

*Anita Pacholik-Żuromska*

Uniwersytet Mikołaja Kopernika  
w Toruniu

Nicolaus Copernicus University  
in Toruń

## **KONEKCJONISTYCZNE MODELE WYJAŚNIANIA PROCESÓW POZNAWCZYCH W KOGNITYWISTYCE**

### **Connectionist Models of Explanation of Cognitive Processes in Cognitive Science**

Słowa kluczowe: konekcjonizm, funkcjonalizm, maszyna Turinga, sztuczne sieci neuronowe, Bayesowska teoria indukcji, enaktywizm.

Key words: connectionism, functionalism, Turing machine, artificial neural networks, Bayesian inference, enactivism.

#### Streszczenie

Celem artykułu jest przegląd i analiza modeli konekcjonistycznych na tle faz rozwoju kognitywistyki. Konekcjonizm, jako druga faza rozwoju kognitywistyki, zaoferował najlepsze narzędzia wyjaśniania i modelowania procesów poznawczych. Został on przedstawiony w relacji do wcześniejszej i późniejszej fazy rozwoju kognitywistyki. Wykazuje się tu również kompatybilność konekcjonizmu z enaktywizmem (trzecią fazą) na gruncie proponowanego modelu wyjaśniania, jak kształtuje się poznanie.

#### Abstract

The aim of this paper is an overview and analysis of the connectionist models on the basis of the milestones in the development of cognitive science. It is claimed that connectionism, as the second phase of cognitive science, offers the best tools of explanation and modelling of cognition. There is also indicated the compatibility of connectionism and enactivism (the third phase) on the basis of the proposed models of explanation.

## Spory o kognitywistykę

Kognitywistyka jest jeszcze stosunkowo młodą dyscypliną – na tyle, że problem stanowi sama jej definicja, co wynika przede wszystkim w jej niejednorodności. Kognitywistyka bowiem to nie jedna dyscyplina, a właściwie ich zbiór. Punkt sporny stanowi pytanie, czy zbiór ten rozpatrywać jako dystrybutywny, czy kolektywny. Czy zatem kognitywistyka to dyscyplina korzystająca z dorobku różnych nauk o poznaniu, czy też ich konglomerat<sup>1</sup>. Nie rozstrzygając tego, czy jest to nauka, dyscyplina, czy dziedzina o charakterze inter-, trans- czy multidyscyplinarnym, warto przytoczyć definicję Włodzisława Ducha głoszącą, że „zajmuje się ona wszystkimi zjawiskami dotyczącymi umysłu, szczególnie zagadnieniami dotyczącymi sposobu postrzegania bodźców i oddziaływania umysłu ze światem i innymi umysłami, a jej głównym zadaniem jest próba syntezy wiadomości z różnych dziedzin [...] i tworzenie modeli umysłu zgodnych z wiedzą dostępną z wszystkich źródeł”<sup>2</sup>.

Trafna wydaje się też konstatacja Jana Woleńskiego, że kognitywistyka to teoria poznawania w odróżnieniu od epistemologii, czyli teorii poznania<sup>3</sup>. Stwierdzenie to dało asumpt do dyskusji, czy kognitywistyka i epistemologia uzupełniają się, czy też kognitywistyka, jako związana z naukami ścisłymi, wyeliminuje potrzebę uprawiania epistemologii. Innymi słowy, pojawiło się pytanie, które formułuje m.in. Urszula Żegleń, czy możliwy jest mariaż kognitywistyki z epistemologią<sup>4</sup>. Badacze wciąż się wahają, jakiej odpowiedzi udzielić. Kolejne zjazdy Polskiego Towarzystwa Kognitywistycznego pokazują, że głosy propo-nentów i oponentów na temat kompatybilności kognitywistyki i epistemologii, czy – szerzej – filozofii, równoważą się. I być może właśnie ten stan jest najodpowiedniejszy.

Podane wyżej definicje kognitywistyki są dość ogólne – na tyle, by nie umieszczać jej w konkretnym nurcie badawczym. Mimo że kognitywistyka ma ok. 60–70 lat, to nurty te zdążyły już się wyodrębnić. Jedne przeminęły, inne ewoluowały, a jeszcze inne wciąż mocno się trzymają. W dość arbitralnym, ale też uzasadnionym podziale można wyodrębnić trzy fazy rozwoju kognitywistyki: komputacjonizm oparty

<sup>1</sup> M. Miłkowski, *Epistemologia znaturalizowana*, (w:) *Przewodnik po epistemologii*, red. R. Ziemińska, WAM, Kraków 2013, s. 511.

<sup>2</sup> W. Duch, *Czym jest kognitywistyka*, „Kognitywistyka i Media w Edukacji” 1998, nr 1, s. 9–50.

<sup>3</sup> J. Woleński, *Metamatematyka a epistemologia*, PWN, Warszawa 1993, s. 114–115.

<sup>4</sup> U. M. Żegleń, *Epistemologia a kognitywistyka*, (w:) *Przewodnik po epistemologii...*, s. 457–493.

na komputerowej metaforze umysłu i odwołujący się do maszyny Turinga jako modelu umysłu, konekcjonizm<sup>5</sup> zasadzający się, jak jego poprzednik na paradygmacie obliczeniowości, jednak promujący obliczanie równoległe, a nie szeregowe oraz enaktywizm głoszący, że w modelowaniu umysłu należy wziąć pod uwagę to, iż umysł i świat to jeden całościowy system. Podział ten jest arbitralny ze względu na uznaniowe przyjęcie liczby owych faz. Ktoś patrzący bardziej szczegółowo, mógłby powiedzieć, że tych faz jest więcej. Jednak uzasadnieniem dla owych mniej lub bardziej ogólnie potraktowanych kamieni milowych kognitywistyki jest to, że wiązały się (jak to zazwyczaj bywa) z postępem wyznaczonym przez rozwój nauki i techniki. W wypadku kognitywistyki rozwój ten wynikał z przyjęcia konkretnego paradygmatu, niekiedy nazywanego przy użyciu metafor nawiązujących do aktualnych fascynacji technologicznych.

## **Pierwsza faza kognitywistyki**

Pierwsza faza kognitywistyki nawiązywała do teorii automatów i pierwszych maszyn liczących. Flagowy tekst tego paradygmatu, zwanego obliczeniowym, to *Maszyna licząca a inteligencja* Alana Turinga<sup>6</sup>. Sam tekst jest późniejszy niż pierwsze wystąpienia matematyków i informatyków, w których dopiero formułowali oni tezy stanowiące podwaliny paradygmatu obliczeniowego. Generalnie rzecz biorąc działanie umysłu porównywane było w nim do działania programu komputerowego, stąd nierzadko mówiono o „komputerowej metaforze umysłu”. Przy czym należy wyraźnie zaznaczyć, że pierwsze komputery diame-

---

<sup>5</sup> Istnieją dwa sposoby zapisu tej nazwy: „konekcjonizm” lub „koneksjonizm” (ang. *connectionism*). Lingwistyczne argumenty przemawiają za zapisem „koneksjonizm”, ponieważ zachowuje się wtedy to samo znaczenie pochodne od słowa „koneksja” (ang. *connection*) i oddaje charakter nurtu, który podkreśla powiązanie procesów przetwarzania informacji. Taki zapis stosuje m.in. W. Duch i tak jest to tłumaczone w artykule D.E. Rumelhart, *Architektura umysłu. Podejście koneksyjne*, tłum. H. Grzegołoska-Klarkowska, (w:) *Modele umysłu*, red. Z. Chlewiński, Wyd. Nauk. PWN, Warszawa 1999. Jednak z drugiej strony „koneksjonizm” jest wieloznaczny, ponieważ odnosi się też do zasady koneksji wprowadzonej przez Johna Searle’a, mówiącej o związku świadomości z intencjonalnością – zob. J. R. Searle, *Świadomość, inwersja wyjaśnień i nauki kognitywne*, tłum. E.M. Hunca, (w:) *ibidem*. Żeby zatem odróżnić ten rodzaj koneksjonizmu od koneksjonizmu jako nurtu w kognitywistyce, niektórzy, np. J. Bremer, zapisują nazwę przez „c”. Obydwa zapisy są poprawne.

<sup>6</sup> A. Turing, *Maszyna licząca a inteligencja*, tłum. M. Szczubiałka, (w:) *Fragmenty filozofii analitycznej. Filozofia umysłu*, red. B. Chwedeńczuk, Spacja, Warszawa 1995.

tralnie różniły się od tego, z czym dziś mamy do czynienia. Przede wszystkim u ich podstaw leżały modele teoretyczne, które z czasem znalazły implementację na fizyczne struktury o mocy obliczeniowej tak niewielkiej, że każdy smartfon je przewyższa. Można powiedzieć, że modele teoretyczne wyprzedzały dalece możliwości techniczne. Jednak z czasem to, co prognozował Turing, stało się faktem, a nawet dalece przekroczyło jego oczekiwania. W fazę pierwszą wpisują się stanowiska funkcjonalistyczne Jerry'ego A. Fodora czy Roberta Sticha, a jej złote czasy przypadają na lata 80.

Definicja kognitywistyki wiążąca się z tym paradygmatem głosi, że „kognitywistyka jest interdyscyplinarną nauką, która bada umysł jako system przetwarzania informacji”<sup>7</sup>. Pojęcie umysłu jako systemu przetwarzania informacji jest tu oczywiście kluczowe. Zostanie ono przejęte i rozwinięte przez następne fazy rozwoju kognitywistyki. W tym momencie istotne jest tu zaznaczenie, na czym owo przetwarzanie informacji miałyby polegać.

Metafora komputerowa odwoływała się raczej do układów szeregowych przetwarzania danego sygnału, nawiązując do obwodów elektrycznych, w których koniec pierwszego łączy się z początkiem drugiego i nie rozdziela się. Prąd przepływa wtedy kolejno przez wszystkie następujące po sobie elementy – w uproszczeniu z punktu A przez B do C. W teoretycznych modelach automatów mowa już nie o przepływie prądu (choć ich implementacja na tym polega), ale o przepływie informacji. „Automaty są matematycznymi modelami urządzeń przetwarzających informację dyskretną. Przetwarzane informacji ma dwa zasadnicze zadania: rozpoznawanie [akceptowanie] języków oraz obliczanie funkcji częściowych ze zbioru  $\Sigma^*$  w  $\Delta^*$ , gdzie  $\Sigma$ ,  $\Delta$  są alfabetami skończonymi, a  $\Sigma^*$  w  $\Delta^*$  są zbiorami wszystkich słów, odpowiednio nad alfabetami  $\Sigma$ ,  $\Delta$ . W związku z tym wyróżnia się dwa typy automatów: automaty akceptujące [rozpoznające] oraz automaty przetwarzające”<sup>8</sup>.

Na koncepcji działania automatu Noam Chomsky oparł swoją koncepcję gramatyki generatywno-transformacyjnej, początkowo wyjaśniającej kompetencje językowe ludzi, z czasem jednak także to, jak działa umysł. Z koncepcji Chomskiego czerpał Fodor, twórca Reprezentacyjnej Teorii Umysłu<sup>9</sup>. Według niego, stany mentalne czy inaczej – posta-

<sup>7</sup> M. Urchs, *O procesorach i procesach myślowych*, tłum. A. Pacholik-Żuromska, Wyd. Nauk. UMK, Toruń 2009, s. 16.

<sup>8</sup> W. Marciszewski (red.), *Logika formalna. Zarys encyklopedyczny z zastosowaniem do lingwistyki i informatyki*, PWN, Warszawa 1987, s. 200.

<sup>9</sup> Zob. np. J.A. Fodor, *Jak grać w reprezentację umysłową – poradnik Fodora*, tłum. A. Putko, (w:) *Modele umysłu...*

wy propozycjonalne to relacje między jednostkami, sądami a sposobami przedstawiania: „sposoby przedstawiania są zdaniami w języku myśli [*mentalese*], a zdania są indywiduowane nie tylko przez treść zawartego w nich sądu, ale też przez składnię”<sup>10</sup>.

Stany mentalne to nic innego, jak dyspozycje do działania będące efektem przetworzonej informacji dostarczanej przez zmysły, której nośnikiem są reprezentacje. Umysł zatem jest rodzajem komputera<sup>11</sup>.

## Druga faza kognitywistyki

Na paradygmacie obliczeniowości zasadza się też konekcjonizm, o którym mówi się, że jest to „druga faza Sztucznej Inteligencji”<sup>12</sup>, ale nie tylko. To również druga faza kognitywistyki. Konekcjonizm to stanowisko, w którym modele architektury umysłu opierają się na równoległym przetwarzaniu informacji. „Ten sposób przetwarzania danych – przekształcanie jednego wzorca w inny podczas ich przechodzenia przez liczne skonfigurowane połączenia synaptyczne – jest nazywany *przetwarzaniem równoległym* (PDP – *Parallel Distributed Processing*). Występuje ono powszechnie w świecie zwierząt, i to z dobrze uzasadnionych powodów. Ma olbrzymią przewagę nad znanym z techniki komputerowej procesem przetwarzania szeregowego [...]”<sup>13</sup>.

Konekcjonizmowi towarzyszy już nie metafora komputerowa, ale metafora sieci, w związku z tym, że przetwarzanie informacji ma tu charakter dystrybutywny. Modele konekcjonistyczne powstały jako odpowiedź na pytanie, jak mózg tworzy umysł. Nie było to już zatem pytanie, jak działa umysł, ale raczej jak działa mózg. W budowaniu modeli przetwarzania informacji przez mózgi sieć jest najodpowiedniejszą strukturą odwzorowującą jego pracę, ponieważ „prawie każda część mózgu jest obserwowana przez inną, a często przez wiele innych narządów”<sup>14</sup>. Przetwarzanie informacji odbywa się tu w ten sposób, że wzorce aktywacji neuronów w narządach zmysłowych monitorowane są przez

---

<sup>10</sup> J.A. Fodor, *Ekspersi od wiązków. Język myśleński i jego semantyka*, tłum. M. Gokiel, Aletheia, Warszawa 2001, s. 73.

<sup>11</sup> Por. ibidem, s. 20. Wydaje się, że poprawniejsze byłoby porównanie umysłu do programu komputerowego, a nie hardware’u, za który można by było uznać mózg, będący strukturą, na której realizowany jest program, czyli język myśli wg koncepcji Fodora.

<sup>12</sup> M. Urchs, op. cit., s. 101.

<sup>13</sup> P.M. Churchland, *Mechanizm rozumu, siedlisko duszy. Filozoficzna podróż w głąb mózgu*, tłum. Z. Karaś, Aletheia, Warszawa 2002, s. 20.

<sup>14</sup> Ibidem, s. 17.

kolejne warstwy neuronów, przekształcające daną informację i przekazujące ją do kolejnych struktur w mózgu. Informacja zakodowana w mózgu to wzorzec aktywacji neuronalnej, który może być wzmacniany przez trening poznawczy. Przekazywanie informacji między neuronami polega na przemieszczaniu się neuroprzekaźników do odpowiednich receptorów w synapsach, co powoduje w zależności od neuroprzekaźnika zwiększenie bądź zmniejszenie polaryzacji błony postsynaptycznej. Przy odpowiednim natężeniu hiperpolaryzacji błony powstaje impuls. Należy zaznaczyć, że „komórka otrzymuje w każdej chwili wiele sygnałów od innych komórek zarówno pobudzających jak i hamujących”<sup>15</sup>. Ta wielość w pewnym sensie konkurujących ze sobą sygnałów przestaje dziwić, jeśli się spojrzy na liczbę połączeń synaptycznych w mózgu. To ponad sto trylionów! Siła i konfiguracja tych połączeń indywidualizuje mózg i może się zmieniać na przykład dzięki uczeniu się.

## **Pierwsza faza kontra druga faza kognitywistyki**

Konekcyjnistyczne podejście do umysłu jako superwenującego na fizycznej strukturze mózgu zaowocowało programem badawczym polegającym na budowie sztucznych sieci neuronowych. „Układ nerwowy posługuje się kombinatorycznym systemem reprezentacji, który pozwala na drobiazgową analizę każdej z odczuwalnych zmysłowych subtelności. To pozwala rozróżniać i rozpoznawać o wiele więcej, niż jesteśmy w stanie wyrazić słowami”<sup>16</sup>. Nie dziwi zatem rozwój teorii od automatów – choć skomplikowanych to wciąż przetwarzających dane analogowo – do automatów komórkowych. Jeśli jednak ktoś sądzi, że to stonkowo świeży pomysł, to się myli! Maszyna Turinga, czyli ogólny model automatu, został wprowadzony w 1936 r., automat komórkowy zaś w roku 1948. Jego autorem był John von Neuman. Różnice między obydwoma rodzajami automatów najlepiej oddadzą poniższe opisy. Pierwszy z nich dotyczy maszyny Turinga: „Pamięć nieskończona maszyny Turinga składa się z nieskończonej taśmy podzielonej na komórki oraz głowicy czytająco-zapisującej. W każdej komórce taśmy pamięci nieskończonej może być zapisany tylko jeden symbol z alfabetu tej taśmy. W trakcie jednego ruchu, w zależności od symboli obserwowanych na taśmach i stanu sterowania, maszyna Turinga może zmienić stan

---

<sup>15</sup> P. Jaśkowski, *Neuronauka poznawcza. Jak mózg tworzy umysł*, Vizja Press&IT, Warszawa 2009, s. 25.

<sup>16</sup> P.M. Churchland, op. cit., s. 31.

sterowania, zmienić zawartość komórki obserwowanej na taśmie pamięci nieskończonej oraz przesunąć głowicę o jedną komórkę w prawo lub lewo. To, jaki ruch wykona maszyna, zależy od jej sterowania”<sup>17</sup>.

Maszyna Turinga działa zatem na zasadzie algorytmu. Zawiera dane wejściowe, reguły ich przekształceń i dane wyjściowe (wynik operacji arytmetycznej). Jednak o ile maszyny Turinga bardzo dobrze spisywały się np. w deszyfracji komunikatów językowych, to zawodziły w rozpoznawaniu bardziej skomplikowanych wzorców. W tym wypadku trzeba było nie jednego, ale wielu „synchronicznie współdziałających”<sup>18</sup> automatów tworzących sieć. Takim przykładem jest sieć Petriego – formalny model systemów rozproszonych służących do wykonywania obliczeń równoległych, czy – używając języka konekcjonizmu – równoległego przetwarzania informacji. I to właśnie przetwarzanie równoległe zostało wykorzystane w konekcjonistycznych modelach w postaci sztucznych sieci neuronowych próbujących w jakimś stopniu symulować pracę mózgu.

Na kandydata do odwzorowania wybrano rozpoznawanie wzorców, jak np. twarze. Zadano sobie pytanie, czy sztuczna sieć neuronowa, polegająca na przeniesieniu modelu pracy mózgu na formułę programu realizowanego przez komputer, będzie w stanie nauczyć się rozpoznawania twarzy, jak to czynią ludzie. W tym celu zaczęto od analizy tego, jak widzi ludzki mózg, a następnie podjęto próbę algorytmizacji procesów percepcyjnych i zaimplementowania ich na sztuczne sieci neuronowe. Problemem była tu translacja widzenia stereoskopowego, czyli różnicy między obrazami przekazywanymi przez lewe i prawe oko. Dzięki tej zdolności potrafimy widzieć trzeci wymiar, czyli rozpoznawać głębię. Obrazy przekazywane z obu oczu są synchronizowane w korze wzrokowej. Jak jednak tę neuroanatomie i neurofizjologię widzenia przełożyć na sztuczny system? Czy będzie on potrafił wykryć głębię? Rozpoznać, które obiekty są bliżej, a które dalej? A w wypadku rozpoznawania twarzy, czy system będzie w stanie rozpoznać, że postrzegany obiekt to właśnie twarz, a nie tylko coś, co twarz przypomina?

Zarówno z rozpoznawaniem głębi, jak i twarzy dawno sobie poradzono. W wypadku głębi efekty mamy choćby w postaci filmów i gier w 3D. W wypadku rozpoznawania twarzy sieci, już nie tylko potrafią je identyfikować, ale nawet rozpoznawać emocje. Generalnie dziś już możemy przekładać na algorytmy wiele rodzajów doznań, choć brzmi to podejrzanie, ponieważ doznania mają zawsze komponent subiektywny.

---

<sup>17</sup> W. Marciszewski, op. cit., s. 203

<sup>18</sup> Ibidem, s. 208.

Niemniej jednak, jeśli system zawiera na wejściu odpowiednie dane, będzie w stanie wyprowadzić z nich odpowiedni wniosek, co to jest. Na przykład jedną z metod rozpoznawania głębi w obrazie, jest śledzenie promieni na podstawie odpowiednio zacienionego obrazu. Algorytm tu zastosowany pozwala na obliczenie stopnia zacienienia i wydobycie z otoczenia tych punktów, które są widoczne z każdego miejsca otoczenia. Oczywiście, te najbardziej widoczne mają wartość 1, te niewidoczne 0<sup>19</sup>.

Algorytmy stosowane w modelach percepcji pozwalają grupować podobne „doznania”, czyli cechy i obiekty, blisko siebie, a różne – daleko, tworząc „systematyczną przestrzeń różnic i podobieństw”<sup>20</sup>, kodowaną wektorowo. Poziomy aktywacji neuronów odpowiedzialnych za daną percepcję, np. barwy, układają się w specyficzny wzorec, który można zapisać właśnie w kombinatoryce kodowania wektorowego. „Siatkówka oka wyposażona jest w trzy różne rodzaje światłoczułych komórek nerwowych nazywanych z powodu kształtu czopkami. Każdy z czopków odbiera tylko jedną z trzech określonych długości fal świetlnych. Te światłoczułe czopki projektują wspólnie swoje poziomy wzbudzenia na inne populacje neuronów, także składające się z komórek trzech różnych typów. Owe leżące dalej komórki nerwowe organizują przestrzeń barw, które rzeczywiście widzimy. Tym razem jest to przestrzeń trójwymiarowa, jeden wymiar dla każdego z trzech typów komórek [...]. Tak więc każdy dostrzegany przez człowieka kolor będzie unikalnym rozkładem poziomów aktywacji neuronów w tych trzech typach procesów rywalizacji”<sup>21</sup>. Rywalizacja może nie jest tu dobrym słowem. Chodzi o siłę pobudzeń trzech typów neuronów, która decyduje o sile sygnałów wyjściowych<sup>22</sup>. Wzorce pobudzeń kodowane są właśnie w postaci wektorowej (kwantyfikacja wektorowa) na podstawie przypisanych wartości cech (np. długość fali świetlnej). Konkretna barwa będzie zatem modelowo (w sztucznych sieciach neuronowych) przedstawiana jako punkt w przestrzeni cech.

Dotychczasowe rozważania mają na celu ukazanie, że konekcyonizm zaoferował zarówno badaniom nad Sztuczną Inteligencją, jak

---

<sup>19</sup> Zob. na przykład P. Dornbach, *Implementation of Bidirectional Ray Tracing Algorithm*, [online] <<http://old.cescg.org/CESCG98/PDornbach/index.html>> (dostęp: 25.07.2017). Dziękuję Remigiuszowi Iwańkowi za wykład na temat algorytmów wykorzystywanych w rozpoznawaniu obrazów oraz polecenie lektur.

<sup>20</sup> P.M. Churchland, op. cit., s. 33.

<sup>21</sup> Ibidem, s. 34–35.

<sup>22</sup> Zob. np. R. Tadeusiewicz, *Sztuczne sieci neuronowe*, (w:) *Przewodnik po kognitywistyce*, red. J. Bremer, Wyd. WAM, Kraków 2016.



i kognitywistycznym modelom umysłu więcej niż pierwsza faza kognitywistyki odwołująca się do maszyny Turinga i szeregowego przetwarzania informacji. Na mikropoziomie pozwalał on modelować np. percepcję wzrokową, zaś na makropoziomie wyjaśniał, w jaki sposób owa percepcja wpływa na zachowanie i jak mózg tworzy reprezentację świata – o czym dopiero będzie mowa. Zanim jednak to nastąpi, należy przedstawić (przynajmniej pokrótce) trzecią fazę kognitywistyki, nie tylko dlatego, by dopełnić obraz, ale też dlatego, że ma ona wiele wspólnego z fazą drugą.

### Trzecia faza kognitywistyki

Enaktywizm to nurt, który na przełomie XX i XXI w. zdominował dyskusję kognitywistyczną. Występuje właściwie jako *pars pro toto* wszystkich e-teorii, choć nie można ich ze sobą utożsamiać. Do e-teorii zalicza się enaktywizm (*enactivism*), teorie ucieleśnionego umysłu (*embodied mind*) oraz teorie rozszerzonego umysłu (*extended mind*). Enaktywizm twierdzi, że poznanie konstytuuje się poprzez działanie podmiotu w świecie; teorie ucieleśnionego umysłu głoszą, że ciało podmiotu odgrywa istotną rolę w poznaniu, a teorie rozszerzonego umysłu, że umysł jest nie tylko w ciele, ale też poza nim – jego zawartość jest np. magazynowana w powszechnie dziś wykorzystywanych urządzeniach, jak laptopy czy smartfony. Teorie te, choć różne w szczegółach, zgadzają się co do ogólnego twierdzenia, że podmiot i jego otoczenie zewnętrzne tworzą jeden zwarty system, a adekwatne modele poznania to takie, które uwzględniają środowisko zewnętrzne jako komponent poznania<sup>23</sup>. Ten rodzaj poznania zwany jest e-poznaniem, ponieważ zgodnie z e-teoriami wymaga ono interakcji podmiotu ze środowiskiem zewnętrznym. Poznanie odbywa się poprzez działanie i dla działania, stąd rola takich zmysłów jak propriocepcja i uwaga, odgrywających ważną rolę w tworzeniu precyzyjnego ruchu. W e-poznaniu w ogóle ważne jest znaczenie procesów jako łączników pomiędzy mózgiem a zewnętrznym światem. W tym ujęciu organizmy nie są zwykłymi odbiornikami sygnałów z zewnątrz, są również ich modyfikatorami.

Głosząc e-poznanie, niekoniecznie trzeba być enaktywistą. Niezależnie od szczegółowej teorii e-poznanie powinno implikować eksternalizm i móc się obronić przed kartezjanizmem w postaci mózgow w na-

---

<sup>23</sup> M. Wilson, *The Six Views of Embodied Cognition*, „Psychonomic Bulletin & Review” 2002, nr 9(4), s. 625–636.

czynniu<sup>24</sup>. A zatem wszelkie koncepcje na temat roli uwagi czy propriocepcji w poznaniu, które to procesy łączą podmiot ze światem, będą również podpadać pod e-teorie, jeśli faktycznie wykazują konieczność i efekty interakcji podmiotu ze światem, jak choćby koncepcja powstawania jaźni według Kaliny Christoff<sup>25</sup>.

Enaktywizm wpłynął także na cybernetykę, ponieważ inżynierowie poszukiwali algorytmów pokazujących celowość działania<sup>26</sup>. Problem polegał jednak na tym, jak pogodzić silnie reprezentacjonistyczny konekcyjizm z enaktywizmem odrzucającym tradycyjne podejście do reprezentacji jako nośnika informacji przetwarzanej obliczeniowo.

## Druga faza kontra trzecia faza kognitywistyki

Na pierwszy rzut oka enaktywizm stanowił odwrót od konekcyjizmu ze względu na to, że modele konekcyjistyczne nie uwzględniały środowiska jako istotnego komponentu poznania, ponieważ miało ono charakter zewnętrzny, a modele umysłu oparte na sztucznych sieciach neuronowych dotyczyły przecież wewnętrznych procesów przetwarzania informacji: „Teorie enaktywne mogą się zgadzać co do tego, że zachowania adaptacyjne wymagają tego, by organizmy i struktury kontroli były systematycznie wrażliwe na statystyczne struktury w ich środowisku, ale większość będzie zaprzeczać, że owa wrażliwość implikuje istnienie zastosowania jakiejś wewnętrznej deskrypcji, czy modelu tych probabilistycznych wzorców”<sup>27</sup>. Model probabilistyczny wzorców polega na kodowaniu predykcijnym przyszłych stanów świata na podstawie danych ze świata, a następnie sprawdzaniu stawianych hipotez i ich zmian w wypadku, gdy hipoteza nie pasuje do danego stanu świata. Owe hipotezy można interpretować jako reprezentacje<sup>28</sup>. W kodowaniu predykcijnym reprezentacja jest zatem modelem praw-

---

<sup>24</sup> Nawiązuje się tu do eksperymentu myślowego H. Putnama z artykułu *Mózgi w naczyniu*, (w:) *Wiele twarzy realizmu i inne eseje*, tłum. A. Grobler, PWN, Warszawa 1998.

<sup>25</sup> K. Christoff, D. Cosmeli, D. Legrand, E. Thompson, *Specifying the Self for Cognitive Neuroscience*, „Trends in Cognitive Science” 2011, nr 3(15), doi:10.1016/j.tics2011.01.001.

<sup>26</sup> A. K. Seth, *The Cybernetic Bayesian Brain – From Interoceptive Inference to Sensorimotor Contingencies*, (w:) *Open MIND*, red. T. Metzinger, J. M. Windt, MIT Press 2015, s. 1457.

<sup>27</sup> Ibidem, s. 1466.

<sup>28</sup> J. Hohwy, *Attention and Conscious Percepton in the Hypothesis Testing Brain*, „Frontiers in Psychology” 2012, t. 3, art. 96.

dopodobnego (oczekiwanego) stanu świata. Kodowanie predykcyjne podpada zatem pod konekjonistyczne modele, które również dawały możliwość przewidywania stanów świata, czy zachowań podmiotów<sup>29</sup>.

Kodowanie predykcyjne ewidentnie klóci się z tezami e-poznania. Zgodnie z założeniami tego drugiego, w najbardziej radykalnej formie utworzonej po zebraniu tez z poszczególnych e-teorii, nie tylko poznanie ma być rozszerzone, nie tylko znaczenia, nie tylko procesy poznawcze, ale też reprezentacje. Znaczenia nie są w głowach i reprezentacje również, ponieważ bodziec ze świata jest elementem reprezentacji mentalnej, a nie tylko jej aktywatorem.

Odmienne podejście do reprezentacji przedstawia kodowanie predykcyjne. One nie powstają. One już są. Najwyżej ulegają modyfikacji pod wpływem kontaktu z mniej oczekiwanyimi bodźcami. Takie podejście do poznania stanowi jednak krok wstecz. Implikuje bowiem wrodzonność wszystkich reprezentacji oraz internalizm. Jedyne co w nim atrakcyjne, to model obliczeniowy oparty na Bayesowskiej teorii indukcji.

„Twierdzenie Bayesa określa tzw. prawdopodobieństwo *a posteriori* [...] mianowicie prawdopodobieństwo hipotezy *s* przy założeniu danych *e*”<sup>30</sup>. Ma ono postać:

$$p(s|e) = p(e|s) \cdot p(s) / p(e) \cdot$$

Odwrotność stanowi prawdopodobieństwo *a priori*. Dla porównania warto mu się przyjrzeć bliżej. „Zakłada się, że dokładnie jedna [z tych hipotez] jest prawdziwa. Ich początkowymi prawdopodobieństwami są  $p(s)$  dla  $s \in S$ . Zdanie *e* interpretujemy jako sprawozdanie z wyniku eksperymentu, który miał pomóc w rozstrzygnięciu między hipotezami należącymi do zbioru *S*. Znane jest prawdopodobieństwo *e* przy założeniu prawdziwości każdej z hipotez:  $p(e|s)$ , dla  $s \in S$ . Dla ustalonego *e* prawdopodobieństwo to jest funkcją *s*; nazywa się je zazwyczaj wiarygodnością *s* ze względu na *e*”<sup>31</sup>.

Co zatem atrakcyjniejszego niż prawdopodobieństwo *a priori* oferuje prawdopodobieństwo *a posteriori*? Otóż zgadza się ono z potoczną intuicją, że „wynik obserwacji potwierdza hipotezę wówczas, gdy jest wysoce prawdopodobny na jej gruncie, a poza tym mało prawdopodobny”<sup>32</sup>. Bayesowski mózg będzie zatem działał w ten sposób, że doświadczenie (*feedback* ze świata) będzie konfrontował z uprzednio postawioną hipotezą na temat spodziewanej sytuacji.

<sup>29</sup> Por. Tadeusiewicz, op. cit., s. 517.

<sup>30</sup> W. Marciszewski, op. cit., s. 411.

<sup>31</sup> Ibidem.

<sup>32</sup> Ibidem, s. 412.

Pytanie, czy coś z kodowania predykcyjnego da się ocalić, nie hamując rozwoju kognitywistycznych teorii. Istnieją już owocne próby powiązania go z enaktywizmem. Czyni to np. Anil Seth. Michael Anderson również uważa, że e-poznanie i kodowanie predykcyjne są do pogodzenia<sup>33</sup>. Następująca propozycja również wpisuje się ten pożądany kompatybilizm: Jeśli nieoczekiwany bodziec zewnętrzny zmienia hipotezę (reprezentację) danego elementu świata, to nie dlatego, że ma on jakieś wewnętrzne znaczenie, ale dlatego, że to znaczenie ustalone jest przez kontekst – daną sytuację zewnętrzną, jednak której element stanowi dany organizm. A zatem kontekst to świat plus podmiot, co zgodne jest z tezą zbiorczą e-teorii, że podmiot i świat tworzą zunifikowany system. Kruchy rozejm między enaktywizmem i kodowaniem predykcyjnym w ten sposób może zostać ocalony.

## Podsumowanie

Konekcyjnistyczne modele wyjaśniania procesów poznawczych, takich jak percepcja, opierają się na algorytmach tworzących sztuczne sieci neuronowe. Liczne modele tego typu mają bardzo szerokie zastosowania. Kodowanie predykcyjne wpisuje się również w model konekcyjnistyczny i jest w tym całkiem obiecujące. Dane empiryczne zdają się potwierdzać, że mózg jest swoistą machiną predykcyjną. Jednak od modelu pracy mózgu do modelu poznania daleka droga. Kodowanie predykcyjne nie mówi bowiem nic o tym, jak powstaje efekt poznawczy w postaci stanu mentalnego determinującego nasze wybory, przekonania i działania. Dopiero jego połączenie z tezami e-teorii daje obiecujące efekty, ponieważ zwraca podmiot światu, a nie od niego izoluje. Kolejna faza kognitywistyki utrzymana w tym duchu powinna odpowiedzieć na pytanie, jak możliwe jest dzielenie reprezentacji.

## Literatura

Anderson M. L., *How to Study the Mind: An Introduction to Embodied Cognition*, (w:) *Brain Development in Learning Environments: Embodied and Perceptual Advancements*, red. F. Santoianni, C. Sabatano, Cambridge Scholars Publishing, Newcastle 2007.

---

<sup>33</sup> M. L. Anderson, *How to Study the Mind: An Introduction to Embodied Cognition*, (w:) *Brain Development in Learning Environments: Embodied and Perceptual Advancements*, red. F. Santoianni, C. Sabatano, Cambridge Scholars Publishing, Newcastle 2007.

- Christoff K., Cosmelli D., Legrand D. Thompson E., *Specifying the Self for Cognitive Neuroscience*, „Trends in Cognitive Science”, March 2011, nr 3(15), doi:10.1016/j.tics2011.01.001.
- Churchland P.M., *Mechanizm rozumu, siedlisko duszy. Filozoficzna podróż w głąb mózgu*, tłum. Z. Karaś, Aletheia, Warszawa 2002.
- Dornbach P., *Implementation of bidirectional ray tracing algorithm*, [online] <<http://old.cescg.org/CESCG98/PDornbach/index.html>>.
- Duch, W., *Czym jest kognitywistyka*, „Kognitywistyka i Media w Edukacji” 1998, nr 1, s. 9–50.
- Fodor J.A., *Eksperci od wiązków. Język myślenia i jego semantyka*, tłum. M. Gokieli, Aletheia, Warszawa 2001.
- Fodor J.A., *Jak grać w reprezentacje umysłowe – poradnik Fodora*, tłum. A. Putko, (w:) *Modele umysłu*, red. Z. Chlewiński, PWN, Warszawa 1999.
- Hohwy J., *Attention and Conscious Perceptron in the Hypothesis Testing Brain*, „Frontiers in Psychology” April 2012, t. 3, art. 96.
- Jaśkowski P., *Neuronauka poznawcza. Jak mózg tworzy umysł*, Vizja Press&IT, Warszawa 2009.
- Marciszewski W. (red.), *Logika formalna. Zarys encyklopedyczny z zastosowaniem do lingwistyki i informatyki*, PWN, Warszawa 1987.
- Miłkowski M., *Epistemologia znaturalizowana*, (w:) *Przewodnik po epistemologii*, red. R. Ziemińska, WAM, Kraków 2013.
- Putnam H., *Mózgi w naczyniu*, (w:) *Wiele twarzy realizmu i inne eseje*, tłum. A. Grobler, PWN, Warszawa 1998.
- Rumelhart D.E., *Architektura umysłu. Podejście konekcyjne*, tłum. H. Grzegółowska-Klarkowska, (w:) *Modele umysłu*, red. Z. Chlewiński, Wyd. Nauk. PWN, Warszawa 1999.
- Searle J. R., *Świadomość, inwersja wyjaśnień i nauki kognitywne*, tłum., E.M. Hunca, (w:) *Modele umysłu*, red. Z. Chlewiński, PWN, Warszawa 1999.
- Seth A.K., *The Cybernetic Bayesian Brain – From Interoceptive Inference to Sensorimotor Contingencies*, (w:) *Open MIND*, red. T. Metzinger, J. M. Windt, MIT Press 2015.
- Tadeusiewicz R., *Sztuczne sieci neuronowe*, (w:) *Przewodnik po kognitywistyce*, red. J. Bremer, WAM, Kraków 2016.
- Turing A., *Maszyna licząca a inteligencja*, tłum. M. Szczubiałka, (w:) *Fragmenty filozofii analitycznej. Filozofia umysłu*, red. B. Chwedeńczuk, Spacja, Warszawa 1995.
- Urchs M., *O procesorach i procesach myślowych*, tłum. A. Pacholik-Żuromska, Wyd. Nauk. UMK, Toruń 2009.
- Wilson M., *The Six Views of Embodied Cognition*, „Psychonomic Bulletin & Review” 2002, nr 9(4) s. 625–636.
- Woleński J., *Metamatematyka a epistemologia*, PWN, Warszawa 1993.
- Zegleń U. M., *Epistemologia a kognitywistyka*, (w:) *Przewodnik po epistemologii*, red. R. Ziemińska, WAM, Kraków 2013.

