

Marek W. Bielecki

Czas i stawanie się w sztucznych i naturalnych układach kognitywnych

Filozofia Nauki 3/4, 21-32

1995

Artykuł został zdigitalizowany i opracowany do udostępnienia w internecie przez Muzeum Historii Polski w ramach prac podejmowanych na rzecz zapewnienia otwartego, powszechnego i trwałego dostępu do polskiego dorobku naukowego i kulturalnego. Artykuł jest umieszczony w kolekcji cyfrowej bazhum.muzhp.pl, gromadzącej zawartość polskich czasopism humanistycznych i społecznych.

Tekst jest udostępniony do wykorzystania w ramach dozwolonego użytku.

Marek W. Bielecki

Czas i stawanie się w sztucznych i naturalnych układach kognitywnych

Uwagi wstępne

Problemy dotyczące natury czasu, zmiany, stawania się, trwania i innych pokrewnych aspektów rzeczywistości znajdowały się niewątpliwie w centrum refleksji intelektualnej człowieka od samego zarania ludzkiego myślenia. Po wielu wiekach zdominowanych w tej dziedzinie przez rozważania filozoficzne i religijne do głosu doszła wreszcie nauka, obiecująca wskazanie drogi do rozwiązania, lub wręcz samo rozwiązanie, odwiecznych problemów «nękających» filozofów. W efekcie wielu filozofów zaczęło odwoływać się wprost i często bezkrytycznie do modeli i teorii naukowych, zacierając tym samym granicę pomiędzy filozofią i nauką. Jako charakterystyczny przykład wymienić można rozważania «filozoficzne» na temat pojęcia przyczynowości czy determinizmu, korzystające bezpośrednio z pojęć fizycznych, występujących na przykład w mechanice kwantowej. Autentycznie filozoficzne analizy, w odróżnieniu od czysto naukowych rozważań, powinny wszakże wykraczać poza ramy modeli i teorii naukowych. Spojrzenie na niektóre kategorie dotyka czasowego wymiaru rzeczywistości z perspektywy epistemologicznej — którą przyjmuję w tym artykule — a więc wymaga wyjścia poza modele fizyczne i odwołania się do modeli konstruowanych w psychologii, lingwistyce czy neurofizjologii.

Pobieżne nawet przyjrzenie się różnym rodzajom układów kognitywnych — naturalnym, takim jak zwierzę lub człowiek, czy sztucznym, takim jak robot — czyli układów obdarzonych pewną formą pamięci i zdolnością percepcji, ujawnia zasadnicze znaczenie, jakie wymiar czasowy ma dla zrozumienia ich natury. Można też powiedzieć, że jest i odwrotnie: że właśnie samo rozróżnienie, którego dokonujemy

pomiędzy różnymi rodzajami układów kognitywnych, oparte jest w znacznym stopniu, choć często niejawnie, na tym, jaką rolę pełni w nich ten wymiar.

Celem tego artykułu jest krótkie scharakteryzowanie, jak pojęcia odnoszące się do wymiaru czasowego «funkcjonują» w sztucznych i naturalnych układach kognitywnych. W szczególności koncentruję się na tym, w jaki sposób pojęcia czasu i stawania się reprezentowane są w kognitywistyce, a zwłaszcza w obrębie dwóch dominujących w tej dziedzinie kierunków badań. Dziedzina ta, stanowiąca połączenie elementów psychologii, neurofizjologii, lingwistyki, informatyki i filozofii, zajmuje się naukowym badaniem procesów poznania, przede wszystkim myślenia i percepcji, odwołując się do modelu komputerowego przetwarzania informacji. Dwa dominujące w kognitywistyce kierunki badań — to podejście tradycyjne, przyjmujące założenie o symbolicznej naturze inteligencji, oraz teoria sieci neuronowych. W artykule tym pokazuję między innymi, że żadne z tych podejść nie jest w stanie efektywnie modelować charakterystycznego dla człowieka poczucia czasowości, i stąd nie jest zdolne do adekwatnego modelowania procesów, które w jawny sposób zależą od czasu — takich na przykład jak «uczenie się nienadzorowane» (ang. *unsupervised learning*). Korzystając z idei zaczerpniętych z filozofii, teorii systemów i neurofizjologii — staram się uchwycić te najistotniejsze cechy dynamicznej struktury układu mózg-umysł,¹ które umożliwiają niektórym tworum kognitywnym, takim jak człowiek, usytuowanie się w «strumieniu czasu».

Warto też na wstępie ponownie podkreślić, że nie jest moim zamiarem przeprowadzenie tu metafizycznej lub formalnej analizy natury czasu i stawania się. Można tylko zaznaczyć, że w perspektywie epistemologicznej bardziej naturalne wydaje się przyjęcie procesów niż na przykład zdarzeń jako najbardziej podstawowych elementów rzeczywistości. Czas rozumiem w uproszczeniu jako tzw. czas zegarowy, czyli pewien proces cykliczny, generujący zbiór zdarzeń, na którym zdefiniować można na przykład relację 'wcześniej'. Pojęcie stawania się interpretuję jako tzw. «czas subiektywny», związany z poczuciem upływu czasu, przemijalności chwili czy też «przechodzeniem» przyszłości w teraźniejszość, a następnie — w przeszłość. Takie odróżnienie czasu od stawania się jest dość powszechnie przyjęte w metafizyce i odwołuje się do dobrze znanego, pochodzącego od J. McTaggarta rozróżnienia czasu serii-A i serii-B [McTaggart 1927] czy do arystotelesowskiego rozróżnienia na *chronos* i *kairos*.

Ponadto, w odróżnieniu od egzystencjalistów typu M. Heideggera, dla których czas jest kategorią podstawową i nieredukowalną do innych pojęć, a już najmniej naukowych, przyjmuję, że modele naukowe przy całej swej ograniczoności stanowią jeden z głównych sposobów powiedzenia czegoś w sposób pozapoetycki i precyzyjny o świecie.

¹ Używam wyrażenia „mózg-umysł”, będącego bezpośrednim tłumaczeniem angielskiego „*the brain-mind*” aby zasygnalizować, że nie przesądzam tu niczego odnośnie do natury związku pomiędzy ciałem i umysłem.

cie i człowieku.² Z drugiej zaś strony, niezbędne zwykle wykroczenie poza cechujące takie modele ograniczenia dokonywać się powinno nie przez ich odrzucenie i zastąpienie «poezją», lecz przez ich «rozszerzenie» pojęciowe, wynikające z określonej ogólnej filozoficznej koncepcji rzeczywistości. Konieczność takiego pojęciowego wyjścia poza modele naukowe jest szczególnie oczywista w odniesieniu do koncepcji dotyczących natury człowieka, w tym i jego «wymiaru czasowego», albowiem w teoriach naukowych odnoszących się do ludzkiej rzeczywistości nieuchronnie pojawia się niedopuszczalne w nauce błędne koło w rozumowaniu.

Pojęcie czasu w klasycznym podejściu do natury procesów kognitywnych

W klasycznym podejściu do natury procesów kognitywnych przyjmuje się, że inteligencja i w ogólności procesy kognitywne są w swej istocie algorytmicznymi procesami przetwarzania informacji reprezentowanej w formie symbolicznej. W ramach takiego ujęcia pojęcie czasu jest zwykle «zarytmetyzowane», tzn. reprezentowane jako zmiana o dyskretnych wartościach liczbowych, występująca w jakimś systemie logiki temporalnej. W efekcie, takie kategorie jak zmiana, przyczynowość czy uczenie się — a więc kategorie, które jawnie zależą od czasu — sprowadzone zostają do bezczasowych, niezmiennych relacji logicznych. Pomimo znacznych wysiłków logików, konstruujących formalne modele i teorie, i informatyków, piszących imponujące programy komputerowe, nie osiągnięto dotychczas istotnego postępu w filozoficznym czy naukowym zrozumieniu natury czasowości i innych związanych z nim kategorii epistemologicznych, i w umiejętności ich reprezentowania w odniesieniu do układów kognitywnych. Niektórzy badacze w dziedzinie teorii sztucznej inteligencji i kognitywistyki są wszakże optymistyczni i utrzymują, że dzięki logikom niemonotonicznym będziemy wreszcie w stanie „ująć w formie mechanicznej procesy rozumowania, odnoszące się do świata realnego, w sposób zarówno precyzyjny, jak i efektywny” [Shoham 1988, s. 173]. Inni badacze są ostrożniejsi i dostrzegają, że błędem większości klasycznych ujęć w dziedzinie teorii sztucznej inteligencji jest to, iż ignorują one fakt, że samo rozumowanie jest także zachodzącym w czasie procesem (dochodzenia do wniosków), i aby móc reprezentować ludzkie rozumowanie w sposób bardziej adekwatny w sztucznych układach kognitywnych, takich jak komputer, proces ten powinien być modelowany równocześnie z samym procesem wyciągania wniosków [Elgot-Drapkin *et al.* 1991]). Y. Shoham (*op. cit.*) wierzy wszakże, że już obecnie dostępne narzędzia formalne są wystarczające do tego, aby „efektywność procesu przewidywania przyszłości nie odbywała się na koszt jego wiarygodności”. Jeśli jednak potraktować poważnie to, co

² Skłonny jestem nawet przychylić się do tezy silniejszej, głoszonej przez P. Churchlanda, który uważa, że zarówno w dziedzinie filozofii umysłu, jak i filozofii nauki, nie jest już możliwe przeprowadzenie istotnych analiz bez odwołania się do całego szeregu dyscyplin badających system umysł-mózg, takich jak psychologia kognitywna, «obliczeniowa neurologia» (ang. *computational neuroscience*) czy teoria sztucznej inteligencji w naukach komputerowych [Churchland 1989].

się w nauce twierdzi na temat natury procesów zachodzących w tak skomplikowanym układzie, jak umysł-mózg, który ma być «autorem» ludzkich procesów rozumowania, i na temat przewidywalności jego stanów, to wiara taka wydaje się mało usasadniona.

Zanim zajmę się charakterystyką podstawowych własności złożonych układów dynamicznych, takich jak umysł-mózg, przedstawię pokrótce próby ujęcia czasowości w drugim głównym kierunku badań w kognitywistyce, opierającym się na modelu sieci neuronowych, a więc na modelu inspirowanym przez biologiczną wiedzę na temat budowy mózgu.

Pojęcie czasu w modelu sieci neuronowych

Zwolennicy modelu sieci neuronowych podkreślają przede wszystkim «kruchosc» (ang. *brittleness*) — tj. brak uniwersalności i plastyczności — oraz inne ograniczenia systemów konstruowanych w ramach klasycznego podejścia, i dlatego opowiadają się za wolnym od tych problemów ujęciem, opierającym się na modelu reprezentowania zjawisk kognitywnych za pomocą sieci neuronowych. Niektórzy z nich uważają, że kategorie kognitywne mogą być reprezentowane przez geometryczne rozkłady poziomów aktywacji węzłów sieci, w której każdy rozkład odpowiadałby pewnej kategorii. Inni, tacy jak P. Churchland [Churchland 1995], pokładają nadzieję na ostateczne rozwiązanie problemu skonstruowania sztucznej inteligencji, a w szczególności reprezentowania relacji czasowych, w tzw. «sieciach rekurencyjnych» (ang. *recurrent networks*) z ich cyklami granicznymi. Churchland twierdzi, że tradycyjne sieci neuronowe typu *feedforward*,³ które zdolne są do uczenia się na przykładach i odtwarzania zarówno samego porządku czasowego takiego procesu, jak i jego wyników (np. w wypadku NETtalk, nauczenia się zarówno kolejności sylab występujących w przykładzie, jak i wypowiedzenia całego ich ciągu). Sieci takie nie są w stanie rozpoznać, że jedna sekwencja czasowa może być odwróceniem drugiej, np. że jeden ciąg sylab czy znaków tworzących pewne słowo może być dokładnym odwróceniem innego ciągu reprezentującego to samo słowo, ale zapisanego czy wypowiedzianego wstecz. Sieć taka jak NETtalk, pisze Churchland, „nie wie nic o uporządkowaniu czasowym” i „posiada jedynie zdolność dokonywania obliczeń w pewnym porządku czasowym, ale nie obliczeń dotyczących samego takiego porządku” (*ibid.*, s. 98). To, czego, jego zdaniem, brakuje tego typu sieciom, aby były zdolne do reprezentowania porządków czasowych — to jakaś forma krótkoterminowej pamięci, a najlepszym sposobem zorientowania się, od czego własność taka może zależeć, jest przyjrzenie się, jak sieci takie różnią się od mózgu. W mózgu zaś poza ścieżkami typu *feedforward* — czyli «wstępującymi» (ang. *ascending pathways*) — dominującymi zwykle w sieciach, występują też silne powiązania przeciwbieżne od «wyższych» warstw neuronowych do «niższych» poprzez ścieżki typu *feedbackwards* czyli «zstępujące» (ang. *descending*) lub «rekuren-

³Dobrze znanym przykładem jest tu sieć NETtalk, zaprojektowana przez T. Sejnowskiego, zdolna do nauczenia się i powtórzenia stosunkowo prostych zdań.

cyjne» (ang. *recurrent*). Umożliwić ma to ponoć mózgowi «dostęp» nie tylko do bodźców odbieranych przez aparat percepcyjny w danym momencie czasu, ale również do tych bodźców, które zostały odebrane ułamek sekundy wcześniej i częściowo już przetworzone przez układ nerwowy. W ten sposób sieć obdarzona ścieżkami zstępującymi nie jest już „więzieniem nieskończenie cienkiej Płaszczyzny Teraźniejszości”, ale jej „kognitywny zakres rozciąga się o kilka ułamków sekund w Rozciągniętą Przeszłość” (*ibid.*, s. 100)

Następne pytanie — to pytanie o sposób, w jaki mózg koduje informację dotyczącą relacji czasowych. Według Churchlanda dokonuje się to zapewne na wiele sposobów, ale jeden z nich jest całkiem oczywisty. Procesem mózgowym kodującym następstwa czasowe występujące w otaczającym świecie jest cykliczny proces aktywacji grup neuronów. W teorii złożonych układów dynamicznych proces taki określany jest nazwą «cyklu granicznego» [Stewart 1994]. Zdaniem Churchlanda całkowita gama zachowań układów kognitywnych może być wygenerowana przez złożenie odpowiednich «prototypowych» liniowych i cyklicznych procesów motoryczno-percepcyjnych, które w przestrzeni fazowej, reprezentującej poziom aktywacji neuronów, przedstawione mogą być matematycznie jako atraktory, mające charakter odpowiednio tzw. punktów stałych i cykli granicznych. Te pierwsze odgrywają decydującą rolę w procesach rozpoznawania prototypowych kształtów czy obiektów, takich jak np. twarz, czyli rozpoznawania niezmiennych konfiguracji cech fizycznych. Te drugie umożliwiają procesy rozpoznawania prototypowych sekwencji czasowych różnych konfiguracji fizycznych, np. sekwencji odpowiadających „procesowi mrugnięcia okiem, skradania się kota czy tańczenia dwojga ludzi” (*op. cit.*). Jak pisze Churchland, „sieć rekurencyjna zdolna jest rozpoznać nieograniczoną liczbę zewnętrznych struktur rozciągniętych zarówno czasowo, jak i przestrzennie. Pomimo że jej krótkoterminowa «pamięć» sięga jedynie ułamek sekundy w przeszłość, dobrze wytrenowana sieć rekurencyjna jest w stanie reprezentować ciągi czasowe o dowolnej rozciągłości” [Churchland 1995, s. 104]. Cechę tę sieci takie zawdzięczać mają swojej zdolności do rekurencyjnej modulacji wewnętrznej aktywności, co umożliwia im generowanie długich ciągów reprezentowanych w postaci wektorów stanów wewnętrznej aktywacji bez konieczności stymulacji przez bodźce zewnętrzne. Rekurencyjne procesy tego typu występują na wielu poziomach funkcjonowania organizmów, szczególnie w sferze procesów motorycznych. Odgrywają też one kluczową rolę w procesach percepcji. Organizm «rozpoznaje» jakiś obiekt, gdy w odpowiedniej populacji neuronów pojawia się wektor aktywacji «zbliżony» do wektora «prototypowego» dla rozpoznanego obiektu. W analogiczny sposób organizm «rozpoznaje» jakiś proces, gdy w odpowiedniej populacji neuronów pojawia się sekwencja wektorów aktywacji, zbliżona do prototypowej dla rozpoznanego procesu. Kategorie odpowiadające obiektom i procesom tworzą w przestrzeni wektorów aktywacji bogate hierarchie — z tym, że te pierwsze reprezentowane są przez punkty, a te drugie — przez linie. Sieci rekurencyjne są też, według Churchlanda, otwarte na przyszłość, tj. są one w stanie «antycypować» konsekwencje pewnych wydarzeń, dzięki swojej zdolności do

przyspieszania lub skracania sekwencji stanów wewnętrznych reprezentujących procesy zewnętrzne. Jak pisze Churchland, „to, jak daleko w przyszłość jest w stanie taka sieć sięgnąć, zależy od przynajmniej dwóch rzeczy: rozciągłości czasowej i adekwatności prototypowych procesów przyczynowych występujących w świecie oraz od zdolności sieci do nauczenia się ich i późniejszego rozpoznawania ich wczesnych stanów” [Churchland 1995, s. 106-107]. We wszystkich wypadkach, czy to w odniesieniu do „boksera zdającego sobie sprawę z tego, że za chwilę spadnie na niego cios przeciwnika, czy też astronoma zdającego sobie sprawę z tego, że za 5 miliardów lat słońce przestanie świecić”, poczucie czasowości ma to samo źródło — w „sekwencji wektorów aktywacji generowanych przez dobrze wytrenowaną sieć rekurencyjną” (*ibid.*, s. 107).

Pomysł odwołania się do procesów oscylacyjnych dla wyjaśnienia natury procesów kognitywnych nie jest nowy. Pojawił się on już wyraźnie pod koniec ubiegłego wieku w pracach W. Jamesa z dziedziny psychologii. Jednym ze współczesnych najlepiej rozwiniętych ujęć, wywodzących się z tej tradycji, jest teoria przetwarzania informacji kognitywnej, określana nazwą ATR (*Adaptive Resonance Theory*), zaproponowana pod koniec lat siedemdziesiątych przez S. Grossberga i od tego czasu wciąż przez niego i jego współpracowników udoskonalana (cf. [Grossberg 1982], [Carpenter & Grossberg 1992]). Kluczowa idea teorii ART jest wynikiem spostrzeżenia, że konflikt pomiędzy stabilnością i plastycznością sieci neuronowej może być rozwiązany w sieciach o architekturze, składającej się z co najmniej trzech poziomów, gdzie poziom wyjściowy połączony jest z poziomem wejściowym, umożliwiając w ten sposób powstanie sprzężenia zwrotnego. Prowadzi to w efekcie do pojawienia się dynamik wewnętrznych o charakterze oscylacyjnym, które nie są zdeterminowane bezpośrednio przez dane wejściowe. Dynamiki takie umożliwiają kodowanie przez sieć regularności nie tylko przestrzennych, ale i czasoprzestrzennych, np. występujących w procesach rozpoznawania mowy czy generowania procesów motorycznych. Według tej teorii — funkcjonalnymi jednostkami zarówno krótkoterminowej, jak i długoterminowej pamięci, są przestrzenne rozkłady aktywacji neuronów (węzłów sieci). Analizy modelu matematycznego wskazują na to, że rozkłady takie mają tendencję do synchronizowania się ze sobą w czasie. Synchronizacja pojawia się w efekcie wzbudzenia przez bodziec drgań rezonansowych w sieci, które oddziałują ze sobą nielokalnie i nieliniowo, generując w rezultacie fale stojące. W języku teorii układów dynamicznych — fale stojące są atraktorami typu cyklu granicznego. Kategorie kognitywne mogą być reprezentowane w postaci „stojących fal rezonansowych, powstających, gdy sprzężone ze sobą komórki sieci neuronowej oscylują w fazie” [Carpenter & Grossberg 1992, s. 388]. Sieci takie „zdolne są do uczenia się stabilnych kategorii rozpoznawczych w odpowiedzi na dowolną sekwencję danych wejściowych” (*ibid.*, s. 369).

Grossberg odwołuje się «filozoficznie» do E. Macha i — podobnie jak ten — wyraża nadzieję, że jego badania przyczynią się do przewyciężenia dualizmu ducha i materii. Mrzonki takie nie dziwią, gdy snute są przez naukowców, ostatnio na przykład

przez fizyka R. Penrose'a czy biologa F. Cricka. Niezrozumiałe wydawać by się mogło natomiast, dlaczego również wielu filozofów, m. in. P. Churchland, upatruje w naukowych modelach mózgu klucza do rozwiązania „tajemnicy rozumu i duszy”. Staje się to wszakże jaśniejsze, gdy uświadomić sobie, że filozofowie tacy, sytuujący się w ramach tzw. «silnej wersji epistemologii znaturalizowanej», przyjmują, że odkrycia i teorie naukowe posiadają bezpośredni sens epistemologiczny i ogólnie biorąc znaczenie filozoficzne. Nic dziwnego przeto, że często ich „sowa Minerwy wylatuje dopiero o zmierzchu”. W wypadku Churchlanda nastąpiło to ponad dwadzieścia lat po tym, kiedy Grossberg i inni rozwinęli badania dynamiki sieci neuronowych w terminach punktów stałych i cykli granicznych.

Nie znaczy to, że neguję sensowność poszukiwania w teoriach i modelach naukowych inspiracji dla rozważań filozoficznych. Najlepszym współczesnym przykładem sensowności takich poszukiwań może być fakt inspiracji czerpanej z komputerowego modelu przetwarzania informacji. Ważne jest wszakże, aby zdawać sobie sprawę z filozoficznych ograniczeń takich modeli. W przeciwnym razie nietrudno wypowiedzieć całkiem serio absurdalną tezę w rodzaju tej, że „mózg musi mieć znaczną kontrolę na tym jak widzi i słyszy różne rzeczy” [Churchland 1995, s. 113].

Jak dotychczas, żaden z naukowych, komputerowych («symbolicznych»), «konekcyjnych», czy jakichś innych modeli procesów kognitywnych nie był w stanie wyjaśnić takich zjawisk psychologicznych, jak zdolność do niemal natychmiastowego uchwycenia całości rozciągniętego w czasie procesu, np. rozpoznania melodii. Ani model Grossberga ani rozważania Churchlanda na temat sieci rekurencyjnych nie dostarczają odpowiedzi na pytanie, w jaki sposób — ani nawet nie jest jasne, co to znaczy, że — sieci są w stanie identyfikować («rozpoznawać») poszczególne wektory aktywacji czy ich ciągi, i «antycypować» dalsze fazy jakiegoś procesu. Już ponad sto lat temu James zauważył, że odwołanie się do izomorfizmu pomiędzy następstwem stanów mózgu i następstwem stanów świadomości nie jest w stanie wyjaśnić zdolności umysłu do natychmiastowego rozpoznania samej sekwencji czasowej stanów świadomości. Zdolność takiego rozpoznania wymaga bowiem, jak się wydaje, uprzedniego istnienia jakiejś formy świadomości porządku czasowego, a więc poczucia stawania się. Inaczej mówiąc, nie ma w tych koncepcjach niczego, co pozwalałoby wyjść poza «mechanistyczny» model czasu-zegara, w ramach którego nie ma miejsca na poczucie stawania się, a więc przeświadczenie o istnieniu przeszłości, teraźniejszości i przyszłości.

Moje własne koncepcje co do natury układów kognitywnych i roli, którą odgrywa w nich wymiar czasowy, zostały także zainspirowane ideami wyrażonymi uprzednio przez naukowców. Podobnie jak Grossberg, podkreślam nieliniowy, nielokalny i samoorganizujący charakter procesów zachodzących w mózgu-umyśle. Również Churchland zauważa, że sieci rekurencyjne wykazują cechę nieliniowości w tym sensie, że ciąg wektorów aktywacji, reprezentowany za pomocą trajektorii w przestrzeni aktywacji sieci, może się czasami okazać niezwykle wrażliwy na bardzo małe zaburzenia czy to zewnętrzne, czy wewnętrzne, prowadząc w efekcie do znacznego wzmocnienia

małych różnic i nieprzewidywalności zachowania takiego układu. Ani Churchland, ani Grossberg nie piszą — choć są zapewne tego świadomi, bo jest to fakt dobrze znany z badań neurofizjologicznych — że dynamika procesów elektromagnetycznych w mózgu ma charakter chaotyczny w sensie tzw. «chaosu deterministycznego», którego jedną z cech charakterystycznych jest duża wrażliwość na małe zaburzenia (cf. [Skarda & Freeman 1987], [Basar 1990]). Do tego właśnie aspektu dynamiki mózgu-umysłu nawiązuję w następnej części artykułu, gdzie szkicuję własną koncepcję natury procesów zachodzących w mózgu-umyśle. Inaczej wszakże niż Churchland i jemu podobni, nie upatruję w tych koncepcjach ostatecznego rozwiązania odwiecznych problemów filozoficznych, a jedynie źródła stymulujących metafor, poszerzających nasze rozumienie natury człowieka i świata. Staram się też wyjść w swoich rozważaniach poza zakres zaproponowanych modeli naukowych.

Czas i stawanie się w układzie mózg-umysł

Racjonalistyczny nurt w filozofii obfituje w argumenty na rzecz nieredukowalnej natury intencjonalności, tj. niesprowadzalności znaczeń wyrażen językowych do zjawisk opisywanych w kategoriach fizycznych, biologicznych, psychologicznych czy socjologicznych. Jak można, zapytują racjonałiści, sensownie mówić o «wyprowadzeniu» znaczeń z takich zjawisk w sytuacji, gdy nasz dostęp do tychże jest możliwy wyłącznie dzięki znaczeniom i poprzez nie? Jeśli nawet zgodzić się z tym «negatywnym» werdyktem, który wydaje się szczególnie przekonujący w odniesieniu do pojęć związanych z czasem, to ciągle jeszcze można mieć nadzieję, że analiza funkcjonowania układu mózg-umysł da nam wgląd w istotę różnicy pomiędzy tym układem i układami innymi, naturalnymi i sztucznymi, którym nie można przypisać posiadania świadomości — w tym również poczucia upływu czasu. Czy badanie dotyczy przy tym samych takich «obiektów», jak mózg i umysł, i ich dynamiki, czy też, jak zapewne wolałby powiedzieć racjonalista, dynamiki związków znaczeniowych pewnych wyrażen — nie ma to wpływu na szczegółowy kształt mojej koncepcji.

Jednym z głównych nierozwiązanych problemów zarówno konekcjonizmu, jak i symbolicznej sztucznej inteligencji, jest modelowanie procesów złożonych — takich, jak ciągłe «nienadzorowane uczenie się», którą to zdolnością obdarzone są naturalne układy kognitywne w rodzaju człowieka. Sztuczne układy kognitywne — takie jak sieci neuronowe czy roboty operujące według z góry zadanego programu — zdolne są tylko «nauczyć się» określonego zespołu kategorii pod wpływem bodźców zadanych w fazie treningowej, ale gdy później natykają się na bodźce czy sytuacje całkowicie nowe, nie są w stanie wyjść poza «wyczone» kategorie i efektywnie je uogólnić lub przeformułować. Wynika to moim zdaniem z tego, że człowiek — inaczej niż zwierzę, a tym bardziej robot — nie jest wyłącznie «pogrążony w strumieniu czasu», ale jest też sam źródłem «stawania się».

Decydującymi własnościami, odpowiedzialnymi za ludzką zdolność stałego uczenia się, są: złożoność, samoorganizujący charakter oraz plastyczność, a zarazem stabilność

dynamiki procesów mózgowych. Cechy te nie są adekwatnie reprezentowane w zaproponowanym przez Grossberga języku, ograniczającym złożoność procesów mózgowych do dynamik typu punktu stałego i cyklu granicznego. Adekwatniejszym podejściem jest opis tych procesów za pomocą dynamik o charakterze hipercyklicznym i chaotycznym, bowiem dopiero te dynamiki wykazują dostateczny stopień złożoności, plastyczności i stabilności. Takie właśnie podejście zaproponowane zostało m. in. przez neurofizjologa W. Freemana, który uważa, że wrażenia zmysłowe — np. zapachy czy smaki — i inne stany kognitywne mogą być utożsamione z pewnymi elektromagnetycznymi procesami zachodzącymi w mózgu i mającymi charakter tzw. «dziwnych» (chaotycznych) atraktorów (cf. [Freeman 1991], [Skarda & Freeman 1987]). Podejście to — mimo że bardziej adekwatne od rozwijanego przez Grossberga — cierpi na poważny niedostatek filozoficzno-metodologiczny.⁴

Model reprezentujący zjawiska kognitywne, jako dynamiczne procesy o charakterze chaotycznym, wydaje się w ogólnym zarysie trafny. Aby jednak mógł posłużyć do lepszego uchwycenia natury czasowości, cechującej różne rodzaje układów kognitywnych, potrzebne są co najmniej dwa dodatkowe elementy: idea samoorganizacji i idea tzw. «chaotycznej hierarchii» [Rössler 1991]. Pojęcie samoorganizacji odnosi się do procesu globalnego uporządkowania, pojawiającego się spontanicznie w złożonych, niemal homogenicznych układach, składających się z wielu oddziałujących ze sobą fluktuujących podukładów. Najbardziej znane przykłady samoorganizacji — to zachowanie się konsumentów na rynku, domen w materiałach magnetycznych i neuronów w mózgu.

Istnieje wiele dowodów na to, że u podstaw wszystkich istotnych procesów biologicznych a także umysłowych leżą odpowiednie zjawiska cykliczne, określane zwykle jako «zegary» biologiczne lub wyznaczniki rytmu. Zegary takie są systemami biologicznymi, których rytm i cała dynamika są zwykle wynikiem synchronizacji składających się na nie jeszcze bardziej podstawowych procesów cyklicznych. Synchronizacja pojawia się tu najczęściej w efekcie samoorganizacji elementarnych procesów konstytuujących dany system-zegar, i jego nielokalnych, nieliniowych oddziaływań z innymi strukturalnie lub funkcjonalnie mu podobnymi systemami-zegarami [Winfrey 1990]. Wyznaczany przez takie systemy rytm nie jest niezmienny, gdyż mogą mieć na niego wpływ zarówno inne strukturalnie lub funkcjonalnie analogiczne systemy, jak i dynamika całego metasystemu, którego są częścią.

Taka właśnie hierarchiczna budowa wydaje się cechować układ umysł-mózg [Holden 1991]. W sensie strukturalnym można w nim bowiem wyróżnić kilka poziomów, przy czym każdy zawiera sprzężone ze sobą określone układy oscylacyjne. Biologicznie najniższy poziom składa się z indywidualnych neuronów, których oscylacyjne

⁴Na filozoficzne trudności tego podejścia zwracałem uwagę już kilkakrotnie, m. in. w 1991 r. na IX Kongresie Logiki, Metodologii i Filozofii Nauki. Nie ma tu jednak miejsca, aby wątek ten rozwinąć.

wyładowania synchronizują się zwykle ze sobą w wyniku samoorganizacji w ramach pewnych zgrupowań («domen») neuronów. Całe takie zgrupowania synchronizują się z kolei ze sobą nawzajem, tworząc zespoły jeszcze wyższego rzędu, które sięgają w końcu najwyższego poziomu, poziomu półkul mózgowych. Sprzężenie pomiędzy zespołami neuronów może mieć charakter czysto strukturalny, poprzez bezpośrednie połączenia neuronowe, lub funkcjonalny — w rezultacie samoorganizującego procesu synchronizacji drgań i innych cech dynamicznych. Ponadto istnieje też w mózgu swoisty centralny wyznacznik rytmu całego organizmu w postaci grupy komórek nerwowych, określanych nazwą „jądra skrzyżowania podwzgórza” („*nucleus supra-chiasmaticus*” — SCN) [Hartwig 1984], które połączone są z jednej strony ze źrenicą oka, mogącą dzięki temu wpływać na ich wewnętrzny rytm, a z drugiej strony z przysadką mózgową, która reguluje wydzielanie hormonu melatoniny, wpływającego z kolei na aktywność mózgu — w tym również w formie sprzężenia zwrotnego na same SCN. Funkcjonowanie mózgu jest z pewnością jeszcze bardziej skomplikowane niż to, które rysuje się na podstawie aktualnie dostępnej wiedzy. Niezależnie od strukturalnych czy funkcjonalnych szczegółów procesów zachodzących w mózgu — najważniejsze, z punktu widzenia ich dynamiki, jest istnienie hierarchii samoorganizujących się poziomów i sprzężenia zwrotnego pomiędzy nimi.

W czysto dynamicznym sensie — układ umysł-mózg, jak również każdy jego funkcjonalnie lub strukturalnie wyróżniony poziom, może być reprezentowany za pomocą rodziny atraktorów, np. chaotycznych. Zarówno każda rodzina, jak i każdy atraktor mogą być scharakteryzowane za pomocą pewnych dynamicznych własności — takich jak na przykład «uogólniona faza» lub «rytm» — które są zdeterminowane relatywnie w odniesieniu do «uogólnionej fazy» i «rytmu» super-systemu, również reprezentowanego przez atraktor, lub ich rodzinę.⁵ Z tego punktu widzenia bodziec nie tyle, jak pisze Freeman (*op. cit.*), «wywołuje» w elektromagnetycznych procesach mózgu dynamiczną reakcję o charakterze atraktora, ile raczej zaburza «uogólnioną fazę», «rytm» i inne dynamiczne własności już istniejących atraktorów. W efekcie powoduje to, że zarówno każdy z atraktorów jak i cała ich rodzina może zacząć «ewoluować adiabatycznie»,⁶ który to sam proces też może być reprezentowany za pomocą atraktorów. Z kolei takie metapoziomowe atraktory mogą, jak sądzę, sprzęgać się ze sobą w procesie samoorganizacji i, synchronizując w ten sposób swoje «rytmy» i koordynując inne własności dynamiczne, generować w rezultacie dynamiki kolejnego, wyższego rzędu. Proces taki prowadzi do powstania dynamicznej struktury hierarchicznej, która jest, jak przypuszczam, rodzajem dynamicznej hierarchii analogicznej do chaotycznej hierarchii Rösslera [Kaneko 1991].

⁵W wypadku atraktorów typu cyklu granicznego, reprezentujących proste zegary biologiczne, takie uogólnione własności redukują się do «zwykłej» fazy i rytmu, które zdefiniować można, tak jak czyni to A. Winfree (*op. cit.*), i których dynamikę badać można w odpowiedniej przestrzeni faz.

⁶Ideę tę czerpię z [Ruelle 1987].

Kilka cech procesów kognitywnych jest wyraźnie uwypuklonych w zarysowanym tu modelu dynamiki mózgu-umysłu. «Strumień świadomości», o którym pisze James i inni psychologowie, utożsamić można z całą hierarchią dynamiczną. Poszczególne stany kognitywne, takie jak wrażenia zmysłowe, odpowiadają wtedy nie jak w podejściu Freemana atraktorom chaotycznym, bezpośrednio «wzbudzonym» przez bodźce, ale pojawieniu się kolejnego poziomu dynamicznego, będącego rezultatem zaburzenia przez bodziec dynamiki atraktorów «wypełniających» mózgu-umysł. W tej sytuacji tożsamość każdego takiego stanu zależna będzie od całej «historii» oddziaływań układu z otoczeniem. Inaczej więc niż w podejściu Freemana nie da się tu powiedzieć, że jeśli ten sam atraktor pojawia się ponownie w procesach mózgowych, to odpowiadać mu musi ten sam stan kognitywny — na przykład to samo wrażenie zmysłowe.

Uwagi końcowe

Zarysowany model dynamiki mózgu-umysłu nie daje wyjaśnienia poczucia czasowości (stawania się). Nie sądzę, aby było to w ogóle możliwe bez odwołania się do innych pojęć czasowych, takich jak pamięć czy oczekiwanie. Wskazuje on jedynie, choć może tylko metaforycznie, na to, jak czas serii-B, tj. czas procesów «cyklicznych», występujących na pewnym poziomie dynamicznej hierarchii, prowadzi do pojawienia się czasu serii-A, czasu «stawania się» w procesie tworzenia się następnego poziomu hierarchii. Dopiero przez odniesienie do tego następnego poziomu można w ogóle mówić o swoistej «cykliczności» procesów niższego poziomu w sensie czasu B. Każdy kolejny poziom dynamicznej hierarchii cechują jego własne charakterystyczne procesy «cykliczne» i odpowiadający im czas B, prowadzące ponownie do następnego poziomu hierarchii i następnego czasu A, itd. W sensie «metafizycznym» można powiedzieć, że model ten sugeruje koncepcję czasu jako tworu wielowymiarowego i «emergentnego».

Nakreślony model dynamiki mózgu-umysłu wskazuje też, że postulowanie korelacji pomiędzy stanami dynamicznymi mózgu a stanami kognitywnymi jest nawet w naukowym sensie dość wątpliwą propozycją, nie wspominając już o sensie filozoficznym. Co więcej, nie wydaje mi się sensowne mówienie, że intencjonalność (wyrażeń językowych) może być, jak twierdzi J. Searle, «produkowana» przez jakieś «siły sprawcze» mózgu [Searle 1992]. Problem intencjonalności pozostaje kluczowym nierozwiązanym problemem filozoficznym i poruszanie go nie było moim celem. Starłem się raczej zarysować model dynamiki procesów mózgowych, cechującej się większym stopniem plastyczności, stabilności i otwartości niż dynamiki zaproponowane w innych modelach. Model taki otwiera, jak sądzę, perspektywę na bardziej filozoficznie inspirowane próby zrozumienia charakterystycznego dla człowieka poczucia czasowości i takich zależnych od tego poczucia cech, jak zdolności nieustannego uczenia się. Można na koniec zadać pytanie, jaka — z punktu widzenia umysłu — jest specyfika dynamiki ludzkiego mózgu, że umożliwia ona pojawienie się umysłu, co nie ma miejsca w

wypadku zwierzęcia czy robota. Wydaje się, że jest nią możliwość „wspinania się po szczeblach hierarchii chaotycznej, która jest niedostępna zwierzętom czy sztucznym systemom kognitywnym”.⁷

BIBLIOGRAFIA

- Basar, E. (red.)
1990 — *Chaos in Brain Function*, Berlin, Springer.
- Beier, G. & Klein, M. (red.)
1991 — *A Chaotic Hierarchy*, Singapore, World Scientific.
- Carpenter, G. & Grossberg, S. (red.)
1992 — *Neural Networks for Vision and Image Processing*, Cambridge, MA, MIT Press.
- Churchland, P. M.
1995 — *The Engine of Reason, the Seat of the Soul. A Philosophical Journey into the Brain*, Cambridge, MA, MIT Press.
1989 — *A Neurocomputational Perspective. The Nature of Mind and the Structure of Science*, Cambridge, MA, MIT Press.
- Elgot-Drapkin J., Miller M., & Perils D.
1991 — „Memory, Reason, and Time: The Step-Logic Approach”, [w:] R. Cummins and J. Pollock (red.) *Philosophy and AI: Essays at the Interface*, Cambridge, MA, MIT Press, s. 79-104.
- Freeman, W. J.
1991 — „The Physiology of Perception”, *Scientific American* 2, 78-85.
- Grossberg, S.
1982 — *Studies of Mind and Brain: Neural Principles of Learning, Perception, Development, Cognition, and Motor Control*, Boston, Reidel Press.
- Holden, A.
1991 — „Structural, Functional and Dynamical Hierarchies in Neural Networks”, [w:] Beier & Klein, *op. cit.*, s. 187-197.
- Kaneko, K.
1991 — „Climbing Up Dynamical Hierarchy”, [w:] Beier & Klein, *op. cit.*, s. 235-250.
- McTaggart, J.
1927 — *The Nature of Existence*, Cambridge, Cambridge UP.
- Rössler, O. E.
1991 — „The Chaotic Hierarchy”, [w:] Beier & Klein, *op. cit.*, s. 25-47.
- Ruelle, D.
1987 — „Diagnosis of Dynamical Systems with Fluctuating Parameters”, [w:] M. V. Berry *et al.* (red.), *Dynamical Chaos*, London: The Royal Society.
- Searle, J.
1992 — *The Rediscovery of the Mind*, Cambridge, MA, MIT Press.
- Shoham, Y.
1988 — *Reasoning about Change. Time and Causation from the Standpoint of Artificial Intelligence*, Cambridge, MA, MIT Press.
- Skarda, Ch. A. & Freeman, W. J.
1987 — „How Brains Make Chaos in Order to Make Sense of the World”, *The Behavioral and Brain Sciences* 10, s. 161-195.
- Stewart, I.
1994 — *Czy Bóg gra w kości? Nowa matematyka chaosu*, Warszawa, PWN.
- Winfrey, A. T.
1990 — *The Geometry of Biological Time*, Berlin, Springer.

⁷Składam gorące podziękowania Marcie Calvo, doktorantce na Wydziale Chemii UW, za wskazanie pewnych braków tego artykułu, które starałem się w miarę możliwości usunąć.