

M a r c i n J a ż y ń s k i

Hume, metafora mózgowa i myślenie życzeniowe

Słowa kluczowe: behawioryzm, Hume, język, metafora komputerowa, naturalizm, koneksjonizm, prawa umysłowe, ucieleśnione poznanie, umysł, poznanie, sztuczna inteligencja

1. Teoria praw procesów umysłowych Hume’a

W *Badaniach dotyczących rozumu ludzkiego* Hume pisze: „Możemy przeto podzielić wszystkie percepcje umysłu na dwie klasy czyli rodzaje, które różnią się od siebie stopniem swej siły i żywotności. Słabsze i mniej żywe percepcje nazywają się zwykle myślami lub ideami”. Ponadto idee „[...] naśladują, lub odtwarzają percepcje zmysłów”. I dalej: „A jednak, chociaż myśl nasza zdaje się posiadać tę nieograniczoną swobodę, przekonamy się po bliższym zbadaniu, że zamknięta jest w obrębie bardzo ciasnych granic, i że cała ta twórcza władza umysłu obejmuje jedynie zdolność łączenia, przekształcania, powiększania lub umniejszania materiału dostarczonego nam przez zmysły i doświadczenie [...], jednym słowem, cały materiał myślenia pochodzi albo z wrażeń zewnętrznych, albo wewnętrznych, a tylko połączenie ich i układ zależą od umysłu i woli”¹.

Wedle Hume’a czynności umysłowe polegają na manipulowaniu materiałem empirycznym zgromadzonym w impresjach i ideach. Operacje te przebiegają według pewnych reguł. Pytanie, jakie to reguły, dotyczy Hume’owskich praw procesów umysłowych. Ogólną odpowiedź na nie dają § 27 i § 28 *Badan*. Mówią one, iż istnieją trzy prawa, pod jakie podpadają wszystkie procesy umysłowe w znaczeniu operacji na ideach, tj. prawa regulujące związki między

¹ D. Hume, *Badania dotyczące rozumu ludzkiego*, tłum. J. Łukasiewicz, K. Twardowski, De Agostini Polska Sp. z o.o. 2001, § 18–25.

ideami, mianowicie podobieństwo, styczność w czasie lub przestrzeni i przyczyna lub skutek.

Stanowisko Hume'a interpretuję naturalistycznie. Wedle takiej interpretacji umysł należy do świata fizycznego, a sposób jego badania wpisuje się w paradygmat nauk o poznawaniu. Stawiam pytania następujące. Na czym polega udział Hume'a we współczesnej wersji naturalistycznego ujęcia umysłu? Jakie są trudności tego programu badawczego? I w końcu, czy jego zwolennikom, do których sam się zaliczam, nie pomyliły się życzenia z rzeczywistością?

2. Hume i koneksjonizm

Na czym polega związek Hume'a z koneksjonizmem? By odpowiedzieć, zestawmy ze sobą dwa twierdzenia rekonstruujące stanowiska Hume'a i koneksjonisty.

(1) Hume: przebieg procesów umysłowych/poznawczych opiera się na związkach pomiędzy ideami, które są kopiami impresji, przez co „zawierają” materiał empiryczny w formie reprezentacji własności obiektów i jako takie podlegają trzem prawom kojarzenia idei: styczności, podobieństwa, przyczyny i skutku.

(2) Koneksjonizm: przebieg procesów umysłowych/poznawczych polega na operacjach, jakie zachodzą na elementach reprezentujących własności obiektów. Istotne są tu, podobnie jak wyżej, po pierwsze, reprezentacje własności, oraz po drugie, typy związków między nimi.

Tym co łączy (1) i (2) w kwestii praw umysłowych jest twierdzenie, że koneksje zachodzące pomiędzy reprezentacjami własności konstytuują prawa przebiegu procesów umysłowych. Zbieżność praw Hume'a z prawami określającymi związki reprezentacji własności w modelu koneksjonistycznym stanie się jeszcze lepiej widoczna, gdy porównamy Hume'a zasady kojarzenia idei oraz niektóre własności sieci neuropodobnej.

Hume'owskie zasady mówią, że idee, jakich zazwyczaj doświadczamy razem, zostają skojarzone w umyśle. Jeżeli pojawia się pierwsza, pociąga to za sobą wystąpienie drugiej. To zasada kojarzenia idei przez przyleganie i styczność w czasie i przestrzeni. Natomiast zasada kojarzenia przez podobieństwo mówi, że jeżeli dwie idee są podobne, wtedy to, co kojarzy się z pierwszą, kojarzy się z drugą. W modelu koneksjonistycznym założeniom tym odpowiadają dwie własności sieci neuropodobnej. Prawu styczności odpowiada schemat zachowania sieci podczas treningu i uczenia się: po otrzymaniu danych wejściowych i informacji o poprawnych danych wyjściowych sieć będzie starała się powtórzyć taki schemat połączeń w przyszłości. Z kolei prawo podo-

bieństwa może być realizowane przez mechanizm automatycznego uogólnienia wedle zasady: jeżeli A ma cechy x, y , to B , które jest podobne do A , też ma prawdopodobnie cechy x, y . Przy czym uogólnienia z jednego przypadku obejmują inne, podobne przypadki.

W obu wypadkach mechanizmy umysłowej przyczynowości mają charakter asocjacyjny. Powstawanie skojarzeń opiera się na bliskości czasoprzestrzennej – Hume – czy też statystyce pobudzeń sieci, a nie na przykład na związkach między treściami myśli.

Jeśli rzeczywiście koneksjonizm jest realizacją idei asocjacionizmu, to być może, oprócz niewątpliwych zalet, dziedziczy także wady i trudności asocjacionistycznego modelu procesów umysłowych.

3. Naturalizm

Koneksjonizm można zinterpretować jako jeden ze stylów współczesnej architektury umysłu. W końcu drugiej połowy XX w. nazwano go „metaforą mózgową” dla odróżnienia od konkurencyjnego, nieco starszego trendu, tak zwanej metafory komputerowej². Mimo pewnych różnic oba należą do naturalistycznego programu badawczego, w którym umysł wyjaśnia się w sposób komputacyjno-mechanicystyczny, czasem wykorzystując narzędzia wypracowane przez behawioryzm. Trzecim stylem architektury umysłu jest szybko rozwijająca się teoria i praktyka ucieleśnionego poznania.

Celem przedsięwzięcia naturalistycznego jest model i symulacja w sztucznych systemach wyższych czynności poznawczych, jak rozwiązywanie problemów czy reprezentacja wiedzy, oraz niższych, odpowiedzialnych za aktywność sensomotoryczną.

W dalszej części zastanowię się nad tym, co rzeczywiście udaje się wyjaśnić dzięki tej redukcjonistycznej strategii? Otóż obawiam się, że wbrew naszym nadziejom i życzeniom, o wiele mniej, niż chcielibyśmy.

² Jak zauważył w rozmowie dr Witek Hensel, gdyby dziś w obecności zwolennika architektury klasycznej lub koneksjonisty użyć określenia „metafora”, podziałaloby to na niego jak płachta na byka. O metaforze komputerowej i mózgowej można było mówić w czasach Turinga (jeśli użycie terminu „metafora mózgowa” w kontekście rozważań Turinga wydaje się dziwne, warto przypomnieć sobie koniec jego klasycznego artykułu *Maszyna myśląca a inteligencja*, w: B. Chwedeńczuk (red.), *Filozofia umysłu*, Fundacja Aletheia 1995, gdzie rozważa on trening sieci neuropodobnej). Współcześnie są to silne programy badawcze, a przynajmniej za takie powszechnie uchodzą. Witek Hensel dodaje, że „używanie tej «metaforycznej» terminologii nie rozjaśnia też niczego – sugeruje tylko, że sprawa ma charakter nie naukowy, a li tylko filozoficzny (by nie rzec, artystyczny)”. Idąc za jego radą, w dalszej części będę unikał tych metaforycznych sformułowań. Natomiast użycie wyrażenia „metafora mózgowa” w tytule ma znaczący charakter symboliczny, z którego nie rezygnuję.

3.1. Architektura umysłu

Klasyczna teoria architektury umysłu – dawniej metafora komputerowa – opiera się na porównaniu oprogramowania komputera (*software*) do działania umysłu, zaś fizycznego środowiska, w którym funkcjonują programy (*hardware*), do fizycznego, np. neurobiologicznego, realizatora umysłu, jakim jest mózg i układ nerwowy. Podstawowe założenia tego stanowiska to funkcjonalizm, komputacja, wieloraka realizacja i teza o fizycznym systemie symboli. W skrócie mówią one, że:

- (a) Umysł jest tym, co robi mózg, lub jest funkcją mózgu – funkcjonalizm.
- (b) Procesy przetwarzania informacji mają charakter obliczeniowy, tj. polegają na operacjach na symbolach – komputacja.
- (c) Procesy umysłowe traktuje się tu jako procesy przetwarzania informacji, przy czym przyjmuje się, że swoiste cechy nośnika informacji nie wpływają na jego funkcję i przez to nie grają roli w procesach informacyjnych – wieloraka realizacja stanów umysłowych w fizycznie różnych środowiskach.
- (d) Zarówno umysł, jak i komputer realizujący program stanowią dwa przypadki ogólnej klasy systemów zdolnych do inteligentnego zachowania – teza o fizycznym systemie symboli (Herbert Simon i Allen Newell).

Z kolei koneksjonizm to program badawczy funkcji poznawczych i ogólna koncepcja umysłu ludzkiego, która proponuje alternatywny wobec architektury klasycznej model działania systemu poznawczego, oparty na tak zwanych sieciach neuropodobnych. Dwie idee leżące u podstaw modeli sieciowych mówią, że:

- (a) Architektura i działanie umysłu jest wyznaczone przez architekturę i działanie mózgu. Zatem model umysłu powinien opierać się na modelu działania mózgu, stąd zamiast o metaforze komputerowej koneksjoniści mówili o metaforze mózgowej.
- (b) Klasyczne modele umysłu obecne np. w silnej wersji sztucznej inteligencji ignorują architekturę mózgu. W związku z tym, wyróżnionym poziomem wyjaśniania nie powinien być poziom funkcjonalno-symboliczny, ale poziom implementacyjny.

Założenie, na jakim opiera się taka strategia, głosi, że funkcjonowanie systemu najprawdopodobniej wynika z jego wewnętrznej organizacji oraz jej własności. Przy czym cechy fizyczne implementacji nie są bez znaczenia, a grają tu zasadniczą funkcjonalną rolę.

Tabela. Najważniejsze różnice między stylami architektury umysłu

Architektura klasyczna	Koneksjonizm (dawniej asocjacionizm)
Model umysłu oparty na maszynie Turinga	Model umysłu oparty na architekturze mózgu
Procesy umysłowe przedstawiane jako obliczeniowe manipulacje na symbolach, przebiegające wedle reguł logiki	Działanie sieci i reakcja na wyjściu nie zależy od programu, ale od przebiegu procesu rozchodzenia się pobudzenia w sieci i wag połączeń
Sekwencyjne przetwarzanie informacji	Rozproszone i równoległe przetwarzanie informacji
Symboliczna reprezentacja wiedzy	Subsymboliczna reprezentacja wiedzy
	Zdolność uczenia się
	Łagodna degradacja błędu

3.2. Hume i behawiorystyczna interpretacja koneksjonizmu

Jak powiedziałem wcześniej, naturalistyczna teoria umysłu w swojej koneksjonistycznej odmianie wykorzystuje narzędzia i osiągnięcia behawioryzmu. W tym nurcie można umieścić Hume'owski asocjacionizm. Sądzę, że to spostrzeżenie wzbudzi, delikatnie mówiąc, silne kontrowersje. Już się z niego tłumaczę.

Wyjaśnienie koneksjonistyczne jest w oczywisty sposób silnie związane z asocjacionizmem poprzez prawa kojarzenia idei. Natomiast asocjacionizm i koneksjonizm łączy z behawioryzmem prawo efektu, które mówi, że siła nawyku zależy od częstości pozytywnego wzmocnienia w przeszłości. Weźmy na przykład klasyczne badania Pawłowa. Pawłow postawił sobie następujące pytanie: skoro odruchy warunkowe nie są wrodzone, to w jaki sposób nabywają je zwierzęta i ludzie? Jego teza głosiła, że jeśli dany bodziec często jest obecny w środowisku psa podczas jego karmienia, to bodziec ten w mózgu zwierzęcia zostanie skojarzony z jedzeniem – będzie sygnalizować zbliżające się pożywienie³. Zbieżność teorii Hume'a i behawioryzmu polega na tym, że mechanizm warunkowania funkcjonuje dzięki mechanizmom kojarzenia idei przez bliskość czasoprzestrzenną. W eksperymencie Pawłowa bodźce początkowo neutralne, jak odgłos kroków, które „w naturalny sposób” nie powodowały wydzielania się śliny, po pewnym czasie i pewnej liczbie powtórzeń

³ Źródło: R.R. Hock, *40 prac badawczych, które zmieniły oblicze psychologii*, tłum. E. Wojtych, Gdańskie Wydawnictwo Psychologiczne 2009, „Nie tylko o śliniących się psach”.

– styczność czasoprzestrzenna – zostały skojarzone przez psy z pożywieniem i w konsekwencji wywoływały reakcję wydzielania śliny⁴.

Co można wyjaśnić w ten sposób? Wiele. Wykorzystując klasyczne warunkowanie, wyjaśnia się między innymi takie zachowania człowieka, jak na przykład powstawanie emocji, fobii, lęków, reakcje na reklamę czy też mechanizmy podniecenia seksualnego.

4. Cele badawcze

Jakie zadania stoją przed naturalistycznym programem badawczym?

Jak wspomniałem, pierwszym celem jest model wyjaśniający strukturę i funkcjonowanie ludzkich procesów umysłowych lub wiarygodna, adekwatna symulacja tzw. niższych i wyższych czynności poznawczych. Wyższe czynności poznawcze to na przykład rozwiązywanie problemów, myślenie, reprezentacja wiedzy i analiza języka naturalnego. Z kolei niższe czynności poznawcze, które bada koneksjonizm i teorie ucieleśnionego poznania, są realizowane równoległe przez organizmy biologiczne w bardzo krótkim czasie. Odpowiadają one za percepcję, postrzeganie obiektów, kategoryzację, szybkie skojarzenia, aktywność sensomotoryczną. Twierdzi się przy tym, iż wyższe czynności poznawcze muszą być jakoś oparte na niższych. Drugie wyzwanie stanowi stworzenie sztucznych systemów, których poziom kompetencji poznawczych w różnego typu zadaniach wymagających inteligencji będzie wyższy niż ludzki.

Obliczeniowo zorientowany naturalista staje przed pewnym dylematem. Otóż wedle klasycznego modelu architektury umysłu myślenie jest procesem podporządkowanym regułom logiki i języka, przebiegającym przy użyciu symboli. Jest więc czymś naturalnym ujmowanie go jako procesu o charakterze obliczeniowym. Z kolei przyjmując nastawienie koneksjonistyczne, twierdzi się, iż myślenie ma niewiele wspólnego z „twardą” logiką. Nie odbywa się pod kontrolą sztywnych reguł określających kolejne kroki operowania jednoznaczными danymi – symbolami. Zarówno myślenie, język, jak i „dane” są pełne niejasności, wieloznaczności i nieracjonalności czy niekonsekwencji. Myślenie jest procesem mechanicznym, komputacyjnym, ale mechanizm ten ma charakter skojarzeniowy i probabilistyczny.

Niektórzy obliczeniowo zorientowani naturaliści zamierzają rozwiązać ten dylemat eklektycznie⁵. Uznają oni, że myślenie oparte na regułach logiki klasycznej – sztywno określających porządek sekwencji kolejnych kroków – powstaje na fundamencie procesów umysłowych, których strukturę lepiej

⁴ Tamże, s. 95.

⁵ Takie rozwiązanie jest też najbliższe autorowi tego artykułu, który popadając w pewne zmieszanie, zdał sobie sprawę z tego, że jak dotąd ma ono charakter życzeniowy.

charakteryzuje rozmyta logika i statystyka – typowe dla modelu rozproszonego przetwarzania informacji.

Problem w tym, że póki co takie rozwiązanie ładnie wygląda na papierze. W rzeczywistości nie jest to rozwiązanie, ale życzenie lub postulat rozwiązań⁶.

W badanie tych problemów zaangażowane są narzędzia wypracowane przez kilka dziedzin naukowych: psychologię poznawczą i behawiorystyczną, neurofizjologię lub neuropsychologię, informatykę, teorię sztucznej inteligencji, a także filozofię umysłu i języka.

Gdzie w tym naturalistycznym wyjaśnieniu umysłu jest miejsce Hume'owskiego asocjacionizmu? Jedną z odpowiedzi przynosi neurocybernetyka⁷. Wedle tej (inter)dyscypliny badawczej funkcje poznawcze są wspomagane przez pamięć kilku typów. Pamięć rozpoznawcza jest odpowiedzialna za identyfikowanie znanych obiektów i pozwala na spostrzeganie innego niż oczekiwany stanu rzeczy. Pamięć proceduralna wzmacnia umiejętności manualne i określa kolejność kroków działania. Pamięć semantyczna potrzebna jest przy interpretacji sensu komunikatu oraz przy wykorzystywaniu złożonych struktur wiedzy. Pamięć robocza kombinuje ze sobą fragmentaryczne informacje w większe całości. Oprócz wspomnianych typów pamięci występuje jeszcze pamięć skojarzeniowa. Jej funkcja polega na automatycznym, stosunkowo nieskomplikowanym wnioskowaniu. Ten typ pamięci bierze udział w realizacji procesów warunkowania klasycznego. Podejrzewam, że tu można by umieścić asocjacionizm Hume'a i badania Pawłowa. Z tym poważnym zastrzeżeniem, że Hume sądził, iż jego teoria opisuje prawa rządzące wszystkimi procesami umysłowymi, a nie jedynie przyczynia się do wyjaśnienia mechanizmów pamięci skojarzeniowej.

5. Rozczarowanie

Jakie trudności napotyka realizacja tego programu? Wbrew naszym optymistycznym przewidywaniom okazuje się, że nie potrafimy jeszcze symulować i modelować komputacyjnie niższych, stosunkowo prostych funkcji poznaw-

⁶ Jak zauważa Włodzisław Duch: „Dobre szanse na realizację tego programu daje technika. W technice takie czynności są podstawą działania robotów. Inteligentne roboty mają się rozwijać podobnie jak dzieci, przyswajając sobie sens symbolicznych przekazów i poleceń poprzez działanie w realnym świecie – prowadzi to do koncepcji filozofii ucieleśnienia poznania z tezą: sens wszystkich symboli określają specyficzne sekwencje działań sensomotorycznych”, w: R. Tadeusiewicz (red.), *Neurocybernetyka teoretyczna*, Wydawnictwo Uniwersytetu Warszawskiego 2009, s. 273.

⁷ Źródło: R. Tadeusiewicz (red.), *Neurocybernetyka teoretyczna*, tam część autorstwa Włodzisława Ducha.

czych, jak percepcja czy identyfikacja obiektów, w takim stopniu, jak robią to prymitywne zwierzęta. Podobnie jest z symulacją wyższych czynności poznawczych, jak na przykład rozumieniem języka naturalnego, wymagającym wykorzystywania złożonych struktur wiedzy, podejmowaniem decyzji czy rozwiązywaniem problemów. Także one wymykają się próbom naśladowania i modelowania komputerowego⁸.

Dlaczego tak jest? Na pewno jedną z przyczyn tego stanu jest bardzo wysoka złożoność naturalnych systemów poznawczych. To znaczy komputerowe symulacje funkcji poznawczych są nadal zbyt proste i ubogie w porównaniu ze skomplikowanymi naturalnymi systemami poznawczymi. Jak pisze neurokognitywista Włodzisław Duch: „Nie wystarczą tu proste inspiracje biologiczne, dotyczące działania poszczególnych neuronów lub ich niewielkich grup. Konieczne jest uwzględnienie wielkoskalowej architektury mózgu i tworzenie modeli funkcji wyspecjalizowanych obszarów”⁹.

Oprócz tej ogólnej przyczyny kłopotów, główne odmiany naturalizmu komputacyjnego borykają się ze swoimi wewnętrznymi problemami. Dotyczą one na przykład reprezentacji indywiduum, kompozycyjności i produktywności myśli, ich spójności semantycznej czy problemu RAM, o którym już wspominałem.

5.1. Trudności koneksjonizmu

a) Kłopot z indywiduum

Trudność ta pojawia się, ponieważ w sieciach neuronowych możliwe jest reprezentowanie poszczególnych własności przysługujących w rzeczywistości obiektom, ale nie samych indywidualnych obiektów. Na przykład, możemy co prawda reprezentować „jarzynowość”, ale nie indywidualną jarzynę. Kłopot w tym, że w sieci dwie jarzyny mogą być reprezentowane tak samo, mimo tego, że są to dwa odmienne przedmioty.

Jeżeli na co dzień posługujemy się pewnym intuicyjnym, abstrakcyjnym pojęciem indywiduum, które rozumiemy jako różne od zbioru własności przysługujących indywiduum, to w sztucznym, koneksyjnym systemie, reprezentującym obiekty jedynie za pomocą cech, nie ma odpowiednika takiego intuicyjnego pojęcia. Jeśli tak, wówczas w modelu sieciowym brakuje czegoś ważnego.

⁸ Tamże.

⁹ Tamże, s. 271.

b) Kłopoty z kompozycyjnością i produktywnością

Kompozycyjność, czyli składanie „częstkowych” znaczeń w większe struktury znaczeniowe, wykonywane wedle jakichś reguł i w jakimś celu, należy do istotnych własności naszego umysłu. Dla ilustracji tego problemu posłużę się przykładem Stevena Pinkera. Weźmy zdanie „Dziecko zjadło cielę”. Pozornie można zaprogramować cechy kompozycyjności w sieci neuronowej poprzez wpisanie w nią elementów reprezentujących odpowiednio „dziecko”, „cielę”, „jeść” i powiązanie ich wzajemnymi związkami. Pojawiają się jednak dwie trudności.

(1) Taki system nie określa tego, czy cielę zjadło dziecko, czy dziecko zjadło cielę, czy może cielę i dziecko jadły. Należy jeszcze określić charakter takich związków, tzn. przypisać elementom ‘cielę’, ‘dziecko’, ‘jeść’ odpowiednie role.

(2) Żeby to uczynić, musielibyśmy każdej kombinacji pojęć i ról – (a) „dziecko zjadło cielę”, (b) „cielę zjadło dziecko”, przypisać odpowiedni węzeł – neuron (a), (b) itd., ale wtedy liczba kombinacji myśli, rosnąc wykładniczo, przekroczy liczbę neuronów w mózgu.

W tym wypadku koneksjonizm narusza dwie cechy naszych procesów umysłowych – kompozycyjność i produktywność myśli (oraz zdań), a oprócz tego, wbrew zapewnieniom samych koneksjonistów, jest nierealistyczny czy niewiarygodny biologicznie.

Istnieje jednak rozwiązanie konwencjonalne. Pytanie brzmi: jak wyjaśnić to, że nasze myśli (i zdania) odznaczają się cechą kompozycyjności oraz że tworzą one zbiór produktywny? Odpowiedź komputacjonisty jest następująca: nasze myśli i język mają cechy kompozycyjności i produktywności, bo obok skończonego słownika – zbioru symboli, jakimi manipulują nasze umysły – posiadamy także skończony zbiór reguł kombinacji symbolicznych, co daje nam możliwość tworzenia potencjalnie nieskończenie wielu myśli i zdań.

c) Spójność semantyczna myśli

Kolejny problem koneksjonizmu stanowi modelowanie takich procesów umysłowych jak rozumowania, tj. przechodzenie od prawdziwych przesłanek do prawdziwych wniosków. Wedle asocjacionizmu i koneksjonizmu procesami umysłowymi rządzą prawa kojarzenia idei, w związku z tym pojawia się pytanie sformułowane na przykład przez Fodora: „w jaki sposób obiekty przedstawiane w ideach i obiekty przyczynowo przez nie wywoływane (myśli wywoływane przez treść innych myśli) mają być w spójny sposób ze sobą powiązane?”¹⁰.

¹⁰ J. Fodor, *Jak grać w reprezentację umysłową*, w: Z. Chlewiński (red.), *Modele Umysłu*, PWN 1999, s. 37–43.

Problem polega na tym, że procesy umysłowe tego typu wymykają się prawom asocjacyjnym. Myśli występujące w rozumowaniach w roli przesłanek i wniosków mogą nie występować w takich wzajemnych relacjach, o jakich mówi asocjacionizm i współczesny koneksjonizm, tj. styczności i podobieństwa. Rozumowanie nie jest kojarzeniem idei. Biorąc to pod uwagę, wydaje się, że model mechanizmu procesów umysłowych takich jak myślenie, jako polegających na przekształceniach symboli wedle reguł logiki, jest lepszym kandydatem do wyjaśnienia mechanizmu rozumowania i semantycznej spójności myśli.

5.2. Trudności metafory komputerowej

a) Komputer i mózg

Wedle metafory komputerowej nie tylko umysł, ale i mózg można opisać poprawnie jako maszynę syntaktyczną. W tym ujęciu mózg to biologiczny realizator umysłu, maszyna syntaktyczna przetwarzająca informacje na sposób obliczeniowy. Łagodnie rzecz ujmując, należy powiedzieć, że jest to dyskusyjna kwestia. Z perspektywy paradygmatu przetwarzania informacji główny problem przybiera tu postać pytania, czy funkcją mózgu jest przetwarzanie informacji w takim samym znaczeniu, w jakim mówimy o przetwarzaniu informacji przez maszynę cyfrową? Jaki jest mechanizm realizacji tej funkcji w mózgu, a jaki w maszynie cyfrowej? Czy zauważalne różnice w „mechanice” przetwarzania informacji w komputerze i w mózgu pozwalają na utożsamienie tych procesów? Następne pytania dotyczą tego, jak te dwa systemy są powiązane ze środowiskiem? Czy i tu jest analogia?

b) Uniwersalność i wieloraka realizacja

Inny problem dotyczy organizacji przechowywania i dostępu do informacji przez mózg i komputer. Otóż w tym względzie różnią się one na podstawowym poziomie. Teoretycznie maszyna Turinga czy sieć neuropodobna może być zrealizowana w różnych środowiskach fizycznych, jak mózg czy komputer. Jest to możliwe – znów teoretycznie – dzięki tej ich własności, jaką jest uniwersalność, ale, jak zauważa Włodzisław Duch, „jak dotąd znane nam praktyczne architektury obliczeniowe ograniczają na wiele sposobów możliwości przetwarzania informacji. W niektórych zastosowaniach komputery przetwarzają informacje lepiej niż mózg, ale nadal w wielu zastosowaniach jest on niedoścignionym wzorem. [...] Nie wiemy, czy architektury kognitywne, zaimplementowane na konwencjonalnych komputerach, kiedykolwiek osiągną

podobne do ludzkiego mózgu możliwości zarówno pod względem niższych, jak i wyższych czynności poznawczych”¹¹.

5.3. Dialog w języku naturalnym i rozumienie języka. Problem RAM

Problem RAM stanowi empiryczne potwierdzenie głęboko ukrytej wady, jaka tkwi w klasycznej teorii architektury umysłu. Dotyczy on zagadnienia rozumienia języka naturalnego przez komputer. Można go przedstawić jako problem symulowania przez maszyny syntaktyczne właściwego ludziom odczytania i rozumienia znaczenia wyrażen języka potocznego, przy czym zasadniczą rolę odgrywa tu wpływ kontekstu na zmiany znaczenia wyrażen języka w zależności od sytuacji. O ile my radzimy sobie z tym zjawiskiem zupełnie dobrze, maszynom sprawia to duży kłopot. Trudność ta powinna zostać przezwyciężona, aby móc mówić o poprawnym komputerowym modelowaniu procesów umysłowych. Gdyby to przedsięwzięcie powiodło się, mielibyśmy kolejną, i to bardzo dobrą podstawę do przyjęcia obliczeniowej teorii umysłu.

Główne pytanie brzmi: czy jest możliwa formalizacja ludzkich procesów poznawczych? Podstawowym zadaniem, które musi zostać wykonane, by osiągnąć ten cel, jest zaprojektowanie systemu komputerowej reprezentacji wiedzy, jaka jest niezbędna dla rozumienia i używania języka naturalnego. Rozwiązaniu tego zadania miały służyć komputerowe systemy RAM czy schematy pojęciowe. Komputer z takim systemem musi potrafić między innymi określić poprawnie kontekst, w jakim należy rozumieć komunikat językowy, oraz wybrać właściwy sposób interpretacji wypowiedzi interlokutora, który pozwoli mu na rozumienie dalszych komunikatów. Kłopot w tym, jak zaprogramować komputer, by był on do tego zdolny? W codziennej komunikacji językowej sposób interpretacji komunikatów jest współwyznaczany przez wiele różnych elementów, jak językowy i pozajęzykowy kontekst, gestykulacja, mimika, intonacja głosu i zachowanie.

Dotychczas skonstruowane, sztuczne systemy reprezentacji wiedzy mają pewną cechę konstrukcyjną, która uniemożliwia im reprezentowanie wiedzy kontekstowej. Programy te operują na skończonej strukturze – „liście” swoich elementów – pojęć i relacji między pojęciami, znaczących obiekty i ich własności. Raz zaprojektowana architektura takiego systemu pozostaje niezmienna. Wszystkie relacje (kombinacje) jego elementów muszą być przewidziane i zaprojektowane. To z kolei ma konsekwencje w postaci małej plastyczności sztucznego systemu w stosunku do tej, jaką wykazuje umysł ludzki w interpretacji znaczeń języka naturalnego. Trzeba zaznaczyć, że w kwestii możliwo-

¹¹ R. Tadeusiewicz (red.), *Neurocybernetyka teoretyczna*, Wydawnictwo Uniwersytetu Warszawskiego 2009, s. 274.

ści tworzenia nowych związków znaczeniowych czy spontanicznych asocjacji modele sieciowe wykazują w tym względzie dużo większą elastyczność niż systemy RAM. Niemniej problem pozostaje.

Co sprawia, że potrafimy tworzyć i rozpoznawać niespotykane wcześniej, nowe związki znaczeniowe? Między innymi zależy to od naszej umiejętności korzystania z bieżącego doświadczenia. Inaczej mówiąc – uczenia się. Patrząc z tej perspektywy, można wiele spodziewać się w kwestii wiernego modelowania reprezentacji wiedzy i rozumienia znaczenia języka naturalnego po takich systemach, które mają zdolność dynamicznego zmieniania swojej organizacji wiedzy pod wpływem interakcji ze środowiskiem, w jakim się znajdują, i zadań, jakie wykonują. Czyli krótko mówiąc, od takich, które potrafią się uczyć.

5.4. Test Turinga

Zwycięstwo komputera w grze w udawanie jest operacyjnym kryterium inteligencji maszynowej. Pomijając kontrowersje wokół wiarygodności kryterium operacyjnego, trzeba stwierdzić, że od dwudziestu lat nie nastąpił duży postęp w rozwiązywaniu tego testu. Konkurencyjne programy o szablonowej konstrukcji nadal nie rozumieją tego, jakie jest znaczenie zdań występujących w dialogu¹². Co prawda znajdują one swoje zastosowania komercyjne w wielu „konwersacjach” dotyczących wąskich dziedzin, jak oferta danej firmy, niemniej semantyczna analiza dialogu nadal pozostaje poza ich zasięgiem, a tym samym nie ma co mówić o rozumieniu języka naturalnego przez komputery.

6. Ucieleśniony umysł

Stosunkowo nowym stylem naturalistycznej architektury umysłu jest tak zwana koncepcja ucieleśnionego poznania. Patrząc z tej perspektywy, obu strategiom, koneksjonistycznej i klasycznej, można zarzucić to, że są niepełne, gdyż ignorują funkcjonalną rolę interakcji między mózgiem i ciałem. O ile w modelu klasycznym świadomie pomija się charakter fizycznego realizatora umysłu, to w przypadku koneksjonizmu błąd ten wygląda na wynik niekonsekwencji. Główna teza filozofii ucieleśnionego poznania mówi, że treść złożonych stanów i procesów umysłowych – jak reprezentacje mentalne, przekonania i rozumowania – cechujących się abstrakcyjnością, powstaje na podstawie reprezentacji percepcyjno-ruchowych i działania organizmu w realnym świecie. Podobnie znaczenie i sens symboli używanych w komunikacji wyznacza

¹² Źródło: R. Tadeusiewicz (red.), *Neurocybernetyka teoretyczna*, tam część autorstwa Włodzisława Ducha, Wydawnictwo Uniwersytetu Warszawskiego 2009.

aktywność sensomotoryczna. Ta sama aktywność jest też współodpowiedzialna za uczenie się.

Natomiast w kwestii architektury umysłu stanowisko ucieleśnionego poznania głosi, że:

(1) Procesy mózgowe nie są niezależne od procesów somatycznych. Istotna jest tu rola biochemicznej i nerwowej komunikacji mózg–ciało.

(2) Sama interakcja umysł/mózg–środowisko to za mało dla wyjaśnienia umysłu, bowiem interakcja taka odbywa się za pośrednictwem ciała. Jeśli model umysłu ma uwzględniać wpływ środowiska na kształt procesów umysłowych, powinien także uwzględniać to, że w takiej interakcji uczestniczy cały organizm, a zatem powinien ujmować także rolę ciała.

A zatem fizykochemiczne podłoże zjawisk umysłowych ma kapitalne znaczenie dla przebiegu i charakteru procesów umysłowych. Co prawda umysł jest tym, co robi mózg (i ciało), ale w dosłownym sensie. A skoro tak, to model odizolowanej od ciała i środowiska maszyny syntaktycznej po prostu nie będzie modelem umysłu ludzkiego. Procesy neurobiologiczne nie przebiegają niezależnie od procesów uznawanych za typowo somatyczne. Funkcjonowanie sieci neuronalnej, odpowiedzialne za działanie umysłu, polega na ciągłej komunikacji i interakcji mózg–ciało. Jeżeli więc traktujemy umysł jak system przetwarzania informacji, należałoby uznać, że także ciało bierze udział w tym procesie.

Koncepcja ucieleśnionego poznania jest atrakcyjna chociażby przez swoje aspiracje do neurobiologicznej wiarygodności. Jednak i ona nie jest wolna od wątpliwości. Na jej gruncie pozostaje wiele nierozwiązanych trudności natury empirycznej i teoretycznej. Główny problem to wyjaśnienie mechanizmów powstawania reprezentacji mentalnych, stanów i procesów umysłowych – jak złożone przekonania – na podstawie stosunkowo mało skomplikowanych reprezentacji percepcyjno-ruchowych.

Krótkie zakończenie. Myślenie życzeniowe

Biorąc pod uwagę zarysowane tu, daleko nie wszystkie, kłopoty naturalistycznego ujęcia umysłu, trudno oprzeć się wrażeniu, że ta postawa przypomina czasem myślenie życzeniowe. Uznajemy za wyjaśnione coś, co jeszcze takie nie jest. Chcielibyśmy, by naturalistyczny obraz umysłu był prawdziwy, ale póki co pozostaje to naszym życzeniem. Czy powinniśmy zarzucić naturalistyczne nastawienie w badaniach nad umysłem? W żadnym razie, ponieważ jak dotąd nie mamy lepszego pomysłu. Możliwe jednak, że wbrew naszym mniemaniom nadal po omacku poruszamy się na grząskim gruncie metafory komputerowej czy mózgowej i nie bardzo wiadomo, gdzie tak naprawdę jesteśmy.

Bibliografia

- Chlewiński Z. (red.), *Modele Umysłu*, PWN 1999.
- Damasio A., *Błąd Kartezjusza. Emocje, rozum i ludzki mózg*, tłum. M. Karpiński, Rebis 1999.
- Devlin K., *Żegnaj Kartezjuszu. Rozstanie z logiką w poszukiwaniu nowej kosmologii umysłu*, tłum. B. Stanosz, Prószyński i S-ka 2000.
- Hock R.R., *40 prac badawczych, które zmieniły oblicze psychologii*, tłum. E. Wojtych, Gdańskie Wydawnictwo Psychologiczne 2009.
- Hume D., *Badania dotyczące rozumu ludzkiego*, tłum. J. Łukasiewicz, K. Twardowski, De Agostini Polska Sp. z o.o. 2001.
- Tadeusiewicz R. (red.), *Neurocybernetyka teoretyczna*, Wydawnictwo Uniwersytetu Warszawskiego 2009.

Streszczenie

Hume zaproponował teorię praw rządzących procesami umysłowymi, której współczesnym rozwinięciem jest koneksjonizm. Inaczej nazywa się go „metaforą mózgową” dla odróżnienia od „metafory komputerowej”. Mimo pewnych różnic obie strategie należą do jednego, naturalistycznego programu badawczego, w którym umysł wyjaśnia się w sposób komputacyjno-mechanicystyczny, czasem wykorzystując narzędzia wypracowane przez behawioryzm. Celem tego przedsięwzięcia jest model i symulacja w sztucznych systemach wyższych czynności poznawczych, jak rozwiązywanie problemów czy reprezentacja wiedzy, oraz niższych, odpowiedzialnych za aktywność sensomotoryczną. Stawiam pytanie o to, co rzeczywiście udało się wyjaśnić przy zastosowaniu tej redukcjonistycznej strategii. Otóż, wbrew naszym nadziejom i życzeniom, o wiele mniej, niż nam się wydawało.