TEZA CHURCHA JAKO HIPOTEZA EMPIRYCZNA

Badania nad Tezą Churcha (dalej TC) nabierają powoli rozpędu. Pojawiają się różne jej sformułowania. Dlatego zawsze należy podać sposób jej rozumienia. Oto sformułowanie, które jest, według mnie, najbliższe temu, jakie podał Church:

[TC] Pojęcie *funkcji efektywnie obliczalnej*[1] jest identyczne z pojęciem *funkcji rekurencyjnej*.

Przekonanie o tym, że tak należy rozumieć TC opieram na następującym tekście samego Churcha:

> Teraz zdefiniujemy pojęcie (notion) [...] *efektywnie obliczalnej* funkcji liczb całkowitych dodatnich przez identyfikację go z pojęciem funkcji rekurencyjnej liczb całkowitych dodatnich (lub z λ-definiowalną funkcją liczb całkowitych dodatnich).[2]

Frege uważał, że przybliżeniem identyczności pojęć jest ich materialna równoważność.

Dla dowolnych pojęć F, G:

Pojęcie F jest identyczne z pojęciem G wtedy i tylko wtedy, gdy $\forall x\ (Fx \equiv Gx)$[3].

Symbol Fx czytamy ‘obiekt x podpada pod pojecie F’. Dzisiejsze rozumienie prawej strony powyższej równoważności zostało zdominowane przez logikę pierwszego rzędu i teorię mnogości. Prowadzi to do uznania za równoważne zdania wyrażającego materialną równoważność pojęć (i tym samym, za Fregem, ich identyczności) ze zdaniem o równości ich ekstensji (zbiorów).[4]

W niniejszej pracy termin *hipoteza* rozumiany jest jako zdanie, które w sensie absolutnym posiada wartość prawdy lub fałszu, ale na obecnym etapie rozwoju nauka nie jest w stanie tej wartości określić. Od czasu Poppera cechą wyznaczającą *naukowość* jakiejś teorii czy hipotezy jest jej *falsyfikowalność*. Z tym, ze tutaj będziemy *falsyfikowalność* rozumieć szerzej, jako możliwość wykazania fałszywości zdań. Jest to cecha konieczna empiryczności, ale nie jest ona wystarczająca, gdyż pod tak pojętą cechę podpadają hipotezy matematyczne (np. hipoteza kontinuum). Hipotezę naukową nazwiemy *empiryczną* wtedy i tylko wtedy, gdy posiada ona zawartość *empiryczną*. Zawartość empiryczna hipotezy (teorii) - to zbiór zdań bazowych (obserwacyjnych), które można wyprowadzić (za pomocą metod logicznych) z hipotezy. Zdanie obserwacyjne - to takie, którego wartość logiczna może być określona na podstawie spostrzeżeń zmysłowych (konfrontacji z rzeczywistością). Takie zdania zawierają terminy obserwacyjne. Prawdziwość tych zdań wyznaczona jest poprzez kontyngentne

---

[1] W niniejszej pracy, o ile nie zostanie jasno stwierdzone inaczej, gdy pojawi się termin *funkcja* znaczyć on będzie *funkcja określona w liczbach całkowitych nieujemnych*, czyli w liczbach naturalnych (z zerem).

[2] Ten fragment pochodzi z paragrafu siódmego pracy ‘An Unsolvable Problem of Elementary Number Theory’. W paragrafie pierwszym jest podobnie - w przypisie trzecim pojawia się sformułowanie tej definicji: ‘... funkcja liczb całkowitych dodatnich będzie się nazywać efektywnie obliczalną, jeśli jest λ-definiowalną w sensie paragrafu 2 poniżej [...]’.

[3] Frege miał nie rozróżnić znaków ≡ i = (identyczności i równoważności). W pracy z 1879 roku używał znaku ≡, zaś w pismach z 1893 i 1903 roku używał znaku =. Por. Jan H. Alnes, ‘Sense and Basic Law V in Frege’s Logicism’, *Nordic Journal of Philosophical Logic*, 4(1999), ss. 1-30, szczególnie s. 4.

[4] W systemie Fregego przejście od identyczności ekstensji pojęć do identyczności pojęć było możliwe dzięki tzw. Basic Law V. Jak wiadomo to prawo dawało sprzeczność w obrębie systemu Fregego logiki drugiego rzędu. Por. uwagi na ten temat w: S. Shapiro, ‘Foundations without Foundationalism’, Oxford 2002, ss. 16-17.

własności Uniwersum – świata w którym żyjemy. Można powiedzieć także (ujmując rzecz od strony epistemologii), że są to zdania aposterioryczne.

1. Próbka poglądów o empirycznym charakterze TC.

Spośród *ojców obliczalności* zwolennikiem pojmowania TC jako hipotezy empirycznej był Emil Post. W swym trzystronicowym artykule 'Finite Combinatory Processes. Formulation I' podaje swą analizę procesu obliczania, która - choć niezależna - co do istoty jest identyczna z analizą Turinga. Do redakcji *Journal of Symbolic Logic* praca Posta dotarła dnia 7 października 1936 roku i jest późniejsza od słynnej pracy angielskiego logika, choć została wcześniej opublikowana. W ostatnim akapicie Post pisze:

> Autor spodziewa się, że obecne sformułowanie okaże się być logicznie równoważnym z rekurencyjnością w sensie rozwiniętym przez Gödla-Churcha.[...] Jego zadaniem nie jest jedynie zaprezentowanie systemu o pewnej logicznej potencji, ale również, w tej wąskiej dziedzinie, psychologicznej odpowiedniości (psychological fidelity). W tym drugim sensie są rozważane coraz to szersze i szersze sformułowania. Z drugiej strony naszym celem będzie wykazanie, ze wszystkie one są równoważne logicznie sformułowaniu I. W tym momencie proponujemy tę konkluzję jako *hipotezę roboczą*. I dla nas (to our mind) taką również jest identyfikacja efektywnej obliczalności z rekurencyjnoscia dokonana przez Churcha.

Tutaj pojawia się przypis numer osiem:

> [...] W rzeczywistości praca dokonana przez Churcha i innych przenosi (carries) tę identyfikację znacząco poza (considerably beyond) stadium (stage) hipotezy roboczej. Jednak maskowanie tej identyfikacji za pomocą definicji ukrywa ten fakt, że zostało odkryte fundamentalne ograniczenie na moc matematyczną (mathematicizing power) Homo Sapiens i przysłania nam konieczność jej ciągłej weryfikacji.

I dalej, w tekście głównym:

> Z tej hipotezy [...] wyrasta dzieło Gödla-Churcha. Sukces tego programu mógłby, według nas, zmienić tę hipotezę raczej w *prawo natury*, a nie w definicję czy też aksjomat.

Przytoczone teksty wskazują na to, że Post uznawał:

- TC nie jest ani definicją ani aksjomatem.[5]
- TC jest hipotezą roboczą wymagającej nieustannej weryfikacji.
- TC została przez Churcha i innych podniesiona do rangi obowiązującego prawa natury.
- TC nakłada ograniczenia na (psychologiczne) matematyczne możliwości Człowieka.

To pojmowanie TC było niezgodne z rozumieniem jej przez Churcha, dla którego TC była definicją. Dlatego Church, w swej recenzji cytowanej pracy Posta, odnosi się krytycznie do pojmowania TC jako *hipotezy roboczej* wymagającej ciągłej weryfikacji. Uważa, że pojęcie *efektywnej obliczalności* nie posiada precyzyjnego sensu i dlatego ujęcie TC jako *hipotezy roboczej* nie posiada również dokładnego znaczenia. Obliczalność, na którą nałożono

---

[5] Sposób, w jaki rozumiał Post definicje, nie jest jasny. Wydaje się jednak, że rozumiał je jako to, co czasami nazywa się *definicją syntetyczną*. Doktorant Posta – Martin Davis – uważał podobnie jak jego nauczyciel: 'Jak możemy wykluczyć możliwość stanięcia któregoś dnia (być może za sprawą pozaziemskich przybyszów) wobec (być może ekstremalnie złożonego) urządzenia (device) lub 'wyroczni', która będzie 'obliczać' jakąś nieobliczalną (w sensie Turinga AO) funkcję?', M. Davis, 'Computability and unsolvability', Dover, New York 1982, s. 11.

warunek skończoności (obliczalność maszynowa), miałaby być adekwatną reprezentacją intuicyjnego pojęcia obliczalności. Wtedy, wedle Churcha, znika potrzeba przyjęcia takiej hipotezy roboczej. Nie odniósł się natomiast wcale do poglądu Posta wyrażonego w cytowanym przypisie.
Ta wypowiedź Churcha na temat TC, nie licząc artykułu w którym została sformułowana, jest jedną z nielicznych. Wydaje się, że później zmodyfikował on nieco swoje poglądy na ten temat. W roku 1940 pisał tak:

> Formalna definicja efektywnej obliczalności [...], i adekwatność tej definicji dla reprezentowania **empirycznego** (podkreślenie moje AO.) pojęcia efektywnej obliczalności znajduje mocne wsparcie w ostatnich rezultatach Turinga [6].

Nie jest natomiast całkiem jasne, jakiego rodzaju definicją miałaby być dla niego sama TC. W przypisie 168 swego 'Introduction to Mathematical Logic' Church rozróżnia cztery typy definicji funkcjonujących w obrębie logiki: definicje będące skrótami (rodzaj definicji syntetycznych?)[7], definicje eksplikujące notację języka (podobne do definicji realnych), reguły semantyczne (metajęzykowe) ustalające interpretację języka systemu[8], definicje rozszerzające język jakieś systemu formalnego (są częścią języka przedmiotowego i muszą spełniać warunki poprawności sformułowane przez Leśniewskiego). Wydaje się, że TC podpadać może tylko pod drugi typ definicji, a to z tego powodu, że Church próbował podać 'usprawiedliwienie' dla swej tezy. Jeśli to przyjąć, to dziwi sprzeciw Churcha wobec ujęcia TC przez Posta, bowiem dałoby się uzgodnić takie dwa stanowiska w kwestii charakteru TC.

Inny logik, o mocnym zacięciu filozoficznym, Hao Wang zwrócił uwagę na empiryczny charakter TC. W książce 'A Survey of Mathematical Logic' pisał:

> Wydaje się, że w pojęciu efektywności sedno sprawy tkwi w terminach mechanicznych i równocześnie mieści się idealizacja, która prowadzi do nieskończoności (there is a core in mechanical terms, and at the same time, there is an idealization which brings in infinity). To czego jakiś fizyczny obiekt może dokonać niezawodnie (reliably) i systematycznie, mogłoby wydać się efektywne, bez względu na to czy rozumiemy sam ten proces, czy też nie. Z tego powodu wydaje się być empirycznym pytanie o to, czy wszystkie efektywne funkcje są rekurencyjne[9].

W prywatnej rozmowie z Thomasem Wang miał powiedzieć, że w powyższym fragmencie chciał uczynić wiarygodną empiryczną interpretację TC, i nie był przekonany co do jej poprawności.[10] Wedle Wanga, jeśli się usunie 'empiryczny element' (z pojęcia efektywnej obliczalności), to istnieje możliwość dowodu TC. Biorąc TC z jej empirycznym komponentem może być tak, że zaobserwujemy rzeczywistą 'maszynę', która będzie obliczać nierekurencyjną funkcję. Thomas rozważa taką hipotetyczną maszynę M, która miałaby obliczać nierekurencyjną funkcję. Zwraca uwagę na to, że maszyna M musiałaby mieć nieskończony czas pracy, gdyż każda funkcja o skończonej dziedzinie w zbiorze liczb naturalnych jest rekurencyjna, a dodatek musielibyśmy mieć możliwość ją tak długo

---

[6] Por. Alonzo Church, 'The Concept of Random Sequence' , *Bulletin of the American Mathematical Society*, 46(1940), s. 133

[7] W tym kontekście Church nie mówi o definicjach syntetycznych. Przy drugim rodzaju definicji wspomina *definicje realne*, jednak nie używa tej nazwy i zaznacza, że chce przez to uniknąć skojarzeń i presupozycji związanych z tym terminem.

[8] Church uważał, że mamy do czynienia z językiem sformalizowanym o ile podana jest jego interpretacja. Jest to pogląd atrakcyjny filozoficznie, choć obecnie rzadko uznawany.

[9] Hao Wang, *op. cit.* s. 87.

[10] Por. William J. Thomas, 'Church's Thesis and Philosophy', nieopublikowana praca doktorska, Case Western Reserve University, 1972, s. 22.

obserwować. Nawet jeśli weźmiemy pod uwagę tzw. 'maszynę Zenona'[11], to matematyk - dla wykazania, że oblicza ona funkcję nierekurencyjną - musiałby mieć możliwość prześledzenia nieskończenie długiego dowodu. Te argumenty wydają się być celne w odniesieniu do maszyny M, która miałaby realnie istnieć. Wang podaje przykład takiej funkcji, zasugerowanej mu przez Speckera: f(n) = 0, gdy n-tego dnia (licząc np. od dzisiaj) wystąpiło gdzieś na świecie trzęsienie ziemi; f(n) = 1, w przeciwnym razie. Nie bardzo wiadomo jak można by wykazać rekurencyjność czy też nierekurencyjność funkcji f. Problematyczne jest także to, iż - ściśle rzecz biorąc - nie mamy do czynienia z funkcją określoną na całym zbiorze liczb naturalnych, zatem należałoby przyjąć istnienie takiej funkcji przez założenie istnienia platońskich obiektów. Funkcje obliczalne przez systemy fizyczne nazywać będziemy f-obliczalnymi[12].

Wersja TC idąca po linii uwagi Wanga to tzw. *fizyczna* wersja TC[13]:

[FTC] Dowolna funkcja określona w liczbach naturalnych, która jest f-obliczalna, jest obliczalna T-obliczalna[14].

Wymieniony fizyczny system (rzeczywisty lub możliwy) - to taki, którego: a) stany zajmują skończoną przestrzeń, b) jego okres funkcjonowania jest skończony w czasie, c) system funkcjonuje zgodnie z prawami fizyki.
Obliczalność w takim systemie polega na tym, że:

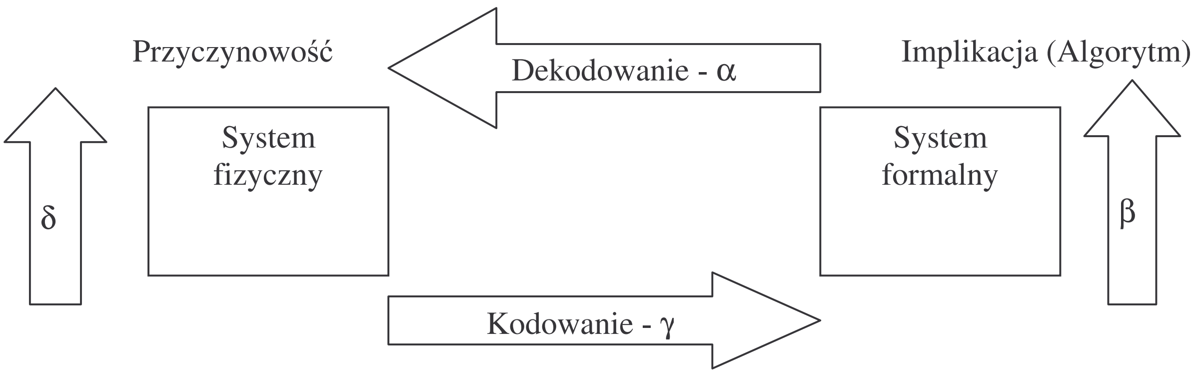
> [...] jego funkcjonowanie przeprowadza go od jednego zbioru stanów 'wejściowych' do innego zbioru stanów 'wyjściowych'.[...] Stany są oznaczone w pewien [sposób] sposób, system jest przygotowany w stanie wejściowym i potem, po wykonaniu ruchu, stan wyjściowy zostaje zmierzony[15].

W przybliżeniu można to tak opisać: bierzemy pod uwagę jakiś system fizyczny o którym możemy powiedzieć, że w sensie podanym przez Deutscha 'oblicza' jakąś funkcję. Zakładamy, że zachowanie systemów fizycznych ma charakter przyczynowy. Rosen[16] uważa, że koniecznym założeniem uprawiania teoretycznej nauki jest symulacja przyczynowych związków zachodzących w systemach fizycznych (materialnych), za pomocą implikacji zachodzących pomiędzy zdaniami opisującymi te systemy. Jeśli zachowanie takiego systemu jest powtarzalne i zachowanie systemu potrafimy mierzyć (obserwować), to możemy poszukiwać formalizmu (przeliczalnie rekurencyjnego), który modelowałby ów system. Jeśli taki system znajdziemy i da się przechodzić od systemu fizycznego do systemu formalnego i z powrotem, to możemy mówić o owym systemie jako o realizacji systemu formalnego (zaś o

---

[11] W nawiązaniu do paradoksów Zenona z Elei. Takie 'maszyny' wykonują n+1-krok w czasie o połowę krótszym niż n-ty krok.
[12] Podobne rozważania prowadziłem w pracy: A. Olszewski, 'Teza Churcha a platonizm', *Zagadnienia Filozoficzne w Nauce*, 34(1999), ss. 96-100.
[13] Wersja Kreisla [FTC] jest taka: *The behavior of any discrete physical system evolving according to local mechanical laws is recursive*. G. Kreisel, 'Mathematical Logic', [w:] 'Lectures on Modern Mathematics', Saaty (ed.), Wiley 1965, ss. 95-195. Por. odcinek 2.714.
[14] Jak widać, z TC znika pojęcie funkcji rekurencyjnej, na korzyść maszyn Turinga. Ma to związek tym, że maszyny Turinga związane są z klasą funkcji częściowo rekurencyjnych. Ciekawą pracą na temat TC jest: 'A Speculative Church-Turing Theorem', U. Boker, N. Dersowitz. Znajdujemy tam interesujące sformułowanie TC: *Modele obliczeniowe są równoważne, lub słabsze, niż maszyny Turinga* (które nawiązuje do sformułowania Churcha z abstraktu opublikowanego w 1935 roku). Tam też, jak postulują autorzy, znajduje się dowód TC.
[15] D. Deutsch, 'Quantum theory, Church-Turing principle and the universal quantum computer', *Proceedings of the Royal Society, Series A*, 400(1985), ss. 97-117.
[16] Por. R. Rosen, 'Effective Processes and Natural Law', [w:] 'The Universal Turing Machine. A Half-Century Survey', R. Herken (ed.), Springer, Wien, New York 1994, ss. 485- 498.

systemie formalnym jako o symulacji (reprezentacji) systemu fizycznego). Można to przedstawić w postaci rysunku:[17]



Jeśli strzałki powyższego schematu rozumieć jako zależności funkcyjne α, β, γ, δ, to zachodzi relacja modelowania pomiędzy danym systemem fizycznym i systemem formalnym o ile spełniony jest warunek przemienności diagramu: $\delta = \alpha \circ \beta \circ \gamma$[18]. [FTC] nie jest, jak uważa Fitz, *rozszerzeniem* TC, lecz jej szczególnym przypadkiem. W pojęciu funkcji efektywnie obliczalnej jest przecież miejsce dla systemów fizycznych. Teza odwrotna do [FTC] ma także empiryczny charakter, gdyż być może któraś z funkcji rekurencyjnych (częściowo rekurencyjnych) nie jest obliczalna przez żaden system fizyczny. [FTC] doprowadziła fizyków do rozważenia systemów fizycznych przekraczających możliwości maszyn Turinga. Należy tutaj zaliczyć: tzw. maszyny Zenona (przyspieszające)[19], obliczanie analogowe, obliczanie kwantowe (quantum computation), kwantowe procesy (quantum processes), maszyny wykorzystujące efekty relatywistyczne, biologiczne obliczanie i inne[20].

Jako kolejnego zwolennika empiryczności TC należy wymienić Robina Gandy. Swoją koncepcję prezentuje w pracy 'Church's Thesis and Principles for Mechanisms'[21]. Jego rozważania dotyczą obliczalności przez maszynę rozumianą, jak sam pisze, *w sensie dziewiętnastowiecznym*. Przykładowo podaje 'Analityczny Silnik' Babbage'a. Ograniczenia na zakres terminu *maszyna* są takie: a) ***finityzm*** - wyłącza z rozważań maszyny analogowe. Jedynym 'fizycznym' ograniczeniem jest ograniczenie dolne - na liniowe wymiary atomowych części maszyny oraz ograniczenie górne - na prędkość rozchodzenia się zmian (prędkość światła), b) ***dyskretność*** - proces obliczania można opisać za pomocą terminów dyskretnych, c) ***determinizm*** – zachowanie maszyny jest jednoznacznie zdeterminowane przez jej stan początkowy.
Gandy chce, żeby jego opis maszyny stosował się do dowolnych urządzeń: mechanicznych, elektrycznych i pojęciowych.
Formułuje cztery zasady, które mają definiować maszyny: (I) dysponujemy zbiorem S opisów stanów maszyny. Zbiór ten jest podzbiorem dziedzicznie skończonych zbiorów (hereditarily

[17] Rosen, *op. cit.,* s. 491.
[18] Por. Hartmut Fitz, 'Church's Thesis. A Philosophical Critique of the Foundations of Modern Computability Theory', praca magisterska, Berlin, 2001, s. 94.
[19] Możliwość istnienia systemu fizycznego, który wykonuje nieskończoną liczbę kroków (w jego czasie wewnętrznym) w skończonym czasie (zewnętrznym względem systemu) miał rozważać już Hermann Weyl. Por. P. H. Potgieter, 'Zeno Machines and Hypercomputations', złożone do *Elsevier Science*.
[20] Por. na przykład Potgieter, *op. cit.*; J. Mycka, 'Empiryczne aspekty teorii obliczalności'; Fitz, *op. cit.*; R. Penrose, 'Nowy umysł cesarza', PWN, Warszawa 1995.
[21] Robin Gandy, 'Church's Thesis and Principles for Mechanisms', [w:] J. Barwise, H. J. Keisler, K. Kunen (eds.), "The Kleene Symposium', North-Holland, 1980, ss. 123-148.

finite sets - HF), nadbudowanym nad nieskończonym zbiorem etykiet (labels), wraz z określoną na nim funkcją przejścia; następne trzy zasady nakładają ograniczenia na S i funkcję przejścia $F:S \rightarrow S$: (II) teoriomnogościowy poziom budowy maszyny jest ograniczony tzn. $\exists k \ (S \subset HF_k)$; (III) dla każdego opisu istnieje ograniczenie na liczbę elementów, z których zbudowana jest maszyna; (IV) zasada lokalnej przyczynowości – każdy stan zależy od ograniczonego (lokalnego) stanu poprzedniego. Gandy wykazuje, że każda funkcja obliczana przez maszynę spełniającą te zasady (za Kreislem nazywane m-obliczalne) jest obliczalna przez maszynę Turinga. Jeśli zaś osłabi się którąś z zasad, to maszyna obliczać będzie dowolną funkcję – będzie dopuszczać *wolną wolę*.

[MTC] Dowolna funkcja określona w liczbach naturalnych, która jest obliczalna przez skończoną, dyskretną i deterministyczną maszynę (m-obliczalna), jest obliczalna przez maszynę Turinga (T-obliczalna).

MTC posiada ścisły dowód, jest twierdzeniem matematycznym. Jej empiryczny charakter polega na założenia, że każda maszyna musi spełniać warunki wyrażone przez zasady (I-IV). Jest to jednak kolejna *teza*[22]. Swój argument Gandy tak formułuje:

[G1] **Teza Gandy'ego:** Dowolne dyskretne, deterministyczne i mechaniczne urządzenie (device) spełnia zasady (I) – (IV).
[G2] **Twierdzenie Gandy'ego:** Każda funkcja obliczalna za pomocą urządzenia spełniającego zasady (I) – (IV) jest również T-obliczalna.
[G3] **Wniosek:** [MTC].[23]

Ogólnie można powiedzieć, że jeśli jakieś sformułowanie TC posiada ścisły dowód, to istnieje jakaś wcześniejsza *teza*, którą przyjęto na podstawie intuicji bez ścisłego dowodu. W literaturze anglosaskiej te właśnie założenia nazywane są czasami *Centralnymi Tezami* argumentacji za TC.

Kleene sformułował dwa dodatkowe *heurystyczne* argumenty za prawdziwością TC. Pierwszy z nich (nazwiemy go argumentem ***z braku kontrprzykładu***) mówi, że TC jest prawdziwa ponieważ każda funkcja efektywnie obliczalna i każda operacja, pozwalająca efektywnie definiować funkcję z innych funkcji, okazała się być ogólnie rekurencyjna. Zbadane zostały różnorodne funkcje i ich klasy. Założeniem tych badań było wyczerpanie wszystkich znanych typów funkcji. Metody badań właściwie wykluczają to, że można znaleźć funkcję, która będzie efektywnie obliczalna i nie będzie można (za pomocą tych metod) przekształcić jej w funkcję rekurencyjną. Kreisel[24] podał przykład takiej funkcji: niech będzie dany system formalny arytmetyki konstruktywnie poprawny, dla którego konstruktywnie wyliczymy jego dowody. Definiujemy funkcję f taką, że $f(n) = 0$, gdy konkluzja n-tego dowodu albo nie ma postaci egzystencjalnej albo ma, ale dowód nie podaje świadka (witness) dla tej kwantyfikacji; $f(n) = (m+1)$, gdy konkluzja dowodu ma postać egzystencjalną i dowód podaje m jako świadka. Okazało się, że funkcja f jest mechanicznie obliczalna. Ten przykład jest interesujący, gdyż (jak zauważa Odifreddi) następuje w nim przejście od formalnej derywacji do *mentalnego aktu* dowodzenia. Odifreddi uważa, że ta definicja nie jest nawet

[22] Fenomen *tez* jest w obrębie nauki dość powszechny. Shagrir sygnalizuje, że istnieje problem z interpretacją maszyny Gandy'ego. Por. Oron Shagrir, 'Effective Computation by Humans and Machines', *Minds and Machines*, 12(2002), ss. 234-236.
[23] Por. Gandy, *op. cit.*, ss. 126 i 123.
[24] Kreisel, *op. cit.*, nr. 2.35.

sensowna, z teoriomnogościowego punktu widzenia[25], Polski logik Józef Pepis, który zajmował się TC i korespondował w tej sprawie z Churchem, odnośnie argumentu z kontrprzykładu pisze tak:

> Co się zaś tyczy tej hypotezy (zachowuję oryginalną pisownię autora A.O.), to autorowie wyżej wspomniani (Church i Turing A.O.) nie dają żadnych przekonywujących argumentów na jej poparcie, lecz opierają się li tylko na **empirycznym** (podkreślenie moje A.O.) fakcie, że nie znane są dotychczas prócz funkcji rekursywnych żadne inne funkcje „obliczalne". [...] Wobec takiego stanu rzeczy kwestia całkowitej rozwiązywalności zagadnienia rozstrzygalności dla systemu węższego rachunku funkcyjnego pozostaje **kwestią otwartą** (podkreślenie A.O.). [26]

Drugi z argumentów Kleenego (nazywa się go argumentem ***ze zbieżności sformułowań***) powołuje się na pewien rodzaj *stabilności* pojęcia funkcji efektywnie obliczalnej. Różnorodne sformułowania formalne tego pojęcia okazały się być równoważnymi. Kreisel podkreśla, że pomimo równoważności sformułowań nie można wykluczyć *systematycznego błędu* w tych ujęciach oraz tego, że pojęcie - o które nam naprawdę chodzi (efektywnej obliczalności) - nie występuje pośród pojęć równoważnych[27].

Podsumowując powyższe rozważania można chyba stwierdzić, że trzy przytoczone wersje TC – [TC], [FTC] oraz [MTC] – posiadają konsekwencje zawierające terminy obserwacyjne. Są to odpowiednio: *pojęcia*, *systemy fizyczne* i *maszyny*. Ten drugi i częściowo trzeci termin mają desygnaty wyraźnie należące do Uniwersum. Desygnatami pierwszego terminu są pojęcia - tajemnicze obiekty funkcjonujące w umyśle. Dla funkcji h-obliczalnych zdania obserwacyjne mają przez to nieco innych charakter. Obserwacja będzie introspekcją i analizą pojęć. Jej zmysłowy (spostrzeżeniowy) charakter polega na koniecznym pojawieniu się opisu obliczania funkcji w tym co Kant nazywał zmysłowymi formami naoczności, w szczególności w formie przestrzeni. Ogólnie, jeśli weźmiemy dowolny obiekt (algorytm, system fizyczny, maszynę skończoną) i wykażemy, że funkcja obliczana przez taki obiekt okaże się być rekurencyjna, to będzie to cząstkowe potwierdzenie dla odpowiedniej wersji TC. Zaś znalezienie kontrprzykładu (w dowolnej z trzech dziedzin) sfalsyfikuje odpowiednią wersję TC. Znaczy to, że żaden człowiek (system fizyczny, maszyna) nie powinien umieć efektywnie obliczać wartości funkcji (dla dowolnego argumentu), która nie jest rekurencyjna[28].

2) **Argumenty Churcha i Turinga za TC**.

Przyjrzyjmy się dokładnie, podanemu przez samego Churcha, uzasadnieniu dla TC. Pojawia się ono w paragrafie siódmym 'An Unsolvable Problem of Elementary Number Theory'[29], i jest to jego drugi argument za TC.

[C1] **Definicja**: Funkcja f jest efektywnie obliczalna wtw istnieje system formalny F w którym jest ona słabo reprezentowalna tzn. jednoargumentowa funkcja f jest obliczalna w

---

[25] Por. P. Odifreddi, 'Kreisel's Church', [w:] 'Kreiseliana', P.Odiferddi (ed.). AK Peters 1996, ss. 389-415; szczególnie 399-400.

[26] Józef Pepis, 'O zagadnieniu rozstrzygalności w zakresie węższego rachunku funkcyjnego', Lwów 1937, ss. 169-170.

[27] Por. Kreisel, *op. cit.*, nr. 2.715.

[28] Bardzo intereującą pracą o empirycznym charakterze TC jest: S. Hansson, 'Church's Thesis as an Empirical Hypothesis', *International Logic Review*, 16(1985), ss. 96-101.

[29] *American Journal of Mathematics*, 58(1936), ss. 345-363. Przedrukowane w antologii Davisa ss. 88-107.

pewnym systemie F wtw istnieje formuła A w języku systemu F taka, że dla każdego $n \in N$: $\vdash_F A(\underline{n})=\underline{m}$ wtw $f(n) = m$[30].
[C2] **Centralna Teza Churcha**[31]: Relacja Prov(x,y) 'x jest dowodem y' w F jest (nawet pierwotnie) rekurencyjna. Znaczy to, że zbiór aksjomatów i reguł inferencji są rekurencyjnie przeliczalne oraz dla każdej reguły istnieje funkcja rekurencyjna g taka, że g(n, x) = y; gdzie y jest numerem gödlowskim formuły, którą uzyskuje się w F przez zastosowanie n-tej reguły systemu F do formuł o numerze gödlowskim x.
[C3] **Wniosek (TC)**: Każda funkcja efektywnie obliczalna jest rekurencyjna.

Argument ten ma charakter dedukcyjny. TC wynika logicznie z przyjętych przesłanek. Church nie podaje jednak żadnego uzasadnienia dla [C2][32]. Załóżmy, że istnieje efektywnie obliczalna i nierekurencyjna funkcja f. Dołączmy do aksjomatów jakiejś odpowiednio bogatej arytmetyki AR, jako aksjomaty, wszystkie zdania o postaci $P(\underline{n},\underline{m})$ w taki sposób, że: $\vdash_{AR'}$ $P(\underline{n},\underline{m})$ wtw f(n)=m, gdzie P jest dwuargumentowym, nowym predykatem. Aksjomatyka systemu pozostanie efektywna i sam system - również. Ponieważ nie dołączamy żadnej nowej reguły, zbiór reguł jest efektywnie enumerowalny i każda reguła pozostaje efektywna. Funkcja f byłaby w tej hipotetycznej arytmetyce AR' słabo reprezentowalna. TC jest wtedy oczywiście fałszywa. Nie można zatem w [C2] zamienić słowa 'rekurencyjny' na słowo 'efektywny'. Jest to mało prawdopodobne, ale jednak możliwe, że efektywna, lecz nierekurencyjna funkcja istnieje. Dlatego Post, który interpretował TC w kategoriach psychologii kognitywistycznej uważał, że dla weryfikacji TC należałoby zbadać **wszystkie możliwe** sposoby, w jakich umysł ludzki mógłby sformułować skończony proces.[33] Chodzi o wyczerpującą analizę czasowo-przestrzennych możliwości symbolizacji i 'symbolicznej' manipulacji przez człowieka.[34] Analiza ta dotyczy człowieka jako kogoś obecnego w rzeczywistości – Uniwersum. W podobnym duchu wypowiadał się Kreisel gdy przypuszczał, że dopiero zbudowanie systemu formalnego obejmującego **całą** matematykę mogłoby rozstrzygnąć TC w odniesieniu do funkcji obliczalnych przez człowieka (h-obliczalnych).[35] Post postulował stworzenie *teorii pojęć* (Theory of Conception or Concepts). Ta nauka, jak się zdaje, miałby mieć charakter empiryczny i zdawać sprawę z faktycznych własności ludzkiego umysłu. W podobnym duchu wypowiada się Gandy, gdy pisze o Gödlu:

> Obiekcja GÖDLA może zostać odpowiednio usprawiedliwiona poprzez teorię inteligencji. Jak on przyjmuje, nasze obecne rozumienie ludzkiego umysłu jest dalekie od bycia wystarczająco wnikliwym (penetrating) dla skonstruowania takiej teorii. Dla tego celu wiedza dostarczana przez introspekcję, historię idei, eksperymentalną psychologię, neurofiziologię i sztuczną inteligencję wydaje się być za skromna (seems meagre indeed). Należy zachować otwarty umysł.[36]

---

[30] Oron Shagrir, *op.cit.*, *Minds and Machines*, 12(2002) wymaga by funkcja była reprezentowalna. Ten warunek nie jest jednak dobry, ponieważ w systemie sprzecznym każda funkcja jest reprezentowalna i TC przez to byłaby fałszywa.
[31] Nazwa pochodzi od W. Sieg, 'Step by Recursive Step: Church's Analysis of Effective Calculability', *The Bulletin of Symbolic Logic*, 3(1997), ss. 154-180.
[32] Por. W. Sieg, *op. cit.*, s. 165.
[33] Por. Fitz, *op. cit.*, ss. 37-39. Por. E. Post, Appendix do pracy 'Absolutely Unsolvable Problems and Relatively Undecidable Propositions', przedruk [w:] antologii Davisa, ss. 340-433. W przypisach o numerach 1 i 9 Post podkreśla konieczność zbadania *wszystkich* sposobów symbolizacji przez człowieka.
[34] Por. Post, 'Appendix', s. 426.
[35] Por. G. Kreisel, 'Which number theoretic problems can be solved in recursive progressions on Π^1_1 – paths through *O*?', *The Journal of Symbolic Logic*, 37(1972), ss. 311-334, szczególnie s. 316. Istnienie takiego systemu nie jest, według Kreisela, sprzeczne z twierdzeniem Gödla o niezupełności.
[36] Gandy, *op. cit.*, s. 124.

Wspomniana obiekcja GÖDLA dotyczyła TC w tym sensie mianowicie, iż sprzeciwiał się jej psychologicznej interpretacji. Uważał, że rezultaty uzyskane przy założeniu TC należy interpretować jako ograniczające możliwości czystego formalizmu w matematyce.[37] Analiza Turinga była dla niego wystarczającym wsparciem dla TC, ale zastrzegając interpretację TC w następującej postaci:

[GTC] **Każda funkcja (określona w liczbach naturalnych) obliczalna za pomocą mechanicznej procedury jest obliczalna przez maszynę Turinga**.[38]

Gödel uważał jednak, że możliwa jest efektywna komunikacja pomiędzy ludźmi odwołująca się do treści (sensu) wypowiedzi, a nie tylko do kombinatorycznych (czasowo-przestrzennych) relacji pomiędzy kombinacjami znaków.[39]
Powyższe rozważania miały za zadania wykazać, że Church arbitralnie założył [C2] podczas gdy w miejsce tej przesłanki należałoby wstawić

[C2']: *Relacja dowodliwości* ***dowolnego*** *systemu spełniającego* [C1] *jest rekurencyjna.*

Taka przesłanka ma już charakter empiryczny, gdyż jej prawdziwość jest uzależniona od faktycznych własności dowolnych, możliwych i odpowiednich systemów formalnych.

Przyjrzymy się obecnie argumentowi Turinga z pracy 'On Computable Numbers, with an Application to the Entscheidungsproblem':[40]

[T1] **Centralna Teza Turinga**: Ludzki komputer (komputor) spełnia trzy warunki ograniczające.
[T2] **Twierdzenie Turinga**: Dowolna funkcja, która może zostać obliczona przez komputor spełniający trzy ograniczenia, jest obliczalna przez maszynę Turinga.
[T3] **Wniosek (TC)**: Każda funkcja obliczalna przez komputor, jest obliczalna przez maszynę Turinga.[41]

Wspomniane trzy warunki to: ***determinizm*** (każda konfiguracja procesu obliczania określa jednoznacznie następny krok obliczeń), ***ograniczoność*** (na liczbę bezpośrednio rozpoznawalnych symbolicznych konfiguracji i na liczbę stanów wewnętrznych maszyny), ***lokalność*** (zmianie podlegają tylko bezpośrednio obserwowane konfiguracje, a odległość nowo obserwowanych konfiguracji od obserwowanych bezpośrednio jest ograniczona).

W analizie Turinga obliczalności przez człowieka (która okazała się być tak przekonywająca dla Gödla), a którą przyjął jako [T1] w swej argumentacji, mieści się pewien wątek empiryczny. Ograniczenia nałożone na komputor mają odpowiadać **rzeczywistym** ograniczeniom na obliczającego człowieka nałożonym przez naturę. Czy tak rzeczywiście jest? Gödel właśnie wysunął krytykę analizy Turinga odnoszącej się do skończonej liczby stanów wewnętrznych komputora. Gödel zwracał uwagę na to, że choć tych stanów jest skończenie wiele, lecz ich liczba zmierza do nieskończoności, bo umysł jest w swej istocie

[37] Por. 'Postscriptum' w antologii Davisa, ss. 71-73.
[38] Por. Hao Wang, 'From Mathematics to Philosophy', Routledge & Keagan, New York, 1974, s. 84.
[39] Por. K. Gödel, 'Über eine bisher noch nicht benützte Erweiterung des finiten Standpunktes', *Dialectica*, 12(1958), ss. 280-287. W rozmowie z Wangiem Gödel zauważył, że dla poprawności do argumentu Turinga należy dodać dwie przesłanki: 1. Nie istnieje umysł niezależny od materii; 2. Mózg funkcjonuje jak komputer cyfrowy. Por. Wang, *op. cit.*, s. 326.
[40] *Proceedings of the London Mathematical Society*, (2)42 (1936), ss. 230-265. Przedruk [w:] Davis, ss. 115-154.
[41] Shagrir, *op. cit.* ss. 224-226.

dynamiczny (constantly developing), a nie statyczny[42]. Ten zarzut Gödla ma wyraźnie empiryczną proweniencję, odwołuje się do **faktycznych** własności umysłu. Obecnie zwraca się uwagę na to, że analiza Turinga presuponuje *mechanistyczne założenie*.[43]

3) Teza Churcha a teza o niesprzeczność arytmetyki liczb naturalnych (TN).

Pomiędzy TC a jednym z głównych i nierozwiązanych problemów teorii arytmetyki liczb naturalnych – problemem jej niesprzeczności - istnieje podobieństwo, które postaram się wskazać. Przez arytmetykę liczb naturalnych rozumiem wszystkie teorie liczb naturalnych, które są rozszerzeniami arytmetyki *R*, badanej przez Tarskiego, Mostowskiego i Robinsona, w której reprezentowalne są wszystkie funkcje rekurencyjne (o ile *R* jest niesprzeczna).

- Zarówno TC jak i TN nie posiadają (absolutnych) dowodów.[44]
- Ich dostępne dowody (akceptowalne matematycznie) zakładają to czego dowodzą.
- Inne dowody odwołują się do intuicji.
- W obu przypadkach działa, właściwie jako jedyny, empiryczny argument z braku kontrprzykładu.

Church, w liście do Gödla z 27.07.1932 r. pisze:

> W rzeczywistości jedynym dowodem na niesprzeczność systemu *Principia Mathematica* jest **empiryczne świadectwo** (podkreślenie moje A.O.) pochodzące stąd, że system ten tak długo jest w użyciu, tak wiele twierdzeń zostało w nim wyprowadzonych i nikt nie znalazł sprzeczności.

- Duże działy matematyki (arytmetyka liczb naturalnych i teoria obliczalności) uprawia się przy założeniu tych tez.
- Ich fałszywość miałaby daleko idące konsekwencje.

Gdyby się okazało, że TN jest fałszywa, to wtedy badania nad arytmetyką okazałyby się całkowicie trywialne, choć twierdzenia udowodnione pozostałyby prawdziwe.
W przypadku TC wymienia się następujące konsekwencje jej prawdziwości: dla Computer Science – moc obliczeniowa maszyn jest ograniczona raz na zawsze przez TC; dla fizyki – TC ogranicza do rekurencyjnych relacje pomiędzy obserwablami[45]; dla filozofii – umysł matematyczny jest ograniczony przez rekursywność (Post); dla kognitywistyki – preferuje obliczeniowy model umysłu. Fałszywość TC miałaby np. takie skutki, że *Entscheidungsproblem* byłby nadal otwarty i dziesiąty problem Hilberta również.

- Społeczność matematyków i logików jest raczej przekonana co do prawdziwości zarówno TC jak i TN.

[42] Por. Davis, *op. cit*., s. 73.
[43] Por. Fitz, *op. cit.*, s. 22.
[44] Tarski, pytany, czy dowód niesprzeczności arytmetyki podany przez Gentzena (używający pozaskończonej indukcji do ε_0) wzmocnił jego przekonanie o niesprzeczności tej teorii, miał odpowiedzieć: *By about an epsilon*. Por. Neil Tennant, 'Deflationism and the Gödel Phenomena: Reply to Ketland', *Mind*, 114(453) (2005), s. 96.
[45] Dokładniej znaczy to, że [PTC] ogranicza do funkcji rekurencyjnych klasę funkcji obliczalnych przez systemy fizyczne. Zobacz fragment niniejszej pracy, gdzie referowano [PTC].

- Wiedza o niesprzeczności arytmetyki (czyli prawdziwość TN) ma charakter intuicyjny, również uzasadnienie dla TC odwołuje się do intuicji. Obie tezy są jakby *zagadnieniami brzegowymi* matematycznych teorii.

Kurt Gödel wprowadził rozróżnienie na matematykę obiektywną i matematykę subiektywną. Ta pierwsza - jest klasą (zbiorem?) zdań opisujących obiektywną rzeczywistość matematyczną, zaś druga - dotyczy zbioru zdań dowodliwych w ramach pewnego systemu formalnego. Korzystając z tej dystynkcji zdefiniujmy zbiór zdań należących do matematyki obiektywnej Z:

$$Z := \{T:\ \neg T \vdash \neg TC\},$$

gdzie T jest twierdzeniem lub hipotezą matematyki obiektywnej, zaś $\vdash$ jest relacją wyprowadzalności za pomocą logiki klasycznej[46]. Zbiór Z jest niepusty, gdyż TC oczywiście należy do Z. Kleene[47] wykazał, że pierwsze twierdzenie Gödla o niezupełności również należy do Z. Wydaje się, że twierdzenie Tarskiego o niedefiniowalności prawdy należy do zbioru Z[48]. Intuicyjnie rzecz ujmując, elementami zbioru Z są te twierdzenia matematyki obiektywnej, które zależą od TC. Do zbioru Z należą również **wszystkie** twierdzenia o (absolutnej) nierozstrzygalności, w tym twierdzenie o nierozstrzygalności logiki pierwszego rzędu[49].
Problemem otwartym jest to, czy TN należy do zbioru twierdzeń Z. TN jest twierdzeniem matematyki obiektywnej.[50]

4) Zarys koncepcji empirycznego badania pojęć, w szczególności TC.

W niniejszej pracy przyjęto za podstawowe dość szczególne sformułowanie TC, oparte na identyczności dwóch pojęć. Przypuszczam, że sformułowania TC w postaci [MTC] oraz [FTC] wywodzą się z potrzeby uczynienia [TC] bardziej podatną na intersubiektywny wgląd. Pojęcia są niezbyt wdzięcznym przedmiotem badań, gdyż niewiele o nich wiadomo. Ich istnienie wydaje się być bezsporne. Jeśli przyjąć, że TC ma mieć charakter empiryczny, to należy wskazać jakąś intersubiektywną drogę docierania do pojęć. Analiza pojęć często obarczona jest pewną dozą subiektywizmu, a uzewnętrznienie pojęć w postaci, na przykład, systemu formalnego związane jest z pytaniem o adekwatność ujęcia. Jest to niejako droga z umysłu do form naoczności, używając terminologii Kanta. Możliwa wydaje się także droga odwrotna, polegająca na *indukowaniu* pojęć. Weźmy dla przykładu pojęcie *implikacji materialnej*. Przeciętny użytkownik języka polskiego ma jakieś pojęcie tego spójnika. Spróbujmy pokazać w jaki sposób można będzie indukować w umyśle pojęcie implikacji materialnej. Ustalmy jakiś system formalny (w języku z jednym funktorem dwuargumentowym) charakteryzujący spójnik implikacji. Weźmy następnie jakiś program pozwalający na generowanie tez tego systemu wraz z dowodami. Program ten dodatkowo ma liczyć wszystkie użycia formuł pojawiające się w dowodach. Podobnie postąpił Frege (dla logiki pierwszego rzędu) w zakończeniu swojego *Begriffschrift*. Dowodzi tam 133 tez logiki i

[46] Jednakże, o ile mi wiadomo, nie istnieje ogólna definicja relacji wynikania pomiędzy zdaniami matematyki obiektywnej.
[47] S. C. Kleene, 'Reflection on Church's Thesis', *Notre Dame Journal of Formal Logic*, 28(1987), s. 491.
[48] Por. Adam Olszewski, 'Teza Churcha a definicja prawdy Tarskiego', *Analecta Cracoviensia*, 33(2001), ss. 171-176.
[49] Jest tutaj pewna subtelność. Jeśli weźmiemy pod uwagę twierdzenia o nierozstrzygalności w sensie nierekurencyjności, to rozważane twierdzenia nie należą do zbioru Z.

podaje liczbę użycia poszczególnych formuł w dowodach tez. I tak dla przykładu, jego teza o numerze 5 (stosuję współczesną notację):

$$((q \to r) \to ((p \to q) \to (p \to r)));$$

pojawia się w dowodach piętnastu tez. Natomiast teza o numerze 3:

$$((p \to q) \to ((r \to (p \to q)) \to ((r \to p) \to (r \to q))));$$

pojawia się tylko w dowodzie jednej tezy. Dla zbudowania pojęcia implikacji intuicyjnie widać różnicę w istotności tych dwóch przytoczonych tez. Pierwsza z tez jest skomutowanym sylogizmem hipotetycznym, a zatem nawiązuje do bardzo ważnej własności implikacji, polegającej na przechodniości tego funktora, podczas gdy druga teza jest zaledwie poprzedzonym sylogizmem Fregego. Sam sylogizm Fregego jest ważną tezą wyrażającą intuicyjnie rozdzielność implikacji względem siebie samej, i pojawia się w wyliczeniu Fregego pięciokrotnie. Indukcja pojęcia implikacji materialnej miałaby polegać na informowaniu jakiegoś osobnika (który takiego pojęcia nie posiada) o tych własnościach implikacji, które najczęściej występują w dowodach tez. Przypuszczenie (empirycznie weryfikowalne) jest takie, że istnieje taka **skończona** liczba tez implikacyjnych, pozwalająca zbudować pojęcie implikacji. Liczba ta jest graniczna w tym sensie, że zapoznanie się z innymi tezami nie zmienia już samego pojęcia implikacji – jest ono już gotowe (ustalone). Empiryczna weryfikowalność mogłaby polegać na tym, że użytkownik (mający tak zbudowane pojęcie implikacji) umiałby bezbłędnie rozpoznawać to, czy inne formuły języka systemu, które mu nie zostały wcześniej przestawione są, czy też nie są własnościami implikacji.

Stosując ten pomysł do pojęcia funkcji rekurencyjnej należałoby policzyć użycie pewnych funkcji użytych do zdefiniowania innych. Idąc tym tropem można by indukować ogólne pojęcie funkcji rekurencyjnej. Tak samo można by potraktować funkcje efektywnie obliczalne. Indukowanie może mieć tę zaletę, że nie potrzeba znajomości **wszystkich** funkcji dla zbudowania ogólnego pojęcia takiej funkcji. Wystarczą funkcje najbardziej charakterystyczne.