

Dokąd zmierza inteligencja obliczeniowa?

Włodzisław Duch

Katedra Informatyki Stosowanej, Uniwersytet Mikołaja Kopernika,
ul. Grudziądzka 5, 87-100 Toruń
www.phys.uni.torun.pl/kmk

Abstrakt. Jakie ważne problemy stoją przed metodami inteligencji obliczeniowej? Co jest jej największym wyzwaniem? Jakie metody można uznać za najbardziej obiecujące? Jakie perspektywy?

1. Problemy z definicjami.

Inteligencja obliczeniowa (computational intelligence, CI) obejmuje obecnie wiele dziedzin, które rozwinęły się bardzo szybko w ostatnich dziesięcioleciach. Inspiracje biologiczne dotyczące funkcjonowania układu nerwowego przyczyniły się do powstania sieci neuronowych, teoria ewolucji i genetyka zainspirowała powstanie algorytmów ewolucyjnych, obserwacje etnologiczne doprowadziły do algorytmów rojowych i mrówkowych, a immunologia do powstania algorytmów wzorowanych na działaniu układu odpornościowego. Psychologia i lingwistyka przyczyniły się do powstania teorii logicznych, uwzględniających niepewność i nieprecyzyjne określenie informacji, takich jak logika rozmyta, logika przybliżona czy teoria posybilistyczna. Teorie wychodzące poza logikę dwuwartościową można określić jako „miękkie obliczeniowo”, dopuszczające stopniowanie wartości.

Trudno jest znaleźć dobry polski odpowiednik oddający ideę *soft computing*, gdyż nie jest ona jasno zdefiniowana również w języku angielskim. Tworzenie nowych dziedzin na podstawie wyboru metod, które mają wejść w zakres danej dziedziny, nie jest szczęśliwym pomysłem. Na jakiej podstawie mamy np. zaliczyć algorytmy genetyczne do „miękkiego obliczania”, odrzucając metodę stopniowego studzenia, stosowaną do rozwiązywania tych samych zagadnień? Działy nauki (oraz większość działów matematyki) definiuje się określając problemy, jakimi się dana dziedzina zajmuje. Informatyka zajmuje się badaniem własności algorytmów. Problemy obliczeniowe, które nie dają się efektywnie zalgorytmizować wymagają szczególnego podejścia. Inteligencja obliczeniowa zajmuje się teorią i metodami rozwiązywania problemów, które nie są efektywnie algorytmizowalne.

Z takimi problemami organizmy biologiczne mają do czynienia na co dzień. Niektóre z tych problemów dają się rozwiązać dzięki zdolności do adaptacji, czy to za pomocą powolnych mechanizmów ewolucji, czy też mechanizmów uczenia się układu immunologicznego lub nerwowego. Przeżycie organizmu w warunkach konkurencji wymaga kategoryzacji sygnałów dochodzących ze zmysłów, rozpoznawania obiektów i kontroli własnego organizmu. Pamięć semantyczna i wyższe

czynności poznawcze, związane ze zdolnością do rozumowania, planowania, tworzenia teorii i posługiwania się językiem, oparte są na wykorzystaniu reprezentacji symbolicznych. Część inteligencji obliczeniowej, zajmującej się tą problematyką, nazywa się „sztuczną inteligencją” (AI). Od czasu neopozytywizmu utrwaliło się przekonanie o niezależności i uniwersalności procesów myślowych od ich realizacji sprzętowej, czyli procesów zachodzących w mózgu. Przekonanie to stało się podstawą funkcjonalizmu a później kognitywizmu, który głosił, że fizyczny system symboliczny wystarczy do modelowania umysłu (Newell 1981). W ten sposób sztuczna inteligencja weszła w skład nauk kognitywnych, zmierzających do zrozumienia sposobu działania układów poznawczych, w tym ludzkich umysłów.

2. AI i cała reszta.

Sztuczna inteligencja zajmowała się początkowo głównie systemami symbolicznymi, rozwijając się w oderwaniu od zagadnień związanych z rozpoznawaniem struktur i sterowaniem, będących modelami działań sensomotorycznych. Było to podejście w oczywisty sposób niewystarczające dla robotyków jak i ekspertów od badań nad mózgiem, próbujących zrozumieć neuropsychologiczne efekty uszkodzeń poszczególnych jego struktur. Postępy w takich dziedzinach jak analiza języka naturalnego czy dowodzenie twierdzeń matematycznych były powolne. Niestety, nadal utrzymuje się wyraźny podział ekspertami od sztucznej inteligencji i pozostałych działów inteligencji obliczeniowej: nie ma wspólnych konferencji, a pisma specjalistyczne zajmują się dość odmienną problematyką. Nawet w obszarze systemów uczących się nakrywanie jest niewielkie. Uczenie maszynowe w AI oparte jest na metodach indukcyjnych, poszukujących w analizowanych danych reguł symbolicznych, które można wykorzystać do rozumowania (Mitchell 1997; Russell i Norvig 1995; Cichosz 2000). Co prawda program do gry w warcaby, napisany przez Samuela w 1959 roku, uczył się dobierając wagi w heurystycznych funkcjach oceny, szybko przewyższając umiejętnościami swojego twórcę. Jednakże niewiele późniejszych programów do gier planszowych, dowodzenia twierdzeń czy analizy tekstów zawierało elementy uczenia się. Tak złożone systemy AI jak SOAR (Newell 1990) czy Act-R (Anderson 1993, 1995) wykorzystują ciekawe sposoby uczenia przydatne do rozwiązywania problemów, ale całkiem odmienne od metod inteligencji obliczeniowej.

W 1998 roku na *World Congress of Computational Intelligence (WCCI)*, w Anchorage (Alaska) brałem udział w zorganizowanym przez Lee Gilesa i Cliffa Lau panelu dyskusyjnym poświęconym wyzwaniom, stojącym przed inteligencją obliczeniową w nadchodzącym stuleciu. W panelu uczestniczył również David Fogel, Walter Freeman, Steve Grossberg, Paul Werbos i kilku innych ekspertów. Na pytanie z sali, co uczestnicy myślą o teście Turinga, Steve Grossberg natychmiast stwierdził, że nie ma on żadnego znaczenia i jest to „zagadnienie AI dla ludzi od AI”. Według niego samo pojęcie reprezentacji wiedzy, podstawowe dla sztucznej inteligencji, nie ma sensu, gdyż w mózgu nie ma reprezentacji, są tylko mamy „adaptacyjne stany rezonansowe”. Po dłuższej dyskusji udało mi się zmusić

dyskutantów do sformułowania innego wielkiego wyzwania, które stoi przed inteligencją obliczeniową: ma nim być stworzenie sztucznego szczura, robota mogącego przetrwać we wrogim środowisku naturalnym. Wymaga to oczywiście rozwiązania wielu zagadnień dotyczących percepcji, kontroli i do pewnego stopnia planowania, ale nie zastąpi testu Turinga, który dotyczy wyższych czynności poznawczych.

Dopiero w ostatniej dekadzie XX wieku pojawiły się sygnały świadczące o tym, że obydwie te poziomy – wyższych czynności poznawczych i związanych z nimi abstrakcyjnych procesów myślowych, oraz działań sensomotorycznych i afektywnych – są w istocie nierozłączne. Po pierwsze, okazało się, że płaty czołowe, odpowiedzialne za planowanie, kluczowe dla zrozumienia wyższych czynności poznawczych, bez sprzężenia z podkorowymi ośrodkami odpowiedzialnym za zachowania afektywne nie spełniają swojej roli (por. Damasio 1996). Emocje konieczne są do sprawnego podejmowania decyzji. Spowodowało to rozwój projektów dotyczących „systemów afektywnych” i ich zastosowania w robotyce. Filozofia kognitywna powinna traktować paradygmat symboliczny jako przybliżenie do opisu stanów mentalnych, nie pozwalające na pełne zrozumienie wielu zjawisk umysłowych nawet na poziomie wyższych czynności poznawczych. „Bycie-w-świecie” Heideggera, działania intencjonalne, są głębszą podstawą do zrozumienia umysłu niż procesy symboliczne. Po drugie lingwiści i filozofowie języka po raz pierwszy zaczęli docierać do sedna zagadnień związanych z semantyką, odnosząc je do metafor wiążących znaczenie pojęć z działaniem i posiadaniem ciała (por. Lakoff, Johnson 1999). Pojęcia matematyczne nie są tu wyjątkiem, kognitywna filozofia matematyki (Lakoff i Nunez 2000) wniosła tu całkiem nowe spojrzenie, wykraczając poza jałowe spory Platoników z konstruktywistami.

Dlaczego takie, wydawałoby się czysto teoretyczne zagadnienia, mają znaczenie dla przyszłości inteligencji obliczeniowej? Znalazły one już swój wyraz w niektórych projektach dotyczących robotyki, np. idei „inteligencji behawioralnej”, której ucieleśnieniem jest projekt Cog (Brooks, Stein 1994, Adams i inni, 2000), chociaż robotyka kognitywna dopiero powstaje. Oderwanie percepcji i kontroli od rozumowania może znacznie utrudnić tworzenie inteligentnych systemów nawet dla potrzeb matematyki. Najważniejsze zagadnienia CI powinny więc dotyczyć integracji procesów rozpoznawania, konstruowania modeli mentalnych, i rozumowania w oparciu o te modele. Jednym z celów AI jest konstruowanie inteligentnych agentów, autonomicznych programów komputerowych potrafiących gromadzić użyteczną wiedzę i potrafiących użyć w tym celu różnorodne środki. W CI dominują zagadnienia związane z analizą sygnałów, rozpoznawaniem wzorców i sterowaniem, które można w takich agentach wykorzystać.

3. Dziwna historia sieci neuronowych.

Sieci neuronowe są obecnie bardzo obszerną dziedziną, mającą oprócz rozbudowanej teorii uczenia zastosowania zarówno do modelowania realnych układów biologicznych, jak i zastosowania czysto praktyczne, w zagadnieniach aproksymacji i klasyfikacji. Dla modeli neurobiologicznych nie ma tu alternatywy. Po-

czątkowo inspiracją było tworzenie modeli imitujących układy biologiczne; perceptron wzorowany był na konstrukcji oka. Jednakże wkrótce okazało się, że układy neuronowe mogą się przydać w zastosowaniach technicznych, dzięki swojej zdolności do adaptacji. W latach 1960 znano już wiele nieliniowych metod aproksymacji opartych na funkcjach bazowych, takich jak klasyfikatory Gaussowskie czy klasyfikatory oparte na funkcjach potencjalnych. Aproksymacja w oparciu o funkcje sklepane lub eksponencjalne była powszechnie stosowana w fizyce i innych gałęziach nauki. Znalezienie dobrych parametrów w układach nieliniowych było jednak trudnym zadaniem.

Po okresie entuzjazmu trwającego od końca lat 1950 przez całą dekadę okazało się, że proste sieci neuronowe mają spore ograniczenia. Minsky i Papert przedstawili w książce „Perceptrons” (Minsky i Papert 1969) dokładną analizę jednowarstwowych perceptronów, znajdując szereg trudnych problemów, których nie dało się za pomocą takich układów rozwiązać. Niektóre z nich, np. problem spójności grafu, są nadal poza zasięgiem perceptronów wielowarstwowych, chociaż można je rozwiązać za pomocą sieci wysyłających impulsy, w których pojawia się synchronizacja (Wang, 2000). Praca Hopfielda z 1982 roku pokazała, jak bardzo prosty układ dynamiczny w postaci sieci elementów bipolarnych realizuje funkcje pamięci skojarzeniowej (Hopfielda, 1982). Wywołało to duży wzrost zainteresowania takimi układami wśród fizyków (np. książka Hertz, Krough, Palmer 1993) i psychologów. Praca Rumelharta, Hintona i Williama (1986) zapoczątkowała z kolei szybki rozwój zastosowań algorytmu wstecznej propagacji do analizy danych. Inspiracje nadeszły więc ze strony psychologów (wszyscy trzej autorzy są psychologami). Nawet fizycy patrzyli na sieci neuronowe jak na magiczne urządzenia wyobrażając sobie, że zamiast interpolacji zachodzą w nich jakieś tajemne procesy poznawcze (przykłady są np. w pracy Duch, Dierksen 1994). Z rozwojem sieci neuronowych wiązano ogromne nadzieje, na konferencje specjalistyczne przyjeżdżało tysiące osób.

W tej atmosferze odkrywanie związków z teorią aproksymacji, prawdopodobieństwa, statystyką i metodami rozpoznawania wzorców, które rozpoczęło się już przy końcu lat 1980, trwało dość długo. Jako wielkie odkrycie potraktowano np. zastosowanie radialnych funkcji bazowych do aproksymacji (czyli wprowadzenie sieci RBF), sprowadzające się w praktyce do klasyfikatorów i aproksymatorów Gaussowskich, znanych od ponad 20 lat. Do tej pory panuje tendencja by nazywać wszystkie sieci, korzystające z funkcji bazowych, nie mających nic wspólnego z funkcjami radialnymi, sieciami typu RBF. Nadzieje na szybkie powstanie systemów myślących, dysponujących ogólną inteligencją, wykorzystujących sieci neuronowe i algorytmy ewolucyjne, nie sprawdziły się między innymi z powodu nadmiernego skoncentrowania się na rozpoznawaniu wzorców i zagadnieniach optymalizacji.

Co nowego wniosły metody inteligencji obliczeniowej w tym zakresie? W porównaniu z wieloma innymi metodami klasyfikacji, aproksymacji i klasteryzacji rezultaty nie były początkowo nadzwyczajne (por. np. projekt Statlog, Michie i inn. 1994). Wśród 22 metod użytych w tym projekcie sieci MLP nie znalazły się ani razu, a sieci RBF znalazły się tylko raz w pierwszej trójce najlepszych metod. Samoorganizujące się mapy Kohonena okazały się wyjątkowo złymi klasyfikato-

rami, a wizualizacja przez nie oferowana nie oddaje relacji pomiędzy opisywanymi obiektami czy ich skupieniami (Duch, Naud 1996).

Od tego czasu rozwój teorii regularyzacji i wyboru modeli pozwolił na znaczne ulepszenie wyników. Pewną nowością są algorytmy konstruktywistyczne, lub bardziej ogólne algorytmy ontogeniczne, dostosowujące swoją złożoność do analizowanych danych, zwiększające i zmniejszające liczbę elementów przetwarzających sieci zależnie od potrzeb. Okazało się, że sieciowa organizacja obliczeń jest często bardzo wygodna i nazwę „sieci neuronowe” zaczęto stosować do wszelkiego rodzaju algorytmów, w których przepływ obliczeń przedstawić można w postaci węzłów sieci. Każdą sumę związać można z elementem sieci zbierającym dochodzące do niego informacje i przetwarzającym je za pomocą jakiejś funkcji bazowej. Prowadzi to do całkiem ogólnych modeli graficznych (Jordan i Sejnowski, 2001), stanowiąc pomost z popularnymi w statystyce sieciami Bayesowskimi. Z drugiej strony bardzo rozwinęły się czerpiące inspiracje z prac nad perceptronami metody SVM (Support Vector Machines, np. Vapnik 1998). Trudno traktować je jako sieci neuronowe, chociaż większość prac na ten temat pojawia się w pismach i na konferencjach zajmujących się właśnie problematyką sieci neuronowych.

Czego można się spodziewać po próbach połączenia sieci neuronowych z innymi metodami inteligencji obliczeniowej i metodami statystycznymi? Nietrudno jest połączyć jedną z kilkudziesięciu znanych metod klasteryzacji, z jedną z kilkudziesięciu metod optymalizacji (szczególnie dużą liczbę wariantów utworzyć można korzystając z algorytmów genetycznych i rojowych), przedstawić to jako algorytm sieciowy i dodać jakąś metodę interpretacji wyników. Można w ten sposób utworzyć bardzo wiele nowych algorytmów, a do każdego znaleźć jakieś dane, na których dany algorytm będzie dobrze działał. Problemem staje się stworzenie mechanizmów meta-uczenia, pozwalające na automatyczne łączenie ze sobą różnych metod analizy danych w poszukiwaniu ich najlepszej kombinacji. Pewne próby w tym kierunku zrobiono w oparciu o schemat metod opartych na podobieństwie (Duch 2000), ale zagadnienie to wymaga wielu dalszych badań.

Sieci neuronowe stały się niezwykle popularne, stworzono bardzo wiele programów komputerowych implementujące różne sieciowe algorytmy, co spowodowało lawinę zastosowań. W efekcie poszukuje się raczej zastosowań do istniejących narzędzi, niż narzędzi przydatnych do rozwiązywania konkretnych problemów. Algorytmy genetyczne stały się dla metod inteligencji obliczeniowej dominującą, niemalże jedyną metodą optymalizacji globalnej. Nie dzieje się tak w żadnej innej dziedzinie nauki. Na temat ewolucyjnych metod optymalizacji sieci neuronowych napisano niezliczone prace, podczas gdy trudno się doszukać prac próbujących stosować inne metody globalnej optymalizacji do optymalizacji sieci.

Metodologia testowania algorytmów inteligencji obliczeniowej pozostawia wiele do życzenia. Nawet w zakrojonym na dużą skalę projekcie Statlog (Michie i inn. 1994) popełniono wiele błędów: nie podano tak podstawowych informacji jak ocena wariancji używanych metod, użyto tylko najprostszyc wersji sieci neuronowych i innych algorytmów, nie wszystkie z używanych baz danych są publicznie dostępne, trudno więc dokonać porównań. Pomimo wysiłków wprowadzenia serii testów o wzrastającej trudności (np. testów PROBEN) nie są one powszechnie stosowane. Baza danych z UCI (Mertz, Murphy) nie zawiera zbioru testów o

systematycznie rosnących trudnościach i zawsze można w niej znaleźć jakiś zbiór, na którym wyniki dla nowej metody będą dobre. Dostępne tam dane mają czasem całkiem odmienne partycje treningowe od testowych, nie pozwalając żadnej poprawnie działającej metodzie (tj. zakładającej, że dane do testowania uzyskano z tego samego procesu co dane treningowe) uzyskać dobrego wyniku. Najlepsze wyniki osiągają wówczas metody przypadkowo mające korzystne obciążenie (bias) dla tego typu danych. W efekcie nadal nie wiemy, jakie metody należy stosować do analizy danych by osiągnąć najlepsze wyniki. Jednym ze spektakularnych wczesnych sukcesów zastosowań sieci neuronowych był algorytm Net-Talk, uczący się wymowy słów języka angielskiego (Sejnowski i Rosenberg 1987). Lepszy wynik osiągnąć można za pomocą prostego algorytmu najbliższego sąsiada, nie wymagającego adaptacji tysięcy parametrów (Waltz 1995).

Można się więc zastanawiać, co zostanie w dłuższym okresie czasu z tego obszaru inteligencji obliczeniowej. Teoria informacji, prawdopodobieństwa, parametrów ukrytych, teoria optymalizacji ryzyka i inne działy statystyki dają pewną podstawę dla tworzenia modeli danych. Zastosowania metod inteligencji obliczeniowej wkraczające na obszar statystyki, rachunku prawdopodobieństwa czy teorii optymalizacji mają silną konkurencję w metodach wyrastających z tych dziedzin. Niestety brakuje powszechnie dostępnych programów komputerowych realizujących algorytmy tego rodzaju. Statystycy posługują się często niezbyt poza tym środowiskiem znanym językiem *S*, a wiele ciekawych algorytmów nie doczekało się wcale publicznie dostępnej implementacji. Nie ma programów do regularyzowanej analizy dyskryminacyjnej, ciekawy algorytm Alloc80, który okazał się najlepszy w testach Statlog (Michie 1994), nie jest dostępny, trudno więc zrobić porównanie z wynikami dobrych metod statystycznych. Wiele metod klasteryzacji i wizualizacji również nie doczekało się dostępnego oprogramowania, podczas gdy wariantów sieci Kohonena jest bardzo wiele.

Z drugiej strony rozwój teorii i algorytmów uczenia się z krytykiem, pamięci asocjacyjnych opartych na sieciach z rekurencją, sieci wysyłających impulsy, metod modelowania niepewności, teorii uczenia się (*Computational Learning Theory*, por. Cichosz 2000), wyboru modeli i wielu innych idei jest trwałym i ważnym wkładem inteligencji obliczeniowej nie znajdującym na razie konkurencji w innych działach matematyki.

4. Nowe wyzwania.

Zastosowania metod inteligencji obliczeniowej ograniczone są często do problemów, którymi zajmuje się rozpoznawanie struktur (pattern recognition). Większość prac skupia się przy tym nad zagadnieniami zdefiniowanymi w ramach paradygmatu przestrzeni cech, określającej własności obiektów. Sieci neuronowe potrzebują danych w postaci wektorów liczb o ustalonej liczbie składowych. Odpowiada to funkcjom kory zmysłowej, podświadomym mechanizmom rozpoznawania podstawowych cech obiektów, wykrywaniu cech wyższego rzędu i kategoryzacji na tej podstawie. Mózgi zajmują się wyłącznie sygnałami, mającymi struk-

ture czasoprzestrzenną, sekwencjami sygnałów, podczas gdy metody CI najczęściej danymi statycznymi.

Tymczasem wiele problemów nie da się w ogóle przedstawić w tej postaci. Należą do nich zagadnienia wymagające złożonych metod reprezentacji wiedzy, opis obiektów o zmiennej strukturze (organizacji, przedsiębiorstw, cząsteczek chemicznych), sekwencji symboli (liter, wyrazów, zdań, par zasad DNA lub aminokwasów białek), zmieniającego się stanu obiektów (pacjenta, gier planszowych, gier wojennych). Niektóre z tych zagadnień wchodzą w zakres zainteresowań sztucznej inteligencji. Niezwykle ambitne projekty, takie jak *General Problem Solver* (Newell 1990, Winston 1992), od początku wytyczyły w tej dziedzinie dobrze określone cele. Stworzenie programu wykazującego się ogólną inteligencją okazało się bardzo trudne, jednakże również inteligencja ludzka nie okazała się tak uniwersalna, jak początkowo sądzono. Uczenie się rozwiązywania problemów w jednym kontekście nie prowadzi automatycznie do osiągnięcia lepszych rezultatów dla podobnych problemów w odmiennym kontekście (por. Anderson 1995). W dobrze określonej dziedzinie daje się utworzyć ontologie zawierające opis używanych pojęć i utworzyć bazę wiedzy w oparciu o powiązania między nimi. Przykładem systemu, którego kompetencje znacznie przewyższają możliwości ludzkiego intelektu jest EcoSys (Karp 2001), zawierający oparty na regułach produkcji model procesów metabolicznych i genetycznych zachodzących w bakterii *Escherische Coli*.

Takie zastosowania stawiają przed inteligencją obliczeniową szereg wyzwań. Jak wykorzystać wiedzę zdobytą w oparciu o analizę danych do systematycznego rozumowania? Stworzenie systemu do wspomagania diagnoz medycznych to jedynie pierwszy krok do planowania i monitorowania terapii. Takie działania wymagają rozważenia szeregu wariantów, a więc procesów szukania optymalnych rozwiązań. Najłatwiej jest je wykonać w systemach opartych na regułach. Jeśli z danych można wyciągnąć niewielką liczbę stosunkowo prostych reguł to da się je wykorzystać w algorytmie planującym. Zrozumienie danych, zarówno w sensie odkrywania reguł logiki klasycznej lub rozmytej, lub też szukania prototypów wystarczających do kategoryzacji przez podobieństwo, nie było dotychczas celem statystyki. W tym celu zastosować można wiele metod inteligencji obliczeniowej (Duch, Adamczak, Grąbczewski 2001).

Jakich metody można użyć, jeśli liczba cech, istotnych dla opisu danych z analizowanej bazy nie jest ustalona, a więc nie można się posłużyć paradygmatem przestrzeni wektorowej? W niektórych przypadkach można mieć nadzieję, że problem da się przeanalizować w kilku przestrzeniach, np. wstępnych testów po których nastąpią bardziej zaawansowane testy różnego rodzaju, zależnie od wyników oceny początkowych testów. Potrzebne więc będą różne modele, za pomocą których da się otrzymać końcowy rezultat. Nie zawsze jest to jednak wystarczające. Cząsteczki chemiczne można w bardzo uproszczony sposób zapisać w postaci grafów, których struktury da się analizować za pomocą sieci z rekurencją (Frasconi i inn. 1998). W nieco bardziej ogólny sposób można zdefiniować operatory przekształcające obiekty lub stany opisu problemu w siebie i obliczyć podobieństwa jako sumę kosztów elementarnych operacji. W tym przypadku koszty mogą być parametrami adaptacyjnymi, pozwalającymi na upodobnienie obiektów należących do tej samej klasy do siebie (Marczak i inn. 2002). Mając daną macierz

podobieństw można do niej zastosować wiele metod klasyfikacji, np. metody oparte na podobieństwie lub analizę dyskryminacyjną Fishera.

W realnych sytuacjach znamy zwykle tylko powiązania elementów, regularności wiążące kilka zmiennych, które można się nauczyć na prostych przykładach. Jak korzystać z wiedzy na temat podproblemów przy rozwiązywaniu złożonego zadania? Ekspert analizując formuły algebraiczne lub patrząc na sytuację na sali sądowej czy na szachownicy korzysta w intuicyjny sposób z takiej wiedzy prowadząc dłuższe rozumowanie. Nawet jeśli problem da się opisać w przestrzeni cech to początkowo znamy tylko nieliczne z nich i na podstawie fragmentarycznej wiedzy kolejno dopełniamy brakujące elementy. Wykorzystanie takiej wiedzy jako heurystyk pozwala uniknąć eksplozji kombinatorycznej w procesach szukania rozwiązań (Duch i Diercksen 1995, Duch 1997).

5. Nowe spojrzenie.

Najbardziej ambitnym celem CI jest budowa robota zdolnego nie tylko do przeżycia we wrogim środowisku, wymagającego wysokiej kompetencji behawioralnej, lecz również zdolnego do przejścia testu Turinga, a więc posiadającego wysokie kompetencje kognitywne. Budowa tego rodzaju systemu wymaga integracji wszystkich gałęzi inteligencji obliczeniowej, zarówno tych zajmujących się procesami sensomotorycznymi, rozpoznawaniem prostych bodźców, percepcją jak i wyższymi czynnościami poznawczymi. Konieczne jest nowe spojrzenie na to zagadnienie. Z jednej strony mamy komunikację pomiędzy neuronami, pozwalającą na detekcję elementarnych cech sygnału wzrokowego czy słuchowego, a z drugiej komunikację pomiędzy ludźmi na poziomie symbolicznym. Na te dwa skrajne i na wszystkie pośrednie poziomy można spojrzeć w podobny sposób.

Mózg nie jest wielką, jednorodną siecią neuronową, lecz ma specyficzną, modułarną i hierarchiczną strukturę. Symboliczne metody AI próbują w sposób przybliżony opisać działanie tego systemu na najwyższym poziomie, a sieci neuronowe robią to samo na poziomie najniższym. Badanie poziomów pośrednich jest również interesujące, choć rzadko uprawiane. Na różnych poziomach złożoności mamy do czynienia z jednostkami przetwarzającymi informację, oddziaływanymi ze sobą w procesach wymiany informacji, i tworzącymi pewne całości, które uważać można za emergentne, autonomiczne jednostki wyższego poziomu. Te same ogólne zasady opisują działanie systemów na różnych poziomach złożoności.

Najprostszym poziomem rozważanym przez inteligencję obliczeniową jest perceptron (neurobiologiczne modele obliczeniowe dotyczą również niższych poziomów, ale są to modele zjawisk biofizycznych, a nie poznawczych). Wiedza perceptronu sprowadza się do jednego parametru, progu pobudzenia, podczas gdy wagi synaptyczne pełnią rolę parametrów definiujących oddziaływania z innymi perceptronami lub światem zewnętrznym.

Perceptrony połączone razem tworzą różne struktury sieciowe, w najprostszym przypadku perceptrony wielowarstwowe (MLP), zdolne do nauczenia się dowolnych odwzorowań. Struktura połączeń sieci biologicznych nie jest w pełni ustalo-

na i zmienia się formując pod-sieci elementów oddziaływujących ze sobą. Proste elementy nie wiedzą, przed jakim zadaniem stoi taka podsieć, potrafią jedynie dostosowywać przepływ informacji wykorzystując lokalne oddziaływania pomiędzy połączonymi ze sobą elementami np. za pomocą reguły Hebba czy wstecznej propagacji błędów. Reguły uczenia pełnią tu rolę lokalnego krytyka zachowania, mającego wpływ na sposób oddziaływań i parametry wewnętrzne elementów przetwarzających informację. Oddziaływania i wiedza wewnętrzna określają rodzaj problemów, jakie dany system może rozwiązać. Zdolność do uogólniania pojawiająca się w sieciach jest pierwszym krokiem w kierunku twórczych zachowań.

Prosty perceptron definiuje hiperpłaszczyznę rozróżniającą dwa rodzaje sygnałów, poniżej i powyżej progu. Dodanie nowych parametrów wewnętrznych pozwala na realizację nowych zadań. Jednym z problemów, których perceptrony nie potrafi rozwiązać, jest problem spójności grafów, postawiony przez Minskiego i Paperta (1969). Dodanie jednego parametru reprezentującego fazę i pozwalającego na synchronizację grupy elementów pozwala na rozwiązanie problemu (Kunstman i inn. 1994). Czy to wystarczy do klasyfikacji innych topologicznych niezmienników? Jak scharakteryzować klasy złożoności problemów w zależności od rodzaju elementów przetwarzających i sposobu ich oddziaływania? Na razie jest tu więcej pytań niż odpowiedzi.

Kolumny korowe składają się z dużej liczby neuronów działających często jako jedna jednostka (czasami nazywana Hebbowskim zbiorem neuronów). Różne konfiguracje pobudzeń takich elementów tworzą podsieci kodujące złożone formy rozpoznawania lub działania. Kolumny traktowane jako jednostki przetwarzające wyższego rzędu składają się z perceptronów, które reagują na proste aspekty dochodzących do nich sygnałów. Do ich opisu nie wystarczy jeden parametr wewnętrzny neuronów progowych. Kolumny kory reagują raczej na podobieństwo struktury dochodzących do nich sygnałów do struktur wcześniej poznanych. Jak opisać takie zachowanie jedną funkcją? Zamiast funkcji sigmoidalnych z jednym parametrem wewnętrznym można użyć funkcji transferu o większej liczbie parametrów. Funkcje Gaussa mają N parametrów wewnętrznych (pozycja Gaussa, określająca jego pole recepcyjne) oraz tyle samo parametrów oddziaływania, związanych z odwrotnością dyspersji. Bardziej złożone funkcje (Duch i Jankowski 1999) mogą modelować złożone obszary decyzji, realizowane dzięki kombinacji wielu sigmoidalnych neuronów w kolumnie. Sieci o nietypowych funkcjach transferu mogą przy mniejszej całkowitej liczbie parametrów (a więc mniejszej złożoności samej sieci) osiągnąć podobne rezultaty jak standardowe sieci MLP czy RBF. Niewiele dotychczas włożono wysiłku w badanie własności takich sieci, chociaż jest rzeczą oczywistą, że funkcje transferu mają wielki wpływ na szybkość zbieżności algorytmów uczenia. Sieci wykorzystujące różne funkcje transferu w różnych węzłach (sieci heterogeniczne) umożliwiają odkrycie prostszych modeli danych niż standardowe sieci MLP. Pierwsze kroki w kierunku budowy takich systemów zostały niedawno zrobione (Duch i Jankowski 2001, Duch, Adamczak i Diercksen 2001, Jankowski i Duch 2001, Grąbczewski i Duch 2002), ale znalezienie sprawnych metod uczenia dla takich systemów pozostaje wyzwaniem.

Prawdziwe neurony wysyłają impulsy i tworzą sieci atraktorowe, pamięci skojarzeniowe, które należy opisywać używając języka układów dynamicznych. Pobudzenia takich sieci reprezentują wartości poszczególnych cech sygnału wejściowego, a ich dynamika scharakteryzowana jest przez położenie i rozmiar basenów atrakcji w przestrzeni cech. Węzły sieci reprezentujące lokalne sieci dynamiczne powinny modelować te baseny atrakcji (Duch 1997). Takie węzły, realizujące złożone funkcje transferu, mogą do pewnego odtworzyć zachowanie się pamięci skojarzeniowych, zwykle modelowanych za pomocą autoasocjacyjnych sieci z rekurencją. Autoasocjację, a w szczególności dopełnianie wartości brakujących Nie są to sieci jednokierunkowe, dla dopełnienia wartości brakujących potrzebne jest sprzężenie zwrotne od jednostek, których aktywność uznajemy za sygnał wyjściowy, do jednostek wejściowych. Ustalone wartości wejściowe \mathbf{X}_n nie będą zmieniane, podczas gdy wartości nieokreślone \mathbf{X}_u powinny dążyć do wartości maksymalizujących prawdopodobieństwo pojawienia się wektora $\mathbf{X}=(\mathbf{X}_n, \mathbf{X}_u)$. Można to osiągnąć startując z przypadkowych wartości nieznanymi zmiennymi \mathbf{X}_u i następnie szukając maksymalnego pobudzenia kolejnych jednostek wejściowych. Dla niektórych funkcji transferu (np. funkcji Gaussa) da się to prosto obliczyć. Dopuszczając niewielkie zmiany wartości wektora \mathbf{X}_n można w ten sposób również uwzględnić błędy w danych.

Zaletą sieci tego rodzaju jest oczywista: automatycznie uzupełniają wartości brakujące, potrafią odpowiadać na różne pytania dopełniając brakujące fragmenty, mogą się uczyć na niepełnych danych i danych niesklasyfikowanych tak dobierając parametry, by jednostki wyjściowe osiągały maksymalne pobudzenia. Możliwości sieci realizujących pamięci asocjacyjne i sieci do aproksymacji (takich jak MLP) warto połączyć w jednym modelu. Jeśli użyć separowalnych funkcji transferu $G(\mathbf{X})=\prod_i G_i(x_i)$, to każda ze składowych $G_i(x_i)$ może być interpretowana jako funkcja przynależności w sensie logiki rozmytej. Relacje pomiędzy położeniami lokalnych maksimów mogą być interpretowane jako podobieństwa związanych z tymi maksymami prototypów. Chociaż ogólna idea została opisana już w 1996 roku (Duch i Dierksen, FSM) nadal o tego typu sieciach niewiele wiadomo, a systematyczne modelowanie basenów atrakcji sieci dynamicznych nie zostało nigdy zrobione.

Wizualizacja procesów neurodynamicznych wymaga znacznej redukcji wymiarowości. Istotne wymiary związane są z wielkościami obserwowalnymi, a więc przestrzenią danych wejściowych i wyjściowych, ale nie zawsze uda się znaleźć dobrą aproksymację basenów atrakcji sieci atraktorowych (Amit 1995) w takiej zredukowanej przestrzeni, potrzebna może być pewna liczba wymiarów ukrytych. Taką redukcję można osiągnąć stosując rozmyte uogólnienie metody dynamiki symbolicznej. Stan sieci atraktorowej opisany jest przez wektor pobudzeń $\mathbf{X}(t)$ w przestrzeni fazowej aktywności wszystkich neuronów. W dynamice symbolicznej (Bedford i inn. 1991; Hsu 1994) dzieli się tę przestrzeń na rozłączne obszary, każdemu przypisuje symboliczną nazwę, i zapisuje w dyskretnych momentach czasu, do którego z obszarów należy $\mathbf{X}(t_i)$. Interesujące informacje dotyczą atraktorów, a więc obszarów przestrzeni, w których układ można znaleźć najczęściej. Rozmyte uogólnienie polega na zastąpieniu przynależności tak/nie do każdego z symbolicznie opisanych obszarów przez stopień przynależności opisany np. funkcją Gaussa. Zamiast pojedynczego symbolu przypisanego wektorowi $\mathbf{X}(t_i)$ mamy

więc wektor $F(\mathbf{X}(t_i))$ określający stopień przynależności do różnych obszarów w danym momencie czasu. Wektor ten zdefiniowany jest w zredukowanej przestrzeni o liczbie wymiarów równej liczbie funkcji przynależności. Przy odpowiednim wyborze położenia tych funkcji można w niskowymiarowej przestrzeni zachować wiele istotnych cech charakteryzujących neurodynamikę. Należą do nich położenia basenów atrakcji i prawdopodobieństwa przejść między nimi. Wybór optymalnej liczby i położenia funkcji przynależności można zrobić przy pomocy sieci opartych na radialnych funkcjach bazowych z silną regularyzacją i wyborem cech dla każdego węzła. Aktywność poszczególnych funkcji $F_k(\mathbf{X}(t_i))$ daje trajektorie w zredukowanej przestrzeni, dające się przybliżyć przez rozkłady gęstości prawdopodobieństwa znalezienia układu w określonym punkcie przestrzeni F_k . Jest to dokładniejszy sposób opisu niż przejście od sieci dynamicznych do automatów skończonych.

Chociaż stosowanie złożonych elementów zwiększa wewnętrzne stopnie swobody oddziaływania pomiędzy elementami są ustalone przez architekturę sieci. Usuwanie i dodawanie elementów sieci nie zmienia na przykład przestrzeni wejściowej. Tymczasem zwierzęta mają bardzo dużą liczbę receptorów zmysłowych i są zdolne do zwracania uwagi na bardzo różne kombinacje dochodzących do ich mózgowych sygnałów. Sieci atraktorowe są „kombinatorycznie produktywne”, zdolne do aktywacji bardzo wielkiej liczby konfiguracji swoich modułów. Sieci jednokierunkowe, nawet ze złożonymi elementami, mają ustalone ścieżki przepływu danych. Nawet jeśli elementy tych sieci będą miały złożone stany wewnętrzne nie wystarczy to do modelowania różnorodnych oddziaływań modułów korowych, oddziaływań wybierających z wielkiej sieci konfiguracje oddziaływujących ze sobą modułów, które mogą coś wniesić do interpretacji otrzymanego sygnału.

W jaki sposób z sieci nieskończenie wielu możliwych oddziaływań wyłania się kombinacja kompetentnych modułów, współpracujących ze sobą? Stan wewnętrzny kolumn korowych (modułów sieci) zmienia się w wyniku wcześniejszego pobudzenia (efekty torowania, znane z badań nad pamięcią), a więc niedawno używane moduły mają większe szanse się uaktywnić. Wstępne, niedoskonałe mechanizmy rozpoznawcze, działające na poziomie starej kory układu limbicznego i wzgórza, kontrolujące mechanizmy emocjonalne, decydują o dostępności neuromodulatorów, a więc o hamowaniu lub pobudzaniu różnych obszarów kory nowej. Sztuczne sieci neuronowe uwzględniają zwykle szybkie procesy rozpoznawania i wolne procesy uczenia się, związane z plastycznością mózgu. Uproszczony model kontroli kory przez układ limbiczny powinien uwzględnić fakt, że parametry wewnętrzne (wagi, czyli interakcje między elementami) sieci nie są ustalone, ale zależą od wstępnej oceny sygnału wejściowego, $P(\mathbf{X})$ oraz od ukrytych parametrów wewnętrznych $P(H(\mathbf{X}))$ charakteryzujących stan wewnętrzny modułów korowych. Każdy moduł powinien ocenić, na ile jest kompetentny w danej sytuacji, i jeśli może coś użytecznego wniesić dodać swój wkład do rozwiązania oddziałując z innymi modułami.

Idea ta może mieć szerokie zastosowanie jako zasada budowy sieci modularnych, od sieci prostych elementów po sieci oddziaływujących ze sobą agentów programowych. Ostatnio zastosowaliśmy ją do tworzenia komitetów kompetentnych klasyfikatorów (Duch, Irtet i Grudziński, w druku). Komitet traktować można jako sieć złożoną z podsieci, lub też jako sieć złożonych elementów przetwa-

rzających, z których każdy jest pewnym modelem adaptującym się lub indywidualną podsiecią. W zastosowaniu do klasyfikacji rezultaty działania $O(\mathbf{X};M_i)$ każdego modelu M_i mają wkład z wagą W_i do końcowego wyniku. Jest to więc architektura typu perceptronu, który łączy ze sobą wyniki poszczególnych modeli, ale wagi nie są w nim ustalone, gdyż są modulowane przez rezultaty wstępnego rozpoznawania za pomocą czynników multiplikatywnych $F(\mathbf{X};M_i)$. W obszarach przestrzeni cech, w których model M_i nie wykazuje kompetencji czynniki te są małe, hamując działanie odpowiednich modułów i dopuszczając do głosu tylko kompetentne modele. Efektywne wagi zależą od chwilowego stanu sieci $W_i(\mathbf{X}) = W_i F(\mathbf{X};M_i)$ pozwalając na tworzenie wirtualnych podsieci, z różnymi ścieżkami przepływu informacji. Zastosowanie tego podejścia do tworzenia komitetów przyniosło bardzo dobre rezultaty (Duch i Itert, w przygotowaniu).

Modulacja aktywności modułów wymaga równoległego dostępu do nich. Informacja rozsyłana jest w mózgu dzięki specyficznej architekturze połączeń, w tym bezpośrednim projekcjom z ośrodków podkorowych układu limbicznego do kory nowej. Zapewne istotną rolę gra tu pamięć robocza, którą Newman i Baars (1993) uważają za mechanizm dystrybucji informacji do wyspecjalizowanych modułów wewnątrz mózgu. Jest to na razie mechanizm dość tajemniczy, w szczególności rola globalnej aktywności EEG jest tu całkiem nieznaną. Przejście od skojarzeń, które realizują prostych modele neuronowe, do sekwencyjnego przetwarzania (myślenie, rozwiązywanie problemów) wymaga uwzględniania sprzężeń zwrotnych, a więc zastosowania sieci z rekurencją. Sieci złożone z modułów zmieniają swoje stany wewnętrzne (lokalną wiedzę każdego modułu) i swoje oddziaływania (modulując wielkość wag) do wymagań przepływu informacji przez system.

Na tym poziomie konieczne są procesy systematycznego szukania, jednakże neurobiologiczne mechanizmy za nimi stojące nie są znane. Wiedza zgromadzona przez moduły może służyć jako heurystyki pomagające w intuicyjnym rozumowaniu takiego systemu. Jak pokazano w pracy (Duch i Diercksen 1995) złożony problem wymagający podejścia kombinatorycznego można całkiem łatwo rozwiązać za pomocą procesów szukania brakujących wartości tak, by pobudzeniu uległa jak największa liczba kolejnych modułów. Użyto w tym przypadku modułów sieci FSM, które specjalizowały się w jakościowym rozpoznawaniu relacji pomiędzy trzema wielkościami elektrycznymi, wynikających z prawa Ohma i prawa Kirchhoffa. Przykład oparty był na analizie obwodu elektrycznego, przedstawionej w książce grupy PDP (McClelland i inn. 1986) w kilku rozdziałach. Wiele relacji typu $\Delta A = f(\Delta B, \Delta C)$, reprezentujących zmiany wartości zmiennych A, B, C , ma identyczną reprezentację w przestrzeni cech i może być realizowana przez ten sam moduł. Znając relacje pomiędzy trójkami zmiennych możemy wyciągnąć wnioski o zachowaniu się całego obwodu opisywanego przez 7 zmiennych i 5 praw, które muszą być spełnione. Rozpoznanie tych praw, a następnie aktywacja 5 modułów (wszystkie są identyczne), pozwala na znalezienie odpowiedzi w przypadku zagadnienia, którego rozwiązanie jest trudne (Duch 1997) i wymaga kombinatorycznej produktywności.

Modularne sieci, takie jak FSM, mogą więc być użyte jako ważne heurystyki w rozwiązywaniu problemów wymagających systematycznego rozumowania. Rozwiązanie znaleźć można za pomocą systematycznego szukania, w którym każdy

logiczny krok (szukanie odmiennych brakujących wartości) wspierane jest przez wiedzę „intuicyjną”, rozstrzygającą co warto dalej rozważać, a co od razu można odrzucić. Czy można w ten sposób rozwiązać typowe problemy stawiane programom rozumującym w sztucznej inteligencji? Być może, ale tworzenie systemów sieciowych, które mogłyby sprostać systemom ekspertowym pozostaje nadal wyzwaniem.

Systemy, którymi zajmuje się sztuczna inteligencja korzystają z wiedzy o złożonej strukturze, wykorzystując złożone sposoby reprezentacji wiedzy. Sieci neuronowe działają w przestrzeniach cech, a więc wykorzystują prostą reprezentację wiedzy, stanowiącą podstawę teorii rozpoznawania struktur (Duda, Hart i Stork 2001). Tego typu reprezentacja nie gra zapewne większej roli w mózgu. Nawet w obrębie jednej modalności zmysłowej, np. na poziomie percepcji wzrokowej, nie posługujemy się przestrzeniami cech, chociaż oceny podobieństwa i dyskryminacja może być wystarczającą podstawą do eksploracji wzrokowej świata (O'Regan i Noë, w druku). Na poziomie abstrakcyjnych procesów rozumowania lub analizy sensu zdań reprezentacja wektorowa nie wydaje się być użyteczna. Potrzebne są złożone struktury wiedzy, którymi zajmuje się sztuczna inteligencja. W jaki sposób można te procesy zrozumieć i modelować za pomocą sieci neuronowych?

Dość ogólne podejście do przetwarzania danych o złożonej strukturze, oparte na sieciach z rekurencją i ukrytych modelach Markova, wprowadzono w pracy (Frasconi i inn. 1998). Nie wydaje się ono łatwe w praktycznym stosowaniu, ani też dobrze uzasadnione neurobiologicznie. Przedstawię tu znacznie prostsze, i jak się wydaje bardziej ogólne, podejście. Dwa najczęściej stosowane schematy reprezentacji wiedzy w sztucznej inteligencji oparte są na przestrzeni stanów i opisie redukcji problemów (Rich i Knight 1990, Winston 1992, Nillson 1995). Stan początkowy przedstawiany jest jako stan bazy danych lub opis problemu, a stan końcowy (rozwiązanie) jest również stanem bazy danych lub prostym problemem o znanym rozwiązaniu. Definiuje się zbiór operatorów, które przekształcają obiekt początkowy (stan, opis problemu) w obiekt końcowy. Rozwiązanie jest serią przekształceń, przy czym stany pośrednie przechowujemy w pamięci roboczej lub zapisujemy na kartce w procesie szukania rozwiązania. Każda z operacji związana jest z pewnymi kosztami. W szczególności koszty mogą być stałe a ważna może być jedynie liczba operacji. Najbardziej cenione są rozwiązania najprostsze. Podstawową operacją powinno więc być ocenianie podobieństwa lub też odległości od pożądanego rozwiązania, rozumianej jako koszty przekształceń. Takie oceny można dokonać sprawnie za pomocą algorytmów programowania dynamicznego.

Podobieństwo złożonych obiektów do siebie można związać z najtańszymi kosztami przekształcenia obiektów w siebie. Podobieństwo wystarczy do dokonania kategoryzacji, a po jego ocenie pierwotne cechy nie są już potrzebne. Z najprostszych cech obrazu padającego na komórki siatkówki układ wzrokowy tworzy złożone cechy wyższego rzędu, odmienne dla różnego typu obiektów. Rozpoznając różne typy obiektów zwracamy uwagę na całkiem inne cechy, a zbiór wszystkich możliwych cech, które mogą mieć widziane obiekty, trudno sobie nawet wyobrazić. Chociaż operacje mentalne na wyższym poziomie abstrakcji nie są jeszcze znane bez wątplenia ocena podobieństwa i relacji pomiędzy złożonymi obiektami pełni w nich istotną rolę. Rozumowanie oparte na podobieństwie nie zawsze daje się sprowadzić do rozumowania logicznego, przypomina ono bardziej

rozumowanie „intuicyjne”. Reguły logiki klasycznej i rozmytej mają ograniczone możliwości ekspresji (Duch, Adamczak i Grąbczewski 2001), wykorzystanie ocen podobieństwa do prototypów stwarza większe możliwości (Duch i Grudziński, 2001). Ogólny schemat metod opartych na podobieństwie obejmuje większość znanych sieci neuronowych jako szczególny przypadek (Duch 2000). Takie metody pozwalają na rozpoznawanie wzorców wykorzystując jedynie podobieństwo, nie trzeba więc zakładać istnienia przestrzeni cech i wektorowej reprezentacji wiedzy (por. Pękalska Pacilik i Duin 2001).

Goldfarb i Nigam (1994) przedstawili ciekawą krytykę wykorzystywania przestrzeni wektorowych jako podstawy dla metod indukcyjnej generalizacji struktury klas. System ewoluujących transformacji zakłada wprowadzanie nowych operatorów transformacji obiektów oraz metod oceny ich podobieństwa w celu porównywania obiektów o odmiennej strukturze. Przydaje się to na przykład w chemii lub biologii molekularnej, gdzie obiekty mają całkiem odmienne struktury, chociaż ich własności (funkcje lub oddziaływania) powodują, że eksperci zaliczają je do tej samej klasy. Systemy ekspertowe do rozpoznawania i syntezy struktur chemicznych w niewielkim stopniu korzystają z reprezentacji w przestrzeni cech.

Sztuczna inteligencja zajmuje się symbolicznym przetwarzaniem informacji, w znacznym stopniu ignorując zagadnienia związane z percepcją, analizą scen wzrokowych, słuchowych czy sygnałów węchowych, czyli analizą wzorców o czasoprzestrzennej strukturze. Są to zagadnienia trudne, niektóre z nich mogą wymagać stosowania sieci neuronowych używających impulsów (Wang 2000, Kunstman i inn. 1994). Istnieje kilka pism specjalizujących się w modelach impulsujących neuronów o poprawnych własnościach neurobiologicznych, warto też wymienić próby wykorzystania inspiracji odkrytych dzięki modelom neurobiologicznym, ich modularnej organizacji, synchronizacji procesów uczenia się i pamięci, do rozwiązywania praktycznych zagadnień (Wermter i inn. 2001). Szczególnie interesujące są uproszczone modele przetwarzające czasoprzestrzenne wzorce pobudzeń układów sensorycznych, wykorzystujące rezultaty działania do podejmowania decyzji na poziomie wyższych czynności poznawczych. Takie modele pozwalają na naturalne pojawienie się niezmienniczości rozpoznawanych struktur. W modelu, który zaproponował Hopfield i Brody (2001) niezmienniczość względem przesunięcia w czasie i różnic w intensywności bodźców wynika z krótkotrwałej synchronizacji potencjałów wysyłanych przez grupy neuronów.

Rozpoznawanie i kategoryzacja czasoprzestrzennych wzorców pozwala na przypisywanie im symbolicznych nazw, chociaż używanie symboli nie zawsze jest dobrym przybliżeniem do opisu zjawisk neurodynamicznych. Przejście od sieci z rekurencją do automatów skończonych i symboli można wykonać na kilka sposobów: szukania reguł przejść między stanami sieci rekurencyjnych, uczenie się zachowań dyskretnych przez takie sieci, kodowanie automatów skończonych w sieciach neuronowych (Giles i Gori 1998, Sun i Giles 2001, Wermter i inn. 2001, Michel i Liu 2002). Na ten temat napisano wiele prac, ale prawie wszystkie zakładają jedynie dwa stany wewnętrzne elementów sieci lub automatów skończonych, aktywny i nieaktywny. Ogranicza to w oczywisty sposób możliwości takich sieci. Niewiele wiadomo o możliwościach automatów ze złożonymi stanami wewnętrznymi. Sekwencyjne procesy w modularnych sieciach złożonych z podsieci powinny z grubsza odpowiadać procesom przetwarzania informacji w korze mózgu.

Procesy takie można próbować aproksymować za pomocą wielostanowych, rozmytych automatów. Złożone elementy sieci z lokalną pamięcią mogą przetwarzać informację w sposób szeregowy, krok po kroku, zależnie od przeszłości. Poszczególne moduły mogą się specjalizować w rozpoznawaniu korelacji działania kilku elementów ze sobą, rozwiązywaniu fragmentów problemu. Zachowanie całości podlega też pewnym regułom, które pełnią rolę gramatyki systemu symbolicznego, realizowanego przez automat skończony.

Ewolujące sieci mogą dodawać i usuwać węzły i połączenia w czasie uczenia. Reprezentacja wewnętrzna wiedzy może wykorzystywać liczne redundante podzbiory cech, zamiast jednej sieci działającej w przestrzeni wektorowej. W ten sposób w sieci obecnych jest bardzo wiele mikrocech, z których tworzy się różne cechy wyższego rzędu dzięki dynamicznej konfiguracji modułów sieci, w zależności od zadania, nad którym pracuje system (skupienia uwagi). Jak dotychczas jest to tylko pomysł teoretyczny, nie udało się jeszcze skonstruować takich sieci.

Złożone moduły powinny wymieniać informację za pomocą wielowymiarowych połączeń. Proste skojarzenia mogą być reprezentowane przez zwykłe połączenia pomiędzy prostymi neuronami. Złożone fragmenty wiedzy realizowane przez moduły oddziałują z innymi poprzez zbiór połączeń, których aktywność zależy od informacji kontekstowej i może być modelowana przez reguły rozmyte. Ten sam schemat można użyć na wyższym poziomie przetwarzania informacji, zakładając, że moduły są teraz wyspecjalizowanymi podsystemami, odpowiadającymi większym obszarom lub strukturom mózgu, tworząc supersieci, odpowiadające funkcjom całych mózgów. Zasady działania są za każdym razem podobne: sieci oddziaływujących ze sobą modułów dostosowujące się do przepływu informacji i zmieniające swoją wewnętrzną wiedzę i sposób oddziaływania z innymi modułami. Sprawne algorytmy uczenia się znane są tylko dla sieci najprostszych elementów. Proces uczenia się prowadzi do powstania nowych, skomplikowanych sposobów zachowania. Jedną z zasad budowy takich sieci jest maksymalizacja pojemności informacyjnej systemu: jeśli supersieć nie może modelować wszystkich dochodzących do niej relacji powinna dodać do siebie dodatkowe moduły, specjalizujące się w uczeniu się faktów i relacji, które są w danym momencie potrzebne. Komunikacja pomiędzy takimi systemami może zachodzić na poziomie symbolicznym, tak jak pomiędzy agentami programowymi (Russel i Norvig 1995). Teoria agentów jest dobrym punktem startu dla rozważań nad sposobem budowy takich sieci.

Wszystkie systemy, które realizują obecnie wyższe czynności poznawcze, rozwiązujące problemy czy analizujące wypowiedzi w języku naturalnym, oparte są na technologii systemów ekspertowych (jednakże Treister-Goren i Hutchens, 2000, opisali system spontanicznie uczący się znaczenia otrzymywanych i wysyłanych symboli, który zapewne da się zrealizować w postaci sieci neuronowej). System CYC (www.cyc.com) zawierający ponad milion faktów i dziesiątki tysięcy koncepcji nie używa sieci neuronowych ani żadnych inspiracji kognitywnych, ograniczając się do metod symbolicznej reprezentacji wiedzy. Inne modele AI, które odniosły znaczny sukces, systemy Soar (Newel 1990) i Act (Anderson 1995), również opierają się wyłącznie na podejściu symbolicznym. Czy można je ulepszyć wykorzystując subsymboliczne podejścia wykorzystujące metody inteligencji obliczeniowej? Sieci Bayesowskie i modele graficzne mogą stanowić po-

most pomiędzy tymi technologiami. Jedynym systemem hybrydowym, który wykorzystywał sieci neuronowe dla analizy tekstów, był DISCERN (Miikkuleinen 1993). Choć wykorzystano w nim szereg interesujących idei system ten przestał się rozwijać.

Bardzo złożone supersieci, takie jak indywidualne mózgi, można też traktować jako jednostki oddziaływujące ze sobą i tworzące struktury wyższego rzędu, takie jak grupy ekspertów, instytucje, uniwersytety, wykorzystujące ogromną wiedzę, wymaganą do rozwiązywania problemów, z którymi borykają się współczesne społeczeństwa. Burza mózgów jest przykładem takich oddziaływań, które mogą przyczynić się do powstania nowych idei, ocenianych i analizowanych przez grupy ekspertów. Najtrudniejszym zadaniem jest tworzenie nowych idei, twórcze działania wymagające nowych kombinacji znanych elementów, generalizacji wiedzy na nowe sposoby. Proces ten nie musi się różnić w zasadniczy sposób od generalizacji wiedzy na niskim poziomie, w sieciach neuronowych, chociaż zachodzi na znacznie wyższym poziomie złożoności. Prawdziwa trudność tworzenia takich systemów może być związana z koniecznością szczegółowej reprezentacji ogromnej wiedzy, pozwalającej na dodawanie nowych kombinacji znanych elementów i tworzenie nowych koncepcji.

Literatura

- Adams, B., Breazeal C., Brooks, R., Scassellati, B.: Humanoid Robots: A New Kind of Tool, IEEE Intelligent Systems **15** (2000) 25-31
- Amit D.J.: The Hebbian paradigm reintegrated: local reverberations as internal representations. *Brain and Behavioral Science* 18 (1995) 617-657
- Anderson, J.R.: *Rules of the Mind*. Erlbaum, Hillsdale, N.J. (1993)
- Anderson, J.R.: *Learning and Memory*. J. Wiley and Sons, New York (1995)
- Brooks R.A, Stein L.A.: Building brains for bodies. *Autonomous Robots* 1 (1994) 7-25
- Bedford, T., Keane M., Series, C.: *Ergodic theory, symbolic dynamics and hyperbolic spaces*. Oxford University Press, Oxford, UK (1991)
- Cichosz, P.: *Systemy uczące się*. Wydawnictwo Naukowo-Techniczne 2000.
- McClelland, J.L, Rumelhart D.E. and the PDP research group.: *Parallel distributed processing*. The MIT Press, Cambridge, MA (1987)
- Damasio A.R.: *Descartes' Error: Emotion, Reason and the Human Brain*. Papermac (1996); tłum. Błąd Kartezjusza, Rebis, Poznań 1999
- Duch, W.: Similarity-Based Methods. *Control and Cybernetics* 4 (2000) 937-968
- Duch, W.: Platonic model of mind as an approximation to neurodynamics. In: *Brain-like computing and intelligent information systems*, ed. S. Amari, N. Kasabov. Springer, Singapore (1997) 491-512
- Duch, W., Adamczak, R., Dierksen, G.H.F.: Constructive density estimation network based on several different separable transfer functions. 9th European Symposium on Artificial Neural Networks (ESANN), Brugge. De-facto publications (2001) 107-112

- Duch, W., Adamczak, R., Grąbczewski, K.: Methodology of extraction, optimization and application of crisp and fuzzy logical rules. *IEEE Transactions on Neural Networks* 12 (2001) 277-306
- Duch, W., Diercksen, G.H.F.: Neural networks as tools to solve problems in physics and chemistry. *Computer Physics Communications* 82 (1994) 91-103
- Duch, W., Diercksen, G.H.F.: Feature Space Mapping as a universal adaptive system. *Computer Physics Communications* 87 (1995) 341-371
- Duch, W., Grudziński, K., Prototype based rules - new way to understand the data. *Int. Joint Conference on Neural Networks*, Washington D.C., July 2001, 1858-1863
- Duch, W., Itert, L., Grudziński, K.: Competent undemocratic committees. *Int. Conf. on Neural Networks and Soft Computing*, Zakopane, Poland (w druku, 2002)
- Duch W., Jankowski, N.: Survey of neural transfer functions. *Neural Computing Surveys* 2 (1999) 163-213
- Duch, W., Jankowski, N.: Transfer functions: hidden possibilities for better neural networks. *9th European Symposium on Artificial Neural Networks (ESANN)*, Brugge. De-facto publications (2001) 81-94
- Duch, W. Naud, A. Simplexes, Multi-Dimensional Scaling and Self-Organized Mapping. *Proc. 8th Joint EPS-APS Int. Conf. on Physics Computing '96*, Kraków 1996
- Duda, R.O, Hart, P.E, Stork, D.G.: *Pattern Classification*, 2nd Ed, John Wiley & Sons, New York (2001)
- Frasconi, P., Gori, M., Sperduti, A.: A General Framework for Adaptive Processing of Data Structures. *IEEE Transactions on Neural Networks* 9 (1998) 768-786
- Giles, L.C., Gori, M. (Eds): *Adaptive procesing of sequences and data structures*. Springer, Berlin (1998)
- Goldfarb, L. Nigam, S.: The unified learning paradigm: A foundation for AI. In: V.Honovar, L.Uhr, Eds. *Artificial Intelligence and Neural Networks: Steps Toward Principled Integration*. Academic Press, Boston (1994)
- Grąbczewski K, Duch W, Heterogenous forests of decision trees. *Springer Lecture Notes in Computer Science* 2415 (2002) 504-509.
- Hertz, J., Krogh, A, Palmer R.G, *Wstęp do teorii obliczeń neuronowych*. WNT, Warszawa 1993
- Hopfield, J.J.: Neural networks and physical systems with emergent collective computational abilities, *Proc. National Academy of Science USA*, 79 (1982) 2554-2558
- Hopfield, J.J., Brody, C.D.: What is a moment? Transient synchrony as a collective mechanism for spatiotemporal integration. *PNAS* 98 (2001) 1282-1287
- Hsu C.S.: Global analysis by cell mapping, *J. of Bifurcation and Chaos* 2 (1994) 727-771

- Jankowski, N., Duch W.: Optimal transfer function neural networks. 9th European Symposium on Artificial Neural Networks (ESANN), Brugge. De-facto publications (2001) 101-106
- Jordan, M., Sejnowski, T.J. (Eds.) Graphical Models. Foundations of Neural Computation. MIT Press 2001.
- Karp, P.D.: Pathway databases: a case study in computational symbolic theories. *Science* 293 (2001) 2040-2044
- Kunstman, N., Hillermeier C., Rabus, B., Tavan P.: An associative memory that can form hypotheses: a phase-coded neural network. *Biological Cybernetics* 72 (1994) 119-132
- Lakoff, G., Johnson, M.: *Philosophy in the Flesh: The Embodied Mind and its Challenge to Western Thought*. New York: Basic Books 1999
- Lakoff, G., Núñez, R.E.: *Where Mathematics Comes From: How the Embodied Mind Brings Mathematics into Being*. New York: Basic Books, 2000
- Marczak, M., Duch, W., Grudziński, K., Naud, A.: (2002) Transformation Distances, Strings and Identification of DNA Promoters. *Int. Conf. on Neural Networks and Soft Computing*, Zakopane, Poland (in print, 2002)
- Mertz, C.J., Murphy, P.M.: UCI repository of machine learning databases, <http://www.ics.uci.edu/pub/machine-learningdatabases>
- Michie, D, Spiegelharter D.J, Taylor C.C, *Machine learning, Neural and Statistical Classification*. Elis Horwood, London 1994.
- Miikkulainen, R. *Subsymbolic natural language processing: an integrated model of scripts, lexicon and memory*. MIT Press, Cambridge, MA, 1993.
- Minsky M., Papert S.: *Perceptrons*. MIT Press, Cambridge, MA (1969), 2nd ed. (1988)
- Michel, A.N, Liu, D. *Qualitative Analysis and Synthesis of Recurrent Neural Networks*. New York: Marcel Dekker, Inc., 2002
- Mitchell T.: *Machine learning*. McGraw Hill (1997)
- Newman J, Baars B.J.: Neural Global Workspace Model. *Concepts in Neuroscience* 4 (1993) 255-290
- Newell, A.: *Unified Theories of Cognition*. Cambridge, MA: Harvard University Press (1990)
- Nilsson, N. J. (1998). *Artificial Intelligence: A New Synthesis*. San Mateo, CA: Morgan Kaufmann.
- O'Regan, J.K., Noë, A.: A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences* 24(5) (2001, in print)
- Pekalska, E., Pacilik, P., Duin, R.P.W.: A generalized kernel approach to dissimilarity-based classification. *J. Machine Learning Research* 2 (2001) 175-211
- Rich E., Knight K.: *Artificial Intelligence*. McGraw Hill Inc, Int'l Edition (1991)
- Rumelhart, D.E., Hinton, G.E., Williams R.J.: Learning representations by back-propagating errors, *Nature* **323** (1986) 533-536

- Russell, S. J., Norvig, P.: Artificial Intelligence: A Modern Approach. Prentice-Hall, Englewood Cliffs, N.J. (1995)
- Sejnowski, T.J., Rosenberg, C.R.: Parallel networks that learn to pronounce English text. *Journal of Complex Systems* 1 (1987) 145-168.
- Sun, R., Giles, L. (Eds): Sequence learning. Springer Verlag, Berlin (2001)
- Treister-Goren, A., Hutchens, J.L.: Creating AI: A unique interplay between the development of learning algorithms and their education. Technical Report, AI Enterprises, Tel-Aviv 2000. Available from <http://www.a-i.com>
- Waltz D.L.: Memory-based reasoning. W: M. A. Arbib, ed, *The Handbook of Brain Theory and Neural Networks*. MIT Press 1995, str. 568-570.
- Wang, D.: On Connectedness: A Solution Based on Oscillatory Correlation. *Neural Computation* 12 (2000) 131–139
- Wermter, S., Austin, J., Willshaw, D., (Eds.): Emergent neural computational architectures based on neuroscience. *Towards neuroscience-inspired computing*. Springer, Berlin (2001)
- Winston P.: *Artificial Intelligence*. 3rd ed, Addison Wesley (1992)
- Vapnik, V, *Statistical learning theory*. New York: John Wiley & Sons, 1998.