

MARCIN WRÓBEL
Uniwersytet Wrocławski

Trzy teorie umysłu

Tekst niniejszy jest próbą skonfrontowania z sobą najważniejszych aktualnie koncepcji zmierzających do stworzenia całościowej koncepcji umysłu. Jak można wywnioskować z tytułu, będę zajmował się trzema teoriami, konkurującymi obecnie o palmę pierwszeństwa w filozofii umysłu. Nie oznacza to bynajmniej, że liczba ta wyczerpuje wszystkie istniejące propozycje, wskazuje jednak na te pomysły, które osiągnęły znaczący sukces, mierzony między innymi ich innowacyjnością i przydatnością do zastosowań praktycznych. Postaram się zwięźle i przystępnie pokazać historyczne i teoretyczne korzenie każdej, ich zarówno silne, jak i słabe punkty oraz okoliczności, które utrzymują je ciągle w ścisłej czołówce peletonu i nie pozostają z tyłu. Pośród nich znajdują się rozliczne ważne zdobycze nauk szczegółowych, które uzyskano dzięki dogłębnemu eksplorowaniu lub próbie rewizji założeń ogólnych tych właśnie teorii.

Wydawać by się mogło, że skoro tytułowe teorie konkurują z sobą o miano najtrafniejszej, to są nieporównywalne i zupełnie różne. W praktyce okazuje się jednak, że zachodzą między nimi dość silne związki pokrewieństwa. Nierzadko udzielają one zbliżonych odpowiedzi na te same pytania i postulowane rozbieżności między nimi okazują się bardziej deklaracjami ich apologetów niż rzeczywistymi różnicami. Zazwyczaj o inności decydują detale lub zaangażowanie w inną dziedzinę przedmiotową, którą nie zajmują się rywalki. Ich uzgodnienie między sobą wydaje się możliwe, choć z pewnością niełatwe, nieco miejsca poświęcę więc próbie rozważenia ewentualnych korzyści, które mogłyby płynąć z wykonania takiego przedsięwzięcia.

1. Poznać znaczy zrozumieć

Zagadnienie ludzkiego umysłu jest jedną z naukowych zagadek, która od dawna skutecznie opiera się wszelkim próbom ostatecznego rozwiązania. Zdając sobie

sprawę, iż stopień komplikacji tego problemu prawdopodobnie przerasta możliwości jakiegokolwiek pojedynczej dyscypliny nauki, badacze różnych specjalności połączyli swoje wysiłki. Ten skonsolidowany multidyscyplinarny front naukowy znany jest dzisiaj pod nazwą kognitywistyki.

Wewnątrz tego, nieco nieformalnego, ruchu da się wyróżnić kilka podgrup zainteresowanych odmiennymi aspektami problemu umysłu. Mają one nie tylko odmienne obiekty zainteresowań, ale także często bardzo różne metodologie i perspektywy teoretyczne. Utrudnia to w znacznym stopniu komunikację między nimi i w efekcie transfer wyników wewnątrz całego ruchu jest dość skromny. Jednak mimo tych przeszkód współpraca istnieje i osiągnięto wiele spektakularnych wyników naukowych.

Zamieszczam tu krótką charakterystykę wspomnianych podgrup, opartą na bardzo prostej typologii.

Grupa Bio: w ramach nauk o poznaniu znajdujemy szerokie grono specjalistów z nauk przyrodniczych, eksplorujących wszelkie „żywe” komponenty, biorące udział w generowaniu świadomości¹. Zaliczyłbym tutaj wąskie biologiczne specjalności, jak neurobiologia, neurofizjologia, neurofarmakologia, psychologia poznawcza, a także subdyscypliny medyczne — neurologię i neuropsychologię. Przedstawiciele tych nauk atakują zagadnienie umysłu metodą „bottom up”, wychodząc od fizjologii pojedynczych neuronów, następnie całych ich grup aż do rozległych szlaków nerwowych, przebiegających przez wiele specjalizowanych ośrodków mózgowych. Zasadniczym zadaniem do wykonania jest dla nich odtworzenie mechanizmów generowania świadomości, wyjaśnienie, w jaki sposób spójna aktywacja odległych obszarów mózgu umożliwia wykonywanie ruchów dowolnych, kierowanie uwagą czy ustalenie stabilnej wewnętrznej reprezentacji otoczenia. Ważnym elementem wiedzy potrzebnej do wykonania tego zadania jest powiększający się zasób informacji o szczegółach anatomii funkcjonalnej — lokalizacji specjalizowanych grup komórek nerwowych, realizujących określone zadania, jak przetwarzanie mowy, utrzymywanie stanu czuwania, obracanie obiektu w przestrzeni czy wykonywanie obliczeń arytmetycznych².

¹ Związek między świadomością a umysłem nie jest do końca jasny. Zazwyczaj przyjmuje się, że umysł jest zbiorczą nazwą wszelkich wyższych procesów psychicznych, enigmatycznym tworem umożliwiającym samo ich zaistnienie. Świadomość uważa się za najważniejszą postać, w jakiej manifestuje się działanie umysłu. Umożliwia ona kontrolę wykonania i ocenę jakości zachowań oraz stwarza prywatną perspektywę poznawczą. Przedstawiciele nauk biologicznych zazwyczaj zajmują się raczej świadomością jako funkcją działania mniejszego bądź większego zbioru komórek nerwowych, dlatego też w ich siatce pojęciowej umysłu zazwyczaj się nie znajduje. Z kolei teoretycznie nastawieni przedstawiciele nauk inżynierskich i formalnych chętniej mówią o umyśle, próbując modelować jego działanie, zostawiając świadomość biologizującym przyrodnikom.

² Podejście to nosi nazwę hipotezy lokalizacyjnej i jest najaktualniejszym stanowiskiem neuroanatomów w kwestii związku mózgu z umysłem. Za jej pioniera uważa się Wezaliusza (XV/XVI w.), a za współczesnych „ojców-założycieli” badaczy: Paula Broca, Manfreda Wernickego (XIX w.) oraz Karla Brodmanna (XX w.). Teoria ta zdetronizowała dwie inne ważne koncepcje: doktrynę komór (Galen II w., Nemezjusz i św. Augustyn, IV/V w., Awicenna X/XI w., Kartezjusz XVII w.) oraz frenologię (Thomas Reid, Douglas Stewart XVIII w., Franz Joseph Gall, Johan Caspar Spurzheim XIX w.).

Jako że nie wszystkie badania da się konkluzywnie przeprowadzać na ludziach, w sukurs anatomom przychodzą badacze świata zwierząt — prymatolodzy, zoolodzy, psychologowie, anatomicy i psychologowie porównawczy. Spójnej perspektywy teoretycznej starają się im dostarczać socjobiolodzy, etolodzy poznawczy i behawioralni, biolodzy ewolucyjni oraz genetycy. W ich wypadku porusza się między innymi zagadnienie adaptacyjności świadomości jako cechy biologicznej, genezy jej powstania, sposobu utrwalania i dziedziczenia oraz problem ekskluzywności treści mentalnych dla świata ludzi.

Grupa Robo: drugą frakcję stanowią specjaliści od nauk robotycznych i formalnych. Cybernetycy, robotycy i elektronicy zajmują się głównie implementacją modeli poznawczych „wydestylowanych” z żywych organizmów układów nieożywionych oraz organizowaniem struktur fizycznych zdolnych do przetwarzania informacji, jakich nie sposób spotkać w naturze. Ważną gałęzią zainteresowań jest modelowanie złożonych ludzkich zdolności poznawczych i konstruowanie maszyn zdolnych do takich zachowań. W całej galerii nowinek technicznych na uwagę zasługują niewątpliwie udane konstrukcje w rodzaju robota zdolnego do widzenia stereoskopowego i kategoryzowania postrzeganych obiektów, maszyny czytającej na głos, układu zdolnego do rozpoznawania ludzkich twarzy oraz emocji czy wreszcie małego robota uczącego się adaptacyjnie patrzeć na centralną część swojego pola widzenia.

Nad teoretyczną spójnością całego technicznego przedsięwzięcia czuwają eksperci od nauk formalnych — logicy, matematycy, informatycy oraz lingwiści. Ich zadania są bardzo różne, obejmują szerokie spektrum zagadnień, począwszy od konceptualizowania modeli przepływu informacji w układach fizycznych, przez tworzenie inteligentnego oprogramowania, modelowanie wzorców przetwarzania języków naturalnych i sformalizowanych czy reprezentowanie otoczenia aż po projektowanie zaawansowanych algorytmów rozwiązywania problemów i układów decyzyjnych zdolnych do egzekwowania ruchów dowolnych, planowania strategicznego czy złożonego, adaptacyjnego uczenia się. Tutaj stosuje się taktykę „top-down”, wychodząc od ogólnych schematów działania i abstrakcyjnych modeli funkcjonalnych umysłu, schodząc w dół do ich fizycznych implementacji.

Grupa Teo: tutaj swoje miejsce mają przedstawiciele tych specjalności, którzy ze względu na przedmiot swoich badań nie mieszczą się w paradygmacie nauk ścisłych czy też formalnych, a też są zainteresowani stworzeniem efektywnej teorii umysłu. Zaliczyłbym do nich psychologów rozwojowych i ogólnych, socjologów i pedagogów³ oraz filozofów umysłu rzecz jasna.

Zwracam uwagę czytelnika na dysproporcję między liczbą filozofów zaangażowanych w tworzenie zaktualizowanych koncepcji a liczbą tych związanych z aktualną propozycją.

³ Obecność pedagogów może być pewnym zaskoczeniem, jednak uwidacznia ważny trend, jaki niedawno zrodził się w łonie tej nauki. Oto stosowny, inspirujący cytat na potwierdzenie tezy: „W ostatnich kilku latach zostały zapoczątkowane kontakty między pedagogami a uczonymi zajmującymi się mózgiem. Wiosną 2000 roku jedna z autorek pracowała 3 miesiące w Parliamentary Office of Science and Technology (POST), oddelegowana tam w ramach doktoratu z neurobiologii w londyńskim University College. Do obowiązków POST należy dostarczanie brytyjskiej Izbie Gmin i Izbie Lordów sprawozdań na temat aktualnych problemów naukowych [...]. Uczestnicząc w tego rodzaju debatach, w czerwcu

Dla psychologów walor terapeutyczny pełnej teorii umysłu jest nie do przecenienia, socjolodzy mogliby na jej podstawie zabrać się za skuteczną inżynierię społeczną, pedagodzy zaś kultywować skuteczną i dostosowaną do indywidualnych potrzeb ucznia edukację. Dlatego zadania tej grupy koncentrują się wokół tworzenia modeli działania umysłu w aspekcie ponadindywidualnym i prób implementowania osiągnięć całego ruchu do praktyki społecznej.

Cóż zatem pozostaje do zrobienia filozofom? Kilka bardzo ważnych zadań: krytyczna rewizja podstaw teoretycznych i spójna wizja całości wysiłków ruchu, nowatorskie metodologie oparte na ścisłej (i efektywnej) ontologii zjawisk psychicznych, próby integracji wyników, robocze teorie świadomości, wreszcie trafna, wieloaspektowa teoria umysłu zdolna pomieścić zarówno wyniki z empirycznych nauk szczegółowych, jak i abstrakcyjne modele formalne.

2. Myślące kalkulatory i inteligentne odkurzacze

Można przewidywać, iż skoro teorie mają być trzy i wyróżniłem trzy podgrupy naukowców wewnątrz nauk o poznaniu, to zapewne jest tak, że bionaukowcy mają swoją wersję, robonaukowcy swoją, a teonaukowcy własną. Nie jest to przewidywanie słuszne, gdyż w rzeczywistości sprawa jest daleko bardziej zagmatwana.

Pierwsza z trzech najważniejszych wizji umysłu ma kilka niezależnych źródeł. Wszystkie one jednak mają ważną cechę wspólną — nie traktują umysłu substancjalnie, przełamując klasyczny dualistyczny paradygmat pochodzący wprost od Kartezjusza. Bardzo zgrabne jej sformułowanie podaje Steven Pinker:

To jest książka o mózgu, ale nie będę dużo mówił o neuronach, hormonach i neuroprzekaźnikach. Powodem jest fakt, że umysł nie jest mózgiem, lecz tym, co mózg robi, i to nie wszystkim [...]. Mózg zawdzięcza wyjątkowy status swoim dokonaniom, dzięki którym możemy widzieć, myśleć, czuć, dokonywać wyborów i działać. Jego szczególną cechą jest przetwarzanie informacji, czyli komputacja. Informacja i komputacja zawarte są we wzorach danych i relacjach logiki, niezależnych od fizycznego medium, które się nimi posługuje. Odkrycie to, które jako pierwsi sformułowali matematyk Alan Turing, informatycy Alan Newell, Herbert Simon i Marvin Minsky oraz filozofowie Hilary Putnam i Jerry Fodor, nazywane jest teraz komputacyjną teorią umysłu⁴.

Takie postawienie sprawy ma kilka istotnych konsekwencji. Przede wszystkim umysł jest tu wykonywaniem pewnych czynności — komputacją (obliczaniem lub,

2000 roku opracowałyśmy raport dla Economic and Social Research Council (ESRC), który miał pokazać, że odkrycia w dziedzinie neurobiologii mogą być inspiracją programu badawczego w dziedzinie edukacji [...]. W dyskusjach z nauczycielami i badaczami zajmującymi się wychowaniem okazało się, ku naszemu zaskoczeniu, że prawie nie istnieje literatura na temat powiązań nauki o mózgu z edukacją. Naukowcy wiedzą zaś obecnie całkiem sporo o uczeniu się — o tym, jak przebiega rozwój komórek mózgu przed urodzeniem i po nim; jak niemowlęta uczą się widzieć, słyszeć, mówić i chodzić; jak dzieci nabywają moralności i uczą się funkcjonować w otoczeniu społecznym; i jak dorosły mózg jest w stanie nadal uczyć się i rozwijać. Byłyśmy zdziwione, że mimo coraz większej wiedzy na ten temat i jej wyraźnych odniesień do zasad pedagogiki istnieje tak niewiele powiązań między badaniami nad mózgiem a zasadami i praktyką edukacji". S.J. Blakemore, U. Frith, *Jak uczy się mózg*, tłum. R. Andruszko, Kraków 2008, s. 2.

⁴ S. Pinker, *Jak działa umysł*, tłum. M. Koraszewska, Warszawa 2002, s. 33–34.

lepiej — przetwarzaniem informacji). Co więcej, efektywnym opisem danej aktywności umysłowej jest właśnie realizowana funkcja matematyczna. Z tych dwóch powodów nie ma żadnego znaczenia, kto wykonuje daną czynność — zarówno człowiek, jak i kalkulator, obliczając, ile to jest $2 + 2$, robią dokładnie to samo. Jako że realizują identyczną funkcję matematyczną, możemy powiedzieć, że wykonują taką samą operację umysłową.

Co ten zestaw poglądów twierdzi o umyśle? Przede wszystkim to, że nie jest on elitarną cechą konkretnej konfiguracji specyficznych komórek u pewnego gatunku węglowców, lecz własnością układu fizycznego, zdolnego realizować rozmaite funkcje matematyczne. Staje się pewnym wyabstrahowanym konstruktem teoretycznym, którego nie sposób dostrzec pod mikroskopem. Mamy tu do czynienia z wyraźnym odejściem od myślenia substancjalnego — znika filozoficzny problem, czy umysł to dusza, ciepłe tchnienie czy emergentna własność drugiego rzędu, wyłaniająca się na skutek działania centralnego układu nerwowego. Dla komputacjonisty pytanie to nie ma żadnego znaczenia, gdyż odpowiedź na nie nic nam nowego nie powie. Znacznie ważniejsze jest zagadnienie, co umysł robi i w jaki sposób. Znamionuje ono podejście funkcjonalistyczne, zainteresowane konfiguracjami przepływu informacji (czy też przekształcania symboli), oderwanymi od swojego fizycznego nośnika⁵. W ten sposób na scenę wkracza matematyka i logika, dzięki którym można próbować konstruować abstrakcyjne modele przepływu informacji. Główną sprawą staje się nie dogłębne zbadanie fizyki czy fizjologii interesujących obiektów, lecz zidentyfikowanie algorytmów, odpowiedzialnych za ich niezwykle umysłowe właściwości. Chodzi o te wzorce przekształceń, które umożliwiają podejmowanie decyzji, rozwiązywanie problemów, symulowanie kontrfaktycznych sytuacji czy reprezentowanie otoczenia. Propagatorzy przekonują, że to wszystko potrafią nie tylko ludzie.

Komputacjonizm ma szerokie grono zwolenników z wielu powodów. Pośród nich z pewnością ważne miejsce zajmują jego naukowe walory. Ich lista nie jest trudna do ułożenia (choć nie wszyscy zgadzają się z poglądem, iż są to akurat zalety):

- jasno zdefiniowany przedmiot badań;
- wyraźnie sformułowane założenia i podstawowe twierdzenia;
- precyzyjnie sformułowana terminologia, wyrażona w jednoznacznym języku;
- wyrafinowana i kreatywna metodologia;
- wspólne zaplecze teoretyczne: fizyka, matematyka i logika;
- wysoki stopień abstrakcyjności i znaczna formalizacja teorii;
- prostota i elegancja;
- właściwie nieograniczone możliwości rozwoju.

Do tej listy zalet należałoby dopisać wielkimi literami STOSOWALNOŚĆ! Sądzę, że w znacznie wyższym stopniu o atrakcyjności tej właśnie wizji umysłu decyduje ogromny sukces dyscypliny, która zajmuje się wprowadzaniem jej założeń

⁵ Alan Turing w wywiadzie radiowym z 1952 r. stwierdził: „Nie interesuje nas fakt, że mózg ma konsystencję zimnej owsianki. Nie powiemy: »Ta maszyna jest całkiem twarda, czyli nie jest mózgiem, a więc nie może myśleć«”. H.A. Turing, *Krytyczna analiza myśli filozofa, który stworzył podwaliny teorii sztucznej inteligencji*, tłum. J. Nowotniak, [w:] *Wielcy filozofowie, Derrida, Turing, Wittgenstein*, Warszawa 1998, s. 58.

teoretycznych w „życie”, czyli robotyki i sztucznej inteligencji. Nieustanny skok jakościowy w konstruowaniu automatów o coraz większych możliwościach utwierdza w przekonaniu, że obrany kierunek badań jest słuszny.

Wielu ludzi nie czuje się pewnie, słysząc o tym, że buduje się automaty imitujące ludzi. Jest to wynik pewnego nieporozumienia. Nie chodzi wcale o to, by zbudować sztucznego człowieka, ale maszynę, która ucieleśniałaby umysł w takim samym stopniu co my, posiadała podobne zdolności. Biorąc poprawkę na komputacyjną wizję tego, czym jest umysł, nie ma w tym nic zdroźnego⁶. Zgodnie z funkcjonalistycznymi założeniami buduje się automaty zdolne robić to samo co ludzie, ale po swojemu. Uważa się, że i tak w wielu aspektach nas nie dogonią:

Są pewne rzeczy, których żadna maszyna prawdopodobnie nie dokona, gdyż wymagają one mocy świadomości. Turing przedstawił wybór działań rzekomo niemożliwych dla maszyny: być miłym, zasobnym, pięknym, przyjaznym [...] przejawiać inicjatywę, mieć poczucie humoru, odróżniać dobro od zła, popełniać błędy [...] zakochiwać się, delektować truskawkami z bitą śmietaną [...] sprawić, by ktoś zakochał się we mnie, uczyć poprzez doświadczenie [...] używać poprawnie słów, być podmiotem swoich własnych myśli [...] zrobić coś naprawdę nowego.

To dobra lista; 50 lat później maszyny nie potrafią zrobić większości z nich⁷.

A co takiego potrafią już maszyny? Kolekcja ambitnych i spektakularnych projektów obejmuje na razie takie pozycje, jak choćby Cog, skonstruowany w 1993 r. przez Rodneya Brooksa, profesora MIT. Mający ruchomą głowę, tors i ramiona, jest zdolny do samodzielnego uczenia się, także od ludzi, przez wchodzenie w interakcje (również społeczne) z obiektami (potrafi między innymi lokalizować twarz i oczy rozmówcy, dzielić z nim uwagę, zwracać się ku wskazywanym przezeń obiektom i imitować kiwnięcia głową)⁸.

Jego następcą jest Kismet, zaprojektowany przez współpracowniczkę Brooksa Cyntherę Breazeal. Choć składa się tylko z głowy i ruchomej szyi, zdolny jest do manifestowania nastrojów, które są kombinacją trzech zmiennych — postawy (otwartość na nowe bodźce), pobudzenia (poziom stymulacji) oraz nastawienia (ogólny poziom szczęścia). Manifestuje on swoje chęci przez poruszanie głową, mruganie powiekami, zmianę dystansu do obiektu oraz poszukiwanie lub ignorowanie go (ze szczególnym uwzględnieniem ludzkich twarzy). Potrafi więc zachęcać lub unikać interakcji, w zależności od samopoczucia⁹.

⁶ Najpełniejszy wykład doktryny komputacyjnej znajdzie czytelnik w: P.S. Churchland, T.J. Sejnowski, *The Computational Brain*, London 1993.

⁷ S. Blackmore, *Consciousness. An Introduction*, London 2003, s. 36, tłum. własne. Autorka cytuje we fragmencie pracy: A. Turing, *Computing machinery and intelligence*, „Mind” 51 (1950), s. 433–460; wyd. polskie *Maszyna licząca a inteligencja*, tłum. M. Szczubiałka, [w:] *Filozofia umysłu*, Warszawa 1995.

⁸ Obszerniejszy opis konstrukcji można znaleźć w: S. Blackmore, *Consciousness. An Introduction*, s. 211–224. Więcej informacji o projekcie Cog: R.A. Brooks et al., *The Cog project: Building a humanoid robot*, [w:] C. Nehaniv (ed.), *Computation for Metaphors, Analogy and Agents*, „Lecture notes in artificial intelligence” 1562, s. 52–87, New York 1998.

⁹ Kismet jest także opisywany przez S. Blackmore (zob. przyp. 8). Źródłowa bibliografia Kismet: C.L. Breazeal, *Designing Sociable Robots*, Cambridge 2001.

Do listy tego typu konstrukcji można dopisać wysoce samodzielne marsjańskie łaziki do zbierania próbek skał, zdolne do lokomocji układy wyposażone w wysięgniki, manipulujące przedmiotami i adaptacyjnie uczące się kategoryzować napotkane obiekty oraz wiele, wiele innych. Wszystkie te projekty, jakkolwiek by one były zaawansowane technologicznie, na własny rachunek muszą skonfrontować się z dwoma fundamentalnymi pytaniami. Pierwsze z nich nasuwa się samo: „czy maszyna może myśleć?”¹⁰. Właściwie nie do końca wiadomo, o co pytamy, lecz jeśli odpowiedź ma być usytuowana w paradygmacie komputacyjnym, to zapewne brzmi ona: owszem, mogą, jeśli tylko realizują odpowiednie funkcje formalne.

Sceptyków taka odpowiedź z pewnością nie zadowoli. Często argumentuje się przeciw komputacjonistom, używając figur pokroju tych, które widnieją w tytule tego podrozdziału: „Czy nie jest nonsensem twierdzić, że kalkulator myśli, rozwiązując zadania arytmetyczne, a mój odkurzacz ma inteligencję, bo potrafi odróżniać płyny od pyłów?”. Gorliwi zwolennicy komputacjonizmu potrafią odpowiedzieć, iż z pewnością nie jest to nonsensem, przecież na przykład termostat pokojowy też ma przekonania. Dokładnie trzy: jest za zimno, temperatura jest odpowiednia oraz jest zbyt ciepło. Na ich podstawie działa, regulując temperaturę w pokoju.

Komputacyjna wizja umysłu znacznie odbiega od wyobrażeń potocznych, związanych z odczuwaniem emocji, prowadzeniem wewnętrznego monologu, odczuwaniem stanów własnego ciała, stresem przed egzaminem, napadami chęci na czekoladę, zapachem herbaty, uśmiechaniem się do znajomych czy płaczem na filmach. Bardzo łatwo dojść do przekonania, że coś tu jest nie w porządku. W końcu szukamy teorii ludzkiego umysłu, a dostajemy twierdzenie, że człowiek i kalkulator robią to samo. Z tego typu wątpliwości bierze się drugie, nie wiem, czy nie ważniejsze pytanie, przed którym stają konstruktorzy robotów i programiści. Dobitnie sytuację tę ujął noblista Francis Crick:

Zainteresowanie zagadnieniem widzenia może mieć wiele różnych powodów. Są tacy, którzy chcą zbudować maszynę do widzenia, która potrafiłaby widzieć równie dobrze, a nawet lepiej od nas i mogłaby być wykorzystywana do potrzeb domowych, przemysłowych czy wojskowych. Mało interesuje ich sposób, w jaki nasze mózgi spełniają to zadanie, co najwyżej jako inspiracja. Maszyna do widzenia nie musi być wierną imitacją mózgu, wszak samolot też nie musi machać skrzydłami [...]. Wysuwano argument, że skoro do zrozumienia lotu zarówno ptaków, jak i samolotów, konieczna jest znajomość praw aerodynamiki, to aby zrozumieć proces widzenia, powinniśmy poszukiwać rządzących nim ogólnych zasad¹¹.

Gdzieś między wierszami czai się niewypowiedziane wprost pytanie: co te wszystkie badania nad robotami mają wspólnego z ludźmi? Czy badanie umysłu

¹⁰ Zapewne wielu czytelnikom przywołuje ono takie filmy jak *Terminator 1* czy *Matrix* lub literaturę science fiction z lat 70. XX w. Zagadnienie nie jest jednak tak błahe, jak można sądzić. Zwolenników tezy, że problem ten jest zdecydowanie bardziej *fiction* niż *science*, odsyłam do twórczości Stanisława Lema. Zwłaszcza lektura *Golema XIV* pokazuje, że problem jest poważny i z pewnością naukowy.

¹¹ F. Crick, *Zdumiewająca hipoteza, czyli nauka w poszukiwaniu duszy*, tłum. B. Chacińska-Abramowicz, M. Abramowicz, Warszawa 1997, s. 107–108.

nie powinno zmierzać do zbudowania samolotu, który potrafi machać skrzydłami? W końcu pytamy o umysł ludzki, a nie o formalny abstrakt, który pasuje do kalkulatora, odkurzacza, komputera, kury, małpy i człowieka. Mimo uczynienia wielu kroków zmierzających do modelowania ludzkiego umysłu w maszynach pozostaje wątpliwość, czy były to kroki we właściwym kierunku.

3. Sieci w szkole

Nie wszystkim zwolennikom podejścia funkcjonalistycznego podobały się daleko idące deklaracje komputacjonistów. Dla wielu z nich stwierdzenie, iż układy żywe i sztuczne, choć rzeczywiście wykonują te same operacje, dysponują takim samym umysłem, wydało się mało wiarygodne. Zrezygnowano więc ze zbyt mocnego, jak się wydaje, założenia o obliczalności procesów psychicznych i możliwości ich wykonywania przez odpowiednie algorytmy. Zwrócono się w nieco inną stronę, starając się precyzyjniej modelować mózg i jego pracę, by móc lepiej poznać zjawiska, jakie zachodzą w nim na poziomie wymiany informacji. W końcu pojawiła się technologia umożliwiająca wprowadzenie tych zamiarów w życie, opierająca się częściowo na teorii systemów i cybernetycznych modelach dystrybucji informacji Norberta Wienera¹²:

Lata 80. były świadkiem rozkwitu „koneksjonizmu”, nowego podejścia, opartego na sztucznych sieciach neuronowych (SSN¹³) i rozproszonego przetwarzania jednoczesnego. Częścią motywacji stojącej za sieciami neuronowymi była chęć bliższego zamodelowania ludzkiego mózgu, choć nawet teraz SSN są bardzo proste w porównaniu z komórkami ludzkiego mózgu. Wielką różnicą w porównaniu z tradycyjną SI jest fakt, iż SSN nie są programowane, lecz trenowane¹⁴.

Podczas swoich prac nad konstruowaniem coraz lepszych sieci i ich trenowaniem koneksjoniści szybko zauważyli pewne zastanawiające zjawiska. Przede wszystkim okazało się, że fizyczna realizacja danej funkcji ma ogromne znaczenie. Dzięki pewnym własnościom strukturalnym płatanina sztucznych krzemowych połączeń chwilami zachowuje się podobnie jak skłębione komórki nerwowe. Najważniejszym ich pokrewieństwem jest zdolność do adaptacyjnego uczenia się i brak sztywnych reguł działania.

Ważnym argumentem przeciwko komputacyjnym wytworom jest sztywność ich działania. Roboty i komputery działają na podstawie programu, który daje im ograniczone pole manewru i zmusza do pracy krokowej — wykonywania tylko jednego procesu w danym momencie. Z własnych obserwacji wiemy, że ludzie mają możliwość robienia wielu rzeczy jednocześnie, dlatego też „inność” projektów komputacyjnych jest łatwo odczuwalna dla przeciętnego posiadacza umysłu. Sieci pozbawione są tych ograniczeń. Dzięki wielu połączeniom między swoimi elementami są w stanie działać, wykorzystując rozproszone przetwarzanie jednoczesne — po-

¹² Najważniejszą pozycją dla całego kierunku jest bez wątpienia: N. Wiener, *Cybernetyka, czyli sterowanie i komunikacja w zwierzęciu i maszynie*, Warszawa 1971.

¹³ Polska kalka angielskiego skrótu ANN (*artificial neuronal networks*).

¹⁴ S. Blackmore, *Consciousness. An Introduction*, s. 187.

szczególne części zajmują się osobnymi procesami i w efekcie całość realizuje kilka procesów naraz. Zdolność do wielotorowego przepływu informacji leży u podstaw adaptowania się sieci do bodźców, które napotyka¹⁵. Ta dość tajemnicza własność eliminuje konieczność pisania sztywnego zestawu instrukcji, jak informacja ma być analizowana na poszczególnych poziomach.

I znów należy zapytać się, co koneksjonizm ma do powiedzenia na temat umysłu. Przede wszystkim pokazuje, że własności mentalne u ludzi mają niewątpliwie związek z płataniną neuronów w czaszce, a co więcej, sposób, w jaki jest wykonana, ma prawdopodobnie fundamentalne znaczenie dla sposobu działania umysłu. Eksponuje także wielozadaniowość i wielofunkcyjność umysłu, kładąc silny nacisk na wykonywanie wielu czynności jednocześnie. Oznacza to, że pod względem zdolności mentalnych możemy porównywać tylko obiekty o podobnej (skomplikowanej) budowie, opartej na tych samych zasadach działania. Kura i kalkulator z tego punktu widzenia okazują się o wiele mniej interesującymi obiektami, niż można sądzić, natomiast małpy naczelne, delfiny i ośmiornice stają się niezmiernie ciekawe. Ponadto koneksjonizm podaje też atrakcyjną propozycję wyjaśnienia mechanizmów nabywania wiedzy o świecie opartej na wieloaspektowej kategoryzacji obiektów¹⁶. Ważnym elementem tej wizji umysłu jest także uwypuklenie indywidualnych różnic pomiędzy jego poszczególnymi posiadaczami, uzyskiwanych na drodze kontaktów z odmiennym zespołem danych z zewnątrz i innych sposobów samodzielnego ich przetwarzania (czyli przez życie)¹⁷.

Do niewątpliwych zalet tego podejścia należy bardzo udane modelowanie niektórych właściwości ludzkiego mózgu i umysłu. Samorzutne uczenie się sieci podczas treningu jest godnym uwagi analogonem pamięci. W układzie nerwowym szlak pamięciowy jest realizowany przez tak zwane wzmocnienie synaptyczne, które powoduje zwiększenie prawdopodobieństwa odpalenia połączonych z sobą komórek. Optymalizowanie trafności odpowiedzi sieci podczas jej kalibrowania oparte jest na ustaleniu wag (znaczenia/siły) poszczególnych węzłów. W ten sposób informacja jest dystrybuowana poprzez połączone wzajemnie elementy z większym

¹⁵ Trening sieci polega zazwyczaj na dostarczaniu informacji określonego rodzaju na wejścia i umieszczanie poprawnych wyników na wyjściach. Sieć po kilku takich sesjach „sama” konfiguruje swoje połączenia tak, by uzyskiwać poprawne wyniki, wykorzystując dostępny zasób danych. Dzięki temu, gdy natrafi na nowy obiekt znanego już rodzaju, potrafi samoistnie wygenerować odpowiednią informację wyjściową.

¹⁶ Wyobraźmy sobie, że mamy taką oto listę obiektów: jabłko, krowa, młotek, krzesło, ser żółty i jajko. Otrzymujemy zadanie pogrupowania ich pod kątem szeroko rozumianego podobieństwa. Jedno z rozwiązań to dwa zbiory: a) {jabłko, krowa, ser, jajko}; {młotek, krzesło}, inne, równie dobre: b) {krowa, ser}; {jajko}; {jabłko, młotek, krzesło} lub c) {krowa, krzesło}; {młotek, jabłko, jajko}. Klucz do każdego z podanych rozwiązań jest inny (odpowiednio: a) naturalne/sztuczne; b) wspólny przodek; c) można na tym usiąść/zdecydowanie nie siadamy), ale wszystkie są równie poprawne. Ze względu na brak sztywnych reguł uczenia się kategoryzujemy przedmioty na różne sposoby na podstawie arbitralnych preferencji.

¹⁷ Zwięzły wykład podejścia koneksjonistycznego można znaleźć w: A. Clark, R. Lutz (eds), *Connectionism in Context*, New York 1992. Za najważniejszą inspirację dla całego kierunku uważa się powszechnie psychologiczną publikację: D.O. Hebb, *Organization of Human Behavior*, New York 1964.

prawdopodobieństwem niż przez inne fragmenty sieci¹⁸. Analogia ta, choć niekoniecznie zupełna, sugeruje, że sieci neuronowe wraz ze swoim stopniowym rozwojem mogą stać się wiarygodnym modelem centralnego układu nerwowego i pozwolić na dokładniejsze zbadanie detali zależności mózg/umysł. Dodatkowo trafność koneksjonizmu zwiększa osłabienie siły funkcjonalizmu, na którym częściowo się on opiera. Autor słynnego podziału na silną i słabą SI pisze:

Obszarem, na którym spotkały się ze sobą nauka o sztucznej inteligencji i filozoficzna teoria funkcjonalizmu, była idea, że umysł jest tylko programem komputerowym. Pogląd ów ochrzciłem mianem „mocnej wersji sztucznej inteligencji” [...]. Zdroworozsądkowy zarzut wobec mocnej wersji SI głosi po prostu, że komputerowy model umysłu nie uwzględnia podstawowych aspektów umysłu, takich jak świadomość i intencjonalność¹⁹.

Głównym zatem problemem silnej SI jest intuicyjnie niewiarygodny pogląd, iż maszyna i człowiek dysponują takim samym umysłem. Zwolennicy podejścia sieciowego zadowalają się skromniejszym roszczeniem: aktywność ich sieci potrafi momentami efektywnie imitować działanie ludzkiego umysłu. W ten sposób słabsza SI staje się w obliczu kolejnych sukcesów coraz trudniejsza do zignorowania²⁰.

A o jakich sukcesach mowa? W roku 1987 Terry Sejnowski i Charles Rosenberg zbudowali NETtalk, sieć zdolną do uczenia się czytać (w założeniach miała ona odtwarzać angielską wymowę na podstawie transkrypcji fonetycznej dwóch tekstów). Aby ocenić wyniki, podłączono ją do syntezy mowy²¹. Eksperyment powiódł się w znacznym stopniu, mimo bardzo ograniczonego zestawu treningowego.

¹⁸ Ważnym problemem jest arbitralne ustalenie początkowych wartości siły połączeń. Poczytny propagator podejścia sieciowego Paul Churchland pisze: „Ten problem trzeba było rozwiązać. Na szczęście istnieje ogólna metoda znajdowania rozwiązań dla problemów transformujących, metoda powszechnie dziś stosowana przez naukowców tworzących sztuczne sieci nerwowe. Jest ona biologicznie realistyczna pod jednym względem — polega mianowicie na stopniowym dostrajaniu połączeń synaptycznych w odpowiedzi na bodźce zewnętrzne. Niestety, pod każdym innym względem jest biologicznie nierealistyczna. Jednak z braku czegoś lepszego dobre i to. Dopóki nie odkryjemy, na czym polegają procesy uczenia się zachodzące w mózgu, musimy zadowolić się ich namiastką. Ta tymczasowa metoda nosi nazwę »dostrajanie synaps przez sukcesywne zmniejszanie błędu« [*successive backpropagation of errors*]”. P. Churchland, *Mechanizm rozumu, siedlisko duszy. Filozoficzna podróż w głąb mózgu*, tłum. Z. Karaś, Warszawa 2002, s. 55. W polskiej literaturze technicznej opisywana w cytacie metoda pojawia się także pod nazwą wstecznej propagacji błędu.

¹⁹ J. Searle, *Umysł na nowo odkryty*, tłum. T. Baszniak, Warszawa 1999, s. 70–71.

²⁰ Najmocniejsze argumenty przeciwko SI we wszystkich postaciach wytoczył John Searle w swoim artykule *Mind, brains and programs*, „Behavioral and Brain Sciences” 3 (1980), s. 417–457.

²¹ Francis Crick tak opisuje swoje wrażenia: „Fascynujące jest słuchanie, jak urządzenie uczy się »czytać po angielsku«. Ponieważ początkowe połączenia są przypadkowe, najpierw słychać jedynie bezsensowny ciąg dźwięków. NETtalk szybko uczy się odróżniać samogłoski od spółgłosek, ale początkowo zna tylko jedną samogłoskę i jedną spółgłoskę, co daje wrażenie paplaniny. Następnie zaczyna rozpoznawać granice między słowami, dzięki czemu pojawiają się ciągi pseudosłów. Po mniej więcej dziesięciokrotnym przejściu przez zestaw uczący słowa stają się zrozumiałe i brzmią bardzo podobnie jak w mowie małego dziecka”. F. Crick, *Zdumiewająca hipoteza, czyli nauka w poszukiwaniu duszy*, s. 254. Oryginalna bibliografia projektu: T.J. Sejnowski, C.R. Rosenberg, *Parallel networks that learn to pronounce English text*, „Complex Systems” 1 (1987), s. 145–168.

Inną ciekawą konstrukcją jest sieć do rozpoznawania twarzy, opracowana przez Garrisona Cottrella z Uniwersytetu Kalifornijskiego. Wykorzystując sześćdziesiąt cztery zdjęcia jedenastu twarzy oraz trzynaście zdjęć innych obiektów, chciał on nauczyć sieć odróżniać twarze od innych obiektów, separować twarze męskie i żeńskie oraz dopasowywać twarze do imion ich właścicieli. Sieć poddano kilku niezależnym testom — odróżniania twarzy i nietwarzy oraz dopasowywania imion (100% skuteczności); rozpoznawania znanych twarzy na podstawie nowych, nieoglądanych wcześniej zdjęć (98%); odróżniania twarzy od nietwarzy i dopasowywania płci (100% w pierwszej części, 81% w drugiej); rozpoznawania znanych twarzy na znanych zdjęciach przysłoniętych poziomym paskiem na jednej piątej powierzchni (najgorszy wynik to 71% trafień po zasłonięciu czoła)²².

Mimo tych znaczących osiągnięć część krytycznych argumentów wobec mocnej wersji sztucznej inteligencji można dalej z powodzeniem zastosować do jej słabszej wersji — podejścia sieciowego. Najważniejszy z nich to wątpliwość dotycząca zasadności ekstrapolowania wyników badań nad modelowaniem umysłu w sieciach do ludzkich zdolności mentalnych. Mimo programowej deklaracji konstruktorów sieci i upodabniania swoich konstrukcji do naszego układu nerwowego moment, w którym będzie można zasadnie uznać, że rzeczywiście się to udało, jest jeszcze daleko. Póki co, analogie zachodzą głównie na poziomie formalnym i funkcjonalnym. Fizyczne prawa wiążące współpracę komórek sieci odzwierciedlają fizjologię neuronów i fizykochemiczne zasady ich wzajemnych oddziaływań w znikomym stopniu. Dopóki nie powstaną sieci wiernie symulujące współpracę komórek nerwowych, z uwzględnieniem ich fizjologii i funkcjonalnej specyfiki, ich konstruktorzy będą musieli obawiać się argumentacji tego typu²³. Inną, nieco mniej ważną, wadą jest, iż nie ma jak dotąd hipotezy, która potrafiłaby wyjaśnić wyczerpująco zasadę samorzutnego uczenia się sieci i grupowania się komórek w tak zwane warstwy ukryte, które specjalizują się w analizowaniu konkretnych aspektów bodźców. Innymi słowy konstruktorzy wiedzą, jak sprowokować pewne zjawiska, ale pozostają one nieco poza ich kontrolą i dają zaskakujące wyniki. Można się spodziewać, iż wyjaśnienie tego fenomenu pozwoli zbliżyć kolejne projekty do ich żywych pierwowzorów. Ponadto, mimo pewnej bliskości zasad działania sieci i żywych zespołów komórek nerwowych, założenia, na których projektuje się sieci, wydają się nadal zbyt redukcyjne. Nie uwzględniając pewnych aspektów działania układu nerwowego (oraz jego

²² Głębszy opis czytelnik może znaleźć w: P.M. Churchland, *Mechanizm rozumu, siedlisko duszy. Filozoficzna podróż w głąb mózgu*, s. 53–58. Oryginalna bibliografia do badań: G. Cottrell, *Extracting features from faces using compression networks: Face, identity, emotion and gender recognition using holons*, [w:] D. Touretzky et al. (eds), *Connectionist Models: Proceedings of the 1990 Summer School*, San Mateo 1991.

²³ Inny noblista, Gerald Edelman, formułuje zarzuty wobec funkcjonalizmu stojącego za komputacjonizmem i koneksjonizmem w bardziej dosadnej formie: „W jaki sposób powstał ten »odcieśniony« i formalny, funkcjonalistyczny obraz umysłu? Jak ktokolwiek mógł przyjąć tak abstrakcyjną formę rozumowania i aktywności umysłowej?”. G.M. Edelman, *Przenikliwe powietrze, jasny ogień. O materii umysłu*, tłum. J. Rączaszek, Warszawa 1998, s. 313.

związku z układem wydzielania wewnętrznego), mogą nie być zdolne do modelowania takich interesujących aspektów ludzkiego umysłu, jak: odczuwanie własnych stanów mentalnych, intencjonalność, własna perspektywa poznawcza, planowanie strategiczne czy podejmowanie nieracjonalnych decyzji na podstawie skonfliktowanych wartości. Udana rekonstrukcja tych własności w SNN byłaby z pewnością wielkim zaskoczeniem i jednocześnie bolesnym ciosem dla naturalistów oraz innych przeciwników słabej SI.

4. Pokaż, czego nie potrafisz, a powiem ci, co nie działa

Ostatnie z podejść, które zamierzam przedstawić, różni się od omówionych tym, że pewne jego implikacje są akceptowane zarówno przez zwolenników konstruowania robotów oraz sieci, jak i przez biologów oraz filozofów o inklinacjach naturalistycznych. Dla przejrzystości postaram się pokazać sposoby ich wykorzystywania przez różne formacje naukowców we względnej separacji i dopiero potem je z sobą skonfrontować.

Bardzo lapidarną formę wizji modularnej, bo o niej mowa, stosowaną przez robonaukowców podaje Steven Pinker:

Modułuj.

Używaj podprogramów.

Każdy moduł powinien dobrze wykonywać jedną rzecz.

Każdy moduł powinien być użyteczny.

Zlokalizuj wejścia i wyjścia w podprogramach²⁴.

Jest do dość przejrzysty zbiór zasad, który pozwala zoptymalizować procesy rozwiązywania problemów przez jakikolwiek złożony układ zdolny do wykonywania obliczeń. Zachęca on do dzielenia wszelkich operacji na mniejsze części i konstruowania autonomicznych modułów specjalizujących się w przetwarzaniu konkretnych aspektów poszukiwanych rozwiązań. W ten sposób złożone problemy są rozbijane na podproblemy i przetwarzane w różnych modułach, co pozwala na zwiększenie wydajności przez wykonywanie kilku operacji jednocześnie. Tym samym nakłada się na projektantów układów mechanicznych wymóg zaprojektowania rozwiązań sprzętowych odpowiedzialnych za poszczególne fragmenty obróbki danych. W robotyce osiąga się to na przykład przez oddzielenie układu odpowiedzialnego za lokomocję od modułu percepcyjnego i decyzyjnego, który integruje pracę obydwu poprzednich. W sieciach popularnym rozwiązaniem jest łączenie poszczególnych komórek w warstwy, które dostrajają się tak, by zajmować się kodowaniem jednego z aspektów interesujących poznawczo obiektów. Ich wysoki stopień komplikacji związany jest między innymi z koniecznością zaprojektowania osobnych modułów dla każdego rodzaju zmiennej, jaką może rozpoznać dana konstrukcja²⁵. Ciekawym

²⁴ S. Pinker, *Jak działa umysł*, s. 103.

²⁵ W wypadku sieci rozpoznającej twarze projektanci wprowadzili aż osiemdziesiąt różnych własności, na które reagowała sieć. Zastanawiające, że ludzie dobrze radzą sobie z przetwarzaniem około dziesięciu zmiennych jednocześnie, a uzyskują znacząco lepsze wyniki niż obliczeniowo przytłaczające je sieci.

wymogiem (który nie pojawia się w wersji biologicznej) jest zastrzeżenie budowania tylko takich modułów, które będą brały udział w rzeczywistych procesach obliczeniowych i nie okażą się zbędne.

Bionaukowcy mają własną efektywną wersję modularyzmu, która odniosła w XX w. bardzo duży sukces. Doprowadziła między innymi do powstania neuropsychologii, wąsko wyspecjalizowanej subdyscypliny badającej rozmiary zaburzeń zachowania:

Przedmiot dyscypliny określa się jako badanie mózgowych mechanizmów zachowania (*brain-behaviour relations*) na podstawie analizy zaburzeń zachowania występujących w przypadkach funkcjonalnych i (lub) strukturalnych zmian w mózgu człowieka [...]. U podstaw tych badań znajduje się „metoda psychoanatomiczna”, oznaczająca, według Walsha, ujmowanie związku: mózg–zachowanie w formie modelu rozbudowanego układu funkcjonalnego w znaczeniu nadanym temu terminowi przez Łuriego²⁶.

Widać dość jasno, że związek między zachowaniem podmiotu a aktywnością danego ośrodka mózgowego jest tu ujmowany bardzo ściśle. Podejście to stanowi pewien analogon specjalizacji funkcjonalnej konstrukcji mechanicznych, z tą istotną różnicą, że postulat o osobnym realizowaniu poszczególnych procesów przez układ nerwowy stanowi samą podstawę tej nauki. Dlatego też przyrost wiedzy następuje tu w kierunku odwrotnym — z obserwowanego zachowania wnioskuje się redukcynnie o zaburzeniu czynności danego układu funkcjonalnego, przy założeniu, że hipoteza lokalizacyjna (lub mocniej — cała anatomia funkcjonalna) jest teorią poprawnie zbudowaną. Dodatkowe detale, opisując specyfikę badania neuropsychologicznego, wyjaśnia założyciel całego kierunku, Aleksander Łuria:

Badanie to ma dwa cele. Po pierwsze, mamy nadzieję, że znalezienie uszkodzeń mózgu odpowiedzialnych za specyficzne zaburzenia zachowania pozwoli nam opracować metodę wczesnej diagnozy i precyzyjnej lokalizacji uszkodzeń mózgu [...]. Po drugie, badanie neuropsychologiczne powinno dać nam analizę czynnikową, prowadzącą do lepszego zrozumienia elementów składowych złożonych funkcji psychicznych, za które odpowiedzialne jest działanie różnych części mózgu²⁷.

W przywołanym fragmencie daje się zauważyć dodatkowy komponent doktryny modularnej. Jest nim topograficzne rozproszenie układów funkcjonalnych²⁸, które otwiera możliwość kompensacji upośledzonej zdolności dzięki neuroplastyczności. Częściowa zmiana specjalizacji i samoorganizacja jest zjawiskiem, którego próżno szukać u komputacjonistów, u koneksjonistów zaś występuje sporadycznie.

Jednym z najważniejszych atutów podejścia modularnego jest jego zgodność z wiedzą biologiczną o funkcjonowaniu układu nerwowego, podejściem inżynierów i konstruktorów do projektowania, wykonywania i programowania swoich wytworów oraz potocznymi intuicjami na temat umysłu, biorącymi się z jego codziennego używania. Modularyzm umożliwia przekonujące wyjaśnienie, dlaczego określone uszko-

²⁶ K. Walsh, *Neuropsychologia kliniczna*, tłum. B. Mroziak, Warszawa 1997, s. 5.

²⁷ A. Luria, *The functional organization of the brain*, „Scientific American” 222 (1977), s. 66.

²⁸ Co oznacza w praktyce, że efektywnym kryterium zaliczania konkretnych grup komórek do danego układu nie jest ich styczność przestrzenna z pozostałymi jego elementami, lecz ich wspólna aktywność podczas wykonywania specyficznego zadania. Powoduje to znaczną komplikację badań i wyjaśnienia, dlaczego anatomia funkcjonalna mózgu nie jest jeszcze wiedzą kompletną.

dzenia powodują statystycznie powtarzalne upośledzenia zachowania²⁹, i stosowanie skutecznych terapii objawowych oraz zajęć kompensacyjnych, pod warunkiem odwracalności uszkodzenia. W prosty sposób tłumaczy także zdolność umysłu do zajmowania się wieloma zadaniami jednocześnie. Uwzględniając historyczność każdego mózgu, można sformułować teorię dotyczącą różnic indywidualnych między ludźmi, zezwalając jednocześnie na budowanie ogólnego modelu aktywności mózgowia, nad którym dałoby się sensownie nadbudować badania nad ogólną teorią umysłu ludzkiego³⁰.

W obliczu tak wielu fundamentalnych zastosowań myślenia modularnego, jak choćby diagnostyka i terapia medyczna, neurochirurgia, aż po załogowe loty kosmiczne, można się zastanawiać, dlaczego wizja modularna nie jest zwyciężczynią wyścigu po laur najlepszej teorii umysłu. Powodów takiego stanu rzeczy jest zapewne kilka. Jak sądzę, najważniejszym z nich jest stosunkowo ograniczona liczba rozstrzygnięć, które niesie z sobą jej przyjęcie. W rzeczywistości stanowi ona po prostu zbiór twierdzeń dotyczących zasad ogólnej organizacji funkcjonalnej dowolnego typu obiektów, wystarczająco skomplikowanych, aby były zdolne do plastycznego zachowania na podstawie przetwarzania informacji z otoczenia. Modularyzm w ogóle nie rozstrzyga dziedziny przedmiotowej, do której należy go adaptować. Równie dobrze spisuje się na polu neurologii jak robotyki czy projektowania mikroprocesorów i alarmów samochodowych.

Dlatego też wymóg specjalizowania poszczególnych części układów wykonawczych bardzo szybko stał się użytecznym dodatkiem zarówno w teoriach komputacjonistów, jak i koneksjonistów. Widzieliśmy także, jak dobrze radzą sobie dzięki takiemu podejściu neurologi i neuropsycholodzy, którzy mogli dodatkowo skupić się na projektowaniu jeszcze lepszych metod neuroobrazowania. Nie podając żadnych wiążących rozstrzygnięć ontologicznych, staje się ogólnym opisem, zaskakująco trafnym, sposobu, w jaki rzeczywiste obiekty, zarówno te ożywione, jak i czysto mechaniczne, radzą sobie z kanalizowaniem informacji.

5. Modularny koneksjonizm komputacyjny

Prawdopodobnie jest to teoria, którą da się zbudować, i co ważniejsze — z powodzeniem wprowadzić w życie. Z pewnością zabieg ten nie będzie polegał na prostym wrzuceniu założeń i rozstrzygnięć poszczególnych teorii składowych do jednego worka. Wzajemne dopasowanie będzie musiało pociągnąć za sobą redukcję pewnych roszczeń, osłabienie lub wzmocnienie innych oraz dodanie zupełnie nowych. Nie brak głosów, które twierdzą, że poszczególne pary z tej trójki są swoistymi odmianami tych samych twierdzeń. Broniąc silnej SI, Steven Pinker pisze:

²⁹ Technika pozyskiwania wiedzy na podstawie konfrontowania badań zaburzonych pacjentów ze „zdrową” grupą kontrolną nosi nazwę metody lezyjnej i jest bezpośrednią implementacją hipotezy lokalizacyjnej.

³⁰ Najpełniejszy wykład modularnej wersji teorii umysłu znajdzie czytelnik w: J. Fodor, *The Modularity of Mind: An Essay on Faculty Psychology*, London 1983.

Koneksjonizm nie jest alternatywą komputacyjnej teorii umysłu, ale jej odmianą, która twierdzi, że główny rodzaj przetwarzania informacji odbywający się w umyśle jest funkcją wielu zmiennych. Koneksjonizm nie jest niezbędną poprawką teorii, że umysł jest jak komputer, wyposażony w bezbłędną i seryjną jednostkę centralnego przetwarzania o dużej szybkości, nie ma bowiem zwolenników takiej teorii³¹.

Także projektowanie sieci wymaga od samego początku myślenia modularnego, gdyż realizowanie określonych zadań zmusza do dokonania wyboru, na jakie bodźce będzie ona wrażliwa, a które będzie zupełnie pomijać.

Co będzie zawierać taka teoria? Trudno na to pytanie jednoznacznie odpowiedzieć, ale jeśli za wskazówki przyjąć najmocniejsze strony wszystkich trzech dotychczasowych rozwiązań, można pokusić się o pewne przewidywania. Z pewnością znajdują się tam pewne ważne elementy myślenia funkcjonalistycznego — podstawą już dziś jest przekonanie, że umysł może być realizowany z powodzeniem w obiektach skończonego stanu, tak jak przewidywał to Turing. Jednak fizyczna implementacja tego poglądu będzie raczej zawierała sieciową tendencję do kopiowania naturalnych układów neuronalnych. Być może uda się nawet zamodelować ich fizjologię za pomocą wysublimowanych algorytmów, zdolnych do chwilowych zmian sposobów przeszukiwania poszczególnych obszarów przestrzeni obliczeniowej (co mogłoby odzwierciedlać fallibilność decyzyjną lub osłabienie pamięci pod wpływem nagłego silnego stresu). Przyjęcie perspektywy, iż sztuczne twory mogą chwilami „naśladować” działanie układu nerwowego, z pewnością zachęciłoby do współpracy wielu naturalistów i biologów, co spowodowałoby być może jakiś przełom technologiczny, umożliwiając chociażby rozwinięcie zaawansowanych badań nad efektywnymi implantami.

Najważniejszą trudnością do pokonania przy budowaniu zintegrowanej teorii umysłu jest zaspokojenie roszczeń zarówno robonaukowców, którzy są w stanie budować bardzo złożone modele funkcjonowania umysłu, jak i bionaukowców, mających coraz lepsze rozeznanie w roli anatomii i fizjologii na pracę umysłu. Trzeba więc tylko, i aż, opracować sposób, w jaki funkcjonalne metody *top-down* miałyby się spotkać z naturalistycznymi modelami *bottom-up*, najlepiej gdzieś w połowie drogi, wiążąc z sobą dwie fundamentalne dziedziny przedmiotowe, w których bada się umysł: procesy fizyczne obiektów nieożywionych i fizjologiczne obiektów żywych. Wtedy to teoria ludzkiego umysłu będzie kompletna.

Three Theories of Mind

Summary

The article tries to give the reader more detailed insight into three most successful and influential theoretical approaches to describe mind completely. All of them were founded during frantic studies of cognitive sciences movement, starting in the late 1960s century and still improving these days. The main thesis is that there is a stall

³¹ S. Pinker, *Jak działa umysł*, s. 127.

in the race for the title of the most precise theory of mind, and none of the presented views possess the ability to overcome the competitors once and for all.

However, each of these theories, despite being the pretender for the title, not come without its limitations. That brings about the construction of the whole argumentation: historical and scientific origins, main assets and latest impressive breakthroughs, followed by its serious flaws, pointed out by the opponents, provided for each theory separately.

There was even little room to put some information about the developing of the whole cognitive sciences movement, resulting in its diversity today that partially causes the inability to provide the single view of mind. The biological, robotics and theoretical divisions include different 'how to find mind' approach and so are unable to work upon the very same project.

As one can easily figure out, the paper covers the bond between the *computational, connectional and modular* theories of mind and its influence over the development of cognitive sciences. The computational part includes some basic ideas in AI and autonomic decision making robot building, with the most impressive achievements in this field. The connectional part provides the reader with some theory behind ANN and the finest examples of networks modeling advanced human abilities. The modular part gives insight into the functional specialization of its parts for 'multi-task capable constructs' and the great success of 'separate brain part-separate behavioral activity' diagnostic method, designed for patients with serious disabilities after head injury, utilized in neuropsychology.

In conclusion there is a call for putting more effort in designing the great unification theory, one that would contain the main assets of each of presented theories, but without its limitations. This complex view should be far more effective and convincing, eventually giving the final multi-aspect theory of human mind.