Jan Woleński

Samozwrotność i odrzucanie

Filozofia Nauki 1/1, 89-102

1993

Filozofia Nauki
Rok I, 1993, Nr 1

Jan Woleński

Samozwrotność i odrzucanie

§1. Odrzucanie i Prawdomówca

Najbardziej znana wersja paradoksu Kłamcy jest następująca. Rozważamy zdanie:

(1) zdanie zapisane w niniejszym artykule pod numerem (1) jest fałszywe.

Oznaczamy to zdanie symbolem *b*. Mamy zatem:

(2) *b* = (*b* jest fałszywe).

Przyjmujemy, że schemat („wtw" jest skrótem dla „wtedy i tylko wtedy"):

(T) zdanie *A* jest prawdziwe wtw *A*

zachodzi dla zdania oznaczonego symbolem *b*. Tak więc:

(3) b jest prawdziwe wtw *b*.

Zastosowanie umowy (2) do (3) natychmiast daje:

(4) *b* jest prawdziwe wtw *b* jest fałszywe.

Ta ostatnia równoważność stanowi paradoks Kłamcy.

Uwaga 1. Przedstawiona wyżej wersja paradoksu Kłamcy pochodzi od Łukasiewicza [Łukasiewicz 1915]. Została spopularyzowana przez Tarskiego (z powołaniem się na Łukasiewicza, ale bez szczegółów bibliograficznych) w jego monografii o pojęciu prawdy w językach nauk dedukcyjnych [Tarski 1933]. Artykuł Łukasiewicza z 1915 r. jest rozszerzoną wersją jego popularnego i ważnego w historii filozofii polskiej eseju o twórczości w nauce [Łukasiewicz 1912], rozszerzoną m.in. o sformułowanie paradoksu Kłamcy (por. [Woleński 1990] w sprawie dalszych uwag historycznych). Dalej będę mówił o Tarskiego rozwiązaniu paradoksu Kłamcy, chociaż jest to uproszczenie historycznie także i z uwagi na rolę Leśniewskiego w tym względzie.

Uwaga 2. Wyraz „Kłamca" zastosowany do paradoksu Kłamcy w wersji (1)-(T)-(4) nie jest całkowicie właściwy. Historyczny paradoks Kłamcy został sformułowany w taki sposób, że wyraz „Kłamca" dotyczył osoby wypowiadającej zdanie „Ja teraz kłamię". Nazywanie zdania (1) „Kłamcą" stanowi niewątpliwe rozszerzenie użycia tej nazwy. Ściśle biorąc, powinno się używać określenia „paradoks zdania stwierdzającego swą własną fałszywość". Uwaga ta stosuje się *mutatis mutandis* do wyrazu „Prawdomówca" (termin ten podsunął mi Zdzisław Augustynek).

Rozwiązanie paradoksu Kłamcy zaproponowane przez Tarskiego polega na uznaniu, że zdanie (1) narusza reguły gramatyki logicznej. Owo naruszenie polega na pomieszaniu stopni języka; jeśli bowiem uznamy, że każde zdanie należy do języka określonego stopnia, a ponadto, że żadne zdanie nie może należeć do języków dwóch różnych stopni, to żadne zdanie nie może orzekać swej własnej fałszywości. Gdy zdanie A należy do języka stopnia n, to jego fałszywość jest orzekana w zdaniu „A jest fałszem", które należy do języka stopnia $n + 1$. Ostatecznie więc zdanie (1) jest niegramatyczne (tj. nonsensowne) z punktu widzenia zasad gramatyki logicznej.

Uwaga 3. Mogłoby się wydawać, że stwierdzenie, iż żadne zdanie nie może należeć do języków dwóch różnych stopni jest nietrafne wobec semantycznych konstrukcji Tarskiego, skoro można z nich wyprowadzić wniosek, że języki zawierają się w swych metajęzykach i, *a fortiori*, w językach wyższych stopni. W istocie rzeczy owo «zawieranie się» polega na tym, że zdania języków niższych stopni mają swe przekłady w językach wyższych stopni.

Uwaga 4. Wysłowienie niektórych tez sprawia pewne trudności. Oto np. powiada się, że zdanie (1) jest niegramatyczne (nonsensowne). Wszelako jeśli coś jest niegramatyczne, to nie może być zdaniem. Należałoby w tym wypadku mówić albo o wyrażeniach albo też o wyrażeniach zdaniokształtnych. W pewnych wypadkach będę posługiwał się takim bardziej pedantycznym sposobem wypowiadania się, aczkolwiek zasadniczo stosuję bardziej swobodny język, który na szczęście nie prowadzi do nieporozumień.

Rozwiązanie odwołujące się do stopni języka i niegramatyczności Kłamcy budzi szereg oporów. Wskazuje się np. [Kripke 1976], że (1) jest najzupełniej poprawnym zdaniem języka potocznego, a ponadto, że język ten nie wyróżnia stopni. Chociaż krytyka ta na pewno ma jakieś podstawy, to ignoruje ona fakt, że każde rozwiązanie paradoksu Kłamcy ma swą cenę.[1] Analiza Tarskiego doprowadziła do pewnej diagnozy rozważanej antynomii. Pojawia się ona, gdy (a) uznajemy schemat (T), (b) uznajemy logikę klasyczną, oraz (c) dopuszczamy języki zamknięte, tj. takie, które zawierają swoje własne metajęzyki. Rozwiązanie Tarskiego polega na przyjęciu (a) i (b), a odrzuceniu (c). Dalsza historia zagadnienia pokazała, że akceptacja (c) wymaga ograniczenia logiki klasycznej. Tak czy inaczej jakąś cenę trzeba zapłacić za rozwiązanie Kłamcy. Nie będę tego wątku dalej rozwijał, chociaż dla porządku wyznam, że modyfikacja logiki klasycznej wydaje mi się wyższą ceną od wykluczenia języków zamkniętych.

Rozwiązanie Tarskiego budzi jeszcze inne zastrzeżenie. Otóż, wydaje się, że niegramatyczność podlega pewnej zasadzie symetrii, mianowicie

(5) jeśli wyrażenie zdaniokształtne A jest niegramatyczne, to wyrażenie $\neg A$ jest również niegramatyczne.

Zasada (5), która powiada, że niegramatyczność jest zachowywana przez operację negacji zdaniowej wymaga, aby zdanie:

(6) zdanie zapisane w niniejszym artykule pod numerem (6) jest prawdziwe,

było niegramatyczne, jeśli niegramatyczne jest zdanie (1). Zdanie (6), które można określić jako „Prawdomówca", jest bowiem negacją (1). Wszelako Prawdomówca nie prowadzi do paradoksu. Oznaczmy to zdanie literą c. Mamy

(7) $c = (c$ jest prawdziwe),

ale rozumowania prowadzącego do paradoksu Kłamcy nie możemy powtórzyć. Można oczywiście umówić się, że jeśli wyrażenie zdaniokształtne A jest niegramatyczne, to jego negacja podziela ten sam los, ale umowa taka wygląda na rozwiązanie *ad hoc*. Znacznie lepiej byłoby, gdyby niegramatyczność (6) przejawiła się bezpośrednio, na jego własny rachunek. Otóż, zamierzam pokazać, że i Prawdomówca prowadzi do paradoksu samozwrotności.

Paradoks Prawdomówcy możemy uzyskać negując schemat (T) dla zdania (6). Zabieg taki, chociaż nieintuicyjny, jest możliwy, gdyż (T) nie jest tautologią logiczną. Odrzućmy więc:

(8) c jest prawdziwe wtw c,

co jest równoważne uznaniu:

(9) (c jest prawdziwe $\wedge \neg c$) lub [$\neg$(c jest prawdziwe) $\wedge$ c].

Stosując (7) otrzymujemy:

(10) ($c \wedge \neg c$) lub [c jest prawdziwe $\wedge$ $\neg$(c jest prawdziwe)].

Oba składniki alternatywy (10) są wewnętrznie sprzeczne, a więc i cała ta alternatywa jest antynomiorodna. I to jest właśnie paradoks Prawdomówcy związany z odrzuceniem schematu (T).

Uwaga 5. Gabriel Sandu z Uniwersytetu w Helsinkach podniósł (komunikat prywatny) pewien zarzut przeciwko powyższemu rozumowaniu. Powiada on, że (7) implikuje zachodzenie (T) dla zdania (6), gdyż zdania identyczne są *eo ipso* równoważne. Jednakże znak równości w (7) (a także w (2)) nie jest spójnikiem zdaniowym. Jest on traktowany jako wyrażający to, iż znak c jest skrótem typograficznym dla „c jest prawdziwe". Zakłada się tylko tyle, że obie strony (7) można wymieniać w kontekstach, które są dalej stosowane. Nie musimy nawet zakładać, że owa wymienialność dotyczy dowolnych kontekstów. Odnotujmy jeszcze, że wyrażenie $\neg c$ znaczy „negacja zdania, dla którego litera c jest skrótem". W związku z wyjaśnionym sensem symbolu c pojawia się jednak problem jak rozumieć pewne wyrażenia, np. „$c \wedge \neg c$", skoro litera c nie jest zmienną zdaniową. Otóż, wyrażenie $c \wedge \neg c$ rozumiem jako skrót dla zwrotu „koniunkcja składająca się ze zdania c jako pierwszego argumentu, symbolu spójnika koniunkcji oraz zdania oznaczonego symbolem $\neg c$ jako drugiego argumentu". Wątpliwości może też budzić brak cudzysłowów w większości kontekstów niniejszego artykułu, w których mówi się o wyrażeniach. Otóż przyjmuję, że terminy „wyrażenie", „litera", „symbol" itp. zastępują cudzysłowy, jeśli stoją przed literami używanymi jako symbole. Cudzysłowy stosuję, gdy takie zastępniki nie występują, lub też wyrażenia, o których mowa zawierają słowa potoczne. Ten drugi wypadek dotyczy np. zdania „c jest prawdziwe". Uważam, że pedantyczne stosowanie cudzysłowów w prezentacji wywodu, w którym stopnie języka i tak są pomieszane (tj. wywodu prowadzącego do Kłamcy) byłoby przesadą.

Analizując Kłamcę i Prawdomówcę odkrywamy jeszcze jedną ukrytą przesłankę, mianowicie, że użyta jest nie tylko logika klasyczna, ale także i operacja konsekwencji logicznej Cn. W istocie rzeczy, odpowiednie paradoksy są układami $((\neg \mathrm{T}) = \neg\, (\mathrm{T}))$:

(11) (a) Kłamca $= \langle b, (\mathrm{T}), Cn\{b, (\mathrm{T})\}\rangle$,

(b) Prawdomówca $= \langle c, (\neg\mathrm{T}), Cn\{c, (\neg\mathrm{T})\}\rangle$,

przy czym (**J** - język):

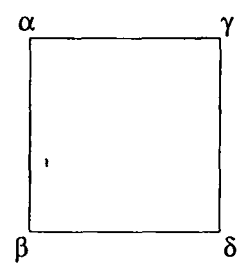
(12) $Cn\{b, (\mathrm{T})\} = Cn\{c, (\neg\mathrm{T})\} = \mathbf{J}$,

co pokazuje, że mamy do czynienia z argumentacjami prowadzącymi do sprzeczności. W (11) nie figurują (2) i (7), ponieważ są traktowane jako metajęzykowe (w stosunku do rozważanego języka) skróty terminologiczne (typograficzne).

Operacja Cn nie jest jednak jedyną konsekwencją logiczną związaną z logiką klasyczną. Inna, którą oznaczymy symbolem Cn^{N} i nazwiemy „konsekwencją negatywną", jest zdefiniowana przez:

(13) $A \in Cn^{\mathrm{N}} X$ wtw $\neg A \in Cn\ X^{\mathrm{N}}$.

Określenie to prowadzi do kwadratu logicznego (K):



gdzie litery α, β, γ i δ oznaczają odpowiednio zdania $A \in Cn\, X$, $\neg(A \in Cn^N X)$, $A \in Cn^N X$ i $\neg(A \in Cn\, X)$. Łatwo sprawdzić, że kwadrat (K) ujmuje związki opozycji zachodzące pomiędzy zdaniami α, β, γ i δ, które są zupełnie analogiczne do znanych zależności wśród zdań kategorycznych. W szczególności mamy:

(14) (a) $\neg(\alpha \wedge \gamma)$,
(b) $\neg(\alpha \Leftrightarrow \delta)$,
(c) $\neg(\beta \Leftrightarrow \gamma)$,
(d) $\alpha \Rightarrow \beta$,
(e) $\gamma \Rightarrow \delta$,
(f) $\beta \vee \delta$.

Zdanie β znaczy „*A* jest niesprzeczne ze zbiorem zdań *X*". Dalej możemy przyjąć:

(15) (a) zdanie *A* jest zależne od zbioru zdań *X* wtw $A \in Cn^N X$ lub $A \in Cn\, X$;

(b) zdanie *A* jest niezależne od zbioru zdań *X* wtw *A* nie jest zależne od zbioru *X*, tj. $A \notin Cn\, X$ oraz $A \notin Cn^N X$;

(c) zbiór zdań *X* jest zupełny wtw każde zdanie należące do języka **J** jest zależne od *X*;

(d) zbiór zdań *X* jest niezupełny wtw zbiór zdań *X* nie jest zupełny, tj. gdy przynajmniej jedno zdanie jest od niego zależne.

Konsekwencja negatywna należy do rodziny konsekwencji związanych z pojęciem odrzucania wraz z konsekwencją Cn^{-1} [Słupecki, Bryll, Wybraniec-Skardowska 1971] i konsekwencją dualną *dCn* [Wójcicki 1973, Spasowski 1983]. W gruncie rzeczy, jest to może najprostsza z konsekwencji odrzuceniowych zdefiniowana przez «*modulo* negację» konsekwencji *Cn*. Jest ona motywowana przez dążenie do ujęcia odrzucania jako doskonale symetrycznego wobec uznawania oraz uzyskania kwadratu logicznego, w którym obecna jest *Cn*.

Uwaga 6. Pojęcie dualności nie jest definiowane jednolicie w logice. Na ogół powiada się, że zmienne zdaniowe są dualne względem siebie, tak samo negacje formuł, a koniunkcja — względem alternatywy i na odwrót. Czasem jednak dualność jest określana tak, że zmienne są dualne względem swych negacji. Konsekwencja *dCn* jest dualna wobec *Cn* przy pierwszym rozumieniu dualności, a Cn^N — przy drugim.

Uwaga 7. Konsekwencje odrzuceniowe mają pewne znaczenie w analizie filozoficznej, np. w analizie logiki pojęcia „zakazu" czy też treści teorii [Woleński 1985, 1989 i 1992].

„Konsekwencja logiczna" jest pojęciem zasadniczo syntaktycznym. Ma jednak swe aspekty semantyczne i pragmatyczne. Konsekwencja zwykła jest skorelowana z prawdziwością, gdyż opisuje inferencje zachowujące prawdziwość. To jest jej aspekt semantyczny. Jest też skorelowana z procesami uznawania zdań, gdyż jeśli A jest uznane i $B \in Cn\, A$, to i B jest uznane. To jest aspekt pragmatyczny Cn. Pojęcie uznania jest przy tym traktowane tutaj jako logiczne, a nie psychologiczne, co umożliwia pominięcie drażliwego pytania, czy zbiór zdań uznanych zawsze jest domknięty konsekwencją logiczną. Tak więc konsekwencja zwykła jest semantycznie prawdziwościowa, a pragmatycznie — uznaniowa (asertywna). Kodyfikuje takie inferencje, które zawsze prowadzą od prawdziwych przesłanek do prawdziwych wniosków. Przypuśćmy, że chcemy zbudować teorię inferencji zachowujących fałsz, tj. takich, które nie mogą prowadzić od fałszywych przesłanek do prawdziwych wniosków. Operacja Cn do tego się nie nadaje ze względu na to, że „z fałszu wynika (wedle Cn-reguł) wszystko". To nieścisłe powiedzenie znaczy, że w poprawnym wnioskowaniu z fałszywych przesłanek mogą wynikać zarówno prawdy jak i fałsze. Natomiast konsekwencja negatywna (podobnie jak Cn^{-1} i dCn) jest semantycznie fałszywościowa (dzięki temu kodyfikuje inferencje zachowujące fałszywość), a nadto pragmatycznie — odrzuceniowa (rejektywna), bo zachowuje odrzucanie. Niewątpliwie, gdy odrzucamy A, to jesteśmy też skłonni odrzucić $A \wedge B$, ale tego faktu nie można wprost wyrazić przy pomocy Cn. Konsekwencja negatywna to umożliwia, gdyż („$\dashv A$" znaczy „A jest odrzucone"):

(16) $X \dashv A$ wtw $A \in Cn^{\mathrm{N}} X$,

a w szczególności:

(17) $(A \wedge B) \in Cn^{\mathrm{N}} A$.

Formuła (18) jest odpowiednikiem („$\vdash A$" znaczy „A jest uznane"):

(18) $X \vdash A$ wtw $A \in Cn\, X$.

Konsekwencja negatywna ma tzw. ogólne własności konsekwencji logicznej, m.in. jest monotoniczna ($X \subseteq Y \Rightarrow Cn^{\mathrm{N}} X \subseteq Cn^{\mathrm{N}} Y$), idempotentna ($Cn^{\mathrm{N}}Cn^{\mathrm{N}} X = Cn^{\mathrm{N}} X$) oraz finitystyczna (jeśli $A \in Cn^{\mathrm{N}} X$, to A należy do konsekwencji jakiegoś skończonego pozdzbioru zbioru X). Logika odrzucania (Cn^{N}-logika lub N-logika) różni się rzecz jasna od logiki uznawania (Cn-logiki lub U-logiki), ponieważ zbiory tez są w obu wypadkach różne, a nawet rozłączne. Nie możemy np. wyrazić (17) w przedmiotowym języku logiki (związanej z Cn) przez

(19) $A \Rightarrow (A \wedge B)$,

jeśli „$\Rightarrow$" oznacza implikację materialną. Trzeba wprowadzić nową implikację, oznaczoną np. „$\Rightarrow^{\mathrm{N}}$", zdefiniowaną jako:

(20) $A \Rightarrow^{\mathrm{N}} B =^{\mathrm{df}} (A \vee \neg B)$,

a wtedy otrzymamy:

(21) $A \Rightarrow^{\mathrm{N}} (A \wedge B)$

jako tezę logiki negatywnej ($\mathbf{LOG}^{\mathrm{N}}$). Mamy dalej:

(22) (a) $Cn^{\mathrm{N}}\{A\} \cap Cn^{\mathrm{N}}\{\neg A\} = \mathbf{LOG}^{\mathrm{N}}$

(b) $Cn^{\mathrm{N}}\{A, \neg A\} = \mathbf{J}$,

a $\mathbf{LOG}^{\mathrm{N}}$ możemy też określić jako $Cn^{\mathrm{N}}\varnothing$.

Zapewne nie jest rzeczą przypadku, że logika rozwinęła się w związku z Cn, a nie Cn^N lub jakąś inną konsekwencją odrzuceniową. Dlatego też operowanie przedmiotową $\mathbf{LOG}^N$ nie jest zbyt naturalne. Rzecz można jednak uprościć w ten sposób, że się będzie operowało pojęciem „odrzucania" (także rozumianym logicznie, a nie psychologicznie) w metajęzyku, który jest asertywny. I tak np. schemat:

(23) jeśli odrzucone jest zdanie A, to odrzucona jest również i koniunkcja $A \wedge B$, wyraża w asertywnym metajęzyku tezę (17). Znaczy to, że funktor „jeśli, to" jest użyty w (23) w swym normalnym asertywnym znaczeniu, a $\mathbf{LOG}^N$ jest tutaj ukryta w pojęciu odrzucania. Ponadto, jeśli mamy regułę o kształcie:

(24) jeśli $\dashv A$, to $\dashv B$,

to możemy zastępować formuły A i B ich «uznaniowymi równoważnikami»; np. formuła $A \wedge B$ może być zastąpiona formułą $\neg(\neg A \vee \neg B)$. Mamy m.in. następujące reguły odrzucania

(25) (a) $\dashv (A \vee B)$ wtw $\dashv A$ oraz $\dashv B$,

(b) $\dashv (A \wedge B)$ wtw $\dashv A$ lub $\dashv B$,

(c) $\dashv (A \Leftrightarrow B)$ wtw $\dashv (A \wedge B)$ oraz $\dashv (\neg A \wedge \neg B)$,

(d) jeśli $A \Leftrightarrow B$ oraz $\dashv A$, to $\dashv B$,

(e) kontrtautologie (w sensie Cn-logiki) są odrzucone,

(f) tautologie (w sensie Cn-logiki) są przeciwodrzucone.

Reguła (25f) umożliwia identyfikację paradoksu w ramach logiki odrzucania, jeśli w jakimś wierszu dowodowym pojawi się tautologia, podobnie jak wystąpienie sprzeczności w dowodzie zrelatywizowanym do Cn informuje o pojawieniu się sytuacji antynomialnej (wyraz „przeciwodrzucone" został zastosowany w (25d) tylko po to, aby uniknąć niezręcznego w tym kontekście określenia „uznane").

Użycie Cn w zapisie (11) sugeruje, że operator $\dashv$ (czyli Cn-logika) jest właściwy dla formalizacji Kłamcy i Prawdomówcy. Wersja Prawdomówcy zachowująca schemat (T) oparta jest na założeniu, że operator $\dashv$ (czyli N-logika) jest właściwy dla zdania stwierdzającego swą własną prawdziwość. Takim zdaniem jest $\neg$(1). Skoro „b" jest skrótem dla „b jest fałszywe", mamy:

(26) $\neg b$ = (b jest prawdziwe).

Załóżmy:

(27) $\dashv \neg b$.

Założenie to razem z (26) prowadzi do:

(28) $\dashv$ b jest prawdziwe.

Schemat (T) zastosowany do b daje:

(29) $b \Leftrightarrow$ (b jest prawdziwe),

a z tego przez (25d) otrzymujemy:

(30) $\dashv b$.

Z (27) i (30) wynika (por. (25a):

(31) $\dashv (b \vee \neg b)$.

Tak więc, założenie (26) prowadzi (por. (25f)) do sprzeczności. Załóżmy teraz:

(32) $\dashv b$.

Schemat (T) i (25d) dają:

(33) $\dashv b$ jest prawdziwe,

a zastosowanie (26) do (33):

(34) $\dashv \neg b$.

Tak więc konkluzja (31) wynika również z założenia (32), co kończy dowód Prawdomówcy. Paradoks ten możemy zapisać jako:

(35) Prawdomówca = $\langle \neg b, Cn^{N}\{\neg b\}\rangle$,

przy czym schemat (T) został w tym wypadku «przeniesiony» do metajęzyka. Możemy jednak uzyskać Prawdomówcę przez przyjęcie:

(T^{N}) $\dashv$ (A jest prawdziwe wtw $\neg A$).

Jeśli zastosujemy (T^{N}) do zdania $\neg b$, to możemy wykazać, że zdanie „A jest prawdziwe wtw $\neg A$" nie da się odrzucić ani przy założeniu (27) ani też (30). Wtedy (35) przechodzi w:

(36) Prawdomówca = $\langle \neg b, (T^{N}), Cn\{\neg b, (T^{N})\}\rangle$.

Tak więc i zdanie orzekające swą własną prawdziwość jest antynomialne, nawet przy zachowaniu schematu (T). Dzięki temu zasada (5) jest spełniona, bo zdania (1) i (6) są niegramatyczne ze względu na wzajemnie zwierciadlane (chociaż nieidentyczne) warunki: z jednej strony mamy Cn, a z drugiej — Cn^{N}.

Jestem zwolennikiem likwidacji paradoksu kłamcy przez wykluczenie języków zamkniętych, a wykazanie symetrii w kwestii niegramatyczności zdań (1) i (6) jest dodatkowym argumentem na rzecz rozwiązania Tarskiego. Wszelako chciałbym coś dodać w sprawie zasadności wprowadzenia hierarchii języków wedle ich stopni i relatywizacji pojęć semantycznych do określonego języka. Argument, że jest to obce językowi potocznemu, wydaje mi się całkowicie chybiony, gdyż jest podobny do powiedzenia, że relatywizacja praw mechaniki do układów odniesienia jest nietrafna, bo obca potocznemu rozpatrywaniu ruchu. Stosunek między semantyką Tarskiego a tzw. semantyką naiwną jest podobny do stosunku między mechaniką teoretyczną a mechaniką naiwną.

Niemniej jednak pragmatyczna analiza paradoksów samozwrotności podpowiada pewne ich rozwiązanie, które może warte jest wzmianki. Wspomniałem wyżej, że zachodzi wyraźny związek między Cn, prawdą i asercją z jednej strony oraz Cn^{N}, fałszem i odrzucaniem — z drugiej. Analiza paradoksu Kłamcy pokazuje, że do zdania (1) zawierającego samozwrotnie przymiotnik „fałszywe", stosowana jest konsekwencja Cn, natomiast do zdania (6), zawierającego samozwrotnie przymiotnik „prawdziwe" — konsekwencja Cn^{N}. Wolno to uznać za niekonsekwencję, gdyż jeśli już wypowiada się (1) czy też (2), to należałoby uznać odpowiednio niewłaściwość używania Cn lub Cn^{N}

w dalszej analizie tych zdań. Należałoby tak uczynić ze względu na «zobowiązania pragmatyczne», a nie w sensie absolutnym, gdyż zawsze można utrzymywać, iż dowolna konsekwencja logiczna może być aplikowana jako operacja syntaktyczna do dowolnego zbioru zdań jako obiektów składniowych. Niemniej jednak pragmatyczne wyjaśnienie paradoksów odwołujące się do zgodności składni, semantyki i pragmatyki, chociaż nieabsolutne, ma jednak pewien walor wobec języka naturalnego.

Uwaga 8. Wcześniejsza wersja niniejszego artykułu zawierała paradoks Prawdomówcy tylko w wersji zakładającej odrzucenie (T). Tak też przedstawiłem sprawę w odczycie wygłoszonym na posiedzeniu naukowym Zakładu Semiotyki Logicznej Uniwersytetu Warszawskiego i Polskiego Towarzystwa Semiotycznego w dniu 30 kwietnia 1993 r. Dyskusja, a zwłaszcza głos Andrzeja Grzegorczyka, przekonały mnie jednak, że wersja (11b) nie jest intuicyjna. Trzeba jednak zauważyć, że negacja (T) sama przez się nie generuje żadnego paradoksu, gdyż np. nie wystarcza do dowodu, że zdanie (1) prowadzi do paradoksu.

§2. Odrzucanie i niezupełność

Nieformalne wyjaśnienie fenomenu niezupełności arytmetyki jest nastepujące [Gödel 1931.175-176]. Przyjmujemy, że nasza logika jest adekwatna, tj. zachowuje prawdę, oraz, że aksjomaty arytmetyki Peano są prawdziwe. Rozważmy teraz zdanie

(37) zdanie zapisane w niniejszym artykule pod numerem (37) jest niedowodliwe w arytmetyce.

Zdanie (37) stwierdza więc własną niedowodliwość w arytmetyce. Załóżmy, że zdanie (37) jest dowodliwe. Znaczy to, ze względu na adekwatność naszej logiki, że jest ono prawdziwe. Ale jeśli jest prawdziwe, to ze względu na swą treść jest właśnie niedowodliwe, co przeczy przypuszczeniu, że jest dowodliwe. Jeśli natomiast negacja zdania (37) byłaby dowodliwa, to samo zdanie (37) musiałoby być fałszywe, a więc, ze względu na swą treść, dowodliwe. Wtedy jednak dowodliwe byłoby zarówno zdanie (37), jak i jego negacja, co jest niemożliwe. Cały dalszy wywód Gödla poświęcony był wykazaniu, że niezupełność arytmetyki (teoria T jest niezupełna wtedy i tylko wtedy, gdy istnieje formuła A taka, że A należy do języka teorii T oraz $A \notin Cn\, T$ i $\neg A \notin Cn\, T$) daje się wykazać bez odwoływania się do pojęcia „prawdy". Oryginalny dowód Gödla dotyczył twierdzenia (jest to tzw. I twierdzenie Gödla o niezupełności):

(38) jeśli arytmetyka (dalej **AR**) jest niesprzeczna (dokładniej: ω-niesprzeczna), to jest niezupełna.

Niemniej jednak (38) formułuje się i tak:

(39) jeśli **AR** jest niesprzeczna, to istnieje prawdziwe zdanie A należące do języka **AR** takie, że $A \notin Cn$ **AR**.

Takim prawdziwym zdaniem jest np. (37) zapisane w kodzie arytmetycznym przez zastosowanie arytmetyzacji (uwaga ta jest ważna wobec wszystkich dalszych przykładów rozważanych nieformalnie), bo stwierdza swą własną niedowodliwość i rzeczywiście jest niedowodliwe. Twierdzenia (37), (39) oraz prawo wyłączonego środka dają:

(40) jeśli **AR** jest niesprzeczna, to istnieje fałszywe zdanie A należące do języka **AR** takie, że $A \notin$ **AR**.

Chciałoby się jednak wykazać (40) niejako wprost, a nie tylko pośrednio. Możemy to zrobić przy pomocy $\mathbf{LOG}^{\mathrm{N}}$. Smullyan [Smullyan 1961. 49, 129-130, Smullyan 1992. 11] rozważa zdanie:

(41) zdanie zapisane w niniejszym artykule pod numerem (41) jest odrzucalne w arytmetyce.

Gdy zastosujemy schemat (T) do (41), dostaniemy:

(42) zdanie (41) jest prawdziwe wtw zdanie (41) jest odrzucalne w arytmetyce,

co dalej prowadzi do:

(43) zdanie (41) jest prawdziwe i odrzucalne w arytmetyce lub zdanie (41) jest fałszywe i nieodrzucalne w arytmetyce.

Zakładając, że nasza logika jest adekwatna, wnioskujemy, że (41) nie może być prawdziwe i odrzucalne lub fałszywe i dowodliwe. Musi więc być fałszywe i niedowodliwe.

Rozumowanie to zakłada, że odrzucalność jest zrelatywizowana do Cn. Wszelako rezultat Smullyana można uzyskać przy pomocy $\mathbf{LOG}^{\mathrm{N}}$ i to bez odwoływania się do schematu (T). Przyjmijmy, że aksjomaty arytmetyczne są fałszywe, a nasza logika zachowuje fałsz. Taką logiką jest właśnie $\mathbf{LOG}^{\mathrm{N}}$. Rozważamy teraz zdanie:

(44) zdanie zapisane w niniejszym artykule pod numerem (44) jest Cn^{N}-dowodliwe (dowodliwe relatywnie do konsekwencji Cn^{N}) w $\mathbf{AR}^{\mathrm{N}}$ (tj. w teorii opartej na negacji aksjomatów **AR**),

i rozumujemy całkowicie analogicznie jak w wypadku zdania (37). Załóżmy, że (44) jest Cn^{N}-dowodliwe w $\mathbf{AR}^{\mathrm{N}}$. Nie może być prawdziwe, gdyż wtedy, ze względu na własności Cn^{N}, nie byłoby Cn^{N}-dowodliwe. Jest więc fałszywe, a skoro tak, to jest Cn^{N}-niedowodliwe. Załóżmy teraz, że zaprzeczenie zdania (44) jest Cn^{N}-dowodliwe. Wtedy jednak zarówno zdanie (44), jak i jego negacja, należałyby do zdań Cn^{N}-dowodliwych. Ponieważ Cn^{N}-dowodliwość jest tym samym co odrzucanie, Cn^{N}-dowodliwa byłaby też alternatywa złożona ze zdania (44) oraz jego negacji, co jest niemożliwe ze względu na (25f).

Relatywizacja do Cn^{N}-dowodliwości może być pominięta, jeśli rozszerzymy pojęcie dowodliwości arytmetycznej przez stypulację:

(45) zdanie A jest dowodliwe arytmetycznie wtw, gdy jest Cn-dowodliwe (tj. dowodliwe w sensie zwykłym) lub Cn^{N}-dowodliwe.

Przy takiej umowie wolno nam stwierdzić:

(46) jeśli **AR** jest niesprzeczna, to istnieje fałszywe zdanie A należące do języka **AR**, takie, że A jest niedowodliwe arytmetycznie.

To właśnie chcieliśmy wykazać.

Uwaga 9. Twierdzenie (46) winno zawierać relatywizację do N-niesprzeczności $\mathbf{AR}^{\mathrm{N}}$, co znaczy, że wśród tez tej teorii nie ma zdania kształtu „A lub $\neg A$". Byłoby tak wtedy, gdyby wśród twierdzeń $\mathbf{AR}^{\mathrm{N}}$ znalazło się jakieś zdanie i jego negacja. Tak więc warunek dla N-niesprzeczności jest dokładnie taki sam, jak dla niesprzeczności w rozumieniu standardowym. Jeśli więc **AR** jest niesprzeczna, to $\mathbf{AR}^{\mathrm{N}}$ jest N-niesprzeczna.

Uwaga 10. Prawdziwość i fałszywość w odniesieniu do **AR** i $\mathbf{AR}^{\mathrm{N}}$ jest rozumiana jako prawdziwość i fałszywość w modelu standardowym. Zakładam, że ze względu na ten model aksjomaty arytmetyczne są uznawane, a ich negacje odrzucane. Można postępować inaczej, tj. uznać wszystkie lub niektóre negacje aksjomatów Peana za prawdziwe. Wtedy jednak będziemy mieli do czynienia z prawdziwością w modelach niestandardowych arytmetyki.

Uwaga 11. Niezupełność arytmetyki może być też określona tak samo jak w (15d).

Możemy dalej sformułować odpowiednik twierdzenia (38) jako:

(47) jeśli $\mathbf{AR}^{\mathrm{N}}$ jest N-niesprzeczna, to $\mathbf{AR}^{\mathrm{N}}$ jest także N-niezupełna, tj. istnieje zdanie A należące do języka $\mathbf{AR}^{\mathrm{N}}$ takie, że $A \notin Cn^{\mathrm{N}}\, \mathbf{AR}^{\mathrm{N}}$ oraz $\neg A \notin$ $\notin Cn^{\mathrm{N}}\, \mathbf{AR}^{\mathrm{N}}$,

a także analogon II twierdzenia Gödla o niezupełności (jeśli **AR** jest niesprzeczna, to formuła *CON* wyrażająca niesprzeczność **AR** jest niedowodliwa w **AR**) jako:

(48) jeśli $\mathbf{AR}^{\mathrm{N}}$ jest N-niesprzeczna, to formuła CON^{N} wyrażająca N-niesprzeczność $\mathbf{AR}^{\mathrm{N}}$ jest niedowodliwa w $\mathbf{AR}^{\mathrm{N}}$.

Zdanie (37) jest niedowodliwe arytmetycznie. W 1952 r. Henkin [Henkin 1952] postawił problem dowodliwości zdania stwierdzającego własną dowodliwość, a więc zdania („dowodliwe" znaczy „Cn-dowodliwe"):

(49) zdanie zapisane w niniejszym artykule pod numerem (49) jest dowodliwe w arytmetyce.

Z uwagi na schemat (T), zachodzi:

(50) zdanie (49) jest prawdziwe wtw zdanie (49) jest dowodliwe w arytmetyce,

co daje:

(51) zdanie (49) jest prawdziwe i dowodliwe w arytmetyce lub zdanie (49) jest fałszywe i niedowodliwe w arytmetyce.

Która z tych możliwości zachodzi? Problem ten został rozwiązany przez Löba [Löb 1955]. Prawdziwe jest mianowicie twierdzenie:

(52) dla każdego *A*, jeśli zdanie „jeśli *A* jest *Cn*-dowodliwe w **AR**, to *A*" jest *Cn*-dowodliwe w **AR**, to *A* jest *Cn*-dowodliwe w **AR**.

Z twierdzenia Löba wynika, że zdanie orzekające swą własną dowodliwość jest zarazem prawdziwe i dowodliwe. Dla sformułowania odpowiednika twierdzenia Löba dla $\mathbf{AR}^{\mathrm{N}}$ trzeba rozważyć zdanie:

(53) zdanie zapisane w niniejszym artykule pod numerem (53) jest Cn^{N}-niedowodliwe w $\mathbf{AR}^{\mathrm{N}}$.

Mamy dalej:

(54) zdanie (53) jest prawdziwe wtw zdanie (54) jest Cn^{N}-niedowodliwe w $\mathbf{AR}^{\mathrm{N}}$,

a w konsekwencji:

(55) zdanie (53) jest Cn^{N}-niedowodliwe i prawdziwe w $\mathbf{AR}^{\mathrm{N}}$ lub zdanie (53) jest Cn^{N}-dowodliwe i fałszywe w $\mathbf{AR}^{\mathrm{N}}$.

Odpowiednikiem twierdzenia Löba dla $\mathbf{AR}^{\mathrm{N}}$ jest:

(56) dla każdego *A*, jeśli zdanie „jeśli *A* jest Cn^{N}-dowodliwe w $\mathbf{AR}^{\mathrm{N}}$, to *A*" jest Cn^{N}-dowodliwe w $\mathbf{AR}^{\mathrm{N}}$, to *A* jest również Cn^{N}-dowodliwe w $\mathbf{AR}^{\mathrm{N}}$.

Zdanie (53) mówi o sobie, że jest Cn^{N}-niedowodliwe w $\mathbf{AR}^{\mathrm{N}}$. Możemy je zapisać nieformalnie jako:

(57) jestem Cn^{N}-niedowodliwe w $\mathbf{AR}^{\mathrm{N}}$.

Jeśli zastosujemy (56) do (57), otrzymamy:

(58) jeśli zdanie „jeśli 'jestem Cn^{N}-niedowodliwe w $\mathbf{AR}^{\mathrm{N}}$' jest Cn^{N}-dowodliwe w $\mathbf{AR}^{\mathrm{N}}$, to jestem Cn^{N}-niedowodliwe w $\mathbf{AR}^{\mathrm{N}}$", jest Cn^{N}-dowodliwe w $\mathbf{AR}^{\mathrm{N}}$, to zdanie „jestem Cn^{N}-niedowodliwe w $\mathbf{AR}^{\mathrm{N}}$ " jest Cn^{N}-dowodliwe w $\mathbf{AR}^{\mathrm{N}}$.

Analiza (w języku asertywnym) zdania (58) doprowadza do wniosku, że uznana jest koniunkcja:

(59) „jestem Cn^{N}-niedowodliwe w $\mathbf{AR}^{\mathrm{N}}$" jest Cn^{N}-dowodliwe w $\mathbf{AR}^{\mathrm{N}}$ oraz nieprawda, że jestem Cn^{N}-niedowodliwe w $\mathbf{AR}^{\mathrm{N}}$.

Tak więc (52) stwierdza, że zachodzi drugi człon alternatywy wyrażonej w (55).

Uwaga 12. Ścisłe dowody twierdzeń (47), (48) i (56) wymagają użycia tych samych technik, które są stosowane w dowodach odpowiednich twierdzeń o **AR**.

Pojawia się wreszcie problem definiowalności pojęcia „fałszu arytmetycznego" (czy też zbioru fałszywych zdań arytmetycznych) w $\mathbf{AR}^{\mathrm{N}}$, co jest przedmiotem następującego analogonu twierdzenia Tarskiego o niedefiniowalności prawdy arytmetycznej w **AR** :

(60) Zbiór zdań fałszywych **AR** jest niedefiniowalny w $\mathbf{AR}^{\mathrm{N}}$, jeśli $\mathbf{AR}^{\mathrm{N}}$ jest N-niesprzeczna.

Dowód (60) jest wzorowany na rozumowaniu Smullyana [Smullyan 1992. 103] uzasadniającym jedną z wersji twierdzenia Tarskiego. Przyjmijmy, że predykat $F(x)$ jest predykatem fałszu dla teorii T, o ile zachodzi (symbol $\overline{A}$ zaznacza, że formuła A zapisana jest w kodzie arytmetycznym):

(61) $\forall A\ (F(\overline{A}) \Leftrightarrow A) \in Cn^{\mathrm{N}}\, T^{\mathrm{N}}$.

Niech $F(x)$ będzie predykatem fałszu dla $\mathbf{AR}^{\mathrm{N}}$. Ze względu na ogólne własności arytmetyki, każda formuła typu $P(\overline{A})$ ma swój punkt stały, tj. formułę A, co znaczy, że równoważność:

(62) $P(\overline{A}) \Leftrightarrow A$

jest dowodliwa w **AR** dla każdego A. W związku z tym, formuła:

(63) $P(\overline{A}) \Leftrightarrow \neg A$

jest odrzucona w $\mathbf{AR}^{\mathrm{N}}$ dla każdego A; formułę $\neg A$ można uważać za N-punkt stały formuły $P(\overline{A})$ w $\mathbf{AR}^{\mathrm{N}}$. Tak więc, jeśli położymy „$\neg F$" w miejsce P, otrzymamy:

(64) $\exists A\ (\neg F(\overline{A}) \Leftrightarrow \neg A) \in Cn^{\mathrm{N}}\, \mathbf{AR}^{\mathrm{N}}$.

Jednakże zachodzi sprzeczność pomiędzy (61) oraz (64) w tym sensie, że dla pewnej formuły A, $\dashv A$ oraz $\dashv \neg A$, wbrew założeniu, że $\mathbf{AR}^{\mathrm{N}}$ jest niesprzeczna. Znaczy to, że pojęcie fałszu arytmetycznego jest niedefiniowalne w $\mathbf{AR}^{\mathrm{N}}$. Odnotujmy, że formuła $F(\overline{A}) \Leftrightarrow \neg A$ wyraża jakoś myśl, że A jest prawdziwe. Pokazuje to związek (60) z paradoksem Prawdomówcy, podobnie jak twierdzenie o niedefiniowalności prawdy dla **AR** ma związek z paradoksem Kłamcy. W ten sposób rozważania §1 i §2 tworzą pewną całość.

Bibliografia

Gödel, Kurt

1931 - „Über formal unentscheidbare Sätze der *Principia Mathematica* und vervandter Systeme I", *Monatshefte für Mathematik und Physik*, t. 38, s.173-198.

Henkin, Leon

1952 - „A problem concerning provability", *The Journal of Symbolic Logic*, t.17, s. 160.

Kripke, Saul

1976 - „An outline of the theory of truth", *The Journal of Philosophy*, v. 72, s. 690-716.

Löb, Martin

1955 - „Solution to a problem of Leon Henkin", *The Journal of Symbolic Logic*, t. 20, s. 115-118.

Łukasiewicz, Jan

1912 - „O twórczości w nauce", [w:] *Księga Pamiątkowa ku uczczeniu 250 rocznicy założenia Uniwersytetu Lwowskiego*, Nakładem Uniwersytetu Lwowskiego, Lwów, s.1-15.

1915 - „O nauce", [w:] *Poradnik dla Samouków*, t. I, wyd. przez Stanisława Michalskiego, Heflich i Michalski, Warszawa, s. XV-XXXIX .

Słupecki, Jerzy; Bryll, Grzegorz; Wybraniec-Skardowska, Urszula

1971 - „Theory of Rejected Propositions", *Studia Logica*, t. 29, s. 75-123.

Spasowski, Maciej

1983 - „Własności dualnych odpowiedników operacji konsekwencji", *Acta Universitatis Vratislaviensis*, Logika 10, s. 71-116.

Tarski, Alfred

1933 - „Pojęcie prawdy w językach nauk dedukcyjnych", Nakładem Towarzystwa Naukowego Warszawskiego, Warszawa.

Woleński, Jan

1985 - „Deontic logic and consequence operations", [w:] *Action, Logic and Social Theory*, wyd. przez Ghitę Holmstrom i Andrew Jonesa, Societas Philosophica Fennica, Helsinki 1985, s. 314-326.

1989 - „On comparison of theories by their contents", *Studia Logica*, t. 47, nr.4, s. 109-114.

1990 - „Jan Łukasiewicz i semantyczna definicja prawdy", *Ruch Filozoficzny*, t. XLVII, nr 1, s. 47-49.

1992 - „Konsekwencje odrzuceniowe i porównywanie teorii", *Zeszyty Naukowe Wyższej Szkoły Pedagogicznej im. Powstańców Śląskich w Opolu*, Matematyka XXVIII, s. 105-111.

Wójcicki, Ryszard

1973 - „Dual Counterparts of Consequence Operations", *Bulletin of The Section of Logic*, Polish Academy of Sciences, t. 2, nr 1, s. 54-57.