



MASZYNA WIELCE OŚWIECONA. UMYŚŁ LUDZKI Z PERSPEKTYWY NAUK KOGNITYWNYCH

*Skoro jednak wszystkie zdolności duszy zależą tak bardzo
od swoistej budowy mózgu i całego ciała,
że najwidoczniej są tej budowy wynikiem,
mamy tu do czynienia z maszyną wielce oświeconą.*

Julien Offray de La Mettrie
(przeł. Stefan Rudniański)

Gdyby legendarny Marsjanin miał możliwość dłuższego obserwowania zachowania Ziemiaków, sam będąc dla nich niewidoczny, musiałby się zastanowić nad budową i oprogramowaniem „komputera” sterującego owym zachowaniem. Być może pod wpływem tych obserwacji sam podjąłby próbę skonstruowania komputera zdolnego do idealnego naśladowania zachowania Ziemiaków. Jako przedstawiciel cywilizacji technicznie zaawansowanej mógłby odnieść sukces. Dysponując odpowiednim komputerem, zbudowałby następnie robota, z zewnątrz zupełnie nieodróżnialnego od ludzi. Robot, sterowany „marsjańskim” komputerem, zachowywałby się pod każdym względem jak człowiek. Prawdziwi ludzie nie byłiby w stanie odróżnić go na podstawie zachowania od swoich bliźnich, tym bardziej że niczego by przecież nie podejrzewali. Robot zachowywałby się jak człowiek, mówiłby jak człowiek, okazywałby emocje jak człowiek – jak człowiek. Zdałby więc test, który przyjęto nazywać testem Turinga: jego zachowanie byłoby nieodróżnialne od zachowania ludzi. W końcu umarłby jak człowiek i dopiero sekcja „zwłok” by wykazała, że jego ciało różni się zasadniczo od naszego. Mogłoby to być ciało funkcjonujące na podłożu krzemu, a nie białka; mogłoby zresztą przejawiać dowolną budowę i „fizjologię” – istotne w tym momencie jest tylko to, że nie byłoby to ciało ludzkie.

Pojawia się wobec tego szereg pytań. Czy robot tylko zachowywał się jak człowiek, skutecznie oszukując ludzi, czy też myślał, czuł i przeżywał jak jeden z nas? A jeśli myślał i czuł jak człowiek, to czy wiedział, że jest „obcy”? Czy w ogóle

miał świadomość własnego istnienia i własnej odrębności? Czy czuł się podmiotem i sprawcą własnych działań, czy też był tylko bezmyślnym wykonawcą programów, zapisanym w sterującym nim komputerze? Czy był zdolny do odróżniania dobra od zła? Wolał słuchać Mozarta czy też Schoenberga, i czy w ogóle dostrzegał estetyczną – a nie informacyjną – wartość dźwięków? Czy widząc herbacianą różę, doznawał tych samych wrażeń, których my doznajemy w identycznej sytuacji?

Kognitywistyka jako nauka o umyśle

Pytania tego rodzaju są ważne nie tylko dla filozofii umysłu, ale również dla psychologii poznawczej, neurobiologii, lingwistyki, antropologii i nauk komputerowych. Innymi słowy, są to centralne zagadnienia interdyscyplinarnej nauki o poznaniu, zwanej **kognitywistyką** (*cognitive science*). Kognitywiści postępują podobnie jak nasz legendarny Marsjanin. Z jednej strony, obserwują rozmaite przejawy zachowania się ludzi, zwłaszcza zachowania inteligentnego, starając się na tej podstawie dociec, jakie procesy poznawcze rozgrywają się w umyśle istoty zachowującej się inteligentnie. Oczywiście pojawia się tu problem, co to znaczy zachowanie inteligentne. I choć nie sposób wytyczyć ostrej granicy między inteligencją a jej brakiem, typowe przypadki zwykle nie wzbudzają kontrowersji. Rozwiązywanie problemów i działania twórcze – to niewątpliwie przejawy inteligencji, podobnie jak czytanie, mówienie, a nawet niektóre czynności *stricte* motoryczne. Z punktu widzenia ogólnej teorii umysłu jest niewątpliwie bardzo ważne, aby zrozumieć, jakie procesy poznawcze odpowiadają za tego rodzaju czynności, czyli jak musi być zbudowany i oprogramowany „komputer”, który każda z inteligentnych, ludzkich istot nosi na karku w postaci półtora kilograma galaretowatej masy.

Z drugiej strony, kognitywiści próbują sami budować inteligentne komputery. Nie czekają z tym do czasu aż dogłębnie zrozumieją funkcjonowanie mózgu, konstruując programy komputerowe i sztuczne sieci neuronowe obdarzone swoistą architekturą, niekoniecznie zgodną z wiedzą o budowie i funkcjonowaniu mózgu, i wybiórczymi zdolnościami – niekoniecznie porównywalnymi z tym, czego potrafi dokonać człowiek. Niekiedy takie sztuczne układy dysponują nader wątpliwymi zdolnościami intelektualnymi, jeśli porównać je do możliwości umysłu ludzkiego. Ale też wartość takiego postępowania nie polega na skonstruowaniu komputera, który byłby w stanie zastąpić człowieka, lecz na sprawdzeniu, czy jest możliwe symulowanie ludzkich zdolności poznawczych w układach sztucznych. Jeżeli byłoby to możliwe, budowa sztucznej inteligencji byłaby dobrym sposobem badania inteligencji naturalnej. Jeśli bowiem maszyna zachowuje się tak jak człowiek, a zasady jej funkcjonowania są nam – jako konstruktorom – doskonale znane, możemy na tej podstawie wnioskować, że być może na podobnej zasadzie funkcjonuje umysł ludzki. Wnioskując o nieznanym, czyli o zasadach działania umysłu, na podstawie znanego, czyli napisanego przez nas programu komputerowego, przybliżamy się do poznania tajemnicy umysłu. Oczywiście tego rodzaju rozumowanie wymaga przyjęcia założenia, że podo-

bieństwo skutków w postaci zachowania określonego rodzaju oznacza podobieństwo procesów sterujących owym zachowaniem. Założenie to jest mocno kontrowersyjne, ponieważ ekwiwalentność skutków nie oznacza identyczności przyczyn i uwarunkowań. Jednak sam fakt skonstruowania programu komputerowego lub sieci neuropodobnej, udatnie symulującej działanie umysłu ludzkiego, byłby niezwykle osiągnięciem poznawczym, oznaczałby bowiem możliwość zrealizowania inteligentnego procesu poznawczego w postaci maszynowej.

Kognitywistyka to zarazem **konstrukcja i dekonstrukcja**. Konstrukcja polega na budowaniu inteligentnych maszyn, dekonstrukcja – na rozkładaniu na czynniki pierwsze „maszyn” naturalnych. W tym pierwszym specjalizują się konstruktorzy sztucznej inteligencji, w tym drugim – psychologowie poznawczy, neurobiolodzy i przedstawiciele innych nauk empirycznych. Konstruktorom inteligentnych maszyn przyświeca zwykle cel ściśle inżynierski: aby maszyna umiała zrobić coś, co uchodzi za czynność intelektualną (na przykład rozpoznała podpis na czeku, a zwłaszcza odróżniła podpis prawdziwy od fałszywego). Niekiedy konstruktorom chodzi o coś więcej – aby maszyna pokonała człowieka na jego polu (na przykład wygrała w szachy z arcymistrzem). Tak czy inaczej, konstruktorzy sztucznej inteligencji starają się oprogramować maszynę tak, aby wykonywała zadane jej czynności, niekoniecznie przy tym zwracając uwagę na to, czy maszyna wiernie naśladuje człowieka. Z kolei dekonstruktorzy biorą na warsztat „maszynę” już istniejącą, starając się zrozumieć, jak działa, a zwłaszcza – do czego służy. To właśnie miał na myśli Steven Pinker [1997], mówiąc, że kognitywistyka to „inżynieria na opak” (*reverse engineering*). Prawdziwy inżynier tworzy urządzenie zdolne do wykonywania określonej klasy zadań, „inżynier na opak” natomiast stara się zrozumieć budowę i przeznaczenie „maszyny”, którą znalazł w gotowej postaci.

Metody owej dekonstrukcji są liczne i zróżnicowane. Do najważniejszych należy badanie procesów poznawczych zdrowego człowieka w laboratorium (psychologia poznawcza). Niezwykle ważną rolę pełnią też badania nad „maszynami” losowo uszkodzonymi (neuropsychologia, neurochirurgia), nic bowiem nie ujawnia z taką mocą sensu jakiegoś „urządzenia” w ogólnej konstrukcji systemu niż jego awaria. Ważną rolę odgrywają też badania nad językiem, określanym niekiedy jako okno, przez które można zerknąć w czyjś umysł. Kognitywiści wierzą, że w przyszłości zabiegi konstruktorskie i dekonstruktorskie zejdą się w jednym punkcie, którym będzie naukowa teoria umysłu, poparta faktami empirycznymi i uwiarygodniona możliwością zbudowania sztucznej wersji umysłu. Taką teorię będzie można wyrazić w terminach architektury opisującej budowę możliwego do skonstruowania komputera, jak też w terminach licznych procedur sterujących owym komputerem. Tym samym stanie się dopuszczalne stwierdzenie, że nawet jeśli nie jest pewne, czy umysł ludzki działa tak jak skonstruowana i oprogramowana przez nas maszyna, jest całkiem możliwe, że działa tak właśnie. Byłaby to, jak zauważył de La Mettrie [1751/1984], „maszyna wielce oświecona”.

Gra idzie o bardzo wysoką stawkę, mianowicie o to, czy jest możliwe, aby umysł ludzki poznał sam siebie, i by uczynił to za pomocą narzędzi poznawczych naukowego przyrodoznawstwa. Cel jest przynajmniej tak stary, jak stara jest filozofia, szczególnie

teoria poznania i filozofia umysłu. Jednak rozwój nauk tworzących kognitywistykę, zwłaszcza neurobiologii i nauk komputerowych, nadały temu odwiecznemu problemowi nowy sens, a badaniom nad nim – niespotykany dotąd impet. Nawet jeśli okaże się – co prawdopodobne – że cel w postaci naukowej teorii umysłu nie zostanie osiągnięty, postęp w dziedzinie badań nad umysłem dokona się niewątpliwie. I już się dokonuje. To, co obserwujemy od kilku dziesięcioleci, nosi wręcz nazwę **rewolucji kognitywnej**.

Sens tej rewolucji, przynajmniej z punktu widzenia psychologii, nie zawsze jest dostrzegany i właściwie interpretowany. Psychologia naukowa powstała pod koniec XIX wieku jako empiryczne badanie stanów świadomości, głównie percepcji. Dość szybko jednak okazało się, że introspekcyjne zeznania na temat własnych stanów świadomości nie przynoszą prawie żadnej istotnej informacji co do istoty i charakteru ludzkiego poznania. Dzisiaj wiemy, że klasyczna psychologia introspekcyjna nie mogła odnieść sukcesu, ponieważ ludzie nie są świadomi większości swoich procesów poznawczych. Ścisłej rzecz biorąc, są świadomi skutków niektórych procesów – w postaci wrażeń, spostrzeżeń, uczuć – ale nie samych procesów, a zwłaszcza ich struktury i dynamiki. Każdy czytelnik tego tekstu z łatwością odpowie na pytanie, jak się nazywa obecny prezydent RP. Jest więc świadomy faktu, kto jest prezydentem, jak też może być świadomy, iż dysponuje tą wiedzą. Ale prawie nikt z czytelników, z wyjątkiem profesjonalnych badaczy pamięci, nie umiałby opisać procesu odpowiedzialnego za przywołanie z magazynu pamięciowego wiedzy o nazwisku obecnego prezydenta RP. Zresztą profesjonaliści też nie dysponują wiedzą pewną, a jedynie teoriami, hipotezami i pewną ilością danych empirycznych na ten temat. Subiektywnie łatwe i bezwysiłkowe zadanie przywołania z pamięci długotrwałej dobrze znanego faktu jest w istocie zależne od bardzo złożonego i ciągle słabo poznanego mechanizmu, którego nie są w stanie symulować najlepsze z dostępnych obecnie komputerów. Problem nie leży zresztą w mocy obliczeniowej komputerów, ale w braku dobrej, powszechnie uznanej i ugruntowanej w faktach teorii procesu przywołania pamięciowego, która mogłaby stanowić punkt wyjścia do tworzenia modeli obliczeniowych.

Fiasko psychologii introspekcyjnej spowodowało przewrót behawiorystyczny. Przedmiotem badania miało być od tej pory zachowanie, a z tej perspektywy stan świadomości, a nawet samo jej istnienie, był obojętny. To, co pośredniczy między bodźcem a reakcją, potraktowano jak „czarną skrzynkę”, z definicji niedostępną obiektywnemu badaniu. Behawiorysty udoskonaili metodologię badań i sprawili, że psychologia stała się dyscypliną bliską naukom przyrodniczym. W wielu obszarach, zwłaszcza jeśli chodzi o procesy uczenia się, znacząco poszerzyli wiedzę o człowieku. Jednak programowa rezygnacja z badania umysłu sprawiła, że psychologia behawiorystyczna stała się po prostu mało interesująca. Nie po to ktoś chce zostać psychologiem, aby usłyszeć, że nie wolno mu badać psychiki, bo nie ma po temu wystarczająco dobrych narzędzi poznawczych. Co więcej, dla większości psychologów stało się oczywiste, że „mentalistyczne” kategorie pojęciowe, takie jak świadomość czy uwaga, są niezbędne do zrozumienia zachowania – właśnie zachowania! Nacisk na badanie zachowania, a nie umysłu, mógł mieć rozmaite uzasadnienie – ideologiczne, teoretyczne czy wyłącznie metodologiczne. Ale właśnie dzięki przeorientowaniu psychologii z badania świadomości na badanie zachowania stało się jasne, że pełny

Maszyna wielce oświecona. Umysł ludzki z perspektywy nauk kognitywnych

opis i wyjaśnienie zachowania wymaga otwarcia „czarnej skrzynki”, czyli uwzględnienia w badaniu procesów umysłowych. Wtedy właśnie wybuchła rewolucja kognitywna, znaczonej pracami pionierów: Brunera [por. Bruner, 1978; Bruner, Goodnow, Austin, 1956], Kelly’ego [1955] i Neissera [1967].

Blaski i cienie behawioryzmu są ciągle przedmiotem gorącej debaty. Tęsknota psychologów i przedstawicieli innych nauk społecznych do „unaukowienia” ich dyscyplin jest zrozumiała, zważywszy na fakt, że koncepcje psychologiczne zwykle nie zasługują na miano teorii, a większość „praw” psychologicznych to raczej statystyczne tendencje. Nasuwa się pytanie, czy propozycje behawiorystów skutecznie przybliżają nas do celu, jakim jest uczynienie z psychologii „prawdziwej” nauki. Być może rację ma Neisser [1992], że behawioryści polegli, ponieważ przyjęli niewłaściwy model nauki. Próbowali mianowicie zbudować naukę o zachowaniu na wzór fizyki, podczas gdy psychologia – jak sądzi autor – jest „częścią biologii”. Opis zachowania w kategoriach ścisłych zależności między bodźcami a reakcjami, być może z uwzględnieniem rozmaitych „zmiennych pośredniczących”, byłby niewątpliwie czymś fascynującym, gdyby był możliwy. Ale jeśli nie jest możliwy, trzeba szukać innego modelu opisu zachowania i psychiki człowieka, takiego, w którym centralną rolę pełniłyby kategorie odnoszące się do procesów poznawczych i stanów umysłu. Psychologia poznawcza wraz z kognitywistyką przyjęły taki właśnie model budowania teorii umysłu.

Rewolucja kognitywna nie polegała na prostym powrocie do badania świadomości. Przeciwnie, świadomość wracała do psychologii kognitywnej bardzo powoli, a i dziś nie przez wszystkich jest uważana za konieczną i ważną kategorię pojęciową. Sens tak zwanej rewolucji kognitywnej polegał bowiem na tym, aby badać procesy umysłowe z wykorzystaniem metodologii nauk przyrodniczych, czyli mniej więcej tak, jak się bada zjawisko metabolizmu lub powstawanie plam na Słońcu. Wbrew powtarzanym niekiedy sloganom psychologia poznawcza nie wymaga postrzegania człowieka jako istoty świadomej, odpowiedzialnej i racjonalnej. Być może ludzie tacy właśnie są, a przynajmniej bywają. Ale nawet gdyby byli całkowicie pozbawieni świadomości i racjonalności, nie moglibyśmy czuć się zwolnieni z naukowej refleksji nad zjawiskami poznawczymi – świadomymi czy nieświadomymi, racjonalnymi lub irracjonalnymi. Dlatego psychologia kognitywna szuka właściwych modeli teoretycznego opisu mechanizmów funkcjonowania umysłu, ale próbuje weryfikować je z wykorzystaniem obiektywnego aparatu metodologicznego, typowego dla nauk przyrodniczych. Nie trzeba wnikać w stan świadomości szczura, aby badać zasady poznawczej reprezentacji przestrzeni u przedstawicieli tego gatunku. Skoro tak, nie trzeba też wnikać w świadomość człowieka, aby badać procesy przywoływania informacji z pamięci długotrwałej czy jakiegokolwiek inne procesy poznawcze. Co więcej, nie wolno nam nawet próbować wnikania w stany świadomości osoby badanej, bo wtedy z badaczy przekształcimy się w wyznawców „psychologii potocznej”¹.

¹ W ten sposób (*folk psychology*) określa się naiwne, zdroworozsądkowe koncepcje człowieka, psychiki i natury ludzkiej [Dennett, 1987; Grzegorek, 2006].

Sztuka polega więc na zastosowaniu właściwych procedur eksperymentalnych i kontrolowaniu ważnych zmiennych, co niewątpliwie przejęto jako dziedzictwo behawioryzmu. Ale w przeciwieństwie do behawiorystów, którzy „czarnej skrzynki” się bali i uwzględniali ją w swoich modelach tylko wtedy, gdy inaczej nie mogli zrozumieć zachowania, kognitywiści ośmielili się ową skrzynkę otworzyć, skupiając się na obiektywnym badaniu zachodzących w niej procesów. Nastąpiło więc odwrócenie celów i środków: dla behawiorysty czynniki „mentalistyczne” są kategorią pomocniczą w opisie zachowania, dla kognitywisty badanie zachowania jest sposobem obiektywnego dotarcia do procesów umysłowych.

Nie znaczy to, że świadomość nie ma żadnego znaczenia dla psychologii poznawczej i kognitywistyki. Zeznania introspekcyjne mogą być wartościowym punktem wyjścia do konstruowania modeli teoretycznych, choć nie powinny być sposobem ich weryfikacji. Ponadto ludzie mają poczucie podmiotowości i są przekonani o dysponowaniu świadomością, o czym wprost mówią; w takim razie jest to fakt, który sam w sobie domaga się teoretycznego wyjaśnienia. Wreszcie, bez różnych form i rodzajów świadomości, a także nieświadomości, nie moglibyśmy zrozumieć zachowania ludzi w licznych sytuacjach i podczas wykonywania rozmaitych zadań poznawczych. Tak więc świadomość jest dla nauki problemem, choć badanie jej stanów nie może być wiarygodnym źródłem wiedzy o umyśle. Nieadekwatność introspekcyjnej metody badawczej wynika ze wspomnianego już faktu, że o własnych procesach poznawczych i mechanizmach funkcjonowania intelektu ludzie po prostu nie mogą nic sensownego powiedzieć, ponieważ nie mają do tych procesów i mechanizmów wystarczającego dostępu.

Duchy w maszynie umysłu?

Tytuł niniejszego rozdziału jest zamierzonym nawiązaniem do klasycznego dzieła de La Mettriego [1751/1984], ponieważ odzwierciedla dążenie do mechanistycznego, „maszynowego” opisu umysłu ludzkiego. Nie mamy jednak pewności, czy taki opis jest możliwy, i czy w konsekwencji możliwe będzie zbudowanie sztucznego umysłu. Wątpliwości tego rodzaju nie był, zdaje się, pozbawiony sam de La Mettrie. Gdy rozważa on możliwość stworzenia „mówiącego automatu”, formułuje opinię, że „(...) zbudowanie takiej maszyny nie jest niepodobieństwem, zwłaszcza gdyby się do tego zabrał jakiś nowy Prometeusz” [La Mettrie, de, 1751/1984: 80]. Mocne, materialistyczne stanowisko w kwestii skonstruowania sztucznego umysłu opatrzone więc warunkiem, że potrzebny byłby w tym celu wysiłek iście prometejski. W niniejszej publikacji nie przyjmuję z góry założenia o tym, że umysł jest automatem. Próbuje jedynie szukać odpowiedzi na pytanie, czy naukowa teoria umysłu jest możliwa, albo inaczej – w jaki sposób należałoby opisać umysł ludzki, aby taka teoria stała się możliwa w przyszłości.

Dążąc do odpowiedzi na to pytanie, kognitywista powinien przyjąć kilka podstawowych **reguł metodologicznych**. Po pierwsze, powinien podjąć próbę skatalogo-

wania inteligentnych czynności człowieka. Lista takich zachowań będzie stanowić wyjściową bazę empiryczną, dzięki której wyznaczy się „warunki brzegowe” konieczne do spełnienia przez każdą teorię umysłu. Jest to postawa metodologiczna bliska behawioryzmowi, z jego koncentracją na zachowaniu i unikaniem interpretacji kognitywnych. Po drugie, należy zastanowić się, co wiadomo na temat funkcjonowania umysłu ludzkiego na podstawie dotychczasowych badań empirycznych, przede wszystkim prowadzonych w ramach psychologii poznawczej. Chodzi o zbadanie, do jakiego stopnia wiedza na temat struktur i procesów poznawczych pozwala zrozumieć inteligentne zachowania opisane w pierwszym kroku. W drugim kroku obowiązuje więc postępowanie zgodne z metodologicznymi dyrektywami kognitywizmu. Po trzecie, kognitywista powinien dążyć do sformułowania teorii umysłu zdolnej do adekwatnego opisu i wyjaśnienia inteligentnych czynności (krok pierwszy), ale wykorzystującej empiryczną wiedzę na temat funkcjonowania umysłu ludzkiego (krok drugi). Tworząc taką teorię, można rozważać możliwość skonstruowania sztucznego systemu zdolnego do realizowania czynności inteligentnych. W ten sposób powstanie możliwość łączenia podejścia dekonstrukcjonistycznego z podejściem konstrukcjonistycznym. W rezultacie takiego programu badawczego być może pojawi się możliwość sformułowania odpowiedzi na pytanie, czy architektura systemu poznawczego jest w ogóle możliwa do zaimplementowania w systemach sztucznych.

Realizując ten zamiar, kognitywista musi korzystać z bazy empirycznej gromadzonej przez psychologów. Oprócz psychologii poznawczej, co oczywiste, musi czerpać z dorobku dwóch jeszcze dziedzin: psychologii ewolucyjnej i psychologii rozwojowej. Pierwsza z tych dyscyplin pozwala teoretyzować na temat ewolucyjnych wyjaśnień i sensu czynności poznawczych. Spojrzenie z perspektywy psychologii ewolucyjnej pozwala lepiej zrozumieć funkcjonowanie „maszyny”, którą zastaliśmy i którą zamierzamy zdekonstruować. Podstawowe pytanie, do jakiego skłania podejście ewolucyjne, jest pytaniem o funkcję przystosowawczą określonej zdolności poznawczej, struktury lub procesu. Założywszy, że umysł jest „narzędziem” poznania, możemy zadać pytanie, do czego służą poszczególne części lub moduły owego narzędzia, i jak ewoluowały. Korzyści poznawcze, wynikające z ewolucyjnego myślenia o psychice w zakresie badań nad emocjami i zachowaniami społecznymi człowieka [Buss, 2001], nakazują podjąć analogiczną próbę w odniesieniu do badań nad umysłem. Psychologia rozwojowa natomiast dostarcza danych na temat rozwoju funkcji poznawczych w ontogenezie. Uzupełnia zatem wiedzę o zasadach filogenetycznego rozwoju funkcji poznawczych – wiedzą o rozwoju osobniczym. A wiedza o rozwoju pozwala lepiej zrozumieć sens, funkcję i strukturę analizowanego zjawiska.

Próbując stworzyć teorię umysłu, kognitywista nie uniknie ani podstawowych problemów definicyjnych, ani pytań i dylematów natury filozoficznej. Zacznijmy od tych pierwszych. Termin „umysł” znaczy tyle, co **inteligentny system poznawczy**. Jest to więc dość szerokie rozumienie tego pojęcia, obejmujące zarówno świadomość, jak i procesy nieświadome. Problem świadomości musi być oczywiście podjęty jako jedno z centralnych zagadnień kognitywistyki. Jednak olbrzymia ilość danych, wskazujących na nieświadome przetwarzanie informacji przez człowieka, nakazuje rozszerzyć znaczenie pojęcia „umysł” daleko poza procesy związane ze świadomością.

Umysł w tym rozumieniu obejmuje również uczucia oraz ich fizjologiczno-mózgowe podłoże, czyli emocje. Coraz więcej danych wskazuje na to, że różne aspekty emocji wchodzi w interakcję z procesami ściśle poznawczymi, takimi jak spostrzeganie, myślenie i pamięć. Co więcej, zakres i rodzaj uświadomienia emocji przez podmiot jest problemem nader interesującym z punktu widzenia teorii świadomości [LeDoux, 2000]. Jak zauważył Tomasz Maruszewski [2001], niemożność dalszego ignorowania emocji i uczuć przyczyniła się do zasadniczej zmiany sposobu myślenia psychologów poznawczych o umyśle.

Tak więc umysł to nie tylko procesy świadome, i nie tylko poznanie w ścisłym sensie, ale również emocje, uczucia i nastroje. Można by powiedzieć, że umysł jest systemem przetwarzania informacji, gdyby nie fakt, że przetwarzanie informacji zachodzi również na znacznie niższych, nieinteligentnych piętrach regulacji, których sens polega na zachowaniu homeostazy fizjologicznej. Z tego punktu widzenia regulacja i wzajemna kontrola poziomu cukru i insuliny w organizmie jest procesem odbioru, przetwarzania i wysyłania informacji. Co więcej, jeśli przetwarzanie informacji zdefiniujemy jako zmianę stanu na wyjściu, w stosunku do stanu na wejściu, w wyniku zastosowania procedur kodowania i dekodowania, systemem przetwarzania informacji jest każda pojedyncza komórka nerwowa. Umysł jest systemem zdolnym do przetwarzania informacji, ale zdolność ta nie jest specyficzną różnicą, dzięki której moglibyśmy go zdefiniować. Nie można zwłaszcza zawęzić funkcjonowania umysłu do przetwarzania informacji przez człowieka, bo przyjęta perspektywa metodologiczna dopuszcza możliwość skonstruowania umysłu sztucznego, a przynajmniej z góry jej nie wyklucza. Trzeba ponadto wziąć pod uwagę „umysły innych zwierząt” [Pisula, 2003], zwłaszcza gdy przyjmuje się ewolucjonistyczny punkt widzenia. W tej sytuacji najlepiej będzie przyjąć, że umysł to system poznawczy, zdolny do sterowania inteligentnymi czynnościami. Jakie to czynności, i w jaki sposób mogą być sterowane, to główne pytania kognitywistyki.

Co do kwestii natury filozoficznej, należy – choćby w sposób najbardziej wstępny – odnieść się do problemu psychofizycznego, funkcjonalizmu i obliczeniowej teorii umysłu. W kwestii relacji „umysł–ciało” większość kognitywistów to **materialiści**, ponieważ dopuszczają możliwość naturalistycznego opisu umysłu. Ale jest to materializm bardziej metodologiczny niż światopoglądowy. Podejście naukowe wymaga przyjęcia założenia o możliwości wyjaśnienia badanego zjawiska w terminach materialistycznych. Uczony nie musi być materialistą światopoglądowym, jeśli uzna, że niektóre obszary rzeczywistości ze swej natury wymykają się analizom naukowym. Ale tam, gdzie prowadzi się badania naukowe, trzeba przynajmniej roboczo przyjąć założenie, że przedmiot badania jest wyjaśnialny w terminach materialistycznych. Fakt, że przedmiotem badania jest umysł, czyni takie założenie trudniejszym do przyjęcia i bardziej kontrowersyjnym. Ale nie ma innej możliwości, jeśli chcemy uniknąć pułapek związanych z dualizmem. Inaczej mówiąc, nie możemy się zgodzić na żadne „duchy w maszynie umysłu” [Kosslyn, 1983], ponieważ w przeciwnym razie powinniśmy od razu zrezygnować z jakichkolwiek pretensji do stworzenia naukowej teorii umysłu. Umiarkowany charakter materializmu, który tu przyjmujemy, wynika ponadto z przekonania, że czynności i procesy umysłowe są pochodne wobec

Maszyna wielce oświecona. Umysł ludzki z perspektywy nauk kognitywnych

procesów mózgowych, ale do nich nieredukowalne. Innymi słowy, nie jest możliwe opisanie umysłu w terminach procesów fizjologicznych (czy jeszcze niżej: biochemicznych lub bioelektrycznych), choć właśnie owe procesy, ich przebieg i sposób zorganizowania, są przyczyną, podłożem i koniecznym warunkiem pojawienia się umysłu.

Kognitywista jest z reguły **umiarkowanym funkcjonalistą**, ponieważ odpowiada mu ujmowanie umysłu w kategoriach funkcji, a nie struktur fizycznych. Umysły, które znamy, są pochodną mózgu, ale nie można wykluczyć, że istnieją już w tej chwili – w odległych galaktykach – umysły funkcjonujące na innym podłożu. Co więcej, z faktu, że wszystkie znane nam umysły są „formą istnienia białka”, nie wynika, że nie mogą istnieć umysły, będące „formą istnienia krzemu” lub innej substancji. Gdyby funkcje poznawcze mogły być realizowane poza mózgiem, czy wręcz poza strukturami organicznymi, wzrosłyby szanse na stworzenie naukowej teorii umysłu. Oznaczałoby to bowiem, że poznanie może być implementowane w systemach sztucznych, a te byłyby przecież dziełem badaczy. Badacze i konstruktorzy sztucznych umysłów, wiedząc, jak one działają, mieliby w ręku mocny argument: mogliby twierdzić, że umysł ludzki funkcjonuje podobnie jak skonstruowana przez nich maszyna. Podejście funkcjonalistyczne jest więc stanowiskiem metodologicznym, którego przyjęcie ułatwia osiągnięcie celu w postaci naukowej teorii umysłu. W każdym razie podejście takie ułatwia demystyfikację umysłu jako zjawiska rzekomo niepoznawalnego metodami naukowymi.

Zarazem trzeba się chyba zgodzić na umiarkowaną wersję funkcjonalizmu, ponieważ wydaje się oczywiste, że podobne lub identyczne funkcje poznawcze, pełnione przez układy materialne o bardzo różnej budowie, mogą mieć w istocie całkiem różną strukturę i przebieg. Weźmy banalny przykład: człowiek emituje słowa w języku naturalnym i radio też je emituje. Nie chodzi w tej chwili o złożony proces kodowania znaczeń, intencji i stanów świadomości w postaci dźwięków języka naturalnego, co jest istotą mowy, lecz o końcowy etap tegoż procesu, polegający na emisji słów. Techniczne rozwiązanie problemu emisji głosu jest zupełnie inne w wypadku aparatu głosowego człowieka niż w wypadku głośnika radiowego. Skutek jest ten sam, proces i realizujące go struktury materialne – zupełnie różne. Podobieństwo skutków nie oznacza podobieństwa przyczyn, więc gdyby nawet kiedyś udało się skonstruować sztuczny umysł, nie byłoby możliwe udowodnienie, że umysł naturalny funkcjonuje tak samo czy nawet podobnie. W tym miejscu możliwości funkcjonalistycznego opisu umysłu wyczerpują się. Ale to, co dzięki tego rodzaju postępowaniu uda się, być może, w przyszłości osiągnąć, nakazuje nie rezygnować z góry z funkcjonalistycznego sposobu podejścia do problemu.

Stanowisko większości kognitywistów w kwestii funkcjonalizmu dobrze ilustruje klasyczny przykład pieniędzy [Wilson, 1999]. Pieniądz nie jest tym, z czego go zrobiono, to znaczy nie jest papierem, metalem ani plastikiem. Istota pieniądza tkwi w jego funkcji jako uniwersalnego miernika wartości materialnych. Ale funkcja ta nie mogłaby być pełniona, gdyby nie owe materialne nośniki. Inaczej mówiąc, każdy pieniądz wymaga materialnego podłoża, ale wartość i rola pieniądza nie jest do owego podłoża redukowalna. Pieniądz pełni ważną rolę w kauzalnym łańcuchu zda-

rzeń, bez udziału pieniądza niemożliwe są, na przykład, transakcje kupna–sprzedaży. Pieniądz ma więc siłę sprawczą, może wpływać na zachowanie się ludzi, a także na położenie i losy przedmiotów materialnych. Bez wiedzy o tym, że istnieje pieniądz i pełni określone funkcje, nie byłibyśmy w stanie zrozumieć wielu zachowań człowieka. Z drugiej strony, funkcja uniwersalnego miernika wartości może być pełniona poza pieniądzem, na przykład przez muszelki, koraliki lub żetony, a współcześnie – przez „wirtualny” zapis na koncie bankowym. Ponadto ta sama funkcja jest z powodzeniem pełniona przez różne waluty. Z tego, że funkcja miernika wartości może być „implementowana” w różnych nośnikach materialnych, wynika, że różne przedmioty są funkcjonalnie ekwiwalentne. Nie wynika stąd jednak, że są ekwiwalentne pod innym względem, ani że ekwiwalencja funkcji oznacza identyczność sposobu jej pełnienia. Tak więc, zgodnie z przyjętym tu stanowiskiem funkcjonalistycznym, umysł jest ujmowany poprzez jego funkcje, które wymagają materialnego podłoża i które – przynajmniej potencjalnie – mogą być implementowane w różnych układach fizykalnych. Nie oznacza to, że umysł da się zredukować do zjawisk materialnych, ani że ta sama funkcja poznawcza, pełniona przez różne systemy, jest pełniona w ten sam sposób.

Problem funkcjonalizmu wiąże się ściśle z problemem tak zwanej **obliczeniowej teorii umysłu** (*computational theory of mind*). Obliczeniowa teoria umysłu (OTM) powstała w wyniku prac wielu uczonych, przede wszystkim Fodora [1975] i Pylyshyna [1980, 1984], znajdujących się pod wpływem dzieła Turinga [1950]. OTM zakłada, że umysł jest maszyną, a jego czynności – serią obliczeń, czyli operacji polegających na przetwarzaniu informacji. Mocna wersja tej teorii zakłada, że umysł jest tak zwaną uniwersalną maszyną Turinga, czyli urządzeniem zdolnym do odczytywania i zapisywania nieskończonego ciągu informacji według skończonego zbioru jawnych (*explicit*) reguł. Wyobraźmy sobie nieskończenie długą taśmę z zapisanymi na niej symbolami. Urządzenie wyposażone jest ponadto w czytnik, zdolny do identyfikowania symbolu znajdującego się w polu odczytu. Dysponując skończonym zestawem reguł przetwarzania informacji, maszyna Turinga jest w stanie dowolnie długo przesunąć czytnik w przód lub w tył, odczytywać napotkany po takim ruchu symbol i przekształcić go w inny symbol lub dokonać nowego ruchu czytnika. Reguły, jakimi rządzi się taka maszyna, mają postać „jeśli... to..., w przeciwnym razie...”. Jest to dobrze znana programistom komputerowym formuła „*if... then... else*”. Reguła może brzmieć, na przykład, następująco: „jeśli widzisz kwadrat, zamień go na trójkąt, w przeciwnym razie cofnij się o dwa pola”. Inna reguła może nakazywać: „jeśli widzisz koło, przejdź jedno pole do przodu, w przeciwnym razie zamień widziany symbol na kwadrat”. Maszyna Turinga nie jest maszyną w sensie fizycznym; jest automatem „wirtualnym”, zdolnym do przetwarzania informacji według jawnych, ścisłych reguł.

Pytanie, czy umysł ludzki jest maszyną Turinga, i czy naukowa teoria umysłu musi być teorią obliczeniową, jest przedmiotem nieustających polemik. Pinker [1997] nie dopuszcza innej możliwości, ponieważ, jego zdaniem, tylko OTM pozwala przezwyciężyć pułapkę dualizmu. Jeśli mój zamiar położenia się na kanapie sprawia, że moje ciało rzeczywiście zmienia położenie w przestrzeni fizycznej, w jaki sposób niema-

terialne zamiary mojego umysłu mogą oddziaływać na materialne ruchy mojego ciała? Oto klasyczny dylemat stosunku duszy do ciała, nazwany problemem psychofizycznym, niemający – jak się wydaje – rozsądnego rozwiązania przy tak zadanym pytaniu. Ale problem znika, gdy przyjmiemy, że mój zamiar też jest materialny: jest mianowicie serią obliczeń, wykonywanych przez określone partie mojego mózgu. Kończącym fragmentem owych obliczeń jest rozkaz wydany mięśniom, sterującym ruchami ciała. Jeśli umysł to swoisty „komputer”, nie ma dwóch bytów, z których jeden byłby materialny, a drugi niematerialny. Jest jeden materialny byt, ale cała jego istota, sens i sposób działania sprowadza się do przetwarzania informacji.

Inni badacze, na przykład „późny” Neisser [1994], Searle [1992], a zwłaszcza Penrose [2000], formułują liczne wątpliwości co do możliwości zbudowania obliczeniowej teorii umysłu. Najważniejsze zastrzeżenia wynikają z faktu, że uniwersalna maszyna Turinga jest systemem przetwarzającym symbole według jawnych i ścisłych reguł, czyli algorytmów. Co więcej, jest to maszyna działająca szeregowo. Tradycyjne badania nad sztuczną inteligencją [np. Newell, Simon, 1972] polegają właśnie na emulowaniu ludzkich czynności inteligentnych poprzez tworzenie algorytmów. Tymczasem wiemy z badań nad mózgiem, że działa on równolegle, niekoniecznie na poziomie symbolicznym i często przy braku jawnych reguł przetwarzania informacji. Dlatego pojawiły się ujęcia alternatywne w postaci tak zwanych modeli konekcyjnistycznych, polegających na tworzeniu sieci neuropodobnych (*neural networks*). Sieć neuropodobna nie otrzymuje jawnych instrukcji działania, choć modyfikuje swe zachowanie pod wpływem doświadczenia i informacji zwrotnych. Ponadto działa na poziomie subsymbolicznym oraz przetwarza informacje w sposób rozproszony i równoległy [Rumelhart, McClelland i PDP Research Group, 1986]. Twórcy modeli konekcyjnistycznych skupiają uwagę głównie na architekturze sieci, a nie na syntaktycznych regułach przetwarzania symboli. Modelowanie sieciowe uważane jest obecnie za alternatywę, a nawet opozycję, w stosunku do „starej, dobrej sztucznej inteligencji”, opartej na algorytmicznym przetwarzaniu symboli. Być może w przyszłości dojdzie do unifikacji obu podejść, zresztą już w tej chwili obserwuje się rozwój tak zwanych hybrydowych modeli poznania [Jordan, Russell, 1999]. Nie można też wykluczyć pojawienia się takich rozwiązań w zakresie modelowania funkcji poznawczych, o których obecnie nawet nam się nie śni.

Inne zastrzeżenia wobec OTM wynikają z faktu, że reguły przetwarzania symboli same w sobie nie zapewniają systemowi inteligencji. Ilustruje to tak zwany **paradoks chińskiego pokoju** [Searle, 1980]. Przypuśćmy, że ktoś nieznający języka chińskiego został zamknięty w pokoju, do którego są mu dostarczane teksty po chińsku. Człowiek ten dysponuje zapisanym w księgach pełnym zestawem reguł syntaktycznych, pozwalających tworzyć poprawne gramatycznie zdania w języku chińskim. Posługując się tymi regułami, formułuje sensowne odpowiedzi na dostarczane mu komunikaty. Chińczycy, znajdujący się na zewnątrz pokoju, nie domyślają się, że zamknięty w nim człowiek nie zna chińskiego, ponieważ jego odpowiedzi – dzięki znajomości kompletu reguł syntaktycznych – nie są odróżnialne od tych, których by udzielił rodowity Chińczyk. Czy człowiek zamknięty w chińskim pokoju zna język chiński?

Czy rozumie to, co mu przekazują? Czy rozumie to, co sam formułuje w odpowiedzi na otrzymany komunikat? Czy jest inteligentnym użytkownikiem języka chińskiego?

Searle udziela na powyższe pytania odpowiedzi negatywnej. Ale podstawowy problem, jak się wydaje, tkwi w mocno dyskusyjnym założeniu, że zachowanie naiwnego językowo człowieka, który by dysponował pełnym zestawem reguł syntaktycznych danego języka, nie byłoby odróżnialne od zachowania rodowitych użytkowników tego języka. Po pierwsze, odpowiedzi udzielane przez tę osobę musiałyby trwać bardzo długo, tak długo, jak tego wymaga dopasowanie otrzymanego na wejściu komunikatu do właściwej odpowiedzi za pomocą olbrzymiej ilości reguł składniowych zapisanych w księgach. Z pewnością Chińczycy, czekający na zewnątrz, powzieliby poważne wątpliwości, czy interlokutor jest jednym z nich, a w każdym razie – czy jest inteligentny. Po drugie, i znacznie ważniejsze, jest wysoce wątpliwe, czy odpowiedzi udzielane przez naiwną językowo osobę byłyby rzeczywiście sensowne. Być może byłyby poprawne składniowo, ale przypuszczalnie brakowałoby im sensu i dostosowania do kontekstu. Umiejętność formułowania poprawnych syntaktycznie komunikatów nie implikuje bowiem umiejętności formułowania znaczeń. Semantyczna strona języka nie jest redukowalna do składni, albowiem wymaga uwzględnienia relacji umysł–świat. Znaczenie jest relacją, która wiąże umysł ze znakiem (na przykład słowem) i przedmiotem oznaczanym. Dlatego paradoks chińskiego pokoju wydaje się źle skonstruowany. Ten pouczający eksperyment myślowy nie powinien więc być używany jako argument na rzecz tezy, że umysłu nie da się opisać jako maszyny ani że nie da się zbudować sztucznej maszyny, zachowującej się jak człowiek. Jest natomiast przekonującym argumentem, że umysł nie może być maszyną Turinga, a więc taką, której sposób działania ogranicza się do przetwarzania symboli według jawnych reguł algorytmicznych.

Wydaje się, że również w kwestii OTM możliwe jest przyjęcie stanowiska umiarkowanego. Stanowisko to sprowadza się do przekonania, że umysł jest maszyną, ale niekoniecznie maszyną Turinga. Umysł jest maszyną w tym sensie, że polega na przetwarzaniu informacji przez określone byty materialne – mózg, układy krzemowe, optyczne lub jakiegokolwiek inne. Nie istnieje informacja, której nie odpowiadałby jakiś substrat materialny, na przykład mózgowy. Ale też informacja nie jest redukowalna do stanów fizykalnych swego materialnego podłoża. Materialne podłoże umysłu musi charakteryzować się architekturą umożliwiającą przetwarzanie informacji. Sieć neuronów w mózgu spełnia ten warunek, podczas gdy budowa kości nie zezwala na przetwarzanie informacji. Oto dlaczego myślimy mózgiem, a nie kośćmi. Przyszłe próby zbudowania sztucznego umysłu będą – być może – wykorzystywać inne nośniki materialne, ale zawsze takie, które zezwalają systemowi na przetwarzanie informacji. Jeśli umysł jest maszyną, można ją, po pierwsze, dekonstruować, a po drugie – konstruować na nowo w układach sztucznych. Tym samym można realizować program badawczy kognitywistyki.

Jednakże umysł niekoniecznie jest maszyną Turinga, pozostaje bowiem w relacji do swego otoczenia, i tylko w tej relacji pełni swe inteligentne funkcje. Jest mało prawdopodobne, aby umysł ludzki ewoluował tylko po to, by przetwarzać symbole według reguł algorytmicznych. Ewoluował, aby ułatwić swemu nosicielowi regula-

Maszyna wielce oświecona. Umysł ludzki z perspektywy nauk kognitywnych

cję stosunków z otoczeniem, zatem tylko w relacji do otoczenia może ujawnić swoją kompetencję, przydatność i inteligencję. Umysł może przetwarzać symbole, jeśli to zwiększa jego inteligencję; ale sama umiejętność przetwarzania symboli to jeszcze nie inteligencja. Jeśli więc definiujemy umysł jako **inteligentny** system przetwarzania informacji, musimy przyjąć, że natura wyposażyła go w umiejętność samodzielnego i dostosowanego do okoliczności działania w świecie. Jest to ważna dyrektywa dla prób konstrukcji umysłu sztucznego, a tym bardziej – dla prób tworzenia naukowej teorii umysłu.

Podsumowanie

Gdyby więc znalazł się jakiś nowy Prometeusz, chcący skonstruować sztuczny umysł, musiałby pamiętać o kilku podstawowych zasadach. Po pierwsze, musiałby uważnie obserwować działanie umysłów naturalnych, pamiętając wszakże, iż umysł sztuczny nie musi być ich prostym naśladownictwem. Po drugie, byłby świadomy tego, że sukces w postaci skonstruowania sztucznego umysłu niekoniecznie oznacza pełne zrozumienie umysłu naturalnego, ponieważ podobieństwo funkcjonalne nie implikuje podobieństwa innego rodzaju. Po trzecie, wiedziałby doskonale, że jego zadaniem jest skonstruowanie sztucznego umysłu jako systemu zdolnego do inteligentnego przetwarzania informacji, choć miałby duże problemy ze zdefiniowaniem ostrej granicy między działaniem inteligentnym a nieinteligentnym. Po czwarte, musiałby przyjąć, że sztuczny umysł jest maszyną przetwarzającą informację, choć nie byłby skazany na konkretne rozwiązania teoretyczne i techniczne w kwestii tego, co to za maszyna i jak działa. Po piąte, pamiętałby o tym, że umysł przetwarza informację nie po to, by je przetwarzać według z góry określonego algorytmu, ale w jakimś celu. Nadrzędnym celem każdego umysłu naturalnego jest umożliwienie inteligentnej regulacji stosunków swojego właściciela z otoczeniem, zatem umysł sztuczny tylko wtedy zasługiwałby na swą nazwę, gdyby czynił swego nosiciela – na przykład robota – bardziej inteligentnym w działaniu w określonym środowisku.

Czy jednak po skonstruowaniu „maszyny wielce oświeconej” nowy Prometeusz mógłby z satysfakcją stwierdzić, że przynajmniej jeden problem nękający przez wieki filozofów, mianowicie problem stosunku duszy do ciała czy też umysłu do mózgu, definitywnie przestał być problemem?

Literatura cytowana

- Bruner J.S. (1978), *Poza dostarczone informacje: Studia z psychologii poznawania*, przeł. B. Mroziak, Warszawa: PWN.
- Bruner J.S., Goodnow J., Austin G.A. (1956), *A study of thinking*, New York: John Wiley and Sons.
- Buss D.M. (2001), *Psychologia ewolucyjna*, przeł. M. Orski, Gdańsk: GWP.

Edward Nęcka

- Dennett D.C. (1987), *The intentional stance*, Cambridge, MA: The MIT Press.
- Fodor J. (1975), *The language of thought*, New York: Crowell.
- Grzegorek T. (2006), *Potoczne ujęcie zachowania. Krytyczna analiza relacji między psychologią i tzw. folk psychology*, Kraków: Uniwersytet Jagielloński (nieopublikowana praca doktorska).
- Jordan M.I., Russell S. (1999), *Computational intelligence* [w:] R.A. Wilson, F.C. Keil (red.), *The MIT encyclopedia of the cognitive sciences*, Cambridge, MA: The MIT Press.
- Kelly G.A. (1955), *The psychology of personal constructs*, New York: Norton.
- Kosslyn S.M. (1983), *Ghosts in the mind's machine: Creating and using images in the brain*, New York: Norton.
- La Mettrie de J.O. (1984/1751), *Człowiek–maszyna*, przeł. S. Rudniański, wyd. 2, Warszawa: PWN.
- LeDoux J. (2000), *Mózg emocjonalny*, przeł. A. Jankowski, Poznań: Media Rodzina.
- Maruszewski T. (2001), *Subtelny czar psychologii poznawczej*, „Przegląd Psychologiczny”, t. 44, nr 1, 57–71.
- Neisser U. (1967), *Cognitive psychology*, New York: Appleton-Century-Crofts.
- Neisser U. (1992), *Two Themes in the Study of Cognition* [w:] H.L. Pick, Jr. P. van den Broek, D.C. Knill (red.), *Cognition: Conceptual and methodological issues* (333–340), Washington DC: American Psychological Association.
- Neisser U. (1994), *Multiple Systems: A new Approach to Cognitive Theory*, „European Journal of Cognitive Psychology”, 6, 225–242.
- Newell A., Simon, H.A. (1972), *Human problem solving*, Englewood Cliffs, NJ: Prentice-Hall.
- Penroes R. (2000), *Cienie umysłu. Poszukiwanie naukowej teorii świadomości*, przeł. P. Amsterdamski, Poznań: Zysk i S-ka.
- Pinker S. (1997), *How the mind works*, New York: W.W. Norton.
- Pisula W. (2003), *Psychologia zachowań eksploracyjnych zwierząt*, Gdańsk: GWP.
- Pylyshyn Z. (1980), *Computation and Cognition: Issues in the Foundation of Cognitive Science*, „Behavioral and Brain Sciences”, 3, 111–132.
- Pylyshyn Z. (1984), *Computation and cognition: Toward a foundation of cognitive science*, Cambridge, MA: The MIT Press.
- Rumelhart D.E., McClelland J.L. i PDP Research Group (1986), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (vol. 1 and 2), Cambridge, MA: The MIT Press.
- Searle J.R. (1980), *Minds, Brains, and Programs*, „Behavioral and Brain Sciences”, 3, 417–457.
- Searle J. (1992), *The Rediscovery of the Mind*, Cambridge, MA: The MIT Press.
- Turing A. (1950), *Computing, Machinery and Intelligence*, „Mind”, 59, 433–460.
- Wilson R.A. (1999), *Philosophy* [w:] R.A. Wilson, F.C. Keil (red.), *The MIT encyclopedia of the cognitive sciences*, Cambridge, MA: The MIT Press.