

**Marzena Fornal\***

Akademia Humanistyczno-Ekonomiczna w Łodzi

 <https://orcid.org/0000-0002-0263-9081>e-mail: [marzena.fornal@gmail.com](mailto:marzena.fornal@gmail.com)

## Czy istnieją czynności umysłu, których nie można reprezentować za pomocą maszyn Turinga?

[https://doi.org/10.25312/2083-2923.23\\_08mfo](https://doi.org/10.25312/2083-2923.23_08mfo)

**Streszczenie:** Artykuł stanowi próbę odpowiedzi na pytanie, czy jesteśmy w stanie wskazać takie czynności umysłu, które nie są możliwe do reprezentacji za pomocą maszyny Turinga? Jest to zatem pytanie o to, czy wszystkie nasze stany mentalne posiadają obliczeniową naturę. Problem ten będzie rozważany w odniesieniu do tak zwanego argumentu gödłowskiego, opierającego się na dwóch twierdzeniach Gödla: 1) o niezupełności oraz 2) o niedowodliwości niesprzeczności, skierowanego przeciwko obliczeniowym teoriom umysłu. Argument ten w wersji zaprezentowanej przez Johna Randolpha Lucasa zostanie poddany krytycznej analizie, która doprowadzi do pozytywnych wniosków zawartych w końcowej części artykułu.

**Słowa kluczowe:** komputacjonizm, obliczeniowa teoria umysłu, argument gödłowski, John Randolph Lucas, Roger Penrose, maszyna Turinga, Kurt Gödel

---

\* Marzena Fornal – doktor nauk humanistycznych, absolwentka filozofii na Uniwersytecie Łódzkim. Specjalizuje się w filozofii umysłu, w szczególności we współczesnych koncepcjach wywodzących się z filozofii analitycznej, kognitywistyki oraz szeroko rozumianych neuronauk. W swoich badaniach autorka szczególną uwagę poświęca problemowi świadomości, jej statusowi ontycznemu, genezie i możliwościom jej naukowego poznania oraz problemowi psychofizycznemu w teorii emergencji. Uczestniczka wielu konferencji i autorka publikacji z zakresu filozofii umysłu.

## Wstęp

Starając się odpowiedzieć na pytanie zawarte w temacie, pozwolę sobie ograniczyć się jedynie do tak zwanego argumentu gödłowskiego, stosowanego przez wielu przeciwników mechanicystycznej teorii umysłu<sup>1</sup>. W ramach wstępu niezbędne jest przedstawienie dwóch twierdzeń Gödla<sup>2</sup> kluczowych dla tej linii argumentacyjnej. Twierdzenie pierwsze o niezupełności mówi o tym, że dla żadnego systemu formalnego, zawierającego przynajmniej prostą arytmetykę (tj. liczby naturalne oraz działania dodawania i mnożenia), nie jest możliwe wygenerowanie zbioru wszystkich twierdzeń prawdziwych (dla tego systemu) w obrębie tego systemu. Innymi słowy, zbiór zdań dowodliwych tego systemu nie pokrywa się ze zbiorem zdań prawdziwych. Jeżeli zbiór zdań dowodliwych jest mniejszy od zbioru zdań prawdziwych to mamy do czynienia z systemem niesprzecznym (takim, w którym żadne zdanie fałszywe nie może być dowiedzione). Zgodnie z tym dla systemu takiego będą istniały zdania prawdziwe niedowodliwe, czyli takie, które z metapoziomu (spoza systemu) uznajemy jako prawdziwe, ale nie jesteśmy w stanie ich dowieść wewnątrz tego systemu. Z kolei, gdy zbiór zdań dowodliwych jest większy, oznacza to, że system taki jest spreczny<sup>3</sup>.

Drugie twierdzenie Gödla o niedowodliwości niesprzeczności jest konsekwencją pierwszego. Zakłada ono, że w ramach żadnego systemu formalnego zawierającego przynajmniej prostą arytmetykę nie da się dowieść jego niesprzeczności. Aby taki dowód przeprowadzić, niezbędny byłby szerszy system, którego jednak niesprzeczności w ramach niego samego również nie można byłoby dowieść. W tym celu konieczna byłaby budowa kolejnego metapoziomu i tak *ad infinitum*.

## Lucas i Penrose przeciwko komputacjonizmowi

Oba twierdzenia Gödla możemy zastosować do maszyn cybernetycznych (obliczeniowych), równoważnych maszynom Turinga, ponieważ w rzeczywistości są one konkretną realizacją danego systemu formalnego. Owa stosowalność tych twierdzeń dla maszyn Turinga stanowi rdzeń argumentacji Lucasa i Penrose'a przeciwko mechanicystycznym teoriom umysłu, a zwłaszcza przeciwko komputacjonizmowi.

---

<sup>1</sup> Jednym z pominiętych zagadnień w tej pracy jest tak zwany problem stopu, który związany jest ze sprawą przewidywania przyszłości w systemach mechanicznych. Nie istnieje maszyna Turinga rozstrzygająca w skończonej liczbie kroków, czy dowolna maszyna Turinga zakończy pracę. Analiza tego problemu, jak się wydaje, mogłaby (przynajmniej pozornie) również dostarczyć odpowiedzi na pytanie zawarte w temacie pracy.

<sup>2</sup> K. Gödel, *Über formal unentscheidbare Sätze der 'Principia Mathematica' und verwandter Systeme*, „Monatshefte für Mathematik und Physik”, Vol. 37, s. 173–198.

<sup>3</sup> Oczywiście, można rozszerzyć definicję systemów formalnych tak, że twierdzenie Gödla nie będzie dla nich zachodzić. Jednak takie niestandardowe systemy nigdy nie będą rozstrzygalne, to znaczy ich algorytm wnioskowania nie dałby się zaprogramować na maszynie Turinga lub ich zbiór aksjomatów nie dałby się taką maszyną wygenerować.

Argument Penrose'a<sup>4</sup> jest w rzeczywistości „odświeżoną” wersją argumentu Lucasa, jednak aby tę linię argumentacyjną uznać za uprawnioną, jesteśmy zobligowani do przyjęcia kilku silnych tez metafizycznych, w tym: 1. Niektóre stany umysłowe nie mają charakteru obliczeniowego (algorytmicznego), mogą być jednak wyjaśnione za pomocą teorii fizycznych (głównie mechaniki kwantowej i teorii względności) – teza redukcjonizmu fizykalistycznego, 2. Istnieją przynajmniej trzy poziomy ontologiczne, a jednym z nich jest poziom bytów matematycznych istniejących realnie – teza realizmu matematycznego. W dalszej części pracy będę odnosić się jedynie do argumentacji Lucasa, gdyż moim zdaniem założenia początkowe tej linii argumentacyjnej są bardziej neutralne metafizycznie.

Przypomnijmy, że główną tezę komputacjonistycznej teorii umysłu jest twierdzenie, że maszyna licząca o stanach nieciągłych (dyskretnych), taka jak uniwersalna maszyna Turinga<sup>5</sup>, jest dobrym narzędziem do naśladowania zdolności ludzkiego umysłu. W skrajnej wersji twierdzi się zaś, że uniwersalna maszyna Turinga jest równoważna ludzkiemu umysłowi<sup>6</sup>. Częstokroć zwolennicy komputacjonizmu, korzystając z tak zwanej analogii komputerowej, określają umysł jako software (tudzież mindware), czyli oprogramowanie, zaś mózgi – jako narzędzie do realizacji owego oprogramowania, czyli hardware. Analogia ta bierze się stąd, że komputer cyfrowy jest uznawany za uniwersalną maszynę Turinga, tj. realizuje pewien konkretny system formalny, jednak może implementować dowolny system formalny realizowalny przez inny komputer cyfrowy. Kolejnym wnioskiem, jaki wyciągnęli komputacjoniści z dostrzeżenia analogii pomiędzy funkcjonowaniem maszyny cyfrowej a procesem kognicji, było stwierdzenie, że procesy poznawcze człowieka mają charakter obliczeniowy (algorytmiczny), co znaczy, że procesy te mogą zostać wyjaśnione za pomocą skończonego ciągu symboli powiązanych ze sobą za pomocą reguł zaimplementowanych w systemie przetwarzającym dane. Jest to wniosek o charakterze mechanicystycznym i redukcjonistycznym – utożsamiający stany mentalne ze stanami układu fizycznego.

## Argument Lucasa – analiza krytyczna

Lucas zakładał, że umysł jest niealgorytmiczny, ale niesprzeczny, przez co nie może być symulowany adekwatnie za pomocą maszyny cyfrowej. Jednak we współczesnych dyskusjach w filozofii umysłu nie jest to paradygmat jedyny uprawniony. Dopuszczalne są cztery możliwości: 1. Umysł jest algorytmiczny i niesprzeczny; 2. Umysł jest

<sup>4</sup> R. Penrose, *Nowy umysł cesarza*, Wydawnictwo Naukowe PWN, Warszawa 1995.

<sup>5</sup> A. Turing, *Maszyna licząca a inteligencja*, [w:] B. Chwedeńczuk (red.), *Filozofia umysłu*, Fundacja „Aletheia”: „Spacja”, Warszawa 1995.

<sup>6</sup> Zapożyczając terminologię od Johna Searle’a, pierwsza teza jest tezą słabej sztucznej inteligencji (słabej AI), druga z kolei to teza mocnej sztucznej inteligencji (mocnej AI). Zob. J. Searle, *Umysły, mózgi i programy*, [w:] B. Chwedeńczuk (red.), *Filozofia umysłu*, Fundacja „Aletheia”: „Spacja”, Warszawa 1995.

algorytmiczny i sprzeczny; 3. Umysł jest niealgorytmiczny i niesprzeczny; 4. Umysł jest niealgorytmiczny i sprzeczny. Zwolennicy komputacjonizmu i innych teorii mechanicznych będą opowiadać się za możliwością pierwszą. Przeciwnicy argumentu gödłowskiego starają się zazwyczaj wykazać, że: a) zarówno umysły, jak i maszyny podlegają wynikającym z twierdzeń Gödla ograniczeniom (Dennett<sup>7</sup>, Wang<sup>8</sup>) albo b) umysły ludzkie są sprzeczne, przez co twierdzenia limitacyjne Gödla nie są doń stosowalne (Putnam<sup>9</sup>, Turing<sup>10</sup>). Strategia a) opiera się głównie na wykazaniu trudności w dowodzeniu niesprzeczności umysłu. Jak się wydaje, strategia b) nie wyklucza całkowicie symulacji umysłu przy pomocy komputera, choć sprzeczność jest tutaj równoważna błędemu algorytmowi.

Argumentację Lucasa możemy streścić w czterech тезach:

1. Argument stosuje się do maszyn „cybernetycznych”, czyli takich, które równoważne są systemom formalnym. Każda taka maszyna ma skończoną liczbę stanów i interakcji. System formalny, jaki jest realizowany przez taką maszynę, jest dany poprzez wypisane w określonym języku formalnym aksjomaty i reguły wnioskowania. Obliczanie, czyli sekwencja operacji dokonywanych przez daną maszynę, odpowiada takiemu systemowi formalnemu<sup>11</sup>.
2. Jeżeli jakaś maszyna cybernetyczna ma być modelem ludzkiego umysłu, to „musi zawierać mechanizm, który może oznajmiał prawdy arytmetyki”<sup>12</sup>. Wyrażenia, które maszyna może „przedłożyć, jako prawdziwe”, odpowiadają twierdzeniom systemu.
3. Możemy wskazać formułę Gödla, która nie będzie dowodliwa w tym systemie. Musimy przy tym założyć, że system, o którym mowa, jest niesprzeczny, a przynajmniej jego część arytmetyczna. Gdybyśmy założyli jego sprzeczność, to nie tylko formuła Gödla, ale każde wyrażenie byłoby dowodliwe w tym systemie. Co więcej, maszyna, która realizowałaby taki sprzeczny system, nie mogłaby zostać uznana za model ludzkiego umysłu, gdyż Lucas przyjmuje tezę, że ludzki umysł jest niesprzeczny (a przynajmniej jego część arytmetyczna).

<sup>7</sup> Zob. D. Dennett, *Consciousness Explained*, Little, Brown and Company, Boston 1991; tenże, *Darwin Dangerous Idea. Evolution and The Meanings of Life*, Simon & Schuster Paperbacks, New York–Toronto–London–Sydney 1995.

<sup>8</sup> Zob. H. Wang, *A Logical Journey: From Gödel to Philosophy*, A Bradford Book, 1997.

<sup>9</sup> H. Putnam, *The Nature of Mental States*, [w:] tegoż, *Mind, Language and Reality*, Vol. 2, Harvard University Press, Cambridge 1975.

<sup>10</sup> A. Turing, *On Computable Numbers, with an Application to the Entscheidungsproblem*, „Proceedings of the London Mathematical Society” 1936, Vol. 42(1).

<sup>11</sup> Uznając, że maszyna taka jest adekwatnym modelem umysłu, stajemy przed problemem zdeterminowania. Albowiem każda maszyna cyfrowa jest programowana na podstawie zestawu instrukcji, a więc jej zachowanie jest z góry zdeterminowane przez zestaw reguł początkowych, a co za tym idzie jej wszelkie działania są możliwe do przewidzenia. Nawet wprowadzenie czynnika losowości nie usuwa tego problemu, wszak losowy wybór spośród alternatyw i tak jest skończonym zbiorem określonych działań, więc z góry wszystkie możliwości jesteśmy w stanie przewidzieć.

<sup>12</sup> J.R. Lucas, *Umysły, maszyny i Gödel*, „Hybris” 2009, nr 8, s. 101.

4. Ludzki umysł widzi prawdziwość formuły Gödla. Co więcej, możemy prześledzić konstrukcję takiej formuły i przekonać się o jej niedowodliwości w systemie oraz o jej prawdziwości. Prawdziwość jej jest w istocie wynikiem, a nawet wyrazem jej niedowodliwości. Nasz umysł umie więc coś, czego nie umie maszyna. Nie da się symulować maszynowo wszystkich czynności umysłu naraz. Umysł nie może być równoważny jakiegokolwiek maszynie<sup>13</sup>.

Lucas tym samym stara się dowieść, że ludzie uposażeni są w specjalną zdolność, którą Penrose nazywa intuicją matematyczną. Podstawową cechą intuicji matematycznej jest jej niealgorytmiczność, wszak ściśle związana jest ona z rozumieniem, a rozumienie nie polega na wykonywaniu algorytmu. Jeżeli jednak chcemy uznawać istotę ludzką za istotę rozumną, a więc też racjonalną, musimy przyjąć twierdzenie o jej niesprzeczności.

Jak pisze Lucas: „Gdybyśmy rzeczywiście byli maszynami sprzecznymi, powinniśmy pozostać zadowoleni z naszych sprzeczności i ochoczo uznawać oba ich członki<sup>14</sup>”. Umysły ludzkie są selektywne: choć omylne, zawsze dążą do wybrania twierdzeń, które wydają się prawdziwe i z reguły odrzucają twierdzenia, które wydają się fałszywe – w tym ujawnia się nasze dążenie do poznania prawdy, poszukiwania wiedzy jasnej i wyraźnej.

Ludzki umysł jako omylna, lecz samokorygująca się maszyna nadal jednak podlega ograniczeniom wynikającym z twierdzenia Gödla. Jeżeli umysł rzeczywiście jest niespreczny, to nie jest w stanie przeprowadzić formalnego dowodu swej własnej niesprzeczności. Nie istnieją jednak racje ku temu, aby twierdzić, że nie jest w stanie wyjść poza wewnętrzny system formalny i budować wnioski o swej niesprzeczności w sposób nieformalny, na przykład poprzez zdroworozsądkowe stwierdzenia. Takie zdroworozsądkowe stwierdzenia, co prawda, nigdy nie będą mogły zostać sformalizowane całkowicie, lecz to właśnie podtrzymuje limitacje nałożone przez Gödla. Od refleksyjności ludzkiego umysłu oczekujemy, że będzie mógł orzekać o sobie samym. Zatem stwierdzenia o własnej niesprzeczności są stwierdzeniami jak najbardziej uprawomocnionymi. Co więcej, nie tylko uprawomocnionymi, lecz również pożądanymi, bo tylko przy założeniu niesprzeczności ludzkiego umysłu jakakolwiek nauka jest w ogóle możliwa. Tezy „ja jestem niespreczny” nie jesteśmy w stanie dowieść, jednakowoż sensowne jest zakładać jej prawdziwość. To, co odróżnia nas od maszyn, to zdolność do wydawania samozwrotnych sądów. Samoświadomość bowiem polega na zdolności orzekania o swoich własnych procesach. Istota świadoma potrafi poradzić sobie z formułą gödłowską właśnie dlatego, że jest zdolna do samozwrotnego odniesienia, a więc potrafi rozważać siebie samą, swoje działanie i swoje myśli<sup>15</sup>.

<sup>13</sup> Zob. S. Krajewski, *Twierdzenie Gödla i jego interpretacje filozoficzne*, Wydawnictwo Instytutu Filozofii i Socjologii PAN, Warszawa 2003, s. 10–11.

<sup>14</sup> J.R. Lucas, dz. cyt., s. 109.

<sup>15</sup> Teza mechanicyzmu stara się traktować umysł redukcyjnie, twierdząc, że umysł działa w oparciu o mechaniczne zasady, a więc jest maszyną, której działanie jest zdeterminowane przez

Zatem jeżeli chcielibyśmy, opierając się na argumentach Lucasa, odpowiedzieć na pytanie: Czy istnieją czynności umysłu, których nie można reprezentować za pomocą maszyny Turinga? – to odpowiedź brzmiałaby: a) umysł ludzki posiada tak zwaną intuicję matematyczną oraz operuje pojęciami semantycznymi (posiada przynajmniej pojęcie prawdy), maszyna zaś dokonuje jedynie operacji syntaktycznych na symbolach formalnych; b) umysł ludzki posiada świadomość, która umożliwia mu wyprowadzanie zdań prawdziwych samowrotnych, czyli sądów o sobie samym, maszyna zaś tego nie potrafi.

Przeanalizujmy jednak trudności, jakie wiążą się z tymi czterema przesłankami. Co do przesłanki pierwszej – pomimo niejasności związanych z zakresem pojęcia *maszyna* możemy przyjąć, że argumentacja Lucasa odnosi się przynajmniej do szerokiej klasy maszyn, które są równoważne maszynie Turinga. Wnioskowanie Lucasa może zostać obalone, jeżeli wykażemy, że istnieją maszyny, które nie spełniają wyżej podanej definicji, na przykład poprzez wskazanie jakiegoś ogólniejszego pojęcia maszyny. Jeśli chodzi o przesłankę drugą, mamy tutaj do czynienia z problemem natury pojęciowej. Lucas bowiem naprzemiennie stosuje pojęcia *dowodliwość* oraz *prawdziwość*. Pojęcie prawdziwości stosuje zarówno do maszyn, jak i do ludzi, jednak w tym przypadku człowiek, który stoi „na zewnątrz” maszyny, dysponuje zewnętrznym pojęciem prawdy, które jest silniejsze od tego, jakim dysponuje maszyna (gdzie dowodliwość jest równoważna prawdziwości). Czy można zatem odnosić w sposób sensowny „taki rodzaj prawdziwości” do maszyny? Jak się wydaje, kwestia ta związana jest z problemem, na który wskazywał John Searle w swoim słynnym argumentie „chińskiego pokoju”, mianowicie czy maszyna może operować pojęciami semantycznymi, takimi jak pojęcie prawdy. Jeżeli założymy, że jedynie ludzki umysł ma dostęp do semantycznej treści pojęć, jako jedyny rozumie sens wyrażań, maszyna zaś operuje tylko na symbolach (operacje syntaktyczne bez semantyki), to teza mechanicystyczna zostaje obalona bez potrzeby odnoszenia się do twierdzenia Gödla.

Oдноśnie do przesłanki trzeciej musimy wskazać na dwa problemy: a) skąd mamy pewność, że dana maszyna jest niesprzeczna; b) skąd mamy pewność, że ludzki umysł jest niesprzeczny. Co tyczy się maszyn cyfrowych, jedynym sposobem na ustalenie, czy maszyna taka jest niesprzeczna, jest jej dokładne sprawdzenie. Przy sprawdzaniu takiej maszyny zawsze jednak istnieje możliwość, że jakaś sprzeczność nie została wykryta. Oдноśnie do niesprzeczności ludzkich umysłów, to podobnie jak w przypadku maszyn nigdy nie możemy być w pełni pewni, że udało nam się wyeliminować niesprzeczność naszych twierdzeń. Przyjęcie tezy o sprzeczności umysłu wydaje się jednak podważać paradygmat, jakoby procesy umysłowe były związane ściśle z logiką klasyczną, co z kolei budzi naturalny sprzeciw. Jednym z argumentów za przyjęciem tezy o sprzeczności umysłu jest zaobserwowane wśród psychologów ewolucyjnych

---

działania jej części, stanu początkowego i budowy. Lucas zaś starał się wykazać, że ludzki umysł jest czymś emergentnym i niepodlegającym redukcji – nie może być on traktowany jako prosta suma swoich części. Zob. J.R. Lucas, dz. cyt., s. 116.

zjawisko samooszukiwania się (ang. *self deception*). Z logicznego punktu widzenia w procesie samooszukiwania osoba twierdzi zarazem, że  $p \wedge \neg p$ , a zatem zgodnie z prawem logicznym ( $p \wedge \neg p \rightarrow q$ ) z jej przekonań wynika zdanie dowolne. Tak jednak się nie dzieje, ponieważ cały ten proces (który cechuje automatyzm) odbywa się nieświadomie. Nie można tutaj mówić o sądach w sensie logicznym, chociaż można mówić o pewnym przetwarzaniu danych, gdyż proces ten zachodzi na poziomie układu nerwowego. Zaznaczyć należy, że posiadanie sprzecznych przekonań nie jest na pewno tym samym, co sprzeczność całego umysłu, jednak podobnie jak w przypadku systemów formalnych pomiędzy sprzecznościami zachodzić musi istotny związek. Wbrew przypuszczeniom posiadanie sprzecznych przekonań nie doprowadza do przepełnienia systemu, to znaczy ani nie doprowadza do zawieszenia takiego systemu (system umie sobie radzić z pojawiającymi się sprzecznościami, chociaż posiada normatywną regułę ich unikania), ani nie doprowadza do sytuacji, gdzie system uznaje jakiegokolwiek zdania za prawdziwe<sup>16</sup>. Zwolennicy koncepcji sprzeczności umysłu muszą jednak uporać się z przynajmniej dwiema kwestiami: 1. Jakiego typu mechanizmy generują sprzeczność? 2. Jaka logika powinna zastąpić logikę klasyczną? Czy miałyby to być któraś z logik parakonsystentnych?<sup>17</sup>.

W związku z przesłanką czwartą należy zauważyć, że prawdziwość formuły Gödla jest kwestią naszego założenia, nie zaś jakiegoś uprzywilejowanego wglądu. Problem prawdziwości formuły Gödla odnosi się bezpośrednio do kwestii tego, czy wiemy, że rozważany system jest niesprzeczny. Jak się wydaje, nie jesteśmy w stanie tego udowodnić w sposób pewny. Lucas jako zwolennik metafizycznego zastosowania twierdzenia Gödla uważa, że zdanie Gödla jest niedowodliwe w omawianym systemie, dla którego jest ono skonstruowane, ale „my, stojąc na zewnątrz systemu, potrafimy zobaczyć jego prawdziwość<sup>18</sup>”. Jak się jednak wydaje, dostrzeżenie prawdziwości formuły Gödla nie jest jakąś specyficzną ludzką zdolnością, lecz zwyczajną prawdziwością matematyczną. Możemy bowiem stworzyć maszynę będącą na zewnątrz maszyny wyjściowej, która będzie w stanie dowieść prawdziwości formuły Gödla<sup>19</sup>.

<sup>16</sup> Sprzeczność dostrzegalna jest również pomiędzy różnymi obszarami aktywności intelektualnej człowieka. Zdarza się nierzadko, że jedna osoba dysponuje sprzecznymi zestawami przekonań, które odnosić się będą do dwóch różnych dziedzin intelektualnej działalności człowieka. Przykładowo, może ona przyjmować za prawdziwe jednocześnie zbiory prawd religijnych oraz zbiór twierdzeń naukowych, które wydają się wzajemnie się wykluczać (tak zwana teoria dwu prawd).

<sup>17</sup> Por. W. Grygiel, *Jak uniesprzecznicić sprzeczność umysłu?*, „Zagadnienia Filozoficzne w Nauce” 2010, nr XLVII.

<sup>18</sup> J.R. Lucas, dz. cyt., s. 107.

<sup>19</sup> Krajewski wskazuje na jeszcze jeden niepokojący element w argumentcie Lucasa. Mianowicie stosuje się on tylko do ludzi logicznie wykształconych. Skoro kluczowym krokiem jest stwierdzenie prawdziwości zdania Gödla, to ile osób wchodzi w grę? Zob. S. Krajewski, dz. cyt., s. 33.



## Wnioski

Podsumowując, nie sędzę, aby argument gödłowski mógł rozstrzygnąć kwestię sprzeczności/niesprzeczności umysłu. Jeżeli umysł jest niesprzeczny, to zgodnie z twierdzeniem limitacyjnym, nie jesteśmy w stanie tego dowieść. Jednocześnie nie możemy wykluczyć nie tylko pojedynczych sprzeczności, ale również sprzeczności globalnej umysłu. Zakładając, że ludzki umysł ma budowę modułową, taka globalna sprzeczność mogłaby pojawiać się na poziomie integracji lokalnych modułów obliczeniowych, a więc dla nas może pozostawać sprzecznością nieujawniającą się. Nie wydaje się również, aby Lucas całkowicie obalił tezę mechanicystyczną, udowadniając, że niektórych procesów umysłowych po prostu nie można z powodzeniem symulować na maszynach cyfrowych<sup>20</sup>. Jeżeli teza o niealgorytmiczności umysłu okazałaby się słuszna, wskazywałoby to na to, że w świecie istnieje pewien obszar, który opiera się cyfrowej reprezentacji. Do podobnych wniosków doszedł już sam Alan Turing, który analizując moc swoich maszyn, uznał, że istnieją funkcje, których wartości nie są w stanie obliczyć. To jednak motywuje do poszukiwania innych metod obliczania, takich jak chociażby sieci neuronowe, które mogłyby okazać się adekwatnym modelem ludzkiego umysłu.

Zgadzam się natomiast z Lucasem co do niezbędności intuicji w matematyce i w ogóle w myśleniu ludzkim – jest to coś, co decyduje o swoistości ludzkiego umysłu. Równie istotna jest refleksyjność i zdolność do samoodnoszenia się. Nie wydaje mi się jednak, aby zastosowanie argumentu gödłowskiego było niezbędne do zwrócenia uwagi na te właśnie konstytutywne cechy ludzkiego umysłu.

## Bibliografia

- Dennett D., *Consciousness Explained*, Little, Brown and Company, Boston 1991.
- Dennett D., *Darwin Dangerous Idea. Evolution and The Meanings of Life*, Simon & Schuster Paperbacks, New York–Toronto–London–Sydney 1995.
- Gödel K., *Über formal unentscheidbare Sätze der 'Principia Mathematica' und verwandter Systeme*, „Monatshefte für Mathematik und Physik”, Vol. 37.
- Grygiel W., *Jak uniesprzecznici sprzeczność umysłu?*, „Zagadnienia Filozoficzne w Nauce” 2010, nr XLVII.
- Krajewski S., *Twierdzenie Gödla i jego interpretacje filozoficzne*, Wydawnictwo Instytutu Filozofii i Socjologii PAN, Warszawa 2003.
- Lucas J.R., *Umysły, maszyny i Gödel*, „Hybris” 2009, nr 8.
- Penrose R., *Cienie umysłu. Poszukiwanie naukowej teorii świadomości*, Zysk i S-ka, Poznań 2000.

---

<sup>20</sup> Choć zgadzam się, że nie może tego uczynić maszyna Turinga.



- Penrose R., *Nowy umysł cesarza*, Wydawnictwo Naukowe PWN, Warszawa 1995.
- Putnam H., *The Nature of Mental States*, [w:] tegoż, *Mind, Language and Reality*, Vol. 2, Harvard University Press, Cambridge 1975.
- Searle J., *Umysły, mózgi i programy*, [w:] B. Chwedeńczuk (red.), *Filozofia umysłu*, Fundacja „Aletheia”: „Spacja”, Warszawa 1995.
- Turing A., *Maszyna licząca a inteligencja*, [w:] B. Chwedeńczuk (red.), *Filozofia umysłu*, Fundacja „Aletheia”: „Spacja”, Warszawa 1995.
- Turing A., *On Computable Numbers, with an Application to the Entscheidungsproblem*, „Proceedings of the London Mathematical Society” 1936, Vol. 42(1).
- Wang H., *A Logical Journey: From Gödel to Philosophy*, A Bradford Book, 1997.

### **Are there activities of the mind that cannot be represented by Turing machines?**

**Abstract:** The article is an attempt to answer the question whether we are able to identify such activities of the mind that are not possible to be represented by a Turing machine? Thus, it is a question of whether all our mental states have a computational nature. This problem will be considered in relation to the so-called Gödel argument, based on two Gödel theorems: 1. on incompleteness and 2. on the incompleteness of non-contradiction, directed against computational theories of mind. This argument, as presented by John Randolph Lucas, will be critically analyzed, which will lead to positive conclusions in the final part of the article.

**Keywords:** computationalism, computational theory of mind, Gödelian argument, John Randolph Lucas, Roger Penrose, Turing machine, Kurt Gödel

### **About the Author**

Marzena Fornal – doctor of humanities, graduate of philosophy at the University of Lodz. She specializes in the philosophy of mind, in particular in contemporary concepts arising from analytical philosophy, cognitive science and broadly understood neurosciences. She devoted special attention to the problem of consciousness, its ontic status, genesis, and the possibilities of its scientific cognition as well as psychophysical problem in theory of emergence. Participant of many conferences and author of publications in the field of philosophy of mind.