

# Życie wewnętrzne komputerów

Wiele jest ciekawych pytań, na które nie znamy jeszcze odpowiedzi. Wraz z rozwojem nauki liczba takich pytań nieustannie rośnie i nie mamy już nadziei, by pojedynczy człowiek mógł ogarnąć całą wiedzę dotyczącą jednej dziedziny nauki. Pełne zrozumienie nawet najmniejszej komórki żywego organizmu, składającej się z tysięcy różnych białek wchodzących z sobą w skomplikowane reakcje, może przekraczać możliwości umysłu człowieka. Początek XXI wieku ma być erą biologii tak jak początek XX wieku był erą fizyki. Zrozumienie, z czego składa się świat materii i jakie zachodzą w nim przemiany to domena fizyki, najbardziej podstawowej ze wszystkich nauk przyrodniczych. Na czym jednak polega samo rozumienie i czym jest to, co potrafi rozumieć?

Wśród wszystkich problemów badanych przez naukę jest jeden, którego rozwiązanie stanowić będzie klucz do dalszego postępu cywilizacji i zmieni ją w sposób zupełnie nieprzewidywalny. Problemem tym jest zrozumienie sposobu działania mózgu, a w szczególności zachowań inteligentnych związanych z myśleniem. Kongres Stanów Zjednoczonych docenił ten problem i dlatego ogłosił ostatnią dekadę XX wieku Dekadą Mózgu. Prezydent Bush w czasie inauguracji tej dekady odczytał proklamację zaczynającą się od słów:

Trzyfuntowa masa komórek nerwowych i ich wypustek, kierująca naszymi działaniami, jest najwspanialszym, a zarazem najbardziej tajemniczym, produktem aktu stworzenia.

Przygotowano również europejski program badań nad mózgiem. Postępy neurobiologii w ostatnich latach są ogromne, coraz więcej wiemy o biochemicznej strukturze układu nerwowego, o roli czynników genetycznych w rozwoju mózgu i ich wpływie na choroby układu nerwowego. Polska ma piękne tradycje w badaniach neurofizjologicznych (szkoła neurofizjologii prof. Jerzego Konorskiego znana jest szeroko na świecie a Instytut Nenczkiego istnieje od ponad 70 lat), neurobiologicznych, neurofarmakologicznych, neuropatologicznych i neurochemicznych, lecz wielkie braki w badaniach nad sztuczną inteligencją i modelowaniem układu nerwowego. Dziedziny te nie są wcale wymieniane wśród nauk wchodzących w skład polskiego programu badań nad mózgiem. Tymczasem Czechy wydają własne pismo „Neural Network World” i nawet mała Estonia ma swoje towarzystwo skupiające specjalistów w zakresie modelowania sieci neuronowych, „Neural Network Society”. W Polsce pierwsza krajowa konferencja „Sieci Neuronowe i ich Zastosowania” odbyła się dopiero w 1994 r. w miejscowości Kule koło Częstochowy. Konferencja była bardzo udana, zgromadziła ponad sto osób zainteresowanych problematyką sieci neuronowych i w 1995 roku powinno dojść do utworzenia krajowego towarzystwa skupiającego specjalistów w tej dziedzinie. Założono również lokalną listę dyskusyjną w Internecie.

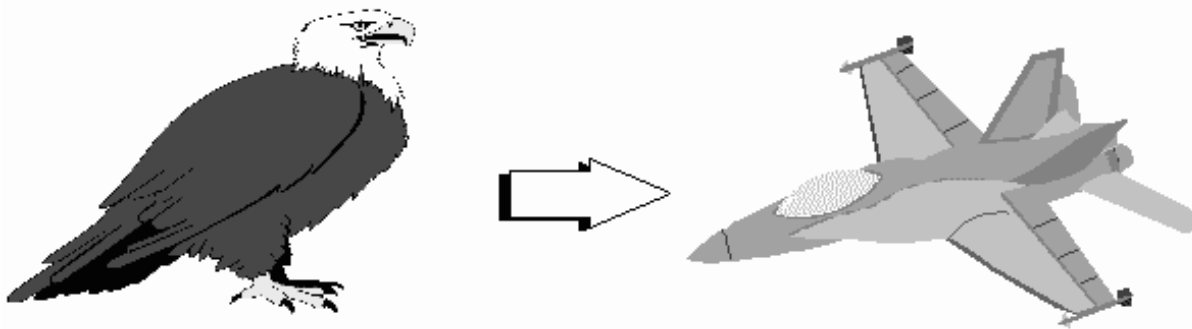
## Złożoność mózgu

O mózgu, wbrew powszechnym opiniom, wiemy już bardzo dużo. Niestety, chociaż wszystkie prawdy, które udaje się ludziom zrozumieć są proste, to proste są tylko dla specjalistów znających język, w którym łatwo jest je wyrazić. W przypadku mózgu język ten, korzystając z pojęć wprowadzonych przez wiele nauk szczegółowych, dopiero się tworzy a próba popularnego przedstawienia wiedzy współczesnej na ten temat musi być z konieczności bardzo fragmentaryczna.

Podstawowym problemem jest złożoność mózgu. Składa się on z wielu struktur i z miliardów komórek zwanych neuronami. W całym mózgu jest ich około 100 miliardów a sama kora mózgowa (neocortex) liczy sobie około 10 miliardów neuronów, z których każdy łączy się średnio z około 10 tysiącami innych neuronów, dając  $10^{14}$  zmiennych parametrów adaptacyjnych - połączeń między neuronami mózgu. Świat symboli, odczuć, wszystkich doznań człowieka, zwany przez Karla Poppera światem B, określony jest w przestrzeni tych parametrów. Wiedza zgromadzona w jednym mózgu jest jednak zaledwie niewielkim fragmentem wiedzy zdobytej przez ludzkość. Do czasu wynalezienia pisma, a więc metod utrwalania wiedzy, jej przyrost był bardzo powolny. Dlatego stopień złożoności zgromadzonej przez ludzkość wiedzy należy przemnożyć przez liczbę mózgów aktywnie uczestniczących w jej gromadzeniu i powiększaniu. Liczba ta nie przekracza miliarda, a więc całkowitą złożoność świata kultury, świata C, jak nazywa go Karl Popper, świata symboli stworzonych przez umysł ludzi można szacować z pewnym nadmiarem na  $10^{23}$  a więc około liczby Avogadro.

Ogromna złożoność całego mózgu to jeszcze nie wszystko. Pojedynczy neuron, komórka układu nerwowego, to struktura niezwykle złożona. Wiemy już ogromnie dużo o biochemicznej strukturze neuronów, o sposobach przekazywania sygnałów pomiędzy neuronami, o związkach układu hormonalnego z układem nerwowym. Zachowanie człowieka zdeterminowane jest przez sposób postrzegania świata uwarunkowany strukturą układu nerwowego, a ta z kolei rozwija się pobudzana hormonami we wczesnym okresie embrionalnym. Zakłócenia działania mechanizmów biochemicznych prowadzą do oczywistych skutków, takich jak mikrocefalia, czyli mikroskopijne, nierozwinięte mózgi, mogą też prowadzić do bardziej subtelnych zmian zachowań, np. niezgodności płci mózgu i płci genetycznej, a w efekcie do silnego pragnienia zmiany operacyjnej płci, a w słabszej formie do homoseksualizmu. Nasza filozoficzna, a tym bardziej moralna perspektywa człowieka nie wchłonęła jeszcze tych faktów. Sądzimy na podstawie pozorów: jeśli ma wąsy, to musi być mężczyzna.

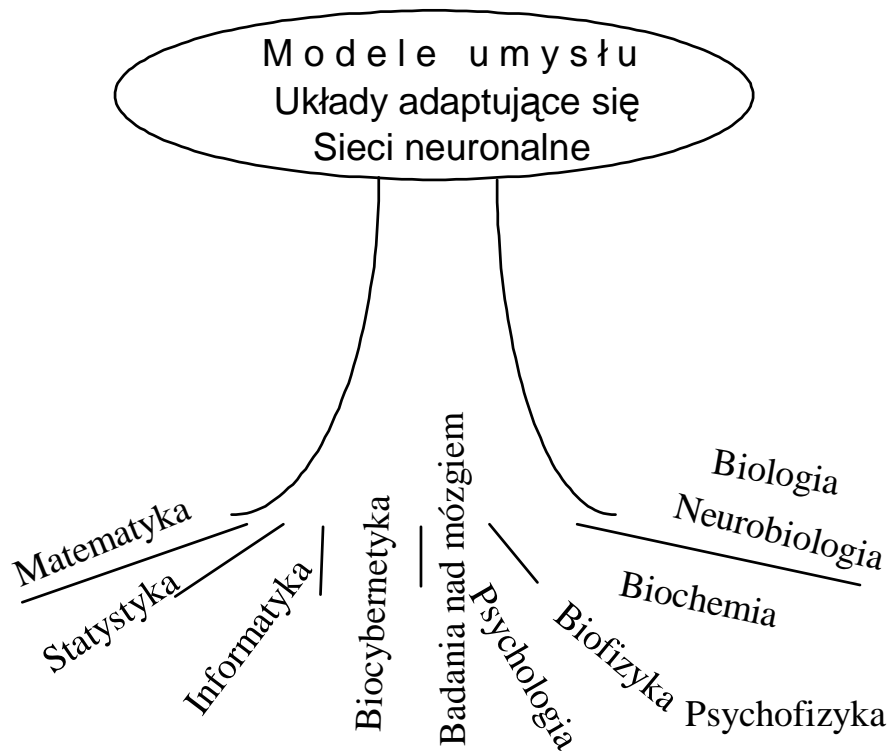
Wydawałoby się, iż w obliczu owej ogromnej złożoności biochemicznej neuronów i złożoności neurobiologicznej całego mózgu na wszelkie próby budowania sztucznego umysłu spoglądać możemy z pobłażliwością. Mikroprocesory działają co prawda bardzo szybko, milion razy szybciej niż neurony, ale złożoność mikroprocesorów jest w przybliżeniu równa liczbie ich elementów składowych, a ta wynosi „zaledwie” kilka milionów. Czynnikiem określający złożoność najbardziej skomplikowanych superkomputerów, złożonych z setek mikroprocesorów, jest miliony razy mniejszy od tego, który charakteryzuje nasz mózg. Powstaje jednak centralne pytanie: czy do powstania umysłu, do powstania świata wewnętrznego, cała ta złożoność jest istotnie konieczna? Czy nie można, znając pewne ogólne zasady działania układu nerwowego, zbudować lepiej działającego modelu? Do budowy urządzeń technicznych nie jest bowiem konieczna szczegółowa teoria. Przykładem może tu być samolot. Do tej pory nie mamy pełnej teorii lotu ptaków: nie można przecież ich lotu testować w komorach aerodynamicznych a budowę skrzydeł, upierzenie, sposób poruszania się i wszystkie inne cechy biologiczne trudno jest opisać przy pomocy prostych modeli. Kiedy bracia Wright rozpoczynali swoje próby lotów o prawach aerodynamiki wiedzano bardzo niewiele i wielu ludzi było przekonanych, że latanie jest niemożliwe. Obecnie okazuje się, że lata spadochron ślizgowy, lotnia, szybowiec, skrzydło typu delta, wiele typów rakiet. Jeszcze bardziej skrajnym przykładem odejścia od biologicznego pierwowzoru może być rower, motocykl i samochód:



pojazdy te zastąpiły konie, chociaż z ich budową nie mają nic wspólnego. Czy możliwe jest zbudowanie sztucznego umysłu w oparciu o uproszczoną teorię, którą obecnie dysponujemy?

Rozumiemy już wprawdzie pewne bardzo szczegółowe mechanizmy związane z procesami uczenia się, analizy obrazu przez układ wzrokowy, analizy mowy przez układ słuchowy być może jednak wiedza ta, z technicznego punktu widzenia, dla stworzenia sztucznej inteligencji nie jest konieczna. Warto tu posłużyć się inną analogią: zjawiska chemiczne, w tym również biochemiczne, zredukować można do procesów fizycznych opisywanych przez mechanikę kwantową. Chemicy, a tym bardziej biolodzy, nie stosują jednak na codzień metod mechaniki kwantowej gdyż mają do czynienia ze zbyt złożonymi problemami, by rozwiązać je za pomocą metod ścisłych. Modele fenomenologiczne, oparte na empirycznych obserwacjach dotyczących klasyfikacji zjawisk, nie dają się w prosty sposób zredukować do podstawowych teorii fizycznych. Czy można takie modele fenomenologiczne zastosować również do umysłu? Podjęto wiele takich prób, poczynając od psychoanalizy i wyrosłych z niej kierunków, jednak fenomenologia całkowicie oderwana od rzeczywistości biologicznej nie pozwala na stworzenie użytecznych teorii szczegółowych (psychoanalitycy oczywiście nie zgodziliby się z taką oceną).

Inne spojrzenie na umysł, na złożoność świata koncepcji, daje nam psychologia. Człowiek rozróżnia od kilku do kilkudziesięciu tysięcy symboli. W przypadku bardziej wyrafinowanych umysłów możemy więc mówić o „przestrzeni konceptualnej” złożonej z około 100.000 symboli. Jeśli każdy z tych symboli przyjmowałby tylko 2 wartości to liczba ich możliwych zestawień byłaby równa liczbie rogów hipersześcianu o 100.000 wymiarów,



czyli  $2^{100000}$ , tj. około  $10^{30000}$ . Liczby te są niewyobrażalnie duże, dla przykładu liczba cząstek we Wszechświecie nie przekracza  $10^{100}$ . Przestrzeń konceptualna jest więc ogromna, z praktycznego punktu widzenia nieskończona i prawie pusta, gdyż większość rogów hipersześcianu nie odpowiada żadnej sensownej koncepcji, której warto przypisać jakąś nazwę.

W jaki sposób tworzą się koncepcje dotyczące rzeczywistości? Można na to patrzeć od dwóch stron: różnicowania się rzeczywistości wyłaniającej się z nicności w procesie tworzenia się układu nerwowego u niemowląt z jednej strony, i doświadczenia całości istnienia, doświadczenia określonego przez psychologów transpersonalnych jako „oceaniczne”, pojawiające się w czasie stanów medytacyjnych. Rzeczywistość subiektywna, odbierana w takich stanach całościowo, rozpada się na różne modalności: dźwięki, zapachy, smaki, dotyk, wrażenia koloru, kształtu, ruchu.

Uczenie się mowy znacznie przyspiesza proces utrwalania klasyfikacji wrażeń i pojęć. Klasyfikacja jest jednym z podstawowych zadań młodego umysłu, stąd najczęstsze pytania to: „jak to się nazywa” i „co to jest”. W mózgu mamy dostatecznie dużo neuronów, których połączenia tworzą się w pierwszych latach życia, w okresie intensywnej nauki, by dokonać takiej klasyfikacji. Ilustruje to znakomicie rysunek wycinków kory mózgowej odpowiedzialnej za bodźce wizualne. Połączenia neuronów są dość plastyczne lecz jeśli mózg nie uczy się nowych rzeczy to usiłuje on zoptymalizować swoją budowę usuwając zbędne połączenia i zbędne neurony. Przeciennie między 20 a 70 rokiem życia tracimy z tego powodu około 20% masy swojego mózgu! Jest to wynikiem naturalnej tendencji organizmu do optymalizacji zużycia energii. Zagadnienia te dyskutowane są w specjalnym numerze „Świata Nauki” z listopada 1992 roku, stanowiącym bardzo dobre wprowadzenie do tematyki „umysł a mózg”.

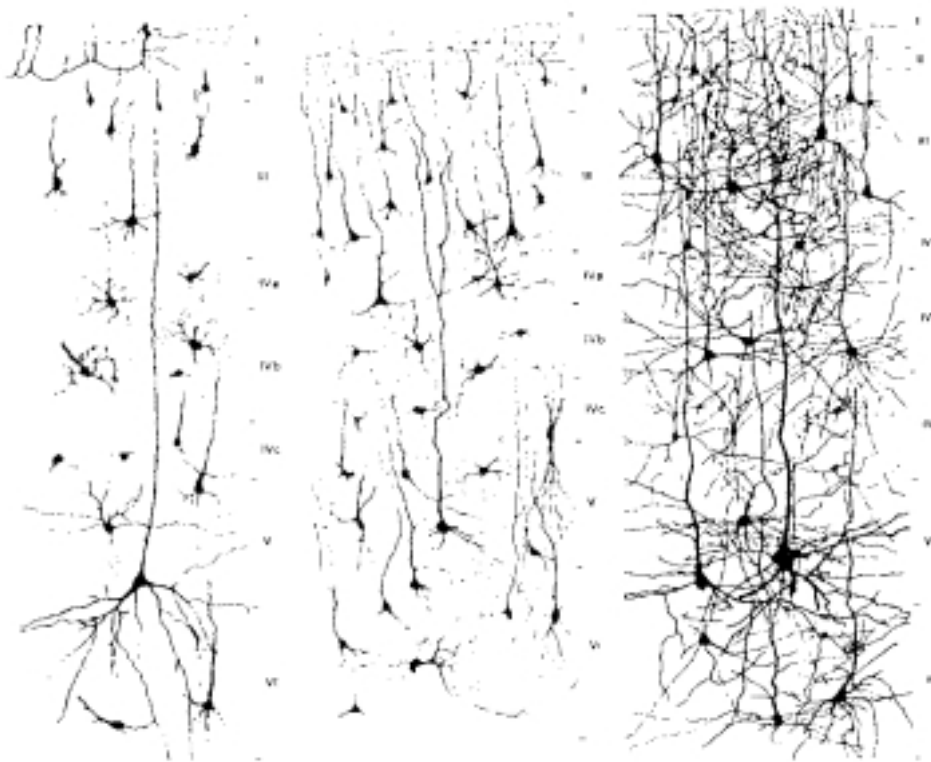
Komputery były często przyrównywane do mózgu ale nie jest to właściwe porównanie. Znacznie lepszą metaforą lub modelem mózgu jest układ adaptujący się, w szczególności model fragmentu sieci neuronalnej. Jest to system, który ma dostatecznie dużo parametrów wewnętrznych by przez ich modyfikacje nauczyć się odwzorowania dochodzących do niego danych wejściowych (bodźców) na pożądane wyniki. Układy takie wytwarzają w swoim wnętrzu pewną reprezentację rzeczywistości. Zanim jednak do tego dojdziemy przyjrzyjmy się nieco bliżej możliwościom umysłowym człowieka.

## Ograniczenia umysłu człowieka

Historia rozwoju kultury jest bardzo krótka. Zapisana historia liczy sobie około pięć tysięcy lat. W miarę upływu czasu ludzkość tworzyła coraz bardziej złożone sztuczne środowisko życia, do którego nie wszyscy ludzie potrafili się przystosować. Niektórzy badacze, np. prof. Julian Jaynes z Uniwersytetu w Princeton, twierdzą nawet, że świadomość ludzka rozwinęła się w czasach historycznych i w najstarszych przekazach pisemnych, pochodzących z czasów, kiedy Bóg rozmawiał z człowiekiem a diabły i anioły widywano często, dopatrują się zupełnie odmiennego widzenia świata, dominacji działań nieświadomych. Zauważmy, jak złożone są formy życia społecznego niektórych zwierząt, których nie uznajemy za świadome. Początki tworzenia zorganizowanych społeczeństw mogły przebiegać nieświadomie. Gdyby świadomość zrodziła się we wcześniejszym okresie cywilizacja rozwinęła by się również wcześniej. Ocenia się, że w Stanach Zjednoczonych około połowa ludności to funkcjonalni analfabeci. Popularność prymitywnych sposobów rozumienia rzeczywistości, poglądów antynaukowych i fundamentalistycznych, w znacznej mierze wynika z trudności z jakimi wielu umysłem przychodzi zrozumienia świata, którego obraz stał się zbyt złożony. Rozwój cywilizacji przebiega zawsze na granicy możliwości pojmowania przez umysł człowieka.

Jeśli mamy do czynienia z człowiekiem o uszkodzonym mózgu, jeśli przyroda nie da nam do dyspozycji 10 miliardów neuronów a tylko miliard, jak to się dzieje w przypadku wad rozwojowych, wyższe czynności psychiczne, takie jak kojarzenie faktów, w zasadzie zanikają. Częstość chorób psychicznych, objawów nieprawidłowości w funkcjonowaniu naszej maszynerii mózgowej, depresji, manii, samobójstw i schizofrenii ciągle rośnie (świadczą o tym dane epidemiologiczne z ostatnich kilkudziesięciu lat). Wiele z tych chorób związana jest z mechanizmem przekazywania sygnałów pomiędzy neuronami przez neurotransmitery. Stosunkowo proste związki chemiczne, takie jak węglan litu, skutecznie leczący stany maniakalne, działają na procesy synaptyczne. Mniej groźne zaburzenia, prowadzące do dysleksji i podobnych problemów, występują u znacznej części społeczeństwa i związane są prawdopodobnie z błędami powstałymi w okresie rozwoju biologicznego mózgu. Właściwie trudno jest znaleźć mózg bez wad.

Nawet najlepiej działające mózgi podlegają jednak silnym ograniczeniom. Niektórzy ludzie potrafią nauczyć się długich list słów korzystając z metod mnemotechnicznych, nie wydaje się jednak, by prowadziło to do sensownych generalizacji. Biegłe posługiwanie się językiem o ponad 100 tysiącach słów nie wydaje się możliwe i trzeba sobie uświadomić, że nasze indywidualne odwzorowanie rzeczywistości nigdy nie stanie się dużo subtelniejsze. Nasza klasyfikacja zjawisk, regiony decyzji określające używane pojęcia, pozostaną nie ty



toporne ale i dość dowolne, zależne od punktu widzenia i sposobu aktywnego oddziaływania człowieka na rzeczywistość. Słyszysz się czasami plotki na temat rzekomego wykorzystania tylko niewielkiego procentu naszych możliwości umysłowych. Są nawet całe organizacje obiecujące, że dzięki ich technikom nauki czy serii wtajemniczeń na kursach przez nie organizowanych nauczymy się lepiej wykorzystywać swój umysł. Niestety, są to jedynie plotki i łatwo się przekonać, że intensywne myślenie absorbuje nasz umysł na tyle, iż zatrzymujemy się bez ruchu i przestajemy zwracać uwagę na otoczenie. Niewielka podzielność uwagi, jaką człowiek dysponuje, jest najlepszym dowodem ograniczonych możliwości przetwarzania informacji przez mózg. Polepszenie zdolności koncentracji pomaga oczywiście w myśleniu i nauce, istnieją jednak pewne naturalne bariery, podobnie jak w sporcie, które trudno jest umysłowi ludzkiemu przekroczyć. Umysł człowieka, produkt ewolucji nastawionej na przetrwanie, a nie obiektywny opis świata, ma liczne ograniczenia. Przede wszystkim nasza wyobraźnia ograniczona jest do trójwymiarowej przestrzeni i jeśli mamy do czynienia z danymi, zależnymi od czterech lub więcej parametrów (a więc wymagającymi geometrycznej interpretacji w czterech lub więcej wymiarach), nie jest nam łatwo dostrzec powiązań między nimi. Podobnie jest z wykrywaniem pewnych cech rzeczywistości: jeśli mamy korelację więcej niż kilku parametrów trudno jest nam dane zjawisko „ogarnąć umysłem”. Z badań psychologicznych wynika, że większość ludzi może zdawać sobie sprawę z 7 elementów jednocześnie, przy czym w wyjątkowych wypadkach może to być do 9 elementów. Tymczasem nawet małe modele neuropodobne, wykorzystujące kilkanaście lub kilkaset pseudoneuronów, radzą sobie lepiej niż mózg, z jego 10 miliardami neuronów, z niektórymi problemami wymagającymi kojarzenia abstrakcyjnych danych. Mózg jest zdumiewająco mało efektywny przy kojarzeniu abstrakcyjnych danych. Nauka tabliczki mnożenia w pierwszej klasie szkoły trwa wiele miesięcy, nauka czytania, a więc rozpoznawania kształtów liter, podobnie.

Skąd się biorą te ograniczenia ludzkiego umysłu? Podobnie jak i ograniczenia w naszych możliwościach fizycznych, są one wynikiem budowy biologicznej. Neurony nie są najlepszej jakości, przekaz sygnałów bardzo skomplikowany, kodowanie impulsowe informacji nie jest optymalne, skomplikowany jest ich sposób odżywiania się ... Krótko mówiąc, „dobry Bóg już zrobił co mógł, teraz trzeba zawołać fachowca”. Przyroda wytworzyła z budulca, którym dysponowała, cudowne struktury, ale z punktu widzenia inżyniera niepotrzebnie skomplikowane. Jeśli zastąpić elementy biologiczne elektronicznymi zachowując podobną organizację całości możemy mieć nadzieję, że tak zbudowany system będzie miliony razy szybszy od ludzkiego mózgu a jego wewnętrzna reprezentacja rzeczywistości pozbawiona będzie ograniczeń układów biologicznych.

## Przykład klasyfikacji przez układy neuronowe

Kiedy patrzymy na fragment rozwijającej się kory mózgu - na rysunku 2 przedstawiono neurony kory wzrokowej człowieka w wieku 3, 15 i 24 miesięcy od momentu poczęcia - widać, jak z niezróżnicowanego świata pierwotnych wrażeń zaczynają się wyłaniać kształty i coraz bardziej subtelne klasyfikacje. Człowiek uczy się rozróżniać coraz bardziej subtelne kształty, kolory, ruch, a każdy jego postęp w tym względzie utrwalany jest w strukturze połączeń kory wzrokowej. Struktura sieci neuronowej mózgu zdeterminowana jest przez czynniki genetyczne tylko w najbardziej ogólnym zarysie, dopiero wpływ otoczenia określa jej szczegóły. Doświadczenia na zwierzętach potwierdzają, że pozbawienie możliwości widzenia noworodka w pierwszych tygodniach życia może w nieodwracalny sposób wpłynąć na formowanie się połączeń neuronów i w efekcie spowodować całkowitą utratę wzroku, pomimo sprawnego działania samych oczu.

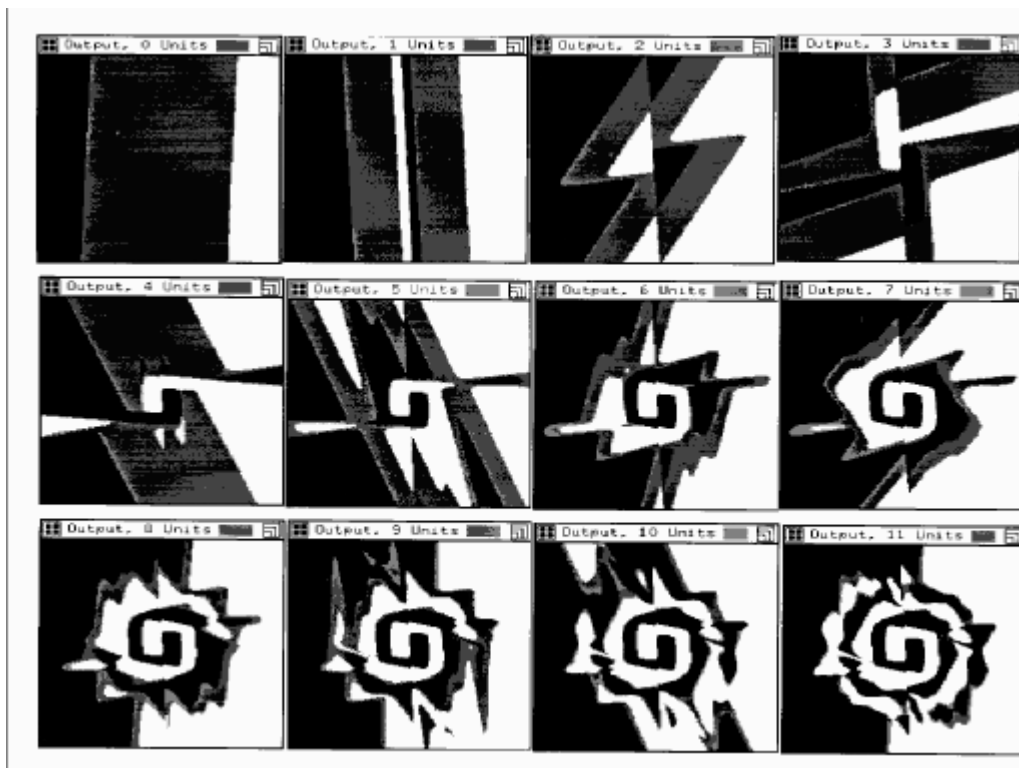
Zwykle zakłada się, że neurony działają progowo, tzn. sumują dochodzące do ich ciała sygnały i jeśli całkowite pobudzenie w danym momencie przekracza pewien próg wysyłają sygnał dalej. Jak wygląda klasyfikacja, formowanie sobie obrazu świata, przy pomocy takich sieci?

Można to prześledzić na przykładzie klasyfikacji dwóch zbiorów punktów. Przypuśćmy, że mamy około 100 punktów leżących na płaszczyźnie należących do jednej klasy (np. jaśniejszych) i tyle samo punktów należących do drugiej klasy (np. ciemniejszych). Zadaniem układu adaptującego się jest nie tylko nauczyć się klasyfikacji punktów, lecz również w sensowny sposób uogólnić klasyfikację na pobliskie punkty. Możemy to szczegółowo prześledzić na przykładzie problemu dwóch spiral - jeśli punkty, należące do danej klasy układają się na spirali lub innej prostej figurze geometrycznej, łatwo jest przewidzieć, do której klasy przypisać punkty na rysunku (por. Rys. 3). W jaki sposób przedstawić można proces klasyfikacji, dokonywany przez sieć neuronalną? Modyfikując wewnętrzne parametry układu adaptującego się - w tym przypadku typowego modelu sieci neuronalnej - rozbijamy pierwotną całość na coraz bardziej wyrafinowane rejony decyzji.

Na serii przedstawionych na Rys. 4 obrazków widzimy przykładowy przebieg procesu klasyfikacji przy użyciu algorytmu, który zwiększa stopniowo liczbę neuronów poświęconych temu zagadnieniu (inne algorytmy prowadzą do podobnych serii obrazków, chociaż końcowa klasyfikacja może nie być tak dobra). Początkowo widać rozróżnienie tylko na dwa obszary; w miarę dodawania neuronów wewnętrzny obraz rzeczywistości staje się coraz bardziej wyrafinowany, mamy coraz więcej płaszczyzn decyzji pozwalających na dokładniejsze odróżnianie elementów rzeczywistości. Płaszczyzny decyzji pozostawiają nieskończone obszary daleko od danych punktów, obszary, w których w zasadzie nie mamy podstaw, by dokonać klasyfikacji, jednak sieć tego typu nie może się powstrzymać przed przypisaniem ich do jakiejś klasy. Cóż, znamy takie przykłady również wśród ludzi, którym wydaje się, iż wszystko wiedzą.

Jeśli punkty te należą do wielowymiarowej przestrzeni zadanie jest dla nas dużo trudniejsze. Światopogląd człowieka na poziomie abstrakcyjnym kształtuje się podobnie, przez tworzenie regionów decyzji. Każdy nowy symbol czy pojęcie, które sobie przyswajamy, wymaga wyodrębnienia się nowego obszaru klasyfikacji w naszej wewnętrznej przestrzeni, wymaga zbadania relacji nowego pojęcia względem innych, już znanych. Tworzenie obszarów decyzji, „układanie sobie w głowie” nowej informacji, jest jednym z najważniejszych sposobów kształtowania wewnętrznego obrazu rzeczywistości.

Co stoi na przeszkodzie budowania dużych modeli neuronowych? Obwody scalone projektowane były z myślą o szybkości, a nie złożoności. Układ nerwowy nie pracuje zbyt szybko, jego czas charakterystyczny wynosi około milisekundy, prawie milion razy więcej niż czas cyklu szybkich układów półprzewodnikowych. Z drugiej strony złożoność projektowanych obwodów scalonych była do niedawna zbliżona do równoważnej im liczby tranzystorów. Potrzebne są obwody wolniejsze ale dopuszczające dużo większą liczbę wewnętrznych połączeń. International Computer Science Institute w Berkeley zapowiada na 1995 rok ukończenie neuronalnego superkomputera CN1 opartego na milionie elementów przetwarzających posiadających miliard połączeń. Chociaż do złożoności ludzkiego mózgu jeszcze mu daleko superkomputer neuronalny powinien pozwolić, w takich zastosowaniach związanych z klasyfikacją jak rozpoznawanie mowy czy pisma ręcznego, na zmniejszenie stopnia błędów do poziomu człowieka. Powinien pozwolić również na eksperymenty z dość



Rys. 4 Tworzenie się wewnętrznej reprezentacji faktów dla zagadnienia dwóch spiral

## Uczenie się

Nie chciałbym stwarzać wrażenia, że powstanie sztucznej inteligencji to tylko kwestia skali symulacji układów neuropodobnych lub liczby parametrów do adaptacji, chociaż w niektórych zagadnieniach jest to sprawa kluczowa. Alternatywne, bardziej klasyczne podejścia do sztucznej inteligencji oparte są na systemach doradczych (ekspertowych), wykorzystujących najczęściej „logikę pierwszego rzędu”. Wewnętrzny model rzeczywistości zawarty jest wówczas w regułach, czyniących ze zbioru faktów wiedzę. Ocenia się, iż dla realizacji zdrowego rozsądku potrzeba około 100 milionów reguł i projekt tworzenia takiego systemu (znany pod akronimem C.Y.C) jest już bardzo zaawansowany. W ostatnich kilku latach poczyniono jednak ogromne postępy w sposobach uczenia układów adaptujących się i ta droga do stworzenia sztucznej inteligencji wydaje mi się bardziej naturalna. Wyróżnia się 2 podstawowe sposoby uczenia się: bez nadzoru, polegający na samoczynnym wykrywaniu regularności w danych, oraz uczenie z nadzorem, polegającym na korekcji błędnych odpowiedzi.

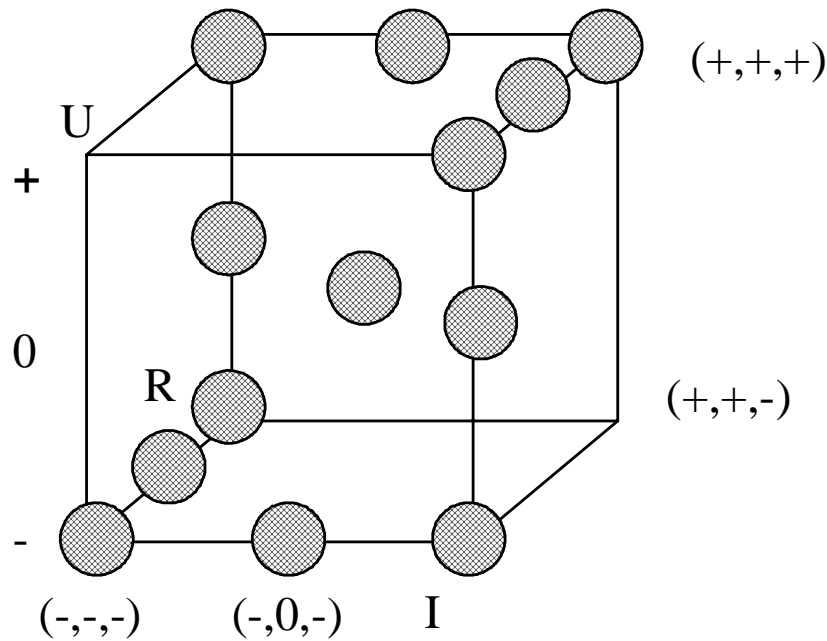
Uczenie bez nadzoru to uczenie się samodzielne, dominujące szczególnie w pierwszych latach życia. Prowadzi ono do stosunkowo szybkiego rozwoju struktury połączeń neuronów kory mózgu utrwalając wewnętrzny obraz regularności dochodzących do nas danych zmysłowych. Efektem jest powstanie intuicji, umiejscowienie się pewnych wrażeń w relacji do innych i ich klasyfikacja. W późniejszym okresie życia uczenie bez nadzoru gra również ważną rolę w formowaniu się opinii, zachodzi jednak znacznie wolniej, a u niektórych osób zanika. Wiedzę, w ten sposób zdobytą, można w dobrym przybliżeniu określić jako wiedzę statyczną, związaną z intuicją. Czym bowiem jest intuicja? Trudnym do zwerbalizowania procesem polegającym na „wycuciu sytuacji”, na ocenie faktu opartej jedynie na zgromadzonej wiedzy, a więc na wewnętrznym obrazie świata. Można postawić następującą hipotezę: intuicja jest rezultatem oceny opartej na porównaniu danego faktu czy sytuacji z koncepcjami utrwalonymi w przestrzeni wewnętrznej; porównaniu nie wymagającym żadnego rozumowania a jedynie klasyfikacji, dokonywanej przez mózg w zupełnie nieświadomy sposób. Intuicje mogą oczywiście być całkowicie błędne.

Uczenie się z nadzorem polega na bezustannej korekcji błędnych odpowiedzi przez zewnętrznego nauczyciela, którym może być człowiek, tekst książki lub program komputerowy. Z technicznego punktu widzenia uczenie się z nadzorem jest stosunkowo łatwe i polega na budowaniu aproksymacji do nadchodzących danych. Modele sieci neuronowych nadają się do tego, gdyż dostarczają dużą liczbę zmiennych parametrów (parametrów adaptacyjnych), dzięki którym do danych można dopasować dowolną funkcję. W konkretnych zastosowaniach, do problemów technicznych i w naukach przyrodniczych, w większości przypadków lepsze rezultaty osiągnąć można stosując specjalne modele matematyczne, nie wymagające tak wielu parametrów. Zaletą sieci neuronalnych uczonych z nadzorem jest jednak ich uniwersalność. W konkretnych przypadkach również możliwości uczenia się i przewidywania człowieka da się udoskonalić stosując specyficzne modele matematyczne (np. symulacje komputerowe) do opisu zjawisk.

## Model FSM

Obydwa sposoby uczenia się zilustruję na przykładzie rozwijanego od niedawna w naszym zespole modelu FSM (Feature Space Mapping). Model ten pozwala na bezpośrednie uchwycenie złożonych koncepcji, tworzenie asocjacji, samoorganizację wiedzy a także korzystanie z praw ogólnych i rozumowanie w oparciu o takie prawa. Można go traktować jako model neuronowy, oparty na neuronach działających w sposób zlokalizowany, wykrywających pewne cechy w docierających danych, można go też traktować jak system doradczy posługujący się logiką rozmytą. Jest on pewnym przybliżeniem do prawdziwej dynamiki działania mózgu, z przyczyn technicznych trudnej do bezpośredniego modelowania.

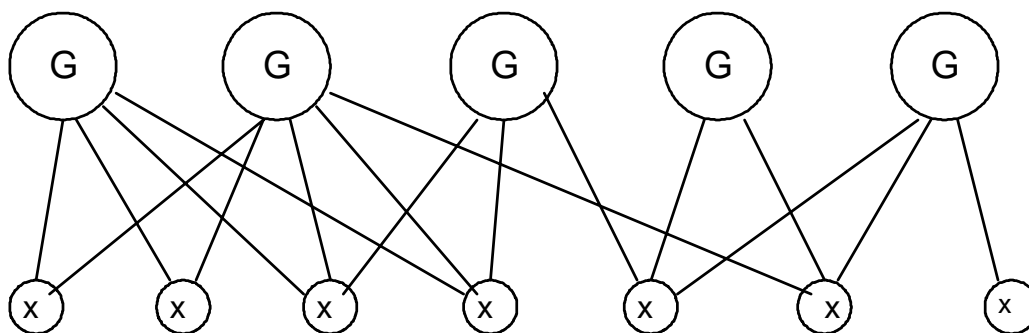
Model FSM rozpatrywać można na dwóch poziomach: koncepcyjnym, odpowiadającym wewnętrznej przestrzeni wyobrażeń, czyli umysłowi, i technicznym, odpowiadającym mózgowi. Na poziomie koncepcyjnym definiujemy przestrzeń koncepcji przez dodawanie dla każdej nowej koncepcji czy symbolu nowego wymiaru. nym na ich oznaczenie; w przestrzeni koncepcji możliwe są również pojęcia czy fakty, trudne do zwerbalizowania, stanowiące mieszanekę doznań. Przestrzeń możliwych koncepcji jest bardzo duża i prawie pusta, tylko w nielicznych obszarach pojawiają się w niej fakty. Na poziomie technicznym (Rys. 5) jest to układ wielu węzłów (neuronów) połączonych z wejściami i wyjściami. Pojawienie się jakichś cech wzbudza w węzłach G opisujących koncepcje najbardziej prawdopodobne brakujące cechy opisu, a więc możliwe dopełnienia danej sytuacji.



Rys. 5 Przykład sieci węzłów realizujących model FSM: koncepcje reprezentowane są przez węzły G, cechy obiektów przez sygnały wejściowe x. System dokonuje klasyfikacji i rozpoznaje obiekty na podstawie częściowego opisu.

Koncepcja napięcia wiąże się z wieloma innymi koncepcjami: pola elektrycznego, potencjału, siły, wielkości, zmiany. Prawo Ohma wiąże bezpośrednio jedynie te dwie ostatnie cechy napięcia z natężeniem i oporem. Wielkość napięcia równa jest wielkości natężenia razy wielkość oporu. Intuicyjnie oznacza to, że jeśli dwie wielkości w prawie Ohma są stałe to trzecia jest również stała, jeśli rosną lub maleją to trzecia też. Część intuicji, związanych z prawem Ohma, przedstawić można na rysunku jako fakty na krawędziach sześcianu, którego boki reprezentują tendencje do zmian wartości napięcia, natężenia i oporu. Taka reprezentacja prawa Ohma i podobna reprezentacja prawa Kirchoffa pozwala na analizę jakościową bardziej złożonych obwodów elektrycznych.

Oparcie się na reprezentacji symbolu jako skojarzenia wielu czynników, wielu cech prymitywnych, odnoszących się do wrażeń oznacza, że będziemy korzystali z iloczynu zbiorów, a w modelu FSM z nakrywania się



Cechy, własności, koncepcje prymitywne

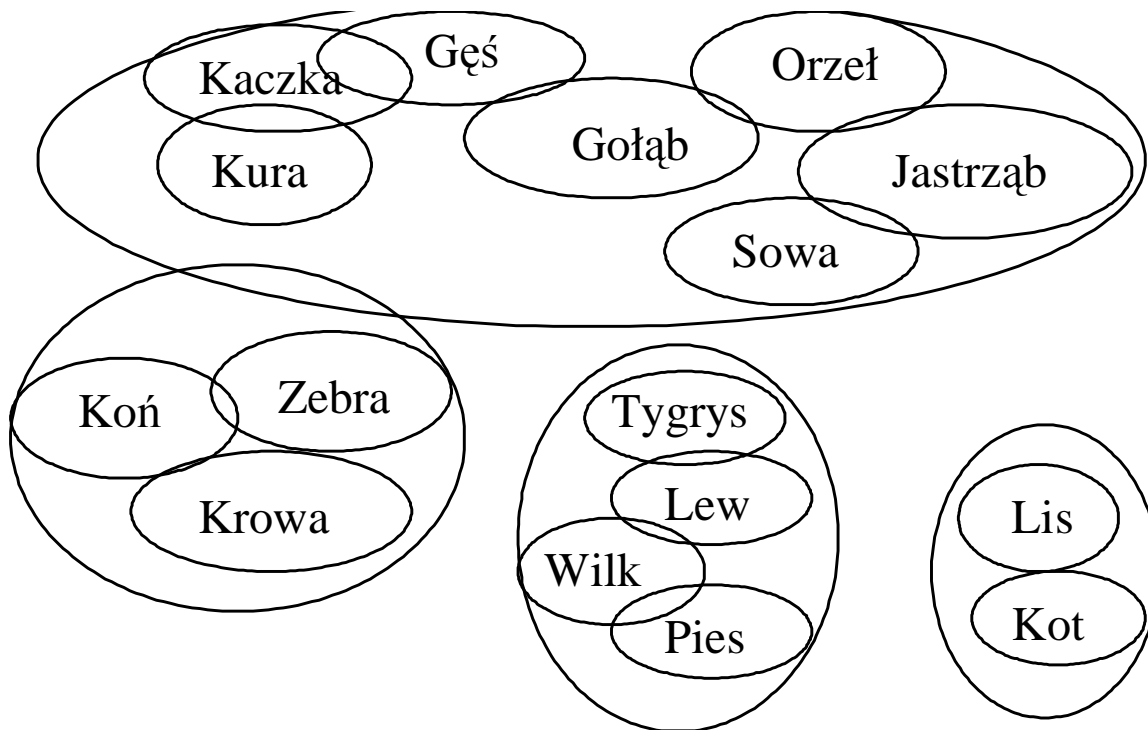


rozmytych funkcji niezależnych zmiennych opisujących każdy z wymiarów. Z technicznego punktu widzenia neuron czy grupa neuronów odpowiedzialna za daną koncepcję nie ma połączeń z wejściami, które dla tej koncepcji nie są istotne.

Model FSM dopuszcza również budowanie wielopoziomowych hierarchii koncepcji oraz stopniowe skupianie się na istotnych faktach. Wielopoziomowe hierarchie wiążą się z istnieniem wspólnych cech i różnymi stopniami abstrakcji. Na przykład, w przestrzeni koncepcji istnieje dość obszerny zbiór pojęć który odnosi się do zwierząt. Zwierzęta dzielą się na różne grupy i gatunki. Podziału tego dokonujemy na podstawie różnych cech zwierząt. FSM umożliwia, na podstawie opisów tekstowych, stworzenie wewnętrznej reprezentacji symboli opisujących zwierzęta. Nazwa zwierzęcia, razem z jego cechami, kodowana jest jako fakt w przestrzeni koncepcji. Każdy nowy opis, nowe zdanie, zmienia wzajemne relacje faktów, przyciągając te leżące w pobliżu i mające wspólne cechy (wymiary) z faktem, do którego odnosi się dany opis.

Jako przykład przeanalizujemy zdania, w których występują następujące zwierzęta: gołąb, kura, kaczka, gęś, sowa, jastrząb, orzeł, lis, pies, wilk, kot, tygrys, lew, koń, zebra, krowa; w sumie 16 zwierząt. Typowym opisem będzie: koń biega; koń ma kopyta; koń jest duży. 13 cech, pojawiających się w tych opisach to: mały, średni, duży; ma 2 nogi, 4 nogi, włosy, kopyta, grzywę, pióra; lubi polować, biegać, latać, pływać. Wynikiem analizy 76 zdań opisujących te zwierzęta jest zbiór faktów mieszczących się w  $13+16=29$  wymiarowej przestrzeni. Po wielokrotnym rozpatrzeniu tekstów fakty te ustawiły się względem siebie w pewnych wzajemnych relacjach. Jeśli rzutujemy te fakty na płaszczyznę zachowując odległości między nimi otrzymamy naturalną klasyfikację nazw zwierząt, przedstawioną na rysunku 7. Psycholodzy poznawczy ustalili, że istotnie tego typu relacje odzwierciedlają poczucie podobieństwa pomiędzy poszczególnymi zwierzętami. Wyraża się to w testach w częstości i szybkości kojarzenia oraz w sposobie grupowania różnych zwierząt ze sobą. Człowiek ucząc się na przykładach w sposób spontaniczny tworzy sobie wyobrażenie o prototypach.

Jaka część naszej wiedzy zawarta jest w takim statycznym, zorganizowanym układzie faktów naszej przestrzeni koncepcyjnej? Próbowano przy pomocy układów neuronowych analizować zdania, stworzone w oparciu o pewną sztuczną gramatykę, podając sieci próbki zdań poprawnych i niepoprawnych. Chociaż system nie miał pojęcia o regułach gramatycznych wyrobił sobie jednak wyczucie poprawności gramatycznej i potrafił określić również dla nowych zdań, czy są one poprawnie skonstruowane. Czy możliwa jest w ten sposób analiza bardziej subtelnych znaczeń? Teksty, które analizuje system, zawierają wiedzę definiującą wzajemne relacje symboli w nich występujących. Odpowiednie wzajemne relacje symboli w przestrzeni wewnętrznej pozwalają na pojawianie się właściwych skojarzeń i intuicyjnych odpowiedzi. Uczenie się języka ojczystego nie wymaga



uczenia się jego gramatyki. Kora mózgowa nie przetwarza informacji w oparciu o system reguł gramatycznych. Nie znam żadnych badań świadczących o tym, że nauka gramatyki w szkole przyczynia się do bardziej logicznego myślenia lub bardziej poprawnego języka. Zaczynam nawet podejrzewać, że jest to pozostałość po tradycji średniowiecznej, równie potrzebna jak wyrostek robaczkowy ...

Logika w połączeniu z intuicją możliwa jest do realizacji dzięki narzuceniu pewnych ograniczeń na możliwe wartości parametrów adaptujących się. Intuicja i instynkt to już nie tylko symbole zasłaniające naszą ignorancję. Można wyraźnie oddzielić wiedzę statyczną, nie wymagającą sekwencyjnych procesów myślowych, związanej z wiedzą utrwaloną w strukturze naszej przestrzeni koncepcji od wiedzy dynamicznej, wymagającej aktywnego rozumowania.

## Dokąd zmierzamy?

Po raz pierwszy widać szansę stworzenia systemów komputerowych, które będą miały „życie wewnętrzne”, a przynajmniej wewnętrzną reprezentację rzeczywistości. Będzie to zawsze reprezentacja subiektywna, zależna od materiału, który damy systemowi do analizy. Nie budowaliśmy dotychczas systemów, które bezpośrednio tworzyły sobie taki obraz, a jedynie systemy przetwarzające dane, rozwiązujące zadania algorytmiczne. Systemy adaptujące się, działające w oparciu o przestrzeń konceptualną, będą bardzo odmienne od komputerów znanych nam obecnie, będą posiadały swój własny „charakter”, indywidualność.

Dokąd doprowadzą nas modele układów adaptujących się? Nie wiemy, możemy się jednak obawiać, że przekonanie o wyjątkowości mózgu człowieka, podobnie jak inne przekonania antropocentryczne o unikalności naszego miejsca we Wszechświecie i szczególnej roli naszej cywilizacji, przekonania, z którymi czuliśmy się dobrze i których wyzbywaliśmy się w bolesny sposób, jest jeszcze jednym złudzeniem. Spójrzmy, jak szybki był rozwój nauki w ostatniej dekadzie, jak wiele już wiemy o sposobie działania mózgu, jak wiele potrafimy szczegółowo modelować, jak zmienił się świat komputerów... Zapytajmy, jak długo jeszcze myśl ludzka nie znajdzie konkurencji a nasz świat wewnętrzny będzie najbardziej wyrafinowaną reprezentacją rzeczywistości? Kiedy się nad tym zastanawiam przychodzą mi na myśl słowa T.S. Eliota:

*Uliczki ciągną się jak nudna dyskusja  
Chytra w zamysłach  
Doprowadzenia ciebie do tego przytłaczającego pytania...*

....

*I zaiste nadejdzie czas  
By pomyśleć: „Czy się ośmielę?”, „Czy się ośmielę?”  
Czas, by się odwrócić, zstąpić ze schodów,  
Z tysią płamą na czubku mej głowy -*

....

*Czy się ośmielę  
Niepokoić wszechświat?*

T.S. Eliot (1917),  
The Love-Song of Alfred J. Prufrock

## Literatura

W języku polskim brakuje niestety książek na temat nauk poznawczych, modelowania umysłu, niewiele napisano również o sieciach neuronowych. Dlatego polecić mogę jedynie kilka książek w języku angielskim wiążących się ściśle z tematem tego artykułu. Numer specjalny „Świata Nauki” z listopada 1992 roku „Umysł a mózg” zawiera najlepsze wprowadzenie do wielu aspektów modelowania umysłu. Dodałem również kilka pozycji dotyczących sztucznej inteligencji, chociaż metody tej dziedziny są bardzo odmienne od opisanego przeze mnie kierunku badań.

J.D. Bolter, *Człowiek Turinga* (PIW 1990) - to popularna książka zawierająca rozważania na temat sztucznej inteligencji i komputerów.

Z. Hippe, *Zastosowanie metod sztucznej inteligencji w chemii* (Wyd. Naukowe PWN 1993) - wbrew tytułowi zawiera szersze i przystępnie napisane ujęcie podstawowych metod sztucznej inteligencji nie tylko w zastosowaniu do chemii.

Z. Bubnicki, *Wstęp do systemów ekspertowych* (PWN 1990) - książka o systemach ekspertowych przeznaczona dla ekspertów.

R. Tadeusiewicz, *Sieci neuronowe* (Akademicka Oficyna Wyd. RM, Warszawa 1993) - książka biocybernetyka, pioniera badań nad sieciami neuronowymi w naszym kraju

J. Hertz, A. Krogh, R.G. Palmer, *Wstęp do teorii obliczeń neuronowych* (WNT, Warszawa 1993) - bardzo dobra ale dość techniczna książka o sieciach neuronowych.

T. Kacprzak, K. Ślot, *Sieci neuronowe komórkowe* (Wyd. Naukowe PWN 1995) - książka omawia bardzo szczególny rodzaj sieci neuronowych będących modelem struktur pola wzrokowego.

KORBICZ, J., OBUCHOWICZ, A. i UCIŃSKI D., *Sztuczne sieci neuronowe. Podstawy i zastosowania*. Akademicka Oficyna Wyd. RM, Warszawa 1993. [Wprowadzenie do teorii i zastosowań sieci neuronowych, zawiera również rozdział o algorytmach genetycznych.]

F. Crick, *The Astonishing hypothesis. The scientific search for the soul*. (Charles Scribner's Sons, N.Y. 1994) - nowa książka odkrywcy struktury DNA, laureata Nagrody Nobla, napisana przystępnie pomimo dość szczegółowego potraktowania pewnych aspektów budowy kory wzrokowej i całego mózgu.

P. Johnson-Laird, *The computer and the mind. An introduction to cognitive science*. (Harvard University Press, Cambridge, Massachusetts, 1988) - przystępnie napisane wprowadzenie do nauk poznawczych z punktu widzenia obliczeniowych możliwości mózgu.

A. Newell, *Unified theories of cognition* (Harvard University Press, Cambridge, Massachusetts, 1990) - ostatnia książka pioniera badań w zakresie sztucznej inteligencji. Jest to jego najbardziej ambitne przedsięwzięcie: próba stworzenia pełnego modelu umysłu i komputerowej realizacji tego modelu.

Artykuł ten powstał z notatek do referatu, wygłoszonego na posiedzeniu Polskiego Towarzystwa Fizycznego w Toruniu, 14.10.1993, pt: „Poprawianie Boga: modele sieci neuronowych.” Ukazał się jako:

Duch W (1994) Życie wewnętrzne komputerów. Toruńskie Studia Dydaktyczne, rok III(6), pp.191-206

Duch W (1995) Życie wewnętrzne komputerów cd. Komputer w Edukacji, 3-4: 19-27

Życie wewnętrzne komputerów - streszczenie

Zrozumienie sposobu działania mózgu jest drogą do stworzenia sztucznego umysłu. Jest to największe wyzwanie, jakie stoi obecnie przed nauką. Modele komputerowe sieci neuronowych dają nadzieję na konstrukcję systemów uczących się posiadających „wewnętrzne wyobrażenie” świata. Systemy takie są modelem subiektywnego, intuicyjnego rozumienia. Jako przykład podano jakościowe rozumienie prawa Ohma oraz automatyczne tworzenie się kategorii na podstawie tekstowego opisu własności zwierząt.