

Wojciech Słomski

Krytyka komputacyjnych teorii umysłu Podejście Rogera Penrose'a

Uwagi wstępne

Nie wymaga wielkiej przenikliwości konstatacja, że w ramach współczesnej filozofii analitycznej trwają ożywione prace badawcze w zakresie dwu wzajemnie przenikających się dziedzin: filozofii umysłu oraz filozofii języka. Badania te posilowane często analizami z zakresu teorii obliczalności, językoznawstwa matematycznego, poznawczo zorientowanej psychologii czy informatyki stawiają sobie za jeden z głównych celów odpowiedź na pytanie o sposób funkcjonowania i własności umysłu, warunki intencjonalnego posługiwania się językiem, a także związek pomiędzy umysłem i światem zewnętrznym oraz sposobem opisu tego świata. Jest przy tym dość oczywiste, że zarysowane problemy nie są nowe – korzeniami sięgają one filozofii Platona i Arystotelesa. Jednakże, jak podkreśla Urszula Żegleń, „związek dzisiejszej filozofii umysłu z dyscyplinami wchodzącymi w zakres nauk kognitywnych powoduje, że w rozmaity sposób formułowane i rozstrzygane w ciągu dziejów filozofii problemy umysłu (w tym także świadomości) zastąpione zostały nowymi pytaniami i rozwiązaniami, podawanymi z pozycji nauk empirycznych, czy w jakimś uzgodnieniu z jej wynikami”¹.

Spojrzenie na wyniki, jakie uzyskują teoretycy umysłu i języka w ostatnich trzydziestu latach pozwala stwierdzić, że jednym z domi-

¹ U. Żegleń: *Wprowadzenie do problematyki filozofii umysłu*. „Kognitywistyka i Media Edukacji” 1998, nr 1, s. 113.

nujących podejść do zagadnienia sposobu funkcjonowania i własności umysłu było powiązanie refleksji filozoficznej z danymi dostarczonymi przez nauki komputerowe i badania logiczno-matematyczne², czy szerzej – badania nad sztuczną inteligencją. Jak pisze Włodzisław Duch: „Badania nad sztuczną inteligencją doprowadziły do nowego spojrzenia na umysł jako maszynę do przetwarzania informacji, wpływając zarówno na powstanie filozofii kognitywnej, jak i psychologii kognitywnej [...]. Systemy przetwarzające informacje, takie jak ludzkie umysły, powinny [zgodnie z tym ujęciem – W.S.] dać się zrozumieć dzięki badaniu algorytmów oraz badaniu sposobu wewnętrznej reprezentacji pojęć, czyli związków symboli z tym, co one reprezentują”³. Przedstawione ujęcie mieści się w komputacyjnym paradygmacie ujmowania natury umysłu⁴. Przedstawię teraz główne założenia tego programu badawczego oraz jedną spośród najbardziej znanych krytyk tego podejścia – krytykę dokonaną przez Rogera Penrose’a. W ostatniej części tekstu podejmę zaś próbę oceny metodologicznej zasadności uwag Penrose’a.

Główne założenia komputacyjnego programu badania umysłu

Za genezę komputacyjnego podejścia do badania umysłu uznaje się zazwyczaj dwa artykuły autorstwa Alana Turinga: opublikowany w 1936 roku *On Computable Numbers with an Applications to „Entscheidungsproblem”* oraz wydany w roku 1950 *Computing Machinery and Intelligence*. Z filozoficznego punktu widzenia na uwagę zasługuje drugi spośród wymienionych tekstów. W nim to Turing zaproponował rodzaj testu mającego na celu rozstrzygnięcie problemu możliwości dokonywania operacji myślowych przez maszyny matematyczne. Założenia tego testu można przedstawić następująco: wyobraźmy sobie grę, w której uczestniczy człowiek, komputer i sędzia. Sędzia nie widzi ani człowieka, ani komputera. Może on tylko korespondować z nimi, mając do

² Zob.: J. Searle: *Umysł, mózg i nauka*. Przeł. J. Bobryk. Warszawa 1995, s. 21–28.

³ W. Duch: *Czym jest kognitywistyka?* „Kognitywistyka i Media w Edukacji” 1998, nr 1, s. 25.

⁴ Zob.: J. Searle: *Minds, Brain and Programs*. „Behavioral and Brain Sciences” 1982, no 3, s. 420.

dyspozycji dwa terminale X i Y . Jeden z nich zapewnia łączność z komputerem, drugi – z człowiekiem, jednakże sędzia nie wie, który terminal łączy go z człowiekiem, a który z komputerem. Zadanie sędziego polega na stwierdzeniu po serii pytań (np. o upodobania, ubiór), czy X = człowiek, a Y = komputer, czy też odwrotnie. Jeśli sędzia nie będzie potrafił jednoznacznie orzec, który spośród terminali łączy go z komputerem, oznacza to, że zaimplementowany w komputerze program wykazuje cechy analogiczne do własności ludzkiego umysłu⁵. Stwierdzenie to stało się podstawą postulowania analogii pomiędzy ludzkim umysłem i programami komputerowymi oraz założeniem możliwości opisu własności ludzkiego umysłu wyłącznie za pomocą zbioru reguł obliczeniowych. Skoro programy komputerowe konstruowane są na podstawie reguł syntaktycznych i programy te są analogiczne pod względem działania do ludzkiego umysłu, to również ludzki umysł musi poddawać się opisowi wyłącznie w kategoriach obliczeniowych. Wnioskowanie tego rodzaju stało się podstawą formułowania filozoficznych teorii umysłu opartych na omówionym wyniku Turinga⁶.

Warto odnotować, że komputacjonistyczna próba ujęcia umysłu była także warunkowana akceptacją rozstrzygnięć wynikających z pracy Turinga *On Computable...* W niej to bowiem Turing zaproponował schemat maszyny, która zdawała się stanowić idealną konstrukcję służącą do rozwiązywania pewnych problemów językowych, a mianowicie – obliczania funkcji danych w postaci symboli (wyrażeń) oraz rozpoznawania i rozstrzygania języków. Innymi słowy, maszyna Turinga jest modelem, który ma za zadanie realizację każdego rodzaju procedury obliczeniowej, rozumianej jako „niezależna od interpretacji, fizycznie interpretowalna manipulacja symbolami. Tylko w ten sposób okazało się możliwe przeprowadzenie jakiegokolwiek obliczania, o jakim matematycy mogli jedynie marzyć, a także spowodowanie, by maszyny wykonywały wiele różnych inteligentnych czynności, do których wcześniej zdolni byli jedynie ludzie (i zwierzęta). Prawdopodobnie można było więc w tych warunkach całkiem naturalnie wnioskować, że ponieważ: (1) nie wiemy, w jaki sposób podmioty poznawcze poznają, i (2) obliczanie może dokonywać tak wiele czynności, do jakich zdolny jest poznający podmiot, to poznanie jest tylko formą obliczania. [...] Ostatecznie, zgodnie z tezą Churcha–Turinga obliczanie jest wyjątkowe i najwyraźniej wszechmocne [...]; cokolwiek mogą zrobić systemy fizycz-

⁵ Zob.: A. Turing: *Maszyny liczące a inteligencja*. Przeł. D. Gajkowicz. W: *Maszyny matematyczne i myślenie*. Red. E. Feigenbaum, J. Feldman. Przeł. D. Gajkowicz i inni. Warszawa 1972, s. 26.

⁶ Zob.: H. Putnam: *Why Functionalism Didn't Work?* In: Idem: *Words and Life*. Cambridge, Mass., 1994, s. 456.

ne, mogą to także wykonywać komputery⁷. Ten sposób myślenia o obliczaniu sprawił, że zdaniem teoretyków sztucznej inteligencji, maszyna Turinga stała się najbardziej adekwatnym modelem eksplikującym naturę procesów poznawczych. W filozofii umysłu podejście to określa się najczęściej mianem funkcjonalizmu komputerowego.

Podejście funkcjonalistyczne zapoczątkował Hilary Putnam w klasycznym już dziś artykule *Minds and Machines*⁸. Opierając się na wynikach testu Turinga i jednocześnie odrzucając substancjalistyczne podejście do umysłu, Putnam usiłował wykazać, że adekwatna teoria umysłu winna pytać o jego funkcjonowanie, nie zaś o jego konstytutywne własności. Zgadza się z wynikami, jakie uzyskała Żegleń, można stwierdzić, że funkcjonalistyczne podejście do umysłu pozwoliło na uniifikujące spojrzenie na systemy, które dokonują operacji poznawczych, czyli tzw. systemy kognitywne. Z punktu widzenia funkcjonalisty mechanicystycznego systemem kognitywnym jest zarówno umysł ludzki (naturalny system poznawczy), jak i program komputerowy (sztuczny system poznawczy)⁹. Zatem w świetle podejścia mechanicystycznego wewnętrzne stany dowolnego systemu kognitywnego (np. wrażenia, przekonania, pragnienia) z konieczności należy traktować jako stany zdeterminowane komputacyjnie. Determinacja ta ma charakter zarówno ontologiczny, jak i przyczynowy. Oznacza to, że przyczyną istnienia stanów wewnętrznych danego systemu są relacje zachodzące pomiędzy jego elementami syntaktycznymi. Ponadto, stan wewnętrzny *A* dowolnego systemu kognitywnego *B* jest tożsamy z konfiguracją elementów syntaktycznych. Mówiąc inaczej, „wedle przyjętej w funkcjonalizmie metafory komputerowej umysł traktowany jest na wzór programu komputerowego (*software*), którego *hardware* stanowi mózg. Mówiąc z dużym uproszczeniem, programy komputerowe w aspekcie funkcjonalnym traktowane są jako plany organizacji mózgu elektronicznego. Przez analogię do funkcjonowania sztucznego systemu kognitywnego rozpatruje się funkcjonowanie naturalnego systemu kognitywnego [...]. Jest to nie tylko podejście funkcjonalne, ale i formalne do umysłu, gdzie opis procesów poznawczych sprowadza się do operacji na symbolach, tj. do syntaktyki”¹⁰.

Przytoczona wypowiedź podnosi najbardziej dyskusyjne założenie komputacyjnych teorii umysłu. Założenie to ma charakter epistemolo-

⁷ S. Harnard: *Computation Is Just Interpretable Manipulation. Cognition Isn't*. „Minds and Machines” 1995, no 4, s. 385.

⁸ Zob.: H. Putnam: *Minds and Machines*. In: *Idem: Philosophical Papers*. Vol. 2. Cambridge, Mass., 1975.

⁹ Por.: U. Żegleń: *Wprowadzenie...*, s. 120.

¹⁰ Ibidem, s. 121.

giczno-metodologiczny. Można je przedstawić jako twierdzenie, że zarówno wewnętrzna organizacja stanów, jak i sposób nabywania informacji są opisywane wyłącznie za pomocą zbioru procedur obliczeniowych. Innymi słowy, wszystkie operacje poznawcze dowolnego systemu można zredukować (przyczynowo oraz definicyjnie) do relacji syntaktycznych¹¹. W ujęciu tym reprezentacje poznawcze są zatem charakteryzowane wyłącznie jako operacje formalne. Operacje te, jak podkreśla między innymi Jerry Fodor, odzwierciedlają kauzalne relacje zachodzące między stanami zewnętrznymi danego systemu: „Każdy stan umysłowy może być powiązany z pewną formułą, która w sposób unikatowy określa jego położenie w sieci poprzez wyszczególnienie jego potencjalnych interakcji z każdym innym stanem umysłowym. [...] należy jednak zauważyć, że również sieć jest wytworzona przez relacje inferencyjne zachodzące między sądami [...]. Zatem to, że zdanie: Cio-cia jest niższa od wujka Wilifreda, pociąga sąd, że wujek Wilifred jest niższy od cioci, jest przypuszczalnie nieprzypadkową własnością tego sądu”¹². Ujęcie tego rodzaju pozwala wnosić, że procesy umysłowe dowolnego systemu kognitywnego mają wyłącznie syntaktyczne reprezentacje. Wydaje się również, że w ramach komputacyjnych badań nad umysłem celowo pomija się treściowe (semantyczne/intencjonalne) aspekty tych procesów. Zabieg tego typu umożliwia obiektywizację opisu funkcjonowania procesów psychicznych¹³. Własności semantyczne lub/i intencjonalne stanowią natomiast funkcję procesów obliczeniowych. Za Searle’em można stwierdzić, że komputacyjne badania nad umysłem za aksjomat przyjmują, że semantyczne własności systemów kognitywnych są pochodną ich syntaktycznej struktury. Do zaistnienia własności semantycznych nie są więc konieczne ani interakcje systemu ze środowiskiem zewnętrznym, ani wyposażenie go we własność świadomości¹⁴.

Jak wskazuje lektura współczesnych prac z zakresu filozofii umysłu, przedstawione tu główne założenia komputacyjnych teorii umysłu budzą wiele kontrowersji. Do najbardziej znanych krytyk należy zaliczyć argument „chińskiego pokoju” zaproponowany przez Searle’a¹⁵, dokonaną przez Putnama krytykę stanowiska funkcjonalistycz-

¹¹ Por.: S. Pinker: *How the Minds Works*. London 1997, s. 6.

¹² J. Fodor: *Fodor’s Guide to Mental Representation*. „Mind” 1985, no 44, s. 81.

¹³ Por.: Idem: *The Language of Thought*. New York 1975, s. 79.

¹⁴ Zob.: J. Searle: *Consciousness Explanatory Inversion and Cognitive Science*. „Behavioral and Brain Sciences” 1990, no 13, s. 586.

¹⁵ Zob.: Idem: *The Chinese Room Revisited*. „Behavioral and Brain Sciences” 1982, no 2, s. 345–348.

nego¹⁶ oraz refutację głównych założeń komputacjonizmu związana z koncepcją Rogera Penrose'a. W następnej części tekstu dokonam omówienia i oceny ostatniego spośród wymienionych ujęć. Wybór ten podyktowany jest faktem, że podejście Searle'a i Putnama zostało w literaturze przedmiotu wielokrotnie i wyczerpująco omówione, w przeciwieństwie do koncepcji Penrose'a, która zasługuje na uwagę również ze względu na wagę i kontrowersyjność jej wyników.

Penrose'a krytyka komputacyjnych podejść do badania umysłu

Podejście Penrose'a do zagadnienia sposobu badania umysłu jest implikowane rozstrzygnięciami z zakresu fizyki matematycznej oraz filozofii Platona. Korelację tych dwóch, jak się zdaje, niewspółmiernych ujęć wyjaśnia on, powołując się na bezdyskusyjny w jego opinii fakt: „Jedną z zadziwiających cech zachowania świata stanowi jego nadzwyczajna zgodność z prawami matematycznymi. Im lepiej rozumiemy świat fizyczny, im głębiej poznajemy prawa natury, tym bardziej wydaje się nam, że świat fizyczny gdzieś wyparowuje i pozostaje tylko matematyka. Im głębiej rozumiemy prawa fizyki, tym dalej wkraczamy w świat matematyki i matematycznych pojęć”¹⁷. Trzeba jednakże dodać, że platonizm Penrose'a sięga głębiej – ma *stricte* ontologiczną charakterystykę. Podobnie jak Karl Popper¹⁸, Penrose dokonuje podziału rzeczywistości na trzy światy. Świat pierwszy i drugi jest zdaniem zarówno Penrose'a i Poppera, odpowiednio: światem zjawisk fizycznych i psychicznych. Natomiast o ile świat trzeci jest dla Poppera światem zawierającym obiektywne wytwory kultury i nauki, to według Penrose'a stanowi on zbiór absolutów platońskich, w szczególności – absolutnych praw matematycznych¹⁹. Owo założenie ontologiczne jest, zdaniem Penrose'a, konieczne, aby móc dokonać eksplikacji funkcjonowa-

¹⁶ Zob.: H. Putnam: *Why Functionalism...*

¹⁷ R. Penrose: *Makroświat, mikroświat i ludzki umysł*. Przeł. P. Amsterdamzki. Warszawa 1997, s. 18–19.

¹⁸ Zob.: K. Popper: *Epistemologia podmiotu poznającego*. W: Idem: *Wiedza obiektywna. Ewolucyjna teoria epistemologiczna*. Przeł. A. Chmielewski. Warszawa 1992, s. 148–206.

¹⁹ Zob.: R. Penrose: *Nowy umysł cesarza. O komputerach, umyśle i prawach fizyki*. Przeł. P. Amsterdamzki. Warszawa 1995, s. 53–62.

nia umysłu. Wypada również w tym miejscu podkreślić, że swoje podstawowe założenia ontologiczne Penrose wzmacnia do postaci następujących stwierdzeń: „[...] sądzę, że cały świat fizyczny można w zasadzie opisać za pomocą matematyki [...]. Jeśli wybierzemy odpowiednie teorie matematyczne, to można za ich pomocą bardzo dokładnie opisać świat fizyczny, a zatem świat ten zachowuje się zgodnie z regułami matematyki. Wobec tego pewna niewielka część świata platońskiego opisuje cały świat fizyczny. Po drugie, sądzę, że nie istnieją niezależne obiekty umysłowe, nie mające oparcia w świecie fizycznym [...]. Po trzecie wreszcie, wierzę, że przynajmniej w zasadzie wszystkie obiekty ze świata platońskiego są dostępne dla naszego umysłu”²⁰.

W świetle przyjętych założeń wydaje się oczywiste, że jedyną prądomocną teorią umysłu może być koncepcja o fizykalistycznej proveniencji. Sformułowanie takiej koncepcji wymaga jednak, zdaniem Penrose’a, oczyszczenia pola badań nad umysłem z błędnych założeń i stwierdzeń²¹. Jedną z teorii stanowiących, wedle Penrose’a, zbiór takich właśnie błędnych założeń są komputacyjne teorie umysłu.

Komputacyjne teorie umysłu Penrose zbiorczo określa mianem obliczeniowego funkcjonalizmu. *Credo* tego rodzaju podejść formułuje następująco: „[...] wszelkie myślenie polega na wykonywaniu pewnych obliczeń, a zatem jeśli jakiś podmiot wykonuje takie obliczenia, to pojawia się świadomość”²². Zgodnie z intencją Penrose’a trzeba wyraźnie podkreślić, że przeciwko podejściom komputacyjnym przemawiają argumenty empiryczne i logiczne. W dalszym wywodzie skupię się wyłącznie na argumentach natury logicznej ze względu na ich doniosłość, a zarazem kontrowersyjność.

Przyjmując definicję obliczania, jaką podał Turing w *Liczbach obliczalnych*, Penrose stara się wykazać, że aby umysł ludzki działał analogicznie do funkcjonowania programów komputerowych, musi on działać w sposób algorytmiczny. W przypadku analizy działania programów komputerowych jedynie pewne problemy (i to tylko problemy matematyczne) można, zdaniem Penrose’a, rozwiązywać za pomocą matematycznych metod efektywnych²³. Przykładem problemu, dla którego istnieje algorytm, jest zagadnienie znalezienia liczby, której nie można wyrazić w postaci sumy trzech liczb kwadratowych. Istnieją jednakże nawet matematyczne problemy, które nie mają efektywnych rozwiązań. Na mocy twierdzenia Langrange’a można np. wykazać, że nie jest moż-

²⁰ R. Penrose: *Makroświat...*, s. 102–103.

²¹ Por.: Idem: *Nowy umysł cesarza...*, s. 19.

²² Idem: *Makroświat...*, s. 105–106.

²³ Zob.: Idem: *Shadows of the Mind*. Oxford 1994, s. 112–150.

liwe znalezienie liczby naturalnej, która nie jest sumą n liczb kwadratowych, ponieważ jeżeli n jest większe od lub równe 4, to program komputerowy nigdy nie zakończy pracy. Efekt ten informatycy często określają mianem martwej pętli. Ponadto refleksja nad naturą hipotezy Goldbacha wskazuje, że nie jest możliwe, by program komputerowy wygenerował ciąg liczb parzystych większych od 2 – liczby, której nie można przedstawić w postaci sumy co najwyżej n liczb pierwszych.

Uogólniając przedstawione przykłady, można powiedzieć, że przeciwko komputacyjnemu ujęciu umysłu przemawiają w głównej mierze ustalenia matematyczne. Penrose, powołując się na niemożność rozwiązania problemu końca pracy dla uniwersalnej maszyny Turinga oraz na twierdzenie Gödla o niezupełności arytmetyki, stara się wskazać, że sposobu działania ludzkiego umysłu nie da się opisać za pomocą zbioru metod matematycznie efektywnych. Wydaje się ponadto, że sposób rozumowania Penrose'a jest analogiczny do metody, jaką w krytyce komputacjonizmu przyjął John Lucas w znanej pracy *Minds, Machines and Gödel*²⁴. Podobnie jak brytyjski logik, tak i Penrose stwierdza, że analiza sposobu uczenia się i wnioskowania zgodnie z zasadami naturalnych systemów poznawczych wskazuje, że nawet rozwiązanie problemów matematycznych zakłada implementację kategorii rozumienia, powstała, być może, w procesie doboru naturalnego. W kwestii tej Penrose pisze: „Nikt nie uczy dziecka zbioru reguł obliczeniowych; zamiast tego umożliwia się dziecku zrozumienie na przykład tego, czym są liczby naturalne. Powiedziałbym, że w ten sposób dziecko nawiązuje pewien kontakt z platońskim światem matematyki”²⁵. Wedle Penrose'a więc twierdzenie Gödla pozwala na konkluzję, że nawet zrozumienie zagadnień matematycznych nie jest implikowane przez zbiór reguł obliczeniowych. Nieobliczalność kategorii rozumienia pozwala wnosić, że umysł ludzki nie stanowi korelatu elementów syntaktycznych wraz z relacjami między nimi. Powołując się na pracę Davida Deutscha²⁶, Penrose argumentuje, że występowania elementów nieobliczeniowych w matematyce i w fizyce (na co wskazuje np. twierdzenie Bergera o zbiorze trzech polikwadratów) jest rozstrzygającym argumentem przeciwko komputacyjnemu ujmowaniu sposobów działania umysłów. Obliczeniowe ujęcie umysłu byłoby zatem, wedle Penrose'a, możliwe tylko wówczas, gdyby istniał sposób na formalizację nieobliczalnych ele-

²⁴ Por.: J. Lucas: *Minds, Machines and Gödel*. „Philosophy” 1961, no 137, s. 112–137.

²⁵ R. Penrose: *Makroświat...*, s. 119.

²⁶ Zob.: D. Deutsch: *Quantum Theory, the Church – Turing Principle and the Universal Quantum Computer*. „Proceedings Royal Society” 1985, Seria A400, s. 97–117.

mentów występujących w strukturze naturalnych podmiotów poznawczych, np. zdrowy rozsądek, wrażliwość estetyczna. Jako że nie wypracowano technik pozwalających na formalizację tego rodzaju stanów kognitywnych/wolicjonalnych, komputacyjna teoria umysłu jest ujęciem błędnym.

Uwagi końcowe

W niniejszych rozważaniach starałem się choćby w zarysie przeanalizować jeden z palących problemów, nurtujących zarówno filozofów, jak i językoznawców, psychologów, matematyków oraz badaczy sztucznej inteligencji. Opierając się na argumentacji Penrose'a, a zarazem zakładając jej słuszność, chciałem wskazać na ograniczenia dotyczące możliwości konstruowania sztucznych podmiotów kognitywnych. Wybór teorii Penrose'a nie był przypadkiem. Penrose bowiem w swej krytyce podstaw komputacjonizmu uderzył w fundamentalne założenia metod charakterystycznych dla tego stanowiska. Dokonując krytyki mającej ugruntowanie matematyczne, wskazał on, że formalnie jednoznaczne opisanie stanów poznawczych dowolnych podmiotów przetwarzających informacje nie jest możliwe ze względu na fakt, że nawet w samej matematyce istnieją problemy nie poddające się opisowi algorytmicznemu. Z tej przyczyny krytykę Penrose'a uznać należy za metodologicznie poprawną. Co ważne, rozstrzygnięcia Penrose'a nie przekreślają możliwości i zasadności konstruowania systemów rozwiązujących jedynie część problemów trudnych do rozwiązania dla podmiotów naturalnych. Systemy te określane często mianem systemów ekspertowych mają zastosowanie zarówno w praktyce badawczej, jak i w codziennym życiu. Nie są one jednak nawet w części aproksymacjami do funkcjonowania ludzkiego umysłu. Ich użyteczność oceniać można w innej perspektywie. Jest to wszakże temat *stricte* informatyczny.

Wojciech Ślowski

The criticism of the computational theories of mind
Roger Penrose's approach

Summary

The author of the article presents how the view on the results achieved by the theoreticians of mind in the last thirty years, leads to the conclusion that one of the major approaches to the issues of the functioning of mind, is connecting philosophical reflections with the data from computer science, logical and mathematical research and the research on artificial intelligence. The author claims that information processing systems should be understandable thanks to the algorithms analyses and the analysis of the internal representation of concepts, that is the relations between symbols and what they symbolize. The author presents major theses of such a research program and one of Roger Penrose's analysis of this approach. At the end of the article the author attempts to methodologically assess Penrose's observations.

Wojciech Ślowski

Kritik an komputativen Geistestheorien
Die Auffassung von Roger Penrose

Zusammenfassung

Der Verfasser zeigt, dass die von Geistestheoretikern innerhalb von den letzten dreißig Jahren erhaltenen Ergebnisse den folgenden Schluss ziehen lassen: eine von den wichtigsten Meinungen über die Funktion des menschlichen Verstandes war die Verbindung der philosophischen Reflexion mit den, durch Computerwissenschaften und logisch-mathematische Forschungen – Forschungen über die künstliche Intelligenz – eingeholten Daten. Er meint also, dass die die Daten verarbeitenden Systeme sich dank der Untersuchung von Logarithmen und dank der Analyse der inneren Vertretung von Begriffen, d. i. der Zusammenhänge der Symbole damit, was sie eigentlich vertreten, verstehen lassen. Der Verfasser präsentiert die Hauptgrundsätze solch eines Forschungsprogramms und eine von den bekanntesten Analysen, die zum ersten Mal von Roger Penrose durchgeführt wurde. Zum Schluss versucht er, Penroses Methode methodologisch zu beurteilen.