Niniejszy tytuł orzeka o samym sobie.

Jędrzej Kołodziejski

Większość z nas zna zapewne bajkową postać Pinokia - ożywionego za sprawą magii drewnianego chłopca, którego nos rósł wtedy i tylko wtedy, gdy ten kłamał. Czy jednak zadaliśmy sobie kiedyś pytanie, co stałoby się, gdyby Pinokio powiedział "Mój nos teraz urośnie."? Jeśli jego nos zacząłby rosnąć, to znaczyłoby, że wypowiedź była prawdziwa – ale w takim wypadku nos nie powinien rosnąć. Jeśli jednak nos by nie urósł, słowa Pinokia byłyby fałszywe – czyli jednak powinien rosnąć. Sprzeczność! Antynomia ta stanowi jedno z wielu wcieleń słynnego *paradoksu kłamcy* sformułowanego w antycznej Grecji przez Eubulidesa. Paradoks dotyczy prawdziwości następującego *zdania kłamcy*:

To zdanie jest fałszywe.

Chwila zastanowienia pokazuje, że z prawdziwości powyższego zdania wynika jego fałszywość, a z fałszywości prawdziwość – nie może być zatem ani prawdziwe ani fałszywe! Co z tym fantem zrobić? Tym, co na pierwszy rzut oka odróżnia zdanie kłamcy od "porządnych" zdań jest samoodnośność. Zamiast bowiem mówić o stołach, czy liczbach naturalnych – jak to zdania mają w zwyczaju – orzeka ono coś o sobie samym. Wydaje się, że to właśnie ta cecha stanowi źródło sprzeczności. Czy siedzącą na ławie oskarżonych samoodnośność uznamy zatem za winną? Jak się przekonamy, sprawa nie jest tak prosta, jak mogłoby się wydawać.

Samoodnośne wyrażenia są wśród nas

Zacznijmy od tego, że samozwrotność występuje w językach naturalnych – takich jak język polski – i ma się tam wcale dobrze. W polszczyźnie mamy wręcz specjalne słowo – "niniejszy" – służące do konstrukcji samoodnośnych wyrażeń. Nie wydaje się kontrowersyjne ani podejrzane stwierdzenie, że niniejszy artykuł napisany jest w języku polskim, lub że niniejsze zdanie jest szesnastym zdaniem tego artykułu (nie licząc tytułów ani wykrzykników).

W języku mamy do czynienia z rozmaitymi rodzajami definicji. Możemy definiować daną rzecz przez wyliczenie jej istotnych cech – takich jak np. posiadanie blatu i nóg przez stół. Inną formą definicji, nazywaną definicją *ostensywną*, jest bezpośrednie wskazanie egzemplarza rozważanej rzeczy – np. stołu. Zdanie:

To jest definicja ostensywna.

jest zatem definicją ostensywną definicji ostensywnej!

Obecność samoodnośnych wyrażeń w języku naturalnym nie oznacza oczywiście, że musimy je akceptować w ścisłych i precyzyjnych teoriach formalnych. Jak się jednak okazuje, czasem nawet największa ostrożność i rygor nie uchronią nas przed samoodnośnością.

Wybierzmy matematyczną strukturę, o której będziemy mówić – dla ustalenia uwagi skupimy się na liczbach naturalnych $\mathbb{N}$ z operacjami dodawania $+$ i mnożenia $\times$, jednak równie dobrze moglibyśmy mówić na przykład o zbiorach, lub zbiorach skończonych. Dla struktury takiej możemy (choć nie zrobimy tego w niniejszym artykule) bardzo precyzyjnie zdefiniować, czym są poprawnie zbudowane formuły logiczne i określić znaczenie każdej takiej formuły. Przy pomocy różnych formuł możemy wyrażać różne stwierdzenia – na przykład formuła[1] $\mathsf{Parzyste}(x) =$ "$\exists_y(y + y = x)$" mówi, że liczba x jest parzysta, zaś $\mathsf{Przemienność} =$"$\forall x(\forall y(x \times y = y \times x))$" wyraża znaną nam ze szkoły przemienność mnożenia liczb naturalnych. Formuły, które *po prostu* coś konstatują (takie jak $\mathsf{Przemienność}$), nie zaś orzekają *o jakiejś zmiennej* (jak to robi $\mathsf{Parzyste}(x)$ o x-ie, które jest prawdziwe o liczbie 2, a fałszywe o liczbie 3) nazwiemy zdaniami. Ponieważ formuły logiczne są skończonymi napisami, możemy ponumerować je wszystkie – każdemu φ przypisując unikatowy numer $\ulcorner\varphi\urcorner \in \mathbb{N}$.[2] Zauważmy, że pozwala nam to interpretować formuły mówiące o liczbach jako mówiące o kodach formuł – czyli pośrednio o samych formułach. Jak pokazał Rudolf Carnap, jeśli tylko sposób w jaki numerujemy formuły posiada kilka naturalnych własności, umożliwia to mówienie formuł *o samych sobie*:

[1] "$\exists x(...)$" to skrót dla "istnieje x takie że...", zaś "$\forall_x(...)$" – dla "każdego x ...".

[2] Numerowanie takie możemy na przykład otrzymać porządkując wszystkie formuły leksykograficznie i n-tej formule przypisując numer n. Alternatywnie, utożsamiając liczby z ich zapisami binarnymi otrzymujemy bijekcję liczb z (rozpoczynającymi się od jedynki) ciągami zer i jedynek.

Lemat Przekątniowy. *Dla każdej formuły $\varphi(x)$, skonstruować można zdanie ψ takie że:*

$$\psi \text{ wtedy o tylko wtedy gdy } \varphi(\ulcorner\psi\urcorner)$$

Zdanie ψ mówi zatem dokładnie tyle, że liczba $\ulcorner\psi\urcorner$ będąca jego własnym numerem posiada własność φ.

Aby zilustrować znaczenie lematu, rozważmy inny paradoks – nazwiemy go paradoksem Gödla – związanego z następującym *zdaniem Gödla*:

Tego zdania nie da się udowodnić.

Czy zdanie Gödla może być fałszywe? Gdyby tak było, fałszem byłoby to, co głosi – a więc zdanie Gödla miałoby dowód. To jednak niemożliwe, nie da się bowiem udowodnić zdania fałszywego. Zatem zdanie Gödla jest prawdziwe – ale to z kolei oznacza, że nie ma dowodu. Mamy zatem przykład zdania prawdziwego, ale niedowodliwego! Dzięki Lematowi Przekątniowemu, ten wyrażony w języku naturalnym paradoks daje się przekuć w ścisły dowód słynnego Twierdzenia Gödla:

Twierdzenie 1 (Gödla)**.** *Dla każdego (rozsądnego) systemu aksjomatów A, istnieje zdanie prawdziwe, ale niedające się z A udowodnić.*

Istotnie, rozważmy następującą własność liczb: “liczba x jest numerem zdania nieposiadającego dowodu w oparciu o aksjomaty A”. Okazuje się, że jeśli tylko A da się sensownie reprezentować, to istnieje formuła $\varphi_A(x)$ wyrażająca powyższą własność. Z Lematu Przekątniowego otrzymujemy wtedy zdanie ψ mówiące, że $\ulcorner\psi\urcorner$ jest numerem zdania nieposiadającego dowodu w oparciu o aksjomaty A. Zdanie ψ orzeka więc *o samym sobie*, że nie ma dowodu w oparciu o aksjomaty A. Argument analogiczny do użytego przy (nieformalnym) zdaniu Gödla pokazuje więc, że jeśli tylko z A nie da się udowodnić nic *fałszywego*, to są również pewne *prawdziwe* stwierdzenia których zeń nie dowiedziemy.

Alibi

No dobrze, widzimy zatem, że samoodnośność występuje “w przyrodzie” i że często, mimo wielkiej ostrożności, nie da się jej uniknąć nawet w systemach formalnych. Ważniejsze jednak jest pytanie, czy to faktycznie ona odpowiada za antynomie. Jak się okazuje, samozwrotność ma alibi: paradoksy podobne do paradoksu kłamcy mogą wystąpić również bez niej. Rozważmy następujące dwa zdania:

Następne zdanie jest prawdziwe.
Poprzednie zdanie jest fałszywe.

Analiza podobna do paradoksu kłamcy pokazuje, że żadne ze zdań nie może być ani prawdziwe, ani fałszywe. Z drugiej strony nietrudno dostrzec, że zdania te – choć żadne z nich nie mówi bezpośrednio o sobie, odnoszą się do siebie pośrednio, poprzez cykliczne powiązanie (Czytelnik może zechcieć zapisać cykl n zdań tworzących razem paradoks). Kolejny przykład pozbawiony jest już jakiejkolwiek autoreferencji – również pośredniej.

Tym razem rozważmy *nieskończony* ciąg zdań, z których każde mówi że wszystkie kolejne są fałszywe:

$$\begin{aligned}
\varphi_1 &= && \textit{Dla każdego } i > 1\textit{, zdanie } \varphi_i \textit{ jest fałszywe.}\\
\varphi_2 &= && \textit{Dla każdego } i > 2\textit{, zdanie } \varphi_i \textit{ jest fałszywe.}\\
&\vdots\\
\varphi_j &= && \textit{Dla każdego } i > j\textit{, zdanie } \varphi_i \textit{ jest fałszywe.}\\
&\vdots
\end{aligned}$$

Zastanówmy się, czy dla jakiegokolwiek j prawdziwe jest φ_j. Gdyby tak było, znaczyłoby to że wszystkie dalsze zdania – a więc w szczególności φ_{j+1} – są fałszywe. Skoro jednak φ_{j+1} mówi, że wszystkie następujące po nim zdania są fałszywe – to fałszywość φ_{j+1} oznacza że dla pewnego $i > j + 1$ zdanie φ_i jest prawdziwe, co przeczy φ_j. Wszystkie zdania w ciągu muszą więc być fałszywe. Ale skoro tak, to prawdą jest to, co głosi każde ze zdań – sprzeczność! Powyższy paradoks – sformułowany po raz pierwszy przez Stephena Yablo – pokazuje, że na trudności takie jak paradoks kłamcy natrafiamy nawet bez cyklicznego odnoszenia się do siebie zdań. Samoodnośność ma alibi – to nie ona jest winna sprzeczności.

Na tropie sprzeczności

Skoro samoodnośność została oczyszczona z zarzutów, pozostajemy z pytaniem: kto zatem jest winny? Pewien trop pojawił się już w artykule. Być może zobaczywszy Lemat Przekątniowy i to, jak pozwala on sformalizować paradoks Gödla, Czytelnik zadaje sobie pytanie, jak w tym kontekście prezentuje się bardziej przecież złowieszczy paradoks kłamcy. Czyżby arytmetyka była sprzeczna? Rozważmy bowiem następującą własność liczb: “liczba x jest kodem zdania *prawdziwego*”. Czy istnieje formuła wyrażająca powyższą własność? Gdyby tak było, moglibyśmy zdefiniować również paradoksalną własność “liczba x jest kodem formuły *fałszywej*”. Stosując Lemat Przekątniowy do wyrażającej ją hipotetycznej formuły $\varphi(x)$ otrzymalibyśmy zdanie ψ mówiące o sobie, że jest fałszywe. Wiemy już jednak, że takie paradoksalne zdanie prowadzi do sprzeczności. Jeśli więc wierzymy, że arytmetyka sprzeczna nie jest, to płynącym stąd wnioskiem jest twierdzenie Tarskiego o niedefiniowalności prawdy:

Twierdzenie 2 (Tarskiego)**.** *Nie istnieje formuła $\varphi(x)$ wyrażająca prawdziwość, tj. taka, że: $\varphi(x)$ wtedy i tylko wtedy gdy liczba x jest kodem formuły prawdziwej.*

Nie znaczy to, że nie da się ściśle opisać prawdy i fałszu dla formuł określonego systemu formalnego (w naszym przykładzie – arytmetyki liczb naturalnych). Opis taki wymaga jednak wyjścia poza sam ten system i przyjęcia *zewnętrznej perspektywy.*

Niczym w dobrym kryminale, godnym Herkulesa Poirot i Sherlocka Holmesa, ostateczny werdykt jest więc zupełnie inny, niż wydawało się na początku. To nie samoodnośność, ale lekkomyślnie nadużywane pojęcie prawdy, choć z pozoru tak niewinne, odpowiada za sprzeczności! Dodajmy na koniec, że Twierdzenie Tarskiego mówi tylko o prawdziwości na gruncie sformalizowanych, matematycznych systemów – takich jak arytmetyka liczb naturalnych. Pojęcie prawdziwości w językach naturalnych – takich jak język polski – znajduje się poza jurysdykcją formalnych wnioskowań i tu, otoczona hordami paradoksów, sprzeczność może pozostać bezkarna.