Funkcje rekurencyjne (12) (JiNoI III)

Jerzy Pogonowski

Zakład Logiki Stosowanej UAM
www.logic.amu.edu.pl
pogon@amu.edu.pl

(wykład dodatkowy)

Plan na dziś

Plan na dziś:

- Minijęzyk Smullyana;
- Twierdzenie Gödla dla minijęzyka;
- Dodatek: Maszyny logiczne Smullyana.

W maju zostanie przedstawiony dowód Twierdzenia Gödla o niezupełności Arytmetyki Peana.
Dla oswojenia się z rozumowaniami przekątniowymi, które odgrywają w tym dowodzie istotną rolę pobawimy się dzisiaj pewnym małym systemem logicznym, skonstruowanym przez Raymonda Smullyana (zob. rozdział 15 w *Szatan, Cantor i nieskończoność oraz inne łamigłówki.*).

Książki z zagadkami logicznymi Raymonda Smullyana

- *Jaki jest tytuł tej książki? Tajemnica Drakuli, zabawy i łamigłówki logiczne.* Warszawa 1993. Przełożył: Bohdan Chwedeńczuk. Trzy wydania polskie.
- *Dama czy tygrys oraz inne zagadki logiczne.* Warszawa 1995, 2004. Przełożył: Bohdan Chwedeńczuk.
- *Szatan, Cantor i nieskończoność oraz inne łamigłówki.* Warszawa 1998. Przełożyli z angielskiego: Anna i Krzysztof Wójtowicz.
- *Przedrzeźniać Przedrzeźniacza. Oraz Inne Zagadki Logiczne Łącznie z Zadziwiającą Przygodą w Krainie Logiki Kombinatorycznej.* Warszawa 2007. Przekład z języka angielskiego: Jerzy Pogonowski.
- *Forever Undecided. A Puzzle Guide to Gödel.* Oxford University Press, 1988. Z angielskiego przełożył Jerzy Pogonowski. Ukaże się w 2007 jako: *Na zawsze nierozstrzygnięte. Zagadkowy Przewodnik po Twierdzeniach Gödla.*

Minijęzyk Smullyana

Rozważmy język o czterech symbolach: $\clubsuit, \spadesuit, \diamondsuit, \heartsuit$.
Wyrażeniem tego języka jest dowolny skończony ciąg tych symboli.

Zbudujemy miniaturowy system S, w którym można dowodzić pewnych wyrażeń tego języka.
Nie będzie przy tym istotne, na czym polega owa dowodliwość.

Interesować nas będzie jedynie jej związek z określoną dla tego języka prawdziwością jego wyrażeń.
Wyrażenia, które nie są prawdziwe w S nazwiemy fałszywymi w S.

Nie będzie istotne, czym jest prawdziwość. Ważne będą jedynie wzajemne związki dowodliwości i prawdziwości.

Minijęzyk Smullyana

Przypiszemy wyrażeniom tego języka następującą interpretację:

- $\spadesuit X$ — stwierdza, że wyrażenie X jest dowodliwe w S;
- $\clubsuit X$ — stwierdza, że wyrażenie XX jest dowodliwe w S;
- $\heartsuit X$ — stwierdza, że wyrażenie X nie jest dowodliwe w S;
- $\diamondsuit X$ — stwierdza, że wyrażenie XX nie jest dowodliwe w S.

Powiemy, że:

- $\spadesuit X$ jest prawdziwe w S, gdy X jest dowodliwe w S;
- $\clubsuit X$ jest prawdziwe w S, gdy XX jest dowodliwe w S;
- $\heartsuit X$ jest prawdziwe w S, gdy X nie jest dowodliwe w S;
- $\diamondsuit X$ jest prawdziwe w S, gdy XX nie jest dowodliwe w S.

Poprawność systemu S

Widać, że ważną cechą systemu S jest jego samozwrotność — można w nim udowodnić różne zdania, stwierdzające, co w tym systemie jest, a co nie jest dowodliwe.
Jedyne założenie, które czynimy o systemie S to założenie jego poprawności: wszystkie zdania dowodliwe w S są prawdziwe w S.

Konsekwencjami tego założenia są:

- W_1 Jeśli $\spadesuit X$ jest dowodliwe w S, to X jest dowodliwe w S.
- W_2 Jeśli $\heartsuit X$ jest dowodliwe w S, to X nie jest dowodliwe w S.
- W_3 Jeśli $\clubsuit X$ jest dowodliwe w S, to XX jest dowodliwe w S.
- W_4 Jeśli $\diamondsuit X$ jest dowodliwe w S, to XX nie jest dowodliwe w S.

Twierdzenie Gödla dla S

Istnieje wyrażenie prawdziwe w S, które nie jest dowodliwe w S.

Dowód. Takim wyrażeniem jest $\diamondsuit\diamondsuit$.
Stwierdza ono, że podwojenie wyrażenia $\diamondsuit$ nie jest dowodliwe.
Podwojeniem $\diamondsuit$ jest $\diamondsuit\diamondsuit$.
Zatem $\diamondsuit\diamondsuit$ jest prawdziwe w S wtedy i tylko wtedy, gdy nie jest dowodliwe w S.
Oznacza to, że $\diamondsuit\diamondsuit$ jest:

- prawdziwe w S i niedowodliwe w S; albo
- fałszywe w S i dowodliwe w S.

Drugi człon tej alternatywy jest wykluczony ze względu na poprawność S: tylko zdania prawdziwe są dowodliwe.
Zatem zachodzi pierwszy człon tej alternatywy.

Uwagi o interpretacji zamierzonej

Uwaga. Jest chyba dla wszystkich oczywiste, że nasza zamierzona interpretacja minijęzyka Smullyana jest jedną z wielu możliwych interpretacji.
To, co jest istotne, to:

- W_0 możliwość interpretowania ciągów symboli;
- warunki W_1—W_4.

Można więc myśleć o innych jeszcze interpretacjach, spełniających te warunki. Dla przykładu, symbol $\spadesuit$ możemy interpretować jako:

- drukowalny (przez jakąś maszynę);
- poznawalny (przez jakiś podmiot);
- akceptowalny (np. przez Watykan).

Anything goes, jeśli tylko spełnione są warunki W_0—W_4.

Operacja sprzężenia

- Sprzężeniem $\spadesuit X$ jest $\heartsuit X$.
- Sprzężeniem $\heartsuit X$ jest $\spadesuit X$.
- Sprzężeniem $\clubsuit X$ jest $\diamondsuit X$.
- Sprzężeniem $\diamondsuit X$ jest $\clubsuit X$.

Sprzężenie X oznaczamy przez $\overline{X}$. Dla dowolnej pary wyrażeń sprzężonych, jedno z nich jest prawdziwe w S, a drugie jest fałszywe w S.
Wyrażenie nazywamy obalalnym w S, gdy jego sprzężenie jest dowodliwe w S. Zatem:

- $\heartsuit X$ obalane w S wtedy i tylko wtedy, gdy $\spadesuit X$ dowodliwe w S.
- $\spadesuit X$ obalane w S wtedy i tylko wtedy, gdy $\heartsuit X$ dowodliwe w S.
- $\clubsuit X$ obalane w S wtedy i tylko wtedy, gdy $\diamondsuit X$ dowodliwe w S.
- $\diamondsuit X$ obalane w S wtedy i tylko wtedy, gdy $\clubsuit X$ dowodliwe w S.

Nierozstrzygalność S

Wyrażenie, które nie jest ani dowodliwe w S, ani obalalne w S nazwiemy nierozstrzygalnym w S.

Pokazaliśmy, że $\diamondsuit\diamondsuit$ jest prawdziwe i niedowodliwe w S.

Z prawdziwości $\diamondsuit\diamondsuit$ wynika, że wyrażenie z nim sprzężone, czyli $\clubsuit\diamondsuit$ jest fałszywe w S.

Stąd, na mocy poprawności S, wyrażenie $\clubsuit\diamondsuit$ jest także niedowodliwe w S.

Stąd i z definicji $\clubsuit$, mamy, iż $\diamondsuit\diamondsuit$ jest niedowodliwe w S.

Oznacza to, iż wyrażenie $\diamondsuit\diamondsuit$ jest nierozstrzygalne w S.

Nierozstrzygalność S

Uwaga. Można przeprowadzić powyższą argumentację wcale nie odwołując się do pojęcia prawdy.

Istotnie, nierozstrzygalność wyrażenia $\diamondsuit\diamondsuit$ udowodnić można bezpośrednio z warunków W_1—W_4.

(1) Przypuśćmy, że $\diamondsuit\diamondsuit$ jest dowodliwe.

Podstawiając za X w W_4 wyrażenie $\diamondsuit$ otrzymujemy, że podwojenie $\diamondsuit$ jest niedowodliwe, co znaczy, że $\diamondsuit\diamondsuit$ jest niedowodliwe. Sprzeczność.

Zatem $\diamondsuit\diamondsuit$ nie jest dowodliwe.

(2) Gdyby dowodliwe było sprzężenie wyrażenia $\diamondsuit\diamondsuit$, czyli wyrażenie $\clubsuit\diamondsuit$, to na mocy warunku W_3 (podstawiamy $\diamondsuit$ za X) wyrażenie $\diamondsuit\diamondsuit$ byłoby dowodliwe.

Pokazaliśmy już jednak, że $\diamondsuit\diamondsuit$ nie jest dowodliwe.

Wynika stąd, że $\clubsuit\diamondsuit$ również nie jest dowodliwe.

Ostatecznie, wyrażenie $\diamondsuit\diamondsuit$ nie jest rozstrzygalne w S.

Niektóre inne ciekawe wyrażenia S

- $\heartsuit\diamondsuit$ stwierdza o sobie, że jest obalalne.
- $\diamondsuit\clubsuit$ stwierdza o sobie, że nie jest obalalne.
- $\clubsuit\clubsuit$ stwierdza o sobie, że jest dowodliwe.

- Dla dowolnego wyrażenia E istnieje wyrażenie X, które stwierdza, że EX jest dowodliwe (tj. X jest prawdziwe wtedy i tylko wtedy, gdy EX jest dowodliwe). Dowód. $X = \heartsuit E \heartsuit$.
- Dla dowolnego wyrażenia E istnieje wyrażenie X, które stwierdza, że EX nie jest dowodliwe. Dowód. $X = \diamondsuit E \diamondsuit$.
- Dla dowolnego wyrażenia E istnieje wyrażenie X, które stwierdza, że $E\overline{X}$ jest dowodliwe. Ćwiczenie.

  Dla dowolnego wyrażenia E istnieje wyrażenie X, które stwierdza, że $E\overline{X}$ nie jest dowodliwe. Ćwiczenie.

Niektóre inne ciekawe wyrażenia S

- Istnieją wyrażenia X i Y takie, że:
  - X stwierdza, że Y jest dowodliwe.
  - Y stwierdza, że X nie jest dowodliwe.

  Co najmniej jedno z wyrażeń X, Y musi być prawdziwe, ale nie ma metody ustalenia, które.
- Istnieją wyrażenia X i Y takie, że:
  - X stwierdza, że Y jest obalalne.
  - Y stwierdza, że X nie jest obalalne.

  Co najmniej jedno z wyrażeń X, Y musi być fałszywe, ale nieobalalne.
- Istnieją wyrażenia X i Y takie, że:
  - X stwierdza, że Y jest dowodliwe.
  - Y stwierdza, że X jest obalalne. Jedno z nich jest prawdziwe, ale niedowodliwe, a drugie fałszywe, ale nieobalalne.

Niektóre inne ciekawe wyrażenia S

- Istnieją wyrażenia X i Y takie, że:
  - X stwierdza, że Y nie jest dowodliwe.
  - Y stwierdza, że X nie jest obalalne.

- Istnieją wyrażenia X, Y i Z takie, że:
  - X stwierdza, że Y jest obalalne.
  - Y stwierdza, że Z jest nieobalalne.
  - Z stwierdza, że X jest dowodliwe.

  Nadto, zachodzi jedna z trzech możliwości:
  - X jest prawdziwe, ale nie jest dowodliwe.
  - Y jest fałszywe, ale nie jest obalalne.
  - Z jest fałszywe, ale nie jest obalalne.

Systemy regularne

Powiemy, że system S, spełniający warunki W_1—W_4 jest regularny, jeśli spełnione są także warunki:

- Jeśli X jest dowodliwe w S, to $\spadesuit X$ jest dowodliwe w S.
- Jeśli XX jest dowodliwe w S, to $\clubsuit X$ jest dowodliwe w S.

Z tej definicji oraz z warunków W_1 i W_3 wynika, że jeśli S jest systemem regularnym, to:

- $\spadesuit X$ jest dowodliwe wtedy i tylko wtedy, gdy X jest dowodliwe.
- $\clubsuit X$ jest dowodliwe wtedy i tylko wtedy, gdy XX jest dowodliwe.

Wyrażenia pozytywne to wyrażenia postaci $\spadesuit X$ lub $\clubsuit X$.
Wyrażenia negatywne to wyrażenia postaci $\heartsuit X$ lub $\diamondsuit X$.

Systemy regularne

- System S jest regularny wtedy i tylko wtedy, gdy wszystkie prawdziwe wyrażenia pozytywne S są dowodliwe w S.
- Jeśli system S jest regularny, to każde fałszywe wyrażenie negatywne jest obalalne w S.
- Jeśli S jest regularny, to dla dowolnego wyrażenia X oraz dowolnego ciągu E symboli $\spadesuit$: wyrażenie EX jest dowodliwe wtedy i tylko wtedy, gdy X jest dowodliwe.
- W systemie regularnym $\clubsuit X$ jest dowodliwe wtedy i tylko wtedy, gdy $\spadesuit XX$ jest dowodliwe.
- Jeśli E jest dowolnym ciągiem symboli $\spadesuit$, to wyrażenie $\diamondsuit E \diamondsuit$ jest prawdziwe i niedowodliwe we wszystkich systemach regularnych. Zatem istnieje nieskończenie wiele wyrażeń prawdziwych i niedowodliwych we wszystkich systemach regularnych.

Koniec 1

To koniec części łatwej, przeznaczonej dla studentek III roku Językoznawstwa i Nauk o Informacji UAM.

W drugiej części odwołujemy się do pewnych własności systemów logiki modalnej.

Część ta jest fragmentem prezentacji **Szczęściarze epistemiczni**, przygotowywanej na konferencję:
Zastosowania Logiki w Filozofii i Podstawach Matematyki, XII.

Maszyny logiczne Smullyana

Smullyan skonstruował cały szereg maszyn logicznych, które drukują zdania „mówiące" coś o nich samych.
Maszyny: Craiga, Fergussona i McCullocha, przedstawione w *Jaki jest tytuł tej książki?* oraz *Dama czy tygrys?* są już znane polskiemu czytelnikowi.
Tu przedstawimy pewną maszynę Smullyana, opisaną w *Forever Undecided.*

Dla pełnego zrozumienia jej działania potrzebna jest znajomość wybranych logik modalnych: logiki epistemicznej oraz logiki dowodliwości (logiki Gödla-Löba).
Zakładamy u audytorium znajomość tego materiału.

Maszyny logiczne Smullyana

Malcolm Fergusson, gdy usłyszał o twierdzeniach Gödla i Löba, z miejsca zabrał się za konstrukcję maszyny, którą z zachwytem pokazał swoim przyjaciołom.
Ku ich zadowoleniu udowodnił, że maszyna jest niesprzeczną i stabilną maszyną typu G, a szczególne upodobanie znalazł w demonstracji, że maszyna, chociaż niesprzeczna, nigdy nie może dowieść własnej niesprzeczności!
Maszyna ilustruje w niezwykle prosty i pouczający sposób podstawowe idee zawarte w Pierwszym oraz Drugim Twierdzeniu Gödla jak również w Twierdzeniu Löba.

Niżej podajemy opis działania maszyny Fergussona oraz pewne ważne fakty jej dotyczące.
Opis pochodzi z rozdziału 26 *Forever Undecided*. W rozdziale tym znajdujemy też opis dwóch innych maszyn, który tu pominiemy.

Maszyna drukuje różnorakie zdania zbudowane z siedemnastu symboli. Pierwsze siedem z tych symboli to następujące:

P	$\perp$	$\rightarrow$	(	)	d	,
1	2	3	4	5	6	7

Pod każdym z tych symboli podpisano jego *numer Gödlowski*.

Pozostałe dziesięć symboli to znane cyfry 1, 2, 3, 4, 5, 6, 7, 8, 9, 0. Tym cyfrom przyporządkowujemy numery Gödlowskie w następujący sposób. Numerem Gödlowskim cyfry 1 jest 89 (8 po której następuje jedna 9); numerem Gödlowskim cyfry 2 jest 899 (8 po której następują dwie 9); i tak dalej, aż do cyfry 0, której numerem Gödlowskim jest 89999999999 (8 po której następuje dziesięć 9).

Tak więc, każdy z siedemnastu symboli uzyskuje numer Gödlowski.

Dla danego wyrażenia złożonego, odnajdujemy jego numer Gödlowski przez zastąpienie każdego symbolu jego numerem Gödlowskim — dla przykładu, numerem Gödlowskim wyrażenia $(P \perp \rightarrow \perp)$ jest 412325. Inny przykład: numerem Gödlowskim $P35$ jest 18999899999.

Dla dowolnego wyrażenia E, przez $\overline{E}$ rozumiemy numer Gödlowski E (zapisany jako ciąg cyfr 1, 2, …, 0).

Nie każda liczba jest numerem Gödlowskim jakiegoś wyrażenia (na przykład, 88 nie jest numerem Gödlowskim żadnego wyrażenia).

Jeśli *n jest* numerem Gödlowskim jakiegoś wyrażenia, to będziemy czasem odwoływać się do tego wyrażenia jako do n-tego wyrażenia. (Dla przykładu, Pd jest szesnastym wyrażeniem, $\perp$ jest drugim wyrażeniem.)

Maszyna jest *samoodnosząca się* (do siebie) w tym sensie, że wyrażenia drukowane przez maszynę stwierdzają, co maszyna może, a czego nie może wydrukować. Wyrażenie nazywamy *drukowalnym*, jeśli maszyna może je wydrukować.

Symbol „P" oznacza „drukowalne" i dla dowolnego wyrażenia E zbudowanego z podanych siedemnastu symboli, jeśli chcemy zapisać zdanie stwierdzające, że E jest drukowalne, to piszemy nie PE, lecz $P\overline{E}$ (tj., P po którym następuje numer Gödlowski E).
Dla przykładu, zdaniem stwierdzającym, że $(P \perp \rightarrow \perp)$ jest drukowalne jest $P\overline{(P \perp \rightarrow \perp)}$ — tj. $P412325$.

Dla dowolnych wyrażeń X oraz Y, Fergusson zdefiniował *diagonalizację X względem Y* jako wyrażenie $(X(\overline{X}, \overline{Y}) \rightarrow Y)$.
Symbol „d" jest skrótem dla „diagonalizacja" — i dla dowolnych wyrażeń X oraz Y, wyrażenie $Pd(\overline{X}, \overline{Y})$ jest zdaniem stwierdzającym, że diagonalizacja X względem Y jest drukowalna.

Zdefiniujemy teraz, co to znaczy, że wyrażenie jest *zdaniem (maszynowym)* i co to znaczy, że zdanie jest *prawdziwe*.

- (1) $\perp$ jest zdaniem i $\perp$ jest fałszywe.
- (2) Dla dowolnego wyrażenia X, wyrażenie $P\overline{X}$ jest zdaniem i jest ono prawdziwe wtedy i tylko wtedy, gdy wyrażenie X jest drukowalne.
- (3) Dla dowolnych wyrażeń X oraz Y, wyrażenie $Pd(\overline{X}, \overline{Y})$ jest zdaniem i jest ono prawdziwe wtedy i tylko wtedy, gdy wyrażenie $(X(\overline{X}, \overline{Y}) \to Y)$ — które jest diagonalizacją X względem Y — jest drukowalne.
- (4) Dla dowolnych zdań X oraz Y, wyrażenie $(X \to Y)$ jest zdaniem i jest ono prawdziwe wtedy i tylko wtedy, gdy albo X nie jest prawdziwe, albo Y jest prawdziwe.

Rozumie się, że żadne wyrażenie nie jest zdaniem (maszynowym), jeśli nie zostało otrzymane zgodnie z powyższymi regułami. Spójniki logiczne $\neg, \wedge, \vee, \equiv$ są definiowane z $\to$ oraz $\perp$ w znany sposób.

Podamy teraz reguły ustalające, co maszyna może wydrukować. Maszyna jest zaprogramowana do kolejnego drukowania nieskończonej listy zdań. Pewne zdania, nazywane *aksjomatami* mogą zostać wydrukowane na każdym etapie tego procesu. Wśród aksjomatów są wszystkie tautologie. (tak więc, dla dowolnej tautologii X, maszyna może wydrukować X kiedy tylko chce, niezależnie od tego, co dotąd wydrukowała lub czego nie wydrukowała w poprzednich etapach.)

Dalej, maszyna jest zaprogramowana tak, że dla dowolnych zdań X oraz Y, jeśli na pewnym etapie maszyna wydrukowała już X oraz $X \rightarrow Y$, to może wydrukować Y. Tak więc, maszyna jest typu 1 (w tym sensie, że zbiór zdań drukowalnych jest typu 1).

Ponieważ jest prawdą, że jeśli X oraz $X \rightarrow Y$ są oba drukowalne, to Y też jest drukowalne, to zdanie $(P\overline{X} \wedge P\overline{(X \rightarrow Y)}) \rightarrow P\overline{Y}$ jest prawdziwe; lub, co na jedno wychodzi, zdanie $P\overline{(X \rightarrow Y)} \rightarrow (P\overline{X} \rightarrow P\overline{Y})$ jest prawdziwe. Maszyna „wie" zatem o prawdziwości wszystkich zdań postaci $P\overline{(X \rightarrow Y)} \rightarrow (P\overline{X} \rightarrow P\overline{Y})$ i przyjmuje je jako aksjomaty. Tak więc, maszyna jest typu 2.

Następnie, jeśli maszyna kiedykolwiek wydrukuje zdanie X, to „wie” ona, że wydrukowała X i prędzej czy później wydrukuje prawdziwe zdanie $P\overline{X}$. (Zdanie $P\overline{X}$ *jest* prawdziwe, ponieważ X zostało wydrukowane.) A więc maszyna jest normalna, a stąd jest typu 3.
Ponieważ maszyna jest normalna, więc dla dowolnego zdania X, zdanie $P\overline{X} \to P\overline{P\overline{X}}$ jest prawdziwe. Czyli maszyna jest początkowo „świadoma” prawdziwości wszystkich takich zdań oraz przyjmuje je jako aksjomaty. Zatem maszyna jest typu 4.

Jest jeszcze jedna rzecz, którą maszyna potrafi robić, a jest to rzecz dość istotna. Dla dowolnych wyrażeń X oraz Y, zdanie $Pd(\overline{X}, \overline{Y})$ jest prawdziwe wtedy i tylko wtedy, gdy $(X(\overline{X}, \overline{Y}) \to Y)$ jest drukowalne, co z kolei zachodzi wtedy i tylko wtedy, gdy zdanie $P\overline{(X(\overline{X}, \overline{Y}) \to Y)}$ jest prawdziwe. Zatem następujące zdanie jest prawdziwe: $Pd(\overline{X}, \overline{Y}) \equiv P\overline{(X(\overline{X}, \overline{Y}) \to Y)}$.
Maszyna „wie” o prawdziwości wszystkich takich zdań i przyjmuje je jako aksjomaty. Te aksjomaty nazywane są *aksjomatami przekątniowymi.*

Aksjomaty i reguły maszyny

Aksjomaty:

- *Grupa 1.* Wszystkie tautologie.
- *Grupa 2.* Wszystkie zdania postaci $P\overline{(X \to Y)} \to (P\overline{X} \to P\overline{Y})$.
- *Grupa 3.* Wszystkie zdania postaci $P\overline{X} \to P\overline{P\overline{X}}$.
- *Grupa 4 (aksjomaty przekątniowe).* Wszystkie zdania postaci $Pd(\overline{X}, \overline{Y}) \equiv P(\overline{X(\overline{X}, \overline{Y}) \to Y})$, gdzie X oraz Y są dowolnymi wyrażeniami (niekoniecznie zdaniami).

Reguły operowania:

- (1) Aksjomaty mogą zostać wydrukowane na każdym etapie.
- (2) Dla dowolnych już wydrukowanych zdań X oraz $(X \to Y)$, maszyna może wydrukować Y.
- (3) Dla dowolnego wydrukowanego już zdania X, maszyna może wydrukować $P\overline{X}$.

Rozumie się, że jedynym sposobem wydrukowania przez maszynę jakiegoś zdania X na pewnym etapie jest zastosowanie się do powyższych reguł. Zatem, X jest drukowalne na danym etapie *tylko* wtedy, gdy zachodzi jeden z następujących trzech warunków: (1) X jest aksjomatem; (2) istnieje zdanie Y takie, że Y oraz $(Y \rightarrow X)$ zostały już wydrukowane na etapie wcześniejszym; (3) istnieje zdanie Y takie, że X jest zdaniem $P\overline{Y}$ oraz Y zostało już wydrukowane na etapie wcześniejszym.

Uwagi. Dla każdego zdania X, niech $\boldsymbol{B}X$ będzie zdaniem $P\overline{X}$. Symbol „$\boldsymbol{B}$" *nie* należy do języka maszyny; używamy go do *mówienia o* maszynie. Używamy „$\boldsymbol{B}$" jako odpowiadającego operacji, która przyporządkowuje każdemu zdaniu X zdanie $P\overline{X}$.
Gdy mówimy, że maszyna jest typu 4, rozumiemy przez to, że jest ona typu 4 ze względu na tę operację $\boldsymbol{B}$. W istocie, bez aksjomatów przekątniowych, system aksjomatyczny tej maszyny jest systemem modalnym K_4.
Zobaczymy wkrótce, że dodanie aksjomatów przekątniowych daje nam pełną moc systemu modalnego G.

Dowodliwość. Zdefiniowaliśmy dla każdego zdania maszyny co to znaczy, że zdanie to jest *prawdziwe*, a więc każde zdanie maszyny wyraża określone zdanie, które może być prawdziwe lub może być fałszywe.

Uwaga. Dotąd *proposition* oddawaliśmy zawsze jako *zdanie*. Teraz mamy:

- *zdania (maszyny)* (w oryginale *sentences*) — zdania języka przedmiotowego,
- oraz zdania metajęzyka (w oryginale *propositions*), tj. języka, w którym *mówimy o* maszynie, jej zdaniach (maszynowych), itp.

W przypadkach, gdy mogłoby to prowadzić do nieporozumień, w dalszym ciągu będziemy dodawać określenie *maszynowe*, gdy mowa będzie o zdaniach drukowanych przez maszynę.

Powiemy, że maszyna *dowodzi* danego zdania, gdy drukuje ona zdanie maszynowe, które wyraża to dane zdanie. Dla przykładu, zdanie maszynowe $\neg P2$ wyraża zdanie stwierdzające, że maszyna jest niesprzeczna (ponieważ 2 jest numerem Gödlowskim $\perp$), a więc jeśli maszyna wydrukowała $\neg P2$, to udowodniła swoją własną niesprzeczność. Gdyby maszyna wydrukowała $P2$, to udowodniłaby swoją własną *sprzeczność*.

Powiemy, że maszyna jest *ścisła*, jeśli wszystkie zdania dowodliwe przez maszynę są prawdziwe.
Powiemy, że maszyna jest *niesprzeczna*, jeśli nie może ona dowieść $\perp$, oraz że jest *stabilna*, jeśli dla każdego zdania (maszynowego) X, jeśli $P\overline{X}$ jest drukowalne, to drukowalne jest też X.

Zwrotność. Przechodzimy teraz do dowodu, że maszyna jest Gödlowska, a faktycznie, zwrotna.

Ważne własności maszyny Fergussona

- (1) Znajdziemy zdanie G takie, że zdanie $G \equiv \neg P\overline{G}$ — tj. zdanie $G \equiv (P\overline{G} \rightarrow \perp)$ — jest drukowalne.

- (2) Pokażemy, że dla dowolnego zdania Y istnieje zdanie X takie, że zdanie $X \equiv (P\overline{X} \rightarrow Y)$ jest drukowalne.

Uwaga. Problem 1 jest szczególnym przypadkiem problemu 2, a więc najpierw rozwiążemy problem 2.

Przypomnijmy, że:

- warunek wspomniany w problemie (2) nazywamy zwrotnością;
- systemem typu G nazywamy system modalny typu 4, w którym dowodliwe są wszystkie zdania postaci $\boldsymbol{B}(\boldsymbol{B}p \rightarrow q) \rightarrow \boldsymbol{B}p$.

Niech Y będzie dowolnym zdaniem. Dla dowolnego wyrażenia Z, zdanie $Pd(\overline{Z}, \overline{Y}) \equiv P\overline{(Z(\overline{Z}, \overline{Y}) \to Y)}$ jest drukowalne (ponieważ jest jednym z aksjomatów przekątniowych).

Weźmiemy za Z wyrażenie Pd i otrzymujemy wtedy, że $Pd(\overline{Pd}, \overline{Y}) \equiv P\overline{(Pd(\overline{Pd}, \overline{Y}) \to Y)}$ jest drukowalne.

Ponieważ maszyna jest typu 1, więc wynika stąd, że następujące zdanie jest drukowalne:

$$(Pd(\overline{Pd}, \overline{Y}) \to Y) \equiv (P\overline{(Pd(\overline{Pd}, \overline{Y}) \to Y)} \to Y)$$

Tak więc, zdanie $X \equiv (P\overline{X} \to Y)$ jest drukowalne, gdzie X jest zdaniem $(Pd(\overline{Pd}, \overline{Y}) \to Y)$.

Problem 1 jest szczególnym przypadkiem problemu 2, gdy za Y weźmiemy $\perp$. Tak więc, zdaniem Gödla G dla tej maszyny jest $Pd(\overline{Pd}, \overline{\perp}) \rightarrow \perp$ — tj., zdanie $(Pd(16, 2) \rightarrow \perp)$.

Co stwierdza zdanie $Pd(16, 2)$?

Mówi ono, że diagonalizacja szesnastego wyrażenia względem drugiego wyrażenia jest drukowalna. Wyrażeniem szesnastym jest Pd, a wyrażeniem drugim jest $\perp$, a więc $Pd(16, 2)$ mówi, że diagonalizacja Pd względem $\perp$ jest drukowalna, ale ta diagonalizacja to zdanie $(Pd(16, 2) \rightarrow \perp)$ — tj. właśnie samo zdanie G!

A więc $Pd(16, 2)$ mówi, że G *jest* drukowalne, a stąd $(Pd(16, 2) \rightarrow \perp)$ — które jest zdaniem G — mówi, że G *nie jest* drukowalne (lub, co na jedno wychodzi, że drukowalność G implikuje fałsz logiczny). Tak więc, G mówi, że G nie jest drukowalne; G jest *prawdziwe* wtedy i tylko wtedy, gdy G nie jest drukowalne.

Zatem G stwierdza swoją własną niedrukowalność. Oto, w miniaturce, pomysłowa idea Gödla otrzymywania samoodniesienia.

Zdanie $G \equiv \neg P\overline{G}$ — tj. zdanie $G \equiv (P\overline{G} \rightarrow \perp)$ — jest nie tylko prawdziwe, ale także drukowalne (jest ono jednym z aksjomatów przekątniowych).

Ponieważ maszyna jest normalna i jest typu 1, wynika stąd na mocy Pierwszego Twierdzenia Gödla o Niezupełności, że jeśli maszyna jest niesprzeczna, to G nie jest drukowalne, a jeśli maszyna jest dodatkowo stabilna, to również $\neg G$ nie jest drukowalne.

A więc, *jeśli* maszyna jest jednocześnie niesprzeczna i stabilna, to zdanie G jest nierozstrzygalne w systemie zdań, które maszyna może wydrukować.

Maszyna jest faktycznie typu 4, a ponieważ jest Gödlowska — zdanie $G \equiv \neg P\overline{G}$ jest drukowalne — więc z Drugiego Twierdzenia Gödla o Niedowodliwości Niesprzeczności wynika, że jeśli maszyna jest niesprzeczna, to nie może ona dowieść swojej własnej niesprzeczności — tj. nie może wydrukować zdania $\neg P2$.

Nadto, jeśli maszyna jest niesprzeczna, to zdanie $\neg P2$ jest *prawdziwe*, a stąd jest innym przykładem zdania prawdziwego, którego maszyna nie może wydrukować.

Co więcej, maszyna jest zwrotna (problem 2), a będąc typu 4, musi być Löbowska (na mocy Twierdzenia Löba), a więc dla dowolnego zdania X, jeśli $P\overline{X} \rightarrow X$ jest drukowalne, to drukowalne jest X. Ponieważ każdy zwrotny Löbowski system typu 4 jest typu G, więc wynika stąd, że maszyna jest typu G.

Czy Maszyna Fergussona jest niesprzeczna?

Poprawność, ścisłość i niesprzeczność Maszyny Fergussona.

Pokazaliśmy, że *jeśli* maszyna Fergussona jest niesprzeczna, to nie może udowodnić swojej własnej niesprzeczności.

Ale skąd wiemy, czy maszyna jest, czy nie jest niesprzeczna?

Udowodnimy teraz, że maszyna jest nie tylko niesprzeczna, ale że jest też całkowicie ścisła — tj., że każde zdanie wydrukowane przez maszynę jest prawdziwe.

Pokazaliśmy już, że wszystkie *aksjomaty* maszyny są prawdziwe, ale prześledźmy uważnie to rozumowanie.

Aksjomaty Grupy 1 są wszystkie tautologiami, a stąd są z pewnością prawdziwe.

Jeśli chodzi o aksjomaty Grupy 2, to powiedzieć, że $P\overline{(X \to Y)} \to (P\overline{X} \to \overline{Y})$ jest prawdziwe to tyle, co powiedzieć, że jeśli oba $P\overline{(X \to Y)}$ oraz $P\overline{X}$ są prawdziwe, to takie jest też $P\overline{Y}$, czyli to samo, co powiedzieć, że jeśli $(X \to Y)$ oraz X są oba drukowalne, to takie jest też Y.
A tak oczywiście jest, na mocy Operacji 2.

Tak więc, aksjomaty Grupy 2 są wszystkie prawdziwe.

Jeśli chodzi o aksjomaty Grupy 3, powiedzieć, że $P\overline{X} \to P\overline{P\overline{X}}$ jest prawdziwe, to tyle, co powiedzieć, że jeśli $P\overline{X}$ jest prawdziwe, to takie jest też $P\overline{P\overline{X}}$.

To z kolei jest tym samym, co powiedzenie, że jeśli X jest drukowalne, to takie jest też $P\overline{X}$ — a tak jest rzeczywiście, na mocy Operacji 3.

Jeśli chodzi o aksjomaty przekątniowe, to $Pd(\overline{X}, \overline{Y})$ jest prawdziwe wtedy i tylko wtedy, gdy $(X(\overline{X}, \overline{Y}) \to Y)$ jest drukowalne, a tak jest wtedy i tylko wtedy, gdy $P\overline{(X(\overline{X}, \overline{Y}) \to Y)}$ jest prawdziwe.

Zatem $Pd(\overline{X}, \overline{Y}) \equiv P\overline{(X(\overline{X}, \overline{Y}) \to Y)}$ jest prawdziwe.

Wiemy teraz, że wszystkie aksjomaty maszyny są prawdziwe, ale musimy pokazać, że wszystkie zdania *drukowalne* są prawdziwe.

Przypomnijmy, że maszyna drukuje zdania na pewnych *etapach*. Chcemy teraz ustanowić następujący lemat, twierdzenie i wniosek:

- **Lemat.** Jeśli X jest zdaniem wydrukowanym na pewnym etapie i wszystkie zdania wydrukowane do tego etapu są prawdziwe, to X jest prawdziwe.
- **Twierdzenie.** Każde zdanie wydrukowane przez maszynę jest prawdziwe.
- **Wniosek.** Maszyna jest jednocześnie niesprzeczna i stabilna.

Dowody.
Najpierw udowodnimy lemat. Załóżmy, że wszystkie dotąd wydrukowane zdania są prawdziwe; mamy pokazać, że X jest prawdziwe.

Przypadek 1. X jest aksjomatem. Wtedy X jest prawdziwe (jak już udowodniliśmy).
Przypadek 2. Istnieje zdanie Y takie, że Y oraz $(Y \rightarrow X)$ zostały już wydrukowane. Wtedy z przyjętego założenia Y oraz $(Y \rightarrow X)$ są oba prawdziwe, a więc X jest prawdziwe.
Przypadek 3. X jest postaci $P\overline{Y}$, gdzie Y jest zdaniem, które już zostało wydrukowane. Ponieważ Y zostało wydrukowane, więc $P\overline{Y}$ jest prawdziwe — tj. X jest prawdziwe.

To kończy dowód lematu.

Dowód Twierdzenia.

Maszyna jest zaprogramowana tak, aby wydrukować wszystkie drukowalne zdania w jakimś określonym ciągu $X_1, X_2, \ldots, X_n, \ldots$ Przez X_n rozumiemy zdanie wydrukowane na etapie n.

Pierwsze zdanie wydrukowane przez maszynę (zdanie X_1) musi być aksjomatem (ponieważ dotąd maszyna nie wydrukowała żadnych zdań), a stąd X_1 musi być prawdziwe.

Jeśli powyższa lista zawierałaby jakiekolwiek zdanie fałszywe, to musiałaby istnieć *najmniejsza* liczba n taka, że X_n jest fałszywe — to jest, musiałoby istnieć *pierwsze* zdanie fałszywe wydrukowane przez maszynę. Wiemy, że n nie jest równe 1 (ponieważ X_1 jest prawdziwe), a zatem n jest większe od 1. Znaczy to, że maszyna drukuje zdanie fałszywe na etapie n, ale na wszystkich wcześniejszych etapach drukowała wyłącznie zdania prawdziwe. Przeczy to jednak lematowi.

Zatem maszyna nigdy nie może wydrukować jakichkolwiek zdań fałszywych.

Dowód Wniosku.

Ponieważ maszyna jest ścisła (na mocy Twierdzenia), więc $\perp$ nigdy nie może zostać wydrukowane, ponieważ $\perp$ jest fałszywe. Zatem maszyna jest niesprzeczna.

Następnie, przypuśćmy, że $P\overline{X}$ jest drukowalne. Wtedy $P\overline{X}$ jest prawdziwe (na mocy Twierdzenia), co oznacza, że X jest drukowalne. Zatem maszyna jest stabilna.

Widzimy teraz, że maszyna Fergussona *jest* niesprzeczna, ale nigdy nie potrafi dowieść swojej niesprzeczności. Tak więc i ty i ja (równie dobrze jak Fergusson) wiemy, że maszyna jest niesprzeczna, ale biedna maszyna wiedzy tej nie ma!

Koniec 2

O dalszych wynikach związanych z „maszynową” interpretacją twierdzeń metalogicznych traktuje rozdział 28 *Forever Undecided*.

Wydawnictwo *Książka i Wiedza* chce opublikować *Na zawsze nierozstrzygnięte* pod koniec roku 2007.