

Marek W. Bielecki

Rozumienie sieci neuronowych

Wstęp

Zdecydowana większość naukowców jest przekonana, że procesy kognitywne cechujące ludzi i zwierzęta są bezpośrednim rezultatem aktywności złożonego układu nerwowego. Wielu filozofów podziela to przekonanie i sądzi, że wszelkie rozważania na temat natury ludzkiej, w tym przede wszystkim zdolności myślenia i rozumienia, odwoływać się powinny wprost do biologicznych własności ludzkiego organizmu. Wyrażają też pogląd, że sama filozofia, jeśli nie chce całkowicie utracić intelektualnego autorytetu i społecznego znaczenia, uwzględniać powinna wyniki lub wręcz przyjąć metody nauk szczegółowych, przede wszystkim neurofizjologii.

Przeświadczenie to znajduje oparcie zarówno w fakcie istnienia korelacji pomiędzy aktywnością układu nerwowego i obserwowalnym zachowaniem organizmu, jak i w znacznych sukcesach w modelowaniu zdolności kognitywnych, przede wszystkim percepcyjnych, za pomocą sztucznych sieci neuronowych, czyli tzw. „ANNs” (z ang.: „Artificial Neural Networks”).

Wielu filozofów, obezwładnionych wizją człowieka jako maszyny biologicznej i nie mających dosyć odwagi, aby w wieku zdominowanym przez naukową wizję świata opowiedzieć się za idealizmem, podąża zwykle za Kartezjuszem i przyjmuje istnienie, różnego od mózgu, niematerialnego umysłu. Poza argumentami natury metafizycznej na rzecz istnienia takiego bytu, filozofowie tacy nierzadko wysuwają też i argumenty natury praktycznej, twierdząc między innymi, że całego szeregu ludzkich zdolności umysłowych, w tym na przykład — jak sądził sam Kartezjusz — zdolności posługiwania się językiem naturalnym, nie tylko nie da się modelować, ale nawet efektywnie symulować przez najbardziej skomplikowane maszyny.

Z drugiej strony, entuzjaści postępu naukowo-technicznego wskazują, jak postęp ten obala jedno po drugim twierdzenia filozofów — ostatnio, na przykład, twierdzenie, że komputer nigdy nie będzie w stanie regularnie wygrywać w sza-

chy z najlepszymi ludzkimi szachistami¹ — i przewidują, że w niezbyt odległej przyszłości powstaną androidalne maszyny wyposażone w sztuczne sieci neuronowe, dostatecznie skomplikowane, by osiągnąć, a w końcu przekroczyć, poziom ludzkich zdolności kognitywnych. Powstałoby wtedy niebezpieczeństwo, że takie „superistoty” odzęgają się od swoich mniej rozumnych przodków.

Pojawia się przeto naturalne pytanie, czy podobny los nie spotka też i innych twierdzeń filozofów, w tym i Kartezjusza, na temat niezdolności modelowania przez maszyny złożonych ludzkich procesów myślowych i ich rezultatów, takich jak rozumienie czy posługiwanie się językiem naturalnym.

W niniejszym artykule koncentruję się na scharakteryzowaniu „kognitywnych zdolności” ANNs, rozważeniu kwestii, na ile sieć taka jest adekwatnym modelem systemu nerwowego, na czym polega istotna różnica pomiędzy funkcjonowaniem umysłu i mózgu i czy w związku z tym sensowne jest uznanie, że ANNs, układ nerwowy, a w szczególności „mózg”, mogą cokolwiek „zrozumieć”².

I. Sztuczne sieci neuronowe (ANNs)

W ciągu ostatnich kilkunastu lat, w badaniach modelowania procesów kognitywnych nastąpiła radykalna zmiana wynikająca ze zdominowania tych badań przez nowy paradygmat zwany „konekcjonizmem”. U podstaw tego paradygmatu leży założenie naukowe, zgodne ze zdrowym rozsądkiem, że wszelkie organizmy żywe rozpoznają występujące w świecie regularności dzięki zdolności rejestrowania i przetwarzania informacji zawartej w tych regularnościach. W ramach tego paradygmatu uznaje się, że w organizmach wyposażonych w system nerwowy informacja „przechowywana” jest w postaci aktywacji określonych grup neuronów, a jej przetwarzanie polega na niealgoritmicznym procesie modyfikacji rozkładu tych aktywacji. Źródłem myślenia i wszelkich zdolności kognitywnych jest proces wzajemnego oddziaływania wielu połączonych ze sobą neurono-podobnych procesorów³. Ponieważ fizyczna natura tych procesorów nie ma znaczenia dla istoty samego procesu, więc uznaje się, że sztuczne sieci neuronowe, tj. sieci zbudowane z neurono-podobnych procesorów, powinny być w stanie modelować kognitywne zdolności układu nerwowego z dowolną dokładnością.

¹ Por. Dreyfus [1979].

² T. Winograd i F. Flores [1986], odwołując się do hermeneutyki i teorii systemów autopoietycznych, dyskutują obszernie naturę ludzkiego rozumienia, oceniając przy tym negatywnie możliwości jego komputerowego modelowania. Uznają oni jednocześnie, że człowiek jest jedynie układem biologicznym, ale omijają kwestię, czy złożone sieci neuronowe mogą wykazywać podobne do ludzkich zdolności kognitywne.

³ Autorzy tekstu, który stał się manifestem konekcjonizmu stwierdzają: „[...] intelligence emerges from the interactions of large numbers of simple processing units.” D.E. Rumelhart, G. Hinton, & J.L. McClelland, *A General Framework for Parallel Distributed Processing*, w: Rumelhart, McClelland [1986], tom 1, s. IX.

Typowa ANNs składa się zazwyczaj z kilkudziesięciu neuronopodobnych jednostek (węzłów sieci), których działanie polega na zbieraniu przychodzącego sygnału i, gdy intensywność wzbudzenia danej jednostki przekroczy jej próg aktywacji, przekazaniu go do wszystkich jednostek, z którymi jest połączona. Większość jednostek powiązana jest wzajemnie za pomocą jednokierunkowych połączeń o zmiennej wadze, reprezentowanej matematycznie przez liczby z przedziału $[-1,1]$. Nowsze rodzaje ANNs różnią się od swych poprzedniczek tym, że zawierają również połączenia wsteczne umożliwiające pojawienie się sprzężeń zwrotnych. Sieci takie określa się niekiedy mianem „rekurencyjnych”.

W typowej sieci rekurencyjnej jednostki ukonstytuowane są na trzech poziomach: wejścia, poziomu niejawnego i wyjścia. Jednostki wejściowe, które aktywowane są w zależności od siły i innych cech sygnału wejściowego, przekazują odpowiednią intensywność tej aktywacji do jednostek poziomu niejawnego. Intensywność ta jest proporcjonalna do poziomu wzbudzenia każdej jednostki wejściowej i siły jej połączeń z odpowiednią jednostką poziomu niejawnego. Jednostki poziomu niejawnego natomiast, otrzymują sygnał nie tylko od jednostek wejściowych i od innych jednostek z tego samego poziomu, ale ponadto wstecznie, w formie sprzężenia zwrotnego od jednostek z poziomu wyjścia. Ostatecznie, aktywacja jednostek poziomu wyjścia zależy od intensywności sygnału przychodzącego od jednostek niższego (niejawnego) poziomu i innych jednostek wyjścia. Sprzężenie zwrotne pomiędzy poziomem wyjścia i poziomem niejawnym umożliwia sieci generowanie wewnętrznych dynamik typu oscylacyjnego, których zachowanie nie jest bezpośrednio zdeterminowane przez sygnał wejściowy i które mogą istnieć nawet przy jego braku.

Rozkład aktywacji jednostek poziomu wyjścia odpowiada informacji „wyjściowej”, która reprezentuje reakcję sieci na „bodziec”. Chwilowy rozkład aktywacji wszystkich jednostek nazywa się „konfiguracją sieci” i uważa zwykle za odpowiadający temu, co sieć „wie” w danej chwili. Natomiast rozkład współczynników wagowych uznaje się za reprezentujący „pamięć długoterminową”. Sądzi się również, że procesy rekurencyjne umożliwiają występowanie w sieci „pamięci krótkoterminowej”, bo informacja przetworzona przez jednostki poziomu niejawnego powraca do nich w chwilę później w zmodyfikowanej formie do kolejnego przetworzenia⁴.

Wśród wielu modeli sztucznych sieci neuronowych dwa ujęcia uważa się za szczególnie trafnie opisujące sposób funkcjonowania biologicznych systemów nerwowych, albowiem oba oparte są na wynikach badań neurofizjologicznych. Jednym z nich jest model zaproponowany przez G. Edelmanna [1982], [1992], zaś drugim model rozwijany przez S. Grossberga i jego współpracowników [1982], [1988].

⁴ Jak pisze Paul Churchland, „This allows the creature to present its current situation in a way that takes into account the situation that immediately preceded it”. (Por. Churchland [1995], s. 100).

Edelman formułuje teorię „wyboru grup neuronowych”, w skrócie — „TNGS” (z ang.: „Theory of Neuronal Group Selection”), której podstawowym założeniem jest to, że procesy kognitywne, na przykład procesy percepcyjne, są rezultatem wzajemnego oddziaływania i koordynacji aktywności grup neuronów w mózgu⁵. Grupy takie, zawierające od 50 do 10 000 neuronów, tworzą swego rodzaju „mapy”, które konstytuują się w procesach zarówno rozwoju anatomicznego, jak i wzmacniania lub osłabiania połączeń synaptycznych. Procesy te są częściowo kontrolowane przez program genetyczny, a częściowo zależą od historii oddziaływań organizmu z otoczeniem. Nie ma też w mózgu żadnej centralnej „mapy nadzorczej”, która sumowałaby wszystkie cząstkowe „informacje” i sterowała zachowaniem organizmu. Znajduje się w nim wszakże struktura dynamiczna wyższego niż poszczególne mapy rzędu, określana przez Edelmana mianem *global mapping*. Składa się ona z wielu lokalnych map i połączeń z takimi częściami mózgu, jak hipokamp czy mózdzek, w których mapy nie występują, ale które umożliwiają sensoryczno-motoryczną aktywność organizmu.

Edelman twierdzi ponadto, że organizm „kategoryzuje” swe otoczenie zawsze ze względu na wewnętrzny „system wartości”, którego podstawy zdeterminowane są ewolucyjnie i którego zadaniem jest zagwarantowanie zdolności organizmu do utrzymania się przy życiu. Podstawy dla takiego „systemu wartości” ułożone są w tych częściach mózgu, które odpowiedzialne są za regulację podstawowych funkcji organicznych, takich jak oddychanie, bicie serca, itp.

Opierając się na swej teorii mózgu, Edelman zaprojektował model sieci neuronowej pod nazwą „Darwin III”, zdolnej do „kategoryzowania” świata w reakcjach na odbierane bodźce⁶. Edelman podkreśla przy tym, że Darwin III jest najmniejszą siecią zdolną do „kategoryzowania” i że mózg prawdziwego zwierzęcia składa się ze znacznie większej ilości sieci tego typu.

Z kolei Grossberg rozwija teorię „adaptacyjnego rezonansu” (ART — „Adaptive Resonance Theory”), która zakłada, że głównym mechanizmem umożliwiającym sieciom neuronowym „kodowanie regularności”, jest kombinacja oscylacyjnych i adaptacyjnych zachowań grup neuronów. Rozkłady aktywacji jednostek sieci synchronizują się z sobą w czasie, w efekcie wzajemnych nielokalnych i nieliniowych oddziaływań. Kategorie kognitywne reprezentowane są w postaci fal stojących, powstających, gdy oddziałujące poprzez sprzężenia zwrotne

⁵ Edelman pisze: „A fundamental premise of the TNGS is that the selective coordination of the complex patterns of interconnection between neuronal groups by reentry is the basis of behavior. Indeed, reentry (combined with memory [...]) is the main basis for a bridge between physiology and psychology”. (Por. Edelman [1992], s. 85).

⁶ Edelman pisze o sieci typu Darwin III: „Its nervous system is organized into several subsystems, each responsible for different aspects of its behavior. [...] After experience with randomly moving objects that it «sees», its eye will follow any object. [...] Similarly, its arm reaches out to «touch» an object, and with each selection of movements it increases its success in achieving this touching”. (Tamże, s. 93).

jednostki oscylują w fazie. Sieci takie, zdaniem Grossberga, zdolne są do uczenia się stabilnych kategorii percepcyjnych. Grossberg wyraża też przekonanie, przyznając się do filozoficznej inspiracji czerpanej z E. Macha, że „ART” pozwoli przezwyciężyć odwieczny filozoficzny problem dualizmu umysł–ciało.

Wielu filozofów-kognitywistów, w tym przede wszystkim Paul i Patricia Churchland, podziela pogląd Edelmana, Grossberga i większości naukowców, że to, co już obecnie wiadomo o budowie i działaniu układu nerwowego wystarcza do zrozumienia, jak układ ten, a zwłaszcza ludzki mózg, czyli w ich mniemaniu — umysł — zdolny jest do kategoryzowania świata. Paul Churchland uważa na przykład, podobnie jak Grossberg, że wszelkie zachowania układów kognitywnych mogą być wygenerowane przez złożenie odpowiednich „prototypowych” liniowych i cyklicznych zmian rozkładów aktywności neuronów, reprezentujących procesy motoryczno-percepcyjne. Zmiany tego typu można matematycznie przedstawić w odpowiedniej przestrzeni fazowej w postaci atraktorów typu punktu stałego i cyklu granicznego. Ten pierwszy typ odgrywać ma decydującą rolę w procesach rozpoznawania prototypowych kształtów czy obiektów takich, jak np. twarz, czyli rozpoznawania stałych konfiguracji cech fizycznych. Ten drugi natomiast, ma umożliwiać procesy rozpoznawania prototypowych sekwencji czasowych różnych konfiguracji takich cech.

Można więc ogólnie powiedzieć, że w ramach tych „konekjonistycznych” ujęć uznaje się, że ANNs nie różnią się w sposobie funkcjonowania od sieci naturalnych i że w związku z tym nie widać zasadniczych przeszkód uniemożliwiających skonstruowanie dostatecznie złożonych sieci, które cechowałyby się zdolnościami kognitywnymi zbliżonymi do ludzkich, a może i nawet je przewyższającymi.

Tę optymistyczną perspektywę zaburzają jedynie niektóre wyniki badań dynamiki złożonego systemu nerwowego oraz pewne ogólne trudności natury epistemologicznej.

II. Naturalne sieci neuronowe

Część neurofizjologów, powołując się na dane empiryczne charakteryzujące dynamikę złożonych układów nerwowych, twierdzi, że nie da się kognitywnych zdolności takich układów, przede wszystkim mózgu, zrozumieć i adekwatnie wyjaśnić przez odwołanie się do tak prostych dynamik, jak punkty stałe i cykle graniczne.

W. Freeman⁷, na przykład, wskazuje na to, że dynamika układu nerwowego, nawet w przypadku stosunkowo prostych organizmów, ma o wiele bardziej złożony charakter od tego, który zakładają Grossberg czy Edelman. Uważa on, że jedynym systemem pojęciowym umożliwiającym zrozumienie funkcjonowania

⁷ Por. Freeman [1991], Skarda & Freeman [1987].

złożonego układu nerwowego, takiego jak mózg, jest teoria chaosu deterministycznego.

W wyniku wieloletnich badań dynamiki układu nerwowego, przede wszystkim układu olfaktorycznego w mózgu królika, Freeman doszedł do wniosku, że kategorie kognitywne, takie jak zapachy, reprezentowane są przez tzw. atraktory chaotyczne i że każdemu atraktorowi odpowiada jeden zapach, z tych które organizm zdolny jest rozpoznać. Akt percepcji, jego zdaniem, polega nie na „odwzorowaniu” bodźca w układzie nerwowym, a na zaburzeniu jego wewnętrznej dynamiki prowadzącym do przeskoku układu z dziedziny jednego atraktora w dziedzinę innego. Każdy taki „przeskok” jest krokiem na trajektorii wytyczającej kierunek „reorganizacji” i „rozwoju” mózgu. Freeman spekuluje przy tym filozoficznie, że świadomość uznać można za „subiektywne doświadczenie tego procesu”.

Model Freemana opisujący *how brains make sense of the world*, pomimo wspierającego go bogactwa danych empirycznych, jest nieadekwatny tak naukowo, jak i filozoficznie. Charakter dynamiki generowanej w mózgu ma bowiem odpowiadać, zdaniem Freemana, nie tyle fizycznym cechom bodźca, co jego „znaczeniu” dla organizmu. Ponieważ „znaczenie” to zależy musi od całej historii oddziaływań organizmu z otoczeniem, to nie da się przewidzieć, jaka dynamika (jaki stan kognitywny) wygenerowany zostanie przez mózg w reakcji na kolejny bodziec i co to w ogóle może znaczyć z punktu widzenia zewnętrznego obserwatora, że coś dla organizmu jest „bodźcem”. W efekcie Freeman oscyluje ciągle pomiędzy naukowym realizmem i „mózgowym solipsyzmem”.

Mimo to Freeman, w odróżnieniu od wielu filozofów-kognitywistów, zdaje się dostrzegać filozoficzne problemy wynikające z założenia, że funkcjonowanie układu nerwowego nie polega na odwzorowywaniu docierających do organizmu bodźców i że układ ten należy uznać za informacyjnie zamknięty. Jednakże nie wysnuwa on z tego wniosków epistemologicznych, zapewne dlatego, że godziłyby one w „obiektywistyczne” podstawy epistemologii nauki i podważały naukowy status jego własnych badań.

Podobne koncepcje na temat informacyjnej zamkniętości układu nerwowego, a nawet całego organizmu, odnaleźć można w bardziej rozwiniętej formie, u H. Maturany i F. Vareli [1980], [1987]. Dostrzegli oni ponadto nieuchronnie samozwrotny charakter wszelkich teorii naukowych dotyczących człowieka, w tym i ich własnej. Nie zdecydowali się jednak pójść w ślady Kanta i przyznać, że teorie takie nie mogą mieć w związku z tym charakteru naukowego.

Poniżej scharakteryzuję różnice pomiędzy umysłem i mózgiem, korzystając z koncepcji autopoietyczności wszelkich złożonych układów dynamicznych, które samoorganizują się w wyniku oddziaływania ze środowiskiem, opracowanej przez Maturanę i Varełę.

III. Umysł i mózg jako autopoietyczne układy hierarchiczne

Umysł, gdy zwraca się w swych analizach ku samemu sobie, „dostrzega” natychmiast, że cechuje go nieustanna zmienność i ewolucja mająca charakter hierarchiczny. Podobne cechy przypisać też wypada każdemu złożonemu układowi nerwowemu, takiemu jak mózg, który najwyraźniej ma charakter hierarchiczny, zarówno w wymiarze strukturalnym, jak i, co neurofizjologowie dostrzegają wszakże znacznie rzadziej, w wymiarze dynamicznym⁸. To właśnie hierarchiczność, zwłaszcza w wymiarze dynamicznym, różni zasadniczo złożony układ nerwowy, traktowany jako system kognitywny, od ANNs.

Ujęciem, które pozwala w sposób bardziej precyzyjny niż inne uchwycić cechy umysłu i odnieść je do odpowiedniego modelu układu nerwowego, jest teoria systemów autopoietycznych. Teoria ta umożliwia w szczególności sprecyzowanie różnic pomiędzy naturalnymi i sztucznymi sieciami neuronowymi.

U podstaw teorii systemów autopoietycznych leży założenie, że każdy złożony system dynamiczny, taki jak społeczeństwo, organizm żywy lub też jakikolwiek jego organ, zachowuje swą tożsamość dzięki samoodtwarzaniu się jego elementów i ich relacji w wyniku samozwrotnego procesu odróżniania się od swego otoczenia i konstytuowania tym samym swoich granic. W ten sposób, pomimo ciągłych zmian wewnętrznych, system samoodtworza się jako „ten sam”.

Zasadnicza różnica pomiędzy ANNs i systemem nerwowym polega na tym, że ANNs nie są systemami autopoietycznymi. Ponadto system nerwowy, traktowany jako układ fizjologiczny, jest systemem autopoietycznym charakteryzującym się innym sposobem samoodtwarzania niż system nerwowy, traktowany jako złożona sieć neuronowa przetwarzająca informację. W odróżnieniu od systemu nerwowego pojętego jako układ fizjologiczny, system nerwowy pojęty jako sieć neuronowa, samoodtworza się dzięki hierarchizowaniu swej wewnętrznej dynamiki i przy braku możliwości takiego hierarchizowania — zamiera.

Mózg, jako organ fizjologiczny, determinuje swe granice przez odróżnienie się od tych podsystemów organizmu, z którymi oddziałuje na poziomie fizjologicznym i które w związku z tym niejako „istnieją dla niego” jedynie jako jego wewnętrzne stany fizjologiczne. Natomiast mózg charakteryzowany jako złożona sieć neuronowa o określonej dynamice, determinuje swoje granice przez różnicowanie swej wewnętrznej dynamiki w reakcji na zaburzające ją bodźce, przede wszystkim te, które docierają z narządów zmysłowych⁹. Naukowcy, a wraz z nimi liczni filozofowie, nie odróżniają mózgu jako organu fizjologicznego od mózgu traktowanego jako układ przetwarzający informację, a tego ostatniego od umysłu. Prowadzi to w konsekwencji do tak absurdalnych prób, jak poszukiwanie świadomości w procesach fizjologicznych zachodzących w szyszynce lub innych

⁸ Por. Holden [1991].

⁹ Model takiej dynamiki naszkicowałem w Bielecki [1995].

częściach mózgu czy wręcz w procesach molekularnych lub kwantowych zachodzących w neuronach.

Umysł wszakże różni się od systemu nerwowego jako sieci neuronowej, w tym również i mózgu, sposobem autoreprodukcji. O umyśle najsensowniej jest mówić jako o hierarchicznym systemie dynamicznym samoodtworzającym się poprzez reprodukcję pojęć, która rozpoczyna się wraz z pojawieniem się samoświadomości, pierwotnie ujawniającej się oczywiście w postaci posługiwania się słowem „ja”. Ten dynamiczny system cechuje się, w odróżnieniu od mózgu, zdolnością do samoopisu, dzięki której to właśnie w umyśle, a nie w mózgu, może dokonać się rozróżnienie pomiędzy umysłem i mózgiem lub też ich utożsamienie.

Hierarchizacja umysłu dokonuje się w procesie „hermeneutyczno-egzystencjalnego” wykraczania poza konstytuujące umysł przedsady i w ten sposób sytuuje go w wymiarze czasowym. Wykraczanie takie odbywać się musi w kierunku najszerszego otoczenia, którym może być jedynie to, co pozapojeciowe¹⁰.

Uwagi końcowe

W oparciu o zarysowaną powyżej autopoietyczną koncepcję złożonych układów dynamicznych narzuca się szereg wniosków na temat sensowności modelowania układu nerwowego za pomocą ANNs oraz modelowania umysłu przez odwoływanie się do struktury i dynamiki mózgu. Można, na przykład, przewidywać, że ze względu na zbyt prostą strukturę istniejących obecnie ANNs i brak możliwości generowania przez nie dynamik typu hierarchicznego, nie grozi ludzkości w najbliższej przyszłości pojawienie się androidów zdolnych cokolwiek „zrozumieć”, a już na pewno nie takich, które zagrozić by mogły ludzkiej dominacji.

Ponadto, nawet jeśli w dalszej przyszłości, postęp naukowy umożliwi skonstruowanie dostatecznie złożonych ANNs, cechujących się zdolnością samoodtworzania, czego logicznie wykluczyć nie sposób, to androidy w nie wyposażone, pozostałyby na poziomie zwierząt. Jeśli jednak traktowane będą „jak ludzie”, to by znaczyło, że umysł, zapoznając swą zdolność do hierarchizowania, zredukował się do mózgu.

Na zakończenie podkreślić należy, że poważne zagrożenie dla człowieka płynie nie ze strony materialnych wytworów rozwoju technologicznego, takich jak przyszłe możliwe inteligentne androidy, ale ze strony samego ludzkiego umysłu usiłującego zredukować się do takiej czy innej kategorii. Na podobieństwo człowieka redukującego siebie do kategorii społecznej, psychologicznej lub biologicznej („swojego ciała”), umysł redukujący się do kategorii biologicznej („mózgu”)

¹⁰ Szerzej autopoietyczną naturę relacji pomiędzy tym, co rozumiane, w mojej terminologii — „umysłem”, a tym, co poza nie(go) wykracza, omawia Marta Calvo w swoim artykule *Rozumienie jako autopoietyczna samoobserwacja*, w tymże numerze „Przeglądu Filozoficznego”.

zapoznaje swą własną autopoietyczną naturę, która polega właśnie na zdolności do nieograniczonego wykraczania poza siebie w hierarchicznym procesie konstytuowania własnych granic. W efekcie umysł taki podważa swój status umysłu i zaburza proces rozumienia samego siebie, czyli ostatecznie swe własne życie. Albowiem, jak trafnie dostrzegł E. Lévinas „Żywotność życia to nieustanne rozszarpywanie tożsamości”¹¹.

Bibliografia

- Bielecki M., [1995], *Czas i stawianie się w sztucznych i naturalnych układach kognitywnych*, „Filozofia Nauki”, nr 4 (12).
- Churchland P.M., [1995], *The Engine of Reason, the Seat of the Soul. A Philosophical Journey into the Brain*, Cambridge, MA, MIT Press.
- Dreyfus H., [1979], *What Computers Can't do. A Critique of Artificial Reason*, 2nd ed., New York, Harper and Row.
- Edelman G., [1989], *The Remembered Present. A Biological Theory of Consciousness*, New York, Basic Books.
- Edelman G., [1992], *Bright Air, Brilliant Fire*, New York, Basic Books.
- Freeman W. J., [1991], *The Physiology of Perception*, „Scientific American”, 2, s. 78-85.
- Grossberg S., [1982], *Studies of Mind and Brain: Neural Principle of Learning, Perception, Development, Cognition and Motor Control*, Boston, D. Reidel.
- Grossberg S. (ed.), [1988], *Neural Networks and Natural Intelligence*, Cambridge, MA, MIT Press.
- Holden A., [1991], *Structural, Functional and Dynamical Hierarchies in Neural Networks*, w: Beier G., Klein M. (eds.) *A Chaotic Hierarchy*, Singapore, World Scientific, s. 187-197.
- Lévinas E., [1986], *De Dieu qui vient à l'idée*, Paris, Librairie Philosophique J. Vrin.
- Maturana H., Varela F., [1980], *Autopoiesis and Cognition*, Dordrecht, D. Reidel.
- Maturana H., Varela F., [1987], *The Tree of Knowledge*, Boston, Shambhala.
- Rumelhart D.E., McClelland J.L. (eds.), [1986], *Parallel Distributed Processing*, Vol. 1-2, Cambridge, MA, MIT.
- 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.
- Winograd T., Flores F., [1986], *Understanding Computers and Cognition*, Reading, MA, Addison-Wesley.

¹¹ „La vivacité de la vie — incessant éclatement de l'identification”. E. Lévinas, *De la conscience à la veille. A partir de Husserl*, w: Lévinas [1986], s. 57.

Understanding in Neural Networks

Many philosophers accept uncritically the scientific image of man according to which humans are biological organisms that differ from animals mainly in that they are endowed with a more complex nervous system. Such systems, including the brain, can be viewed as extremely complex neural systems. Ever since artificial neural networks (ANNs) have emerged as a tool in modelling the activity of the nervous system, the question arises, to what extent ANNs are capable of providing an insight into the nature of human cognitive capabilities. The author refers to the theory of autopoietic systems and W. Freeman's neurophysiological studies to show how artificial neural networks, including the most complex of them, the so-called recursive ANNs, are incapable of providing an accurate model of the human nervous system or exhibiting cognitive abilities. First of all, in contrast to natural neural networks, ANNs are not capable of autopoiesis, i.e. self-reproduction. Secondly, when we look at the brain and the mind as autopoietic systems, it becomes apparent that their specific processes of autopoietic self-organization are fundamentally different from each other.