

## Funkcjonalizm obliczeniowy – kilka uwag z perspektywy ewolucyjnej

### Charakterystyka stanowiska funkcjonalizmu obliczeniowego

Stanowisko, które zamierzam krótko scharakteryzować i poddać krytyce poniżej, stanowi jedną z konkurencyjnych propozycji w sporze o naturę umysłu<sup>1</sup>. Funkcjonalizm głosi, że świadomym stanem mentalnym jest wszystko to, co jest w stanie realizować funkcje, które według naszej wiedzy realizują uświadamiane sobie przez nas stany mentalne. Jeżeli na mózg spojrzeć przede wszystkim jako na biologiczną strukturę, której zadaniem jest przetwarzanie informacji, wyczerpująca charakterystyka składających się na świadomy umysł stanów mentalnych jest możliwa w terminach korelacji pomiędzy sygnałami (danymi) wejściowymi a wyjściowymi mózgu. Gdy świadomość okazuje się jedynie funkcją, którą realizuje mózg, nic nie stoi na przeszkodzie, aby świadome stany przypisać każdemu indywiduum posiadającemu strukturalną złożoność zdolną do odpowiedniego przetwarzania informacji<sup>2</sup>.

Funkcjonalizm obliczeniowy (od teraz: FO) wymaga dodatkowo założenia, że świadomość jest obliczeniowym sposobem przetwarzania informacji – sygnał na „wejściu” i „wyjściu” posiadającego świadomość indywiduum daje się wyrazić jako funkcja matematyczna równoważna formalnie zdefiniowanej efektywnej operacji mechanicznej o skończonej ilości kroków (maszynie Turinga)<sup>3</sup>. Komputer jest urządzeniem elektronicznym (wykorzystującym do swojego działania zjawiska elektryczne) przetwarzającym informacje w sposób obliczeniowy. Komputery traktuje się zwykle jako fizyczne realizacje maszyn Turinga. Klasa takich urządzeń ma tę właściwość, że struktura przetwarzania informacji (od teraz: SPI) każdego z nich może być symulowana przez dowolny komputer o innej specyfikacji obliczeniowej (SPI).

Postulatem omawianej wersji funkcjonalizmu jest więc potraktowanie świadomości jako fenomenu czysto informacyjnego, w dodatku o charakterze obliczeniowym (świadomość = obliczeniowy sposób przetwarzania informacji przez mózg). Prowadzi to do uznania, że mózg w pewnym sensie przypomina komputer, a zachodzące

<sup>1</sup> Termin „umysł” rozumieć należy jako równoznaczny z kartezjańską „substancją myślącą”, „świadomością fenomenalną” czy wreszcie „subiektywnymi stanami jakościowymi”.

<sup>2</sup> Tę właśnie myśl wyraża stwierdzenie, że bycie w świadomym stanie mentalnym w rodzaju przekonań czy pragnień jest byciem w odpowiedniej relacji z symboliczną reprezentacją treści przekonania czy pragnienia. W relacji tej może znaleźć się każde indywiduum posiadające zdolność do odpowiedniego przetwarzania informacji.

<sup>3</sup> Jest to treścią słynnej hipotezy Churcha-Turinga.

w jego obrębie procesy odpowiadające za powstanie stanów mentalnych mogą zostać przez komputer zasymulowane. Świadomy umysł jest więc SPI równoważną pewnemu komputerowemu algorytmowi. Wykonanie ciągu odpowiednich operacji formalnych (przekształceń o charakterze syntaktycznym) prowadzi do powstania stanów świadomych i wytworzenia się umysłu [Putnam 1967].

Następujący schemat pozwala łatwo zorientować się w strukturze omówionych powyżej twierdzeń:

## FUNKCJONALIZM OBLICZENIOWY

- (F1) Stany świadome to odpowiednie korelacje pomiędzy sygnałami wejścia i wyjścia.
- (F2) Korelacje te mają charakter funkcyjny.
- (F3) Funkcje, o których tu mowa, są obliczalne.

Z F1–F3 wynika zaś:

- (F4) Indywidualium  $\alpha$  posiadające stany świadome to tyle, co komputer wykonujący pewien program komputerowy. Innymi słowy: umysł jest programem komputerowym, ciągiem operacji wykonywanych przez odpowiednią maszynę liczącą. „Umysł to zjawisko z poziomu *software* względem mózgu, który odgrywa rolę *hardware*”.

Na koniec jeszcze drobna uwaga terminologiczna. Scharakteryzowane właśnie stanowisko nazywam „funkcjonalizmem obliczeniowym”, lecz w literaturze spotkać się można także z określeniem „funkcjonalizm maszynowy” (*machine state functionalism*) czy „obliczeniowa teoria umysłu” (*computational theory of mind*). Wszystkie je będę traktował jako odnoszące się do tego samego zestawu twierdzeń {F1, F2, F3, F4}.

## Korzyści z przyjęcia funkcjonalizmu obliczeniowego

Twierdzenie F4 wydaje się niezwykle atrakcyjne przynajmniej z trzech powodów. Po pierwsze, udziela takiej odpowiedzi na pytanie o kategorię ontologiczną stanów mentalnych, która znosi kartezjański dualizm pomiędzy duchem a materią. Świadomość jest przecież fenomenem informacyjnym analogicznym do komputerowego *software*. O ile ten ostatni nie stanowi dla materializmu żadnego problemu, podobnie powinniśmy myśleć o stanach mentalnych. Z F4 wynika także, że każdy obiekt wystarczająco złożony, aby przetwarzać informacje w odpowiedni sposób, jest zdolny do wytworzenia świadomości. Stanowi to pożądaný wniosek, gdyż rozsądne wydaje się przyjęcie, że podstawy świadomości nie są wyłącznie białkowe.

Po trzecie wreszcie, FO pozwala wskazać na mechanizm odpowiedzialny za kausalność stanów mentalnych. Pozwala na to interpretacja procesów mózgowych w kategoriach operacji o charakterze algorytmicznym. Umysł jest fizycznym systemem formalnych symboli, które przekształcane są według ustalonych reguł. Symbole te reprezentują myśli, a niezawodne przejście od jednego takiego symbolu do drugiego pozwala wyjaśnić obserwowane wywoływanie jednego stanu świadomego przez inny. Przykładowo, odpowiedź na pytanie, dlaczego przekonaniu, że nadchodzi burza towarzyszy często pragnienie, aby jak najszybciej znaleźć schronienie, polega na postulowaniu istnienia implementowanej w mózgu reguły, łączącej ze sobą wystąpienie odpowiednich przekonań i pragnień. Treścią **hipotezy systemu fizycznych symboli** jest stwierdzenie, że każde inteligentne zachowanie jest w istocie wynikiem działania systemu fizycznych symboli [Newell 1980]<sup>4</sup>.

## Zarzuty pod adresem funkcjonalizmu obliczeniowego

Na pierwszy rzut oka metafora *software* i *hardware* wydaje się obiecującą propozycją wyjaśnienia działania mózgu i procesów poznawczych. Pod jej adresem wysunięto jednak co najmniej kilka krytycznych uwag. Najślynniejsze z nich pochodzą od Johna R. Searle'a [1992].

Argument chińskiego pokoju sprowadza się do uzasadnienia za pomocą fikcyjnego przykładu posługiwania się symbolami języka chińskiego, że do przetwarzania symboli nie jest konieczne ich rozumienie. Zatem, jeżeli umysł to jedynie skończony zbiór zdefiniowanych formalnych operacji na symbolach, zwolennik FO nie jest w stanie wyjaśnić, na czym polega działanie umysłu. Jego efektem jest bowiem rozumienie przetwarzanych treści.

Zresztą, problem dla funkcjonalizmu, nie tylko w wersji obliczeniowej, stanowią również jakości (*qualitative subjective sensations – qualia*) składające się na pokąźną część introspekcyjnie dostępnych stanów mentalnych. Zwolennik funkcjonalizmu postuluje bowiem, że odczuwanie  $x$  jest po prostu tożsame z odbiorem i reakcją na  $x$ , z czym trudno się zgodzić.

Niezależnie od tego, Searle wskazuje na niedostateczne uzasadnienie wynikające z przyjęcia FO twierdzenia, jakoby umysł był fenomenem czysto informacyjnym. Zwolennicy funkcjonalizmu w żaden sposób nie wykazali, że świadome odczuwanie jakościowe nie opiera się na jakichś niezidentyfikowanych jeszcze cechach materii, która jest środowiskiem pojawienia się umysłu. W naszym przypadku jest to biologiczny mózg, ale w grę wchodzi także inne fizyczne realizacje. Szkopuł w tym, że nie jest wcale przesądzone, że dla oddania istoty stanów mentalnych (pragnień, przekonań) oprócz SPI nie będzie konieczne podanie także szczegółów dotyczących fizycznych właściwości medium, które stany mentalne wytwarza. Proponenci FO wydają się rozumieć F4 w ten sposób, iż implikuje ono:

<sup>4</sup> Co należy chyba rozumieć jako równoważne stwierdzeniu, że każde przejawiające inteligentne zachowanie indywiduum traktować można jak maszynę Turinga o ustalonej liście formalnych instrukcji.

(F5) Skopiować stany mentalne indywiduum a to tyle co skopiować jego SPI.

Tymczasem, nie podają oni żadnych przekonujących argumentów na rzecz prawdziwości tego twierdzenia. Są one tym bardziej wymagane, iż chiński pokój oraz problem jakości wskazuje, że prawdziwe może być jego zaprzeczenie.

Z twierdzeniami FO sprzeczne są również wyniki badań nad teorią i zachowaniem się tzw. sieci neuronowych. Hipoteza fizycznego systemu symboli postuluje, że podstawą inteligentnego działania są wehikuły semantyczne mające postać symboli wchodzących w skład operacji formalnych o charakterze algorytmicznym. Lecz sieci neuronowe kodują i przetwarzają informacje za pomocą bogatszych środków niż symbole. Wiedza gromadzona przez sieć neuronową nie jest propozycjonalna, tzn. nie redukuje się do zbioru formuł i operacji na nich. Jeżeli więc mózg jest w ogólnych zarysach siecią neuronową i działa na podobnej zasadzie jak ona, z czym trudno, po drobnych zastrzeżeniach, się nie zgodzić, hipoteza fizycznego systemu symboli jest fałszywa.

## Perspektywa ewolucyjna a funkcjonalizm obliczeniowy

Autorzy zajmujący się problematyką umysłu prowadzą swoje badania zwykle w zupełnym oderwaniu od faktu, że jedyny znany materialny nośnik fenomenu świadomości, czyli biologiczny mózg, jest wynikiem długotrwałego procesu ewolucji życia na Ziemi. W takim razie szczególności sposobu przetwarzania informacji przez mózg powinny mieć ewolucyjne pochodzenie, co nakłada poważne ograniczenia na klasę możliwych wyjaśnień natury umysłu. Z należytego uwzględnienia perspektywy ewolucyjnej w filozofii umysłu wynika kolejny szereg uwag krytycznych pod adresem FO.

Zgodnie z obowiązującym kanonem spojrzenia na biologiczne życie, cechy organizmów kształtowane są poprzez trzy rodzaje czynników: oddziaływanie ze środowiskiem, oddziaływanie z innymi organizmami (w tym mieści się zarówno rywalizacja żywiciel-pasożyt, jak i symbiotyczne związki komórek bakteryjnych prowadzące do powstania komórek eukariotycznych) oraz rywalizację genów (teoria samolubnego genu). Wszystkie wymienione naciski selekcyjne mają jednak charakter przypadkowy. Konsekwencje tego faktu dla filozofii umysłu wydają się doniosłe.

**GENEZA UMYŚLU (BAG-OF-TRICKS).** Jeżeli umysł jest równoważny pewnemu algorytmowi, który realizuje mózg, powstaje pytanie, czy struktura realizująca algorytm jest w stanie powstać w wyniku przypadkowych nacisków selekcyjnych? Na ile prawdopodobne jest, że splot przypadkowych okoliczności doprowadzi do wykształcenia się organizmu lub jakiegoś organu, którego zachowanie się będzie równoważne wykonywaniu zestawu ściśle zdefiniowanych kroków o skończonej liczbie? Pytanie to można rozumieć na dwa sposoby.

W bardziej ogólnej wersji jest to pytanie o to, czy zachowanie doskonale algorytmiczne może powstać przez przypadek w ciągu (powiedzmy) 3 mld lat? Można

pokusić się o oszacowanie stosownego prawdopodobieństwa, lecz z pewnością nie będzie ono zbyt wysokie, co stanowi samo w sobie dość istotny problem dla FO. Wersja mniej ogólna pytania dotyczy możliwości powstania kierowanego algorytmem ludzkiego mózgu w drodze ewolucyjnych zmian układu nerwowego.

FO przyjmuje, że działaniu mózgu odpowiada SPI, którą można zakodować w postaci jednego ciągu instrukcji dla komputera o architekturze równoważnej maszynie Turinga. FO ujmuje zatem umysł jako jedną całość, jednolity system przetwarzania informacji posiadający określoną specyfikację. Tymczasem badania nad ewolucją centralnego układu nerwowego oraz badania struktury i funkcjonowania mózgu pokazują, że wyższe funkcje umysłowe, w tym świadomość, są efektem stopniowej akumulacji ewolucyjnych zmian, zaś mózg jest zbiorem mniejszych podsystemów, z których każdy wyspecjalizowany jest w określonym zadaniu. Spojrzenie na mózg jako dość przypadkowy zestaw współpracujących podsystemów zdobywa sobie coraz większe uznanie wśród naukowców, także za sprawą modeli komputerowych i badań nad zachowującymi się inteligentnie robotami [Clark 1997, 2001]. Umysł interpretowany jest jako system złożony z wielu podsystemów przetwarzania informacji również na gruncie psychologii ewolucyjnej [Tooby, Cosmides 1992; Pinker 1997]. Wszystko to podaje w wątpliwość założenie, jakoby SPI mózgu redukowało się do jednego wyczerpującego zbioru instrukcji. Po pierwsze, zespół relacji między różnymi częściami mózgu może okazać się na tyle bogaty, że nie sposób będzie go zakodować w takiej postaci. Po drugie, opisanie za pomocą algorytmu skomplikowanej sieci relacji pomiędzy ewolucyjnie różnymi podzespołami mózgu, wykształconymi w różnym czasie i celu, implikowałoby istnienie wyróżnionego, centralnego poziomu przetwarzania informacji w mózgu. Jeśli koncepcja *bag-of-tricks* jest trafna, istnienie takiego poziomu jest wykluczone.

**EFEKTYWNE PRZETWARZANIE INFORMACJI.** Kolejny powód wątplenia w ewolucyjne pochodzenie algorytmicznego umysłu stanowi fakt, iż z punktu widzenia zdolności do przeżycia organizmu, poszukiwanie najbardziej efektywnego sposobu przetwarzania informacji w warunkach przypadkowych zmian, tak charakterystycznych dla środowiska przyrodniczego na Ziemi, niekoniecznie wymaga posłużenia się jakimś algorytmem. Przypadkowe błądzenie w pejzażu adaptacyjnym *per saldo* okazać się może znacznie bardziej opłacalne z perspektywy organizmu niż poszukiwanie rozwiązania na drodze predefiniowanych reguł. Część badaczy twierdzi, że odpowiedzią na tak postawiony zarzut może być istnienie algorytmów wykorzystujących naturalne generatory liczb losowych do przełamywania ograniczeń algorytmicznego przetwarzania informacji i zwiększenia ich faktycznej efektywności. Nie wskazują oni jednak na ani jeden przypadek realnego istnienia takiego generatora w przyrodzie i wykorzystywania go przez organizm. Co więcej, wyjaśnienie zdolności adaptacyjnych organizmów wydaje się o wiele bardziej prawdopodobne w kategoriach wykorzystania przez nie niealgorytmicznych i w tym sensie przypadkowych mechanizmów przetwarzania informacji niż algorytmów korzystających z generatorów liczb losowych, do tego o nieustalonym, jak na razie, pochodzeniu.

**EMERGENCJA.** Emergencja stanowi dodatkowy powód do wątpienia w szanse powodzenia programu FO. W odpowiednich warunkach, w układach o odpowiednim poziomie złożoności, zachodzić mogą procesy samoorganizacji prowadzące do powstania własności emergentnych. Cechą takich własności jest niemożliwość zredukowania ich do zachowania się elementów składowych układu posiadającego własności emergentne. Nie da się wykluczyć, że systematyczne zachowanie się organizmu jest właśnie zjawiskiem emergentnym, wyłaniającym się z interakcji pomiędzy obwodami i ośrodkami nerwowymi. Lecz jeżeli umysł jest po prostu przejawem samoorganizacji układu neuronów, próżno oczekiwać, że może zostać w pełni wyjaśniony w kategoriach ich działania. Ewolucyjne pochodzenie układu nerwowego może zostać potraktowane jako argument na rzecz stwierdzenia, że w przypadku neuronów istotnie mamy do czynienia z procesem samoorganizacji.

**POSZERZONY UMYŚL.** Ostatni z problemów, na które chciałbym zwrócić uwagę, wiąże się z koncepcją tzw. poszerzonego umysłu [Clark, Chalmers 1998]. Jej centralny punkt stanowi teza, iż mózg wykorzystuje do przetwarzania informacji fizyczne właściwości środowiska i samego organizmu. Świadczą o tym liczne badania dotyczące rzeczywistej realizacji przez mózg takich funkcji jak śledzenie obiektów wzrokiem czy określanie przybliżonej odległości od nich. A także obserwacje udziału sposobów organizacji otoczenia w zapamiętywaniu oraz wykonywaniu czynności o charakterze eksperckim. Wszystko to skłania do przyjęcia, że umysł rozumiany jako struktura przetwarzania informacji realizowana przez mózg nie ogranicza się do obszaru czaszki, lecz obejmuje również ciało i jego fizyczne otoczenie. Jednak skoro fizyczne otoczenie mózgu jest faktycznie wykorzystywane jako część procesu rozwiązywania problemów, trudno doszukiwać się całości danych na temat aktywności poznawczej i zachowań adaptacyjnych wyłącznie w obrębie neuronów.

Każda teoria umysłu musi więc opisywać rzeczywistość SPI mózgu, czyli brać pod uwagę rzeczywiste środowisko fizyczne umysłu. A to oznacza, że zwolennicy FO przyjmować muszą, iż zasadniczo możliwy jest algorytmiczny opis nie tylko zachowania się samych neuronów, ale również otaczającej nas rzeczywistości, aktywnie wykorzystywanej przez mózg do tworzenia stanów mentalnych. Założenie takie oznacza, że nie tylko działanie mózgu, ale cała rzeczywistość fizyczna może być symulowana przez odpowiednio skonstruowany komputer. Niewiele osób będzie jednak skłonnych zaakceptować tak radykalną konsekwencję FO<sup>5</sup>.

Niezależnie od tego, koncepcja poszerzonego umysłu stwarza inny, równie poważny problem. Zgodnie z nią rozróżnienie poziomów przetwarzania informacji (*software*) oraz implementacji (*hardware*) jest po prostu błędne – sposoby, w jakie mózg przetwarza informacje, są ściśle zależne od otoczenia, jakim dysponuje mózg. Trudno zatem utrzymywać, że umysł to fenomen czysto informacyjny, który zasadniczo może zostać zrealizowany w dowolnym środowisku fizycznym. Twierdzenie to z pewnością należy rozumieć mniej dosłownie, niż życzyliby tego sobie zwolennicy FO.

<sup>5</sup> Wyjątkiem z pewnością będzie D. Deutsch, twórca teorii kwantowych obliczeń.

## Podsumowanie

Zebrane w tym artykule uwagi miały pokazać, jak słabe są podstawy wiary, inspirowanej rozwojem komputerów oraz pracami nad tzw. sztuczną inteligencją (programami symulującymi zachowanie się złożonych organizmów czy funkcje układu nerwowego), że umysł wyjaśnić można w kategoriach wykonania komputerowo realizowalnej mechanicznej procedury. Oprócz dobrze znanych argumentów odwołujących się do ograniczeń takich procedur, FO rysuje się jako stanowisko co najmniej kłopotliwe również w świetle ewolucyjnego ujęcia źródeł oraz biologicznych realiów ludzkiego mózgu.

## BIBLIOGRAFIA

- Clark A. 2001, *Mindware – An Introduction to The Philosophy of Cognitive Science*, Oxford University Press, New York.
- Clark A., Chalmers D. 1998, *The Extended Mind*, „Analysis” 58, s. 10–23.
- Newell A. 1980, *Physical symbol systems*, „Cognitive Science” 4, s. 135–183.
- Pinker S. 1997, *How the Mind Works*, W.W. Norton, New York.
- Putnam H. 1967, *The Nature of Mental States* [w:] *Art, Mind and Religion*, red. W.H. Capitan i D.D. Merrill, University of Pittsburgh Press, Pittsburgh.
- Searle J. 1992, *The Rediscovery of the Mind*, MIT Press, Cambridge, MA.
- Tooby J., Cosmides L. 1992, *The psychological foundations of culture* [w:] *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*, red. J.H. Barkow, L. Cosmides, J. Tooby, Oxford University Press, New York, s. 19–136.