ROCZNIKI FILOZOFICZNE
Tom LXIII, numer 4 – 2015
DOI: http://dx.doi.org/10.18290/rf.2015.63.4-5

WIESŁAW GALUS

ARCHITEKTURA ŚWIADOMOŚCI
CZĘŚĆ III: WOLA I SENS ISTNIENIA

10. MOTYWACJE DO MYŚLENIA

W znanym kognitywistycznym dowcipie twórca samoświadomego robota relacjonuje pierwszy kontakt z inteligentną maszyną: „powiedziała, że może myśleć, ale na razie podjęła decyzję, że nie będzie...". Ażeby osiągnąć wysoki stopień świadomości, trzeba mieć motywację. Powszechnie uważa się, że główną motywacją do działania inteligentnych systemów, włączając w to systemy kierowane mózgami naturalnymi, zwierzęcymi, i sztucznymi AI, jest unikanie bólu i poszukiwanie przyjemności. W modelowaniu matematycznym metodami AI zagadnienie to można uprościć, przyjmując, że ból i przyjemność to parametr matematyczny, rożnie definiowany, w zależności od parametrów fizycznych opisujących adaptację systemu do środowiska. Taki uogólniony ból parametryczny sprowadza się czasem do jednego parametru, gdzie przyjemność także nazywamy bólem, tyle, że ze znakiem ujemnym.

W mózgach naturalnych mechanizmy przyjemności i bólu mają o wiele bardziej zróżnicowaną strukturę, wykształconą ewolucyjnie. W wyniku badań neurologicznych i psychologicznych prowadzonych w ostatnich latach, teoria bólu i przyjemności jest dość dobrze rozwinięta [zob. np. WYLLER 2005].

BÓL

Według popularnego stanowiska perceptualnego poczucie bólu (i innych cielesnych odczuć, jak np. łaskotki, swędzenie, orgazm) jest lokalną reakcją

Dr inż. WIESŁAW GALUS – pracownik Eutherm Sp. z o.o.; adres do korespondencji: ul. Zachodzącego Słońca 46, 01-495 Warszawa; e-mail: w.galus@eutherm.eu

na obiektywne pobudzenie tego regionu, podobne do rzeczywistego lub wyimaginowanego uszkodzenia tkanki [PITCHER 1970]. W szczególności silne bodźce, przekraczające próg akceptowalności systemów receptorowych, mogą być odczuwane jako ból [PAIN 2005].

Znaczenie, jakie przypisujemy odczuwaniu bólu, wiąże się z tym, że jest on doznaniem związanym ze świadomością. Jest jednym z fundamentalnych odczuć powstałym na wczesnym etapie ewolucji. Generacja bólu zaczyna się zwykle od pobudzenia specjalnych receptorów bólu (nocyceptorów) w peryferyjnych częściach systemu nerwowego, reagujących na bodźce potencjalnie uszkadzające tkankę. Są one przekazywane do centralnego systemu nerwowego przez specjalnie dedykowane włókna nerwowe, poprzez rdzeniowy system modulujący informację do śródmózgowia i wzgórza. We wzgórzu wyspecjalizowane, somatotropicznie organizowane pola przetwarzają te pobudzenia i kierują do wyższych pól mózgowych. Wyspecjalizowany mechanizm percepcji można uważać za odrębny zmysł angażujący odrębną modalność sensoryczną. Z tego punktu widzenia mechanizm odczuwania bólu może mieć charakter percepcyjny. Dla wyjaśnienia fenomenu bólu konieczne jest zbadanie relacji pomiędzy postrzeganiem bólu (postrzeganie wymaga kategoryzacji), jego doświadczaniem (poczuciem), identyfikacją i lokalizacją źródła bólu oraz adekwatnością tworzonych reprezentacji mentalnych. [AYDEDE 2009]. Także interpretacja reprezentacjonistyczna budzi wątpliwości, ponieważ adekwatność reprezentacji wymaga konceptu-pojęcia bólu, wykształconego uprzednio w toku doświadczeń życiowych. U ludzi, nawet we wczesnym stadium rozwoju, obserwujemy wrażliwość na ból. Zdolność do odczuwania bólu wydaje się cechą wrodzoną, o małej plastyczności, co sprzeczne jest z postulowaną koniecznością uprzedniego wykształcenia pojęcia bólu. Z punktu widzenia wielu filozofów umysłu ból może być traktowany tak jak i inne qualia. Z reguły więc postawa filozoficzna wobec poczucia bólu przenoszona jest automatycznie na zjawisko świadomości. Tym zagadnieniem jednakże zajmiemy się dopiero w następnym rozdziale. Obecnie zauważyć należy, że krytyka perceptualizmu i reprezentacjonalizmu przeprowadzona być może z zupełnie innych, na pierwszy rzut oka, redukcjonistycznych pozycji. Dyskusja nad qualiami, z których jednym może być ból, traci sens, gdy uświadomimy sobie, że odczuwanie bólu wiąże się z równoczesnym indukowaniem dzięki FSS impresjonów poczucia bólu, źródła bólu, jego przyczyny i lokalizacji, czyli wyobrażenia części ciała, gdzie ból może być odczuwany. Zgodnie z hipotezą FSS te impresjony mogą

asocjować, a ich pobudzenie może dawać efekt halucynacji, złudzenia bólu. Może też indukować impresjony niewłaściwych lokalizacji bólu. Te asocjacje mogą być uświadamiane w dowolnych kombinacjach i permutacjach, tworząc nieprzebrane bogactwo odczuć. Wówczas także można subiektywnie oceniać adekwatność kojarzenia i asocjowania impresjonów reprezentujących pojęcia używane do raportowania odczuwanych wrażeń związanych z bólem. Asocjacje impresjonów generowanych przez nocyceptory tworzyć mogą impresjony wyższych rzędów, na przykład reprezentujących rodzaj uszkodzenia tkanki. Weryfikacja sensoryczna może zmieniać to wyobrażenie o źródłach bólu. Porównanie impresjonów według mechanizmu NPM z wcześniejszymi doświadczeniami, zbliża znów ten proces do postrzegania perceptualnego. Może nawet ulec zmianie odczucie bólu. Ból jest więc zjawiskiem złożonym, obejmującym zarówno proste jego postrzeganie, jak i odczuwanie na wysokim poziomie pojęciowym i emocjonalnym.

Dla skutecznego działania, zdobywania wiedzy o otoczeniu i uzyskiwania świadomości istotna jest możliwość wykorzystania bólu i wszelkiego poczucia dyskomfortu jako motywacji do uczenia i samouczenia. Będzie to możliwe, gdy reprezentacje mentalne, impresjony, tworzone w procesie uczenia będą kojarzone z impresjonami reprezentującymi karę lub nagrodę. Szczególnie łatwo jest realizować takie zadanie w przypadku uczenia pod nadzorem, gdy o motywacji decyduje nauczyciel. W przypadku samouczenia mechanizm motywacji zależy od subiektywnej oceny rezultatów działania. [SCHULTZ, DAYAN, MONTAGUE 1997; SCHULTZ & DICKINSON 2000]. Organizmy żywe stykają się nieustannie z presją motywacyjną w postaci poczucia bólu lub przyjemności. Odczuwanie bólu jest ich najbardziej powszechnym odczuciem, bezwzględnie karzącym zachowania w niedostatecznym stopniu adaptujące system do środowiska.

Zdajemy sobie sprawę, że uproszczone podejście konstruktorów AI, ograniczające motywacje inteligentnych systemów do uogólnionego bólu parametrycznego, dalece nie wyczerpuje złożoności czynników motywujących do działania zwierzęta i ludzi. Matematycznie symetryczna część doznań o źródłach cielesnych wiąże się z odczuwanie przyjemności. Niektóre badania psychologiczne wskazują, że przyjemność, wyzwalana mechanizmem nagrody, może być skuteczniejszą motywacją do uczenia i samouczenia niż ból.

NEURONALNY MECHANIZM NAGRODY

Pojawiło się wiele doniesień wskazujących, że neurony szlaków dopaminergicznych (przepływ informacji w tych obwodach reguluje dopamina) biorą udział w odczuwaniu przyjemności. Sygnalizują także błąd w przewidywaniu otrzymania oczekiwanej nagrody w trakcie procedury uczenia (*reinforcement learning*).Towarzyszy im pobudzenie sygnałami dopaminowymi [SCHULTZ i in. 1997]. Uwalnianie dopaminy może występować w procesach eksploracji oraz pojawiania się elementów nowości w postrzeganych scenach [DAYAN & BELLEINE 2002; KAKADE & DAYAN 2002]. W szczególności sygnał dopaminowy może powiadamiać o „błędzie przewidywania", a nie tylko „błędzie przewidywania nagrody" [HORVITZ 2000, BUNZECK i in. 2012]. Hipokamp sygnalizuje wówczas efekt nowości, prążkowie – efekt nagrody, a środkowa kora oczodołowo-czołowa (mOFC) sygnalizuje zarówno nowość, jak i nagrodę. To sugeruje, że sygnał nowości płynący z płata skroniowego-przyśrodkowego obejmującego hipokamp jest interpretowany jako oczekiwanie nagrody przez mOFC generujący sygnał nagrody w prążkowiu. W ten sposób mechanizm nagrody może kontrolować formowanie pamięci długoterminowej w hipokampie. Dodatkowym argumentem potwierdzającym taką możliwość jest filogenetyczny rozwój mózgu, podczas którego funkcja hipokampa uległa zasadniczej zmianie. Na wczesnych etapach ewolucji, kiedy kora nie spełniała tak szerokich funkcji jak w później wykształconych mózgach zwierząt wyżej rozwiniętych umysłowo, procesem zapamiętywania kierował układ limbiczny, a miejscem zapamiętywania informacji był starszy ewolucyjnie hipokamp. Obecnie funkcje te w dużej mierze przejęły płaty kory czołowej i przedczołowej [BUCKNER i in. 1999; FLETCHER & HENSON 2001]. Neuronalny mechanizm nagrody jest więc stosunkowo dobrze zlokalizowany, choć wiele jego funkcjonalnych właściwości budzi wciąż wątpliwości [MORTEN i in. 2010, 2011].

O ile poznane mechanizmy wyjaśnić mogą mechanizmy Arystotelesowskiej *hedonii*, to *eudajmonii*, czyli poczucia sensu związanego z racjonalnym objaśnieniem rzeczywistości, poszukać powinniśmy w mechanizmach NPM-FSS.

Opisane powyżej mechanizmy kary i nagrody, bólu i przyjemności mają inny charakter niż analizowane przez nas funkcje sieci neuronowych, prowadzące do świadomości. Nie są one immanentną częścią struktury sieci neuronowej, typowej dla kory nowej, lecz mają wyraźnie zlokalizowane ośrodki i szlaki transmisyjne. Należy przypuszczać, że są one związane ze starszymi strukturami mózgu wykształconymi ewolucyjnie. Choć ich

skutki wpływają na działania całego umysłu, zarówno na poziomie świadomym, jak i nieświadomym, to są odrębne od procesów NPM-FSS kształtujących świadomość. Bardziej przypominają mechanizmy postrzegania, charakterystyczne dla innych zmysłów. Z tego względu traktować je można jako dodatkowe zmysły bólu i przyjemności, uzupełniające inne modalności, właściwe dla naturalnych umysłów.

11. CIEKAWOŚĆ, PIERWSZY STOPIEŃ DO NIEBA

Czy to zatem wszystkie motywacje, które kierować mogą postępowaniem świadomego zwierzęcia lub robota? Przecież pamiętamy, że wiele zwierząt, a także ludzi może wykonywać odruchy lub nawet podejmować złożone działania instynktownie. Jaka jest rola instynktu w motywowaniu świadomego działania i automatycznych odruchów? Czy robotom musimy zaszczepić jakieś instynkty, ażeby uzyskały świadomość? Na te pytania zdecydowanej odpowiedzi udzielił Stephen Grossber, a za nim Leonid Perlovsky. Każdy instynkt generuje wymierny sygnał emocjonalny, wskazujący na stopień satysfakcji lub niezadowolenia [GROSSBERG 1982; GROSSBERG & LEVINE 1987, PERLOVSKY 2008c]. Oczywiście, w przypadku zwierząt najbardziej prymitywnych (np. owady) zdolność do odczuwania emocji może być kwestionowana. Lecz można się zastanawiać, czy w tym przypadku należy mówić o instynktach, czy też raczej o reakcjach automatycznych. Do wielu działań, do których człowiek wykorzystuje inteligencję, zwierzętom wystarcza motywacja groźbą kary lub oczekiwaniem nagrody, utożsamiana z bólem lub przyjemnością. Błędy w korzystaniu z odruchów z reguły karane są bólem w szerokim sensie tego słowa. Lecz zdajemy sobie sprawę, że filozofia bólu posługuje się definicją i analizą nawiązującą głównie do bólu cielesnego, odnoszącego się do uszkodzenia tkanek. Wydaje się, że wyższe zwierzęta, szczególnie społeczne, odczuwają bardziej złożone formy bólu. Także ból psychiczny, związany z utratą członków rodziny lub stada. Im wyższy poziom inteligencji, tym większą role odgrywa motywacja bólem i zachowania wyuczone pod wpływem tej motywacji. Natomiast zmniejsza się rola zachowań instynktownych, wrodzonych. Mentalny charakter poczucia bólu i przyjemności wymaga co najmniej pewnego stopnia świadomości umysłu. Reakcje organizmów pozbawionych centralnego układu nerwowego i zdolności komunikacji, są czysto automatyczne, instynktowne. Mogą one demonstrować zachowania o pewnym stopniu inteligencji, począwszy od prostych

tropizmów, a skończywszy na zdolności grupowania się, modyfikacji funkcji fizykochemicznych itp. Mogą nawet wykazywać pewną zdolność uczenia się. Nie mają jednak zdolności planowania, przewidywania skutków własnych działań, budowy dostatecznie złożonego modelu otoczenia i innych działań świadczących o przejawach świadomości. Bez świadomości ból nie może być odczuwany, nie może więc spełniać swojej funkcji motywacyjnej. Stąd podejrzenie, że najbardziej prymitywne zwierzęta o ograniczonej inteligencji nie są zdolne do odczuwania bólu. Ich zachowania mają charakter automatyczny i nie wymagają żadnej motywacji.

HIERARCHIA BÓLU

Ciekawą próbę znalezienia motywacji wyższego rzędu przedstawił Janusz Starzyk [STARZYK 2008; STARZYK & PRASAD 2011]. Zaprezentowana przez niego koncepcja uznaje uogólniony ból za emocję wymagającą zaspokojenia, usatysfakcjonowania. Jest to więc fundamentalna motywacja do inteligentnego działania inteligentnego agenta, czyli inteligencji ucieleśnionej. Współdziałając z mechanizmem uwagi i pamięcią epizodyczną, motywacja bólem pozwala na formułowanie celów i planowanie inteligentnych działań prowadzących do uzyskiwania satysfakcji z uniknięcia bólu. Bezpośrednie jednak działania prowadzące do uniknięcia bólu mogą być z różnych przyczyn niemożliwe. Inteligentny agent może także nauczyć się przewidywać, że choć ból aktualnie nie występuje, to może się on pojawić w przyszłości. Złożone sytuacje, uniemożliwiające uniknięcia dolegliwości bezpośrednich, które autor nazywa bólami prymitywnymi (niskiego rzędu), oraz przewidywanie dyskomfortu w przyszłości, prowadzi do kreowania bólów wyższego rzędu. Według podanego przez autora przykładu zobaczenie pustej lodówki może spowodować dyskomfort w przewidywaniu głodu. Podobnie obejrzenie pustego portfela może nasunąć myśl, że lodówka pozostanie pusta. Tworzy się w ten sposób hierarchia bólów – od prymitywnych, odczuwanych zmysłowo, do bólów abstrakcyjnych, podlegających analizie logicznej i docierających do warstwy pojęciowo-symbolicznej naszej świadomości [STARZYK 2011a; STARZYK i in. 2012]. Kierując się tymi abstrakcyjnymi motywacjami, powinniśmy generować bogactwo zachowań właściwych dla istot o wysokim stopniu świadomości. Powyższy system motywacji zaprojektowany jest do nadawania „sensu istnienia" autonomicznym robotom inteligentnym. I rzeczywiście takie zachowanie obserwuje się w inteligentnych, autonomicznych systemach kognitywnych [HAIKONEN 2007; NUXOLL &

LAIRD 2004, 2012]. Systemom takim stwarzamy motywację w postaci dążenia do osiągania pożądanych „wartości" zdefiniowanych sztucznie, np. minimalizacji sztucznego „bólu" zadawanego parametrycznie jako reakcja na „szkodliwe wpływy wrogiego otoczenia" lub też optymalizacji pożądanych wartości parametrów układów percepcyjnych lub homeostatycznych. Motywacje mogą być także wytwarzane poprzez narzucone z góry funkcje, np. eksploracji terenu. Tak stworzone motywacje nie stawały się jednak immanentną cechą sztucznego umysłu. System dążył do osiągnięcia zadanych przez konstruktora wartości, jednakże nie odczuwał ich, a więc nie miał zdolności zrozumienia wykazywanych instynktów. Wiele jednak umysłów zwierzęcych, o niskim stopniu świadomości, być może ogranicza się do takich właśnie motywacji. Zdajemy sobie sprawę, że motywacje kierujące człowiekiem są znacznie bardziej złożone [OUDEYER & KAPLAN 2008; OUDEYER i in. 2007]. Co kreuje całą subtelną naturę duszy ludzkiej z jej wrażliwością na piękno, poczuciem etyki, miłości, dążeniem do wolności i wolą poznania i tworzenia?

CIEKAWOŚĆ

Kluczową motywacją jest dążenie do osiągnięcia satysfakcji w *eudajmonii*, czyli w czystym akcie poznawania i rozumienia świata, a także samego siebie. Obecnie możemy uzupełnić idee Arystotelesa i wskazać mechanizmy biofizyczne, które do tego prowadzą.

Omówione wyżej właściwości sieci neuronowej ukazują nam strukturę hierarchiczną neuronowych pól modelujących NPM, którą tworzą wysoce zorganizowane struktury mózgu, przekraczające ramy mikrokolumn i kolumn korowych. Są one podłożem biologicznym, w którym tkwią utrwalone mechanizmami molekularnymi impresjony, wytworzone w procesie uczenia. Impresjony są nie tylko reprezentacjami mentalnymi postrzeganych dzięki zmysłom obiektów, ale w wyższych warstwach są reprezentacjami wszelkich myśli, a więc pozyskanej wiedzy symbolicznej, pojęć/konceptów, czyli impresjonów stacjonarnych, przechowywanych w pamięci deklaratywnej, a także reprezentacjami wszelkich wspomnień zdarzeń z historii naszego życia, procesów, przekształceń i transformacji oraz ciągów sygnałów lub symboli w postaci np. muzyki lub mowy. Te ostatnie reprezentacje tworzą opisane wyżej impresjony dynamiczne. Dzięki funkcjonalnym sprzężeniom synaptycznym FSS, impresjony te mogą asocjować, umożliwiając uczenie korela-

cyjne, przypominanie intencjonalne, halucynacje i wyobrażenia, a także planowanie i przewidywanie skutków własnego działania. Pozostaje do wyjaśnienia, co poza zmysłami bólu i przyjemności może pobudzać taki system do generowania wysokich, subtelnych funkcji umysłowych. Przy tym poszukiwać musimy motywacji tkwiącej wewnętrznie w sieci neuronowej, immanentnie z nią związanej (*intrinsic motivation*), a nie tylko wewnętrznych motywacji generowanych oczekiwaniem nagrody lub unikaniem kary (*internal motivations*).

Tego typu motywacja kryje się właśnie w zarysowanej wyżej morfologii i funkcjach sieci neuronowej. W czymże innym mogłaby się kryć? Wszak nie mamy software, zmuszającego nasze mózgi do wykonywania zaprogramowanych działań. Działa ona w następujący sposób. Neuronowe pola modelujące zawierają miliony i miliardy synaps, zgrupowanych w wysepki synaptyczne i agregaty wyższego rzędu. Te matryce sond ładunkowych gotowe są przyjąć pobudzenia ładunkowe, które do nich docierają. Neuronowe pola modelujące gotowe są przesłać te pobudzenia do wyższych pól, jednakże uprzednio muszą je porównać z wzorcami wcześniejszych pobudzeń zachowanych w pamięci. Mówi nam o tym teoria NPM. Lecz z czymże je porównują? Porównują je ze strukturami molekularnymi tworzącymi pamięć trwałą lub krótkotrwałą. Pola modelujące stanowią filtr, który selekcjonuje te impulsy kierunkowe, które pasują do struktury molekularnej wytworzonej w trakcie wcześniejszych doświadczeń.

Jürgen Schmidhuber w swojej pracy w 2003 r. napisał: „Możemy się nauczyć tyko tego, co już prawie wiemy...” [SCHMIDHUBER 2003]. Omawiana właściwość neuronowych pól modulujących może wyjaśnić ten psychologiczny efekt. Na przekazywane impresjony sensoryczne reagują pola NPM dostosowane do ich struktury elektrodynamicznej. Zespoły ukierunkowanych ładunków odbiegające od topologii kanałów jonowych w membranach synaptycznych i dendrytach, nie są w stanie wnikać w szlaki transmisyjne wiązek neuronowych przekazujących je do wyższych pól neuronowych. Natomiast dokonują tego impresjony o właściwej strukturze przestrzennej oraz topologii i w ten sposób następuje ich rozpoznanie. Powielane pobudzenie tworzy drzewo pobudzeń charakterystyczne dla impresjonu. Jak tworzą się bardziej złożone struktury, już opisywaliśmy. Ważniejsze jest, jak muszą się uzupełniać impresjony poprzez ich wyszukiwanie i asocjowanie, żeby tworzyć struktury przestrzenne i topologiczne pasujące do modeli-konceptów

rezydujących w wyższych polach modelujących i pasujące do impresjonów wyższego rzędu. Ten proces ma charakter ciągłego naturalnego procesu, wynikającego z koneksjonistycznej struktury sieci neuronowej. Pola modelujące nieustannie gotowe są przyjmować pobudzenia odpowiadające ich strukturze. W istocie nie robią nic innego. Oscylacyjna działalność mózgu jest i skutkiem interferencji transmitowanych pobudzeń, ale także przyczyną pobudzeń sąsiednich pól. Dzięki temu mamy do czynienia z ciągłą penetracją zasobów pamięciowych i tworzeniem potrzebnych asocjacji. Ich przekazywanie bottom-up następuje szlakami wrodzonych struktur neuronowych opisywanych w rozdziale 9., modyfikowanych, dzięki oddziaływaniu pobudzeń ładunkowych ukierunkowanych zgodnie z teorią neuroelektro-dynamiki NED. W ten sposób w procesie samoorganizacji w wyniku uczenia, na ponadkolumnową strukturę wrodzoną, nakłada się plastyczna, hierarchiczna struktura impresjonów. Tak więc oprócz edelmanizmu neuronowego mamy tu do czynienia z darwinizmem memetycznym, gdzie tworzone dzięki funkcjonalnemu sprzężeniu synaptycznemu aglomeraty zasocjowanych impresjonów „walczą" o przekroczenie filtru uwagi tak, aby mogły dotrzeć do świadomości.

Obraz ten wykazuje pewnego rodzaju niedostatek, polegający na tym, że impresjony najwyższego poziomu, dominujące naszą świadomość, niejako muszą czekać biernie, aż w wyniku selekcji dotrą do nich subimpresjony pobudzające je do aktywności. Przeczy temu psychologia, gdyż zarówno naszą uwagą, jak i ciekawością możemy sterować świadomie i intencjonalnie. Przeczy też temu anatomia mózgu i morfologia sieci, ponieważ struktury te pełne są połączeń wstecznych i potężnych wiązek aksonów zwrotnych, tworzących sprzężenia do niższych pól percepcyjnych. Z dużym prawdopodobieństwem można przyjmować, że sprzężenia te służą do pobudzenia impresjonów niższego rzędu. Te pobudzenia mogą wypełniać lukę i likwidować dysonans poznawczy, dostarczając szczegółów niezbędnych do wypełnienia ogólnych modeli-konceptów zajmujących naszą świadomą uwagę. Jak adresowane są zwrotne impulsy pobudzające? Prawdopodobną zasadą adresowania jest, podobnie jak w przypadku FSS, sprzężenie funkcjonalno-strukturalne. Oznacza to, że sygnały zwrotne przekazywane są szlakami równoległymi do szlaków aferentnych, lecz skierowanych przeciwnie, zstępująco. Organizacja sieci prawdopodobnie obejmuje równoczesne tworzenie obu szlaków w bezpośrednim sąsiedztwie, a być może sprzężonych ładunkowo tak, że pobudzenia zwrotne wracają ścieżkami wykształconymi w pro-

cedurach uczenia i samoorganizacji. Zbadanie zasad zwrotnego pobudzania niższych pól powinno być priorytetem dalszych badań neurofizjologicznych. Powinna być im przypisana waga nie mniejsza niż badanie wstępujących pobudzeń neuronowych.

Wielokrotnie wskazywaliśmy na znaczenie uwagi dla przebiegu procesów molekularnych odpowiedzialnych za myślenie. Opisaliśmy wyżej, jak mechanizm selekcji WTA pozostawia w pamięci roboczej to, co jest w centrum naszej uwagi i co doprowadzone zostaje do naszej świadomości. Czynnikiem decydującym o tej selekcji jest ruch, dynamiczna zmiana sceny, elementy nowości.

Cóż jednak będzie się działo, gdy brak jest dynamicznych zmian w środowisku? Otóż nasz umysł zwiększa czujność. Obniża progi reakcji na sygnały stymulujące, zarówno na poziomie receptorowym jak i analitycznym. Układ wzrokowy, a także inne modalności dokonują skanowania sceny i sygnałów środowiska w poszukiwaniu nawet drobnych zmian lub nieregularności. Procesem tym kieruje znany mechanizm sakad wzrokowych. Ruch sakadyczny gałki ocznej kierowany jest na poszczególne elementy sceny według gęstości informacyjnej i dynamiki obrazu, co odzwierciedla przypuszczalnie ich znaczenie dla postrzegającego podmiotu. Podobnie przebiegać to może na wyższych piętrach hierarchii. Co może kierować tym, procesem skanowania? Po pierwsze spodziewać się tu można występowania mechanizmu sakad mentalnych, zaproponowanego przez Janusza Starzyka [STARZYK 2011b] i kierującego naszą uwagę w sposób przypadkowy na poszczególne elementy. Sam autor powątpiewa, czy mechanizm taki może być istotny w przypadku umysłów naturalnych, na przykład w przypadku człowieka. Traktuje on ten mechanizm, jako wygodny model pobudzania funkcji eksploracyjnych, użyteczny przy budowie inteligentnych robotów. Biorąc jednak pod uwagę jednorodność budowy sieci w obszarach korowych, można zakładać, że jeśli mechanizm ten występuje w korze wzrokowej, to pojawiać się może i w innych polach, powodując skutki przewidziane przez autora. Z punktu widzenia psychologii odpowiadają one błądzeniu myślami w poszukiwaniu tematu, na którym skupić byśmy chcieli naszą uwagę. Na wyższych poziomach percepcji można także przewidywać kierowania uwagi na obiekty odbiegające od utrwalonych uprzednio impresjonów, modeli sceny czy zdarzenia, czyli także wyszukiwania elementów nowości dzięki ruchowi, zmienności lub niezwykłości. Postulowaliśmy wyżej pobudzenia top-down szlakami najbardziej utrwalonymi, czyli ukształtowanymi poprzez

sprzężenia synaptyczne w procesie uczenia, w poszukiwaniu, w zasobach pamięciowych i doznaniach zmysłowych, niezbędnej wiedzy objaśniającej na poziomie symbolicznym nowe problemy pojawiające się przed inteligentnym agentem.

Wszystkie te procesy tworzą szeroki repertuar zachowań, których wspólną cechą są funkcje eksploracyjne, poznawcze, dążenie do zdobycia nowej wiedzy. Zespół emocji prowadzących do opisanych zachowań nazywamy „ciekawością". A więc to ciekawość byłaby tym podstawowym instynktem napędzającym naszą aktywność w kierunku zdobywania wiedzy o świecie i budowy nie tylko cząstkowych impresjonów-modeli-pojęć, ale także generalnego modelu środowiska i świata, w którym żyjemy?

Zdefiniujmy ciekawość jako potrzebę eksploracji dostępnych źródeł wiedzy w celu zaspokojenia potrzeb. Tę cechę w wersji najprymitywniejszej mają nawet jednokomórkowce. Ciekawość jest tożsama z poszukiwanym przez Perlovsky'ego instynktem wiedzy KI. Tym razem jednak widzimy, że nie ma ona postulatywnego charakteru, a jest immanentną cechą sieci neuronowej, zgodnie z wymogiem sformułowanym na wstępie. Organizm musi wykonywać procesy wynikające z jego struktury, musi nieustannie przeszukiwać zasoby pamięciowe, musi asocjować rezydujące w pamięci impresjony w celu wytworzenia agregatów pasujących do struktur impresjonów wyższego rzędu. Ten przymus nadaje ciekawości cechy instynktu. Instynkt ten steruje działaniami poznawczymi organizmu. Widać to na przykładzie sterowania mechanizmem uwagi. Na niższych poziomach instynktowne działanie ciekawości nadaje mechanizmowi charakter automatycznych procesów fizjologicznych i neurologicznych. Na poziomach wyższych przełączanie uwagi ma świadomy charakter intencjonalny i jest skutkiem działania ciekawości. Jak więc łatwo zauważyć, instynktowny charakter ciekawości zaciera się na etapie przetwarzania symbolicznego abstrakcyjnych pojęć w wyższych polach mózgowych, ponieważ działania badawcze, dążenie do pozyskiwania wiedzy, staje się na tyle skomplikowane, że natrafia na wiele ograniczeń funkcjonalnych i fizjologicznych. Są to ograniczenia wynikające ze zdolności intelektualnych, właściwości asocjacyjnych NPM, małej skuteczności FSS, niewydolności pamięci, źle funkcjonujących mechanizmów uwagi itp. Przełamanie tych przeszkód wymaga woli kierowanej świadomością. W efekcie w organizmach żywych występują olbrzymie różnice indywidualne w stopniu wykazywanej ciekawości pomiędzy osobnikami tego samego gatunku.

EMOCJE, MOTYWACJE

Można mieć wątpliwość, czy pospolita „ciekawość" może mieć jakiekolwiek znaczenie wobec potężnych zmysłów bólu i przyjemności. Można sądzić, że te mechanizmy: ciekawość, unikanie bólu, poszukiwanie przyjemności są konkurencyjne, i te silniejsze muszą dominować. Lecz głębsze zastanowienie podpowiada, że istnieje pomiędzy nimi potężny związek, odkryty już przez Arystotelesa. Jak już cytowaliśmy za Grossbergiem na wstępie niniejszego rozdziału, każdy instynkt rodzi emocje. Także Perlovsky, jak pamiętamy, sugerował, że zaspokojenie instynktu wiedzy rodzi emocję pozytywną. Efekt ten został potwierdzony eksperymentalnie przez Nico Bunzecka i współpracowników w pracy „Contextual Interaction Between Novelty and Reward Processing Within the Mesolimbic System" [BUNZECK i in. 2012]. Jak pamiętamy z analizy neuronowych mechanizmów nagrody, przedstawionych powyżej, zobrazowanie mózgu za pomocy fMRI wykazało, że pamięć długoterminowa nowych zdarzeń, zależna od płata skroniowego-przyśrodkowego MTL (*Medial temporal lobe*), jest modulowana przez obwody, które także reagują na układ nagrody. A więc to oczekiwanie nagrody tworzy motywację do poszukiwania nowości. Sprzężenie to możemy za autorami nazwać „nagrodą za eksploracje nowości". W badaniach fMRI tego sprzężenia wykazano, że zachodzące za jego pośrednictwem rozpoznanie i zapamiętywanie scen w pamięci trwałej (na okres 24 godzin) jest silnie skorelowane z aktywnością MTL, brzusznego jądra ogoniastego, a także obszaru istoty czarnej i nakrywki brzusznej SN/VTA (*substantia nigra/ventral tegmental area*). Przy tym główny efekt nowości sygnalizowała aktywność hipokampa, poczucie nagrody – jądro ogoniaste, a mOFC zarówno odczuwanie nowości, jak i nagrody. Te badania sugerują, że sygnał „nowości" odczytywany przez MTL jest interpretowany przez mOFC w kategoriach przewidywania nagrody, co napędza mechanizm nagrody w prążkowiu. Prążkowie, jądro ogoniaste wraz z pokrywą, razem z SN/VTA reguluje zależny od MTL proces formowania pamięci w hipokampie oraz generacji sygnału dodatkowej nagrody z tytułu skutecznych kontekstowych procesów eksploracyjnych.

Lecz procesy eksploracyjne, w których poszukujemy nowości, napędzane są właśnie ciekawością. Tak więc ludzie, ale także i zwierzęta, mogą odczuwać przyjemność, zaspokajając ciekawość. Skądinąd wiemy, że nuda może być uczuciem przykrym. To tę przyjemną emocję nazwał Arystoteles

eudajmonią. Poszukiwanie pełnej satysfakcji wspomaga ciekawość. Przeciwieństwem jest ból rodzący emocję negatywną. Oba instynkty, ból i ciekawość, wykazują olbrzymie podobieństwa:

1. Oba instynkty są naturalne dla istot żywych i dostatecznie silne, aby zadecydować o przeżyciu gatunkowym i osobniczym. Brak reakcji na ból stanowi niebezpieczną chorobę dla ludzi. Dla zwierząt przeżycie bez reakcji na ból jest niemożliwe. Jest rzeczą zrozumiałą i jasno wynikającą z badań nad zachowaniem zwierząt i ludzi, jak fundamentalne znaczenie ma ekspresja „ciekawości" w rozwoju onto- i filogenetycznym. Ciekawość pobudza aktywność osobniczą, jak również decyduje o ekspansji środowiskowej i możliwościach zbadania i eksploracji środowiska, zwiększających lub zmniejszających szanse przeżycia całego gatunku. Konkurencję ewolucyjną wygrywają gatunki o większej ruchliwości napędzanej ciekawością oraz osobniki zdolne do nauczenia się bardziej adekwatnych zachowań, wymagających wiedzy o otoczeniu. Wygasanie ciekawości jest charakterystyczne dla odchodzenia osobników starszych.

2. Oba wykształciły się ewolucyjnie. Presja selekcyjna bólu i przyjemności wydaje się oczywista. Także przejawy ciekawości tworzą przewagę ewolucyjną, ponieważ pobudzają działania eksploracji środowiska, pozwalają gromadzić wiedzę, ułatwiającą zdobywanie zasobów, unikanie niebezpieczeństw a także budowanie modelu środowiska, przez co reakcje osobnicze stają się adekwatne. Zastanawiająca jest powszechność ciekawości, jako przyczyny aktywności. Nawet najprostsze organizmy reagują na dostrzeżone elementy nowości w otoczeniu i podejmują działania, żeby się z nimi zapoznać. Dostrzec to można nawet na poziomie komórkowym, u organizmów jednokomórkowych, w postaci tropizmów i reakcji motorycznych na bodźce świetlne, termiczne lub chemiczne. I nic dziwnego, bo czymże innym są nasze komórki sensoryczne. One w identyczny sposób wykazują wrażliwość na nowe pobudzenia. Gdy bodziec pozostaje niezmienny, wówczas wrażliwość spada i sygnał pobudzenia wygasa (habituacja). Wskazuje to, w jaki sposób wrażliwość komórkowa mogła być ewolucyjnie wykorzystana do utworzenia zmysłów zwierzęcych. Wrażliwość zespołów jednokomórkowców mogła umożliwić reakcje kolektywne po wyewoluowaniu sposobów przekazywania sygnałów pobudzających, co doprowadziło do specjalizacji komórek i powstania systemów wielokomórkowych. Kiedy powstały wyspecjalizowane centralne ośrodki sterowania reakcjami motorycznymi na podstawie postrzeganych

obiektów, dalsza ewolucja przebiegała pomiędzy wyznaczonymi wyżej biegunami: ciekawości – dla wzmożenia aktywności eksploracyjnej i bólu – dla unikania niebezpieczeństw.

3. Dla obu można wskazać mechanizmy neuronowe ich powstawania. W toku dalszej ewolucji wyuczone sygnały niebezpieczeństwa musiały docierać do centralnego ośrodka uruchamiającego zdecydowaną akcję obronną całego organizmu. Odbierany sygnał zagrożenia musiał mieć charakter alarmu, wywołującego stan paniki i przerwanie innych funkcji. Większość czasu organizmy musiały pożytkować na regularną działalność eksploatacyjną, na którą mogła im pozwolić ich budowa i złożoność systemu nerwowego, wrażliwość zmysłów, pojemność pamięci, system kolekcjonowania doświadczeń, czyli zdolność uczenia się.

4. Oba mogą być generowane przez neuronowe pobudzenia zewnętrzne, jak i wewnętrzne. Pobudzenia zewnętrzne integralności cielesnej i przeciążenie wrażliwości zmysłowej skutkują dobrze nam znanym wrażeniem bólu na równi z sygnałami płynącymi z wnętrza naszego organizmu. W przypadku ciekawości bodźce zewnętrzne związane są z działalnością eksploracyjną i pozyskiwaniem wiedzy z zewnątrz, bodźce wewnętrzne zaś pochodzą z działalności penetracji zasobów wewnętrznych skutkujących asocjacjami impresjonów pasujących w przybliżeniu do konfiguracji utworzonych w wyniku wcześniejszych doświadczeń i przechowywanych w pamięci.

Należy uznać, że *de facto* ciekawość i ból to dwa bieguny psychiki istot naturalnych wzajemnie się uzupełniające. Istoty te przejawiają powszechnie ciekawość, jednakże jest ona powściągana strachem. Oczywiście strachem przed bólem (w sensie uogólnionym, czyli wszelkim dyskomfortem). Zgodzić się należy ze stanowiskiem wielu badaczy, że unikanie bólu odnosi się do inteligencji, czyli zachowań bardziej pierwotnych. Podkreśla to behawioralna definicja inteligencji sformułowana przez Janusza Starzyka: „inteligencja to zdolność systemu do uczenia się, jak przetrwać we wrogim środowisku" [STARZYK 2008]. Mogą również istnieć i demonstrowane są, w systemach naturalnych i sztucznych, inteligentne zachowania bez samoświadomości. Przedstawiony mechanizm objawiania się „ciekawości" dowodzi, , w moim przekonaniu, że ten instynkt odnosi się do kształtowania samoświadomości, czyli poczucia charakterystycznego dla najwyżej rozwiniętych istot. To wysoki stopień ciekawości pobudza rozwój osobniczy i gatunkowy.

Wydaje się, że to właśnie ciekawość, czyli potrzeba eksploracji środowiska i zdobywania o nim informacji, jest instynktem bardziej fundamen-

talnym i wcześniejszym ewolucyjnie niż zdolność odczuwania bólu. Przecież wynika ona z potrzeby aktywności, eksploracji środowiska, a także z elementarnych zasad ewolucyjnych, do których należy ekspansja terytorialna gatunku i popęd seksualny. Należy przyjąć, że obie te motywacje powinny działać równolegle, instynkt unikania bólu – w celu zachowania bezpieczeństwa, ciekawość – w celu osiągania wyższych stanów świadomości.

12. KOMPLEKSOWY MODEL UMYSŁU ŚWIADOMEGO

Żeby umysł posiadł wzorce – modele – pojęcia utrwalone w impresjonach wysokiego poziomu, musi zdobyć wiedzę o otaczającym go świecie. Może ona być wrodzona, zdobyta samodzielnie w procesie samouczenia bądź przekazana przez innych osobników społeczności, jeśli umysł należy do członka społeczności. Zasób wiedzy wrodzonej musi być ograniczony w porównaniu z innymi źródłami wiedzy, jeśli oczekujemy, że system kierowany umysłem będzie elastyczny i posiadał będzie duże zdolności adaptacyjne. Tak się dzieje w przypadku umysłów naturalnych funkcjonujących w organizmach żywych. Zwierzęta o wysokim stopniu inteligencji, proporcjonalnie w znacznie mniejszym stopniu korzystają z wiedzy wrodzonej i odruchów wrodzonych a w większym stopniu kierują się wiedzą zdobytą w trakcie rozwoju osobniczego. Opisaliśmy motywacje, które do tego prowadzą. Aby system miał zdolność samouczenia, niezbędne jest jednak spełnienie dodatkowych warunków. Konieczne jest, aby system miał zdolność oddziaływania na środowisko. Musi on być wyposażony nie tylko w sensory, ale i w efektory, urządzenia wykonawcze, które systemowi pozwolą poruszać się lub w inny sposób manipulować środowiskiem, w którym się znajduje. Efektory, motory wykonawcze, zwykle są połączone w jakiś sposób z sensorami tak, aby sensory w trakcie wykonywania manipulacji uzyskiwały lepszy ogląd rezultatów działania. To połączenie zwykliśmy nazywać ciałem, w przypadku zwierząt i ludzi, lub obudową, w przypadku maszyn i robotów. Systemy inteligentne w formie autonomicznego agenta wyposażonego w ciało lub obudowę nazywać będziemy ucieleśnioną inteligencją EI (*Embodied Intelligence*) [HURLEY 1988; NOË 2005, 2009; SPORNS & ALEXANDER 2002]. Wielu badaczy jest zdania, że świadomość i samoświadomość może uzyskać jedynie inteligencja ucieleśniona. Należy zdefiniować zatem, co uznajemy za świadomość i jak się ona ma do inteligencji systemu.

POZIOMY ŚWIADOMOŚCI

Świadomość jest pojęciem niejednoznacznym. Hierarchię stanów świadomości przedstawił między innymi Max Velmans w pracy „Jak definiować Świadomość – i jak nie definiować Świadomości" [VELMANS 2009]. Popularny opis fenomenu świadomości musi przypomnieć, że świadomość jest stopniowalna, i to w podwójnym sensie. Na poziomie fenomenalnym należy wyróżniać stan rozpoznawania bodźców, który prowadzi do świadomego odczucia. Świadome odczucie oznacza, że bodziec jest dostrzeżony, zauważony i rozpoznany. Odbywa się to w niższych warstwach sensorycznych sieci neuronowej. Relacja jest dwukierunkowa. Z jednej strony bodziec musi być uświadomiony, żeby nabrał dla świadomego umysłu jakiegokolwiek znaczenia. Z drugiej strony świadomy umysł musi coś postrzegać. Jeśli nic nie postrzega, to znaczy, że nie jest świadomy. Rozpoznanie bodźców konstytuuje najniższy stopień świadomości, tzw. świadomość percepcyjną. Stopień odczuwania bodźców także jest zmienny i odpowiada stopniowaniu tego rodzaju świadomości [WIERZCHOŃ 2013: 128–130]. Świadomością percepcyjną dysponują wszystkie organizmy wyposażone w centralny układ nerwowy, pamięć i zdolność uczenia. Wyczerpuje to warunki definicji inteligencji sformułowane przez Starzyka. Dowodem na występowanie tego typu świadomości jest elastyczna reakcja na bodźce i jej dostosowywanie do warunków w wyniku kolejnych doświadczeń.

Wyższy stopień świadomości wiąże się z uświadomieniem istnienia przestrzeni wokół podmiotu postrzegającego, tworzenia map tej przestrzeni w dowolnie zmiennych skalach, wyobrażenia sobie wzajemnych relacji przestrzennych i zdolności do planowania działań w tej przestrzeni. Wymaga to uformowania pamięci przestrzennej, geometrycznej. Występuje u zwierząt wyższych o wysokiej inteligencji. Do tego stopnia inteligencji i świadomości aspirują też roboty autonomiczne budowane sztucznie.

Następny pułap nazywany jest samoświadomością (lub świadomością introspekcyjną, według terminologii stosowanej przez Wierzchonia i prac cytowanych przez niego [WIERZCHOŃ 2013]) i dostępny jest tylko dla umysłów zdolnych do myślenia abstrakcyjnego, symbolicznej reprezentacji wiedzy oraz dysponujących językiem symbolicznym. Taki stan świadomości przypisywany jest człowiekowi. Podejrzewa się, że początki samoświado-

mości mogą się ujawniać u niektórych zwierząt dysponujących najprostszym językiem (niekoniecznie werbalnym). Zalicza się do nich małpy człekokształtne, czasem słonie, delfiny lub inne zwierzęta. Umysły sztuczne nie są dotychczas w stanie demonstrować samoświadomości.

Niektórzy badacze i filozofowie umysłu wyróżniają także stan świadomości własnych stanów mentalnych, nazywany świadomością refleksyjną. Dostępny on jest dla ludzi posiadających pełną świadomość, znaczącą wiedzę o sobie i zdolność koncentracji na własnych odczuciach psychicznych.

Wszystkie te formy nie mają ostrych granic rozróżniających je. Mogą płynnie przechodzić od jednego poziomu do drugiego. Człowiek może demonstrować różne stany świadomości w zależności od użytych środków psychotropowych, używek lub stanu zdrowia psychicznego. Bardziej złożone szczeble świadomości tworzyć się będą, jeśli agent będzie miał możliwość dalszego pozyskiwania wiedzy. To zaś oznacza, że umysł musi mieć zdolność samouczenia. Jak wykazaliśmy powyżej, zdolność samouczenia wymaga ciała zdolnego do manipulacji środowiskiem i/lub realizacji funkcji motorycznych. Po prostu autonomiczny agent może eksperymentować ze swym otoczeniem i uczyć się skutecznych zachowań, zgodnie z motywacjami, które nim powodują. Dlatego teza, że świadomość może uzyskiwać tylko umysł połączony z ciałem, czyli inteligencja ucieleśniona, wydaje się udowodniona. Oczywiście, gdy ta świadomość już zaistnieje, funkcja ciała staje się mniej istotna.

Samo pobudzenie grupy neuronów (węzła podstawowego według sugestii Kocha [KOCH 2004]) lub impresjonu reprezentującego postrzegany obiekt (według terminologii Vadakkana) nie wystarcza do uznania, że umysł jest świadomy bodźca, tak jak pobudzenie czujnika nie oznacza, że system alarmowy jest świadomy intruza. Nawet jeśli bodziec powoduje reakcję ruchową agenta autonomicznego, nie możemy twierdzić, że jego postrzeżenie i reakcja jest świadoma. Dla świadomego postrzegania konieczne jest rozumienie. Otóż rozumienie znaczeń pozyskiwanej wiedzy jest tym głębsze, im więcej modalności angażuje agent i im ich więcej asocjuje, tworząc odpowiadające tej wiedzy impresjony wyższych rzędów (reprezentacje wyższego rzędu). Proces samouczenia aktywnego agenta zdolnego do poruszania się, uruchamia korelacje międzymodalne i asocjacje impresjonów, nadające qualiom znaczenia, niedostępne dla umysłu zdolnego wyłącznie do

obserwacji. Na najwyższym poziomie zasocjowane rozlegle impresjony stanowią reprezentację otoczenia i środowiska, w którym agent się porusza. Tak więc stopniowanie świadomości bodźców na poziomie percepcyjnym związane jest z intensywnością transmitowanych bodźców, natomiast stopień organizacji i stanu świadomości zależy od zdolności do uogólniania i sublimowania wiedzy na poszczególnych poziomach jej przetwarzania w strukturze hierarchicznej impresjonów oraz od rozległości i liczby asocjowanych impresjonów.

Doświadczenia związane z ruchem tworzą uogólnione kompleksy epizodów reprezentowanych przez impresjony sytuowane w pamięci przestrzennej, pozwalającej agentowi (lub zwierzęciu) tworzyć model przestrzenny otoczenia, zapamiętywać położenie obiektów w przestrzeni, wykonywać zaplanowane manipulacje itp. System musi posiadać możliwość tworzenia scen w pamięci przestrzennej na podstawie obserwacji oraz sięgania do pamięci przestrzennej, co oznacza indukowanie impresjonów dynamicznych, związanych z doświadczeniami ruchu. Takie przypominanie jest bliskie wyobrażaniu. Manipulacja wyobrażona, dotycząca własnego położenia lub położenia innych figur, prowadzi do wiedzy o symetriach, obrotach, kształtach, czyli do wiedzy geometrycznej. Kojarzenie w trakcie uczenia doświadczeń z zakresu geometrii z symbolami geometrycznymi słownymi lub rysunkowymi prowadzi do opanowania nowego języka. Będzie to język geometrii. Pojęcia geometryczne rzeczywiście lokowane są w obszarach kory mózgowej, w których odnajdujemy także pola ruchowe. Tam też lokowana jest wiedza matematyczna i umiejętność czytania. Wykonywanie tych manipulacji geometrycznych w wyobraźni motywowane może być ciekawością. Tak więc po opanowaniu języka geometrii w procesie uczenia, możemy mieć do czynienia z dalszym doskonaleniem przestrzenno-geometrycznych umiejętności w procesie samouczenia jako skutkiem ciekawości.

Przykładem może być proces nauki jazdy na nartach. Osoba czytająca książki o narciarstwie, oglądająca filmy z zawodów narciarskich, a nawet obserwująca żywych narciarzy na stoku ma inne pojęcie o ewolucjach narciarskich niż narciarz, który je wypraktykował. Przy tym narciarz prawdopodobnie nie będzie potrafił tak przekonywająco opowiedzieć o swoich odczuciach, jak miłośnik narciarskiej literatury. Ich mózgi posługują się innym językiem, wykształconym w toku innych doświadczeń. W mózgach żywych stworzeń John C. Roder i Beechara J. Saab zaobserwowali sprzężenie po-

między zachowaniami eksploracyjnymi (co odpowiada zjawisku ciekawości) a pamięcią przestrzenną na poziomie molekularnym [SAAB i in. 2009]. Jest to doskonałe potwierdzenie Neuro-Elektro-Dynamiki i roli ciekawości jako napędu kształtowania nie tylko pamięci przestrzennej, ale i świadomości.

ZROZUMIEĆ PROCESY ROZUMIENIA

Jak umysł motywowany zmysłami i ciekawością osiąga świadomość? Dotychczas uważano, że sposób, w jaki materia uzyskuje świadomość siebie samej, jest zagadką porównywalną z zagadką istnienia jej samej. Żeby zrozumieć, jak rodzi się świadomość, należy zrozumieć, co to znaczy „rozumienie". Zgodnie z wcześniejszymi rozważaniami widoczne się staje, że zmysły i pamięć dostarczają umysłowi wciąż nowych pobudzeń, a umysł posiada potrzebę ich porównywania i uzgadniania. Jak wykazaliśmy, potrzeba uzgodnienia nowych postrzeżeń ze wzorcami-impresjonami przechowywanymi w pamięci jest immanentną cechą umysłu. Jak zakładamy, pola neuronowe nie mogą czynić nic innego, jak tylko wyszukiwać podobieństwa w docierających wzorcach pobudzeń i przekazywać informacje na wyższe poziomy, jeśli takie podobieństwo zostanie wykryte. Umysł nie może funkcjonować, nie mogąc osiągnąć stanu zgodności tworzonych przez siebie impresjonów-perceptów z impresjonami-modelami rzeczywistości. W celu osiągnięcia tego stanu gotów jest podejmować stały wysiłek poznawczy. Jeśli zaś ten wysiłek nie daje lub nie może dać rezultatu, umysł odczuwa dyskomfort, niepokój, strach, a w skrajnych przypadkach panikę i inne emocje negatywne. Jeśli zgodność wzorców zostanie osiągnięta, to umysł odczuwa emocje pozytywne. Jest to zgodne z koncepcją Perlovsky'ego przedstawioną w rozdziale 3.

- Jeśli porównanie pobudzeń i wzorców mentalnych dotyczy poziomu percepcyjnego, jesteśmy na poziomie percepcyjnym. Pozytywny wynik porównania bodźca z utrwalonymi wcześniej wzorcami oznacza rozpoznanie. Rozumiemy, co postrzegamy.
- Jeśli w obserwowanej scenie rozpoznajemy utrwalone wcześniej mapy przestrzeni, jesteśmy na poziomie świadomości przestrzennej. Rozumiemy, gdzie się znajdujemy, i co się z nami dzieje.
- Jeśli posiadamy wzorzec samego siebie i rozpoznajemy swój ślad w środowisku, w którym się poruszamy, to rozumiemy, że jesteśmy bytem oddziałującym na to środowisko. Odpowiada to samoświadomości.

- Jeśli posiadamy wyobrażenie naszej psychiki i model jej funkcjonowania oraz jesteśmy zdolni do formułowania celu naszych działań i sensu naszego istnienia, zgodnego z tym modelem, to rozumiemy, że posiadamy świadomość, która umożliwia intencjonalne działanie i formułowanie odpowiedzi na trapiące nas pytania egzystencjalne.

We wszystkich tych przypadkach, najważniejsze jest **rozumienie**.

Jeśli więc narzucamy NPM konieczność nieustannego poszukiwania podobieństw między perceptami i uproszczonymi, ogólnymi modelami, dajemy sieciom neuronowym NPM nieusuwalną motywację do poznawania środowiska i samych siebie oraz tworzenia modelu zawierającego wszystkie poznane elementy i określającego relacje między nimi. Jest ona strukturalną właściwością sieci neuronowej tworzącej neuronowe pola modelujące. Ta motywacja to „potrzeba rozumienia". To coś innego niż instynkt wiedzy. Posunęliśmy się krok dalej. Instynkt wiedzy i ciekawość, to narzędzia, którymi „potrzeba rozumienia" posługuje się dla osiągnięcia satysfakcji poznawczej – *eudajmonii*. Ta hipotez wymaga potwierdzenia eksperymentalnego w badaniach biofizycznych i neurologicznych grup neuronów tworzących wspólne pola synaptyczne. Jej potwierdzenie wymaga stwierdzenia, czy grupy neuronów, stanowiące jednocześnie pamięć i procesor przetwarzający docierającą informację, rzeczywiście zajmują się wyłącznie porównywaniem pików kierunkowych reprezentujących percepty ze wzorcami pamięciowymi i przesyłają na wyższe piętra sygnały o intensywności skorelowanej ze stopniem podobieństwa porównywanych sygnałów. A także, czy generują przy tej okazji sygnał chemiczny, elektryczny, czy też biofizyczny sygnalizujący „satysfakcję" z wysokiej wartości funkcji odpowiadającej stopniowi podobieństwa wzorców. Jest to przewidywanie pozwalające na falsyfikację postawionej hipotezy.

„Ciekawość" (eksploracja informacji napływającej z otoczenia oraz eksploracja własnych zasobów pamięciowych), działając aż do skutku, pozwala na samouczenie, zrozumienie obrazu środowiska poprzez jego porównanie z modelem świata wytworzonym w procesie samouczenia, zrozumienie własnej pozycji w tym świecie, i w ten sposób kreuje samoświadomość. Jak wskazano powyżej, to porównanie prowadzi do wysokiej wartości funkcji podobieństwa. Odchodząc od ściśle informatycznego modelu Perlovsky'ego, odpowiada to wysokiej zgodności rozkładu ładunków elektrycznych w polach synaptycznych. Ta teza byłaby spójna z opisanymi mechanizmami neuronowymi, jeżeli przyjmiemy, że taka zgodność odczuwana jest przez mózg jako stan przyjemny, w odróżnieniu od przykrości odczuwanej, gdy

takiej zgodności wzorców osiągnąć nie możemy. Nie można w tym momencie nie docenić roli procesu samouczenia. To doświadczenia całego życia, wszelka pozyskana w toku życia wiedza wypełnia pamięć, tworzy obraz rzeczywistości i model świata. Możemy ją przyswoić, jeśli ją zrozumiemy. Występuje tu pozorna sprzeczność pomiędzy nieusuwalną, immanentną potrzebą rozumienia a sytuacjami, kiedy zrozumienie nie jest możliwe ze względu na brak danych, niedostateczną wiedzę, brak czasu, zmęczenie itp. Jeśli umysł potrafi osiągnąć stan zrozumienia, że pełne zrozumienie rzeczywistości nie jest możliwe, to stara się stworzyć wyjaśnienie zastępcze, w postaci „modelu uzupełniającego" rzeczywistość. Przyjmowane są w tym celu zróżnicowane strategie:

1. Złudzenie zrozumienia – uproszczony model rzeczywistości...

Umysły zawsze posługują się uproszczonym modelem rzeczywistości. Jest to niezbędne dla dokonania kompresji danych o świecie. Pełny model nie byłby modelem, byłby po prostu odwzorowaniem rzeczywistości, co rzadko kiedy bywa użyteczne i wykonalne. Problem powstaje wtedy, gdy stosowany model nie jest adekwatny do potrzeb, a inteligentny agent nie ma świadomości tego faktu.

2. Jestem zbyt głupi, żeby to zrozumieć...

Często sytuacja przerasta możliwości poznawcze, a umysł ma tego świadomość. Nie pozostaje wówczas nic innego, jak przyjęcie do wiadomości swoich własnych ograniczeń. W tym stanie znajduje się większość znanych nam umysłów naturalnych.

3. Nic mnie to nie obchodzi...

Strategia podobna do stanowiska 2. Jednakże pogodzenie się z sytuacją następuje poprzez ignorowanie problemu. Kolejna, poważna część ludzi stosuje z dużą skutecznością tę strategię. Przyjmują ją też niektóre zwierzęta.

4. Pogodzenie się z rzeczywistością – „tak już jest..."

Także podobne podejście do stanowiska 2 i 3, jednak pogodzenie się z rzeczywistością następuje na bazie łagodnej samooceny. Także wiele zwierząt demonstruje taką postawę.

5. Wyjaśnienie mistyczne, wiara w tajemnicę…

Umysł, ambitnie poszukując wyjaśnienia, tworzy fałszywy model lub przyjmuje cudze fałszywe objaśnienia, co obserwujemy jako ucieczkę do mistycyzmu, do wiedzy sztucznie wymyślonej. To jest droga do wierzeń i zabobonów

6. Brak zrozumienia, przerażenie, lęk, choroby psychiczne...

Przedłużający się stan braku rozumienia napływających informacji i obserwowanych zjawisk budzi początkowo niepokój, a często lęk. Dzieje się tak szczególnie, gdy nie tylko nie potrafimy wyjaśnić nowo obserwowanych zjawisk, ale gdy zjawiska te stoją w jaskrawej sprzeczności z dotychczas pozyskaną wiedzą i obrazem świata. Dłuższe trwanie w stanie lęku z tym związanym prowadzi do całego szeregu chorób psychicznych i stanów patologicznych znanych psychologii i psychiatrii.

Zarysowane powyżej reakcje wskazują, jak fundamentalne znaczenie ma potrzeba rozumienia świata, który nas otacza. Dotychczasowe trudności w zdefiniowaniu wielu funkcji umysłu wynikały w dużej mierze z faktu braku zrozumienia, czym jest „rozumienie" i jaki mechanizm biologiczny za ten stan odpowiada. Brakowało też świadomości, jak istotna jest ta potrzeba umysłu i jak silnie wiąże się ona z innymi motywacjami niższego rzędu. Bo przecież brak zrozumienia powoduje dyskomfort, który wyżej nazywaliśmy bólem, a zrozumienie powoduje pobudzenie limbicznego układu nagrody i odczuwanie przyjemności. Te wymienione wyżej niedostatki dowodzą kolejny raz, jak wielkie znaczenie ma „świadomość" i „rozumienie". Arystotelesowska *eudajmonia* to satysfakcja ze zrozumienia rzeczywistości poprzez jej porównanie z formą, czyli modelem świata wytworzonym w procesie uczenia. To dążenie przejawia się w ciekawości napędzającej poznawczą/eksploracyjną aktywność człowieka. W tym dążeniu jesteśmy podporządkowani także *hedonii*, czyli bezwzględnemu dążeniu do przyjemności i unikania bólu.

SKUTKI ŚWIADOMOŚCI

Świadomość, jako stan kwantyfikowalny, skorelowana jest z innymi cechami umysłu. Wśród nich wymienić należy inteligencję, wiedzę i doświadczenie, a także zdolności komunikacyjne. Tę korelację najlepiej uwidacznia pojawiająca się i wzrastająca świadomość w rozwoju filogenetycznym i ontogenetycznym. Świadomość, a potem samoświadomość wykształciła się ewolucyjnie u zwierząt wyższych, ponieważ nawet u prymitywnych zwierząt dawała przewagę ewolucyjną. Byty wykazujące proste tropizmy musiały ustępować wobec zwierząt, które zdolne były do testowania silnie zróżnicowanego środowiska. Bogato udokumentowane procesy adaptacyjne na wszystkich szczeblach ewolucji wskazują na równoległy rozwój zdolności postrzegania układu

nerwowego oraz zachowań inteligentnych. Idąca za tym wzrastająca na ogół ruchliwość osobników odzwierciedla poniekąd tę cechę inteligencji, jaką jest ciekawość otoczenia i podążająca za nią świadomość. U zwierząt wyższych proces ten zakończył się stworzeniem gatunkowych języków komunikacji społecznej jako odpowiedź na najbardziej zróżnicowane środowisko, uwzględniając także interakcję wewnątrz i międzygatunkową. Tylko w tych przypadkach możemy mówić o powstaniu świadomości i zrębów samoświadomości. Świadomości zdolnej do obsługi wzrastającego strumienia informacji wynikającego zarówno z aktywności jak i zróżnicowania środowiska.

Zarówno w rozwoju filogentycznym i ontogentycznym możliwe jest zahamowanie poziomu świadomości na etapie świadomości percepcyjnej lub przestrzennej. Większość zwierząt nie jest w stanie osiągnąć świadomości. Część ludzi nie demonstruje świadomości refleksyjnej. Główną przeszkodą jest brak dostatecznie rozwiniętego języka symbolicznego. Przedstawione schematy rozwoju świadomości wskazują na korelację poziomu świadomości z poziomem kompetencji komunikacyjnych. Rozpoznanie niezwykłych właściwości świadomego umysłu jest niemożliwe bez wykształcenia języka wspólnej komunikacji. Podobne trudności występują w wykryciu i określeniu stopnia świadomości i inteligencji u istot żywych. Trwają niekończące się dyskusje nad samoświadomością małp, delfinów, słoni, psów, kotów i innych zwierząt, choć posiadają one podobne do ludzkich zachowania społeczne i podobny język ciała, reakcje somatyczne, a nawet mimikę, co teoretycznie powinno ułatwić przeprowadzenie stosownych eksperymentów. Można tylko wyobrażać sobie, jakie trudności komunikacyjne powstać mogą w przypadku kontaktu ze świadomą inteligencją obcą lub sztuczną, z istotą nieprzypominającą niczego, co do tej pory widzieliśmy.

Omówienie, w jaki sposób neuronowe pola modelujące generować mogą mowę i rozumieć, co się do nich mówi, przekracza ramy tego artykułu. Warto tu tylko odnotować, że hierarchiczna struktura impresjonów ma zdolność przekształcania wiedzy percepcyjnej w wiedzę abstrakcyjną o charakterze symbolicznym. To umożliwia przekształcenia logiczne tej wiedzy, a operowanie symbolami jest właśnie językiem. Zdolność rozumienia, możliwości korelacyjnego uczenia i zapamiętywania epizodów oraz asocjacje multimodalne pozwalają na rozbudowanie języka stanów mentalnych do poziomu wygodnego języka komunikacji wzajemnej. Operowanie złożonym językiem symbolicznym umożliwia formułowanie pojęć wysoce abstrakcyj-

nych, o większym stopniu ogólności. Wymaga to świadomego logicznego operowania zdaniami tego języka, które to zdania też stanowią epizody zapisane w pamięci epizodycznej ośrodka mowy. Możliwe jest także świadome kierowanie uwagi dzięki formułowanym zdaniom lub według zwerbalizowanego rozumowania logicznego. Wówczas neuronowe pola modelujące zmuszać będą „ciekawość" do poszukiwania złożonych konfiguracji kojarzonych z pojęciami znajdującymi się w centrum „uwagi". W ten sposób demonstrować się będzie świadome zaspokajanie ciekawości, co prowadzić będzie do świadomych działań poznawczych, a w konsekwencji do „rozumienia". Usprawiedliwiony wydaje się sąd, że „tyle świadomości, ile języka". Przy tym świadomość i samoświadomość nie są produktem ubocznym mowy. To mowa jest skutkiem rozwoju świadomości, co zgodne jest z naszymi obserwacjami rozwoju ewolucyjnego i osobniczego.

Prezentowany model umysłu wyklucza istnienie centrum świadomości w mózgu. Podobnie nie wyróżnia się centrum uwagi, a jedynie mechanizm przełączania uwagi, w mózgu nie ma centrum zrozumienia, a jedynie mechanizm rozumienia. Za jednym wyjątkiem: podobnie jak w procesach postrzegania wzrokowego można wskazać obszary, pola neuronowe, będące reprezentacją mentalną (impresjonami) ważnych obiektów zapamiętywanych w pamięci, tak i rozpoznawanym reprezentacjom wyższych funkcji mentalnych można przypisać impresjony stanowiące reprezentacje mentalne tych stanów. Pobudzenie tych impresjonów będzie zawsze odbierane jako odczucie przypisanych im stanów psychicznych. Lokalizacją impresjonów w mózgu są hierarchiczne neuronowe pola modelujące. Pobudzenie całego impresjonu może być indukowane poprzez pobudzenie grupy neuronów, a być może nawet pojedynczego neuronu, tak zwanego neuronu gnostycznego, będącego na szczycie tej hierarchicznej struktury [QUIROGA i in. 2013]. Takie pobudzenie skorelowane jest z odczuwaniem „rozumienia", „świadomości", szczęścia, wzruszenia, strachu, wstrętu etc. W tym sensie, sensie czystej rejestracji odczucia, możemy mówić o polach lub neuronach świadomości, szczęścia i wszystkich znanych nam stanów psychicznych. Te odczucia jednak nie rodzą się w tych polach. Ich pobudzenie jest tylko skorelowane z tymi odczuciami. Stanowisko takie potwierdzone zostało w pracach Mohamada Koubeissi i innych [KOUBEISSI i in. 2014] poprzez wykrycie ośrodka włączającego i wyłączającego świadomość.

Teraz możemy zrozumieć siłę przedstawionego modelu świadomości. Tak jak podstawowym, a być może jedynym obiektywnym prawem materializmu

jest twierdzenie: „materia istnieje", tak podstawowym prawem idealizmu, dotyczącym istnienia ducha, czyli świadomości, jest twierdzenie: „Świadomość istnieje tylko w neuronowych polach modelujących, zdolnych do rozumienia, czyli wykrywania zgodności obserwowanej rzeczywistości z ideami, które tę rzeczywistość mają odzwierciedlać, powstałymi w wyniku wcześniejszych doświadczeń". Nie jest nam znana świadomość bez sieci neuronowej. Także konstruując umysły sztuczne, będziemy musieli odtworzyć heurystykę neuronowych pól modelujących. Potrzeba rozumienia jest nieodłącznym atrybutem umysłu świadomego.

Przedstawiona teoria umysłu, opierająca się na elementach neuro-elektrodynamiki, neuronowych polach modelujących i tworzeniu impresjonów dzięki funkcjonalnemu sprzężeniu synaptycznemu, uwzględniająca motywacje w postaci potrzeby rozumienia, realizowanej dzięki instynktowi ciekawości, jest poprawną, fenomenologiczną teorią świadomości.

Widzimy więc, że świadomość kolekcjonowana jest z okruchów doświadczeń pochodzących z wielowymiarowej, multimodalnej przestrzeni, z aktywnego doświadczania pętli sprzężenia, zmysłów, analitycznego umysłu, oddziaływania na środowisko i zwrotnej obserwacji skutków tego oddziaływania. Jako sumaryczny efekt tych doświadczeń świadomy umysł jest absolutnie unikatowym, niepowtarzalnym tworem, w którym każdy element może wyznaczać jego wartości duchowe. Jest to fenomen emergentny, którego nie można zredukować do żadnego elementu lub pojedynczego procesu tutaj opisywanego. Ze względu na motywacje, które nim kierują, zdolny jest do działań intencjonalnych. Może więc formułować cel i sens własnego istnienia, a także przypisywać te wartości światu, który go otacza.

13. GŁĘBSZE ROZUMIENIE ŚWIADOMOŚCI

Funkcjonowanie pól modelujących prowadzące do rozpoznawania obiektów wyjaśnia powstawanie świadomości. Wyjaśnienie obejmuje sposób funkcjonowania mózgu jako obiektu wykazującego świadomość i podanie przepisu, jak taki obiekt zreplikować, aby te same funkcje mógł wykonywać [CHALMERS 1996]. Większość badaczy (a także osoby reprezentujące niematerialistyczne podejście do problemu świadomości) nie jest jednak usatysfakcjonowana takim prostym funkcjonalnym wyjaśnieniem. Wciąż pozostaje

pytanie, jak to się dzieje, że za pomocą 1,5 kg szarej galarety uzyskujemy świadomość? [ROSENBERG 1996]. Wyjaśnienie problemu „trudnego", czyli sposobu odczuwania świadomości, wymaga przypomnienia, że w hierarchicznej strukturze mózgu stany mentalne odpowiadające rozpoznaniu obserwowanych obiektów i sceny podlegają dalszej analizie. Wyższe pola mózgowe, działające identycznie jak pola modelujące rozpoznające percepty dochodzące z pól zmysłowych, rozpoznają te stany mentalne i kategoryzują je w pojęcia i odczucia o wysokim poziomie abstrakcji czy ogólności, tworząc impresjony najwyższego poziomu. W szczególności poczucie rozumienia obserwowanego zdarzenia na tle ogólnego modelu rzeczywistości, które osiągane być może przez wysoką wartość funkcji podobieństwa reprezentacji mentalnych zawartych w dużych obszarach kory, w których ten model jest odwzorowany i przywołany poprzez asocjacje w dostatecznie wielu aspektach, może być rozpoznawany jako głębokie odczucie świadomości tego, co się wokół nas dzieje. Wręcz wrażenia takie mogą być rejestrowane jako impresjony świadomości, odpowiedzialne za odczuwanie świadomości. Poczucie świadomości miałoby więc charakter qualiów. Wiemy, że wiedza symboliczna jest nieprzydatna, a język symboliczny nie jest w stanie opisać qualiów. Qualia trzeba odczuwać zmysłowo. Podobnie odczuwamy świadomość i rozumienie. Podobnie jak w przypadku innych qualiów możemy opisywać, jak powstają w NPM, lecz nie możemy opisać czym są. Szczególnie komuś, kto nigdy ich nie doświadczył. Natomiast skanując top-down wybrane elementy utworzonego w umyśle modelu, podobnie jak kierując świadomie uwagę wzrokową na poszczególne detale obserwowanej sceny, możemy analizować poszczególne aspekty świadomości tego, co przeżywamy i odczuwamy, uzyskując szczegółowy wgląd we wszelkie subtelności naszej świadomości.

Architektów sztucznych systemów AI gnębi wątpliwość sformułowana najdobitniej przez Davida J. Chalmersa [CHALMERS 1998; ROSENBERG 1996]: „czy maszyny mogą mieć wrażenia i odczucia?", czy mogą zrozumieć to co robią? Odpowiedź sugeruje Włodzisław Duch, dowodząc, że sztuczny intelekt, nazywany przez niego Artilektem, będzie utrzymywał, że jest świadom otoczenia i samego siebie, jeżeli będzie miał zdolność tworzenia mentalnych korelatów stanów świadomych i odróżniania ich od stanów nieświadomych [DUCH 2003a]. Sztuczny umysł realizujący heurystykę neuronowych pól modelujących będzie posiadał zdolność odczuwania wrażeń, tak zwanych qualiów. Qualia to w naszym neurokognitywnym modelu umysłu podobieństwo impresjonów, reprezentujących cechy obiektu. Aspek-

ty podobieństwa cech obiektu mogą być uogólniane i kategoryzowane jak każda inna własność obiektu oraz utrwalone w nowym impresjonie wyższego rzędu, odpowiadającym modelowi tej cechy. Każdorazowe jego przywołanie (poprzez pobudzenie, indukowanie FSS) wywoła zawsze to samo wrażenie. Umysł rozpoznawał będzie qualia.

Zastanawiając się, na ile możemy zrozumieć naszą własną świadomość, musimy przypomnieć, jaka jest istota procesu rozumienia. Impresjony wysokiego poziomu (odpowiednik reprezentacji wyższych rzędów) to rozległe drzewa hierarchicznie połączonych neuronowych pól modelujących, z milionami pól synaptycznych. Dla ich pobudzenia konieczna jest zgodność sygnału pobudzającego z większą częścią tych pól w strukturze całego drzewa stanowiącego impresjon. Pobudzenie może mieć charakter aferentny, od dołu. Wówczas docierające impresjony niższego rzędu muszą wykazywać podobieństwo do impresjonów wyższego rzędu przechowywanych w pamięci trwałej. Jeśli podobieństwo jest dostateczne, wówczas impresjony są zdolne do oddziaływania mechanizmami neuro-elektro-dynamiki i możliwa jest modyfikacja impresjonu wysokiego rzędu, reprezentującego na przykład stan rozpoznania własnej świadomości. Mogą także występować pobudzenia indukowane efatycznie, przez równoległe pola, w których ulokowane są impresjony poczucia „rozumienia", pola mowy, zawierające impresjony słów przeczytanych lub usłyszanych, odnoszące się do „świadomości". Zawsze pobudzenie takiego impresjonu odczuwane będzie jako **zrozumienie**, że jesteśmy świadomi. Te impresjony powstają w trakcie doświadczeń i są uogólnieniem odczuć oraz wyszukaniem wspólnych cech wrażeń doznawanych w trakcie doświadczania rzeczywistości. Nie muszą one być jednoznaczne i ostre, tak jak większość pojęć modeli, przechowywanych w naszej pamięci. Wiele problemów reprezentacjonizmu i samoreprezenentacjonizmu wynika z faktu, że stan świadomości nie jest jednolity. Mogą istnieć zróżnicowane impresjony odpowiadające świadomości percepcyjnej, przestrzennej, samoświadomości i świadomości refleksyjnej. Te impresjony mogą interferować ze sobą i asocjować dzięki temu, że mają postać materialną, pozwalającą na ich „dodawanie", czyli równoczesne pobudzanie szerszych lub węższych obszarów. Dzięki temu możemy mieć zróżnicowane wrażenia i odczuwanie naszej świadomości.

Usuwa to rozterki, paradoksy i sprzeczności subtelnych rozważań reprezentacjonistów, teoretyków reprezentacji wyższego rzędu lub dualizmu. Ich szczegółowe dywagacje nie zdołały w satysfakcjonujący sposób objaśnić od-

czucia świadomości, które pojawia się u świadomych istot. Poczucie takie przypisujemy dzieciom i zwierzętom, które przecież nie posiadają aparatu pojęciowego, pozwalającego rozpoznać wyższego rzędu świadomości u siebie samych, a przecież nie przeszkadza to im być świadomymi.

Te problemy teoretyczne wynikają z abstrahowania od realnych procesów biochemicznych, które za impresjony wysokiego rzędu odpowiadają i zapominaniu, że hierarchiczna struktura NPM dopuszcza inne mechanizmy konstytuowania świadomości percepcyjnej, ograniczającej się do przetwarzania impresjonów w koneksjonistycznej strukturze pól percepcyjnych, a inne mechanizmy na poziomie impresjonów wyższego rzędu. Na tym pierwszym poziomie impresjony nie podlegają operacjom logicznym, natomiast na poziomach wyższego rzędu mogą być kojarzone ze strukturami abstrakcyjnymi, matematycznymi, a więc i logicznymi, i dzięki temu podlegają logicznemu wartościowaniu i przekształcaniu. Te struktury matematyczne, geometryczne i logiczne wykształcone są w trakcie uczenia i zdobywania doświadczeń i lokowane w korze ruchowej i pamięci przestrzennej, jak to opisywaliśmy powyżej.

W niniejszym opisie zasadnicze znaczenie ma odczucie rozumienia docierających impresjonów, rozumienia, że jesteśmy świadomi rozumienia, na czym polega nasza świadomość. Bez teorii umysłu wyjaśniającej proces rozumienia nie można zrozumieć świadomości. Dzieci i zwierzęta rozumieją docierające do nich bodźce. Nie jest im potrzebny aparat analityczny rozkładający subtelności świadomości na czynniki pierwsze. Samoświadomość i świadomość refleksyjna wymaga znajomości złożonych procesów kognitywnych zachodzących w umysłach. Wielu filozofów ledwie zbliża się do zrozumienia, na czym polega ich własna świadomość i jak to się dzieje, że są świadomi. Następnie, dostrzegając symptomy świadomości u zwierząt i dzieci, rozważają, w jaki sposób one mogą być świadome, jeśli wiedzy na temat świadomości nie posiadają. Świadomość jest stopniowalna. Każdy odczuwa taki stopień świadomości, na jaki pozwala mu jego poziom mentalny, złożoność umysłu, poziom poznania świata, wynikający z doświadczeń życiowych i uczenia. Jak już pisaliśmy, stopień świadomości skorelowany jest ze stopniem rozwoju języka komunikacji. Każda istota na tyle może sformułować własne odczucie świadomości, na ile pozwala jej bogactwo języka, którym operuje. U filozofów stopień rozwoju umysłu i języka, a zatem i samoświadomości jest najwyższy, dlatego też mają oni największe trudności ze sformułowaniem swego stanowiska w tej sprawie. Pies czy kot ma świadomość swych odczuć, lecz ani myśli definiować, na czym one polegają.

Takie skrajnie redukcjonistyczne podejście, wynikające z teorii NPM-FSS, nie stoi w sprzeczności z konsekwencjami tej teorii, prowadzącymi do postrzegania świadomości jako emergentnej własności umysłu. Odczuwanie świadomości nie jest tożsame z procesem tworzenia świadomości. Istnienie neuronów lub pól „poczucia samoświadomości" nie ma nic wspólnego ze stanem świadomości, który jest emergentną cechą całego mózgu i nie ma określonej lokalizacji. Istotą świadomości jest zbudowanie modelu rzeczywistości i określenie/zrozumienie swego miejsca w tej rzeczywistości oraz odczucie satysfakcji z tego faktu. Samoświadomość to stan umysłu obejmujący wewnętrzny model samego siebie oddziałującego z wewnętrzną reprezentacją modelu otoczenia zewnętrznego, pozwalający na przewidywanie skutków własnych działań. Uczenie, poprzez planowanie własnych działań, wymaga zespolenia świadomego umysłu z ciałem, czyli obudową i mechanizmami motorycznymi, wykonującymi intencjonalne działania umysłu. Celowe działanie umysłu wynika z powstania w umyśle hierarchii wartości jako immanentnej cechy modelu rzeczywistości. Wyobrażenie skutków działania własnego i innych obiektów kształtujących tę rzeczywistość pozwala na subiektywne zdefiniowanie „dobra". Umysł rozumie, co może być „dobre" dla niego w krótszej i dłuższej perspektywie. Sformułowanie celu długofalowego konstytuuje sens istnienia samego siebie, a w konsekwencji sens istnienia świata, jako narzędzia do spełnienia swojej misji w tym świecie.

Rozwijając tę myśl konsekwentnie, należy stwierdzić, że celowość istnienia materii ma sens tylko w obecności i względem świadomości. Powszechne przekonanie o transcendentnym charakterze świadomości jest uogólnieniem doświadczenia życiowego. Dominacja tej postawy pojawia się w sposób naturalny w procesie wczesnego uczenia poprzez analogię. Jeśli wokół mnie coś się dzieje, to dlatego, że ja to robię. A robię zawsze w jakimś celu. W takim razie, jeśli coś się dzieje bez mojej ingerencji, tzn. że ktoś inny robi to w jakimś celu. Pojawia się myślenie teleologiczne. Domyślnym założeniem takiego myślenia jest istnienie świadomości transcendentnej, nadającej cel, a w takim razie i sens istnienia. Na pytanie: „w jakim celu w Bałtyku pływają dorsze?", prawidłowa odpowiedź powinna brzmieć: „po to, żebyśmy mogli je zjeść", jeśli odpowiadamy z punktu widzenia własnej świadomości. Inna, równie dobra odpowiedź brzmi „żebyśmy mogli zapanować nad całym Bałtykiem", jeśli odpowiadamy z punktu widzenia świadomości dorsza, lub „na chwałę Bożą", jeśli potrafimy w to uwierzyć.

Jest rzeczą zadziwiającą, że jeśli pytamy o cel i sens istnienia materii, to musimy odpowiedzieć sobie, że taki cel nie istnieje, jeśli nie ma świadomości, względem której taki cel moglibyśmy sformułować. Materia jednak ma zdolność samoorganizacji w taki sposób, że świadomość powstaje. Tą formą organizacji jest sieć neuronowa tworząca neuronowe pola modelujące. Można więc przeformułować odpowiedź na pytanie o cel i sens istnienia materii. Sensem i celem istnienia materii jest powstanie świadomości. Ten sens pojawia się w chwili, gdy świadomość powstanie.

14. WOLNA WOLA

Jednym z przejawów świadomości jest wolna wola niezbędna do podejmowania decyzji. Uważa się też, że wybiórcza uwaga musi być kierowana świadomością i zanim podejmiemy decyzję, już sam wybór obiektu na którym skupiamy uwagę. jest procesem wolicjonalnym [DEECKE 2012]. Badania psychologiczne udowodniły, że człowiek posiada zdolność intencjonalnej samokontroli, co jest równoważne kierowaniu się wolną wolą [HOLLAND i in. 2013]. Jak opisaliśmy w rozdziałach 10. i 11., świadoma, ucieleśniona inteligencja posiada motywacje do działania. Te zaś kreują emocje, które z psychologicznego punktu widzenia wspomagają ogniskowanie uwagi i koncentracje na procesie postrzegania i uczenia. Jak wiemy, skupienie uwagi jest niezbędne dla skutecznego zapamiętywania. Jak napisał Luder Deecke: „Bez skupienia uwagi niewiele możemy zapamiętać, a jeszcze mniej sobie przypomnieć" [DEECKE 2012]. Z własnego doświadczenia wiemy, że to czemu towarzyszą emocje, zapamiętujemy nadzwyczaj skutecznie. Badania neurologiczne i wizualizacja pracy mózgu najnowszymi technikami zobrazowania, wskazują, że podczas procesów związanych z zapamiętywaniem układ limbiczny przesyła do hipokampa impulsy odczuwane jako przyjemne, wzmacniając te procesy. Byłoby to eksperymentalne potwierdzenie tezy o sterującej roli emocji w zapamiętywaniu. Mamy tu do czynienia ze splotem funkcji wzajemnie zależnych, którymi zarządza świadomość i emocje. Już to samo tworzy pewną sprzeczność. Dodatkowo także uwaga miałaby być sterowana emocjami i świadomością. Na domiar złego wszystko to wydaje się sprzeczne z naszymi rozważaniami dotyczącymi funkcjonowania mechanizmu uwagi opisanemu w rozdziale 9., gdzie wskazywaliśmy na automatyczne kierowanie uwagi na elementy nowości i dynamiki obserwowanej sceny.

Przedstawiony kompleksowy model umysłu powinien pomóc wyjaśnić, jaka jest relacja między tymi funkcjami umysłu. Czy możemy zaprzeczyć faktom psychologicznym dotyczącym wpływu emocji na uwagę i zapamiętywanie? Oczywiście zaprzeczanie faktom nie jest rozsądne. Z modelu NPM wynika jednak nieco inny mechanizm wiążący te procesy. Jak już opisywaliśmy w rozdziale 12., porównywanie wzorców pobudzeń neuronowych w neuronowych polach modelujących NPM generuje sygnały pobudzające limbiczny system nagrody, kreując poczucie satysfakcji, zadowolenia a niekiedy szczęścia, jeśli zgodność wzorców jest wysoka. Ładunkowe impulsy kierunkowe posiadają wówczas strukturę i topologię pozwalającą na wnikanie w szlaki neuronowe „usłane" przyjaznymi proteinami i przekazywanie pobudzeń do wyższych pól neuronowych oraz pól zasocjowanych dzięki funkcjonalnym sprzężeniom synaptycznym FSS. Tam zapamiętywane są w pamięci trwałej, jeśli filtr uwagi wyselekcjonuje je do poziomu świadomości, nadając im tym samym odpowiednie znaczenie. Jak funkcjonuje ten mechanizm, opisaliśmy w rozdziale 9. Wysokie podobieństwo impresjonów percepcyjnych i impresjonów wyższego rzędu, reprezentujących modele-koncepty, oznacza rozpoznawanie i rozumienie obserwowanej rzeczywistości. To pojawiające się uczucie rozumienia, będzie kategoryzowane podobnie jak wszelkie inne postrzeżenia i wytworzy tym samym wyspecjalizowany impresjon „rozumienia". Dopuśćmy hipotezę, że ten impresjon na wysokim poziomie sprzężony jest z „ośrodkiem przyjemności" uruchamiającym mechanizm nagrody, tak jak to opisano w rozdziale 10. Pamiętając, że w mózgu występują szlaki sprzężeń zwrotnych, z dużym prawdopodobieństwem przyjąć można, że pobudzenie ośrodka przyjemności spowoduje zwrotne pobudzenie pól pamięciowych ścieżkami równoległymi do szlaku pobudzenia aferentnego, zgodnie z opisem z rozdziału 11. Zresztą te właśnie pobudzenia stwierdzono eksperymentalnie w pracy Hansa H. Kornhubera (1973) [KORNHUBER 1973]. Jakież to rozkazy może przekazywać układ limbiczny do ośrodków pamięci? Doświadczenia psychologiczne wskazują, że muszą to być sygnały skutecznego zapamiętywania. Jak jednak może być kodowany ten sygnał? Nasz model jednorodnej sieci neuronowej konsekwentnie dopuszcza jedynie najprostsze przekazywanie impulsów, porównywanie, filtrowanie i znów dalsze przekazywanie. Należy się więc spodziewać, że impulsy ośrodka nagrody pobudzają zwrotnie te impresjony, które zostały już rozpoznane i które te impulsy wcześniej wyzwoliły. Zwrotne pobudzanie powoduje powtarzanie. A powtarzanie jest podstawą zapamiętywania dzięki mechanizmom neuro-elektro-dynamiki opisanych w rozdziale 8.

Jeśli pobudzenia z ośrodka nagrody trafiają do niskich pól kory ruchowej lub pól percepcyjnych, to spowodują one powtarzanie prostych czynności przyjemnych, czasem obsesyjnie. Gdy trafiają do pól pamięci symbolicznej, ułatwiają zapamiętywanie wiedzy symbolicznej, ale także mogą generować przez indukcję FSS działania wolicjonalne, zmierzające do zaplanowanych, świadomych działań, doprowadzających do powtarzania przyjemnych, złożonych zachowań. To wspomaganie zapamiętywania przez powtarzanie, to jeszcze jedna ścieżka wspomagania mechanizmu uwagi i koncentracji na czynnościach, które wcześniej już w polu uwagi się znalazły. Tak więc motywacje generują emocje, które wspomagają ogniskowanie uwagi, którą zarządza wolna wola na poziomie świadomości. W ten sposób początkowy dylemat został rozwiązany na poziomie przepływu ładunków elektrycznych w polach neuronowych.

Powyższe rozumowanie objaśnia także sposób generowania wolnej woli. Istnienie wolnej woli wydaje się sprzeczne z tak głęboko redukcjonistycznym objaśnieniem funkcji umysłowych. Wielu uczonych, szczególnie fizyków i chemików, uważa, że przypisanie procesów umysłowych jakiemukolwiek podłożu materialnemu determinuje przebieg tych procesów i uniemożliwia materialnej istocie posiadanie „wolnej woli". Jeśli jednak przypomnimy sobie, jak przebiegają procesy myślowe, widoczne się staje, że świadomość, rozumienie własnej sytuacji i podejmowanie decyzji następuje na podstawie porównania korelatów mentalnych dziejących się i przewidywanych zdarzeń. Zasady tworzenia wzorców mentalnych przez sieci neuronowe związane są z mnogimi procesami nieliniowymi generującymi rezultaty chaotyczne, a wobec tego nieprzewidywalne [WERNDL 2009]. Jest to immanentna cecha tak złożonej struktury. Wobec tego na ogół nie jesteśmy w stanie przewidzieć decyzji podejmowanych przez mózgi. Na użytek podejmowania decyzji przez mózg, na podstawie porównywania nieprzewidywalnych korelatów mentalnych obiektów rzeczywistych i wyobrażonych, uprawnione wydaje się następujące stanowisko pragmatyczne dotyczące „wolnej woli" systemu: Analizując *a posteriori* sposób utworzenia korelatów, zapewne da się wyjaśnić, jak one powstały i stwierdzić, że system nie miał wyboru i musiał podjąć taką lub inną decyzję. Jednakże *a priori* takie przewidywanie nie jest możliwe. System może dowolnie długo analizować problem, przywoływać nieskończoną liczbę danych, wrażeń, przeczuć, sugestii, zmieniając wagę korelatu. Jego decyzja nie jest przewidywalna i zdeterminowana. „Wybory" dokonywane przez system mają charakter chaosu deterministycznego. Tak więc poczucie „wolnej woli" systemu jest jak najbardziej usprawiedliwione. Czyli świadomy umysł ma *aprio-*

rycznie „wolną wolę", lecz możemy go jej „pozbawić" *a posteriori*. Potwierdzają to eksperymenty psychologiczne, prowadzone z użyciem technik zobrazowania pracy mózgu. W eksperymentach tych wykazano, że aktywność mózgu typowa dla świadomego procesu decyzyjnego obserwowana jest ok. 200 ms przed zapoczątkowaniem czynności, z którą ta aktywność miała związek. [ULLMAN 1996; CRICK 1997]. Zwykłe odruchy nie wymagają działania świadomości. Przebiegają on dzięki bezpośredniemu sprzężeniu impresjonów reprezentujących zagrożenie dla organizmu z impresjonami motorycznymi generującymi odruchową reakcje organizmu. W tym zakresie pozbawieni jesteśmy wolnej woli. Reakcje następują automatycznie. Procesy decyzyjne jednak nawet dla prostych czynności mogą angażować świadomość. Umysł, po zapoczątkowaniu reakcji, ma jeszcze ok 200 ms na jej modyfikację lub – jak nazwał to Benjamin Libet – zgłoszenie veta. Benjamin Libet i Alfred Mele uważają też, że ten okres zwłoki czasowej może umożliwiać uczenie się i usprawnianie reakcji [LIBET i in. 1983; LIBET 1985; MELE 2005; KORNHUBER 1984; DENNETT 2005].

15. BADANIE DUSZY

Przedstawiony model wyjaśnia główne cechy ludzkiej psychiki. Wyjaśnia, jak się tworzy świadomość refleksyjna i samoświadomość, jak umysł formułuje sens i cel swego istnienia, a także sens istnienia otaczającego go świata, w jaki sposób uzyskuje wolną wolę i jak może skutecznie działać dla dobra swojego, swego gatunku i wszelkich umysłów zaklętych w neuronowych polach modelujących. Wyjaśnia, jak potrzeba rozumienia, potrzeba harmonii i piękna może tworzyć sztukę, etykę i dobro. Jak emocje nim kierujące mogą wyzwolić uczucia empatii i miłości. Wyjaśnia też, że dla spełnienia tych funkcji, dla nauczenia się wszystkiego, co dobre i szlachetne, niezbędne jest ciało i możliwość oddziaływania na otoczenie. To bezpośredni dowód na konieczność zjednoczenia ciała i ducha. Mimo że wszystkie te duchowe zjawiska wyjaśniliśmy przepływami ładunków w kanałach jonowych i ich zdegenerowanymi stanami kwantowymi, to zaprezentowany model nie jest modelem redukcyjnym. Jak wielokrotnie wskazywaliśmy, umysł i świadomość to fenomeny emergentne, nieredukowalne do zjawisk fizycznych. Następuje w tym miejscu zespolenie funkcji materialnych z duchem wykreowanym poprzez specyficzne zorganizowanie materii. Ta organizacja jest unikatowa. Jest specjalnym stanem odbiegającym od wszelkich innych form występowania materii. Żadna istota, maszyna, obca inteligencja, człowiek czy zwierzę, które dysponuje modelem środowiska, w którym się

znajduje, rozumie swoje położenie w tym środowisku, potrafi się uczyć i planować swoje działanie na tyle skuteczne, aby przetrwać, nie może się obyć bez hierarchicznej struktury neuronowych pól modelujących.

Czy wytworzona przez neuronowe pola modelujące tak złożona psychika i duchowe jej walory to metafizyczna dusza, opiewana przez wszystkie religie świata? Nie wiadomo. Zależy to pewnie od definicji. Duszy przypisuje się atrybuty nieosiągalne dla opisanych funkcji mózgu. Jedno jest pewne: świat złożonej psychiki, pragnień, marzeń, miłości, solidarności, wolności, moralności kierowany jest działaniem neuro-elektro-dynamiki w neuronowych polach modelujących NPM. Upoważnia to do postulowania jedności ducha i świadomości jako fenomenów emergentnych nieredukowalnych. Oba one pozostawać muszą w ciele. Rozdzielenie ciała i ducha oznacza śmierć systemu świadomego, którym i my jesteśmy. Po śmierci nasze impresjony nie będą już kodowane genetycznie, lecz przeniosą się do krainy wiecznych memów, gdzie będą nadal żyły w przestrzeni społecznej. Dla jednych będzie to pamięć bliskich, dla innych – wdzięczność następnych pokoleń.

Czy przedstawiony model objaśnia budowę ludzkiego mózgu? W żadnym razie nie. Zbiór hipotez pozostaje w dalszym ciągu hipotezą. Hipotetyczny model mózgu świadomego wskazuje jednak, jak może powstawać świadomość bodźców o różnym natężeniu i jak może powstawać wrażenie świadomości (jako nowy rodzaj quale) w sensie „Artilektu" Włodzisława Ducha. W naszych analizach, na tyle, na ile było to możliwe, abstrahowaliśmy od budowy struktur mózgu, licznych wyspecjalizowanych pól mózgowych i szczegółowych opisów procesów elektro- i biochemicznych. Zajmowaliśmy się wyłącznie tymi procesami i elementami budowy, które tworzą jednorodną hierarchiczną strukturę sieci neuronowej, charakterystycznej dla kory nowej i odpowiedzialnej za myślenie, rozumienie i tworzenie świadomości. Szczegółowe badanie mózgu, jego zaburzeń, patologii, innych procesów sterowania organizmem to olbrzymie pole do badań na przyszłe dziesięciolecia dla psychologów, neurologów, biofizyków i wielu innych specjalistów.

Prezentowany model teoretyczny, bazujący na teorii poznania, pozwala jednak wskazać dalsze kierunki badań, istotne z punktu widzenia funkcjonowania procesów umysłowych. Dla modelowania procesów kognitywnych konieczne będzie szczegółowe zbadanie głównych funkcji mikrokolumn, kolumn i neuronowych pól modelujących. Niezbędne badania powinny odkryć:

1. Heurystykę pobudzeń i hamowania w kolumnie neuronowej. Heurystykę selekcji bodźców, czyli mechanizmu uwagi. Istotnym elementem tej heurystyki jest tworzenie sprzężeń zwrotnych, pozwalających podtrzymywać stany aktywne wyselekcjonowanych pobudzeń i tłumienia pobudzeń konkurencyjnych.
2. Heurystykę sprzężeń zwrotnych pomiędzy warstwami sieci neuronowej, przekazywania pobudzeń impresjonów szlakami zstępującymi top-down.
3. Heurystykę asocjacji poziomych, multimodalnych impresjonów w celu budowy wieloaspektowego modelu środowiska.
4. Heurystykę tworzenia impresjonów dynamicznych w pamięci epizodycznej.
5. Mechanizm transkrypcji zmodyfikowanych protein do DNA w celu uzyskania pamięci trwałej.
6. Zaprojektowanie sztucznego neuronu, jako bloku realizującego powyższe heurystyki w przetwarzaniu równoległym. Jego łączenie w szerokie pola zdolne do przetwarzania bieżącego strumienia informacji sensorycznej i łączenie szeregowe w strukturę hierarchiczną zdolną do porównywania wzorców według mechanizmu NPM.
7. Sposoby uczenia i wychowywania inteligentnych agentów, tak, aby stali się oni wartościowymi członkami społeczeństwa.
8. Serwisowanie i konserwacja systemów samoświadomych. Zapobieganie i leczenie chorób psychicznych ludzi, zwierząt i maszyn powinno stać się w przyszłości, jednym z priorytetów służby zdrowia.

Prace w tych kierunkach rozwijają się bardzo intensywnie. Można mieć nadzieję, że przedstawiony model kompleksowy wesprze owe wysiłki, co owocowało będzie lepszym zrozumieniem umysłów naturalnych i przyspieszeniem budowy umysłów sztucznych.

16. PODSUMOWANIE

W wyniku pracy tysięcy uczonych i dziesiątek tysięcy opublikowanych przez nich w ostatnim dziesięcioleciu prac z dziedziny elektrodynamiki kwantowej, biochemii i biofizyki molekularnej, genetyki, fizjologii, badań tkanek, mózgu i umysłu dysponujemy obecnie dostateczną wiedzą, aby dokończyć rozpoczęte przez Arystotelesa 23 wieki temu dzieło i wyjaśnić, jak mózg może uzyskać samoświadomość własnego istnienia. Świadomość nie jest tajemniczym, mistycznym fenomenem wymykającym się definicjom i badaniom empirycznym. Zrozumienie jej istoty i pochodzenia wymaga nie

tylko redukcjonistycznego wyjaśnienia mechanizmów percepcji, przetwarzania informacji oraz generowania decyzji kierujących świadomym podmiotem, ale też wyjaśnienia, jakie motywacje i emocje wyznaczają cele jego działania.

Odkrycia na granicy ludzkiej wiedzy są aktem kreacji nowego postrzegania zjawisk, a więc rozbudowy, poszerzenia i pogłębienia modelu świata, który nas otacza, i jako akty kreacji często wymagają asocjacji mglistych modeli cząstkowych obejmujących fragmenty rzeczywistości, ich generalizacji i uszczegółowienia poprzez dostosowanie ogólnych idei do nowego obszaru ich stosowania. To zaś często przekracza możliwości indywidualnego umysłu i wymaga wsparcia się na ideach rozwijanych przez całą cywilizację w jej historycznym rozwoju. Zdolność tworzenia wiedzy symbolicznej, możliwej do opisania językiem logiki, umożliwia jej przekazywanie i jest fundamentem potęgi tworzonej przez człowieka cywilizacji.

Podobnie objaśnienie zjawiska świadomości, którą może uzyskać materia zorganizowana w opisany tutaj, specyficzny sposób, wymagało wykorzystania idei tworzenia modeli rzeczywistości w postaci impresjonów, poprzez neuronowe pola modelujące dzięki specyficznej heurystyce sieci neuronowych pozwalających na porównywanie perceptów dostarczanych przez nasze zmysły z modelem świata pieczołowicie budowanym przez nasze umysły w toku życiowych doświadczeń.

Podstawową cechą materii jest jej trwanie. Trwanie nie może być utożsamiane z celem jej istnienia. Niezwykła jednak właściwość materii, wyrażająca się w tym, że materia uzyskująca świadomość własnego istnienia wykazuje nadzwyczajną trwałość, przekształca tę właściwość w cel istnienia. Materia uświadomiona zdolna jest do podejmowania celowych działań przedłużających jej trwanie w wysoce zorganizowanej formie. W ten sposób duch wstępuje w materię i czyni z niej stopniowo formę nieśmiertelną.

BIBLIOGRAFIA

AYDEDE, Murat. 2009. „Is feeling pain the perception of something?". *The Journal of Philosophy* 106, 10 (October): 531–567.

BUCKNER, Randy L., William M. KELLEY i Steve PETERSEN. 1999. „Frontal cortex contributes to human memory formation". *Nature Neuroscience* 1: 211–314.

BUNZECK, Nico, Christian F. DOELLER, Ray J. DOLAN i Emrah DUZEL. 2012. „Contextual Interaction Between Novelty and Reward Processing Within the Mesolimbic System", *Human Brain Mapping* 33: 1309–1324.

CHALMERS, David J. 1996. *The Conscious Mind: In Search of a Fundamental Theory.* Oxford: Oxford University Press.

CHALMERS, David J. 1998. „On the search for the neural correlate of consciousness". W: *Toward a Science of Consciousness II: The Second Tucson Discussions and Debates*, edited by Stuart Hameroff, Alfred W. Kaszniak i Alwyn Scott, 219–230. Cambridge: MIT Press.

CRICK, Francis. 1997. *Zdumiewająca hipoteza*, przełożyła Barbara Chacińska-Abrahamowicz. Warszawa: Prószyński i S-ka.

DAYAN, Peter, i Bernard W. BELLEINE. 2002. „Reward, motivation and reinforcement learning". *Neuron* 36, 2: 285–298.

DEECKE, Lüder. 2012. „There Are Conscious and Unconscious Agendas in the Brain and Both Are Important-Our Will Can Be Conscious as Well as Unconscious". *Brain Sciences* 2(3): 405–420.

DENNETT, Daniel C. 2005. *Sweet Dreams: Philosophical Obstacles to a Science of Consciousness.* Cambridge, MA: MIT Press.

Duch, Włodzisław. 2003. „Neurokognitywna teoria świadomości". *Studia z kognitywistyki i filozofii umysłu* 1: 133–154.

FLETCHER, P.C., i Richard N. HENSON. 2001. „Frontal lobes and human memory: Insights from functional neuroimaging". *Brain* 124: 849–889,

GROSSBERG, Stephen. 1982. *Studies of Mind and Brain*. Dordrecht, Holland: D. Reidel.

GROSSBERG, Stephen, i Daniel S. LEVINE. 1987. „Neural dynamics of attentionally modulated Pavlovian conditioning: blocking, inter-stimulus interval, and secondary reinforcement". *Psychobiology* 15(3): 195–240.

HAIKONEN, Pentti O. (2007), *Robot brains. Circuits and systems for conscious machines*. Chichester: Wiley, 2007.

HOLLAND, Owen, Alan DIAMOND, Hugo Gravato MARQUES, Bhargav MITRA i David DEVEREUX. 2013. „Real and apparent biological inspiration in cognitive architectures". B*iologically Inspired Cognitive Architectures* 3 (January): 105–116.

HORVITZ, Jon C. 2000. „Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events". *Neuroscience* 96, 4: 651–656.

HURLEY, S.L. 1998. *Consciousness in Action*. Cambridge, MA: Harvard University Press.

KAKADE, Sham, i Peter Dayan. 2002. „Dopamine: Generalization and bonuses". *Neural Network* 15: 549–559.

KOCH, Christof. 2004. *The Quest for Consciousness: a Neurological Approach*. Austin: Roberts & Company Publishers.

KORNHUBER, Hans H. 1973. „Neural Control of Input into Long Term Memory: Limbic System and Amnestic Syndrome in Man". W: *Memory and Transfer of Information*, edited by H.P. Zippel, 1-22. New York, Plenum Press, 1973.

KORNHUBER, Hans H. (1984). „Attention, readiness for action, and the stages of voluntary decision". *Experimental Brain Research* 9: 420–429.

KOUBEISSI, Mohamad Z., Fabrice BARTOLOMEI, Abdelrahman BELTAGY, Fabienne PICARD. 2014. "Electrical stimulation of a small brain area reversibly disrupts consciousness". *Epilepsy & Behavior* 37 (August): 32–35.

LIBET, Benjamin, Curtis A. GLEASON, Elwood W. WRIGHT i Dennis K. PEARL. 1983. „Time of conscious intention to act in relation to onset of cerebral activity, readiness potential". *Brain* 106: 623–642.

LIBET, Benjamin. 1985. „Unconscious cerebral initiative and the role of conscious will in voluntary action". *The Behavioral and Brain Sciences* 8: 529–566.

MELE, Alfred R. 2005. "Decisions, Intentions, Urges, and Free Will: Why Libet Has Not Shown What He Says He Has". W: *Explanation and Causation: Topics in Contemporary Philosophy*, edited by Joseph Keim Campbell, Michael O'Rourke, Harry Silverstein. Boston, MA: MIT Press.

KRINGELBACH, Morten L., i Kent C. BERRIDGE. 2010. *Pleasures of the Brain*. Oxford: Oxford University Press.

KRINGELBACH, Morten L., i Kent C. BERRIDGE. 2011. „Building a Neuroscience of Pleasure and Well-Being". *Psychology of Well-Being: Theory, Research and Practice*, 1.3.

NOË, Alva. 2005. *Action in Perception*. Cambridge, MA: MIT Press.

NOË, Alva. 2009. *Out of Our Heads: Why You Are Not Your Brain, and Other Lessons from the Biology of Consciousness*. New York: Hill & Wang.

NUXOLL. Andrew M., i John E. LAIRD. 2004. "A Cognitive Model of Episodic Memory Integrated With a General Cognitive Architecture". International Conference on Cognitive Modeling (AAAI).

NUXOLL, Andrew M., i John E. LAIRD. 2012. „Enhancing Intelligent Agents with Episodic Memory, Cognitive Systems Research". W: *Proceedings of the twenty-second AAAI conference on artificial intelligence. Vancouver*, 2007 (w druku), http://faculty.up.edu/nuxoll/pubs/CogSys_NuxollLaird_ver27.pdf

OUDEYER, Pierre-Yves, Frédéric KAPLAN, i Verena V. HAFNER. 2007. „Intrinsic Motivation Systems for Autonomous Mental Development". *Evolutionary Computation, IEEE Transactions on* 11, 2.

OUDEYER, Pierre-Yves, i Frédéric KAPLAN. 2008. "How can we define intrinsic motivation ?", *The 8th International Conference on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems.*

„Pain". 2005. W: *Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Stanford, CA: CSLI Publications.

PERLOVSKY, Leonid I. 2008. „Sapience, Consciousness, and the Knowledge Instinct (Prolegomena to a Physical Theory)". *Toward Artificial Sapience*, edited by Rene V Mayorga i Leonid I. Perlovsky, 33-60. London: Springer.

Pitcher, George. 1970. „Pain Perception". *The Philosophical Review* 79(3): 368–393.

QUIROGA, Rodrigo Quian, Itzhak FRIED i Christof KOCH. 2013. „Brain Cells for Grandmother". *Scientific American* 308: 30–35.

ROSENBERG, Gregg H. 1996. „Rethinking nature: A hard problem within the hard problem". *Journal of Consciousness Studies* 3: 76–88.

SAAB, Bechara J., John GEORGIOU, Arup NATH, Frank J.S. LEE, Min WANG, Aubin MICHALON, Fang LIU, Isabelle M. MANSUY, John C. RODER. 2009, „NCS-1 in the Dentate Gyrus Promotes Exploration, Synaptic Plasticity and Rapid Acquisition of Spatial Memory". *Neuron* 63: 643-656.

SCHMIDHUBER, Jürgen. 2003. „Exploring the predictable". W: *Advances in evolutionary computing*, edited by Ashish Ghosh i Shigeyoshi Tsutsui, 579–612. New York, NY: Springer-Verlag.

SCHULTZ, Wolfram, Peter DAYAN, i P. Read MONTAGUE. 1997. „A neural substrate of prediction and reward", *Science*, 275: 1593–1599.

SCHULTZ, Wolfram, i Anthony DICKINSON. 2000. "Neuronal coding of prediction errors". *Annual review of neuroscience* 23: 473–500.

SPORNS, Olaf, i William H. Alexander. 2002. „Neuromodulation and plasticity in an autonomous robot". *Neural Networks* 15: 761–774.

STARZYK, Janusz A. 2008. "Motivation in embodied intelligence". *Frontiers in Robotics, Automation and Control*, edited by Alexander Zemliak, 83-110. I-Tech Education and Publishing.

STARZYK, Janusz A., i Dilip K. PRASAD. 2011. „A computational model of machine conscious-

ness". *International Journal of Machine Consciousness* 3.02: 255-281..
STARZYK, Janusz A. 2011a. „Motivated Learning for Computational Intelligence". *Machine Learning: Concepts, Methodologies, Tools and Applications: Concepts, Methodologies, Tools and Applications* 120.
STARZYK, Janusz A. 2011b. „Mental saccades in control of cognitive process". *Neural Networks (IJCNN), The 2011 International Joint Conference on. IEEE*.
STARZYK, Janusz A., James T. GRAHAM, Pawel RAIF i Ah-Hwee TAN. 2012. „Motivated Learning for Autonomous Robots Development", *Cognitive Science Research* 14, 1: 10-25.
ULLMAN, Shimon. 1996. *High level vision: object recognition and visual cognition*. Cambridge, MA: MIT Press.
VELMANS, Max. 2009. „How to Define Consciousness and How Not to Define Consciousness". *Journal of Consciousness Studies* 16(5): 139–156.
WERNDL, Charlotte. 2009. „What Are the New Implications of Chaos for Unpredictability?". *The British Journal for the Philosophy of Science* 60.1: 195–220.
WIERZCHOŃ, Michał. 2013. *Granice świadomości. W poszukiwaniu poznawczego modelu subiektywnosci*". Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego.
WYLLER, Truls. 2005. „The place of pain in life". *Philosophy* 80.03: 385–393.

ARCHITEKTURA ŚWIADOMOŚCI
CZĘŚĆ III: WOLA I SENS ISTNIENIA

Streszczenie

W części III przedstawiono motywacje do aktywności autonomicznego systemu inteligentnego w postaci bólu i przyjemności. Wskazano, ze kształtowanie zaawansowanych stanów świadomości takiego systemu wymaga innego rodzaju motywacji w postaci ciekawości i potrzeby zrozumienia. Przedstawiono neuronalne podłoże tych motywacji. Zaproponowano kompleksowy model umysłu świadomego, obejmujący wszystkie poziomy przetwarzania informacji. Przedstawiony model wyjaśnia główne cechy ludzkiej psychiki. Wyjaśnia, jak się tworzy świadomość refleksyjna i samoświadomość, jak umysł formułuje sens i cel swego istnienia, a także sens istnienia otaczającego go świata, w jaki sposób uzyskuje wolną wolę i jak może skutecznie działać dla dobra swojego i dobra świadomości osadzonej w neuronowych polach modelujących. Wyjaśnia, jak potrzeba rozumienia i potrzeba harmonii może tworzyć dobro i inne wartości etyczne. Jakie emocje kierujące umysłem mogą wyzwolić uczucia empatii i altruizmu. Wyjaśnia też, że dla spełnienia tych funkcji, dla nauczenia się wszystkiego, co dobre i moralne, niezbędne jest ciało i możliwość postrzegania oraz oddziaływania na otoczenie. To bezpośredni dowód na konieczność zjednoczenia ciała i ducha aby zaistniały intencjonalnie moralne byty samoświadome.

ARCHITECTURE OF CONSCIOUSNESS
PART THREE: THE WILL AND SENSE OF EXISTENCE

Summary

In the part three, motivations in the form of pain and pleasure for activity of intelligent autonomous system was discussed. It was also shown, that another kind of motivations in the form of curiosity and understanding instinct is necessary for advanced conscious state formation. Neural substrate for this motivations was indicated. Complex conscious mind model was proposed omnirelevant to all levels of information processing. The model explains major human

psyche features. The sources of self-awareness and contemplative consciousness was shown and also described how the sense and goal of existence is formulated. It explains the way how free will can appear and how the mind can work effectively for self-interest and for favor of consciousness embedded in neural modeling fields. It was described how the understanding need and harmony necessity can create the good and another ethic value. How the mind driving emotions can generate empathy and altruism feelings. The model shows, that for all these functions, for learning what is good and moral, the body equipped with both, perception and impact on surrounding is necessary. This is direct evidence that unification of the soul and body is necessary for intentionally moral, self-aware being existence.

Słowa kluczowe: motywacje, ból, przyjemność, ciekawość, instynkt rozumienia, wolna wola, cel i sens istnienia.

Key words: motivations, pain, pleasure, curiosity, understanding instinct, free will, goal and sense of existence.

Information about Author: WIESŁAW GALUS, D.Sc.— Eutherm Sp. z o.o.; address for correspondence: ul. Zachodzącego Słońca 46, PL 01-495 Warszawa; e-mail: w.galus@eutherm.eu