęcia kredencji wiąże kredencje z zakładami: osoba jest przekonana o ϕ w stopniu c wtedy i tylko wtedy, gdy osoba ta jest gotowa do postawienia c jednostek na zakład, który daje wygraną 1 jednostkę, jeżeli ϕ , i zero, jeżeli $\neg\phi$.

Ograniczając się do analizy wyidealizowanych podmiotów racjonalnych (czyli nie nas, ale pewnie czegoś, czym fajnie by było czasem być), to okazuje się, że istnieje silny związek pomiędzy zakładami a teorią prawdopodobieństwa. Można dowieść, że jeżeli kredencje danego podmiotu nie spełniają aksjomatów teorii prawdopodobieństwa, to można skonstruować serię zakładów, z których każdy osobno będzie się wydawał podmiotowi sprawiedliwy, ale na których razem wziętych ta osoba będzie musiała koniec końców stracić (zakłady takie noszą miano **DUTCH BOOKS**). Można też przeprowadzić rozumowanie w drugą stronę i dowieść, że jeżeli kredencje podmiotu spełniają aksjomaty teorii prawdopodobieństwa, to nie można takiej serii zakładów skonstruować. Argumenty wiążące kredencje z prawdopodobieństwem w ten sposób noszą miano **DUTCH BOOK ARGUMENTS**.

Zwolennicy zastosowania teorii prawdopodobieństwa i twierdzenia Bayesa w epistemologii sugerują, że minimalnym warunkiem racjonalności zmiany przekonania jest przestrzeganie **WARUNKOWANIA BAYESIAŃSKIEGO**. Polega ono na tym, że w momencie, gdy dowiadujemy się o nowym świadectwie E , po pierwsze, zmieniamy naszą kredencję w E na 1, a po drugie, dla każdego innego przekonania A nasza kredencja w A powinna równać

się uprzedniej kredencji w A , o ile E . Formalnie mówiąc, jeżeli $\Pr_{\text{stare}}(E) > 0$, to:

$$\begin{aligned}\Pr_{\text{nowe}}(A) &= \Pr_{\text{stare}}(A/E) \\ &= \frac{\Pr_{\text{stare}}(E/A)\Pr_{\text{stare}}(A)}{\Pr_{\text{stare}}(E)}\end{aligned}$$

Istnieje też wariant warunkowania dla sytuacji, w której nie odkrywamy E na pewno i $\Pr_{\text{nowe}}(E) < 1$, ale nie będziemy się tym teraz przejmować.

Klasyczny bayesianizm w epistemologii głosi więc, że dla racjonalnego podmiotu jego wyjściowe kredencje spełniają aksjomaty teorii prawdopodobieństwa, że taki podmiot zmienia stopnie przekonań zgodnie z warunkowaniem bayesiańskim i że nie istnieją inne ograniczenia na racjonalność podmiotu.

Nieklasyczne wersje bayesianizmu konstruowane są przez tych, którzy myślą, że dla racjonalności podmiotu potrzebne są jeszcze jakieś dodatkowe wymogi.

Bayesianizm ma swoje trudności – jest dosyć nowym nurtem, więc większość polemik jest w toku. Postaram się więc tylko wskazać główne obszary tych polemik.

- Po pierwsze, bayesianizm nie mówi nic na temat doboru wyjściowych prawdopodobieństw. Różne podmioty mogą w świetle tych samych danych przypisywać różne prawdopodobieństwa tej samej hipotezie.
- Odpowiedź bayesiańska jest taka, że wraz ze wzrostem ilości świadectwa, prawdopodobieństwa przypisywane hipotezom będą coraz bardziej zbieżne, niezależnie od kredencji wyjściowych. Ściśle rzecz biorąc, to działa dla kredencji różnych od 0 i 1.
- Bayesianizm krytykowany jest też za posługiwanie się zakładami dla analizy stopnia przekonań: nie jest jasne, czy jest to metoda adekwatna.

- Niektórzy krytykują przypisywanie precyzyjnych liczb rzeczywistych stopniom przekonań ludzkich jako nierealistyczne. W odpowiedzi konstruuje się szereg tzw. **TWIERDZENIE O REPREZENTACJI**. Zakłada się, że istnieje tylko pewien porządek na stopniach przekonań, który spełnia pewne własności formalne niewymagające odnoszenia się do liczb rzeczywistych. Następnie dowodzi się, że istnieje miara przyporządkowująca przekonaniom liczby rzeczywiste, która reprezentuje ten sam porządek za pomocą standardowego porządku na liczbach rzeczywistych. W takim razie, twierdzi się, posługiwanie się liczbami rzeczywistymi ma funkcję czysto instrumentalną, gdyż ostatecznie ocena porównawcza sądów pozostaje taka sama.
- Inni jeszcze uznają **DUTCH BOOK ARGUMENTS** za podejrzane. Dlaczego uzasadnienie normy racjonalności poznawczej ma być pragmatyczne? Czy rzeczywiście mam starać się nabierać jakichś przekonań w ten, a nie inny sposób jedynie po to, by uniknąć zakładu, który zmuszony będę przegrać? W odpowiedzi konstruuje się szereg nowych argumentów zastępujących ten argument – zamiast odwoływać się do czynników pragmatycznych, te argumenty posługują się różnymi pojęciami trafności przekonań i zmierzają do wykazania, że przestrzeganie aksjomatów prawdopodobieństwa i warunkowania bayesiańskiego gwarantuje większe szanse na trafność przekonań.
- Innym zjawiskiem jest problem **starych danych**; przynajmniej *prima facie* stare dane nie mogą już wedle bayesianizmu wpłynąć na zmianę naszej kredencji w daną hipotezę, nawet jeżeli dopiero teraz odkryliśmy, że mają one z tą hipotezą związek. Bayesianiści próbują wyjaśnić zjawiska tego typu na rozmaite sposoby.
- Trudno też w obrębie bayesianizmu wyjaśnić rolę istotnych w doborze teorii czynników, takich jak prostota czy jednolitość teorii.

- Nie zawsze prawdopodobieństwa po prawej stronie twierdzenia Bayesa są łatwiejsze do ustalenia. Na przykład $Pr(E/H)$ wydaje się czasem równie tajemnicze jak $Pr(H/E)$ w przypadku ogólnej hipotezy. Nie jest jasne, jak ustalić prawdopodobieństwo obserwacji czarnego kruka, gdy nie wszystkie kruki są czarne, czy prawdopodobieństwo obserwacji Eddingtona, jeżeli teoria względności jest fałszywa.
- Warunkowanie nie wynika z twierdzenia Bayesa. Twierdzenie to bowiem mówi coś o różnych prawdopodobieństwach w tym samym czasie – warunkowanie zaś dotyczy prawdopodobieństw w różnych momentach. Istnieją pewne warianty **DUTCH BOOK ARGUMENTS** dla takich przypadków, ale opierają się na założeniach raczej nietrywialnych i nieapriorycznych.

Niezależnie od tych trudności bayesianizm jest kwitnącym programem badawczym, który znajduje swoje zastosowanie w wielu dyskusjach filozoficznych, w których argumenty można modelować za pomocą pojęcia siły przekonań. W szczególności, dla przykładu, ludzie zajmujący się tego typu rzeczami używają aparatury matematycznej, by argumentować przeciwko sceptycyzmowi, analizować argumenty w filozofii religii, odpowiadać na problemy związane z uzasadnieniem indukcji w filozofii nauki czy modelować różnego rodzaju rozumowania w nauce i filozofii nauki.



TEKSTY POLECANE

1. Warto zapoznać się z książką *Metodologia nauk* Adama Grobiera (2006).
2. Z klasyków dobrze jest przeczytać *Wprowadzenie do filozofii nauki* Rudolfa Carnapa.
3. Przyjemny w lekturze jest podręcznik *An Introduction to Probability and Inductive Logic* autorstwa Iana Hackinga.
4. Ogólny obraz epistemologii formalnej można uzyskać, zapoznawszy się z książką *Mainstream and Formal Epistemology*, której autorem jest Vincent F. Hendricks.

Subiektywizm i bayesianizm

5. Jeżeli chodzi o epistemologię bayesiańską, klasycznym tekstem jest monografia ***Bayesian epistemology***, której autorami są Luc Bovens i Stephan Hartmann.
6. Wartym uwagi jest również wydane w 2015 roku ***A Critical Introduction to Formal Epistemology*** Darrena Bradley'a

KIERUNKOWSKAZ

Przyjrzelіśmy się trochę prawdopodobieństwu i jego filozofii; pora na jakiś czas wrócić do filozofii matematyki. Gdy to uczynimy, wrócimy z kolei do zagadnienia odniesienia terminów jednostkowych. Jeżeli chodzi o filozofię matematyki, zaczniemy od zrozumienia współcześnie standardowej teoriomnogościowej rekonstrukcji arytmetyki liczb naturalnych (a więc siłą rzeczy na początku musimy sobie powiedzieć, jak wygląda teoria mnogości). Dopiero potem dokonamy przeglądu kilku głównych stanowisk w filozofii matematyki.



CZĘŚĆ

IV

Filozofia matematyki

29. Teoria mnogości a arytmetyka

Wróćmy na jakiś czas do naszej opowieści o podstawach matematyki. Teoria ekstensji konstruowana przez Fregego natrafiła na problemy, bo prowadziła do istnienia zbioru Russella, który miał spełniać warunek:

$$\forall x [x \in \mathcal{R} \equiv x \notin x] \quad (20)$$

Warunek ten okazywał się problematyczny, bo prowadził szybko do sprzeczności, gdy stosowaliśmy go do samego zbioru $\mathcal{R}$:

$$\mathcal{R} \in \mathcal{R} \equiv \mathcal{R} \notin \mathcal{R}$$

Teoria typów wprowadzała hierarchię obiektów i ograniczała sensowność formuł typu $x \in y$ do sytuacji, w których x dotyczył typu o jeden poziom niżej niż y . Aby otrzymać arytmetykę, trzeba tam jednak było przyjąć aksjomat nieskończoności, wykluczyć sensowność wielu intuicyjnie bezproblemowych zdań oraz rozmnożyć egzemplarze poszczególnych liczb na różne poziomy (np. liczba 2 na poziomie 2, liczba 2 na poziomie 3 itd.).

Jeżeli się nad tym zastanowić, to paradoks Russella nie jest zagadnieniem, który pokazuje tajemniczość czy problematyczność pojęcia zbioru jako takiego. W zasadzie jest on paradoksem natury logicznej. Nie jest bowiem tak, że pojawia się tylko wtedy, gdy mówimy o relacji należenia. Nie musimy przyjmować żadnych specjalnych aksjomatów, by dowieść jego nieistnienia. Co więcej, istnieją niesprzeczne teorie mnogości, w których możliwe jest $x \in x$, więc syntaktyczne obostrzenia teorii typów wydają się przesadzone.

By zobaczyć, że zbiór Russella nie istnieje z przyczyn czysto logicznych, należy zauważyć, że jeżeli w (20) relację $\in$ zastąpimy dowolną inną relacją $\odot$ definiując jako zbiór $\mathcal{U}$, rezultat będzie analogiczny:

$$\forall x [x \odot \mathcal{U} \equiv \neg x \odot x] \quad (21)$$

i zaowocuje:

$$\mathcal{U} \odot \mathcal{U} \equiv \neg \mathcal{U} \odot \mathcal{U}$$

Być może więc pojęcie zbioru jako takie nie jest problematyczne, tylko po prostu należy skonstruować teorię zbiorów, która nie popada w kłopoty natury logicznej?

Tego typu intuicje leżały u podstaw jednej z głównych reakcji na teorię typów. Teoria typów była na tyle skomplikowana i restrykcyjna, że praktycznie nieużyteczna.

Zamiast niej wysunięto po prostu osobne aksjomaty dla pojęcia zbioru. Tak narodziła się standardowa **TEORIA MNOGOŚCI**. Przyjrzyjmy się jej aksjomatom oraz sposobowi, w jaki otrzymuje się arytmetykę w teorii mnogości.

AKSJOMAT EKSTENSJONALNOŚCI głosi, że jeżeli dwa zbiory mają te same elementy, są tymi samymi zbiorami. Razem z tą własnością identyczności, wedle której obiekty identyczne mają te same własności, można tezę wzmocnić do równoważności: zbiory są identyczne wtedy i tylko wtedy, gdy mają te same elementy.

Jeżeli nasza teoria mnogości dopuszcza obiekty podstawowe, które nie są zbiorami (nazywa się je **URELEMENTAMI**), to mamy również aksjomat, który powiada, że nic nie jest elementem urelementu, oraz aksjomat, który powiada, że istnieje zbiór wszystkich urelementów U .

Definiujemy następnie relację **BYCIA PODZBIOREM**: zbiór A jest podzbiorem zbioru B ($A \subseteq B$) wtedy i tylko wtedy, gdy każdy element zbioru A należy też do zbioru B . W szczególności każdy zbiór jest swoim własnym podzbiorem. Relacja bycia nieiden-

Teoria mnogości a arytmetyka

tycznym podzbiorem nazywana jest relacją bycia **PODZBIOREM WŁAŚCIWYM**. Dwa zbiory są zatem identyczne, jeżeli są wzajemnie swoimi podzbiarami.

Pamiętamy, że do paradoksu Russella prowadziło pełne wyróżnianie: teza, że dla dowolnego warunku istnieje zbiór wszystkich i tylko tych przedmiotów, które spełniają ten warunek. Zamiast niego przyjmuje się zdecydowanie słabszy **AKSJOMAT WYRÓŻNIANIA**. Powiada on, że jeżeli mamy jakiś zbiór A i jakiś warunek ϕ , to istnieje podzbiór tych elementów A , które spełniają ten warunek (oznaczamy go: $\{x \in A \mid \phi(x)\}$). Zamiast więc generować zbiory, aksjomat ten raczej pozwala nam generować podzbiory zbiorów, o których istnieniu już wiemy.

Dla przykładu aksjomat wyróżniania razem z aksjoma-tem, który stwierdza istnienie zbioru wszystkich urelementów, pociąga za sobą istnienie zbioru pustego. Wystarczy bowiem zbiór pusty zdefiniować jako taki podzbiór zbioru urelementów, którego elementy spełniają warunek $x \neq x$. Wyróżniamy więc zbiór pusty ze zbioru urelementów, jako ϕ biorąc warunek sprzeczny niespełniany przez żaden przedmiot. Zbiór pusty oznaczamy jako $\emptyset$.

Możemy przeprowadzać operację iloczynu zbiorów A i B , która polega na braniu zbioru tych elementów, które należą do obu zbiorów. Istnienia iloczynu można dowieść za pomocą aksjomatu wyróżniania, wystarczy wziąć $x \in B$ jako warunek ϕ . Jeżeli bowiem A jest zbiorem, to zbiorem też będzie $\{x \in A \mid x \in B\}$, czyli zbiór tych elementów zbioru A , które należą również do B .

AKSJOMAT PARY powiada, że dla dowolnych obiektów (urelementów lub zbiorów) a i b istnieje zbiór c taki, że $a \in c$ i $b \in c$. Z samego aksjomatu nie wynika, czy to c nie zawiera jeszcze innych elementów. Ale gdy przyjmujemy istnienie takiego zbioru, można dowieść istnienia zbioru $\{a, b\}$ za pomocą aksjomatu wyróżniania, definiując zbiór $\{x \in c \mid x = a \vee x = b\}$.

Z istnienia takiego zbioru c oraz aksjomatu wyróżniania wynika również istnienie zbiorów jednoelementowych $\{a\}$ i $\{b\}$, zdefiniowanych przez $\{x \in c \mid x = a\}$ i $\{x \in c \mid x = b\}$.

Ściśle rzecz biorąc, istnienie zbiorów jednoelementowych wynika także z samego aksjomatu pary, gdy pozwolimy na parę $\{a, a\}$, która jest identyczna z $\{a\}$.

Gdy mamy jakiś zbiór zbiorów, na przykład $\{\{a, b\}, \{c\}\}$, **AKSJOMAT SUMY** mówi nam, że istnieje zbiór zawierający wszystkie elementy zbiorów w tym zbiorze (w tym przypadku chodzi o zbiór, do którego należą co najmniej a , b i c ; istnienie zbioru $\{a, b, c\}$ wynika wtedy na mocy aksjomatu wyróżniania). Suma zbiorów A i B oznaczana jest przez $A \cup B$. Istnienie sumy dwóch dowolnych zbiorów wynika z aksjomatu pary i z aksjomatu sumy. Najpierw aksjomat pary mówi nam, że istnieje rodzina zbiorów $\{A, B\}$, a potem aksjomat sumy mówi nam, że istnieje ich suma.

Standardowa teoria mnogości przyjmuje też **AKSJOMAT UFUNDOWANIA**, który wyklucza nieskończone ciągi typu:

$$\dots a \in b \in a \in b \in a$$

czy nawet zbiory bez dna typu:

$$\dots a_k \in a_{k-1} \in \dots \in a_2 \in a_1 \in a,$$

które jako takie nie są koliste. Aksjomat ten dosłownie powiada, że dowolny niepusty zbiór A posiada element y , który nie ma żadnych elementów należących do A . Nieistnienie ciągów i bezdennych zbiorów wynika z tego aksjomatu.

Wedle jednej z dominujących opinii problem z niektórymi niby-zbiorami, które okazują się nie istnieć – jak zbiór Russella czy zbiór uniwersalny (por. s. 64) – jest taki, że są one za duże; że musiałyby posiadać za dużo elementów. O ile intuicja ta nie zawsze jest jasna, wydaje się trafne twierdzenie, że jeżeli już wiemy, że istnieje jakiś zbiór A , i wiemy, że domniemany zbiór B miałby nie więcej elementów niż A , to można uznać, że B jest rzeczywiście zbiorem. Tę zależność stwierdza **AKSJOMAT ZASTĘPOWANIA**.

Teoria mnogości a arytmetyka

Kolejny aksjomat standardowej teorii mnogości to **AKSJOMAT ZBIORU POTĘGOWEGO**. Jak sama nazwa wskazuje, powiada on, że dla dowolnego zbioru istnieje jego zbiór potęgowy (pojęcie to wyjaśniliśmy już na s. 58).

Zbiór $\{a, b\}$ to ten sam zbiór co $\{b, a\}$. Czasem jednak chcielibyśmy rozróżniać kolejność elementów i mówić o **PARACH I CIĄGACH UPORZĄDKOWANYCH**. Chcielibyśmy, aby to, że ciąg $\langle a, b \rangle$ jest identyczny z ciągiem $\langle c, d \rangle$, implikowało, że $a = c$ i $b = d$. Aby to osiągnąć, wystarczy zdefiniować $\langle a, b \rangle$ jako $\{\{a\}, \{a, b\}\}$. W analogiczny sposób można zdefiniować każdy skończony ciąg przedmiotów (szczegóły znaleźć można w każdym podręczniku do teorii mnogości).

RELACJA n -argumentowa postrzegana jest jako jakiś zbiór n -ek uporządkowanych. Na przykład gdybyśmy traktowali liczby naturalne jako urelementy, to relacja $y = x \times 2$ byłaby zidentyfikowana ze zbiorem par $\{\langle x, y \rangle \mid y = x \times 2\}$:

$$\{\langle 0, 0 \rangle, \langle 1, 2 \rangle, \langle 2, 4 \rangle, \langle 3, 6 \rangle, \langle 4, 8 \rangle, \dots\}$$

FUNKCJA zaś jest taką relacją, która każdemu przedmiotowi przypisuje co najwyżej jeden przedmiot. Na przykład dzielenie przez 2 jest funkcją na liczbach naturalnych, bo każdej liczbie naturalnej przyporządkowuje co najwyżej jedną liczbę naturalną. Warto zwrócić uwagę, że funkcja nie musi przyporządkowywać czegoś każdemu przedmiotowi w rozważanej dziedzinie. Na przykład dzielenie przez 2 nie przyporządkowuje żadnej liczby naturalnej liczbie 1.

Przez jakiś czas (powiedzmy, w pierwszej połowie XX wieku) kontrowersyjnym aksjomatem był **AKSJOMAT WYBORU**. W jednej wersji powiada on, że jeżeli A jest rodziną zbiorów (czyli zbiorem zbiorów), to istnieje funkcja, która z każdego z tych zbiorów wybiera jeden jego element. Innymi słowy, że z każdego stada zbiorów można wybrać po jednym elemencie z każdego zbioru z tego stada.

Głównym powodem, dla którego początkowo niektórzy martwili się o akceptowalność aksjomatu wyboru, była jego **niekon-**

struktywność. Konstruktywiści matematyczni uważali, że aby móc dowieść istnienia jakiegoś obiektu, należy go najpierw zdefiniować za pomocą dostępnych środków. Odrzucali natomiast dowody istnienia, które nie pokazywały dokładnie, o jaki obiekt chodzi. Gdybyśmy dowiedli, że istnieje przedmiot spełniający dany warunek, po prostu pokazując, że założenie, iż nie istnieje taki przedmiot, prowadzi do sprzeczności, konstruktywista narzekałby, że nie pokazaliśmy wcale, jaki przedmiot spełnia ten warunek.

Dowody posługujące się aksjomatem wyboru są niekonstruktywne: pozwalają wnioskować o istnieniu pewnej funkcji, nie mówiąc wcale, jak ona wygląda, czy też nie mówiąc, jak jest zdefiniowana (ani nawet, że jest definiowalna). I rzeczywiście matematyka niekonstruktywna jest znacząco silniejsza od tej, która ogranicza się do metod konstruktywnych. Fakt jednak, że matematyka konstruktywna jest dużo bardziej skomplikowana i nie pozwala na dowody wielu twierdzeń uznawanych przez większość matematyków, sprawił, że nie licząc pewnych wyjątków, konstruktywizm odszedł do lamusa i stanowi raczej hobby niektórych matematyków niż szeroki i popularny program badawczy.

Teraz możemy przyjrzeć się modelowaniu arytmetyki wewnątrz teorii mnogości. Po pierwsze, mówimy, że zbiór A jest **INDUKCYJNY**, jeżeli należy do niego zbiór pusty ($\emptyset \in A$), oraz dla dowolnego elementu $x \in A$ również $x \cup \{x\} \in A$. Zbiór indukcyjny zawiera więc jako elementy:

- po pierwsze, zbiór pusty: $\emptyset$,
- po drugie, $\emptyset \cup \{\emptyset\}$, czyli po prostu $\{\emptyset\}$,
- Po trzecie, $\{\emptyset\} \cup \{\{\emptyset\}\}$, czyli $\{\emptyset, \{\emptyset\}\}$,
- Po czwarte, $\{\emptyset, \{\emptyset\}\} \cup \{\{\emptyset, \{\emptyset\}\}\}$,
czyli $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$,
- itd.

Zbiór indukcyjny posiada więc nieskończoną liczbę elementów. Jak się okazuje, istnienie zbioru indukcyjnego nie jest dowodliwe na podstawie wprowadzonych powyżej aksjoma-

Teoria mnogości a arytmetyka

tów. Dla modelowania arytmetyki potrzebny jest jeszcze jeden aksjomat. **AKSJOMAT NIESKOŃCZONOŚCI** powiada, że istnieje co najmniej jeden zbiór indukcyjny.

To wyczerpuje listę aksjomatów standardowych. Aksjomaty wymienione, bez aksjomatu wyboru, dają nam teorię zwaną **TEORIĄ MNOGOŚCI ZERMELO-FRAENKLA**⁴⁷ (z urelementami). Teoria Zermelo-Fraenkla bez urelementów (w skrócie ZF) zamiast aksjomatów właściwych dla urelementów zawiera **AKSJOMAT ZBIORU PUSTEGO**, który mówi, że istnieje zbiór pusty. ZF z aksjomatem wyboru nazywa się ZFC (C pochodzi od *choice*).

Standardowo zbiór teoriomnogościowych liczb naturalnych definiowany jest jako najmniejszy zbiór indukcyjny i oznaczany przez ω (omega). To znaczy, że ω zawiera te zbiory, które konstruowaliśmy w liście powyżej, i nie zawiera żadnego innego zbioru.

Na s. 45 wprowadziliśmy arytmetykę Peana drugiego rzędu (PA_2). Teraz zajmiemy się arytmetyką Peana rzędu pierwszego (PA_1):

Aksjomat	Znaczenie
$\forall x S(x) \neq 0$	0 nie jest następnikiem żadnej liczby.
$\forall x \forall y [S(x) = S(y) \rightarrow x = y]$	Dwie różne liczby nie mają tego samego następnika.
$\forall x (x + 0 = x)$	Część definicji dodawania.
$\forall x \forall y (x + S(y) = S(x + y))$	Część definicji dodawania.
$\forall x (x \times 0 = 0)$	Część definicji mnożenia.
$\forall x \forall y (x \times S(y) = (x \times y) + x)$	Część definicji mnożenia.
Dla dowolnej formuły φ , $\varphi(0) \wedge \forall x (\varphi(x) \rightarrow \varphi(S(x))) \rightarrow$ $\rightarrow \forall x \varphi(x)]$	Schemat indukcji.

Zasadnicza różnica pomiędzy PA_2 a PA_1 jest taka, że PA_2 zawiera jeden aksjomat indukcji, który kwantyfikuje po wszyst-

⁴⁷ Ernst Zermelo (1871–1953) był niemieckim logikiem i matematykiem, który w 1935 roku zrezygnował ze swojej katedry w ramach protestu przeciwko Hitlerowi. Abraham Fraenkel (1891–1965) był urodzonym w Niemczech żydowskim matematykiem, który po wyemigrowaniu do Izraela głosił, iż Rabinat powinien sprawować zwierzchnictwo nad małżeństwami i rozwodami.

kich podzbiorach zbioru liczb naturalnych. W efekcie aksjomat nieskończoności powiada, że dla **dowolnego zbioru** liczb, jeżeli należy do niego liczba 0 i jeżeli dla dowolnej liczby x , jeżeli x należy do niego, to również $S(x)$ do niego należy, każda liczba naturalna należy do tego zbioru.

W odróżnieniu od tego przypadku schemat indukcji daje nieskończoną liczbę aksjomatów: jest ich tyle, ile formuł, które możemy podstawić za ϕ . Schemat indukcji powiada więc, że dla dowolnej formuły, jeżeli 0 spełnia tę formułę i jeżeli dla dowolnego x , jeżeli x spełnia tę formułę, to $S(x)$ też ją spełnia, to dowolna liczba spełnia tę formułę.

Być może nieco zaskakująco istnieje między tymi wariantami zasadnicza różnica. Aksjomat indukcji mówi coś o **wszystkich zbiorach** liczb naturalnych. Ile ich jest? Mówimy o mocy zbioru potęgowego zbioru liczb naturalnych, więc będzie ich na mocy twierdzenia Cantora więcej niż liczb naturalnych. Schemat indukcji mówi w efekcie tylko o **ZBIORACH DEFINIOWALNYCH** w języku arytmetyki. Tych zaś będzie dużo mniej: co najwyżej tyle samo, ile liczb naturalnych (bo każdy zbiór jest definiowalny za pomocą jakiejś formuły, każda formuła jest skończonym ciągiem symboli, a tych może być co najwyżej tyle, ile liczb naturalnych). Ta różnica między PA_2 a PA_1 powoduje szereg innych fascynujących różnic, o których nie mogę tu mówić ze względu na ograniczenia miejsca i cel książki (żeby wymienić tylko jedną fajną sprawę: PA_1 nie wyklucza modeli niestandardowych z „dzikimi” liczbami większymi od każdej liczby naturalnej; PA_2 takie przypadki wyklucza).

Nasza definicja ω pozwala dowieść schematu indukcji dla zbioru ϕ . Weźmy bowiem dowolne ϕ . Załóżmy, że $\phi(0)$ oraz:

$$\forall x[\phi(x) \rightarrow \phi(x \cup \{x\})]$$

W teorii mnogości rolę następnika x odgrywa następny zbiór w zbiorze indukcyjnym: $x \cup \{x\}$. Chcemy dowieść, że wynika z tego $\forall x \in \omega \phi(x)$. Zdefiniujmy $\Phi = \{x \in \omega \mid \phi(x)\}$. Na mocy aksjomatu

Teoria mnogości a arytmetyka

wyróżniania $\emptyset$ jest zbiorem. Nasze założenia pociągają za sobą to, że $\emptyset$ jest indukcyjne (zawiera bowiem $\emptyset$ i następnik każdego swojego elementu). Jako że ω jest najmniejszym zbiorem indukcyjnym, dostajemy $\omega \subseteq \emptyset$. To jednak daje nam nasz zamierzony wniosek.

Analogicznie możemy dowieść pozostałych aksjomatów. Weźmy dwa pierwsze dla przykładu:

- Aby dowieść, że $\forall x S(x) \neq 0$, dowodzimy raczej, że $n \cup \{n\} \neq \emptyset$. To jednak jest trywialne, bo $n \in n \cup \{n\}$.
- Aby dowieść $\forall x \forall y [S(x) = S(y) \rightarrow x = y]$, dowodzimy raczej, że $\neg \cup \{n\} = m \cup \{m\}$ pociąga za sobą $n = m$. Założmy po-przednik i negację następnika. Zauważmy, że skoro n jest po lewej stronie:

$$\neg \cup \{n\} = m \cup \{m\}$$

to musi też być po prawej. Skoro jednak $n \neq m$, to znaczy, że $n \in \{m\}$. Analogicznie wnioskujemy, że $m \in \{n\}$. To jednak jest wykluczone przez aksjomat ufundowania.

Podobnie jeżeli odpowiednio zdefiniować operację dodawania i mnożenia na zbiorach, możemy dowieść o ω aksjomatów pozostałych. W ten sposób pewna reinterpretacja języka arytmetyki okazuje się podteorią teorii mnogości.

W sumie jeżeli chodzi o rekonstrukcję arytmetyki w teorii mnogości, wszystko śmiga. Czy jednak daje to podstawę do uznania, że filozoficzna ciekawość dotycząca natury wiedzy matematycznej została zaspokojona? Nie bardzo.

KIERUNKOWSKAZ

Teraz przyjrzymy się pobieżnie filozoficznym interpretacjom matematyki. Zaczniemy od klasycznie wyróżnianych: platonizmu, intuicjonizmu i formalizmu (dwa ostatnie są jednak współcześnie nie- zbyt powszechnie akceptowane). Następnie przejdziemy do nieco nowszych tendencji: neologicyzmu, nominalizmu i strukturalizmu (zarówno w jego platonistycznym, jak i nominalistycznym wariacie).



30. Platonizm matematyczny

Założmy więc nawet, że dziedzina zainteresowania matematyków jest w miarę zunifikowana za pomocą teorii mnogości (a już samo to jest dość silnym założeniem). Pozostają dwa ważne pytania:

[PYTANIE EPISTEMOLOGICZNE] Jaka jest natura i prawomocność wiedzy matematycznej?

[PYTANIE ONTOLOGICZNE] Jaki jest status ontologiczny przedmiotów wiedzy matematycznej?

Okazuje się ponadto, że pomiędzy różnymi strategiami odpowiedzi na te pytania zachodzi następująca interakcja: im łatwiejsza wydaje się odpowiedź na jedno z nich, tym trudniejsza wydaje się odpowiedź na drugie. I odwrotnie.

Pozornie najprostsza odpowiedź na pytanie ontologiczne to **MATEMATYCZNY PLATONIZM**, wedle którego przedmioty matematyczne to po prostu przedmioty abstrakcyjne *sui generis*: liczby, zbiory itd. po prostu istnieją w świecie obiektów matematycznych, a ich istnienie i prawdziwość zdań matematycznych są niezależne od naszego języka, naszej myśli czy naszych praktyk.

Niestety, taka odpowiedź na pytanie ontologiczne to dość silne twierdzenie, które wymaga argumentów. Istnieją chyba dwie główne strategie obrony platonizmu.

- Zaczniemy od tego, że zdania matematyczne są prawdziwe i zawierają terminy, które wyglądają jak terminy jednostkowe. Mamy też silną intuicję, że prawdy matematyczne są prawdziwe niezależnie od tego, co my o nich myślimy. Ale zdania zawierające terminy jednostkowe mogą mieć tego typu wartość obiektywną, tylko jeżeli przedmioty, do

których te terminy jednostkowe się odnoszą, rzeczywiście mają własności, które te zdania im przypisują.

- Należy uznawać istnienie tych przedmiotów, których istnienie trzeba założyć, aby wyjaśnić prawdziwość naszych najlepszych teorii naukowych. Nasze najlepsze teorie naukowe posługują się matematyką – aby były prawdziwe, muszą więc istnieć przedmioty matematyczne. Zatem powinniśmy uznawać ich istnienie.

Pierwszy z tych argumentów wywodzi się zasadniczo od Fregego. Argumenty przechodzące jednak od prawdziwości pewnych zdań do istnienia pewnych obiektów nie zawsze są niezawodne (por. rozdz. 17). Niemniej jednak argument jest o tyle ważny, że przed (prawie) każdym, kto chce odrzucić platonizm w filozofii matematyki, stoi wyzwanie wyjaśnienia prawdziwości zdań matematycznych bez odwoływania się do obiektów matematycznych.

Drugi z tych argumentów (zwany też **ARGUMENTEM Z NIEZBYWALNOŚCI**) wywodzi się w pewniej mierze od Hilarego Putnama oraz od Quine'a i jego kryterium **ZAANGAŻOWANIA ONTOLOGICZNEGO** (zwanego też czasem zobowiązaniem ontologicznym). Zdaniem Quine'a, aby wiedzieć, czego istnienie należy uznać, jeżeli uznaje się daną teorię (czyli jakie jest jej zaangażowanie ontologiczne), należy najpierw sformułować tę teorię jako teorię pierwszego rzędu, a następnie zobaczyć, jakie przedmioty muszą znaleźć się w dziedzinie tej teorii, aby była ona prawdziwa. Zastosowane do naszego problemu podejście to głosi, że jeżeli nasze najlepsze teorie naukowe sformułujemy jako teorie pierwszego rzędu, to obiekty matematyczne będą musiały znaleźć się w dziedzinie przedmiotów, o których mówią te teorie.

Ten argument ma pewne pozory precyzji, ale nie jest powalający. Po pierwsze, nacisk na teorie pierwszego rzędu wymaga uzasadnienia. Uzasadnienie, jakie proponuje Quine, to (1) niechęć do modalności i języków nieekstensjonalnych (o pojęciu ekstensjonalności więcej w rozdziale 37), a zatem zamiłowanie do języków, w których występują tylko spójniki klasyczne, oraz (2) niechęć do języków wyższego rzędu.

(1) inspirowane jest argumentami Quine'a przeciwko sensowności i użyteczności pojęć modalnych (które, jak się okazało, nie są zbyt przekonujące); ponadto nigdy nikomu nie udało się żadnej interesującej fizycznej teorii naukowej na serio sformalizować w ekstensjonalnym języku pierwszego rzędu; (2) motywowane jest poglądem, że języki wyższego rzędu niosą ze sobą zobowiązanie do obiektów teoriomnogościowych, a więc nie są czysto logiczne, i faktem, że teoriom wyższego rzędu brakuje pewnych cech formalnych (np. operacja konsekwencji drugiego rzędu nie da się opisać za pomocą skończonego zbioru reguł syntaktycznych, którymi moglibyśmy się posługiwać). Trudno jednak w wyjaśnianiu kryterium zobowiązania ontologicznego już zakładać, że wiemy, jakie jest zobowiązanie ontologiczne teorii wyższego rzędu, a sama niechęć do pewnych własności formalnych również nie jest wybitnie rozstrzygająca.

Ponadto wygląda na to, że dla większości teorii przedstawianych wyjściowo w języku potocznym, nawet jeżeli da się je sformułować w języku pierwszego rzędu, istnieje często wiele alternatywnych sformułowań, zależnie na przykład od tego, które z terminów nazwowych języka potocznego uznamy za rzeczywiste wyrażenia nazwowe, a które tylko za skróty. Nasze formalne rekonstrukcje takich teorii różnić się będą efektem końcowym, jeżeli chodzi o to, co będzie musiało należeć do dziedziny. Nie jest jednak jasne, który z tych sposobów ma być właściwą miarą zobowiązania ontologicznego.

Zagadnienie wskazuje również na konieczność odróżnienia **REALIZMU WARTOŚCI LOGICZNEJ** od **REALIZMU ONTOLOGICZNEGO**. Realizm wartości logicznej odnośnie do jakiejś teorii to pogląd, że zdania języka tej teorii są obiektywnie prawdziwe lub fałszywe. Realizm ontologiczny to pogląd, że rzeczywiście istnieją przedmioty, o których ta teoria zdaje się mówić. Realizm wartości logicznej odnośnie do matematyki nie pociąga za sobą realizmu ontologicznego. Można np. wierzyć, że zdania matematyki są obiektywnie prawdziwe, ale próbować wyjaśniać, na

Intuicjonizm matematyczny

czym polega ich prawdziwość, bez odwoływania się do obiektów matematycznych *sui generis*.

Zaletą realizmu ontologicznego jest to, że jeżeli przedmioty matematyczne istnieją wiecznie, w pozaprzestrzennym i niezmiennym świecie obiektów matematycznych, to wydaje się zrozumiałe, że prawdy o nich są konieczne.

Podstawowy jednak problem z platonizmem jest taki, że udzielając pozornie prostej odpowiedzi na pytanie ontologiczne, nie za dobrze radzi on sobie z odpowiedzią na pytanie epistemologiczne. Jeżeli obiekty matematyczne istnieją poza czasem i przestrzenią jako obiekty abstrakcyjne i przyczynowo nieaktywne, to jakim cudem my, obiekty skończone w czasie i przestrzeni, możemy coś o nich wiedzieć? Platonikom nie udało się za bardzo wyjść poza postulowanie jakiejś tajemniczej intuicji matematycznej (Kurt Gödel) lub sugerowanie, że prawdy matematyczne są prawdziwe o obiektach matematycznych na mocy znaczenia (Frege).

O ile opcja pierwsza pozostawia epistemologię matematyki wielką tajemnicą, o tyle opcja druga sugeruje, że na mocy znaczenia prawdziwe mogą być zdania, z których wynika istnienie jakichś rzeczywistych i przyczynowo niedostępnych obiektów. Wydaje się to przeczyć intuicji, że sam fakt, iż wyrażenia, jakimi się posługujemy, znaczą to, co znaczą, nie może za sobą pociągać rzeczywistego istnienia żadnych przedmiotów, a tym bardziej przedmiotów pozaprzestrzennych i przyczynowo odizolowanych, jakie postuluje platonizm.

31. Intuicjonizm matematyczny

Inną opcją wydaje się intuicjonizm. Intuicjonista uznaje, że obiekty matematyczne to jakiegoś rodzaju obiekty mentalne stworzone przez nasze umysły. Z tego typu pomysłami (wywodzącymi się

od Luitzena E.J. Brouwera) wiąże się pojęcie konstrukcji wykonywanej przez wyidealizowanego matematyka pozbawionego ograniczeń czasu życia, pamięci itd. Taki matematyk może wykonywać dowolnie duże skończone konstrukcje, ale nie jest w stanie wykonywać konstrukcji wymagających nieskończenie wielu kroków. W związku z tym intuicjoniści odrzucają również rzeczywiste istnienie zbiorów nieskończonych (choć dopuszczają możliwość potencjalną: możliwość, że niezależnie od tego, ile obiektów danego typu istnieje, zawsze mogłoby ich istnieć więcej).

Jak już wspominałem na str. 160, konstruktywiści odrzucają niekonstruktywne dowody istnienia. Jak wygląda struktura takiego dowodu? By dowieść jakiejś formuły $\exists x \phi(x)$, zakłada się $\neg \exists x \phi$, dowodzi się, że założenie to prowadzi do sprzeczności, z czego wnioskuje się $\neg \neg \exists x \phi(x)$, a dalej wyprowadza się wniosek, że $\exists x \phi(x)$. Istotnym ruchem jest tutaj użycie schematu $\neg \neg \psi \rightarrow \psi$ (ψ czytamy jako psi), który jest równoważny schematowi wyłączanego środka $\psi \vee \neg \psi$.

To jednak znaczy, że intuicjoniści zmuszeni są do zrewidowania logiki, jaką się posługują. I rzeczywiście odrzucają zasadę wyłączanego środka i równoważne jej zasady logiczne. Logika intuicjonistyczna jest więc słabsza od logiki klasycznej. O ile w przypadku intuicjonistycznej arytmetyki liczb naturalnych teoria intuicjonistyczna nie różni się za bardzo od arytmetyki Peana, o tyle w przypadku bardziej rozwiniętych teorii matematycznych pojawiają się zasadnicze rozbieżności i teorie intuicjonistyczne są zdecydowanie słabsze i bardziej skomplikowane.

Przede wszystkim jednak w przypadku intuicjonizmu pojawia się pytanie, do jakiego stopnia mamy do czynienia z filozofią zastanej matematyki, a do jakiego stopnia z konstrukcją nowej i odmiennej matematyki. Jeżeli naszym celem jest zrozumienie tego, co robiły pokolenia klasycznych matematyków, to filozofia, która owocuje zdecydowanie odmiennymi teoriami matematycznymi, nie dostarcza nam odpowiedzi. Filozoficzne wyjaśnienie matematyki klasycznej zamiast głosić, że należy matematykę klasyczną odrzucić (matematycy nie słuchają za bardzo filozo-

Formalizm matematyczny

fów – predatorowi dzięki!), powinno dostarczyć nam głębszego wglądu w jej funkcjonowanie, przedmiot i znaczenie. Intuicjonizm zaś tego nie robi. Jak już wspomniałem, z tych powodów intuicjonizm od dłuższego czasu nie jest intensywnie studiowany i nie stanowi doniosłego kierunku badań (choć pozostaje przy życiu paru sprytnych intuicjonistów, a pytania o granice metod intuicjonistycznych są do pewnego stopnia również interesujące).

32. Formalizm matematyczny

Kolejnym podejściem jest **FORMALIZM** (stworzony przez Davida Hilberta). Z perspektywy formalistycznej arytmetyka jest teorią zajmującą się własnościami symboli. Zero można interpretować jako jedną pionową kreskę $|$, liczbę 1 jako dwie kreski $||$ itd. Tę część matematyki, która może być tak zinterpretowana, Hilbert uważał za matematykę rzeczywistą; matematykę wyższą, niepoddającą się takiej interpretacji rozumiał instrumentalnie jako tylko formalną grę, polegającą na manipulacji symbolami, mającą na celu dowodzenie prawd rzeczywistej matematyki.

Jedną z kwestii, którą należało rozstrzygnąć, była niesprzeczność matematyki wyższej (np. analizy matematycznej): jeżeli bowiem aparatura ta byłaby sprzeczna, pozwalałaby na dowodzenie wszystkiego (bo w logice klasycznej ze sprzeczności wynika dowolne zdanie). Hilbert wraz ze współpracownikami starał się za pomocą metod czysto arytmetycznych dowodzić niesprzeczności różnych systemów, bez większego powodzenia.

Program Hilberta nabrał trochę rozpędu, ale w roku 1931 Kurt Gödel dowiódł, że jeżeli arytmetyka Peana jest niesprzeczna, to istnieją zdania w języku arytmetyki Peana od niej niezależne: takie zdania – które nie są dowodliwe w arytmetyce i których negacje nie są w niej dowodliwe.

Gödel podał ogólną metodę konstrukcji zdań nierozstrzygalnych dla dowolnej niesprzecznej teorii spełniającej pewne całkiem podstawowe warunki (wystarczającej mocy systemu). To zaś sprawiło, że nie można było problemu rozwiązać po prostu wzmacniając system wyjściowy. Nawet bowiem ten wzmocniony system, o ile byłby niesprzeczny, nie rozstrzygałby wszystkich zdań. To wskazuje na fakt, że powiedzenie, iż zdanie jest prawdziwe, gdy jest dowodliwe w systemie, prowadzi do trudności: dla wszystkich wystarczająco silnych teorii (mniej więcej co najmniej tak silnych jak arytmetyka Peana minus schemat indukcji), jeżeli są one niesprzeczne, istnieją zdania, które wedle tej definicji nie byłyby ani prawdziwe, ani fałszywe.

Co więcej, Gödel dowiódł też **DRUGIEGO TWIERDZENIA O NIEZUPEŁNOŚCI**, które powiada, że żaden niesprzeczny system spełniający te warunki podstawowe nie jest w stanie dowieść swojej własnej niesprzeczności. Warto podkreślić, że gdyby taki system dowodził niesprzeczności jakiejś mocniejszej teorii, dowodziłby też niesprzeczności własnej (bo jeżeli jakiś zbiór jest niesprzeczny, to niesprzeczny jest każdy jego podzbiór). Zatem konsekwencją drugiego twierdzenia jest to, że arytmetyka Peana, jeżeli jest niesprzeczna, nie jest w stanie dowieść niesprzeczności matematyki wyższej. To jednak początkowo stanowiło główny cel programu Hilberta.

Istnieje jednak również filozoficzny problem z formalizmem. Jeżeli matematyka jest tylko nauką o symbolach, w jaki sposób wyjaśnić jej stosowalność w naukach ścisłych? Jeżeli z perspektywy formalistycznej każdy (niesprzeczny) system jest równie dobry, jak wyjaśnić to, że matematycy i naukowcy jedne systemy uznają i ich używają, a inne bezużyteczne lub uznawane za fałszywe odrzucają, a do tego jeszcze pracują głównie w teoriach nie do końca sformalizowanych i bezczelnie przypisują znaczenie uznawanym przez siebie wyrażeniom? Formaliści nigdy nie potrafili w sposób satysfakcjonujący odpowiedzieć na to pytanie. Obecnie nie ma za dużo formalistów, a formalizm nie odgrywa doniosłej roli w filozofii matematyki.

33. Neologicyzm

O logicyzmie Fregego i Russella mówiliśmy już wcześniej. System Fregego okazał się sprzeczny, a system Russella zbyt skomplikowany, ograniczony i słaby, by dostarczyć przyjaznej w użyciu podstawy dla wielu bogatych teorii matematycznych.

Niemniej zwróćmy uwagę na pewien istotny element rozwoju teorii Fregego (por. rozdz. 13). Z Podstawowego Prawa Numer Pięć (które prowadzi do sprzeczności z logiką drugiego rzędu i wyróżnianiem dla pojęć) wyprowadził zasadę Hume'a. Następnie, nie posiłkując się już Podstawowym Prawem Numer Pięć, użył zasady Hume'a, by wyprowadzić arytmetykę Peana drugiego rzędu.

Gdy okazało się, że PPV (Podstawowe Prawo Numer Pięć, zob. s. 47) jest sprzeczne, bo prowadzi do paradoksu Russella, odłożono całą rzecz do szuflady. Dopiero w latach 70. dwudziestego wieku zauważono, że zasada Hume'a:

Liczba pojęcia F (liczba F -ów) jest identyczna z liczbą pojęcia G wtedy i tylko wtedy, gdy pomiędzy pojęciami F i G zachodzi relacja równoliczności

sprzeczna nie jest (jeżeli nie jest sprzeczna analiza matematyczna). Pojawił się więc pomysł oparcia rekonstrukcji teorii matematycznych na bezpiecznych zasadach abstrakcji (więcej o ZASADZIE HUME'A na s. 48) i na unikaniu ryzykownych zasad abstrakcji (takich jak PPV).

NEOLOGICYZM w swojej ortodoksyjnej postaci uznaje, że ZH stanowi wystarczające wyjaśnienie pojęcia liczby, że jest prawdą pojęciową poznawalną *a priori* i że dostarcza w ten sposób wiedzy o obiektywnie istniejących liczbach, które są obiektami *sui generis*. Mniej ortodoksyjne warianty neologicyzmu nie przyjmują realizmu ontologicznego.

Jedną z trudności jest brak ogólności: trudno uogólnić to podejście do innych teorii matematycznych. Jakich zasad abstrakcji użyć, by zrekonstruować takie teorie? Skąd wiadomo, że te zasady abstrakcji będą niesprzeczne?

Inną trudnością jest to, że samo pojęcie ryzykowności zasady abstrakcji jest ryzykowne. Jak odróżnić bezpieczne zasady abstrakcji od niebezpiecznych?

- Można pomyśleć: sprawdzać, które prowadzą do sprzeczności, i przyjmować tylko te, które tego nie robią. Można jednak dowieść, że nie istnieje ogólna metoda sprawdzania, które zasady abstrakcji prowadzą do sprzeczności.
- Nawet gdybyśmy wiedzieli, które nie są sprzeczne, to mamy tzw. **problem złego towarzysza** (*bad company problem*). Niektóre zasady abstrakcji są niesprzeczne i grzeczne, jeśli są osobno, ale robią zadymę i prowadzą do sprzeczności, gdy spotkają się w jednej teorii.

Potrzebne jest więc inne kryterium akceptowalności zasad abstrakcji i w żadnym wypadku nie jest jasne, jak miałyby ono wyglądać. Istnieje kilka propozycji, ale każda ma zwolenników i przeciwników.

34. Nominalizm

NOMINALIŚCI uważają, że obiekty matematyczne nie istnieją. Z jednej strony ułatwiać to może zbudowanie nieczarodziejskiej epistemologii matematyki. Z drugiej jednak strony, nominalista musi wyjaśnić, dlaczego i w jaki sposób działania matematyków i zastosowania matematyki w naukach ścisłych mają sens.

Jedną z propozycji jest **FIKCJONALIZM** (Hartry'ego Fielda). Pomysł polega tutaj na tym, że obiekty matematyczne mają

Strukturalizm matematyczny

takie samo istnienie jak Sherlock Holmes czy predator, a nawet trochę gorsze istnienie niż predator, bo zmyśla się je tylko po to, żeby łatwiej było rozumować.

Field uważa, że na argument z niezbywalności można odpowiedzieć, jeśli najpierw skonstruuje się nominalistyczne wersje teorii fizykalnych (rozważa przykład Newtonowskiej teorii grawitacji), a następnie pokaże się, że dodanie do nich aparatury matematycznej nie zmienia zbioru konsekwencji fizykalnych tej teorii.

Innym ciekawym wariantem jest **NOMINALIZM MODALNY**. Przy tego typu podejściu (Charlesa Chihary) twierdzeniom matematyki przypisuje się obiektywną wartość, ale interpretuje się je jako mające inne warunki prawdziwości.

Zamiast mówić o rzeczywistym istnieniu jakichś obiektów, twierdzenia matematyczne, wedle tego stanowiska, mówią raczej o tym, jakiego typu wyrażenia można w sposób uprawniony wprowadzić.

Na przykład zdanie „Istnieje zbiór pusty” można rozumieć jako głoszące: „Można wprowadzić formułę, która nie jest spełniana przez żaden przedmiot” albo „Można wprowadzić nazwę, która się do niczego nie odnosi”.

Wyzwaniem jest skonstruowanie tego typu teorii dla wielu różnych teorii matematycznych oraz wyjaśnienie natury możliwości, którą się ta metoda posługuje.

Istnieją jeszcze co najmniej dwa ważne podejścia nominalistyczne, ale jako że są one strukturalistyczne, omówię je w następnym rozdziale.

35. Strukturalizm matematyczny

Gdy się nad tym zastanowić, nie jest jasny sam status redukcji liczb do zbiorów (na problem zwrócił uwagę Paul Benacerraf).

Istnieje tak naprawdę nieskończona ilość sposobów, na które możemy zidentyfikować liczby ze zbiorami tak, żeby formalnie rzecz biorąc, wszystko działało. Dla przykładu zamiast ciągu ω (por. s. 161) możemy równie dobrze użyć ciągu:

$$\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\{\{\emptyset\}\}\}, \dots$$

odpowiednio redefiniując nasze rozumienie relacji następnika czy operacji dodawania i mnożenia. To jednak wskazuje, że skoro nie ma powodów, by preferować jedną z tych redukcji, jako mówiącą nam, **czym naprawdę są liczby**, a wszystkie naraz nie mogą być adekwatne, to żadna z nich nie jest filozoficznie pomocna do zrozumienia, czym są liczby.

Stewart Shapiro zauważa, że problem ten nie pojawia się dla teorii matematycznych, które nie sprawiają wrażenia, jakby były teoriami jednego specyficznego modelu. Na przykład nie ma sensu pytać, o której grupie (w sensie matematycznym) jest teoria grup⁴⁸.

Być może więc nawet te teorie matematyczne, które *prima facie* sprawiają wrażenie teorii o jednym specyficznym modelu, wcale o nim nie są? Być może arytmetyka nie jest o jakimś jednym wyśnionym i wybranym zbiorze liczb naturalnych, tylko po prostu o wszystkich obiektach o takiej samej strukturze, co zbiór ω ? To wyjaśniałoby reinterpretowalność arytmetyki Peana na wiele sposobów i pozwalałoby uniknąć wyżej opisanego **problemu identyfikacji**. Wedle tego podejścia arytmetyka dotyczy wspólnej **struktury** wielu różnych **systemów**.

Jednym wariantem strukturalizmu jest **STRUKTURALIZM ANTE REM** (przed rzeczą), wedle którego struktury istniałyby (i istnieją), nawet gdy nie istnieją realizujące je systemy. Na przykład istnieje struktura liczb naturalnych, nawet jeżeli liczba przedmiotów na świecie jest skończona i nie ma żadnego systemu rzeczywistych przedmiotów fizycznych, które mogłyby

⁴⁸ Z drugiej strony, nie jest jasne, że arytmetyka jest teorią tylko jednej struktury. Arytmetyka pierwszego rzędu ma bowiem modele niestandardowe, które nie mają tej samej struktury, co standardowy zbiór liczb naturalnych.

Strukturalizm matematyczny

stanowią realizację tej struktury. Strukturalizm *ante rem* jest, właściwie rzecz biorąc, wariantem platonizmu (tyle, że zamiast o istniejących w niebie platońskim liczbach, mówi się o istniejących w niebie platońskim strukturach) i dlatego natrafia na szereg problemów epistemologicznych napotykanym przez standardową formę platonizmu.

Inną wersją jest **STRUKTURALIZM IN REBUS** (w rzeczach). Głosi on, że jedyne struktury, jakie istnieją, to rzeczywiste systemy złożone z obiektów rzeczywiście istniejących. Jeżeli za obiekty rzeczywiście istniejące uznamy tylko obiekty konkretne, to otrzymamy koncepcję nominalistyczną.

W tej interpretacji zdanie typu „ $2 + 2 = 4$ ” głosi: „w każdym (rzeczywiście istniejącym) ω -systemie S (albo przynajmniej systemie spełniającym warunki narzucone przez aksjomaty arytmetyki Peana), S -drugi element S -dodany do siebie daje S -czwarty element”. Przy takim rozumieniu struktury, jeżeli nie ma żadnego konkretnego ω -systemu, nie istnieje też ω -struktura. Stąd też inna nazwa tego poglądu: **STRUKTURALIZM ELIMINATYWNY**.

Problem ze strukturalizmem eliminatywnym jest taki, że jeżeli rzeczywiście nie istnieje żadna ω -struktura, to każde zdanie zaczynające się od „W każdym (rzeczywiście istniejącym) ω -systemie S ...” będzie prawdziwe (bo poprzednik będzie fałszywy).

To jednak znaczy, że aby otrzymać sensowną matematykę, strukturalista eliminatywny musi założyć, że odpowiednie obiekty istnieją. To sprawia, że chociaż wystarczą mu systemy konkretne, a nie jakieś oderwane od rzeczywistości materialnej obiekty matematyczne, to jednak jego ontologia musi być bardzo bogata.

Kolejnym nominalistycznym wariantem strukturalizmu jest **STRUKTURALIZM MODALNY**. Strukturalista modalny stara się ulepszyć strukturalizm eliminatywny, dodając do swojej interpretacji element modalny i czytając zdania takie, jak „ $2 + 2 = 4$ ” jako „W każdym **MOŻLIWYM** ω -systemie S , S -drugi element S -dodany do siebie daje S -czwarty element”.

Dzięki temu strukturalista modalny nie musi zakładać rzeczywistego istnienia wielu systemów, wystarczy mu ich możliwość. Ceną, jaką za to płaci, jest potrzeba wyjaśnienia pojęcia możliwości, którym się posługuje.



TEKSTY POLECANE

1. Lekturę o filozofii matematyki można zacząć od hasła *Philosophy of Mathematics* (Leon Horsten) w *Stanford Encyclopedia of Philosophy*.
2. Ciekawy przegląd stanowisk w filozofii matematyki znaleźć można w pracach Ch. Chihary: *Constructibility and Mathematical Existence* oraz *A Structural Account of Mathematics*.
3. Zdecydowanie warto zapoznać się z pracami *Filozofia matematyki. Zarys dziejów* oraz wyborem tekstów *Współczesna filozofia matematyki* R. Murawskiego.
4. Dobrym wprowadzeniem do filozofii matematyki pisanym z perspektywy strukturalistycznej jest *Philosophy of Mathematics: Structure and Ontology* S. Shapiry.



KIERUNKOWSKAZ

Omówiliśmy mniej więcej główne nurty filozofii matematyki. Wracamy teraz do wątku odniesienia terminów jednostkowych, poruszonego, w rozdziałach o Fregem i Russellu.

CZĘŚĆ

V

Odrobina filozofii języka

36. Problemy z deskrytywizmem

Deskrytywizm Fregego i Russella, o którym już wspominałem w rozdziale 17, był poglądem raczej dominującym w pierwszej połowie wieku XX. Poważniejsze argumenty przeciw niemu zaczęto wysuwać dopiero potem. Jednym z najważniejszych tekstów, opartym na trzech wykładach z roku 1970, wygłoszonych w Princeton, było *Naming and Necessity* Saula Kripkego.

Deskrytywizm odnośnie do nazw własnych ma dwa ważne warianty. Zajmiemy się najpierw pierwszym:

[D+] Nazwy własne mają to samo znaczenie co przypisane im deskrypcje określone.

Wedle tej koncepcji np. nazwa „Rafał Urbaniak” (dalej w skrócie „RU”) po prostu **znaczy** dla niektórych to samo, co, powiedzmy, „autor najbardziej beużytecznego wprowadzenia do filozofii analitycznej”⁴⁹. Oczywiście nie wszyscy muszą z nazwą „RU” wiązać tę samą deskrypcję. Niektórzy mogą raczej używać deskrypcji „jedyny niesprawiedliwy logik na UG” albo „filozof o najbardziej absurdalnych poglądach w porównaniu do innych filozofów w Gdańsku”, zależnie od swoich przekonań.

W każdym razie, jeżeli [D+] jest prawdą, to musi istnieć przynajmniej jedna deskrypcja określona, która ujmuje znaczenie nazwy „RU”. Załóżmy, że tak jest, i zobaczmy, co się stanie.

Jeżeli „autor najbardziej beużytecznego wprowadzenia do filozofii analitycznej” (w skrócie „ABEFA”) rzeczywiście **znaczy** to samo co „RU”, to zdanie:

⁴⁹ Przepraszam za nawiązywanie do samego siebie w przykładach. Po refleksji uznałem, że jedyną osobą, którą mogę bez obaw obrażać, jestem ja sam.

[RAF 1] Jeżeli RU istnieje, to „RU jest autorem najbardziej bezużytecznego wprowadzenia do filozofii analitycznej” jest konieczne i *a priori* prawdziwe.

Jest tak z tego samego powodu, z jakiego zdaniem *a priori* i koniecznie prawdziwym jest „Jeżeli jakiś trójkąt istnieje, to jest trójbokiem”.

Problem jednak jest taki (to, co teraz rozważymy, nazywa się zazwyczaj **ARGUMENTEM MODALNYM**), że [RAF 1] nie jest ani konieczne, ani *a priori*. Mogłem nie zostać filozofem (faktycznie, w pewnym istotnym momencie w moim życiu poważnie wahałem się między pracą w straży pożarnej, pilotażem helikopterów wojskowych a filozofią; w końcu, dla pieniędzy, wybrałem filozofię). W niektórych światach możliwych zamiast logikiem mogłem zostać myśliwym i polować na predatory. Żeby wiedzieć, że jestem ABEFA, należy mieć pewien kontakt z rzeczywistością; z pewnością nie można tego wiedzieć *a priori* bez przeczytania niniejszej książki.

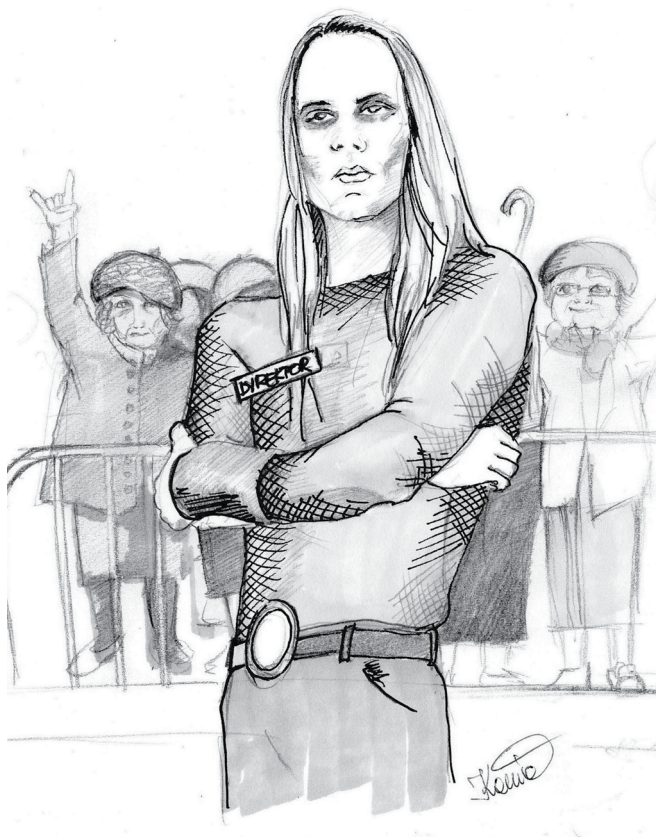
Nie pomoże tutaj zastępowanie deskrypcji ABEFA żadną inną przychodzącą na myśl deskrypcją. Mogę nie być jedynym niesprawiedliwym logikiem na UG (albo mogłem nim nie zostać) i mogę nie być (albo mogłem nie zostać) filozofem o najbardziej absurdalnych poglądach w porównaniu do innych filozofów w Gdańsku. Jeżeli którąkolwiek z tych deskrypcji spełniam, ustalenie, że tak jest, wymaga pewnej wiedzy *a posteriori*, a to, że te deskrypcje spełniam, nie jest konieczne.

Niektórzy (jak John Searle) sugerowali, że deskryptywizm można uratować, mówiąc, że nazwy zamiast odpowiadać jednej deskrypcji, stowarzyszone są raczej z pewnym zbiorem deskrypcji i warunkiem bycia odniesieniem danej nazwy jest spełnianie pewnej znaczącej większości z nich.

To jednak nie ratuje zbyt wiele sytuacji: tak samo jak mogłem nie zostać ABEFA, mogłem też nie spełniać żadnej (albo większości) z pozostałych deskrypcji. Analogiczne rozważania stosują się do innych nazw własnych.

Problemy z deskrytywizmem

Argument wyżej omówiony nasuwa na myśl następujące rozróżnienie. Pewne wyrażenia mogą odnosić się do różnych rzeczy w różnych okolicznościach. Na przykład deskrypcja „dyrektor Radia Maryja” w różnych okolicznościach może być spełniana przez różne obiekty (zamknijmy oczy i wyobraźmy sobie świat, w którym dyrektorem Radia Maryja został Nergal).



Pewne zaś odnoszą się do tych samych przedmiotów w dowolnych okolicznościach, o ile te przedmioty istnieją. Na przykład „RU” odnosi się do tego samego przedmiotu (do mnie) nawet w świecie możliwym, w którym zostałem ćpającym heroinę pseudosatanistą gwałcącym owce i nie mam nic wspólnego z logiką czy filozofią analityczną.



Wyrażenia tego ostatniego typu nazywamy **DESIGNATORAMI SZTYWNYMI**.

[DESIGNATOR SZTYWNY]

Termin jednostkowy ***j* sztywno desygnuje** przedmiot ***x*** wtedy i tylko wtedy, gdy ***j*** desygnuje ***x*** w każdym możliwym świecie, w którym ***x*** istnieje, i nie desygnuje niczego innego niż ***x*** w żadnym świecie możliwym.

Istnieje też w miarę jasny test intuicyjny na to, czy termin jednostkowy ***j*** jest desygnatorem sztywnym. Termin jednostkowy ***j*** jest desygnatorem sztywnym wtedy i tylko wtedy, gdy...

[TEST SZTYWNOŚCI]

...przedmiot, który rzeczywiście jest ***j***, nie mógłby istnieć, nie będąc ***j***, i nic innego niż to, co rzeczywiście jest ***j***, nie mogłoby być ***j***.

Problemy z deskryptywizmem

Tak więc „RU” jest desygnatorem sztywnym, bo obiekt, który rzeczywiście jest RU, nie mógłby istnieć, nie będąc RU, i dlatego, że nic innego niż to, co rzeczywiście jest RU, nie mogłoby być RU.

Należy tu podkreślić, że nie działają argumenty typu:

Przecież RU mógłby nie być RU, bo są światy możliwe, w których po prostu ludzie nazywają go inaczej, na przykład „Heidegger”.

To, że **po zmianie** naszego języka jakieś zdanie (jako obiekt syntaktyczny) ma pewien status modalny, nie ma większego związku ze statusem modalnym naszego zdania wyjściowego w jego rzeczywistym znaczeniu. W przeciwnym wypadku moglibyśmy również argumentować:

Przecież wcale nie jest konieczne, że $2 = 1 + 1$, bo są światy możliwe, w których to zdanie znaczy „Dobrym pomysłem jest mieć w domu tresowanego predatora, bo szybko aportuje” i jest fałszywe.

Jasne, jest prawdą, że w jakimś innym możliwym języku nazywany jestem „Heidegger”, ale nawet w takim świecie dalej jestem sobą, „bo dla każdego x , $x = x$ ” jest prawdą logiczną. Jasne, jeżeli „ $2 = 1 + 1$ ” zmieni znaczenie, to może nie być konieczne. Ale my jesteśmy zainteresowani zdaniem w tym znaczeniu, jakie one mają teraz. Pytamy więc, czy przy naszym znaczeniu wyrażenia „RU”, rzeczywisty RU mógłby nie być RU? Czy coś innego niż nasz rzeczywisty RU mogłoby być RU? I odpowiedź na oba pytania jest przecząca.

Dla kontrastu użyjmy skrótu „ĆPO” dla „ćpającego heroinę pseudosatanisty gwałcącego owce” (i załóżmy, że w rzeczywistości istnieje tylko jeden ĆPO). Czy „ĆPO” jest desygnatorem sztywnym? To zależy od tego, czy przedmiot, który rzeczywiście jest ĆPO, nie mógłby istnieć, nie będąc ĆPO, i czy nic innego niż to, co rzeczywiście jest ĆPO, nie mogłoby być ĆPO.

Ale wiemy, że druga część tego warunku nie jest spełniona. RU (choćaby dlatego, że dostęp do owiec ma mocno ograniczony) w rzeczywistości nie jest $\dot{C}PO$, ale mógłby być.

Pierwsza część warunku też nie jest spełniona, bo nawet jeżeli istnieje obecnie jedyny przedmiot, który jest $\dot{C}PO$ (mam nadzieję, że nie istnieje żaden taki przedmiot), to przecież ta osoba mogła nie zostać $\dot{C}PO$.

Z tego wszystkiego nie wynika, że nie istnieją deskrypcje sztywne. W przypadkach szczególnych mogą się trafić. Na przykład „wynik mnożenia liczby 2 przez liczbę 2” wydaje się odnosić do liczby 4 w każdym możliwym świecie. W każdym bowiem możliwym świecie dwa razy dwa to cztery. Tego typu przypadki są jednak zdecydowanie specyficzne i dotyczą sytuacji, gdy mówimy o obiektach uznawanych za istniejące koniecznie i posiadające te cechy, które posiadają, również koniecznie.

Kolejny wariant deskryptyzmu zastępuje deskrypcje typu „przedmiot, który jest $\dot{C}PO$ ” deskrypcjami typu „przedmiot, który rzeczywiście jest $\dot{C}PO$ ” (zabieg takiej zmiany deskrypcji nazywa się jej **USZTYWNIENIEM**). Trochę to pomaga, bo chociaż (jak już ustaliliśmy) deskrypcja „przedmiot, który jest $\dot{C}PO$ ” nie jest sztywna, to deskrypcja „przedmiot, który rzeczywiście jest $\dot{C}PO$ ” jest sztywna. Dlaczego?

Zastosujmy [TEST SZTYWNOŚCI]. Czy przedmiot, który rzeczywiście jest $\dot{C}PO$, mógłby nie być przedmiotem, który rzeczywiście jest $\dot{C}PO$? Z pewnością mógłby być przedmiotem, który w jakimś świecie możliwym nie jest $\dot{C}PO$. Ale w naszym rzeczywistym świecie (który jest tylko jeden) dalej byłby $\dot{C}PO$. To, czy jest $\dot{C}PO$ w jakimś innym świecie możliwym, nie zmienia jego statusu w rzeczywistości. Czy jakiś przedmiot, który nie jest rzeczywiście $\dot{C}PO$, mógłby być rzeczywiście $\dot{C}PO$? Z pewnością mógłby być $\dot{C}PO$ w jakimś świecie możliwym, ale nie w naszym rzeczywistym, gdzie już wiadomo, że nie jest $\dot{C}PO$.

Dobrze, wiemy więc, że deskrypcje usztywnione są desygntorami sztywnymi. Czy więc nazwy własne nie mogłyby

Problemy z deskrytywizmem

być synonimiczne z takimi usztywnionymi deskrypcjami? Gdyby były synonimiczne, następujące dwa zdania znaczyłyby to samo (zakładając, że „*j*” to nazwa własna, a „*D*” to odpowiadająca jej deskrypcja):

[RAF 2] O ile *j* istnieje, to *j* rzeczywiście jest *D*.

[RAF 3] O ile to, co rzeczywiście jest *D*, istnieje, to to, co rzeczywiście jest *D*, rzeczywiście jest *D*.

Problem jednak jest taki, że drugie zdanie jest *a priori*, podczas gdy dla większości terminów jednostkowych *j* pierwsze zdanie *a priori* nie jest, niezależnie od tego, jakiej deskrypcji *D* użyjemy.

Spróbujmy dla przykładu utożsamić znaczenie nazwy „RU” z deskrypcją „osoba, która w rzeczywistości jest ABEFA”.

Czy jest prawdą *a priori*, co następuje:

[RAF 2'] O ile RU istnieje, to RU rzeczywiście jest osobą, która w rzeczywistości jest ABEFA?

Nie, bo nasza znajomość następnika, że RU jest w naszej rzeczywistości ABEFA, nie jest *a priori*.

Czy jest prawdą *a priori* zdanie następujące:

[RAF 3'] O ile osoba, która w rzeczywistości jest ABEFA, istnieje, to osoba, która w rzeczywistości jest ABEFA, rzeczywiście jest ABEFA?

Tak, bo *a priori* wiemy, że jeżeli ktoś w rzeczywistości jest ABEFA, to rzeczywiście jest ABEFA. Zatem [RAF 2] oraz [RAF 3] różnią się znaczeniem, co wskazuje na to, że nawet usztywniona deskrypcja nie jest równoznaczna z nazwą własną.

Przyjrzyjmy się więc kolejnemu, nieco osłabionemu wariantowi deskrytywizmu:

[D–] Nazwy własne mogą nie być **synonimiczne** z przypisanymi im deskrypcjami, ale jest regułą języka, że deskrypcje te ustalają odniesienie tych nazw.

To znaczy, że język działa tak, że nawet jeżeli nazwa własna nie ma tego samego znaczenia co deskrypcja, z którą jest stowarzyszona, to ma to samo odniesienie na mocy tego, jak funkcjonuje język.

Dla tej słabszej od [D+] interpretacji istotne są następujące tezy, uznawane w przypadku, gdy deskrypcja *D* ustala odniesienie nazwy *n* dla osoby *o*:

1. Osoba *o* jest przekonana, że tylko jeden przedmiot *x* spełnia deskrypcję *D*.
2. Jeżeli *D* stosuje się do dokładnie jednego przedmiotu *x*, to ten *x* jest odniesieniem nazwy *n*.
3. Jeżeli *D* nie stosuje się do dokładnie jednego przedmiotu, to nazwa *n* nie ma odniesienia.
4. Osoba *o* *a priori* wie, że jeżeli nazwa *n* się do czegoś odnosi, to jest prawdą, że *n* jest *D*.

Jest to pogląd trochę słabszy, bo nie stwierdza synonimiczności deskrypcji i nazwy⁵⁰. Niezależnie jednak od argumentu modalnego Kripke uważa, że można do każdego z nich podać kontrprzykłady:

[ad 1] W większości przypadków deskrypcje, które kojarzemy z przedmiotami lub osobami, nie wyznaczają dokładnie jednego przedmiotu i o tym wiemy. Weźmy „Cyceron”. Większość z nas wie tylko, że był on sławnym rzymianinem lub sławnym rzymskim mówcą. To jednak nie jest deskrypcja spełniana tylko przez jedną osobę.

[ad 2] Większość ludzi, o ile w ogóle ma pojęcie o Peano, to myśli, że Peano to człowiek, który sformułował aksjomaty

⁵⁰ Czy stosuje się do niego argument modalny, jest kwestią dyskusyjną, którą czytelnik może z radością porozwagać, wybierając marchewki lub szcztukując swoją żyrafę.

Problemy z deskrytywizmem

arytmetyki Peana. W rzeczywistości jednak aksjomaty te sformułował Dedekind. To jednak nie sprawia, że nazwa „Peano” odnosi się do Dedekinda.

[ad 3] Niektórzy z nazwą „Kripke” kojarzą jedynie deskrypcję „człowiek, który sformułował semantykę Kripkego” (dla logik modalnych). W rzeczywistości jednak semantyka tego typu została mniej więcej jednocześnie sformułowana przez kilka różnych osób (m.in. Hintikkę). Deskrypcja nie stosuje się więc do dokładnie jednego przedmiotu. To jednak nie sprawia, że nazwa „Kripke” nie ma odniesienia.

[ad 4] Czy wiemy *a priori*, że jeżeli nazwa „RU” do czegoś się odnosi, to jest prawdą, że RU jest ABEFA? Nie. To wymaga pewnej wiedzy o rzeczywistości (i o filozofii).

Wariant [D–] jest więc również nieprzekonujący.

Jako alternatywę wobec deskrytywizmu Kripke zarysowuje swoją **PRZYCZYNOWĄ TEORIĘ ODNIESIENIA**. Wedle tej teorii nazwa najpierw zostaje wprowadzona na mocy jakiegoś rodzaju aktu „chrztu”; następnie inni użytkownicy uczą się tej nazwy od pierwszych użytkowników, po prostu używając jej i w zamierzeniu odnosząc się do tego, do czego odnosili się ci pierwsi użytkownicy. W ten sposób nowi użytkownicy uczą się nazwy, nie ucząc się przy okazji żadnej deskrypcji, a może nawet kojarząc sobie z nazwą deskrypcje niewłaściwe.

Samemu sformułowaniu Kripkego daleko jeszcze do pełnej teorii: nie jest wyjaśnione dokładniej, jakiego typu łańcuch przyczynowy stanowi odniesienie w sposób właściwy i jaka jest jego natura. Niewiele jest też powiedziane na temat aktów „chrztu”.

Jeden z głównych nurtów literatury filozofii języka drugiej połowy XX wieku to z jednej strony próby przeformułowania deskrytywizmu tak, aby nie stosowały się do niego powyższe zarzuty, a z drugiej strony krytyka tych wersji i dalszy rozwój przyczynowej teorii odniesienia.

Oto jedna z ciekawych konsekwencji zjawisk opisanych w tym rozdziale. Gwiazdę wieczorną nazwijmy „Hesperus”, a gwiazdę poranną nazwijmy „Phosphorus”. Przez jakiś czas myślano, że Hesperus i Phosphorus to dwie różne gwiazdy. Dopiero po jakimś czasie okazało się, że to jedna i ta sama planeta Wenus.

Zarówno „Hesperus”, jak i „Phosphorus” to nazwy własne i jako takie są desygnatorami sztywnymi: odnoszą się do tego samego przedmiotu w każdym świecie możliwym, w którym ten przedmiot istnieje, i nie odnoszą się do niczego innego w żadnym świecie możliwym.

Skoro tak jest, to jeżeli identyczność „Hesperus” = „Phosphorus” jest prawdziwa, to prawdziwa musi być w każdym świecie możliwym. Zatem jest konieczna. Z drugiej jednak strony odkrycie jej prawdziwości wymagało pewnych badań. Nie jest więc *a priori*. Otrzymujemy więc przykład zdania, które jest konieczne, a nie jest *a priori*.

Kripke rozważa też terminy naturalnorodzajowe, takie jak „woda”, czy „H₂O” i uznaje je za desygnatory sztywne (zdejmuje bowiem jego zdaniem test sztywności). Jeżeli ma rację, to „woda = H₂O” jest identycznością konieczną, chociaż odkrytą dopiero po jakimś czasie. Jeżeli więc zgadzamy się z Kripkem, otrzymujemy kolejny przykład koniecznej identyczności, która nie jest *a priori*.

Ten tok myślenia doprowadził Kripkego do rozwinięcia jednego z ważniejszych w XX wieku argumentów za dualizmem w kwestii umysł – ciało, który przedstawię tutaj tylko w bardzo krótkim zarysie. Terminy takie jak „ból” czy „stan pobudzenia włókien C” uważa Kripke również za terminy naturalnorodzajowe. Gdyby więc identyczność między nimi była prawdziwa, byłaby konieczna. Kripke jednak uważa, że intuicyjnie możemy dostrzec, że nie jest konieczna. Zatem *modus tollendo tollens*, nie jest prawdziwa! Bardziej szczegółowa ocena poglądów Kripkego i ocena tego argumentu będą jednak musiały poczekać na inną okazję.

Problemy z deskrytywizmem

TEKSTY POLECANE



1. Klasycznym tekstem jest *Naming and Necessity* S. Kripkego. Tekst ma wiele wydań.
2. Bardzo przystępnym podręcznikiem do filozofii języka jest *Philosophy of Language: A Contemporary Introduction* W.G. Lycana. Lycan dość szeroko omawia kwestie związane z nazwami własnymi.
3. Innym dobrym podręcznikiem filozofii języka jest *An Introduction to the Philosophy of Language* M. Morrisa.
4. Zdecydowanie warta lektury jest część siódma drugiego tomu *Philosophical Analysis in the Twentieth Century* S. Soamesa.
5. Interesującą alternatywą względem poglądów Kripkego na terminy naturalnorodzajowe są poglądy J. LaPorte'a przedstawione w książce *Natural Kinds and Conceptual Change*.

PROBLEMY Z DESKRYPTYWIZMEM



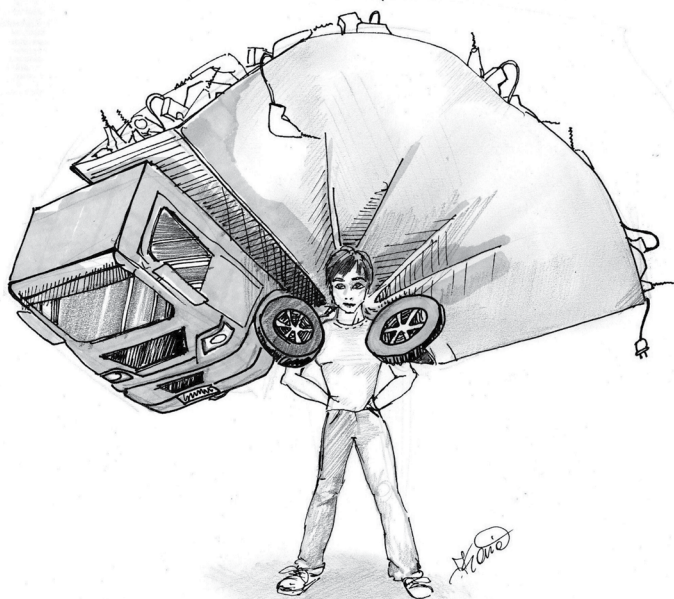
1. Deskrytywizm silny: nazwy własne mają to samo znaczenie, co przypisane im deskrypcje określone.
2. Argument modalny:
 - gdyby nazwa **N** i deskrypcja **D** znaczyły to samo, to „Jeżeli **N** istnieje, to **N** jest **D**” byłoby zdaniem koniecznym.
 - Ale nie jest: **N** może nie spełniać deskrypcji **D**.
 - Zatem **N** i **D** nie znaczą tego samego.
3. Searle: zlepek deskrypcji związanych z nazwą; wystarczy spełniać większość z nich.
4. To nie poprawia sytuacji: można nie spełniać większości z nich.
5. Desygnator sztywny desygnuje dany przedmiot w każdym możliwym świecie, w którym on istnieje, i nie desygnuje niczego innego niż ten przedmiot w żadnym świecie.
6. Nazwy własne są desygnatorami sztywnymi. Deskrypcje w większości nie są.
7. Test sztywności: czy przedmiot, który jest **j**, nie mógłby istnieć, nie będąc **j**, i czy nic innego niż ten przedmiot nie mógłby być **j**?
8. Nie interesują nas scenariusze, w których język się zmienia (np. przedmioty dostają inne nazwy).
9. Niektóre deskrypcje są sztywne (np. matematyczne).

10. Deskrypcje można usztywniać: „przedmiot, który rzeczywiście jest *D*”.
11. Deskrypcje usztywnione też nie są równoznaczne z nazwami, bo „O ile rzeczywiste *D* istnieje, to rzeczywiste *D* jest rzeczywistym *D*” jest *a priori* i konieczne, a „O ile *N* istnieje, to *N* jest rzeczywistym *D*” nie.
12. Słabsza wersja deskryptywizmu: deskrypcja nie jest synonimem nazwy, ale wyznacza jej odniesienie. Problemy:
 - Większość deskrypcji, które kojarzymy, nie wyznacza jednego przedmiotu.
 - Mylne kojarzenie deskrypcji nie zmienia odniesienia nazwy (np. Peano).
 - Brak odniesienia kojarzonej deskrypcji też nie zmienia odniesienia nazwy.
 - Dalej nie ma wiedzy *a priori*, że *N* jest *D*. A powinna być, jeżeli jest regułą języka, że odniesienie nazwy wyznaczane jest przez deskrypcję.
13. Alternatywa: przyczynowa teoria odniesienia.
14. Przykłady zdań koniecznych *a posteriori*.

37. Pojęcie operatora intensjonalnego

Jest fizycznie niemożliwe, żebym uniósł na plecach ciężarówkę załadowaną wiertarkami. Gdyby taka ciężarówka znalazła się nagle na moich plecach, wymusiłaby zmianę mojej geometrii skrzydeł. Nie jest jasne, czy umieszczenie takiej ciężarówki na moich plecach byłoby moralnie dopuszczalne. Niezależnie od tego wszystkiego możemy sobie wyobrazić świat, w którym posiadam jakieś supermoce (poza tymi, które rzeczywiście posiadam), które pozwalają mi na noszenie ciężarówek z wiertarkami. Świat taki jest metafizycznie możliwy, choć niemożliwy fizycznie. Nawet jednak w takim świecie żadna supermoc nie jest w stanie sprawić, żebym rozumiał logikę, nie rozumiejąc logiki.

Pojęcie operatora intensjonalnego



Opis powyższy stanowi zbiór zdań posługujących się rozmaitymi **OPERATORAMI INTENSJONALNYMI**. Operatory intensjonalne (nazwa brzmi mądrze, ale sprawa jest w sumie prosta) to pewne wyrażenia, które służą do budowy zdań (zazwyczaj z innych zdań). Aby wyjaśnić, czym są, prościej zacząć od wytłumaczenia, czym są operatory **EKSTENSJONALNE**.

Dośłownie rzecz biorąc, **EKSTENSJA** jest to, do czego wyrażenie się odnosi: niektórzy (z różnych dziwnych powodów, o których szkoda tu mówić) uważają, że nazwy (takie jak „Zenobiusz”) odnoszą się do przedmiotów (takich jak Zenobiusz we własnej osobie), predykaty (takie jak „jest żyroskopem”) odnoszą się do własności (takich jak sama własność bycia żyroskopem), a zdania (takie jak „Zenobiusz jest żyroskopem”) odnoszą się do **WARTOŚCI LOGICZNYCH**: prawdy lub fałszu, zależnie od tego, czy zdanie jest prawdziwe, czy fałszywe.

ZDANIOWE OPERATORY EKSTENSJONALNE są to więc pewne wyrażenia, za pomocą których budujemy bardziej złożone zdania z innych zdań, np.:

Nieprawda, że...

... i ...

Nie wszystkie takie wyrażenia: jedynie takie, dla których, gdy tylko wiemy, czy zdania składowe są prawdziwe, czy fałszywe, wiemy też od razu, czy prawdziwe są całe wyrażenia złożone. Bardziej pretensjonalny sposób opisywania tej własności to powiedzenie, że wartość logiczna wyrażeń złożonych jest funkcją (= jest wyznaczona przez) wartości logicznej wyrażeń składowych.

NUDNA TABELKA
Jeżeli wiem, że jakieś zdanie A jest prawdziwe, to wiem, że „Nieprawda, że A ” jest fałszywe. I odwrotnie. Jeżeli wiem, że A jest fałszywe, to wiem, że „Nieprawda, że A ” jest prawdziwe.
Jeżeli oba zdania A i B , osobno wzięte, są prawdziwe, to prawdą jest też zdanie „ A i B ”. W innym przypadku jest ono fałszem.

Niestety sprawa nie jest całkiem prosta, bo istnieją operatory zdaniowe, które nie posiadają tej własności. Na przykład sama wiedza na temat tego, czy jest prawdą, że **A**, nie wystarczy do ustalenia, czy „Jest konieczne, że **A**” jest prawdą. Nawet jeżeli **A** jest prawdą, to istnieją przykłady zdań bardziej złożonych o tej samej postaci, ale różniące się wartością logiczną:

A	Czy A ?	Czy jest konieczne, że A ?
$2 + 2 = 4$	Tak	Tak
Kapitan Bomba jest lepszy niż Chuck Norris.	Tak	Tak
Nie ma mnie w biurze.	Tak	Nie

Pojęcie operatora intensjonalnego

Sam fakt, że zdanie jest prawdziwe, nie wystarcza do oceny, czy jest konieczne. Podobne zjawisko dotyczy takich wyrażań jak „Chuck Norris wie, że” (zwanym też **OPERATORAMI EPISTEMICZNYMI**):

A	Czy A?	Czy Chuck Norris wie, że A?
$2 + 2 = 4$	Tak	Prawdopodobnie
Kapitan Bomba jest lepszy niż Chuck Norris.	Tak	Tak
Nie ma mnie w biurze.	Tak	Nie (mam nadzieję)

Ponownie: sama wiedza o tym, czy dane zdanie jest prawdziwe, czy fałszywe, nie wystarcza do ustalenia, czy jest przedmiotem wiedzy Chucka Norrisa.

Operatory intensjonalne są to więc operatory nieekstensjonalne. Stanowią pewne zjawisko językowe, które fajnie by było lepiej zrozumieć, a nie możemy go zrozumieć za pomocą zwykłych reguł, takich jak te w NUDNEJ TABELCE. Jest mało prawdopodobne, żeby ktoś zabrnął tak daleko w tej książce, a nie zgadzał się, że fajnie jest je lepiej zrozumieć. Co zrobić?

Warto zauważyć, że operatory te mają więcej ze sobą wspólnego, niż mogłoby się wydawać. Spójrzmy jeszcze raz na pierwszy akapit, zarysowując pewną hipotezę na temat tego, jak powiązać operatory tam występujące.

Jest fizycznie niemożliwe, żebym uniósł na plecach ciężarówkę załadowaną wiertarkami.

Fizyczna niemożliwość wydaje się związana z niemożliwością logiczną: coś jest fizycznie niemożliwe, jeżeli jest logicznie wykluczone przez prawa przyrody.

Gdyby taka ciężarówka znalazła się nagle na moich plecach, wymusiłaby zmianę mojej geometrii skrzydeł.

Tu także modalność przyczynowa (przynajmniej na pierwszy rzut wiertarką) związana jest z wymuszaniem logicznym przez prawa przyrody i opis stanu początkowego.

Nie jest jasne, czy umieszczenie takiej ciężarówki na moich plecach byłoby moralnie dopuszczalne.

Moralną dopuszczalność można związać z niewykluczeniem logicznym przez zasady moralne.

Możemy sobie wyobrazić świat, w którym posiadam jakieś supermoce (poza tymi, które rzeczywiście posiadam), które pozwalają mi na noszenie ciężarówek z wiertarkami. Świat taki jest metafizycznie możliwy, choć niemożliwy fizycznie.

Możliwość metafizyczna, podobnie jak możliwość fizyczna, interpretowana może być jako niewykluczenie się logiczne z prawdami metafizycznymi.

Żadna supermoc nie jest w stanie sprawić, żebym rozumiał logikę, nie rozumiejąc logiki.

Tutaj mamy do czynienia z niemożliwością ze względu na znaczenie wyrażen (niemożliwością analityczną). Nie jest jasne, czy da się ona zinterpretować w analogiczny sposób, a jej specyficznym rodzajem jest niemożliwość logiczna: niemożliwość ze względu na znaczenie operatorów logicznych, takich jak „nieprawda, że”, „każdy”, „pewien” itd. Głównie tą modalnością (i może trochę możliwością metafizyczną) będziemy się przez chwilę zajmować⁵¹.

51 To, czy wywód powyższy dowodzi, że owe możliwości „redukują się” do innych, jest niejasne, sprawa jest jednak zbyt szczegółowa, żeby teraz ją rozważać.

38. Paradoksy implikacji a modalności

Historycznie najwcześniejsze podejście do operatorów modalnych to podejście aksjomatyczne. Wedle tego podejścia, aby lepiej zrozumieć operator konieczności, należy posłużyć się po prostu aksjomatami i regułami wnioskowania.

Zaczęło się od tego, że logika klasyczna posługuje się **IMPLIKACJĄ MATERIALNĄ** ($p \rightarrow q$ – „jeżeli p , to q ”), która jest fałszywa, jeżeli poprzednik (p) jest prawdziwy, a następnik (q) fałszywy, a prawdziwa we wszystkich innych przypadkach.

CLARENCE IRVING LEWIS (1883–1964), jak i wielu innych myślicieli, zauważył, że takie rozumienie implikacji sprawia, że logicznie poprawne są na przykład następujące rozumowania⁵²:

$$q \models p \rightarrow q \quad (22)$$

$$\neg p \models p \rightarrow q \quad (23)$$

$$p \rightarrow q \models p \wedge r \rightarrow q \quad (24)$$

$$(p \wedge q \rightarrow r) \models (p \rightarrow r) \vee (q \rightarrow r) \quad (25)$$

Gdyby jednak implikacja materialna nadawała się do modelowania okresów warunkowych w języku potocznym, to znałoby, że poprawne logicznie są rozumowania następujące:

- ad (22) Jestem w Gdańsku. Zatem gdybym był w Azerbejdżanie, byłbym w Gdańsku.
- ad (23) Nie jestem w Azerbejdżanie. Zatem gdybym był w Azerbejdżanie, byłbym w Gdańsku.
- ad (24) Jeżeli pogoda będzie ładna, pójdziemy paść predatory. Zatem jeżeli pogoda będzie ładna, a mnie porwą kosmici, pójdziemy paść predatory.

⁵² Logicznie poprawny argument to argument, którego przesłanki gwarantują prawdziwość wniosku. Wynikanie zapisywać będę jako „ $\models$ ”.

- ad (25) Jeżeli predator jest w okolicy i Obcy jest w okolicy, dojdzie do walki między nimi. Zatem albo (jeżeli predator jest w okolicy, dojdzie do walki między nimi), albo (jeżeli Obcy jest w okolicy dojdzie do walki między nimi).

W świetle nieintuicyjności takich rozumowań Lewis na początku XX wieku (1918) starał się skonstruować system, w którym aksjomatyzował nieco silniejszą relację implikacji: implikację konieczną, zwaną też **IMPLIKACJĄ ŚCISŁĄ**. Wyrażenie „ $p \prec q$ ” czytamy „ p ściśle implikuje q ” i jest ono równoważne wyrażeniu „ $\Box(p \rightarrow q)$ ” („jest konieczne, że p materialnie implikuje q ”).

Krótką uwagę o systemach aksjomatycznych. Jednym z klasycznych sposobów opisywania systemu logicznego jest opisanie go na poziomie syntaktycznym: podajemy, jak wygląda język (a właściwie jego gramatyka), którym będziemy się posługiwać, jakie są podstawowe założenia systemu i jak możemy z tych założeń wyprowadzać nowe twierdzenia. Określenie, jak wyprowadzać nowe twierdzenia, również odbywać się musi na poziomie syntaktycznym, a więc mówi się takie rzeczy, jak na przykład „jeżeli twierdzeniem jest wyrażenie ϕ i twierdzeniem jest wyrażenie ψ , to twierdzeniem jest wyrażenie $\phi \wedge \psi$ ” (dołączanie koniunkcji), albo „jeżeli twierdzeniem jest wyrażenie $\phi \rightarrow \psi$ i twierdzeniem jest wyrażenie ϕ , to twierdzeniem jest wyrażenie ψ ”. Dla oznaczania dowodliwości, która rozumiana jest właśnie w ten syntaktyczny sposób, używa się symbolu $\vdash$.

Takie określenie systemu logicznego skonstrastowane jest z podaniem bardziej ogólnej teorii warunków prawdziwości wyrażen rozważanego języka (aby to uczynić, potrzebujemy posłużyć się pojęciem modelu, matematycznego odpowiednika tego, o czym wyrażenia danego języka mogą być prawdziwe) i zdefiniowaniem wynikania jako zachowywania prawdziwości w każdym modelu (zdanie wynika z przesłanek, jeżeli jest prawdziwe w każdym modelu, w którym prawdziwe są przesłanki).

Paradoksy implikacji a modalności

Tak rozumianą relację wynikania czy konsekwencji oznacza się zwykle przez $\models$ i do jej określenia nie trzeba w ogóle podawać żadnych syntaktycznych reguł przeprowadzania dowodów. Dla wielu systemów logicznych wynikanie pokrywa się z dowodliwością, ale nie dla wszystkich.

Lewis zaczął od sformułowania systemu zwanego **S1**, który posiadał następujące aksjomaty:

$$\begin{aligned}(p \wedge q) &\prec (q \wedge p) \\ (p \wedge q) &\prec p \\ p &\prec (p \wedge p) \\ ((p \wedge q) \wedge r) &\prec (p \wedge (q \wedge r)) \\ ((p \prec q) \wedge (q \prec r)) &\prec (p \prec r) \\ (p \wedge (p \prec q)) &\prec q\end{aligned}$$

Niedługo potem dodał⁵³.

$$\Diamond(p \wedge q) \prec \Diamond p$$

otrzymując system **S2**. Jeżeli zamiast tego dodać do **S1** aksjomat:

$$(p \prec q) \prec (\neg \Diamond q \prec \neg \Diamond p)$$

otrzymamy system **S3**. Dodając do **S1** aksjomat:

$$\neg \Diamond \neg p \prec \neg \Diamond \neg \neg \Diamond \neg p$$

otrzymamy system **S4**, a dodając do **S1** aksjomat:

$$\Diamond p \prec \neg \Diamond \neg \Diamond p$$

otrzymamy system **S5**.

⁵³ Wyrażenie „ $\Diamond p$ ” czytamy „jest możliwe, że p ”.

Jeżeli rzeczywiście rozumiemy symbol $\prec$, nie mamy też problemu z rozumieniem konieczności: $\Box q$ możemy zdefiniować na przykład przez $p \vee \neg p \prec q$. Chociaż już w 1932 r. Lewis zmienił nieco zdanie, uznawał $\Diamond$ za operator pierwotny i definiował $p \prec q$ jako $\neg \Diamond (p \wedge \neg q)$ ⁵⁴.

Przez dłuższy czas jednym z ulubionych sportów logików nieklasycznych było konstruowanie rozmaitych aksjomatyzacji modalności, porównywanie ze sobą siły tych systemów i obrona lub krytyka trafności tych lub tamtych aksjomatów.

Powinno już być jasne, na czym polega jeden z głównych problemów z tym podejściem. Bardzo szybko można konstruować nowe aksjomatyzacje i nie jest nawet jasne, w jakiej relacji do siebie pozostają. Czy są równoważne? Czy jedna jest słabsza od drugiej? Czy zbiory ich twierdzeń się krzyżują? Dopóki pozostajemy na poziomie syntaktycznym, odpowiedzi na te pytania są zazwyczaj bardzo trudne. Co gorsza, różnice w aksjomatyzacjach są na tyle drobne, że intuicja szybko staje się bezradna w obliczu całych wiader możliwych dodatkowych aksjomatów.

Wiąże się z tym inna słabość podejścia aksjomatycznego: sam fakt, że możemy sformułować takie, a nie inne aksjomatyzacje, nic nam nie mówi o trafności tychże aksjomatów i stosowalności systemu do jakichkolwiek zjawisk w języku naturalnym. Intuicje odnośnie do trafności tych aksjomatów

54 Przy okazji z upływem czasu okazało się, że implikacja ścisła również nie nadaje się na adekwatne narzędzie analizy okresów warunkowych. Na przykład fakt, że zdania konieczne ściśle implikowane są przez dowolne zdania, a zdania koniecznie fałszywe ściśle implikują dowolne zdania:

$$\begin{aligned}\Box q &\models p \prec q \\ \neg \Diamond p &\models p \prec q\end{aligned}$$

sprawia, że poprawne byłyby następujące rozumowania:

Jest konieczne, że $2 + 2 = 4$. Zatem gdyby matematyka zawiodła i $2 + 2$ równało się 5, $2 + 2$ równałoby się 4.

Nie jest możliwe, żeby znalazł nas tu predator. Zatem gdyby nas znalazł, z łatwością odgryźlibyśmy mu ręce.

Poszukiwanie lepszych analiz okresów warunkowych to zadanie, z którym logicy borykają się od dłuższego czasu; osiągnięto pewien postęp, ale to opowieść na inną okazję.

Paradoksy implikacji a modalności

również nie są wystarczające. Potrzebna jest bardziej rozbudowana i systematyczna teoria znaczenia dla operatorów modalnych. Dopiero w świetle takiej teorii będziemy mogli oceniać dobór aksjomatów.

Podobnie chcielibyśmy, aby nasza aksjomatyzacja nie tylko była trafna w tym sensie, że czegokolwiek możemy dowieść, jest prawdziwe w zamierzonej interpretacji, lecz także chcemy, żeby była zupełna: jeżeli coś jest prawdą logiczną, powinno być dowodliwe w systemie aksjomatycznym. Dopóki jednak nie mamy semantyki, nie mamy żadnej metody rozstrzygania, czy dana aksjomatyzacja realizuje taki cel.

Rozwiązanie pojawiło się dopiero po jakimś czasie, pod koniec lat 50. Była nim semantyka światów możliwych.

CZĘŚĆ
VI
Światy
możliwe

39. Światy możliwe: aparatura

Kluczowym elementem semantyki światów możliwych jest, ku zaskoczeniu czytelnika, pojęcie światów możliwych. Nie wszyscy zgadzają się co do tego, czym światy możliwe są (o tym problemie będziemy jeszcze mówić). Zasadniczo jednak wszyscy zgadzają się co do tego, że świat, w którym istniejemy, mógłby być inny, niż rzeczywiście jest, i każda z tych możliwości to świat możliwy. Większość zgadza się też co do dwóch minimalnych warunków, które taki świat musi spełniać:

- Dla każdego zdania p (lub sądu w sensie logicznym) albo p , albo $\neg p$ jest prawdziwe w tym świecie. Warunek ten nazywa się **warunkiem maksymalności**.
- Dla żadnego zdania p nie jest tak, że zarazem p i $\neg p$ jest prawdziwe w tym świecie. Warunek ten nazywany jest **warunkiem niesprzeczności**.

Kolejnym elementem jest **RELACJA DOSTĘPNOŚCI** pomiędzy światami możliwymi. Świat v jest dostępny ze świata w , jeżeli z perspektywy świata w świat v jest możliwy. W pewnym sensie to, jakie pojęcie możliwości chcemy modelować, ma wpływ na to, jak relacja dostępności będzie wyglądać. Na przykład jeżeli mówimy o możliwości fizycznej, z naszego świata nie jest dostępny fizycznie świat możliwy, w którym podnoszę słonia jedną ręką, choć dostępny jest świat, w którym słoń podnosi mnie trąbą.

W przypadku możliwości logicznej wygląda na to, że relacja dostępności jest nieistotna: wszystkie światy możliwe są logicznie możliwe, a więc wszystkie światy możliwe „widzą” siebie nawzajem. Logiką, która powstaje w rezultacie, jest wspomniana już logika **S5** (o niej też będę głównie mówić). Jeżeli ignorujemy relację dostępności, warunki prawdziwości wyrażen modalnych określa się, mówiąc:

$\Diamond p$ jest prawdziwe w świecie w wtedy i tylko wtedy, gdy istnieje taki świat v , w którym p jest prawdziwe.

Innymi słowy, dane zdanie jest możliwe (w świecie w), jeżeli istnieje świat możliwy (powiedzmy, v), w którym jest ono prawdziwe. Formalnie:

$$w \models \Diamond p \text{ wtw } \exists v v \models p$$

Analogicznie definiuje się konieczność: coś jest konieczne, jeżeli jest prawdą w każdym świecie możliwym:

$$w \models \Box p \text{ wtw } \forall v v \models p$$

Tego typu definicje skutkują szeregiem zalet teoretycznych.

- Po pierwsze, pozwalają sprowadzić wzajemną definiowalność modalności do praw obowiązujących kwantyfikator. Jeżeli $\Diamond p$ jest prawdziwe w świecie w , to p jest prawdziwe w jakimś świecie v , a więc nie jest tak, że $\neg p$ jest prawdziwe w każdym świecie – więc też i nie jest prawdą, że $\neg p$ jest konieczne:

$w \models \Diamond p$	wtw	$\exists v v \models p$
	wtw	$\exists v v \not\models \neg p$
	wtw	$\neg \forall v v \models \neg p$
	wtw	$w \not\models \Box \neg p$

- Po drugie, pozwalają na jednolite wyjaśnienie prawdziwości szeregu prawd modalnych za pomocą kwantyfikacji. Na przykład jeżeli coś jest konieczne w w , to znaczy, że jest prawdziwe we wszystkich światach możliwych. Z tego zaś wynika, że jest też prawdziwe w w :

$w \models \Box p$	wtw	$\forall v v \models p$
	więc	$w \models p$

Światy możliwe: aparatura

Prawem logiki **S5** okazuje się więc:

$$\Box p \rightarrow p \quad (\text{Aksjomat T})$$

- Po trzecie, aparatura światów możliwych razem z relacją dostępności pozwala na dość klarowne usystematyzowanie rozmaitych logik modalnych. Najprostsza logika **K** powstaje, gdy nie zakładamy nic na temat relacji dostępności. Oprócz klasycznej logiki zdań ma tylko jeden aksjomat:

$$\Box(p \rightarrow q) \rightarrow (\Box p \rightarrow \Box q) \quad (\text{Aksjomat K})$$

a oprócz standardowych reguł wnioskowania logiki klasycznej ma tylko **REGUŁĘ NECESYTACJI** lub **UKONIECZNIANIA** (zwaną też **REGUŁĄ GÖDELA**), która powiada, że jeżeli coś jest twierdzeniem logiki, to jest konieczne. Gdy tylko dodamy założenie, że każdy świat widzi sam siebie, automatycznie jako twierdzenie dostajemy coś, co twierdzeniem w **K** nie było, mianowicie (**AKSJOMAT T**). Logika taka aksjomatyzowana jest przez dodanie tej formuły do aksjomatu **K** i istnieje w miarę powszechna zgoda, że nadaje się do modelowania konieczności fizycznej.

- Po czwarte, taka semantyka pozwala na sprawdzanie zupełności systemu aksjomatycznego: gdy tylko mamy semantykę dla języka modalnego, to możemy zastanawiać się, czy nasza aksjomatyzacja pozwala na dowiedzenie wszystkiego, co powinno być dowodliwe. I faktycznie istnieje wiele wyników, w których pokazuje się, że pewne aksjomatyzacje dowodzą wszystkich prawd logicznych w danej semantyce.

Pojęcie światów możliwych okazuje się też przydatne do analizy innych pojęć filozoficznych:

- Sądy w sensie logicznym (czyli znaczenia zdań, które są prawdziwe lub fałszywe) identyfikowane są ze zbiorami

światów możliwych. Intuicyjnie każdy sąd jest identyfikowany ze zbiorem tych i tylko tych światów, w których jest prawdziwy.

- Własności rozumie się często jako funkcję z przedmiotów i światów możliwych w wartości logiczne. Czyni się to, by uniknąć identyfikacji tych własności, które posiadane są w świecie rzeczywistym przez dokładnie te same przedmioty. Na przykład własność bycia człowiekiem, w świecie rzeczywistym posiadana jest przez te same obiekty, które mają własność bycia nieupierzonym dwunogiem. Niemniej istnieją światy możliwe, w których istnieją nieupierzone dwunogi, które nie są ludźmi, i istnieją światy możliwe, w których istnieją ludzie, którzy są upierzeni (to drugie brzmi ciekawiej). Dlatego też patrzenie nie tylko na świat rzeczywisty, lecz także na (inne) światy możliwe pozwala na rozróżnienie tych własności, nawet jeżeli patrzymy tylko na zbiory obiektów je posiadających.
- Rozważane pojęcia pojawiają się bardzo często na przykład w filozofii umysłu. Pojęcie światów możliwych używane jest do wyjaśnienia pojęcia redukowania własności (pierwszy rzut wiertarką: własność *A* redukuje się do własności *B*, gdy nie ma świata możliwego, w którym coś posiada własność *A*, nie posiadając własności *B*) czy pojęcia superweniencji (własności *A* superweniują na własnościach *B*, gdy nie może być różnic we własnościach typu *A* bez różnic we własnościach typu *B*).
- Światy możliwe (wraz z relacją podobieństwa między nimi) występują w większości współczesnych teorii okresów warunkowych. A okresy warunkowe występują gdzie popadnie w filozofii.
- Niemodalne operatory intensjonalne, takie jak „obowiązkowo”, też znalazły swoje miejsce w semantyce światów możliwych. Na przykład relacja dostępności jest zwrotna dla możliwości metafizycznej (a zatem, cokolwiek jest konieczne, jest też prawdziwe). W logikach modelujących

Światy możliwe: potrzeba zrozumienia

zaś obowiązek „□” czyta się raczej jako „obowiązkowo” lub „powinno być tak, że”, a „◇” jako „jest dopuszczalne, że”. Na gruncie takich logik nie będziemy chcieli, żeby formuła „□ $p \rightarrow p$ ” (cokolwiek powinno mieć miejsce, ma miejsce) była tezą. W związku z tym musimy dopuścić modele, w których relacja dostępności nie jest zwrotna; bo też i dostępność oznacza coś innego: świat v jest dostępny ze świata w , jeżeli z punktu widzenia świata w świat v jest dopuszczalny moralnie.

40. Światy możliwe: potrzeba zrozumienia

Mamy więc pewną aparaturę pojęciową – aparaturę światów możliwych – która okazuje się przydatna z różnych powodów. Pozostaje zapytać: dlaczego? Czym są i jak istnieją światy możliwe?

Jedna możliwa odpowiedź to stwierdzenie, że światy możliwe są tylko heurystycznie i eksplanacyjnie pożytecznym mitem, bajeczką, którą fajnie jest się posługiwać, bo podsuwa ciekawe pomysły i pomaga w wyjaśnianiu pewnych rzeczy.

Problem z tą odpowiedzią polega na tym, że nie wiadomo, dlaczego należałoby w naszym aparacie poznawczym posługiwać się mitami i bajeczkami i dlaczego miałyby one pełnić funkcję poznawczą. Meteorologowie nie przyjmują istnienia Neptuna (boga, nie planety), żeby wyjaśnić sobie sztormy, astrofizycy nie przyjmują istnienia czarnoksiężników, żeby wyjaśnić funkcjonowanie czarnych dziur.

Naukowcy mogą przyjmować teorie, które są dosłownie fałszywe (jak mechanika Newtona), żeby coś wyjaśnić. W takich jednak przypadkach nie poprzestają na mówieniu, że jest to pożyteczna bajeczka, tylko wyjaśniają, **dlaczego działa**: poka-

zują, czemu w większości zastosowań, z powodów całkiem nieidiotycznych, daje ona rezultaty wystarczająco trafne, żeby się na nich opierać w życiu codziennym.

Dlaczego więc światy możliwe miałyby działać? Dlaczego teoria światów możliwych miałaby porządkować pole bitwy, pomóc wybierać logiki modalne zależnie od zastosowań, pomagać dowodzić twierdzeń i własności wybranych logik, jeżeli nie jest dosłownie prawdziwa? A może jest dosłownie prawdziwa? Właściwie dlaczego nie?

Jeżeli nie zgodzimy się, że teoria światów możliwych jest po prostu przydatną bajką, możemy być zainteresowani realistyczną teorią światów możliwych, która traktuje tę aparaturę teoretyczną poważnie.

41. Realizm modalny Davida Lewisa

Jedną z osób, która terminologię światów możliwych przyjęła całkiem serio, był David Lewis (nie należy mylić go z C.I. Lewisem, twórcą aksjomatycznych systemów modalnych). Uważał on bowiem, że światy możliwe istnieją tak samo jak nasz świat (z tym, że są od siebie czasoprzestrzennie oddzielone) i że nie różnią się one od niego rodzajem, a tylko tym, co się w nich dzieje.

W pewnym sensie realizm Lewisa jest wariantem redukcjonizmu: modalności, warunki prawdziwości zdań z operatorami modalnymi oraz warunki prawdziwości zdań o światach możliwych stara się on sprowadzić do wydarzeń, które mają (lub nie) takie samo istnienie jak żyrafy czy wiertarki, tyle że czasoprzestennie niedostępnie i dosyć duże.

Powody, by taką koncepcję uznawać, miał następujące. Po pierwsze, niekontrowersyjne jest, że rzeczy mogły potoczyć się

inaczej, niż się potoczyły, że świat mógłby być inny, niż jest. To jednak powoduje, że są pewne stany, w których świat mógłby być, ale w których nie jest; światy, które są jedynie możliwe.

Lewis więc całkiem dosłownie zinterpretował kwantyfikację języka potocznego. Filozoficznie nie jest to ruch zawsze godny polecenia. Dlatego też i Lewis szybko dodaje, że ruch ten jest uprawniony tylko, jeżeli (1) uczynienie go w danym przypadku nie prowadzi do problemów oraz (2) nie istnieją lepsze sposoby interpretacji rozważanych zdań, które nie prowadzą do problemów.

Co do (1), jakie problemy zdaniem Lewisa może napotkać jego realizm?

1. Można powiedzieć, że jest fałszywy, ponieważ istnieje tylko świat rzeczywisty. Ten zarzut jednak po prostu zakłada fałszywość poglądu Lewisa.
2. Może być tak, że ktoś zarzuci realizmowi **ROZRZUTNOŚĆ ONTOLOGICZNĄ**, postulowanie istnienia większej liczby bytów niż trzeba. Lewis tutaj odróżnia **ONTOLOGICZNĄ OSZCZĘDNOŚĆ JAKOŚCIOWĄ** od oszczędności **IŁOŚCIOWEJ**: teoria jest nieoszczędna jakościowo, gdy postuluje istnienie obiektów nowego ontologicznego rodzaju; jest nieoszczędna ilościowo, gdy postuluje istnienie większej ilości obiektów znanego już i akceptowanego rodzaju. O ile rozrzutność jakościową Lewis uznaje za wadę teorii, o tyle rozrzutność ilościowa poważną wadą nie jest. Jego zaś zdaniem, realizm modalny winny jest tylko rozrzutności ilościowej. Postuluje bowiem wyłącznie istnienie innych obiektów tego samego rodzaju co obiekt już znany, mianowicie nasz własny świat.
3. Quine znany jest z zarzutu braku kryteriów indywiduacji obiektów możliwych. Wymyślił kiedyś slogan: **nie ma przedmiotu bez identyczności**, rozumiejąc przez to, że jeżeli chcemy mówić o jakichś indywiduach, to powinniśmy też powiedzieć, w jakiej sytuacji uznalibyśmy dwa takie indywidua za tak naprawdę jedno i to samo (czyli podali tzw. **KRYTERIUM IDENTYCZNOŚCI** dla rozważanych przed-

miotów). Retorycznie więc pyta: Skąd mam wiedzieć, ile możliwych predatorów czai się w tych krzakach⁵⁵? Lewis powiada jednak, że zarzut ten nie stosuje się do niego, bo wedle realizmu modalnego każdy obiekt możliwy jest ograniczony do swojego własnego świata możliwego. Stąd też w tych krzakach (na które teraz patrzę przez okno) nie ma żadnego możliwego predatora, bo wszystkie, w światach, w których istnieją, są u siebie w domu, strojąc choinkę albo robiąc inne rzeczy, które można robić, jak się jest predatorem w swoim własnym świecie.

Odnosnie do kwestii (2), jakie mamy inne sposoby interpretacji dyskursu modalnego?

1. Możemy uznać pojęcie światów możliwych albo operatory modalne za nieanalizowalne i pierwotne. To jednak, powiada Lewis, nie jest teoria, tylko odmowa konstrukcji teorii.
2. Możemy myśleć, że „ p jest możliwe” znaczy coś w stylu „ p jest niesprzeczne”. Wtedy jednak, sugeruje Lewis, powinniśmy zapytać, jak rozumiana jest tutaj niesprzeczność.
 - (a) Jeżeli zdanie niesprzeczne to takie, które może być prawdziwe, lub takie, które nie jest koniecznie fałszywe, to analiza jest kolista, bo posługuje się już modalnościami.
 - (b) Jeżeli zdanie niesprzeczne to takie, którego negacja nie jest twierdzeniem jakiegoś wybranego systemu dedukcyjnego (na przykład jakiegoś sformułowania logiki klasycznej), analiza jest nietrafna. Na przykład wszystkie prawdy arytmetyki liczb naturalnych są konieczne, a dla każdego wartego uwagi aksjomatycznego systemu arytmetyki istnieje fałsz arytmetyczny, którego negacja nie jest twierdzeniem tego systemu⁵⁶.

⁵⁵ Chyba użył innego przykładu.

⁵⁶ Lewis nawiązuje tutaj do niezupełności arytmetyki. Wspomniałem o tym zjawisku na s. 169–170.

- (c) Jeżeli zdanie niesprzeczne to takie, które jest prawdziwe przy pewnej interpretacji stałych pozalogicznych, teoria jest niepoprawna⁵⁷. Nie są bowiem z przyczyn logicznych wykluczone interpretacje, w których „świnia” i „owca” odnoszą się do tych samych własności, a mimo to nie jest możliwe, żeby świnię były owcami.
 - (d) Jeżeli będziemy chcieli poprawić poprzednią teorię, mówiąc, że zdanie jest niesprzeczne, gdy jest prawdziwe przy pewnej **dopuszczalnej** interpretacji wyrażen pozalogicznych, nasza analiza jest kolista. Dopuszczalność bowiem w tym kontekście już zawiera w sobie wbudowaną modalność.
3. Możemy analizować operatory modalne za pomocą światów możliwych, ale uznać światy możliwe za jakiegoś typu obiekty lingwistyczne. Na pierwszy rzut oka nadawałyby się do tego maksymalne niesprzeczne zbiory zdań, tj. takie zbiory zdań, które oprócz niesprzeczności spełniają ten warunek, że dodanie do nich dowolnego zdania sprawi, że będą one sprzeczne. Innymi słowy, zbiór jest maksymalny, jeżeli już „rozstrzyga” prawdziwość każdego zdania. To podejście jednak, zależnie od tego, jak rozumielibyśmy niesprzeczność, byłoby albo nietrafne, albo koliste (z powodów analogicznych do powodów wymienionych w punkcie 2)⁵⁸.

57 W myśl tej definicji na przykład zdanie „Rafał lata” jest prawdziwe, bo ma formę $P(a)$ i istnieje pewne przypisanie własności do P i przedmiotu do a , które daje zdanie prawdziwe, na przykład przypisanie własności bycia logikiem do P i mnie samego do a . Dla kontrastu zdanie „Rafał lata i nie lata” nie jest możliwe, bo ma postać $P(a) \wedge \neg P(a)$ i żadna interpretacja P i a nie da nam zdania prawdziwego o tej postaci.

58 Nieco później, w *Plurality of Worlds*, Lewis dodał jeszcze argument z użyteczności naukowej: światy możliwe należy uznać za rzeczywiste, ponieważ realizm modalny wyjaśnia wiele zjawisk dotyczących modalności, a więc teoria postulująca rzeczywiste światy możliwe jest użyteczna. Niestety, rozważenie tego argumentu wymagałoby wejścia głębiej w filozofię nauki. Wystarczy na razie zauważyć (właściwie to w zgodzie z Lewisem), że teoria użyteczna powinna być uznawana, o ile nie napotyka problemów i o ile nie ma (co najmniej) równie użytecznych teorii konkurencyjnych, które mają mniej problemów.

42. Problemy realizmu modalnego

Zauważmy najpierw, że Lewis broni się przed zarzutem rozrzutności ontologicznej, mówiąc, że realizm jest rozrzutny tylko ilościowo, a nie jakościowo: postuluje bowiem więcej bytów tego samego rodzaju, co byt nam już znany, mianowicie świat. To jednak nie jest wysoce przekonujące: światy możliwe zawierają mogą wiele różnych rzeczy, a w tym i obiekty o wysokim stopniu dzikości ontologicznej. Metafizycznie możliwe są światy z demonami, duchami, światy niematerialne, światy aniołów. Wszystkie one, wedle realizmu, istnieją rzeczywiście, tylko że w jakimś niedostępnym czasoprzestrzennie i przyczynowo miejscu. O ile sam świat jako taki jest, tak samo jak nasz świat, światem, o tyle sam fakt, że istnieć musiałyby w nim przeróżne dziwadła, sprawia, że teoria jest bardziej jakościowo rozrzutna ontologicznie, niż Lewis by chciał.

Po drugie, niezależnie od siły teoretycznej realizmu nie wygląda na to, żeby nadawał się on na **TEORIĘ ZNACZENIA** operatorów modalnych. Gdyby na mocy znaczenia było tak, że gwarantem możliwości jest rzeczywiste istnienie jakichś czasoprzestrzennie i przyczynowo niedostępnych światów, następujące zdanie powinno być analitycznie fałszywe:

Jest możliwe, że w tych krzakach jest predator, chociaż nie istnieje żaden czasoprzestrzennie i przyczynowo niedostępny świat, w którym istnieją krzaki, w których jest predator.

Można się bronić przed tym zarzutem, mówiąc, że zasada przejrzystości znaczenia (pamiętacie Moore'a?) jest i tak podejrzana, ale zasada przejrzystości znaczenia zawodzi zazwyczaj w przypadkach ogólnych i skomplikowanych. Tutaj natomiast moja intuicja dotyczy przypadku konkretnego i jest taka, że gdyby powyższe zdanie miało być analitycznie fałszywe, powin-

Problemy realizmu modalnego

niśmy być w stanie to rozpoznać. Zarzut jednak nie jest ostatecznie powalający, więc idźmy dalej.

Kolejnym problemem jest to, że pojęcie izolacji czasoprzestrzennej i przyczynowej jest dość tajemnicze. W świetle Quine'owskiego marudzenia o kryteriach identyczności Lewis musiał zdecydować się na kryterium identyczności światów możliwych i postawił na tego typu izolację. To jednak wywołuje szereg problemów:

- Jaka jest natura takiej izolacji? Skąd wiemy, że izolacja taka jest możliwa? Co ją wywołuje?
- Jeżeli rzeczywiście inne światy możliwe są tak rzeczywiste jak nasz, tyle że czasoprzestrzennie i przyczynowo oddzielone, jakim cudem jesteśmy w stanie cokolwiek o nich wiedzieć? Nie mamy teleskopów pozwalających zagłądać nam do innych światów możliwych, a sugerowanie, że wiemy coś o predatorach z krwi i kości w innych, rzeczywistych, ale odizolowanych światach, z przyczyn czysto semantycznych albo dzięki intuicji, brzmi co najmniej podejrzanie.
- Wybór izolacji czasoprzestrzennej i przyczynowej nie jest oczywisty: dlaczego nie miałyby istnieć światy, których **fragmenty** są w ten sposób odizolowane?
- Być może możliwe są również światy nieczasoprzestrzenne, zamieszkane jedynie przez anioły. Teoria Lewisa wyklucza jednak możliwość identyfikacji czegoś takiego jako świata.

Następna trudność pojawia się, gdy zauważymy, że wedle Lewisa każde indywiduum jest ograniczone do świata, w którym istnieje. To jednak sprawia, że warunki prawdziwości zdań modalnych dotyczących rzeczywiście istniejących przedmiotów są nieco dziwne.

Dlaczego na przykład jest możliwe, że zagryzam predatora? Nie możemy powiedzieć, że jest tak dlatego, że istnieje świat możliwy, w którym to robię, bo nie ma mnie w żadnym świecie poza tym.

Tego typu problemy sprawiły, że zamiast identyczności przedmiotów pomiędzy światami możliwymi Lewis sugeruje, że istnieje relacja „bycia odpowiednikiem”. Choć nie istnieje w żadnym innym świecie możliwym, to w pewnych światach posiadam odpowiedniki: przedmioty najbardziej podobne do mnie, które w niektórych z tych światów możliwych zagryzają predatory. I właśnie dlatego, że któryś z moich odpowiedników zagryza predatora w jakimś świecie możliwym, jest teraz i tutaj możliwe, że zagryzam predatora.



Tego typu rozwiązanie nie jest pozbawione problemów. Określenie odpowiednika jako indywiduum najbardziej podobnego jest podejrzane, bo w niektórych światach może być tak, że nie można jednoznacznie wybrać odpowiednika spośród kilku kandydatów. Na przykład istnieje świat możliwy, w którym jedna osoba, *A*, ma te cechy co ja, z tym wyjątkiem, że jej lewe oko jest zielone. Druga osoba, *B*, ma te cechy co ja, z tym wyjątkiem, że jej prawe oko jest zielone. Która z nich jest moim odpowiednikiem? Która z nich powinna zagryźć predatora, żeby dla mnie było możliwe zagryzienie predatora?

Innym wyzwaniem jest wyjaśnienie związku pomiędzy tym, co dzieje się z moim odpowiednikiem, a tym, co jest możliwe dla mnie. Jeżeli jestem ścigany przez predatora, a w jakimś odległym, czasoprzestrzennie i przyczynowo niedostępnym świecie jakiś gość, który wygląda jak ja, tylko że ma jedno oko zielone, zagryzł predatora, to jakie pocieszenie to dla mnie stanowi? Jak z tego ma wynikać, że ja mogę też zagryźć predatora?

Same argumenty Lewisa przeciwko alternatywnym teoriom modalności też nie są do końca przekonujące. Zarzuca on na przykład modelowi teoretycznej koncepcji możliwości, że nie wyklucza ona istnienia świniowców (czy też pokrywania się ekstensji predykatu „świnia” z ekstensją predykatu „owca”). To znaczy jednak, po pierwsze, że sam wie, że nie istnieją światy możliwe ze świniowcami – skąd? Wygląda na to, że Lewis ma jakiś uprzywilejowany dostęp poznawczy do światów możliwych, pozwalający mu na ustalenie, które z nich nie istnieją; no chyba, że z góry wie, jakie światy mają być możliwe, a potem postuluje je jako rzeczywiste. To by jednak znaczyło, że jego intuicje dotyczące tego, co jest możliwe, są poznawczo uprzednie w stosunku do tego, co wie on o istnieniu światów możliwych.

Po drugie, znaczy to, że jego zarzut sprowadza się do tego, że logiczna koncepcja możliwości nie obejmuje modalności analitycznych. To, co konieczne logicznie, będzie konieczne analitycznie, ale raczej nie odwrotnie. Czy rzeczywiście zwolennik logicznej koncepcji możliwości nie może określić konieczności analitycznej? Czy nie wystarczy powiedzieć, że coś jest konieczne analitycznie, jeżeli jest konieczne logicznie lub jest logiczną konsekwencją tych zdań, które są prawdziwe ze względu na znaczenie? Gdzie tutaj potrzeba odwoływania się do rzeczywistych światów możliwych?

Wcale także nie jest jasne, dlaczego pojęcie niesprzeczności ma wymagać jakichś modalności. W samej definicji modelu zbioru zdań (= dziedziny razem z funkcjami przypisującymi podzbiory dziedziny predykatom oraz elementy dziedziny stałym indywiduuowym) nie ma żadnych pojęć modalnych,

a spełnialność definiuje się niemodalnie za pomocą istnienia rzeczywistej struktury matematycznej. Jasne, że zbiór będzie niesprzeczny, jeżeli będzie spełnialny w jakimś modelu, ale pojęcie spełnialności nie zawiera tutaj żadnego istotnie modalnego elementu: spełnialność w jakimś modelu znaczy po prostu bycie spełnionym w pewnym istniejącym modelu, który jest po prostu obiektem matematycznym.

Inną sprawą jest to, że niezależnie od spełnienia warunków (1) i (2) argument Lewisa jest nieco pośpieszny. Z tego, że najlepszą metodą modelowania dyskursu o światach możliwych jest semantyka postulująca ich rzeczywiste istnienie (zgódźmy się na chwilę), nie wynika, że światy możliwe istnieją rzeczywiście. Tak samo jak z tego, że najlepszą semantyką dla modelowania dyskursu o predatorze jest taka, która mówi coś o predatorze, nie wynika istnienie predatora. Żeby stwierdzić istnienie predatora, muszę rzeczywiście znaleźć go w krzakach (czy gdziekolwiek indziej) lub przynajmniej empirycznie odkryć jakieś (przyczynowe) ślady jego obecności. Postulowane zaś przez Lewisa czasoprzestrzennie i przyczynowo odcięte światy możliwe z definicji nie poddają się obserwacji czy wnioskowaniu przyczynowemu.

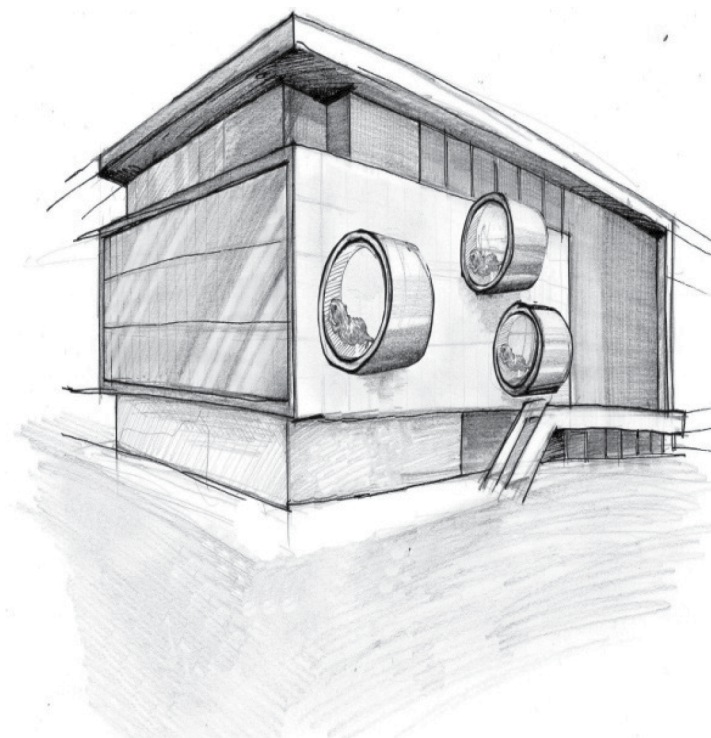
43. Jeśli nie realizm modalny, to co?

Podejście czysto mitologiczne, które postrzega dyskurs o światach możliwych jako pożyteczny mit, nie działa. Podejście realistyczne, które bierze ten dyskurs dosłownie aż do skrajności, nie jest wybitnie przekonujące. Jakie opcje pozostają?

Pamiętamy, że standardowy kiedyś podział podejść do obiektów matematycznych wyróżniał nominalizm, platonizm i konceptualizm? No właśnie, może trafną teorią będzie koncep-

Jeśli nie realizm modalny, to co?

tualizm odnośnie do światów możliwych? Może światy możliwe są tak samo obiektami mentalnymi, jak urojone dwumetrowe chomiki biegające w gigantycznych kołach generujących prąd dla biblioteki Uniwersytetu Gdańskiego?



Problemy jednak z mentalizmem w kwestii światów możliwych pojawiają się bardzo szybko. Po pierwsze, nikt chyba nie ma w głowie kompletnego wyobrażenia świata, które rozstrzyga wszystkie zdania. Jeżeli wyobrażamy sobie świat możliwy, w którym zagryzam predatora, nasze wyobrażenie jest bardzo częściowe i nie mamy zielonego pojęcia, czy w tym świecie możliwym zagryzłem również Dartha Vadera, czy nie.

Po drugie, możliwości jest całe stado. A nawet dwa. Choćby dlatego, że dla każdej liczby rzeczywistej istnieje świat możliwy, w którym wszyscy ludzie nazywają tę liczbę rzeczywistą

„humwumumwa”. Liczba rzeczywiście zaś istniejących stanów mentalnych jest skończona albo co najwyżej przeliczalna (tzn. jest ich co najwyżej tyle, ile liczb naturalnych). To znaczy, że zabraknie rzeczywistych stanów mentalnych, by modelować wszystkie potrzebne nam światy możliwe.

Być może zamiast o rzeczywiście istniejących obiektach mentalnych powinniśmy mówić również o dających się pomyśleć światach możliwych? Bo przecież, chociaż nie każdy świat możliwy jest rzeczywiście pomyślany, to każdy da się pomyśleć (chyba). Może i tak, ale jeśli wybieramy tę opcję, to wbudowujemy sobie pojęcie modalne do naszej teorii: mówimy o **dających się pomyśleć światach**, czyli o **możliwych stanach mentalnych**. A skoro tak, to już zakładamy, że możemy posługiwać się zrozumiałe modalnościami. W szczególności w takim przypadku posługujemy się możliwością, by wyjaśnić, na czym polega możliwość. A to jest już działalność lekko podejrzana intelektualnie. Cóż pozostaje?

Pierwsza opcja to śladami Carnapa, Roberta Adamsa czy Alvina Plantingi identyfikować światy możliwe z pewnymi obecnie istniejącymi obiektami: z (po kolei) maksymalnymi i niesprzecznymi (i) zbiorami zdań, (ii) sędami w sensie logicznym⁵⁹, (iii) stanami rzeczy, lub (iv) własnościami. Wtedy wpłątujemy się w wyjaśnianie pojęcia niesprzeczności bez odwoływania się do pojęć modalnych, ale jak już zauważyliśmy, może sprawa nie jest aż tak beznadziejna, jak chciałby tego Lewis.

Zaleta tego podejścia jest taka, że nie postulujemy istnienia niczego poza tym, czego istnienie i tak już uznajemy w naszym świecie (przy założeniu, że wierzymy w istnienie sądów czy własności).

Pewną trudnością tego podejścia jest **PARADOKS MAKSYMALNOŚCI**, ale ze względu na jego wyrafinowanie techniczne omówię go osobno w rozdziale 44.

⁵⁹ Przypomnijmy, osoby, które wierzą w sady logiczne, identyfikują je ze znaczeniem zdań stwierdzających.

Jeśli nie realizm modalny, to co?

Druga opcja to postąpić tak jak Robert Stalnaker i uznać, że w naszym świecie istnieją takie obiekty abstrakcyjne *sui generis*, jak **sposoby, na jakie mógłby mieć się świat** (*ways the world could be*), i to one właśnie są światami możliwymi. Pogląd ten jest zbliżony do poglądu Plantingi. Zarówno pierwszą, jak i drugą opcję nazywa się **AKTUALIZMEM**, bo starają się one identyfikować światy możliwe z aktualnie istniejącymi przedmiotami.

Podejście Plantingi czy Stalnakera różni się zdecydowanie od podejścia Carnapa i Adamsa. O ile w podejściu Carnapa i Adamsa uznaje się pojęcie zdania lub sądu w sensie logicznym za zrozumiałe i definiuje się świat możliwy jako zbiór takich znanych obiektów, o tyle u Plantingi czy Stalnakera pierwotnie zrozumiałe ma być pojęcie maksymalnej niesprzecznej własności czy sposobu, na jaki mógł mieć się świat, a sąd w sensie logicznym definiuje się jako zbiór światów możliwych (mianowicie tych światów, w których jest prawdziwy).

Definiowanie sądu jako zbioru światów możliwych natrafia na tę trudność, że sądy koniecznie równoważne są tymi samymi zbiorami światów możliwych, a więc teoria ta nie pozwala na rozróżnienie sądu wyrażanego przez zdanie „ $2 + 2 = 4$ ” od sądu wyrażanego przez zdanie „Nie istnieją nieprzestrzenne obiekty kolorowe”, a to już jest poważny problem.

Trzecia opcja, pochodząca od Wittgensteina i Davida Armstronga, to kombinatoryczna teoria światów możliwych, wedle której należy zacząć od podstawowych metafizycznych składników naszej rzeczywistości, jakiegoś rodzaju „atomów”, i postrzegać światy możliwe jako kombinacje tychże.

Zaletą tego podejścia jest, że nie postuluje się istnienia jakichś nowych obiektów, jeżeli mówienie o potencjalnych kombinacjach uznać za niewinne ontologicznie.

Jednym z głównych problemów jest to, że tego typu teoria nie pozwala na możliwości, które intuicyjnie rzecz biorąc, istnieją.

- Jest na przykład możliwe, żeby świat składał się z innego budulca metafizycznego, niż się składa. Nie ma jednak takiej

- rekombinacji rzeczywistego budulca, przy której byłoby prawdą, że świat składa się z innego „budulca metafizycznego”.
- Jest również możliwe, żeby budulca metafizycznego naszego świata, czymkolwiek on jest, było więcej. Żadna jednak rekombinacja budulca nie sprawi, żeby było go więcej.
 - Za możliwe też uznajemy istnienie obiektów czysto duchowych. Jeżeli jednak nasz świat rzeczywisty takowych nie zawiera, nie będzie istnieć rekombinacja, w której takie obiekty zaistnieją.

Drugim problemem tego podejścia jest sama problematyczność pojęcia „ostatecznego budulca wszechświata”. Nie wiemy, czym jest, nie wiemy, czy w ogóle jest. Dlaczego mielibyśmy więc konstruować teorie semantyczne, zakładając tego typu silne tezy metafizyczne?

Ponadto teoria kombinatoryczna w jakiś sposób kwantyfikuje po niezrealizowanych kombinacjach – czy można o czymś takim mówić, nie zakładając już possybilizmu, poglądu, zgodnie z którym już istnieją obiekty czysto możliwe?

Czwarta opcja to **FIKcjONALIZM MODALNY** reprezentowany przez Gideona Rosena, sugerujący, że warunki prawdziwości zdań typu „jest możliwe, że p ” sprowadzają się do czegoś w stylu:

Wedle fikcji modalnego realizmu istnieje świat możliwy, w którym p jest prawdziwe.

W ten sposób, powiada Rosen, zdania modalne dostają te same wartości logiczne, które dostałyby, gdyby realizm modalny był prawdziwy, ale nie płaci się za to uznawaniem na serio modalnego realizmu i natrafianiem na jego trudności.

Wraca tu jednak do nas problem podejścia mitologicznego. Podejmując rozmaite działania, kieruję się między innymi zdaniami modalnymi. Dlaczego miałbym kierować się tymi,

Paradoksy maksymalności

które okazują się prawdziwe w tej, a nie innej fikcji? Dlaczego miałyby to być racjonalne?

Ostatnią opcją jest **ELIMINATYWIZM MODALNY**: pogląd, wedle którego dyskurs, który wydaje się być o światach możliwych, wcale o światach możliwych nie jest. Zwolennik tego typu podejścia stoi przed dosyć poważnym wyzwaniem: musi dostarczyć takiego tłumaczenia zdań o światach możliwych na akceptowalny język nie o światach możliwych, które jest w miarę intuicyjne, owocuje trafną oceną prawdziwości zdań dyskursu o światach możliwych oraz nie daje kontrowersyjnych rezultatów, jeżeli chodzi o nasze intuicyjne oceny zdań modalnych. Jednym z reprezentantów tej opcji jest Chihara. Zaletą tego podejścia jest skromność ontologiczna. Niestety, rzetelne próby tego rodzaju stają się szybko zaawansowane technicznie i nie mamy miejsca na omówienie ich tutaj z taką dokładnością, na jaką zasługują.

44. Paradoksy maksymalności

Niektórzy uważają, że istnieją **sądy w sensie logicznym** – abstrakcyjne obiekty będące znaczeniami zdań twierdzących wygłoszonych w danym kontekście. Tacy ludzie uważają na przykład, że jeżeli patrząc przez okno, wypowiadam zdanie: „Przed budynkiem są dwie żyrafy”, a stojący obok mnie predator, który jak wiadomo, mówi po angielsku, twierdzi: “There are two giraffes in front of the building”, to mimo że wygłaszamy różne zdania, wyrażamy ten sam sąd w sensie logicznym. Chociaż istnienie takich obiektów (jak sądy, nie jak predator czy żyrafy) jest dyskusyjne, załóżmy roboczo, że rzeczywiście istnieją.

Niektóre z tych sądów są prawdziwe, a inne fałszywe. Naturalną intuicją (przynajmniej o tyle, o ile czyjeś intuicje zostały już wypaczone przez kontakt z odrobiną matematyki) jest zatem uznanie, że istnieje również zbiór sądów prawdziwych. Problem polega jednak na tym, że można podać dość silny argument przeciwko temu twierdzeniu.

Chcemy dowieść, że nie istnieje zbiór sądów prawdziwych (oznaczymy go PR). Argument przebiega przez doprowadzenie negacji hipotezy do sprzeczności⁶⁰.

1. Istnieje zbiór sądów prawdziwych, nazwijmy go PR.
założenie (nie wprost)
2. Istnieje zbiór potęgowy zbioru PR, tj. $\wp(\text{PR})$.
teoria mnogości, 1
3. Dla każdego zbioru sądów $B \in \wp(\text{PR})$ istnieje sąd mówiący „ B jest zbiorem sądów”; oznaczmy go s_B .
założenie
4. Jeżeli zarówno B , jak i C należą do $\wp(\text{PR})$ i $B \neq C$, to $s_B \neq s_C$.
założenie
5. Dla każdego $B \in \wp(\text{PR})$, s_B jest prawdziwym sądem.
definicja, 3
6. Istnieje zbiór wszystkich sądów typu s_B . Oznaczmy go $\mathcal{S}$.
teoria mnogości, 5, 1
7. $\mathcal{S} \subseteq \text{PR}$.
definicja, 6, 1

⁶⁰ W tym miejscu przydatne może być zwrócenie do rozdziału 15 o twierdzeniu Cantora: to powinno ułatwić zrozumienie pojęć występujących w rozważanych argumentach.

Paradoksy maksymalności

8. S nie ma więcej elementów niż PR .

teoria mnogości, 7

9. S ma tyle samo elementów co $\mathcal{P}(PR)$.

teoria mnogości, 3, 4

10. $\mathcal{P}(PR)$ ma więcej elementów niż PR .

twierdzenie Cantora

11. S ma więcej elementów niż PR .

teoria mnogości, 9, 10

sprzeczność: 8, 11.

Przesłanka 1 to właśnie założenie, które chcemy obalić, wyprowadzając sprzeczność.

Krok 2 wynika z kroku 1, bo ogólnie w teorii mnogości dla każdego zbioru istnieje zbiór potęgowy tego zbioru⁶¹.

Krok 3 jest raczej intuicyjny. Dla dowolnego zbioru cze-
gokolwiek można powiedzieć, że ten zbiór jest zbiorem tych
rzeczy, których jest zbiorem. W szczególności o dowolnym
zbiorze wiertarek można powiedzieć, że jest zbiorem wiertar-
rek, a o dowolnym zbiorze sądów można powiedzieć, że jest
zbiorem sądów.

Krok 4 obrazuje tylko to, że jeśli zbiory B i C są różne,
to cokolwiek o nich orzekniemy, dostaniemy różne sądy, bo
podmioty tych sądów będą różne. Stąd jeżeli B i C są różnymi
zbiorami sądów, to i sąd „ B jest zbiorem sądów” różni się od
sądu „ C jest zbiorem sądów”.

Krok 5 podkreśla tylko to, co powiedzieliśmy o kroku 3.
Sąd „ B jest zbiorem sądów” nie tylko jest sądem, lecz także jest
sądem prawdziwym.

61 Przypomnijmy: zbiorem potęgowym danego zbioru jest zbiór wszystkich jego podzbiorów. Czytelnik potrzebujący dokładniejszego wyjaśnienia może zajrzeć do rozdziału 15.

Krok 6 wynika z teoriomnogościowego aksjomatu wyróżnienia. Wedle tego aksjomatu, jeżeli istnieje jakiś zbiór A , to istnieje też zbiór tych elementów zbioru A , które spełniają pewien warunek ϕ dla dowolnego warunku ϕ . Stąd też skoro (wedle kroku 1) istnieje zbiór sądów prawdziwych, to istnieje też zbiór sądów prawdziwych należących do S .

Krok 7 opiera się na spostrzeżeniu, że skoro sądy w zbiorze S to sądy prawdziwe, to muszą też należeć do zbioru **wszystkich** sądów prawdziwych.

Krok 8 opiera się na kroku 7 i na teoriomnogościowym fakcie, że podzbiór jakiegoś zbioru nie może mieć więcej elementów niż ten (większy) zbiór.

Krok 9 opiera się na tym, że skoro można (wedle 3 i 4) wzajemnie jednoznacznie skojarzyć elementy zbioru potęgowego $\mathcal{P}(PR)$ z sądami w zbiorze S , to znaczy, że musi ich być tyle samo.

Krok 10 jest po prostu zastosowaniem twierdzenia Cantora, wedle którego dowolny zbiór ma mniej elementów niż jego zbiór potęgowy.

Krok 11 to konsekwencja kroków 9 i 10: skoro S ma tyle samo elementów co $\mathcal{P}(PR)$, a ten z kolei ma więcej elementów niż PR , to i S musi mieć więcej elementów niż PR .

Razem jednak kroki 8 i 11 wykluczają się. To znaczy, że o ile w miarę przekonujące założenia 2, 3 i 4 są prawdziwe, to założenie 1 jest fałszywe i nie istnieje zbiór sądów prawdziwych.

Paradoks ten ma kilka wariantów. Po pierwsze, analogicznie możemy też dowieść nieistnienia zbioru sądów fałszywych oraz nieistnienia zbioru sądów w ogóle.

Po drugie, jeżeli uważamy, że istnieją stany rzeczy i że są one takie, że każdemu sądowi w sensie logicznym odpowiada osobny stan rzeczy, to nie istnieje również zbiór wszystkich stanów rzeczy⁶². To sprawia, że konstruując teorię stanów rze-

⁶² Zadanie: przeformułuj powyższy argument 1–11 tak, aby stosował się bezpośrednio do stanów rzeczy.

Paradoksy maksymalności

czy, należy starać się uniknąć tego założenia (albo uniknąć paradoksu maksymalności w jakiś inny sposób).

Po trzecie, sprawa ma się tylko trochę inaczej, gdy chcemy zastosować paradoks do światów możliwych w sensie Adamsa (por. rozdział 43). Adams identyfikuje światy możliwe z maksymalnymi i niesprzecznymi zbiorami sądów. Można jednak podać argument za tym, że nie istnieje maksymalny niespreczny zbiór sądów.

1' Istnieje jakiś maksymalny niespreczny zbiór sądów, nazwijmy go M .

założenie (nie wprost)

2' Istnieje zbiór potęgowy zbioru M , tj. $\wp(M)$.

teoria mnogości, 1'

3' Dla każdego zbioru sądów $B \in \wp(M)$ istnieje sąd mówiący „ B jest zbiorem sądów”; oznaczmy go μ_B . Zbiór wszystkich takich sądów oznaczmy jako μ .

założenie

4' Jeżeli zarówno B , jak i C należą do $\wp(M)$ i $B \neq C$, to $\mu_B \neq \mu_C$.

założenie

5' Zbiór μ ma tyle samo elementów co $\wp(M)$.

teoria mnogości, 3', 4'

6' Dla każdego zdania typu μ_B , albo μ_B , albo $\neg\mu_B$ należy do M .

maksymalność zbioru, 1'

7' M ma przynajmniej tyle samo elementów co μ .

teoria mnogości, 6'

8' M ma przynajmniej tyle samo elementów co $\wp(M)$.

teoria mnogości, 5', 7'

9' M ma mniej elementów niż $\delta^P(M)$.

twierdzenie Cantora

sprzeczność 8', 9'

Po czwarte, argument tak naprawdę nie wymaga identyfikacji światów możliwych z maksymalnymi niesprzecznymi zbiorami sądów. Wystarczy bowiem tylko uznać, że niezależnie od tego, czym są światy możliwe, każdemu światu możliwemu odpowiada jego zupełny opis, czyli pewien niespreczny i maksymalny zbiór sądów.

Jakie są możliwości wybrnięcia z tego paradoksu? Argument traktuje sądy o zbiorach sądów na równi z wszystkimi sądami, takimi jak te, które już znajdują się w tych zbiorach. Sztuczka polega przecież na tym, że tak wygenerowane sądy (lub ich negacje) wrzucamy z powrotem do zbioru maksymalnego, o którym mówimy. A to prowadzi na przykład do tego, że μ jako podzbiór M też ma odpowiadający sobie sąd μ_μ , który mówi, że μ jest zbiorem sądów. Z definicji wiemy, że μ powinien zawierać również μ_μ , a zatem μ_μ mówi coś o zbiorze, do którego sam należy.

Jeżeli komuś nie odpowiadają tego typu zjawiska, może próbować zrzucić wszystko na problemy z samoodniesieniem i brakiem hierarchizacji sądów. Rozróżnienie poziomów sądów (sądy przedmiotowe, sądy o sądach przedmiotowych, sądy o sądach o sądach przedmiotowych itp.), ostrożnie przeprowadzone, mogłoby pomóc. Wystarczyłoby powiedzieć na przykład, że świat możliwy to maksymalny niespreczny zbiór sądów przedmiotowych (sądów nie o sądach). W takim przypadku uniknie się paradoksu, bo ruch 6' będzie nieuprawniony.

Tak jak w przypadku innych ruchów polegających na zakazie samoodniesienia, pozostaje pytanie, czy nie wylewamy małego predatorka z kąpielą w kwasie.

Paradoksy maksymalności



Wszelkie bowiem ograniczenia tego typu niosą ze sobą dość nieintuicyjne wykluczenie sensowności lub poprawności szeregu sądów.

Inna strategia obrony to uznanie wniosku. Może rzeczywiście należy odrzucić istnienie światów możliwych jako zbiorów i uznać, że są klasami w sensie teoriomnogościowym?

Jeden problem z tym podejściem jest taki, że klasy zazwyczaj nie są zbiorami, bo zawierają za dużo zbiorów – nie ma tutaj podobieństwa do maksymalnych niesprzecznych zbiorów sądów, które w ogóle nie zawierają zbiorów. Zarzut nie jest przekonujący, bo wymaga za dużego podobieństwa; czy nie wystarczy samo to, że w klasie tych sądów hierarchia, który sąd jest o czym, jest wystarczająco poplątana, tak samo jak poplątana byłaby relacja należenia do zbioru w zbiorze wszystkich zbiorów?

Poważniejszym problemem jest to, że semantyka światów możliwych posługuje się **ZBIORAMI** światów możliwych. Jeżeli jednak światy możliwe są klasami, to nie mogą należeć do

żadnego zbioru, a zatem nie mogą być traktowane jako elementy zbioru światów możliwych. Wydaje się więc, że ta opcja znacznie ograniczałaby użyteczność teoretyczną semantyki światów możliwych.

Inna opcja to odrzucić istnienie sądów w sensie logicznym i zamiast tego mówić o zdaniach: napisach w konkretnym języku. W takim przypadku zawiedzie krok 3'. Zdania bowiem są skończonymi ciągami znaków zbudowanych z przeliczalnego alfabetu. Musi ich więc być przeliczalnie wiele⁶³. Natomiast podzbiorów zbioru maksymalnego (który musi być nieskończony) będzie nieprzeliczalnie wiele (czyli więcej). Od razu jest więc jasne, że nie każdy podzbiór zbioru maksymalnego ma odpowiadające mu zdanie.

Kosztem tego ruchu jest to, że już dłużej nie możemy definiować światów możliwych jako zbiorów sądów. Raczej być może musielibyśmy definiować je jako zbiory zdań. To jednak rodzi pewne problemy. Zdań z krwi i kości (czy papieru i tuszu) jest tylko pewna skończona liczba, bo tylko skończona liczba została napisana. To zdecydowanie za mało. Żeby modelować wszystkie możliwości, musimy też mówić o takich zdaniach, które **można** napisać. To jednak sprawia wrażenie, że nasza analiza modalności będzie kolistą. Ocena, czy tak jest rzeczywiście, jest bardziej skomplikowana.

Każda z rozważanych opcji zasługuje na dłuższą dyskusję i niech Moc będzie z każdym, kto chce się w te dyskusje dalej zagłębiać. Bo my nie będziemy. Ważne natomiast jest jeszcze następujące spostrzeżenie. Paradoks w wersji pierwszej w ogóle nie dotyczył światów możliwych, tylko sądów. Źródłem więc paradoksalności są nasze intuicje odnośnie do sądów jako takich, a nie światów możliwych, i to z konstrukcją działającej teorii sądów filozof musi sobie najpierw poradzić, jeżeli chce definiować światy możliwe jako zbiory sądów.

⁶³ Por. uwagi dotyczące pojęcia przeliczalności na s. 61.

Paradoksy maksymalności

TEKSTY POLECANE



1. Świetnym wprowadzeniem do problematyki okresów warunkowych oraz logik nieklasycznych jest *Introduction to NonClassical Logic: From If to Is* Grahama Priesta.
2. Opis systemu Lewisa znajduje się w jego *Survey of Symbolic Logic* z roku 1918. Kolejnym ważnym dziełem Lewisa i C. Langforda jest *Symbolic Logic* z roku 1932.
3. Semantyka światów możliwych pojawiła się w pracach Kripkego (takich jak *A Completeness Theorem in Modal Logic*, *Journal of Symbolic Logic*, 24: 1–14, 1959) i Hintikki (*Knowledge and Belief*, 1962).
4. Świetną książką dotyczącą filozofii światów możliwych jest *Possible Worlds* Johna Diversa. Rozdział 15 dotyczy paradoksów maksymalności.
5. Inną fascynującą książką filozoficzną o światach możliwych, pisaną raczej z perspektywy antyrealistycznej, jest *The Worlds of Possibility* Charlesa Chihary.
6. Klasycznymi już tekstami Davida Lewisa o światach możliwych są jego dwie książki *Counterfactuals* i *On the Plurality of Worlds* oraz artykuł *Possible Worlds* z *Nous* 1976, 10: 65–75.
7. Historycznie najważniejszą pracę o logice modalnej Carnapa jest jego *Meaning and Necessity* z roku 1947.
8. Klasycznym tekstem prezentującym konstrukcję światów możliwych jako zbiorów sądów jest *Theories of Actuality* Roberta Adamsa (*Nous* 1974, 8: 211–231).
9. Najważniejszą pracę Alvina Plantingi jest książka *The Nature of Necessity*, 1976.
10. Kombinatoryczna teoria światów możliwych ma motywację w *Traktacie* Wittgensteina, a w pełni jest sformułowana w książce Davida Armstronga *A Combinatorial Theory of Possibility*, 1989.
11. Fikcjonalizm modalny omówiony jest w *Modal Fictionalism*, *Mind* 1990, 99: 327–354.
12. Poglądy Chihary na światy możliwe znaleźć można w jego książce *Worlds of Possibility*, 1998.
13. Bardzo dobrymi zbiorami artykułów dotyczących filozofii światów możliwych są *The Possible and the Actual. Readings in the Metaphysics of Modality*, 1979, opracowana przez M. J. Loux, oraz *Analytical Metaphy-*

sics. A Collection of Essays. Necessity and Possibility, 1999, pod red. M. Tooleya.

14. Po polsku warto zajrzeć do *Metafizyki w filozofii analitycznej* po red. T. Szubki, Lublin 1996.



POJĘCIE OPERATORA INTENSJONALNEGO

1. Ekstensją nazwy jest przedmiot, ekstensją predykatu zbiór przedmiotów, do których się stosuje, a ekstensją zdania jego wartość logiczna (przyjmijmy roboczo; niektórzy za ekstensję zdań uważają stany rzeczy).
2. Operatory ekstensjonalne tworzą zdania za pomocą innych zdań.
3. Ekstensja (= wartość logiczna) zdań zbudowanych z operatorów ekstensjonalnych jest zdeterminowana przez ekstensję (= wartość logiczną) zdań składowych. Przykłady: koniunkcja, negacja.
4. Operatory intensjonalne to te, które nie są ekstensjonalne.
5. Wartość wyrażenia składowego w takim przypadku nie jest zdeterminowana przez wartość zdania składowego.
6. Przykład: „Jest konieczne, że” wraz z „ $2 + 2 = 4$ ” i „nie ma mnie w biurze”. Oba zdania składowe są prawdziwe, ale odpowiednie zdania złożone różnią się wartością.
7. Spośród wielu modalności możliwość logiczna i analityczna wydają się bardziej podstawowe, bo inne można do nich zredukować.



PARADOKSY IMPLIKACJI A MODALNOŚCI

1. Niektórzy próbują modelować okresy warunkowe za pomocą implikacji materialnej, która jest fałszywa wtedy i tylko wtedy, gdy poprzednik jest prawdziwy, a następnik fałszywy.
2. To owocuje paradoksami implikacji materialnej: rozumowaniami, które zgodnie z logiką klasyczną powinny być poprawne, a które intuicyjnie są absurdalne i niepoprawne.
3. C.I. Lewis w odpowiedzi skonstruował systemy implikacji ścisłej (tj. koniecznej implikacji materialnej), przy okazji wprowadzając operatory modalne.
4. Podejście Lewisa było czysto aksjomatyczne – trudności: porównywanie systemów aksjomatycznych czysto na poziomie składni bardzo trudne, wielość aksjomatyzacji, niejasne kryteria wyboru systemu.

Paradoksy maksymalności

5. Pojawiają się też paradoksy implikacji ścisłej; nieco słabsze od paradoksów implikacji materialnej, ale również problematyczne (np. zdanie konieczne jest następnikiem prawdziwej implikacji ścisłej, niezależnie od poprzednika).

ŚWIATY MOŻLIWE: APARATURA

1. Podstawowe elementy: światy możliwe i relacja dostępności (względnej możliwości) między nimi.
2. Światy możliwe: maksymalne i niesprzeczne (tj. każde zdanie prawdziwe lub fałszywe, ale nie prawdziwe i fałszywe zarazem).
3. Relacja dostępności: zależy od typu modalności, o której mowa, np. niedostępne dla modalności fizycznej są te światy, które mają inną fizykę.
4. System modalności logicznej **S5**: relacja dostępności jest zbędna.
5. W **S5** zdanie jest możliwe wtedy i tylko wtedy, gdy jest prawdziwe w jakimś świecie. Jest konieczne wtedy i tylko wtedy, gdy jest prawdziwe w każdym świecie.
6. Przy takiej semantyce twierdzenia modalne dają się uzasadnić jako wypływające z tego, jak zachowują się kwantyfikatory i wartościowania.
7. Semantyka taka pozwala też na systematyczną klasyfikację i łatwiejszy wybór systemów logicznych. Pozwala też na dowodzenie pełności systemu aksjomatycznego (tj. że system dowodzi wszystkiego, co powinien).
8. Pojęcie światów możliwych przydatne też w innych kwestiach filozoficznych: definicja sądu w sensie logicznym, własności, redukcji ontologicznej, superweniencji; istotna rola w teoriach okresów warunkowych; analiza niemodalnych operatorów intensjonalnych (jak „wiadomo, że” czy „obowiązkowo”).
9. Teoria światów możliwych, która przypisuje im status czysto fikcyjny, nie wyjaśnia, dlaczego teoria ta miałaby być pomocna w rzeczywistości.



REALIZM MODALNY D. LEWISA

1. Teza: światy możliwe rzeczywiście istnieją tak jak nasz, nie różnią się od niego rodzajem, ale są czasoprzestrzennie i przyczynowo odizolowane.
2. Argument: kwantyfikacja w języku potocznym może być brana na serio, jeżeli (1) nie prowadzi to do problemów i (2) nie ma lepszej teorii, któ-



ra nie prowadzi do problemów. Warunki te są spełnione w przypadku dyskursu o światach możliwych.

3. Odnośnie do (1): nie można po prostu zaprzeczać wnioskowi i mówić, że to jest problem; nie można zarzucać rozrzutności ontologicznej, bo jest ona tylko ilościowa – postuluje więcej obiektów znanego już rodzaju, a nie nowy typ przedmiotów; nie można zarzucać niejasnych kryteriów identyczności, bo każde indywiduum żyje tylko u siebie w świecie, a każdy świat indywiduowany jest izolacją czasoprzestrzenną.
4. Odnośnie do (2): uznanie pojęć modalnych za pierwotne to nie teoria; definiowanie niesprzeczności jest koliste (bo posługuje się modalnościami) albo niepoprawne (bo nie radzi sobie z niezupełnością arytmetyki); definiowanie światów możliwych jako niesprzecznych maksymalnych zbiorów sądów zakłada pojęcie niesprzeczności i dziedziczy powyższe problemy.



PROBLEMY REALIZMU MODALNEGO

1. Lewis twierdzi, że realizm jest rozrzutny tylko ilościowo. Ale światy możliwe z różnymi dziwnymi bytami (np. duchami) wskazują na rozrzutność jakościową.
2. Realizm nie daje przekonującej teorii **znaczenia** operatorów modalnych.
3. Pojęcie izolacji czasoprzestrzennej jest dyskusyjne. Jaka jest jej natura? Jak możemy coś wiedzieć o światach odizolowanych? Dlaczego właśnie izolacja ma stanowić zasadę jednostkowania światów możliwych?
4. Lewis zmuszony jest wprowadzić teorię odpowiedników (bo każde indywiduum żyje tylko w swoim świecie). Odpowiedniki jednak mogą nie być jedyne, a poza tym nie jest jasne, dlaczego stan mojego odpowiednika miałby mieć wpływ na mój status modalny.
5. Zarzut Lewisa wobec logicznej koncepcji możliwości (że nie obejmuje ona prawd analitycznych) nie jest trafny. Można przecież posłużyć się pojęciem prawdziwości ze względu na znaczenie.
6. Zarzut Lewisa wobec pojęcia niesprzeczności nie jest przekonujący: teoriomodelowa definicja nie posługuje się pojęciami modalnymi, a jedynie istnieniem obiektów matematycznych.
7. Argument Lewisa za światami jest pośpieszny: z tego, że najlepiej radząca sobie teoria semantyczna jakiegos dyskursu postuluje coś dziwnego,

Paradoksy maksymalności

nie wynika istnienie czegoś dziwnego. Równie dobrze może wskazywać na to, że sam dyskurs jest podejrzany.

ALTERNATYWY WZGLĘDEM REALIZMU MODALNEGO

1. Konceptualizm nie zapewnia wystarczającej ilości stanów mentalnych; chyba że przechodzi w konceptualizm modalny, ale wtedy ma już wbudowaną możliwość.
2. Carnap, Adams: światy możliwe to maksymalne zbiory niesprzeczne zdań lub sądów. Wyzwanie: wyjaśnienie niesprzeczności bez pojęć modalnych. Zaleta: skromność ontologiczna.
3. Plantinga, Stalnaker: maksymalne niesprzeczne własności albo rzeczywiste obiekty abstrakcyjne *sui generis*. Definiowanie sądów jako zbiorów światów możliwych. Problem: identyfikacja sądów koniecznie równoważnych.
4. Wittgenstein, Armstrong: teoria kombinatoryczna. Światy możliwe to kombinacje istniejących „atomów metafizycznych”. Problem: nie otrzymujemy szeregu intuicyjnych możliwości. Możliwości innego budulca, możliwości, że budulca jest więcej, czy możliwości obiektów czysto duchowych. Pojęcie budulca metafizycznego nie jest dosyć jasne.
5. Rosen: fikcjonalizm modalny. Jest możliwe, że **p** wtedy i tylko wtedy, gdy wedle fikcji realizmu istnieje świat możliwy, gdzie **p** jest prawdą. Wracają problemy podejścia mitologicznego.
6. Eliminatywizm modalny: tłumaczenie na język niepostulujący światów możliwych. Sprawy zazwyczaj bardzo techniczne, kierunek raczej nowy i trudno powiedzieć coś definitywnego.



PARADOKSY MAKSYMALNOŚCI

1. Argument pierwszy jest za tezą, że nie istnieje zbiór wszystkich sądów prawdziwych.
2. Przebiega następująco: gdyby taki zbiór istniał, to każdemu jego podzbiorowi moglibyśmy przyporządkować nowy sąd. Zatem sądów musiałoby być tyle, ile podzbiorów zbioru wszystkich sądów. To jednak jest wykluczone przez twierdzenie Cantora.



3. Istnieją analogiczne argumenty za nieistnieniem zbioru wszystkich sądów fałszywych, zbioru wszystkich sądów w ogóle oraz zbioru wszystkich stanów rzeczy.
4. Argument drugi jest za tezą, że nie istnieją maksymalne niesprzeczne zbiory sądów.
5. Przebiega następująco: gdyby istniał, to każdemu jego podzbiorowi moglibyśmy przyporządkować nowy sąd. Albo ten sąd, albo jego negacja musi jednak należeć do wyjściowego zbioru maksymalnego. Zatem zbiór maksymalny musi być co najmniej tak duży, jak własny zbiór potęgowy.
6. Argument działa przeciwko każdej teorii, która przypisuje każdemu światu możliwemu jego osobny maksymalny zbiór sądów w nim prawdziwych.
7. Pewne sposoby wybrnięcia: hierarchizacja sądów, uznanie, że to nie są zbiory, tylko klasy (w sensie teoriomnogościowym), odrzucenie sądów w sensie logicznym. Każda ma swoje słabości: wykluczenie sądów intuicyjnie sensownych, osłabienie teoretycznej przydatności światów możliwych, niemożliwość definiowania światów możliwych jako zbiorów sądów.

45. Zamiast zakończenia

Zagadnienia poruszone w tej książce stanowią wynik mojego wyboru kilku kwestii, których świadomość ułatwić ma czytelnikowi dalsze samodzielne poruszanie się w filozofii analitycznej.

Częścią kompetencji, którą zwykle zakłada się u czytelnika tekstów z filozofii analitycznej, jest chociaż pobieżna znajomość myśli pierwszych filozofów analitycznych. Dlatego kilka pierwszych rozdziałów dotyczyło właśnie tych myślicieli.

Część kompetencji wymaganych wspólnie w tekstach z filozofii analitycznej pisanych na różne tematy dotyczyć będzie kwestii metodologicznych – i dlatego dość silny akcent kładłem na kwestie metodologiczne.

Ponadto pozwoliłem sobie wybrać kilka bardziej interesujących mnie zagadnień i przedstawić krótkie do nich wprowadzenie.

Pozostaje wiele kwestii, których nie poruszyłem, a które na poruszenie w tego typu książce na pewno zasługują. Zdecydowałem jednak pracę nad książką zakończyć i ewentualnie wrócić do jej rozbudowy, jeżeli tekst okaże się w jakikolwiek sposób przydatny (nie liczę traktowania książki jako podkładki pod piwo albo pod myszkę komputerową). Jeżeli czytelnik uzna, że istnieją kwestie, które zasługują na rozbudowę albo poruszenie, bardzo chętnie zapoznam się z jego opinią i postaram się ją uwzględnić w kolejnych wersjach tej książki.



Dr hab. prof. UG Rafał Urbaniak obronił doktorat na Uniwersytecie w Calgary, bo dalej się nie dało. Od tego czasu dręczy studentów na Uniwersytecie Gdańskim oraz pisze nudne teksty o zagadnieniach logicznych, które mało kogo obchodzą, współpracując z Centre for Logic and Philosophy of Science w Gandawie. W międzyczasie, z powodów nie do końca zrozumiałych, mieszkał przez dłuższy czas w Indiach, gdzie powstała pierwsza połowa niniejszej książki. Druga została napisana, gdy po powrocie próbował przyzwyczaić się do nowej starej rzeczywistości i poległ. Książka ta jest wynikiem próby jasnego przedstawienia tego, czego autor żałował, że sam nie wiedział na pierwszym roku studiów filozoficznych.

[...] nie sposób nie zauważyć, że jest to prawdopodobnie jedyny tekst filozoficzny, w którym słowo „predator” pojawia się częściej niż słowo „byt”, „prawda” lub „wiedza”.

Dr Maciej Sendłak, Uniwersytet Szczeciński

Wydawnictwo



Academicon

