

# Czy komputery będą kiedyś świadome?

Włodzisław Duch

Katedra Informatyki Stosowanej UMK

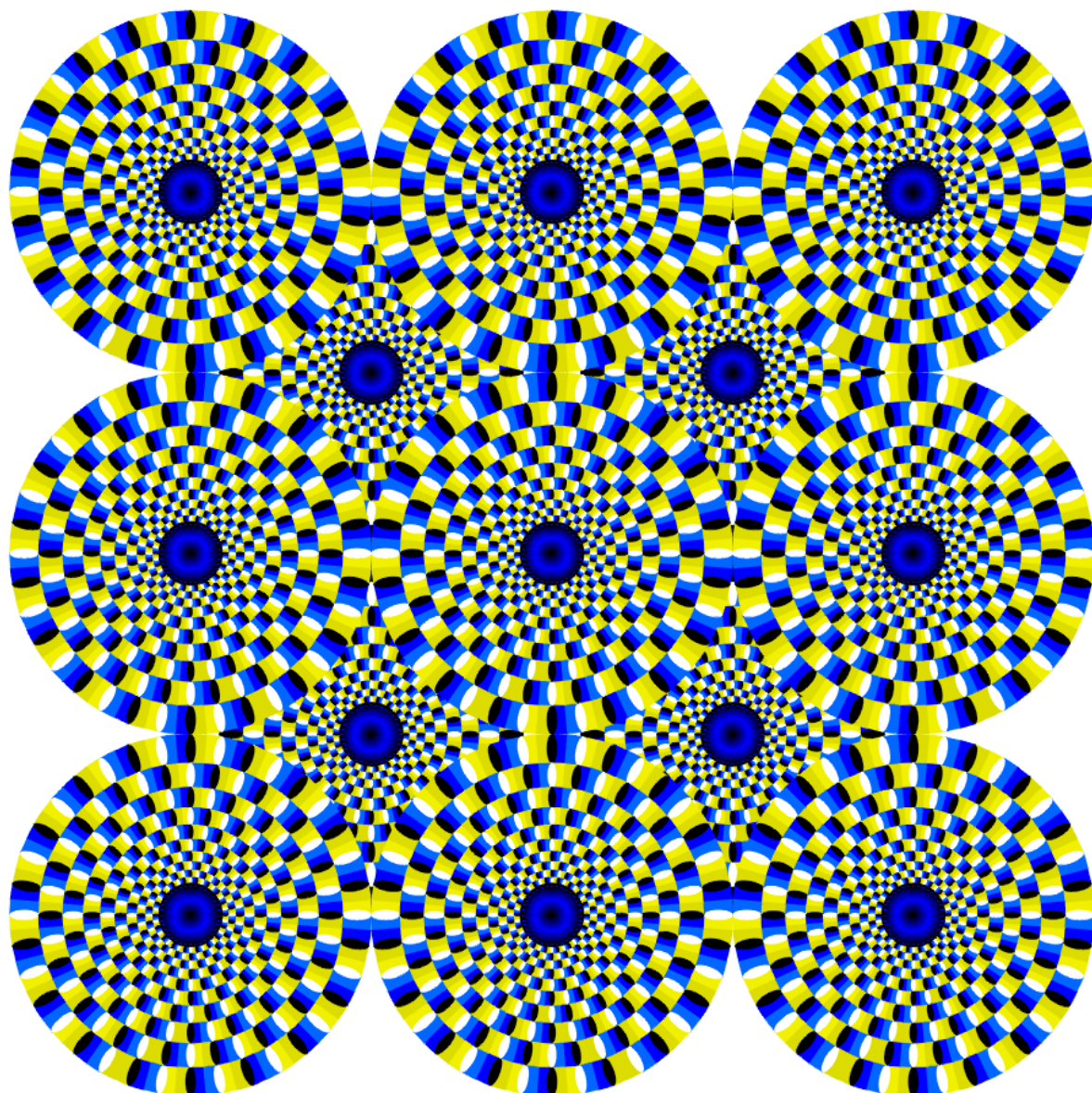
Google: W. Duch

Pełne zrozumienie natury ludzkiej jest obecnie jednym z najważniejszych wyzwań stojących przed nauką, a przyswojenie tej wiedzy będzie nie mniejszym wyzwaniem dla nauk humanistycznych a także tych religii, które chcą być w zgodzie z nauką. Tradycyjne spojrzenie na świat, zaprzeczające istnieniu biologicznie uwarunkowanej natury ludzkiej podsumował Steven Pinker w wydanej w Polsce w 2004 roku książce pod tytułem „Tabula rasa. Spory o naturę ludzką”. W bardziej dosłownym tłumaczeniu książka ta nosi tytuł „Tabula rasa. Współczesne wypieranie się natury ludzkiej”, znacznie silniej akcentujące przekonanie autora, że nadal ulegamy trzem mitom, leżącym u podstaw rozumienia ludzkiej natury: człowiek to niezapisana tablica, wychowanie może ukształtować jego umysł w dowolny sposób; człowiek pierwotny był z natury dobry, zepsuła go cywilizacja; niematerialny duch podejmuje wolne decyzje i kieruje działaniami materialnego ciała.

Genetyka zachowania (zwana również genetyką kognitywną) pokazała, że nie tylko poważne upośledzenia, lecz wiele cech charakteru i osobowości ma biologiczne podłoże. Badanie związków pomiędzy genami a różnego rodzaju zaburzeniami rozwojowymi i chorobami psychicznymi jest obecnie jedną z najszybciej rozwijających się gałęzi nauk o mózgu. Zagadnienie jest oczywiście bardzo skomplikowane, gdyż wymaga zrozumienia roli licznych, współpracujących ze sobą genów odpowiedzialnych za budowę neuronów i innych komórek wpływających na działanie układu nerwowego, powiązania procesów z udziałem dużych grup neuronów z zachowaniem organizmu i jego możliwościami poznawczymi. Końcowym etapem tego procesu jest próba zrozumienia, jak stany mózgu tworzą nasz świat wewnętrzny, nasze myśli, wrażenia, radości i smutki. O ile zrozumienie korelacji aktywności genów z zachowaniem i możliwościami poznawczymi zwierząt i ludzi nie stwarza większych problemów to relacje pomiędzy stanami mózgu i stanami umysłu wydają się wielu ludziom nadal tajemnicze. Dlaczego tak się dzieje? Regulacja wewnętrznego stanu organizmu może odbywać się w większości w sposób automatyczny, nie wymagając podejmowania wielu decyzji. Przetrwanie we wrogim środowisku wymaga przewidywania, planowania i reagowania na różne niebezpieczeństwa, a więc reprezentacji i szybkiego przetwarzania

informacji przydatnej do działania w świecie. Postrzegamy, podejmujemy decyzje i działamy. Skomplikowane mechanizmy poznawcze i emocjonalne, odpowiedzialne za te procesy są głęboko ukryte przed świadomym umysłem. Ich znajomość mogłaby jedynie przeszkodzić w sprawnym działaniu, tak jak myślenie o mechanice działania fortepianu może przeszkodzić pianiście w wykonaniu skomplikowanego pasażu. Komputery również wykonują określone zadania nic nie wiedząc o swojej konstrukcji. Są więc dobre powody, dla których niewiele wiemy o swoich umysłach i sama introspekcja niewiele nam powie.

Dopóki mózg dobrze działa wytwarzany przez niego obraz świata i reprezentacja „ja” w tym świecie wydaje się nie mieć wiele wspólnego z jego pobudzeniami. Wiemy jednak ponad wszelką wątpliwość, że takie powiązanie istnieje. Jak należy je rozumieć? Świat, który widzimy, nie jest rzeczywistym światem, widzimy i odczuwamy jedynie stany swojego umysłu. To, co jednemu z nas śmierdzi, innemu może pachnieć, zapach nie istnieje więc obiektywnie. To, co jednemu smakuje, innemu wydaje się obrzydliwe, smak również nie istnieje obiektywnie. Zgadza się więc, że świat smaków i zapachów jest bardzo subiektywny, ale to samo dotyczy innych zmysłów. Różnimy się słuchem, i to nie tylko słuchem muzycznym: wiele dźwięków mowy wyraźnie odmiennych dla chińskiego czy wietnamskiego ucha brzmi dla nas jednakowo. Dla japońskiego ucha „rój” i „luj” brzmi dokładnie tak samo. Osoby cierpiące na rozmaite afazje akustyczne nie rozpoznają odgłosów zwierząt lub sygnałów akustycznych. Każdy z nas widzi też świat nieco inaczej, dostrzega w różny sposób szczegóły, ruch, kolory. Złudzenia wzrokowe pokazują, jak statyczne obrazy kolorowych kół poruszają się, lub ten sam kolor wywołuje całkiem inne wrażenie w zależności od otoczenia. Zmysł równowagi i czucia też nie jest wyjątkiem i wprowadza nas czasem w błąd, tworząc wrażenie unoszenia się w powietrzu, poczucie czyjeś obecności w pustym pokoju i inne dziwne odczucia.



[Rysunek ruchomy kół, ze strony WWW Akiyoshi Kitaoka]

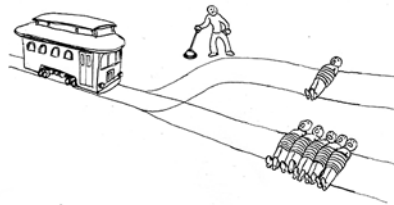
Jaki jest więc naprawdę ten świat? Zależy dla kogo! Postrzegamy wszystko w sposób subiektywny, gdyż odczuwamy tylko stany swojego mózgu. Mniej niż jedno na dziesięć milionów połączeń w mózgu łączy go ze światem zewnętrznym (dla wzroku mamy około 1.2 mln, a dla słuchu zaledwie 30.000 połączeń, podczas gdy liczba wewnętrznych połączeń w mózgu jest rzędu 100 bilionów). Korelacja stanu pierwotnej kory wzrokowej z sygnałami dochodzącymi z siatkówki wynosi zaledwie kilkanaście procent, za pozostałe ponad 80% odpowiadają informacje płynące z innych obszarów kory. Mózg zajmuje się więc przede wszystkim sam sobą, analizuje i interpretuje własne pobudzenia. Niezwykle gęsta sieć połączeń neuronów stwarza przestrzeń, w której możliwe jest bogate życie wewnętrzne.

Przypisujemy podejmowane decyzje swojemu 'ja', ale badając pobudzenia mózgu metodami neurobrazowania lub analizując jego elektryczne potencjały (czyli elektroencefalogram, EEG) można w niektórych eksperymentach dowiedzieć się jaką decyzję podjął mózg zanim dowie się o tym jego właściciel! Mówiąc bardziej precyzyjnie, można zobaczyć decyzję zanim procesy tworzące 'ja' przez ten mózg się o tym dowiedzą! O tym, jaką decyzję podjął mózg 'ja' dowiaduje się obserwując aktywność kory czołowej, a więc obszarów kontrolujących działanie, ruchy ciała. Działając na mózg zmiennym polem magnetycznym można też wpłynąć na podejmowane decyzje w sposób niezauważalny dla uczestniczącej w eksperymencie osoby: np. jeśli ma ona przypadkowo naciskać przyciski raz prawą a raz lewą ręką będzie przez 80% czasu naciskać przycisk prawą ręką, cały czas będąc przekonana, że podejmuje wolne decyzje.

Stosowane obecnie metody podglądania aktywności mózgu są nadal bardzo prymitywne. Początkowo metody te pozwalały odróżnić tylko proste wyobrażenia lub intencje dotyczące ruchu lewą lub prawą ręką lub nogą, ale w 2007 roku pokazano, że można też odróżnić znacznie bardziej subtelne stany mózgu, związane z podjęciem decyzji czy dodać czy odjąć dwie liczby od siebie. W jaki sposób podejmujemy decyzje? Pojawiają się w naszych umysłach i przypisujemy je swojemu 'ja', które w jakiś tajemniczy sposób ma dysponować wolą. Taki opis sytuacji oczywiście niczego nie wyjaśnia. Czasami pojawiają się też decyzje alternatywne, wahamy się pomiędzy kilkoma rozwiązaniami, ale ponieważ nie jesteśmy świadomi większości procesów zachodzących w mózgu nie zdajemy sobie sprawy z tego, skąd się te decyzje biorą. Najbardziej prawdopodobny mechanizm to szybkie, równoległe wyobrażenie sobie różnych sytuacji, zachodzące w korze przedczołowej i ich jednoczesna emocjonalna ocena. Konkurencja pomiędzy konfiguracjami pobudzeń neuronów realizujących te procesy pozostawia tylko te decyzje, które są optymalne z punktu widzenia spodziewanych emocji. Oceny dotyczą nie tylko wyników bezpośrednich działań, lecz mogą uwzględnić też dłuższą perspektywę, dlatego zgadzamy się na chwilowe cierpienie oczekując późniejszej nagrody.

Ten sam mechanizm pozwala nam rozstrzygać moralne dylematy. Pacjenci z uszkodzonymi obszarami brzuszno-przyśrodkowej części kory przedczołowej w prawej półkuli mózgu (rvm-PFC) mają problemy nie tylko z ocenami emocjonalnymi, ale również niektórymi ocenami moralnymi. Typowa sytuacja, w której bada się moralne zachowania obserwując jednocześnie działanie mózgu wygląda tak: wyobraź sobie, że kontrolujesz zwrotnice na kolei i widzisz szybko nadjeżdżający z zakrętu pociąg, a na torach pracują robotnicy. Jeśli nie zrobisz nic zginie pięć osób, jeśli przestawisz zwrotnicę pociąg zabije

tylko jedną osobę. Prawie wszyscy w takiej sytuacji, łącznie z pacjentami z uszkodzonym obszarem rvm-PFC, decydują się na przestawienie zwrotnicy. Różnice dotyczą problemów, w których działanie powoduje dotkliwą stratę osobistą. Wyobraź sobie, że mieszkasz w wiosce, która jest przeszukiwana przez barbarzyńskich najeźdźców; wszyscy schowali się pod podłoga jednego z domów i zginą, jeśli zostaną odkryci. Twoje małe dziecko zaczyna płakać, czy będziesz go powstrzymywać aż się zadusi, czy też pozwolisz mu płakać, a tym samym wszystkim zginąć? W sytuacji takich osobistych konfliktów moralnych ujawniają się wyraźne różnice. Pacjenci z uszkodzonym obszarem rvm-PFC ponad dwukrotnie rzadziej wybierają zaduszenie własnego dziecka. W mniej drastycznych testach nie są zdolni by poświęcić krótkoterminowe zyski dla wysoce prawdopodobnej wygranej w dłuższym czasie. Mózg ma wysoce wyspecjalizowane obszary, których uszkodzenia powodują specyficzne, często bardzo subtelne objawy.



[Rysunek znaleziony gdzieś w sieci, niestety nie pamiętam gdzie]

Postęp w dziedzinie obserwacji aktywności mózgu jest szybki i można sobie wyobrazić, że w przyszłości podglądanie aktywności mózgu pozwoli nam poznać prawdziwe intencje, upodobania, a nawet skryte myśli, których 'ja' tworzone przez tenże mózg jeszcze nie zna! „Poznaj samego siebie” nabierze wówczas całkiem nowego znaczenia. „Tyle o sobie wiemy, ile nas sprawdzono” również. Ponieważ moje 'ja', jeden z wielu procesów, jakie zachodzą w mózgu, nie do końca potrafi przewidzieć reakcje całego mózgu, potrafimy sami siebie zaskoczyć. Czasami sam się siebie pytam, czemu postąpiłem w sposób niezgodny ze swoimi przekonaniem lub oczekiwaniami? Nie wiem, podkusiło mnie jakieś lichy, „to” było silniejsze ode mnie. Czym jest to 'ja' i czym jest „to coś” silniejsze niż 'ja', które się „siebie” pyta? Dopóki myślimy, że 'ja' mam mózg, a nie mózg ma „mnie”, takie paradoksalne stwierdzenia pozostaną niezrozumiałe. Tradycyjnie psychologowie powoływali się tu na pojęcie „podświadomości”, ale nauki o mózgu oferują odmienny punkt widzenia. 'Ja' jest tu wewnętrznym modelem, zbiorem wyobrażeń o sobie pozwalającym na niezbyt doskonale przewidywanie decyzji całego mózgu. Siedliskiem tych procesów jest kora przyśrodkowa, kora zakrętu obręczy, które aktywizuje się w testach werbalnych, przestrzennych,

emocjonalnych, związanych z rozpoznawaniem własnej twarzy, czyli w sytuacjach kontrastujących to, co odnosi się do 'ja' z tym, co odnosi się do innych. Są to dobrze ukryte obszary mózgu, rzadko ulegające uszkodzeniom, pośredniczące w komunikacji pomiędzy licznymi podkorowymi strukturami układu limbicznego, odpowiedzialnego za emocje i motywacje, pniem mózgu regulującym stan całego organizmu, ogólną przytomność i mechanizmy uwagi, oraz innymi obszarami kory mózgu.

'Ja' nie jest wcale monolitem. Na to, co określamy jako 'ja' składa się wiele funkcji. „Proto-ja” związane jest z podstawowymi informacjami o stanie ciała i wewnętrznym stanie mózgu, informacji, które mózg interpretuje jako poczucie istnienia. Jestem, bo czuję. Uściśleniem tych informacji jest orientacja 'ja' w przestrzeni: odczuwam świat z punktu widzenia swojego ciała, które zajmuje określoną pozycję względem innych obiektów. Autobiograficzne 'ja' wiąże się z wspomnieniami o sobie w różnych sytuacjach, pamięcią siebie, która nas określa. Społeczne 'ja' oparte jest na relacjach z innymi ludźmi, rolami, w które się wcielamy. Znaczna część tych relacji jest trudna do werbalizacji, gdyż reakcje mózgu są zbyt złożone, by je opisać za pomocą słów. W szczególności działanie prawej półkuli mózgu oraz interpretacje stanów emocjonalnych, zmieniających się w ciągły sposób trudno jest uchwycić w słowach. W efekcie nie zawsze potrafimy sformułować, o co „nam” chodzi. Mój mózg jest mądrzejszy ode mnie, czyli od jednego ze swoich licznych procesów.

Czy można w pełni zintegrować swój model „ja”, poznając „siebie”? Co by to oznaczało? Opis tego, co dzieje się w mózgu, odróżniając przyczyny od skutków, nie jest możliwy przynajmniej z dwóch powodów. Po pierwsze, żaden z obszarów mózgu nie jest niezależny od pozostałych, ich aktywność jest na tyle silnie sprzężona, iż każdy z obszarów bierze w różnym stopniu udział w podejmowaniu decyzji czy przetwarzaniu informacji. Prosta przyczynowość liniowa typu „najpierw bodziec, potem reakcja” w takiej sytuacji nie ma zastosowania. Nawet w przypadku rozpoznawania przedmiotów obraz z siatkówki oka wędruje nie tylko przez pierwotną korę wzrokową mieszczącą się z tyłu głowy (w płacie potylicznym kory) w stronę dolnej części kory skroniowej, w której następuje rozpoznawanie obiektów. Informacja płynie też w przeciwnym kierunku oraz w inne strony, uzgadniając z czym mamy do czynienia dzięki współpracy pomiędzy korą wzrokową (dostarczającą dokładnych informacji), dolną korą skroniową (dostarczającą wstępnych hipotez) i korą potyliczną (pozwalającą skupić wzrok na miejscu, w którym jest obserwowany przedmiot). Próby komputerowej analizy obrazu, które pomijały ten skomplikowany scenariusz, pomimo wielu lat wysiłków nie zaprowadziły zbyt daleko. Jeśli do rozwiązania problemu potrzebna jest grupa specjalistów, których kompetencje częściowo się nakładają, choroba lub eliminacja

jednego z nich zmieni sposób działania całego zespołu, ale funkcje brakującego eksperta częściowo przejmą jego koledzy, chociaż kompetencja całego zespołu może się zmniejszyć.

Drugim powodem trudności opisu werbalnego działania mózgu są słabe, stopniowe zmiany aktywności w przypadku większości procesów, a więc działania nieświadome, konkurujące ze sobą równoległe scenariusze, pozwalające na szukanie optymalnych, konkurencyjnych interpretacji czy odpowiedzi. Dopiero po wstępnej ocenie emocjonalnej pobudzenia mózgu związane z najlepszymi rozwiązaniami osiągają taki poziom, przy którym inne części mózgu, np. te odpowiedzialne za mowę, mogą je jednoznacznie zidentyfikować. Narastanie neuronalnej aktywności interpretowane jest jako przeczucie, że odpowiedź jest już blisko i za chwilę uda się nam ją rozpoznać, a wtedy możemy sobie z tego „zdać sprawę”, wypowiedzieć się lub podjąć jakieś inne działanie. Dlatego próba pełnego zrozumienia sposobu działania mózgu, w sensie jego werbalnego opisu (np. w stylu psychoanalizy), nie może nigdy się w pełni udać. Możemy jedynie podpierając się wynikami eksperymentów tworzyć coraz doskonalsze modele komputerowe zachodzących w mózgu procesów.

Głębsze poznawanie „siebie” było dotychczas domeną psychoterapii i ścieżek duchowego rozwoju, wywodzących się z różnych tradycji religijnych. Umysł, stawiający sobie cele duchowe, pokazuje swoją autonomię, porównując się do jakiegoś ideału może wpłynąć na swoje działanie, przez trening mentalny zmienić powstające w mózgu odczucia tak, by je w pełni zintegrować ze swoimi wyobrażeniami i celami. Dalaj Lama często mówi o potrzebie kształcenia nie tylko intelektu, ale również „ciepłego serca”, współczucia i zdolności empatii. Znaczna część buddyjskich (w tym tybetańskich) praktyk religijnych była bezpośrednio nakierowana na taki trening umysłu. W nieco bardziej zawaolowany sposób podobne praktyki obecne są w innych tradycjach religijnych. Rozpoznanie centralnej roli emocji w procesach uczenia się i podejmowania decyzji, oraz potrzeba posiadania dobrych wzorców, pokazuje konieczność modyfikacji programów kształcenia. Na razie sprawdzanie reakcji emocjonalnych na różne produkty czy wypowiedzi polityków próbuje się stosować w neuromarketingu. Miejmy nadzieję, że najnowsze odkrycia badaczy mózgu pozwolą nie tylko na manipulację wyborcami czy klientami, ale zaprowadzą nas znacznie dalej na drodze do odkrycia naszych prawdziwych potrzeb, w poszukiwaniu „duchowego siebie”.

Nie ma żadnych wiarygodnych dowodów na istnienie umysłu bez mózgu. Coraz więcej odkryć pokazuje za to, jak pobudzanie różnych obszarów mózgu wywołuje określone wrażenia, jak dziwne efekty spowodować może załamanie się sprawnego działania tych obszarów. Halucynacje wzrokowe, słuchowe czy dziwne odczucia nie zawsze muszą być związane z wyraźnie widocznymi uszkodzeniami, wystarczą lokalne pobudzenia nieco

słabsze niż te, które wywołują ataki padaczki. Do najdziwniejszych zjawisk należą bez wątpienia przeżycia autoskopowe, czyli halucynacje wzrokowe i wrażenie podwojenia jakiegoś organu, kończyny lub całego ciała, widzenia swojego sobowtóra lub wyjścia poza własne ciało i patrzenia na siebie z zewnątrz. Autoskopia spotykana jest najczęściej w schizofrenii, u osób cierpiących na padaczkę, migreny, guzy mózgu, może się też pojawić w stanach ekstremalnych (np. powrotu do normalnego działania mózgu po śmierci klinicznej), lub u osób normalnych w wyniku zaburzeń sposobu reprezentacji przestrzeni. W ostatnich latach porównano ponad sto takich przypadków, a u kilku pacjentów wywołano wrażenia przebywania poza ciałem pobudzając ich mózgi słabym prądem. Okazuje się, że system reprezentacji przestrzeni w mózgu jest dość skomplikowany: inaczej reprezentujemy przestrzeń personalną, bliską ciału, a więc obszar w zasięgu naszego bezpośredniego działania, a inaczej przestrzeń nieco bardziej odległą. Zjawiska autoskopowe związane są z zaburzeniami przestrzeni personalnej z powodu błędnej integracji informacji wzrokowej, kinestetycznej (ruch i równowaga ciała), proprioceptywnej (dostarczanej przez mięśnie i ścięgna) oraz czuciowej. Różne zaburzenia tego układu prowadzą do specyficznych zjawisk autoskopowych.

W przeszłości zjawiska autoskopowe były przedstawiane jako jedne z dowodów istnienia umysłu niezależnie od mózgu. Sny, w których spotykamy zmarłych czy lecimy w odległe strony, interpretowane były jako dowód na wędrówkę dusz, powszechną w religiach szamańskich i w hinduizmie. Mieszkańcy Sumatry wiązali chorego człowieka sznurami by nie uciekła jego dusza. Dusza była tym, co animuje, czyli ożywia martwe ciało. Proste rozwiązania mają na celu stworzenie złudzenia, że coś rozumiemy, ale rzeczywistość jest znacznie bardziej skomplikowana i znacznie ciekawsza. Dzisiaj wszystkie funkcje przypisywane duszy przez teologów znalazły swoje wyjaśnienie, pojęcie to dawno utraciło swoje znaczenie, nawet część teologów próbuje od niego odejść i traktować w sposób metaforyczny. Pojęcie duszy jest jednak głęboko zakodowane w naszych umysłach w postaci rozlicznych metafor, więc z pewnością nie zniknie, niemniej z naukowego punktu widzenia nie ma tu nic do wyjaśnienia. Są niezliczone dowody na to, że mózgi tworzą umysły. Po pierwsze, związek specyficznych uszkodzeń mózgu z zaburzeniami poznawczymi, afektywnymi, zaburzeniami i chorobami osobowości (zwanymi dawniej chorobami duszy) nie pozostawia wątpliwości, że mózg nie jest żadnym odbiornikiem decyzji niematerialnego składnika, lecz te decyzje tworzy. Nie mogę przecież sądzić, że moja dusza podejmuje wolne decyzje, ale moje 'ja' nic o tym nie wie i zamiast przypadkowo wybierać raz lewy, raz prawy przycisk, wybieram ciągle ten sam, bo pole magnetyczne zniekształca odbiór „moich”



decyzji. Taka dusza byłaby jakimś abstrakcyjnym bytem, nie mającym ze mną nic wspólnego. Po drugie, nawet proste neuropodobne modele komputerowe wytwarzają pewne funkcje umysło-podobne, np. charakterystyczne dla naszego myślenia procesy skojarzeniowe, a badanie ich uszkodzeń daje interesujące modele syndromów neuropsychologicznych. Od połowy lat 90. rozwija się dziedzina zwana psychiatrią komputerową, tworząca neuronowe modele wyjaśniające na czym polegają choroby psychiczne i w jaki sposób przejawiają się różne zaburzenia przepływu informacji w mózgu. Im bardziej dokładne symulacje tym więcej funkcji umysłu udaje się w ten sposób stworzyć.

Jak więc należy rozumieć relację umysł – mózg? Na mózg możemy patrzeć jak na substrat, który stwarza przestrzeń dla zdarzeń psychicznych, tak jak konfiguracje atomów są substratem do tego, by tworzyć fizyczne obiekty i zdarzenia z ich udziałem. Na ekranie telewizora mogą pojawiać się różne obrazy dzięki pobudzeniom elementów wysyłających światło. Obrazy na ekranie są tylko konfiguracjami pobudzeń świecących elementów, tak jak obrazy wewnątrz mózgu są konfiguracją pobudzeń mikroobwodów neuronowych. Zasadnicza różnica polega na tym, że neurony ucząc się nabierają skłonności do powtarzania tych samych wzorców pobudzeń, podobnie jak woda spływająca po miękkim zboczach ma tendencję do spływania tymi samymi rowkami i taka erozja tworzy specyficzne struktury w glebie. Takie porównanie pojawia się w indyjskim tekście „Pytania króla Milindy” napisanym około 2100 lat temu. Na pytanie skąd się biorą skłonności u ludzi, buddyjski mędrzec Nagasena odpowiada pytaniem: „Kiedy pada deszcz, któredy płynie woda?” „Będzie płynąć po pochyłościach gruntu.” „A gdyby deszcz spadł ponownie, dokąd by płynęła woda?” „Płynęłaby w tym samym kierunku, co pierwsza woda”. Tak właśnie powstają skłonności a przedstawiony tu opis stanowi podstawę neurologicznej teorii uczenia się, opisaną przez kanadyjskiego psychologa Donalda Hebb’a w 1949 roku. Jego reguła zwykle przedstawiana jest w lapidarnej formie: „neurony, które są często jednocześnie aktywne, silniej się ze sobą łączą”.

W powtarzalnych wzorcach aktywacji sieci neuronowych zawarte są ślady pamięci. Wspomnienia lub wyobrażenia przywołują konfiguracje pobudzeń neuronów zbliżone do tych, które powstały w momencie przeżywania danej sytuacji. Pamięć jest „zamrożoną przeszłością” mózgu. Wynikające z przeżywanych, przypominanych lub wyobrażanych stanów emocje są z jednej strony układem pobudzeń neuronów, lecz z drugiej strony tworzą realne zdarzenia w przestrzeni naszych umysłów; mózg jest tylko substratem pozwalającym im zaistnieć. Fizyka i chemia powie nam jakie siły utrzymują razem atomy, składające się na materiał, z którego zrobiony jest pomnik, jednak nie będzie mogła odpowiedzieć na pytanie

po co go postawiono i dlaczego akurat w tym miejscu. Podobnie neurofizjologia może opisać stany mózgu w terminach pobudzeń neuronów, natomiast to, dlaczego pojawiają się konkretne stany a nie inne wynika z całej osobistej historii, a nawet prehistorii naszego gatunku. Życie wewnętrzne każdego z nas jest niepowtarzalne. Powodów, dla których stara melodia wywołuje we mnie liczne wspomnienia i stany emocjonalne nie można w pełni zrozumieć na poziomie opisu pobudliwości neuronów, potrzebna jest do tego znajomość mojej indywidualnej historii. Sprawny mózg tylko umożliwia powstanie wspomnień i odtworzenie związanych z nim przeżyć, ale informacje o połączeniach neuronów i ich wzajemnej pobudliwości nie wystarczą do nadania sensu powstającym pobudzeniom. Umysł jest strumieniem zdarzeń mentalnych (wewnętrznych interpretacji konfiguracji pobudzeń neuronów) i relacji pomiędzy tymi zdarzeniami. Relacje zakodowane są w mózgu jako tendencje do skojarzeń pomiędzy mentalnymi zdarzeniami. Pomimo materialnego podłoża z wewnętrznej perspektywy umysł nie jest materialny. Jest to jednak całkiem inny rodzaj niematerialności niż wyobrażali sobie średniowieczni czy antyczni filozofowie, dla których niematerialny duch był rodzajem prostej, niezniszczalnej substancji. Złożone życie wewnętrzne wymaga złożonego, dobrze chronionego substratu, jaki znajdujemy wewnątrz czaszki.

Mózg obserwuje sam siebie, zdając sobie sprawę zarówno ze stanu, w jaki otoczenie się w nim „odbija” (metafora umysłu jako lustro pojawiała się w różnych tradycjach filozoficznych), jak i z własnych stanów wewnętrznych. To, co nazywam swoim ‘ja’, ‘moja osobowość’ a także ‘moja świadomość’ to tylko kilka z wielu procesów, które realizuje. Czy świadomość naprawdę jest takim problemem dla nauki jak się to zwykle przedstawia? Każdy sztuczny system o wysokim stopniu złożoności (nazwijmy go artilektem, czyli sztucznym intelektem), zdolny, podobnie jak nasz mózg, do komentowania własnych stanów, musi być też zdolny do „zdawania sobie sprawy” z tego, co się w nim dzieje, a więc do opisu swoich wrażeń, odczuć i reakcji w sposób, który można będzie określić jako świadomy. Artilekt mógłby na przykład powiedzieć: „Cóż to za odcień czerwieni, dawno nie widziałem takiego zachodu Słońca, przypomina mi przyjemne chwile spędzone na dyskusji z moim przyjacielem”. W sztucznym mózgu artilektu nie będzie zaprogramowanych wypowiedzi, będą tworzone na podstawie oceny własnego stanu wewnętrznego opartego na wcześniej zapamiętanych stanach oraz ocenie emocji wywołanych przez informację wzrokową. Świadomość powinna być naturalną własnością systemu zorganizowanego w mózgowopodobny sposób, wynikiem procesu komentowania własnych stanów wewnętrznych. Jeśli tak jest w istocie to świadomość może pojawić się w sztucznych systemach.

Świadomość nie jest czynnikiem sprawczym naszego działania czy myślenia. Nie wiem przecież jak „świadomie” wywołać określone myśli, wpaść na pomysł, mogę jedynie skupić się nad problemem i oczekiwać, że coś mi przyjdzie do głowy. Tak będzie pod warunkiem, że mój mózg został odpowiednio wytrenowany i potrafi skojarzyć ze sobą już znane fakty. Podglądanie mózgu w czasie rozwiązywania problemów, w których pojawia się gwałtownie rozwiązanie (nazywane w psychologii procesem wglądu, lub określane jako proces Eureka!) pokazuje silne zaangażowanie prawej półkuli mózgu, która potrafi analizować problem w sposób intuicyjny, na niewerbalnym, abstrakcyjnym poziomie. Trudności w zdefiniowaniu pojęcia ‘świadomość’ nie przeszkadzają wcale w badaniach, szuka się bowiem takich sytuacji, w których stajemy się nagle świadomi, np. patrząc i niespodziewanie dostrzegając ukryty element obrazu lub alternatywną interpretację w przypadku obrazów niejednoznacznych. Potrafimy obserwować tego typu korelaty świadomości, jak też ocenić z grubsza stan emocjonalny na podstawie rejestrowanej aktywności mózgu. Nie wiemy jeszcze jak odczytać myśli czy też subtelne informacje na temat wrażeń doświadczanych przez daną osobę, ani nawet czy takie informacje zawarte są w sygnałach mierzonych za pomocą obecnie używanych metod. Badania w tej dziedzinie są obecnie prowadzone na szeroką skalę, może to więc być kwestią niezbyt odległej przyszłości.

Modele komputerowe są na razie dość prymitywne, daleko jest do tego by stworzyć sztuczną strukturę, która przypominałaby to, co mamy w mózgu. Trudno nam sobie wyobrazić świadomość w sztucznych urządzeniach, tak jak kiedyś trudno było sobie wyobrazić przesyłanie dźwięków lub obrazów przez druty a potem bez przewodów. Malutkie wgłębienia na płycie DVD niosą informację o dźwięku i obrazie, ale dopiero odtwarzacz i telewizor potrafi je zamienić na rzeczywisty obraz i dźwięk. Sieci neuronowe w mózgu mogą pamiętać wiele wrażeń i faktów ale potrzebny jest cały organizm by „odtworzyć” zawarte w tych sieciach informacje. Artilekty będą miały inne zmysły i inaczej będą oceniać swoje stany wewnętrzne, dlatego ich świat wewnętrzny, a więc i ich świadomość, nie będzie identyczna z naszą (ludzie różnią się tu pod wieloma względami, zwłaszcza w przypadku świadomości refleksyjnej). Ogólne poczucie istnienia może jednak być podobne.

W tej chwili na świecie realizowanych jest kilka dużych projektów zmierzających do stworzenia złożonych, sztucznych systemów, które wykażą jakieś ślady świadomości. W niezbyt odległej perspektywie możliwe staną się symulacje mózgu na poziomie miliardów neuronów i stu trylionów połączeń. Przymiarki do takich symulacji zrobiono już kilka lat temu, ale jedna sekunda symulacji na 27 komputerach trwała 50 dni. W ciągu 10 lat dzięki postępowi w konstrukcji komputerów symulacje o złożoności potrzebnej do opisanie w

szczegółach całego mózgu będą dość rutynowe. W listopadzie 2007 roku zakończono pierwszy etap bardzo ambitnego projektu „Blue Brain”, prowadzonego od 2005 roku przez Politechnikę w Lozannie (EPFL) przy pomocy firmy IBM. Stworzono bardzo dokładny program, który na podstawie obserwacji kolumny kory mózgu symuluje liczącą kilkadziesiąt tysięcy neuronów strukturę ustanawiając pomiędzy neuronami w trzech wymiarach 30 milionów połączeń synaptycznych. Dzięki takim symulacjom możemy się upewnić, czy modele teoretyczne naprawdę wiernie opisują biologiczną rzeczywistość. Tak szczegółowy model kolumny korowej pozwala na robienie doświadczeń w komputerze (*in silico*), integrując w sobie gromadzoną od stu lat wiedzę na temat komórkowej budowy kory mózgu. Następnym krokiem to z jednej strony symulacje neuronów na poziomie molekularnym i genetycznym, potrzebne dla zrozumienia wpływu leków na działanie kory mózgu, a z drugiej strony uproszczenia symulacji indywidualnych kolumn w celu stworzenia modelu całego mózgu.

Zajmie to jeszcze kilka lat, ale na razie wszystko wskazuje na to, że jesteśmy na dobrej drodze.

---

Włodzisław Duch zajmuje się neuroinformatyką kognitywną, jest kierownikiem Katedry Informatyki Stosowanej UMK oraz Prezydentem European Neural Network Society. Jego strona dostępna jest po napisaniu „W Duch” w dowolnej wyszukiwarce WWW.