EDUKACJA FILOZOFICZNA
Vol. 7 1989

JAN ZYGMUNT

WOKÓŁ *PRAWDY I DOWODU* ALFREDA TARSKIEGO. PROPOZYCJA DYDAKTYCZNA DOTYCZĄCA TWIERDZENIA GÖDLA O NIEZUPEŁNOŚCI I TWIERDZENIA TARSKIEGO O NIEDEFINIOWALNOŚCI

I. WSTĘP

Niniejsza praca, co sugeruje już sam jej tytuł, nawiązuje ściśle do artykułu A. Tarskiego *Prawda i dowód*, opublikowanego w polskiej wersji językowej w „Studiach Filozoficznych" [1].

Artykuł Tarskiego jest dość popularnym i nietechnicznym, ale za to z pierwszej ręki, zarysem Tarkiego teorii prawdy i jej konsekwencji metalogicznych, podanej w zestawieniu z epokowymi odkryciami Kurta Gödla z 1931 roku [2].

Wiadomo, że oryginalny wykład teorii prawdy zawarł Tarski w monografii *Pojęcie prawdy w językach nauk dedukcyjnych* [3], zaś rozprawa *The semantic conception of truth and the foundation of semantics* [4] poświęcona jest nieformalnej prezentacji wyników z *Pojęcia prawdy*..., zawiera też rezultaty nowe, zwłaszcza o charakterze filozoficznym i metodologicznym, jak również dużą i bogatą część polemiczną.

Niniejsza praca adresowana jest przede wszystkim do studenta filozofii, który pragnie zapoznać się z rudymentami teorii Tarskiego i Gödla i ich najprostszymi konsekwencjami matematycznymi.

Propozycja nie dotyczy wymienionych przed chwilą publikacji oryginalnych, które dla początkującego studenta mogą być za trudne (za-

[1] A. Tarski: *Prawda i dowód*. „Studia Filozoficzne" 1984, nr 2, s. 9—30. Tekst oryginalny: Idem: *Truth and proof*. „Scientific American" 1969, nr 220(6), s. 63—67.

[2] K. Gödel: *Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme. I.* „Monatshefte für Mathematik und Physik" 1931, nr 38, s. 173—198. Przekłady angielskie m.in. w: *From Frege to Gödel* (red. J. van Heijenoort). Cambridge 1967 oraz K. Gödel: *Collected Works*. Vol. 1 (red. S. Feferman et al.), Oxford 1986.

[3] A. Tarski: *Pojęcie prawdy w językach nauk dedukcyjnych*. „Prace Towarzystwa Naukowego Warszawskiego, Wydział III". Warszawa 1933, nr 34.

[4] Idem: *The semantic conception of truth and the foundation of semantics*. „Philosophy and Phenomenological Research" 1944, nr 4, s. 341—376.

równo ze względów rachunkowych jak i językowych), ale właśnie artykułu *Prawda i dowód*, który jest nieformalny, świetnie napisany i łatwo dostępny. Proponuje się studentowi, aby lekturę *Prawdy i dowodu* uzupełnił niniejszym tekstem, a dopiero później sięgnął do publikacji bardziej zaawansowanych.

Z semantyczną teorią prawdy Tarskiego student filozofii spotyka się już na pierwszym roku studiów, bądź w początkach drugiego, kiedy to definiuje się (przez spełnianie) prawdę w dziedzinie odniesioną do zdań języka pierwszego rzędu, a więc — do języka sformalizowanego. Już wtedy można poruszyć zagadnienie prawdy w odniesieniu do zdań języka naturalnego oraz wzmiankować o twierdzeniach limitacyjnych dla formalizmów pierwszego rzędu, w szczególności — o twierdzeniach Gödla o niepełności arytmetyki formalnej oraz o twierdzeniu Tarskiego o niedefiniowalności prawdy arytmetycznej (tzn. o nieistnieniu predykatu arytmetycznego w języku I rzędu, którego ekstensja byłaby identyczna ze zbiorem numerów prawdziwych zdań arytmetycznych).

Semantyczną teorię prawdy prezentować też można na III roku filozofii w ramach przedmiotu *Metodologia nauk*, szczególnie wtedy, gdy przynajmniej część owego przedmiotu wykładana jest jako formalna metodologia nauk, w ścisłym nawiązaniu do polskiej tradycji zastosowań logiki w metodologii, reprezentowanej pracami Ajdukiewicza, Kokoszyńskiej-Lutmanowej, Suszki, Przełęckiego i Wójcickiego.

Kolejne punkty niniejszej pracy są szczegółowymi nawiązaniami do niektórych wątków *Prawdy i dowodu*.

Punkt II zawiera zestaw ćwiczeń na temat rozróżnień: język vs. metajęzyk, użycie vs. wymienienie wyrażenia. Są to na ogół ćwiczenia elementarne, chociaż niektóre metazagadki nie mają natychmiastowych rozwiązań.

W punkcie III zamieszczono kilka szczegółowych uwag o wyrażeniach samostosowalnych, głównie w nawiązaniu do tej wersji paradoksu kłamcy, którą znajdujemy w *Prawdzie i dowodzie*.

W punkcie IV pokazujemy dwa sposoby konstruowania nazw i zdań samoodnośnych.

W punkcie V przedstawiamy pewien system formalny S, pochodzący od R. M. Smullyana, który pozwala na przeprowadzenie rozumowań przekątniowych, prowadzących do twierdzeń (à la Gödel i Tarski) o niezupełności i niedefiniowalności.

Niniejsze opracowanie jest prawie w całości oparte na słynnym i powszechnie uznanym artykule R. Smullyana *Languages in which self reference is possible* („Journal of Symbolic Logic" 1957, nr 22, s. 55—67) oraz na jego pięknej książce *The lady or the tiger?* (New York 1982).

II. JĘZYK VS. METAJĘZYK; UŻYCIE A WYMIENIENIE WYRAŻENIA

Tytułowe pojęcia punktu II proponujemy zilustrować pięcioma ćwiczeniami.

ĆWICZENIE 1. B. Russel w książce *Mój rozwój filozoficzny*[5] zacytował (wymienił) następujące zdanie:

To, co znaczy jest, *a zatem różni się od* jest, *ponieważ* „jest *jest" byłoby niedorzecznością.*

Przeanalizuj sens zdania Russela. Zapisz jego myśl za pomocą kilku zdań prostych. Które słowa w tym zdaniu są wymienione, a które użyte.

ĆWICZENIE 2. Lewis Carrol w *O tym, co Alicja odkryła po drugiej stronie lustra*[6] bawi czytelnika takim oto dialogiem Alicji z Rycerzem:

— *Pozwól, że zaśpiewam ci piosenkę, żeby cię pocieszyć.*

— *Czy to bardzo długa piosenka? — zapytała Alicja* (...).

— *Jest długa — powiedział Rycerz — ale jest bardzo,* bardzo *piękna. Każdemu, kto usłyszał mnie śpiewającego ją... albo napływały łzy do oczu, albo...*

— *Albo co? — zapytała Alicja, gdyż Rycerz urwał niespodziewanie.*

— *Albo nie napływały. Nazwa tej piosenki nazywa się* Rybie Oczy.

— *To znaczy, tak się nazywa ta piosenka? — powiedziała Alicja, starając się okazać zainteresowanie.*

— *Nie, nie rozumiesz — powiedział Rycerz nieco zbity z tropu.*

— *Tak się* nazywa *jej nazwa. A nazwa brzmi* Staruszek.

— *Więc powinnam była powiedzieć: „Tak się nazywa ta piosenka?" — poprawiła się Alicja.*

— *Nie, nie powinnaś była, to zupełnie inna sprawa. Piosenka nazywa się* Sposoby i Drogi; *ale tak się tylko* nazywa, *wiesz!*

— *W takim razie, cóż to jest za piosenka? — powiedziała Alicja zupełnie już zdezorientowana w tej chwili...*

— *Właśnie zbliżałem się do tego — powiedział Rycerz. — Jest to piosenka* O siedzącym na furtce, *a melodia jest moim własnym wynalazkiem.* (...)

— *Ale melodia ta nie jest jego własnym wynalazkiem — powiedziała*

[5] B. Russel: *Mój rozwój filozoficzny.* Warszawa 1971, s. 66—67.

[6] L. Carrol: *O tym, co Alicja odkryła po drugiej stronie lustra.* Warszawa 1972, s. 121—122.

do siebie — *to* Daję ci wszystko, co ci mogę dać. *Stała słuchając w skupieniu, lecz łzy nie nabiegły jej do oczu.*

Przeanalizuj dialog Alicji z Rycerzem i zwróć uwagę na funkcję, jaką w nim pełnią wyrazy pisane prostym drukiem. Odpowiedz na następujące pytania: Jaka jest nazwa piosenki? Jaki jest jej tytuł? Czy tytuł jest identyczny z nazwą? O czym jest piosenka? Jaka jest jej melodia? Itd. itp.

W poniższych trzech metazagadkach, oraz w ćwiczeniu 6 b. i bajce na s. 21, predykaty *osoba x mówi prawdę, osoba x jest prawdomówna* oraz ich stylistyczne odmiany uważać będziemy za równoznaczne ze zwrotem: *Każde zdanie oznajmujące wypowiedziane przez osobę x jest prawdziwe.*

Analogicznie rozumieć będziemy wyrażenia *x jest kłamcą, x kłamie* oraz ich stylistyczne wersje jako synonimiczne z predykatem: *Każde zdanie wypowiedziane przez osobę x jest fałszywe.*

Przyjęte założenia nie są całkowicie niezgodne z potocznym rozumieniem rozważanych wyrażeń. Są jednak od tego rozumienia różne, bo np. kłamca to także ktoś, kto świadomie drugiego wprowadza w błąd, a prawdomówny — to mówiący zgodnie ze swymi przekonaniami.

ĆWICZENIE 3. (Metazagadka 1). Jeden z dwóch braci bliźniaków, imieniem Jan, popełnił przestępstwo. Rzecz w tym, by zidentyfikować Jana. W śledztwie ustalono, że co najmniej jeden z bliźniaków nigdy nie mówi prawdy. W czasie procesu sędzia zadaje obydwu bliźniakom to samo pytanie: *Czy oskarżony jest Janem?* Pierwszy z nich odpowiada *tak.* Drugi też odpowiedział, tylko nie wiemy, czy *tak*, czy *nie.* I na podstawie obydwu odpowiedzi sędzia zidentyfikował Jana. Pytanie do Czytelnika, który jest Janem — pierwszy, czy drugi?

ĆWICZENIE 4. (Metazagadka 2). Pewną wyspę zamieszkują tubylcy, którzy ze swej natury zawsze mówią prawdę, bądź nigdy nie mówią prawdy. Inaczej: wyspa zamieszkana jest przez ludzi dwóch kategorii — prawdomównych i kłamców. Na wyspę przybył mędrzec i napotkawszy dwóch tubylców, A i B, zapragnął zakwalifikować ich do kategorii, do których napotkani należą. Mieszkańcowi A postawił pytanie:

(1) Czy obaj jesteście prawdomówni?

Ten mu odpowiedział *tak* lub *nie* — ale Czytelnik nie wie co — a po namyśle mędrzec doszedł do logicznego wniosku, że na podstawie usłyszanej odpowiedzi nie zidentyfikuje napotkanych. Wtedy zadał mieszkańcowi A drugie pytanie:

(2) Czy obaj jesteście tej samej kategorii?

Mędrzec usłyszał odpowiedź i zidentyfikował spotkanych tubylców.

Pytanie do Czytelnika: Do jakich kategorii (prawdomównych czy kłamców) należą A i B?

ĆWICZENIE 5. (Metazagadka 3). W sądzie toczy się proces w sprawie trzech podejrzanych. Ustalono, że jeden z nich (ale nie ustalono który) jest prawdomówny, drugi — nigdy nie mówi prawdy, krótko: jest kłamcą, a trzeci jest *normalnym człowiekiem*, który niekiedy mówi prawdę, a niekiedy kłamie. I ten właśnie *normalny człowiek* był szpiegiem. Celem przesłuchań było zidentyfikowanie szpiega. Przyjmijmy, że A, B i C są imionami podejrzanych. Oto ich zeznania:

a) A zeznał dokładnie jedno z dwojga: albo że C jest szpiegiem, albo że C jest kłamcą;

b) B zeznał dokładnie jedno z trojga: albo że A jest prawdomówny, albo że A jest kłamcą, albo że A jest szpiegiem;

c) C zeznał dokładnie jedno z trojga: albo że B jest prawdomówny, albo że B jest kłamcą, albo że B jest szpiegiem.

Na podstawie zeznań sędzia zidentyfikował szpiega. Historię tę opowiedziano pewnemu logikowi, który po wszechstronnym przeanalizowaniu sprawy oświadczył, że nie ma wystarczających danych, by wskazać wśród podejrzanych szpiega. Wtedy powiedziano logikowi, co dokładnie zeznał A. I na tej podstawie logik zidentyfikował szpiega. Pytanie do Czytelnika: Który z podejrzanych jest szpiegiem — A, B, czy C?

III. UWAGI O ZDANIACH SAMOSTOSOWALNYCH

Największą sławą i zainteresowaniem logików cieszą się te wypowiedzi samostosowalne, które pojawiają się przy okazji prezentowania różnych antynomii sematycznych. Z problematyką samostosowalności wyrażeń (inaczej: samoodnoszenia się lub samozwrotności, ang. *self-reference*), w szczególności ze zdaniami samostosowalnymi zetknął się czytelnik w artykule Tarskiego *Prawda i dowód*. W niniejszym paragrafie podamy kilka szczegółów na ten temat. Uwagi nasze nie wyczerpują omawianej kwestii.

Zacznijmy od zdania wypisanego w wierszu oznaczonym cyfrą (1). Powiemy o nim:

(1) To zdanie jest złożone gramatycznie.

Czy zdanie (1) jest samostosowalne w tym sensie, że mówi (stwierdza) coś o sobie samym? Tym *coś* miałaby być złożoność gramatyczna zdania (1). Jest tak pod jednym wszakże warunkiem, że zaimek wskazujący *To*, stojący na początku zdania (1), odnosi się właśnie do zdania (1). Warunek ten ma charakter semantyczny i ustala (w naszym przypadku jednoznacznie) ów przedmiot, do którego wyrażenie okazjonalne *to* się odnosi. Wtedy rzeczywiście uważać możemy, że zdanie (1) mówi o sobie samym, że jest złożone. Jest przy tym zdaniem fałszywym!

Gdybyśmy ustalili inne znaczenie zaimka *to* i przyjęlibyśmy, na przykład, że *to* odnosi się do pierwszego zdania niniejszego paragrafu, wtedy zdania (1) nie uważalibyśmy za samostosowalne. W tym przypadku stwierdzałoby ono o pierwszym zdaniu punktu III niniejszej pracy, że jest złożone. W dodatku (1) byłoby zdaniem prawdziwym.

Weźmy teraz pod rozwagę odpowiednik zdania, którym posługuje się Tarski, kiedy wykłada antynomię kłamcy[7]:

(2) *Zdanie napisane kursywą na stronie 16 w niniejszym artykule nie jest prawdziwe.*

Czy (2) jest zdaniem zwrotnym? Przy jakich ewentualnie założeniach? Załóżmy dla wygody, że *z* jest skrótem (!) zdania (2).
Spójrzmy teraz na następującą równość:

(2.1) Zdanie napisane kursywą na stronie 16 w niniejszym artykule = „*z*".

Na mocy przyjętego przed chwilą założenia prawa strona równości (2.1) jest nazwą cudzysłowową zdania (2). Równość ta stwierdza (mówi), że identyczne są przedmioty, będące desygnatami jej członów. Jest zatem równością prawdziwą, o czym przekonujemy się spoglądając na niniejszą stronę tego artykułu. Fizyczne własności zdania (2) — kształt użytych czcionek, miejsce, w którym rozważane zdanie jest wypisane oraz to, że stronie 16 nie ma żadnego innego zdania napisanego kursywą — sprawiają, że (2.1) jest prawdziwe. A skoro tak, to stwierdzić możemy — w oparciu o elementarne prawo logiki dla identyczności — następującą równość:

(2.2) Zdanie napisane kursywą na stronie 16 w niniejszym artykule nie jest prawdziwe $\langle\equiv\rangle$ „*z*" nie jest prawdziwe.

Dopiero w oparciu o założenie empiryczne (2.1) i wynikającą z niego logicznie równoważność (2.2) uważamy zdanie (2) za samostosowalne. Mówi o sobie samym, że nie jest prawdziwe. Jak wiadomo fakt ten w połączeniu z konwencją (T), będącą eksplikacją klasycznego pojęcia prawdy, prowadzi do sprzeczności zwanej antynomią kłamcy[8].

Przypomnijmy, że definicja pojęcia prawdy (krótko: pojęcie prawdy) spełnia konwencję (T) lub inaczej: definicja jest merytorycznie trafna, gdy logiczną konsekwencją tej definicji są wszystkie równoważności następującej postaci:

---

[7] A. Tarski: *Prawda i dowód*. Op. cit., s. 15, wiersz 6.
[8] Ibidem, s. 13, 15.

(T) X jest prawdziwe $\Leftrightarrow p$,

gdzie w miejscu p występuje dowolne zdanie, zaś X jest nazwą jednostkową tego zdania. W szczególności X może być nazwą cudzysłowową zdania p. Konwencja (T) w tym przypadku jest równoważnością:

„p" jest prawdziwe $\Leftrightarrow p$.

Każdą równoważność postaci (T) uważać można za cząstkową definicję prawdy, w tym sensie, że ustala ona merytorycznie trafne rozumienie terminu *prawdziwy* w odniesieniu do konkretnego, pojedynczego zdania. Oczywiście nie wynika z konwencji (T), że w miejscu p mogą występować tylko zdania prawdziwe. Wprost przeciwnie, mogą tam być i zdania fałszywe. Dokładniej, p jest skrótem dowolnego zdania, które jest elementem dziedziny (zbioru zdań), w której predykat prawdy jest, względnie ma być, określony.

Załóżmy na chwilę, że strona 16 niniejszego artykułu została zredagowana inaczej i nie występują na niej żadne zdania napisane kursywą, zaś wiersz (2) zastąpiony został przez zdanie następujące (napisane zwykłą czcionką):

(2') Zdanie napisane kursywą na stronie 16 w niniejszym artykule jest prawdziwe.

Wtedy zdanie (2') nie odnosiłoby się do siebie samego, lecz do nie istniejącego obiektu: zdania napisanego kursywą na stronie 16 niniejszego artykułu. Odpowiadającą tej sytuacji równość (2.1) — przy założeniu, że „z" jest skrótem (2') — skłonni bylibyśmy uznać za fałszywą, o ile w ogóle mielibyśmy przypisywać jej wartość logiczną. Lewa bowiem strona równości byłaby nazwą pustą, nie posiadającą desygnatu, zaś desygnatem prawej strony byłoby zdanie w wierszu (2').

ĆWICZENIE 6. a) Rozważyć i przeanalizować przypadek, kiedy na stronie 16 prócz zdania (2) znajduje się jeszcze inne zdanie napisane kursywą. Dokładniej: przeprowadzić rozumowanie analogicznie do tego, które prowadzi do antynomii kłamcy [9] i wyjaśnić, dlaczego w tym przypadku nie pojawi się sprzeczność;

b) Mieszkaniec wyspy Krety (Kreteńczyk) imieniem Epimenides powiedział: *Wszyscy Kreteńczycy zawsze mówią nieprawdę*. Wykaż, że to zdanie Epimenidesa jest fałszywe. Jakie dodatkowe założenia należy przyjąć o Kreteńczykach, aby zdanie Epimenidesa prowadziło do sprzeczności (antynomii kłamcy)?

Powyższe rozważania prowadziliśmy z zamiarem postawienia następującego pytania: czy w językach naturalnych (takich jak np. język polski) można wypowiedzieć (skonstruować) zdania samozwrotne (samosto-

[9] Ibidem, s. 15.

sowalne), którym własność zwrotności przysługiwałaby jedynie z uwagi na ich strukturę gramatyczną i *naturalną* semantykę języka, bez potrzeby przyjmowania dodatkowych umów i założeń o charakterze empirycznym bądź semantycznym? Rozpatrywane przez nas zdania (1) i (2) nie były więc zdaniami *par excellance* samostosowalnymi.

W związku z postawionym pytaniem zacytujemy Tarskiego: *W tym celu musimy zanalizować te cechy języka potocznego, które są istotnym źródłem antynomii kłamcy. W trakcie analizy zauważamy natychmiast znamienną cechę języka potocznego, a mianowicie jego wszechobejmujący, uniwersalny charakter. Język potoczny jest uniwersalny i ci, którzy go używają, pragną, by był taki. Ma on dać możność wyrażenia wszystkiego, co w ogóle daje się wyrazić, w jakimkolwiek innym języku; dla spełnienia tego wymagania ulega on ciągłemu rozszerzaniu. W szczególności, język potoczny jest semantycznie uniwersalny w następującym sensie: obok tworów językowych, takich jak zdania i terminy, które są częściami składowymi języka, zawiera on również ich nazwy (jak wiemy, nazwy wyrażeń można tworzyć ujmując je w cudzysłowy); ponadto język ten zawiera terminy semantyczne, takie jak „prawda", „nazwa", „oznaczanie", które — pośrednio lub bezpośrednio — dotyczą stosunku między wyrażeniami i przedmiotami, do których te wyrażenia się odnoszą. W rezultacie, dla każdego zdania języka potocznego możemy sformułować w tym samym języku inne zdanie głoszące, że to pierwsze zdanie jest prawdziwe lub że jest fałszywe. Za pomocą pewnej dodatkowej „sztuczki" możemy nawet zbudować w języku potocznym zdanie, które można by nazwać samoodnośnym, a mianowicie zdanie Z, które stwierdza, że ono samo jest zdaniem prawdziwym lub że jest zdaniem fałszywym. W przypadku, gdy Z stwierdza swą własną fałszywość, proste rozumowanie wykazuje, że zdanie Z jest zarazem prawdziwe i fałszywe, i stajemy wobec antynomii kłamcy*[10].

Naszkicujemy teraz dwie konstrukcje, sztuczki — jak powiedziałby Tarski — które dają nam odpowiedź na postawione wyżej pytanie.

IV. KONSTRUKCJA NAZW I ZDAŃ SAMOSTOSOWALNYCH

Są zasadniczo dwa sposoby (sztuczki) budowania nazw samostosowalnych i zdań samostosowalnych. Pierwszy z nich polega na zastosowaniu metody przekątniowej (diagonalnej), a drugi związany jest z pojęciem normy wyrażenia. W pierwszym posługujemy się syntaktyczną operacją podstawiania, w drugim — konkatenacją wyrażeń.

[10] Ibidem, s. 18—19.

Zacznijmy od uściślenia terminów *nazwa samostosowalna* i *zdanie samostosowalne*. Pomocne nam w tym będzie pojęcie denotowania.

Denotowanie jest dla nas relacją zachodzącą pomiędzy nazwami rozważanego języka a pewną klasą przedmiotów. Wybrane wyrażenie oraz ustalony przedmiot pozostają do siebie w tej relacji, gdy owo wyrażenie jest nazwą jednostkową przedmiotu. Nieco formalniej rzecz całą ujmujemy w postaci następującej definicji:

DEFINICJA 1. Niech „N" będzie wyrażeniem. Mówimy, że „N" *denotuje* przedmiot a wtedy i tylko, gdy N jest identyczne z a, krótko: gdy N = a.

Uwaga. W definicji 1. „N" jest zmienną nazwową, tj. za którą można podstawić poszczególne wyrażenia nazwowe z języka polskiego. Tak więc „N" nie jest nazwą cudzysłowową litery En, lecz markuje miejsce, w którym można wstawić jakąś nazwę. Mamy tutaj dobrą anologię z konwencją (T). Ściśle rzecz ujmując definicja 1 jest cząstkową definicją denotowania: ustala czy owa relacja zachodzi czy nie między ustaloną nazwą „N" i przedmiotem a. Nie można zgeneralizować jej w sposób bezpośredni, tak samo jak nie można zgeneralizować konwencji (T) [11].

Problemu równoważnościowej definicji denotowania i jego ewentualnej redukcji do definicji prawdy nie będziemy w tym miejscu ani analizować ani rozwiązywać. Proponujemy go czytelnikowi jako ćwiczenie.

Nazwa *Stanisław August Poniatowski* denotuje ostatniego króla Polski, bo Stanisław August Poniatowski i ostatni król Polski to rzeczywiście te same osoby. Z definicji 1 wynika, że nazwa cudzysłowowa dowolnego wyrażenia denotuje to wyrażenie. Np. „*Niobe straciła wszystkie swoje dzieci*" denotuje *Niobe straciła wszystkie swoje dzieci*. To, że „N" denotuje a, wyrażamy w terminach logiki tradycyjnej mówiąc, że „N" jest nazwą jednostkową przedmiotu a.

Przeciwdziedzina relacji denotowania może w szczególności zawierać wyrażenia języka, a zatem dopuszczalne są konteksty

„N" denotuje „W"

w którym „W" jest zmienną reprezentującą dowolne wyrażenia. Obserwacja ta usprawiedliwia następującą definicję:

DEFINICJA 2. Nazwa „N" jest *samostosowalna*, gdy „N" denotuje „N".

[11] W. V. Quine: *Filozofia logiki*. Warszawa 1977, s. 23—24.

Jest zatem „N" samostosowalne, gdy N = „N", tzn., gdy „N" jest identyczne ze swoją nazwą cudzysłowową. Nie chodzi tutaj o identyczność graficzną między nazwami: nazwa jest samostosowalna, gdy ona i jej nazwa cudzysłowowa denotują ten sam obiekt.

DEFINICJA 3. *Diagonalizacja wyrażenia* E jest to wyrażenie otrzymane z E przez podstawienie w nim za wszystkie zmienne indywiduowe nazwy cudzysłowowej wyrażenia E.

Uwaga. W powyższej definicji E jest zmienną, za którą można podstawić nazwy wyrażeń, a nie dowolne nazwy. W tym względzie rzeczy mają się inaczej niż w przypadku Definicji 1.

Zauważamy, że każdemu wyrażeniu przyporządkowana jest zawsze dokładnie jedna jego diagonalizacja. Przyporządkowanie to jest więc funkcją, która — jak łatwo zauważyć — jest różnowartościowa.

ĆWICZENIE 7. Udowodnić, że tak jest rzeczywiście.

Ustalamy uwagę na wyrażeniu E_0, w którym „x" jest jedyną zmienną wolną:

(E_0) Diagonalizacja wyrażenia x.

Na mocy definicji 3 zastosowanej do E_0 mamy następującą identyczność:

Diagonalizacja wyrażenia *Diagonalizacja wyrażenia x* =
= *Diagonalizacja wyrażenia „Diagonalizacja wyrażenia x"*.

Stąd i z definicji 2 wnosimy, że wyrażenie: „Diagonalizacja wyrażenia *Diagonalizacja wyrażenia x*" jest nazwą samostosowalną.

Przejdźmy teraz do drugiego sposobu konstruowania nazw samozwrotnych.

DEFINICJA 4. *Normą* wyrażenia E jest wyrażenie będące konkatenacją E oraz nazwy cudzysłowowej wyrażenia E.

Na przykład, normą wyrazu „śpiewa" jest wyrażenie „śpiewa „śpiewa". Normą „Ala ma" jest „Ala ma „Ala ma" ", zaś normą „X kocha" jest wyrażenie „X kocha „X kocha" ".

Ustalamy uwagę na wyrażeniu E_1, które brzmi „normą". Na mocy definicji 4 mamy identyczność:

normą „normą" = „normą „normą" ",

która zaświadcza, że wyrażenie „normą „normą" " jest samostosowalne.

DEFINICJA 5. Zdanie podmiotowo-orzecznikowe postaci „S jest P" nazywamy *samostosowalnym*, gdy jego podmiot „S" denotuje to zdanie.

Dla dowolnego P wskażemy takie S, że zdanie „S jest P" będzie samostosowalne.

Niech „S" będzie skrótem nazwy:

Diagonalizacja wyrażenia *diagonalizacja wyrażenia x jest P*.

Wobec definicji denotowania nazwa ta denotuje następujące zdanie:

Diagonalizacja wyrażenia „*diagonalizacja wyrażenia x jest P*" jest P,

które — co widać z powyższej konstrukcji — jest zdaniem samostosowalnym.

ĆWICZENIE 8. Podaj przykłady takich „P", dla których powyższe zdanie jest a) prawdziwe, b) fałszywe.

Skonstruuj zdanie samostosowalne za pomocą funkcji normy.

V. PEWNE POSTACI TWIERDZENIA GÖDLA O NIEZUPEŁNOŚCI

Zacznijmy od bajki z logicznym morałem. W pewnej krainie żyją ludzie dwóch kategorii. Pierwsi, kategorii P, są prawdomówni, zaś drudzy, kategorii K — to kłamcy. Wśród prawdomównych wyróżnia się pewna grupa mieszkańców, którzy są twierdzeniomówni. Ich ogół oznaczymy przez T. Osobnicy należący do P wypowiadają wyłącznie zdania prawdziwe, zaś osobnicy kategorii K — tylko zdania fałszywe. Twierdzeniomówni wypowiadają wyłącznie zdania udowodnione. Zakładamy jeszcze, że pojęcie dowodu, którym posługują się mieszkańcy owej krainy jest poprawne, co znaczy, że każde zdanie posiadające dowód (twierdzenie) jest prawdziwe.

Zauważmy, że żaden z mieszkańców naszego logico-landu nie może wypowiedzieć zdania (+):

(+) Nie jestem prawdomówny,

oraz żadnego zdania, które jest mu równoważne: np. zdania postaci „$c \in P$", gdzie „c" jest nazwą indywiduową mówiącego.

Weźmy pod uwagę inną wypowiedź (++):

(++) Nie jestem twierdzeniomówny.

Zdania tego nie jest w stanie wypowiedzieć ani żaden kłamca ani żaden twierdzeniomówny. Może je jednak wypowiedzieć prawdomówny, który nie jest twierdzeniomówny, a więc osobnik ze zbioru P—T. Jeśli taki osobnik istnieje, to zdanie (++) przez niego wypowiedziane będzie zdaniem samostosowalnym (stwierdzającym pewną cechę mówiącego) oraz zdaniem prawdziwym. Może się też zdarzyć — bo ta ewentualność nie jest logiczną konsekwencją przyjętych założeń — że zdanie (+ +) nie będzie twierdzeniem.

ĆWICZENIE 9. Uzasadnić spostrzeżenia zawarte w powyższym akapicie.

Przytoczona powiastka zapowiada problematykę, która nas zajmie w dalszej części pracy. Intuicje w niej wyłożone zrekonstruujemy w pewnym systemie formalnym, który okaże się semantycznie niezupełny. Wskażemy w nim bowiem zdanie prawdziwe nie będące twierdzeniem. Zdanie to będzie można zinterpretować jako zdanie głoszące o sobie, że jest właśnie niedowiedlne. Będzie więc ono w pewnym sensie odpowiednikiem zdania (+ +).

Pochodzący od Raymonda Smullyana system formalny S, ze względu na swoją prostotę, pozwala na łatwe zrekonstruowanie schematu argumentacji Gödla. Dlatego schematu a nie kopii rozumowania Gödla, że to, co o (oraz w) oryginalnym systemie Gödla należało najpierw udowodnić jako twierdzenia wstępne, będzie tutaj przyjęte *a priori* w charakterze założeń. W szczególności abstrahować będziemy od wszystkich technikaliów związanych z arytmetyzacją języka. W zaprezentowanym schemacie kluczową rolę odegrają tak zwane rozumowania przekątniowe (diagonalne) — te same, które na terenie języka potocznego prowadzą do antynomii.

Czytelnikowi, który chciałby zapoznać się z przykładami rozumowań przekątniowych proponujemy lekturę: Huntera[12], Lyndona[13] oraz Nagla i Newmana[14].

Zamierzonym modelem (dziedziną) systemu S jest układ następujący:

$$U = \langle N, *, A_1, A_2, A_3, \ldots\rangle,$$

w którym:

(i) N jest zbiorem liczb naturalnych;

(ii) $*$ jest ustalonym działaniem dwuargumentowym w zbiorze N. Wartość funkcji $*$ dla argumentów m i n wynosi $m * n$. Nie zachodzi potrzeba bliższego specyfikowania działania $*$.

(iii) $A_1, A_2, A_3, \ldots$ jest ustalonym ciągiem podzbiorów liczb naturalnych. Nie zakładamy, że jest to ciąg różnowartościowy.

Dla dowolnego zbioru $X \in N$ definiujemy zbiór $X*$ następująco: dla dowolnego $n \in N$

$$n \in X* \Leftrightarrow n*n \in X.$$

---

[12] G. Hunter: *Metalogika*. Warszawa 1982, s. 23—25.
[13] R. Lyndon: *O logice matematycznej*. Warszawa 1968, s. 91—93.
[14] E. Nagel, J. R. Newman: *Twierdzenie Gödla*. Warszawa 1966, s. 45—51.

W alfabecie języka J_S systemu S mamy:
(a) zmienne indywiduowe

$$x_1, x_2, x_3 \ldots,$$

których zakresem zmienności jest zbiór liczb naturalnych N;
(b) stałe indywiduowe (liczebniki)

$$\underline{1}, \underline{2}, \underline{3}, \ldots,$$

które denotują odpowiednio liczby 1, 2, 3, ...;
(c) stałe indywiduowe

$$\underline{A}_1, \underline{A}_2, \underline{A}_3, \ldots,$$

które denotują odpowiednio zbiory A_1, A_2, A_3, ...;
(d) symbol funkcyjny $\underline{*}$ dwuargumentowy, który jest nazwą funkcji $*$ wyszczególnionej w punkcie (ii) powyżej;
(e) symbol relacyjny $\underline{\in}$ dwuargumentowy, który jest nazwą relacji należenia $\in$.

Wśród formuł atomowych w języku J_S są wyrażenia następującej postaci:

$$x_k \underline{\in} \underline{A}_n \text{ oraz } \underline{m} \underline{\in} \underline{A}_l.$$

Te drugie są zdaniami atomowymi.

Podkreślmy, że dla naszych celów nie jest konieczne jednoznaczne charakteryzowanie zbioru wszystkich formuł i zdań w języku J_S. W szczególności nie musimy wymieniać stałych logicznych, które ewentualnie w J_S mogą być.

Podzbiór liczb naturalnych A ma *indeks*, gdy A jest którymś wyrazem ciągu (iii), tj. gdy $A = A_i$ dla pewnego i. Wtedy indeksem jest liczba i. Zbiór A może nie mieć wcale indeksu, bądź może mieć jeden lub więcej indeksów. Zamiast mówić, że zbiór A ma indeks, powiadać będziemy, że zbiór A jest *wyrażalny* w systemie S. W szczegółowych wykładach oryginalnych ujęć twierdzeń Gödla i Tarskiego, wyrażalny może znaczyć np. reprezentowalny.

(Z_1) Zakładamy, że każda formuła w języku J_S ma w nim swój kod. W szczególności kodem zdania „$\underline{m} \underline{\in} \underline{A}_i$" jest term „$\underline{m} \underline{*} \underline{i}$", zaś kodem formuły „$x_k \underline{\in} \underline{A}_i$" jest term „$x_k \underline{*} \underline{i}$ ". *Numerem gödlowskim* formuły jest liczba naturalna, której nazwą (liczebnikiem) jest kod tej formuły.

(Z_2) Zakładamy dalej, że w systemie S określona jest pewna *relacja wyprowadzalności* i związany z nią zbiór *twierdzeń* systemu S. Przyjmujemy, że P jest zbiorem numerów gödlowskich wszystkich twierdzeń systemu S.

(Z_3) Zakładamy też, że dla systemu S zdefiniowano pojęcie prawdy.

Dowolne zdanie języka J_S jest *prawdziwe*, gdy jest ono prawdziwe w zamierzonej dziedzinie $(N, *, A_1, A_2, ...)$. W szczególności, zdanie „$\underline{n} \in \underline{A}_l$" jest prawdziwe, gdy rzeczywiście n jest elementem (należy do) zbioru A_l.

(Z_4) Zakładamy, że system S jest *poprawny*, co znaczy, że każde jego twierdzenie jest prawdziwe.

Na tym kończy się opis systemu S (*de facto* pewnej klasy systemów typu S).

Wyszczególnimy teraz trzy warunki o kluczowym znaczeniu dla wywodu niżej zaprezentowanego. Ich autorem jest Kurt Gödel.

(G_1) Zbiór P numerów gödlowskich twierdzeń systemu S jest w S wyrażalny.

(G_2) Jeśli zbiór liczb naturalnych X jest w S wyrażalny, to jego uzupełnienie, czyli zbiór $\overline{X} = N - X$, jest też w S wyrażalne.

(G_3) Jeśli zbiór liczb naturalnych X jest w S wyrażalny, to X^* jest też wyrażalny.

Odczytajmy te warunki w innej, lecz równoważnej terminologii. Warunek (G_1) głosi, że zbiór P ma indeks, a zatem, że ma w systemie S swoją nazwę: $\underline{A}_p$ dla pewnego p. (G_2) postuluje, żeby rodzina zbiorów $\{A_1, A_2, A_3, ...\}$ była domknięta na mnogościowe uzupełnienia do zbioru N, zaś (G_3) postuluje jej domkniętość na działanie*.

TWIERDZENIE 1 (Gödla o niezupełności). Jeśli system S spełnia warunki (G_1), (G_2) i (G_3), to w języku J_S dadzą się wypowiedzieć zdania prawdziwe nie będące twierdzeniami systemu S.

Dowód. Niech $K = \{n: n*n \notin P\}$. Udowodnimy, że

(1) $$K = (\overline{P})^*.$$

Rzeczywiście jest tak, bo zachodzi ciąg równoważności:

$$n \in K \Leftrightarrow n*n \notin P \Leftrightarrow n*n \in \overline{P} \Leftrightarrow n \in (\overline{P})^*.$$

Zatem — wobec (G_1), (G_2) i (G_3) — zbiór K jest wyrażalny w S, czyli

(2) $$K = A_k \text{ dla pewnego k.}$$

Stąd

(3) $$n \in A_k \Leftrightarrow n*n \notin P, \quad \text{dla dowolnego n.}$$

W szczególności

(4) $$k \in A_k \Leftrightarrow k*k \notin P.$$

Weźmy teraz pod uwagę zdanie $\gamma =$ „$\underline{k} \in \underline{A}_k$". Wtedy

(5) $$\gamma \text{ jest prawdziwe} \Leftrightarrow k \in A_k$$

(6) $$\gamma \text{ jest twierdzeniem w S} \Leftrightarrow k*k \in P.$$

Warunek (5) uzasadniamy przez odwołanie się do założenia (Z_3), zaś warunek (6) — przez odwołanie się do założeń (Z_1) i (Z_2).

Z (5), (6) i (4) mamy następującą równoważność:

(7) γ jest prawdziwe $\Leftrightarrow$ γ nie jest twierdzeniem systemu S.

Gdyby γ było twierdzeniem, to — wobec poprawności formalizacji systemu S — γ byłoby prawdziwe, co w połączeniu z (7) rodziłoby sprzeczność. Zatem γ nie jest twierdzeniem. Stąd i z (7) wnosimy, że γ jest zdaniem prawdziwym, co kończy dowód twierdzenia 1.

ĆWICZENIE 10. Odczytaj intuicyjny sens równości $K = \{n: n*n \notin P\}$ oraz zdania γ. Czy zdanie γ koresponduje ze stwierdzeniem (+ +) występującym w przypowiastce na początku niniejszego paragrafu?

Zmierzać teraz będziemy do udowodnienia twierdzenia Tarskiego o niedefiniowalności prawdy w systemie S. Załóżmy w tym celu, że T jest zbiorem numerów gödlowskich wszystkich zdań w języku J_S, które są prawdziwe. W terminologii *Prawdy i dowodu* moglibyśmy T określić mianem zbioru wszystkich *liczb prawdziwych*. Podstawową kwestią, którą rozważymy, jest następujące pytanie Tarskiego: Czy zbiór T jest wyrażalny w systemie E? Dokładniej, czy system spełniający warunki (G_2) i (G_3) spełnia również odpowiednik warunku (G_1) dla T?

Przyjęcie kilku roboczych definicji pomocne będzie w drobiazgowym przedyskutowaniu tej sprawy.

Niech σ będzie zdaniem w języku J_S, zaś X — jakimś zbiorem liczb naturalnych. Mówimy, że σ jest *zdaniem gödlowskim* zbioru X dokładnie wtedy, gdy 1° σ jest prawdziwe i jego numer gödlowski należy do X, bądź 2° σ jest fałszywe i jego numer gödlowski do X nie należy.

ĆWICZENIE 11. Sformułuj Twierdzenie 1 przy użyciu pojęcia zdania gödlowskiego.

System S jest *gödlowski*, gdy wszystkie zbiory w nim wyrażalne (tj. posiadające indeks) mają swoje zdania gödlowskie.

TWIERDZENIE 2. Jeśli system spełnia (G_3), to jest gödlowski.

Dowód. Niech X będzie zbiorem posiadającym indeks. Wtedy wobec (G_3) zbiór X^* też ma indeks, tzn. dla pewnego b: $X^* = A_b$. Zatem dla dowolnego n mamy:

(1) $$n \in X^* \Leftrightarrow n*n \in X.$$

Niech β będzie zdaniem „$\underline{b} \in A_b$". Pokażemy, że β jest zdaniem gödlowskim zbioru X. Z definicji prawdy, tzn. z warunku (Z_3), mamy:

(2) β jest prawdziwe $\Leftrightarrow b \in A_b$.

Kładąc w (1) $X = A_b$ oraz $n = b$ otrzymujemy

(3) $b \in A_b \Leftrightarrow b * b \in X$.

Ponieważ numerem gödlowskim zdania β jest liczba $b*b$, zatem wobec (2) i (3) wnosimy, że β jest rzeczywiście zdaniem gödlowskim dla X.

ĆWICZENIE 12. Uzasadnij, że każde zdanie jest zdaniem gödlowskim zbioru T. Jakie są zdania gödlowskie dla $\overline{T}$, tj. uzupełnienia zbioru T?

TWIERDZENIE 3 (Tarskiego o niedefiniowalności prawdy). Jeśli system S czyni zadość warunkom (G_2) i (G_3), to zbiór T nie jest w S wyrażalny.

Dowód. Wobec Twierdzenia 2 system S jest gödlowski, tzn. każdy zbiór w nim wyrażalny ma swoje zdanie gödlowskie. Ponieważ $\overline{T}$ nie ma w ogóle zdań gödlowskich, wnosimy stąd, że $\overline{T}$ nie jest zbiorem wyrażalnym. To zaś wobec warunku (G_2) implikuje, że zbiór T nie jest wyrażalny.

ĆWICZENIE 13. Niech system S spełnia warunki (G_1) i (G_2). Załóżmy też — zamiast warunku (G_3) — że S jest gödlowski. Udowodnij że istnieją wtedy zdania prawdziwe w języku J_S nie będące twierdzeniami systemu S. (Z uwagi na Twierdzenie 2 jest to uogólnienie Twierdzenia 1).

ĆWICZENIE 14. Wyprowadź Twierdzenie Gödla (tj. Twierdzenie 1) z Twierdzenia Tarskiego. Przeczytaj w związku z tym stosowny fragment *Prawdy i dowodu*[15].

Ostatni fragment niniejszego artykułu poświęcimy tzw. dualnej postaci dowodu Twierdzenia Gödla o niezupełności. Załóżmy w tym celu, że w języku J_S występuje negacja: spójnik zdaniowy, który semantycznie charakteryzujemy następująco: jeśli σ jest jakimś zdaniem w języku J_S, to

($\sim$) $\sim\sigma$ jest prawdziwe $\Leftrightarrow$ σ nie jest prawdziwe.

Mówimy, że zdanie jest *obalalne* na gruncie systemu S, gdy jego negacja jest twierdzeniem systemu S.

Zauważmy, że w systemie poprawnym, tj. spełniającym założenie (Z_4), żadne zdanie prawdziwe nie jest obalalne. Dalej, jeśli system jest niesprzeczny, to zdania obalalne nie są dowiedlne.

---

[15] A. Tarski: *Prawda i dowód*. Op. cit., s. 27.

Wcześniej przekonaliśmy się, że w każdym systemie spełniającym warunki (G_1)—(G_3) istnieje zdanie gödlowskie zbioru P. Jest nim zdanie γ występujące w dowodzie Twierdzenia 1. Prócz prawdziwości zdanie γ ma jeszcze dwie szczególne własności: jest *nierozstrzygalne* w S. Środkami formalnymi dostępnymi w S nie można udowodnić ani γ ani $\sim \gamma$. Z konstrukcji odczytujemy jego sens: γ stwierdza o sobie samym, że nie jest dowiedlne.

Stawiamy teraz pytanie o *dualną formę* dowodu twierdzenia Gödla o niezupełności: Pod jakimi warunkami istnieje zdanie ϱ o następujących cechach:

(1) ϱ jest fałszywe,

(2) ϱ jest nieobalalne,

(3) ϱ stwierdza swoją obalalność.

Z (1) — wobec poprawności formalizacji — mamy również, że

(4) ϱ jest niedowiedlne.

Niech R będzie zbiorem numerów gödlowskich wszystkich zdań obalalnych w S.

ĆWICZENIE 15. Jakie stosunki zakresowe zachodzą między zbiorami T, P i R?

TWIERDZENIE 4. Jeśli system S spełnia warunek (G_3) oraz zbiór R jest w S wyrażalny, to S jest niezupełny. Dokładniej: istnieje wtedy w języku J_S zdanie ϱ, które jest fałszywe, nieobalalne i stwierdzające swą obalalność.

Dowód. Na mocy założeń wnosimy, że zbiór R^* jest w S wyrażalny. Znaczy to, że istnieje takie r naturalne, że

(1) $$R^* = A_r.$$

Stąd dla dowolnej liczby naturalnej n mamy:

(2) $$n \in R^* \Leftrightarrow n * n \in R.$$

Z (1) i (2) wnioskujemy, że

(3) $$r \in A_r \Leftrightarrow r * r \in R.$$

Weźmy pod uwagę zdanie $\varrho =$ „$\underline{r} \underline{\in} \underline{A}_r$". Wtedy

(4) $$\varrho \text{ jest prawdziwe} \Leftrightarrow r * r \in R.$$

Ale numerem gödlowskim zdania ϱ jest liczba $r * r$. Zatem

(5) $$\varrho \text{ jest prawdziwe} \Leftrightarrow \varrho \text{ jest obalalne}.$$

W istocie ϱ nie jest obalalne. Bo gdyby było, to z uwagi na poprawność formalizacji mielibyśmy, że $\sim \varrho$ jest prawdziwe, a zatem ϱ byłoby fałszywe, co przeczyłoby równoważności (5). Wobec tego

(6) ϱ jest fałszywe i nie jest obalalne.

Z powyższego wynika niezupełność systemu, bo

(7) $\sim \varrho$ jest prawdziwe i $\sim \varrho$ nie jest twierdzeniem systemu S,

co kończy dowód Twierdzenia 4.

Podziękowania. Pragnę wyrazić wdzięczność panu dr. Jackowi Jadackiemu za merytorycznie trafną recenzję, dzięki której niniejsze opracowanie ma o wiele mniej różnego rodzaju uchybień niż miało ich przed wydrukowaniem.