Filozofia Nauki
Rok XIII, 2005, Nr 4(52)

Krzysztof Wójtowicz

Kilka uwag o tezie Churcha

W niniejszym artykule chciałbym przedstawić kilka uwag dotyczących wybranych tez zawartych w artykule Adama Olszewskiego (zamieszczonym w niniejszym numerze *Filozofii Nauki*). Niektóre z tych uwag mają charakter polemiczny, ale korzystając z okazji, zamieszczam również kilka uwag, które nie mają takiego charakteru, ale zostały zainspirowane artykułem A.O. i mogą być traktowane jako komentarz. Nie próbuję natomiast w tym artykule analizować nowych argumentów na rzecz prawdziwości (bądź fałszywości) tezy Churcha.

1. WSTĘP — TEZA CHURCHA

Teza Churcha jest przedmiotem dyskusji już od kilkudziesięciu lat w ramach debat dotyczących filozofii matematyki, filozofii umysłu, sztucznej inteligencji *etc*. Stwierdza ona, że naszemu intuicyjnemu pojęciu efektywnej obliczalności (które dotyczy operacji obliczeniowych dokonywanych na liczbach naturalnych) trafnie odpowiada formalna definicja obliczalności, przyjmowana w teorii obliczeń. Tę samą tezę można wyrazić nieco inaczej, w formie stwierdzenia, że formalna definicja jest trafna merytorycznie — właściwie „chwyta" nasze intuicje związane z procesem obliczania.

Formalnych definicji obliczalności jest wiele, ale wszystkie są sobie równoważne (co zresztą stanowi jeden z argumentów na rzecz tezy Churcha). W niniejszym artykule mówiąc o formalnej definicji będę miał na myśli definicję obliczalności poprzez maszyny Turinga, którą uważam za najbardziej obrazową i — swobodnie mówiąc — najbliższą naszym potocznym intuicjom procedury mechanicznej. Teza Churcha może być więc sformułowana w formie równoważności:

(TC) Funkcja określona na liczbach naturalnych jest (intuicyjnie) efektywnie obliczalna ⇔ jest obliczalna za pomocą maszyny Turinga.

2. CZY TEZA CHURCHA DOTYCZY ZBIORÓW CZY POJĘĆ?

A.O. w swoim artykule przedstawia kilka stwierdzeń dotyczących rozumienia tezy Churcha, które uważa za nieporozumienia. Do takich nieporozumień zalicza traktowanie tezy Churcha jako tezy o identyczności dwóch klas: klasy funkcji efektywnie obliczalnych (w sensie intuicyjnym) i klasy funkcji obliczalnych w sensie Turinga. Zdaniem Olszewskiego jest to błąd, ponieważ teza Churcha winna być interpretowana jako teza o identyczności dwóch **pojęć**, a nie teza o identyczności dwóch klas (zbiorów).

Zauważmy najpierw, że interpretacja tezy Churcha jako stwierdzenia dotyczącego (ekstensjonalnej) identyczności zbiorów jest słabszą interpretacją: jeśli bowiem pojęcia są identyczne, to identyczne muszą być też zbiory obiektów podpadających pod te pojęcia, natomiast implikacja przeciwna nie musi zachodzić[1]. Olszewski za właściwą uznaje więc silniejszą interpretację.

Moim zdaniem, przyjęcie takiego czy innego rozstrzygnięcia tego zagadnienia (tzn. uznanie, że TC dotyczy ona pojęć lub zbiorów) nie ma znaczenia z punktu widzenia dyskusji zasadności tezy Churcha. Rozważmy bowiem tezę Churcha w następującym (standardowym) sformułowaniu:

(TC) Funkcja jest efektywnie obliczalna ⇔ jest obliczalna w sensie Turinga.

Ma ono strukturę równoważności: A(x) ⇔ B(x). TC stanowi szczególny przypadek problemu ogólnego: czy stwierdzając równoważności tego typu mamy na myśli identyczność odpowiednich zbiorów (klas) czy też jakąś inną identyczność (pojęć, znaczeń, sensów, sądów *etc.*). Jest to problem semantyczny o charakterze ogólnym, który nie ma bezpośredniego związku z samą tezą Churcha (choć rozstrzygnięcie tego problemu wyznacza pewną interpretację wszelkich stwierdzeń o charakterze równoważności — w tym również TC[2]). Ewentualna argumentacja na rzecz takiej właśnie interpretacji musiałaby mieć charakter ogólny, problem zaś zasadności tezy Churcha nie miałby dla tej argumentacji znaczenia.

Autor wygłasza pewną ogólną tezę semantyczną, interpretacja zaś TC jest szczegółowym wnioskiem z tej tezy. Zauważmy zresztą, że dokładnie te same wnioski z analiz Autora wynikają dla negacji TC — gdyby zgodzić się, że równoważności typu A(x)⇔B(x) winny być interpretowane jako stwierdzenia dotyczące pojęć, to

[1] Implikacja przeciwna, tj. wnioskowanie o identyczności pojęć na podstawie identyczności zbiorów obiektów „podpadających" pod te pojęcia byłaby po prostu zasadą ekstensjonalności dla pojęć, a więc zasadą (co najmniej!) dyskusyjną.

[2] Dotyczy to jednak także stwierdzeń typu: „zwierzę jest ssakiem ⇔ na pewnym etapie rozwoju żywi się wyłącznie mlekiem matki".

wyznacza to zarówno pewną interpretację TC, jak i jej negacji, w żaden jednak sposób nie determinuje wyboru między nimi.[3]

Swoje rozważania dotyczącego tego problemu (zdaniem A.O. — błędu), Autor kończy stwierdzeniem, że TC jest sformułowana w jakimś języku, który nie został jeszcze do końca rozpoznany. „Nie może to, jak się zdaje, być język teorii zbiorów, gdyż wtedy *pojęcie* byłoby zbiorem. To jednak należy z góry wykluczyć". [Olszewski, s. 118].

W związku z tym stwierdzeniem nasuwa mi się kilka uwag. Zacznę od oczywistej: TC nie jest sformułowana w języku formalnej teorii zbiorów (np. ZFC). W tym sensie Autor ma oczywiście rację, ale z pewnością nie chodzi tu o skonstatowanie oczywistego faktu, że wypowiedzi dotyczące naszych intuicji nie są sformalizowane w języku ZFC. Porusza tutaj natomiast ciekawy problem: jak zidentyfikować język, w którym formułujemy tezy o matematyce, które mają charakter filozoficzny czy metodologiczny. Kiedy więc mówi o tym, że nie jest to język teorii zbiorów, może tu chodzić może jedynie o to, że teoriomnogościowy sposób myślenia (czy: ujęcia problemu — przy całej nieostrości tych terminów) jest z gruntu niewłaściwy przy analizie problemu TC.

Z jakim językiem mamy więc do czynienia, kiedy twierdzimy, że pewne matematyczne definicje stanowią adekwatne precyzacje preteoretycznych pojęć matematycznych (czyli, że „chwytają" nasze intuicje)? Kiedy analizujemy motywacje, które kryją się za poszczególnymi definicjami matematycznymi, dochodzimy np. do wniosku, że epsilonowo-deltowa definicja ciągłości adekwatnie kodyfikuje, „chwyta" nasze intuicje (podobnie jak np. topologiczna definicja ciągłości, definicja prawdy, dowodu, długości krzywej, obszaru spójnego *etc.*). Można powiedzieć, że np. topologiczna definicja homeomorfizmu stanowi właściwą formalizację naszego intuicyjnego pojęcia przekształcenia, w którym nie ma sklejania ani rozrywania, jedynie wyginanie, rozciąganie i ściskanie. Teza głosząca, że definicja homeomorfizmu jest właściwą racjonalną rekonstrukcją rozumianego intuicyjnie pojęcia przekształcenia ciągłego, jest wypowiedzią należącą do (swobodnie mówiąc) dyskursu filozoficznego (czy metodologicznego). Dlaczego jednak miałby być czymś niewłaściwym stwierdzenie, że klasa homeomorfizmów to klasa przekształceń bez sklejania i rozrywania, albo stwierdzenie, że zbiór funkcji rzeczywistych, które intuicyjnie uznamy za ciągłe jest identyczny ze zbiorem funkcji ciągłych w sensie definicji epsilonowo-deltowej?[4] A.O. wygłasza ogólną tezę semantyczną — wynika z niej, że właściwe byłoby tu mówienia o identyczności **pojęć**: intuicyjnego i formalnego pojęcia ciągło-

[3] $\neg$TC stwierdzałaby, że nie zachodzi równoważność: funkcja jest intuicyjnie obliczalna $\Leftrightarrow$ jest obliczalna w sensie Turinga. W terminologii A.O.: pojęcia te nie są identyczne. Nb. negacja wersji mocnej TC nie wyklucza asercji wersji słabej: odmienności pojęciowej może towarzyszyć równość zakresów.

[4] Zauważmy zresztą, ze kiedy np. dowodzimy równoważności epsilonowo-deltowej definicji ciągłości Cauchy'ego i ciągowej definicji Heinego, to dowód ten: (1) odbywa się *de facto* w ramach ekstensjonalnej teorii mnogości; (2) dotyczy identyczności pewnych zbiorów, a nie pojęć.

ści. Wracając zaś do TC, należy zauważyć, że jeśli uznamy traktowanie tezy Churcha za tezę o zbiorach za nieporozumienie, to za nieporozumienie należy też uznać takie traktowanie jakichkolwiek równoważności typu A(x)⇔B(x). To jednak nie ma żadnego związku z samą TC.[5]

3. CZY ISTNIEJĄ DOWODY ZA POMOCĄ TEZY CHURCHA?

Autor wyróżnia dwa sposoby użycia TC w dowodach matematycznych. Pierwszy polega na tym, że po prostu na pewnym etapie dowodu np. rozstrzygalności danego problemu przedstawiamy algorytm opisany intuicyjnie, nieformalnie, a następnie — poprzez odwołanie się do TC — stwierdzamy, że istnieje np. stosowna maszyna Turinga rozwiązująca dany problem. Tak faktycznie jest — i taka jest powszechna praktyka.[6] Moją wątpliwość budzi jednak wprowadzanie tutaj — niejako jako nowej kategorii — kategorii dowodu za pomocą TC, gdyż może to być mylące. Wolałbym mówić raczej o pewnym sposobie postępowania, polegającym na tym, że w praktyce matematycznej stosujemy pewne skróty dowodowe, na które pozwalają sobie wykładowcy (lub autorzy podręczników), mając zaufanie do słuchaczy i czytelników (że w miarę potrzeby będą umieli uzupełnić szczegóły techniczne dowodu za pomocą standardowych technik).[7] Użycie tezy Churcha w takich dowodach ma charakter — jak to żargonowo mawiają matematycy — „machania rękami", które oczywiście w miarę potrzeby można sprecyzować (np. gdybyśmy chcieli zaimplementować dany algorytm opisany nieformalnie). Czy należy w związku z tym wyróżniać w matematyce pewną klasę dowodów — dowody przez „machanie rękami"? Moim zdaniem, wystarczy po prostu odnotować fakt, że w swojej codziennej pracy (np. na wykładach albo referatach prezentujących nowy wynik) matematycy nie przedstawiają kompletnych, sformalizowanych dowodów, ale jedynie ich szkice. Zilustruję moje

---

[5] Widzę tu analogię do następującej sytuacji: przypuśćmy, że ktoś twierdzi, że każdy silnik spalinowy musi emitować pewną ilość dwutlenku węgla. Pada w tym momencie zarzut: błędem jest mówienie o silniku i dwutlenku węgla, bo tak naprawdę należałoby mówić o wiązkach wrażeń (taki zarzut mógłby postawić radykalny fenomenalista). Oczywiście, można dyskutować nad jego tezą, ale nie ma ono żadnego związku merytorycznego z problemem emisji dwutlenku węgla. Podobnie: problem interpretacji tez typu A(x)⇔B(x) nie ma związku merytorycznego z problemem zasadności tezy Churcha.

[6] W praktyce — np. na wykładach — ten sposób jest często używany. Np. mówimy: maszyna Turinga sprawdza na taśmie ....; maszyna Turinga przegląda....; maszyna Turinga sprawdza na jednej taśmie ..., na drugiej ...., a na trzeciej porównuje wyniki *etc.* Oczywiście zawsze pamiętamy o tym, że te luźne stwierdzenia można sformalizować.

[7] Cutland w swoim podręczniku posługuje się tego typu skrótami w dowodach, zauważając jednak wyraźnie, że „każdy, kto odwołuje się do tezy Churcha w ten sposób, powinien w razie potrzeby podać dowód formalny". Cutland N. *Computablility: an introduction to recursive function theory*, Cambridge University Press, 1980, s.70.

uwagi prostym przykładem: przypuśćmy, że sformułujemy i uznamy za wiarygodną Tezę o Homeomorficzności (TH)[8]:

(TH) Jeśli dwie figury na płaszczyźnie są normalne i widać, że są homeomorficzne, to faktycznie są homeomorficzne.[9]

Teza ta stwierdza, że topologiczne pojęcie homeomorfizmu (ograniczamy je do figur na płaszczyźnie) dobrze odpowiada pojęciu intuicyjnemu. Czy jednak jest sens wyróżniać dowody za pomocą TH? Byłyby to po prostu dowody „przez machanie rękami", które nie stanowiłyby jednak nowego typu dowodów matematycznych. Można sobie wyobrazić wykład z topologii, na którym wykładowca stwierdza, że przecież widać, że np. figura w kształcie litery „I" oraz litery „S" są homeomorficzne, natomiast litery „A" i „H" nie są — i że faktu tego nie będzie dalej dowodził. Nie można przecież twierdzić, że jest to jakaś nowa forma dowodu za pomocą Tezy o Homeomorfizmie. Z podobną sytuacją mamy do czynienia w całej matematyce (nie tylko w teorii obliczalności czy topologii) — dowody są prezentowane skrótowo, ale szczegóły techniczne zawsze można uzupełnić. W wypadku teorii obliczeń znaczy to po prostu, że w razie potrzeby można *explicite* zdefiniować stosowną maszynę Turinga.

Nie mam więc wątpliwości co do faktu, że szereg dowodów w matematyce jest prezentowanych w sposób skrótowy. Z takimi dowodami mamy też do czynienia na wykładach z teorii obliczeń — i w tym sensie możemy się zgodzić ze stwierdzeniem, że w dowodach wykorzystujemy TC. Chcę jednak wyraźnie podkreślić, że chodzi tutaj jedynie o skrótową prezentację, a nie o nową jakość w dowodzeniu, i że nie jest ona niczym charakterystycznym dla samej teorii obliczeń, ale powszechną praktyką w matematyce.

Autor jednak wyróżnia też drugie rozumienie terminu „dowód za pomocą TC" — jako „dowód, którego jedną z przesłanek jest TC". Ten drugi typ dowodu istotnie różni się od pierwszego — zdaniem A.O., „użycie TC w drugim rozumieniu jest istotne, bez TC twierdzenia dowieść się nie uda" [Olszewski, s. 119].

Tu się z A.O. kategorycznie nie zgadzam: takich dowodów w matematyce po prostu nie ma! Żadne twierdzenie matematyczne nie jest dowodzone przez odwołanie się do tezy, że pewna definicja formalna stanowi adekwatną formalizację pewnych intuicji.

A.O. uzasadnia swoją tezę posługując się przykładem problemu stopu, podając jego cztery sformułowania (które dla wygody Czytelnika przypomnę w przypisie).[10]

---

[8] Jest to jedynie przykład, do którego nie jestem zbytnio przywiązany. Gdyby ktoś wykazał, że TH jest niewiarygodna, to świadczyłoby to jedynie o złym dobraniu przykładu, ale nie osłabiałoby krytyki.

[9] Np. widzimy, że krążek i plama (bez dziur i doczepionych nitek) są faktycznie homeomorficzne, choć być może np. wypisanie takiego homeomorfizmu byłoby technicznie bardzo skomplikowane.

[10] „(1) O dowolnej maszynie Turinga nie jest rozstrzygalne w sensie Turinga to, że zatrzyma się, gdy na wejściu zadamy jej dowolną liczbę.

Rozważania Autora nie są dla mnie jednak do końca jasne. Niewątpliwie, formalnie problem stopu można sformułować w stylizacji maszyny Turinga. Pozostałe sformułowania nie są twierdzeniami matematycznymi — choć każde z tych sformułowań można traktować jako luźno i nieformalnie sformułowaną odpowiedź na pytanie „co to jest problem stopu i jakie jest jego rozwiązanie". Matematyk swobodnie posługuje się pojęciami maszyny Turinga, algorytmu czy procedury efektywnej (traktując je jako synonimy), choć oczywiście tylko pierwszy z tych terminów ma precyzyjnie zdefiniowany sens. Nie prowadzi to oczywiście do nieporozumień, bo wiemy przecież, że na ewentualne pytanie o to, co to takiego jest algorytm albo procedura efektywna ów matematyk odpowie, że chodzi właśnie o maszyny Turinga i działanie maszyny Turinga. Jeśli więc traktujemy te sformułowania po prostu jako zwykłe, codzienne, „robocze" sformułowania problemu stopu, to różnią się one jedynie stylizacją (po prostu terminy „algorytm", „maszyna Turinga" i „efektywna procedura obliczeniowa" są stosowane wymiennie). I odwrotnie: należy pamiętać, że zdania (2)-(4) wolno nam uznać za sformułowanie **matematycznego** twierdzenia o nierozstrzygalności problemu stopu **tylko** wtedy, gdy terminy te będziemy interpretować w sposób ścisły, matematyczny.

Natomiast przy literalnym odczytaniu tych stwierdzeń ich status jest faktycznie różny: drugie i trzecie są stwierdzeniami, w których występuje zarówno precyzyjny termin matematyczny („maszyna Turinga", „w sensie Turinga"), jak i termin rozumiany jedynie nieformalnie, intuicyjnie, czwarte zaś stwierdzenie ma czysto nieformalny charakter — mowa jest w nim o (intuicyjnie rozumianych) algorytmach i o (intuicyjnie rozumianym) pojęciu efektywnej rozstrzygalności. Tak też interpretuje te zdania A.O., zarazem twierdząc, że „Ciekawym jest tutaj fakt, że twierdzenie czwarte można w ogóle udowodnić" [Olszewski, s. 120]. Jednak podany przez Autora argument nie jest dowodem (o ile — zgodnie z intencją Autora — terminy „algorytm" czy „efektywnie obliczalna funkcja" rozumiemy w intuicyjny, preformalny sposób). Nie może być więc tu mowy o dowodzie matematycznym, co najwyżej o pewnej nieformalnej argumentacji, która ma w nas wzbudzić poczucie, że stwierdzenie (4) jest prawdziwe. Jest ono jednak co najwyżej w tym sensie prawdzi-

---

(2) O dowolnej maszynie Turinga nie jest rozstrzygalne efektywnie to, że zatrzyma się, gdy zadamy jej jakąś liczbę na wejściu.

(3) O dowolnym algorytmie nie jest rozstrzygalne w sensie Turinga to, że zatrzyma się, gdy zadamy mu na wejściu jakąś liczbę.

(4) O dowolnym algorytmie nie jest rozstrzygalne efektywnie to, że zatrzyma się, gdy zadamy mu na wejściu jakąś liczbę." (s. 119-120).

Chcę tu też dodać uwagę stylistyczną, dotyczącą (1) — takie sformułowanie może być nieco mylące, bo może sugerować, że chodzi o to, że kiedy weźmiemy dowolną maszynę Turinga M, to dla żadnej liczby *n* problem, „czy M zatrzyma się dla danej wejściowej *n*?" nie jest rozstrzygalny. Dla jednoznaczności, lepiej sformułować tę wersję np. w postaci twierdzenia: Nie istnieje taka maszyna Turinga M*, która potrafi — dla dowolnej maszyny Turinga M i danej wejściowej *n* — stwierdzić, czy maszyna M zatrzyma się przy danej wejściowej *n*.

we, co stwierdzenie „każda porządna funkcja rozkłada się w szereg Taylora", albo „z wyjątkiem sztucznych konstrukcji i patologicznych przypadków, każde A jest B". Nie są to oczywiście twierdzenia matematyczne, ale można się zgodzić, że matematycy w trakcie luźnych dyskusji mogą wypowiadać i uzasadniać tezy tego typu (podobnie jak w trakcie **swobodnych** dyskusji, mogą mówić o tym, że np. „to się powinno dać zrobić", „raczej tu się nie da skonstruować naturalnego kontrprzykładu", „ta definicja lepiej chwyta intuicje związane z pojęciem P, niż inna" *etc.*).

Olszewski zauważa, że drugie zdanie wydaje się nie być stwierdzeniem matematycznym. Oczywiście nie jest, jeśli termin „efektywna rozstrzygalność" będzie używany w sensie czysto intuicyjnym. Twierdzi jednak zarazem, że „paradoksem jest to, że przyjęcie matematycznej przesłanki TC (podkreślam, że TC uchodzi za zdanie niematematyczne) daje w efekcie matematyczne twierdzenie, którego dowód znamy" [Olszewski, s. 120]. Autor twierdzi, że dopiero po przyjęciu TC rozstrzygnięcie problemu stopu jest negatywne. „Bez przyjęcia TC odpowiedź na tak sformułowane pytanie mogłaby również brzmieć TAK". [Olszewski, s. 120]. To jednak nie stanowi bynajmniej żadnego argumentu na rzecz tezy, że przyjęcie TC jest niezbędne w pewnych dowodach matematycznych, bo argumentacja na rzecz (2) nie jest dowodem — gdyż po prostu (2) nie jest stwierdzeniem matematycznym (co przyznaje przecież A.O.). Również fakt, że przyjęcie TC czyni zdanie (2) zdaniem matematycznym nie jest bynajmniej niczym paradoksalnym. Po prostu pewne swobodne stwierdzenie uzyskuje status twierdzenia matematycznego poprzez zmianę (uściślenie znaczeń) terminów. Powróćmy na chwilę do przykładu homeomorfizmu, i rozważmy stwierdzenie, że jeśli jedna figura przestrzenna da się przeformować na drugą bez sklejania i rozrywania, to mają one taką samą grupę podstawową. W takim sformułowaniu nie jest to oczywiście twierdzenie matematyczne, jednak jeśli termin „przekształcenie bez sklejania i rozrywania" zamienimy na termin „homeomorfizm", to powstaje twierdzenie matematyczne (mówiące o tym, że przestrzenie homeomorficzne mają tę samą grupę podstawową).[11] Nie ma w tym nic dziwnego.[12]

Jak rozumiem, przykład argumentacji na rzecz stwierdzenia (2) miał stanowić argument na rzecz tezy o konieczności użycia TC w dowodach matematycznych. Faktycznie jednak, TC jest użyta na jeden z dwóch sposobów: (i) TC dla uściślenia pewnego terminu; (ii) do uzasadnienia pewnego **niematematycznego** stwierdzenia. W żadnej z tych sytuacji nie ma żadnego paradoksu.

---

[11] Analogia z TC jest prosta: tutaj odwołujemy się do tego, że intuicyjne pojęcie ciągłości (jako przekształcenie bez sklejania i rozrywania) jest podstawą topologicznej definicji.

[12] Inny przykład: przypuśćmy, że ktoś twierdzi, że jeśli funkcja rzeczywista f: $\mathbf{R}\to\mathbf{R}$ jest analityczna, to nie ma ząbków. Oczywiście, to nie jest twierdzenie matematyczne, choć jeśli sformułowanie „nie ma ząbków" zinterpretujemy jako „jest różniczkowalna", powstaje twierdzenie matematyczne. Nie ma tu również żadnego paradoksu.

4. TWIERDZENIA GÖDLA A CHARAKTER DOWODU MATEMATYCZNEGO

Autor sugeruje, że to twierdzenia Gödla wskazują na niezbywalność dowodu treściowego. W tym stwierdzeniu dostrzegam nieporozumienie, którego źródłem jest (jak sądzę) pomieszanie pojęcia dowodu niesformalizowanego (choć potencjalnie poddającego się formalizacji) z pojęciem argumentacji na rzecz prawdziwości danego stwierdzenia (lecz argumentacji nie poddającej się takiej formalizacji, czyli niebędącej dowodem w normalnym sensie tego słowa). Dowody znane nam z praktyki matematycznej są niesformalizowane, ale potencjalnie dają się sformalizować (przynajmniej takie jest przekonanie matematyków). Byłoby jednak absurdem oczekiwanie, że dowody twierdzeń z zakresu algebry, teorii prawdopodobieństwa czy analizy zespolonej zostaną przetłumaczone na język teorii mnogości — taki dowód byłby absolutnie nieczytelny (chociażby ze względu na długość). Zarazem jednak matematycy są przekonani, że ich działalność odbywa się w pewnych ustalonych ramach, i że pojęcie poprawnego dowodu matematycznego nie jest pojęciem socjologicznym, psychologicznym ani pragmatycznym, lecz pojęciem logicznym. Znane są bowiem dobrze ramy dla formalnej idealizacji matematyki — ustala je właśnie teoria mnogości (w tym sensie, że wszelkie pojęcia matematyczne dają się zrekonstruować w teorii mnogości ZFC, zaś dopuszczalne, „legalne" dowody matematyczne mieszczą się w ramach ZFC[13]). W tym sensie dowody treściowe (czyli po prostu dowody) są też dowodami formalizowalnymi. Natomiast ewentualna niezbywalność dowodu treściowego w praktyce matematycznej nie wynika bynajmniej z twierdzeń Gödla, ale po prostu z naszej konstrukcji umysłowej: jesteśmy w stanie uprawiać matematykę w taki sposób, w jaki ją uprawiamy, odwołując się do naszych intuicji, wyobraźni *etc.* Pewne pojęcia matematyczne uważamy za ważne, pewne metody dowodowe za trickowe, pewne konstrukcje za naturalne *etc.* (choć nie mamy formalnych kryteriów „trickowości" dowodu). W kontekście odkrycia faktycznie mamy do czynienia z procesami psychologicznymi wymykającymi się formalnemu opisowi — ale twierdzenia Gödla nie mają z tym faktem nic wspólnego. Nie ma bynajmniej potrzeby twierdzić, że istnieje jakaś odrębna kategoria „dowodów przez wgląd bezpośredni", które wymykają się formalizacjom. Dowody są formalizowalne. Natomiast przez odwołanie się do naszego intuicyjnego rozumienia pewnych pojęć możemy uzasadniać np. prawdziwość (czy ostrożnej: zasadność) nowych aksjomatów. Można wyobrazić sobie sytuację, w której matematycy przyjmą nowy aksjomat w wyniku powszechnej zgody na jego prawdziwość i zgodność z intuicjami. Nie będzie to jednak akceptacja w oparciu o dowód treściowy, ale w oparciu o pozaformalną analizę pojęć.[14]

---

[13] Np. dowód, który wykorzystywałby hipotezę kontinuum, nie jest pełnoprawnym dowodem matematycznym.

[14] Można wyobrazić sobie (choć jest to mało prawdopodobne), że zdanie P=NP jest zdaniem niezależnym od ZFC. Gdyby tak faktycznie było, to jako nowy aksjomat mogłoby zostać przyjęte zdanie P≠NP, jako bardziej zgodne z intuicjami badaczy.

Błędna argumentacja na rzecz niezbywalności dowodu treściowego w oparciu o twierdzenia Gödla miałaby (w pewnym uproszczeniu) taką postać:[15]

(1) Istnieje zdanie α niezależne od danej teorii formalnej T.

(2) Ponieważ α jest niezależne, więc nie istnieje dowód formalny tego zdania.

(3) Jednak my potrafimy uzasadnić jego prawdziwość.

(4) Ale skoro potrafimy wykazać prawdziwość α, więc musi istnieć jakiś dowód treściowy tego zdania.

W takiej argumentacji mamy do czynienia z oczywistym pomieszaniem pojęć. To, że my w jakiś sposób **uzasadniamy** prawdziwość zdania α nie implikuje bynajmniej, że istnieje jakiś **dowód** prawdziwości tego zdania. Fakt, że mamy poczucie prawdziwości danego zdania α, nie czyni jeszcze tego zdania dowodliwym w żadnym sensie tego słowa.[16]

Jeśli chodzi o argumentację odwołującą się do twierdzeń Gödla i zdań gödlowskich, to chcę zwrócić uwagę na ryzyko pewnego nieporozumienia. Naturalna interpretacja metamatematyczna zdania gödlowskiego G to „ja nie mam dowodu". Uznajemy je więc — w naturalny sposób — za zdanie prawdziwe. Należy jednak zauważyć, że przejście od zdania G (które jest zdaniem arytmetycznym, mówiącym coś o pewnych wielomianach) do jego interpretacji „ja nie mam dowodu" odbywa się *via* stosowne kodowania — dlatego mówimy o naturalnej **interpretacji** metamatematycznej tego zdania, a nie o czysto **matematycznej** treści tego zdania.

W wypadku zdań Gödla (a więc oryginalnie zdań arytmetycznych) mówimy o prawdziwości w intuicyjnym sensie, ale z formalnego punktu widzenia, należy mówić o prawdziwości w sposób zrelatywizowany do stosownego modelu dla PA (w którym interpretowane są terminy występujące w zdaniu G). Mówiąc o tym, że zdanie G jest prawdziwe, mamy więc *de facto* na myśli jego prawdziwość, gdy interpretowane jest „w prawdziwych liczbach naturalnych" — czyli w modelu standar-

---

Ten przykład można uznać za nieadekwatny, ponieważ panuje raczej przekonanie, że po prostu zdanie $P \neq NP$ jest prawdą, i że wcześniej czy później zostanie znaleziony jego dowód. Jednak można wyobrazić sobie sytuację, w której panuje (prawie) powszechna zgoda co do tego, że pewne zdanie α jest prawdziwe (specjaliści w danej dziedzinie uznają np., że $\neg\alpha$ miałoby bardzo nienaturalne konsekwencje *etc.*), a jednak okazuje się że α jest niezależne od ZFC. W tej sytuacji naturalne byłoby przyjęcie α jako nowego aksjomatu. Jednak oczywiście nie byłoby to przyjęcie nowego na podstawie dowodu treściowego, ale na podstawie pewnej nieformalnej argumentacji.

[15] Nie twierdzę, że tak argumentuje A.O. (choć dostrzegam w jego artykule ślady takiej argumentacji), jednak — korzystając z okazji — chcę zwrócić uwagę na to nieporozumienie, gdyż zetknąłem się z nim kilkakrotnie.

[16] Np. większość badaczy teorii mnogości ma poczucie, że aksjomat konstruowalności jest fałszywy, tj. że $V \neq L$. Oczywiście, nie ma (i nigdy nie będzie) żadnego dowodu zdania „$V \neq L$" na podstawie aksjomatów ZFC (bo jest ono od ZFC niezależne). Ewentualny dowód „treściowy" tego zdania musiałby opierać się na nowych aksjomatach teorii mnogości, ale wtedy — po przyjęciu nowych aksjomatów A — taki dowód stałby się zwykłym dowodem formalnym zdania „$V \neq L$" na podstawie teorii ZFC+A.

dowym **N**.[17] Nie jest ono jednak oczywiście twierdzeniem PA (jest bowiem niezależne od PA). Natomiast prawdziwość tego zdania w modelu standardowym **N** jest po prostu twierdzeniem ZFC. Istnieje więc formalny dowód twierdzenia „G jest prawdziwe w modelu standardowym **N**", który jest sformułowany w teorii silniejszej od PA.[18] Problem „dowodów treściowych" jest więc problemem dotyczącym psychologii twórczości matematycznej, ale nie niezupełności PA.

5. DOWODY KOMPUTEROWE

A.O. jako jeden z argumentów na rzecz tezy, że pojęcie dowodu matematycznego nie jest pojęciem dobrze określonym przytacza problem dowodów komputerowych. Należy przyznać, że status dowodów komputerowych w matematyce nie został jeszcze do końca przemyślany. Wedle mojej wiedzy, panuje poczucie, że wprawdzie dowody komputerowe są dowodami (np. twierdzenie o czterech barwach uważa się za udowodnione), ale są jakby nieco gorszymi dowodami, pozostawiają uczucie pewnego niedosytu (i to niedosytu nie tylko o charakterze estetycznym). Jednak dowód komputerowy jest dowodem formalnym — i to na ogół znacznie bardziej formalnym niż dowody „ludzkie". Działanie komputera jest dla nas przejrzyste — wykonuje on pewne operacje, które my (w zasadzie) moglibyśmy też wykonać. Problem z dowodami komputerowymi nie polega na tym, że one są mniej formalne, ale na tym, że musimy mieć zaufanie do działania pewnego fizycznego urządzenia, jakim jest komputer.[19] Nasze przekonanie o tym, że twierdzenie o czterech barwach jest prawdziwe oparte jest poniekąd na rachunku prawdopodobieństwa: jest bardzo mało prawdopodobne, aby wszystkie programy i wszystkie komputery, za pomocą których dowodzone było to twierdzenie działały błędnie — a więc możemy przyjąć, że jest to twierdzenie prawdziwe. Nie jest to jednak dowód o innej strukturze, niż zwykły dowód obliczeniowy — po prostu te operacje, które nam zajęłyby 1000000000 lat, komputer wykona w kilka godzin. Problem dowodów komputerowych jest filozoficznie inspirujący, ale nie ma on nic wspólnego z problemem dowodu treściowego ani też z problemem TC.

[17] Należy pamiętać, że istnieją (oczywiście niestandardowe) modele dla PA, w których zdanie G jest **fałszywe**.

[18] Warto tu przypomnieć wyniki dotyczące niezależności od PA pewnych naturalnych zdań o charakterze kombinatorycznym (mam na myśli wyniki Parisa-Harringtona i podobne). Są to wszystko zdania niezależne od PA, ale prawdziwe w sensie modelu standardowego. Czy znaczy to, że aby je udowodnić musimy odwoływać się do dowodu treściowego, który ma inny charakter niż dowód formalny? Dowody tych zdań są formalizowalne w ZFC.

[19] Mówiąc ściśle: musimy mieć zaufanie do tego, że operacje przeprowadzane przez komputer są **zgodne z programem**, ale nie do tego, że są **formalne**. Jest to — swobodnie mówiąc — kwestia zaufania o charakterze empirycznym, a nie logicznym.

6. DOWÓD TEZY CHURCHA?

Autor swoje rozważania na temat ew. dowodu tezy Churcha osadza w kontekście rozważań dotyczących twierdzeń Gödla i braku ogólnego pojęcia dowodu w matematyce. Może powstać wrażenie, że brak dowodu tezy Churcha ma jakiś związek z faktem, że pojęcie dowodu matematycznego jest niejasne.

To wrażenie jest błędne. Brak dowodu tezy Churcha nie ma żadnego związku z tym, że brak jest w matematyce ogólnego, dobrze określonego pojęcia dowodu. Po prostu, TC nie dotyczy samych tylko pojęć matematycznych, ale naszych intuicji, i w tym sensie, niezależnie od tego jak byśmy sprecyzowali pojęcie dowodu (byle sensownie — a nie np. uznali głosowanie za dobrą metodę dowodową), to matematycznego dowodu TC być nie może. Możemy co najwyżej mówić o argumentacji na rzecz TC, ale żadna taka argumentacja nigdy nie stanie się dowodem matematycznym.